University of São Paulo
Medical School of Ribeirão Preto

Ricardo Puziol de Oliveira

Multivariate lifetime models to evaluate long-term survivors in medical studies

Ribeirão Preto
2019

Ricardo Puziol de Oliveira

Multivariate lifetime models to evaluate long-term survivors in medical studies

Ph.D. dissertation presented to the Program of Public Health at Ribeirão Preto Medical School, University of São Paulo to obtain the degree of Doctor Science.

University of São Paulo - MSRP/USP

Supervisor: Jorge Alberto Achcar

Ribeirão Preto

2019

Catalog record electronically elaborated by the author, using the LaTeX template of USPSC class from University of São Paulo.

Oliveira, Ricaro Puziol de
Multivariate lifetime models to evaluate long-term survivors in medical studies/ Ricardo Puziol de Oliveira. – Ribeirão Preto, 2019.
188p.

Dissertation (Ph.D.) – Medical School of Ribeirão Preto. University of São Paulo - MSRP/USP . Supervisor: Achcar, Jorge Alberto.

1. Bayesian approach. 2. Multivariate models. 3. Cure rate. 4. Medical studies. 5. Public health. 6. Regression models. 7. Survival analysis. I. Achcar, Jorge Alberto. II. University of São Paulo. III. Medical School of Ribeirão Preto. IV. Multivariate lifetime models to evaluate long-term survivors in medical studies.

Ficha catalográfica elaborada eletronicamente pelo autor, utilizando o modelo LaTeX da classe USPSC da Universidade de São Paulo.

Oliveira, Ricaro Puziol de
Modelos multivariados para avaliar sobreviventes de longo prazo em estudos médicos / Ricardo Puziol de Oliveira. – Ribeirão Preto, 2019.
188p.

Tese (Doutorado) – Faculdade de Medicina de Ribeirão Preto. Universidade de São Paulo - FMRP/USP. Orientador: Achcar, Jorge Alberto.

1. Análise Bayesiana. 2. Modelos multivariados. 3. Fração de cura. 4. Estudos médicos. 5. Saúde pública. 6. Modelos regressão. 7. Análise de sobrevivência. I. Achcar, Jorge Alberto. II. Universidade de São Paulo. III. Faculdade de Medicina de Ribeirão Preto. IV. Modelos multivariados para avaliar sobreviventes de longo prazo em estudos médicos.

Oliveira, R. P. **Multivariate lifetime models to evaluate long-term survivors in medical studies**. 2019. 188p. Dissertation (Ph.D.) – Medical School of Ribeirão Preto, University of São Paulo - MSRP/USP , 2019.

Examining Board

Prof. Dr. Jorge Alberto Achcar
(Supervisor/President)

Prof. Dr. Altacílio Aparecido Nunes
(1^{st} examiner)

Prof. Dr. Edwin Ortega
(2^{nd} examiner)

Prof. Dr. Josmar Mazucheli
(3^{rd} examiner)

Ribeirão Preto, 2019.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001.

"In this world where time is your enemy, it is my greatest ally. This grand game of life that you think you play in fact plays you. To that I say... Let the games begin!"
Nefarian - World of Warcraft

Acknowledgements

I would like to express my gratitude to my advisor Professor Jorge Alberto Achcar, who not only introduced me to this field of study, but also gave me ideas and suggestions. I thank him also for sharing his knowledge and proposing interesting work themes along my doctorate. I would also like to thank all the members of my thesis committee. Their advice and patience is greatly appreciated.

Special thanks go to Professor Josmar Mazucheli, who was my advisor at my master course, for his support, especially in the manuscripts derived from this thesis. I am also thankful to the Medical School of Ribeirão Preto, the University of São Paulo and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for supporting my work.

Lastly, I would like to thank my family, especially my parents, for their love, support and for being by my side on this journey encouraging me. I would also like to thank my friends for all support and help, especially those now are my research partners.

Abstract

Multivariate survival data are presented in the literature in all shapes and sizes. A common situation is the presence of correlated lifetimes when an individual is followed-up for the occurrence of two or more types of events, or when distinct individuals have dependent event times. In many applications involving these type of data, it is common the use of continuous random variable modeling approach. In this direction, the multivariate normal distribution is the most common used since it has friendly properties such as a readily interpretable dependence structure. Moreover, in most of these studies, there is the presence of covariates such as treatments, group indicators, individual characteristics, or environmental conditions, whose relationship to lifetime is of interest. In this situation, it is needed to assume lifetime regression models. In this way, the well known Cox proportional hazards model and its variations, using the marginal hazard functions employed for the analysis of multivariate survival data in literature are not enough to explain the complete dependence structure of pair of lifetimes on the covariate vector. In this thesis, it is presented some new multivariate lifetime models assuming cure rate structure based on mixture and non-mixture approaches for the analysis of long-term survivors applied to medical studies. The proposed models could be also useful to study the dependence structure of pair of lifetimes on the covariate vector $\mathbf{X}$. The results emerging from this study reinforce the fact that the search of appropriate multivariate lifetime distributions could be extremely difficult depending on the correlation structure of the lifetime data. However, the proposed methodology could be very useful in the medical lifetime data analysis where the interest is the estimation of the fraction of patients in the studied population who never experience the event of interest. In addition, the identification of important covariates was also easily obtained assuming the proposed models even using non-informative priors for the parameters of the model, under a Bayesian approach. The results could be also extended to other cross-over trials in clinical research; reliability analysis in engineering; risk analysis in economics; among many other areas. For reproducible research, the general framework for the computer codes of the proposed modeling approach is also presented which could be carried out using free R or OpenBugs free softwares.

Keywords: Bayesian approach; multivariate models; cancer studies; continuous models; cure rate; dependence structure; discrete models; medical studies; public health; regression models; risk factors; survival analysis.

Resumo

Dados multivariados de sobrevida são apresentados na literatura em muitas formas e direcionamentos de modelagem. Uma situação comum é a presença de tempos de sobrevida correlacionados quando um indivíduo é acompanhado até a ocorrência de dois ou mais tipos de eventos, ou quando indivíduos distintos têm tempos dependentes para o mesmo tipo de evento ocorrendo várias vezes. Em muitas aplicações envolvendo esses tipos de dados, é comum o uso de uma abordagem de modelagem assumindo variáveis aleatórias contínuas. Nessa direção, a distribuição normal multivariada é a mais comumente utilizada uma vez que possui propriedades amigáveis como uma estrutura de dependência prontamente interpretável. Além disso, na maioria desses estudos, há a presença de covariáveis, como tratamentos, indicadores de grupos, características individuais ou condições ambientais, cuja relação com o tempo de vida é de interesse. Nessa situação, é necessário assumir modelos de regressão de longa duração. Dessa forma, o conhecido modelo de riscos proporcionais de Cox e suas variações, utilizando funções de risco marginais usadas para a análise de dados de sobrevida multivariada como observado na literatura, não são suficientes para explicar a estrutura de dependência completa do par de tempos de vida no vetor das covariáveis. Nesta tese, são apresentados alguns novos modelos multivariados de longa duração assumindo uma estrutura de taxa de cura baseada em abordagens de modelos de misturas e não-misturas para a análise de sobreviventes de longo prazo aplicados a dados de estudos médicos. Os modelos propostos também podem ser úteis para estudar a estrutura de dependência do par de tempos de vidas no vetor de covariáveis $\mathbf{X}$. Os resultados que emergiram deste estudo reforçam o fato de que a busca de distribuições multivariadas apropriadas podem ser extremamente difíceis, dependendo da estrutura de correlação dos dados de sobrevida. No entanto, a metodologia proposta poderia ser muito útil na análise dos dados de sobrevida médicos onde o interesse é a estimativa da fração de pacientes na população estudada que nunca experimentaram o evento de interesse. Além disso, a identificação de covariáveis importantes também foi facilmente obtida, assumindo os modelos propostos, mesmo usando distribuições a priori não informativas para os parâmetros do modelo, sob uma abordagem Bayesiana. Os resultados também poderiam ser estendidos a outros tipos de ensaios clínicos; análise de confiabilidade em engenharia; análise de risco em economia; entre muitas outras áreas. Para a pesquisa reprodutível, também é apresentada a estrutura geral para os códigos de computador da abordagem de modelagem proposta que pode ser realizada usando softwares livres R ou OpenBugs.

Palavras-chave: análise Bayesiana; modelos multivariados; estudos de câncer; modelos contínuos; taxa de cura; estrutura de dependência; modelos discretos; estudos médicos.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Motivation

Multivariate lifetime distributions have been extensively studied in the literature since for this class of models time-dependent association measures play a vital role as it is seen in medical studies, toxicology studies, cancer studies and so on. In general, in these situations, it is common the presence of censored data and the presence of a set of covariates associated to each unit where it is common the use of a continuous random variable modeling approach for the lifetime data analysis. It is important to point out that even if the data are discrete, it is common to model the dataset by a continuous distribution due to analytical tractability or facility to get parameter estimators using existing statistical software where the most popular continuous lifetime distributions usually are implemented. However, in many applications it is impossible to measure lifetime length of a device on a continuous scale, as in the on/off switching devices, number of cycles until failure, or the number of accidents in a road (see for example, Kundu and Dey, 2009; Kundu and Gupta, 2010; Kundu, 2014; Kundu and Nekoukhou, 2018).

Most of multivariate lifetime data are derived from studies that most of researchers are interested in waiting time until the occurrence of an event of interest. As an example clinical trials where the study involves following patients for a period of time and monitoring patient's survival to assess the efficacy of a new treatment (Vahidpour, 2016). In general, the event of interest in such studies could be death, cure, heart attack, remission time, reaction time for a treatment, deterioration time of a organ, or adverse reaction; and the follow-up time for the study may range from few weeks to many years. In the literature, this kind of data is called *time-to-event data*. However, in some situations, the event of interest may not occur for some individuals, even after a very long period of follow-up time. In those cases, the standard survival models cannot accurately describe the behavior of all individuals. According to Vahidpour (2016), cure rate models could be useful to be fitted by on time-to-event data with long term survivors.

From these considerations, there is a great motivation to introduce new lifetime distributions with simple mathematical properties and simplifications, especially multivariate ones, to get the inferences of interest and capture the dependence structure among two or more responses associated to each patient which is the main goal of this study. In this direction, special attention has been given on bivariate geometric distributions and bivariate Poisson distributions (see for example Arnold, 1975; Kocherlakota and Kocherlakota, 1992; Kocherlakota, 1995; Basu and Dhar, 1995; Kumar, 2008; Kemp, 2013; Lee and Cha, 2014; Nekoukhou and Kundu, 2017; Kundu and Nekoukhou, 2018) as alternatives to many bivariate continuous models introduced in the literature (see for example, Gumbel, 1960; Freund, 1961; Marshall and Olkin, 1967a,b; Downton, 1970; Hawkes, 1972; Block and Basu, 1974; Hougaard, 1986; Sarkar, 1987; Arnold and Strauss, 1988; Hanagal, 2006; Hanagal and Ahmadi, 2008; Balakrishnan and Lai, 2009).

1.2 Goals

1.2.1 Main Goals

In recent years, studies on multivariate survival models have been growing fast where the use of Bayesian methods of inference have been very popular, especially under MCMC (Markov Chain Monte Carlo) simulation techniques, to get accurate estimators for the parameters of the probability models (see for example, Achcar and Leandro, 1998; dos Santos and Achcar, 2011). Based on that, the main goal of this study is to introduce new multivariate cure rate models using stress and shock models (Marshall and Olkin, 1967b) as well the Marshall-Olkin method to add a parameter to a family of distributions (Marshall and Olkin, 1997); and to explore the accuracy of some multivariate models as extensions of some existing multivariate models presented in the literature as, for example, the bivariate exponential models (Gumbel, 1960; Freund, 1961; Downton, 1970; Block and Basu, 1974) and bivariate geometric models (Arnold, 1975; Basu and Dhar, 1995; Krishna and Pundir, 2009) under a Bayesian approach for modeling of long-term survivors in medical studies.

1.2.2 Specific Goals

- Introduce new multivariate models in presence of cure rate as well the main mathematical properties.
- Provide a great statistical background of Survival and Bayesian analysis; and data analysis involving medical data in presence of risk or prognostic factors.
- Use of open source statistical softwares, as for example R and OpenBUGS softwares, for reproducible research.

1.3 Background

Survival data, *lifetime data, failure time data,* or *time to event data* are terms used to describe data that measure the time to the occurrence of some event which arises in a number of applied fields. In this thesis, a literature review of survival analysis techniques as, for example, the Kaplan-Meier estimator (Kaplan and Meier, 1958), model diagnostics, cure rate models and so on are presented in Chapter 2. Chapter 2 also presents some concepts of Bayesian analysis as, for example, prior distributions, Markov Chains, Gibbs sampling algorithm and so on. In Chapter 3, it is studied some construction methods for bivariate models as the stress and shock methods and the Marshall-Olkin method. The discrete bivariate generalized Rayleigh distribution and a class of bivariate Lindley distributions are proposed using those methods and their main mathematical properties are studied. Chapter 3 also presents an extension of the proposed models and classical bivariate models using the cure rate models based on the mixture approach. A summary of the proposed models concludes the Chapter 3.

In Chapter 4, it is presented three applications with medical data. The first one is related to pelvic sarcomas which are unusual but not rare malignancies and account for only 1% of adult solid tumors. For the data analysis, it is assumed a univariate discrete Weibull model under cure rate based on mixture and non-mixture approaches. The regression approach was also considered to investigate some risk factors for pelvic sarcomas and the residuals of the proposed model were checked by simulated envelopes. The second application is related to tobacco which is one of the biggest public health threats that the world has ever faced, killing more than 7 million people a year according to the World Health Organization (2018). For the statistical analysis, it is considered the proposed bivariate cure rate models based on a possible existing dependence between both times. Finally, the third application is related to diabetic retinopathy disease which is a chronic progressive, potentially sight-threatening disease of the retinal microvasculature associated with the prolonged hyperglycaemia. For the statistical analysis, it is assumed as lifetimes the times to blindness for the eye randomized to laser treatment and the times to blindness for the eye randomized that not received the treatment. A regression model was considered to investigate if the age of the patient was prognosticated with diabetes as a potential risk factor for blindness.

In Chapter 5, it is presented the use of the proposed models in other fields of study as, for example, reliability analysis and multivariate probability modeling. In the reliability context, it is illustrated an industrial scenario involving series systems as well engine winding reliability. In multivariate context, it is presented a new multivariate distribution and its main mathematical properties as well two applications in medical studies. Finally, Chapter 6 end this thesis with general conclusions on the study presented here.

Chapter 2

Literature Review

2.1 An Introduction to Survival Analysis

Survival data, *lifetime data, failure time data,* or *time to event data* are terms used to describe data that measure the time to the occurrence of some event and arises in a number of applied fields, such as medicine, biology, public health, epidemiology, engineering, economics, demography, among many other fields. According to Tutz et al. (2016), in all of these fields, the focus is on the time modeling it takes until specific event occurs, that is, time-to-event data as well on predicting the probability of response, survival, or mean lifetime, comparisons of the survival distributions and the identification of risk and/or prognostic factors related to the responses given by the lifetimes.

The event could be death, as the term suggest, but the event also could be any well-defined circumstance. For example, in medical research, the event could be some life-changing occurrence such as cure from a disease, remission time for a specific cancer, times of exposed individuals until be infected by a disease, the times to deterioration level or times to reaction for a treatment in pairs of lungs, kidneys, eyes or ears of humans; in reliability analysis, the event could be related to break-down, repair of machines, or lifetimes for the n-components of a engineering system. Other possibilities are also presented in toxicology studies where the failure times may be observed for the pups in each litter or studies related to the effects of two drugs on brain activity; felons' time to parole (criminology); duration of first marriage (sociology); length of newspaper or magazine subscription (marketing); and worker's compensation claims (insurance) and their various influencing risk or prognostic factors; among many other applications (see, for example, Cox, 1972; Maller and Zhou, 1996; Klein and Moeschberger, 1997; De Angelis et al., 1999; Fleming and Lin, 2000; Lee and Wang, 2003; Romeu, 2004; Rausand and Arnljot, 2004; Ibrahim et al., 2005; Giolo and Colosimo, 2006; Sreeja and Sankaran, 2008; Singh and Mukhopadhyay, 2011; Hougaard, 2012; Crowder, 2012; Eryilmaz and Tank, 2012; Li and Dhar, 2013; Collett, 2015).

In order to study those events, it is needed to define the time of occurrence of the event of interest as well a time point, say, time 0, from which times are started to be measured. For example, when the time is measured as age, the defining time point is birth; in drugs study, the time point 0 is the time of start of treatment; in disease studies, the time point 0 is the time of diagnosis of the disease; and so on. Once the time of occurrence of an event has been defined, it is common the use of nonparametric, semi-parametric or parametric modeling approach in the data analysis.

In Section 2.1.1, it is presented some examples to illustrate the main features of survival data. These examples present data obtained from various research studies. A large amount of examples can also be found in published literature. Some practical concerns in designing a survival study are addressed in Section 2.1.2. Section 2.1.3 lays out the fundamental statistical concepts needed for modeling survival data, including distribution characterization and commonly used parameters. After that, we discuss the important issue of data incompleteness, that is, the censoring mechanisms in Section 2.1.4. Moreover, in Sections 2.1.5, 2.1.6 and 2.1.7, it is discussed some aspects of nonparametric and parametric estimation procedures for censored data as well some diagnostics procedures. Finally, Section 2.1.8 presents the cure rate model using a parametric approach which is the main goal of study of this thesis.

2.1.1 Examples of Survival Data

In this section, it is presented some examples of survival data related to medical studies. These examples motivate the methodological development in the following chapters. Survival data in these examples are obtained from various types of study designs.

2.1.1.1 Bloodstream Infection

Overuse of antibiotics is a major driver of antibiotic resistance, a growing problem in intensive care units (ICUs) worldwide. To examine the relationship between duration of carbapenem administration and subsequent nosocomial multidrug resistant (MDR) bloodstream infections (BSIs), a prospective observational study was undertaken at the National University Hospital in Singapore for all adult patients admitted to an ICU or high dependency unit (HDU) receiving more than 48 hours of a carbapenem antibiotic (Donaldson et al., 2009; Li and Ma, 2013). During the two-year study period, 415 patients were followed-up for an event of interest. The outcome of interest was the development of BSI. This disease may not directly lead to death. In this study, the research question of interest was whether the duration of carbapenem use would affect the probability of developing BSI among patients in ICU and HDU.

2.1.1.2 Lung Cancer

Lung cancer is the leading cause of cancer mortality as it is associated with a low survival probability. Nonsmall cell lung cancer (NSCLC) is more common than small cell lung cancer. Gene mutations may provide prognostic value to guide proper treatment decisions. In this example, it is considered a research study led by Tan Tock Seng Hospital in Singapore where Lim et al. (2009) used whole genome amplification (WGA) technology to investigate how specific gene mutations affected the survival probability of patients with NSCLC. In this study, 88 advanced-stage NSCLC patients were enrolled, and their low-volume lung biopsies underwent WGA before direct sequencing for EGFR, KRAS, P53, and CMET mutations. These genes have been suggested to be associated with lung cancer in published studies. Each patient was followed-up from the time of diagnosis of cancer to the time of death or the end of the study. The event of interest is the death outcome. The time for a patient to arrive at this outcome is usually called the survival time. The statistical analysis procedure to study the distribution of survival time is termed survival analysis. Specifically, practitioners are interested in determining the probability that a patient with NSCLS survives for a certain length of time, say one year, or five years, after the diagnosis of disease (Li and Ma, 2013).

2.1.1.3 Melanoma Skin Cancer

For this example of survival data, let us consider a study presented in Lee and Wang (2003) related to thirty melanoma patients (stages 2 to 4) which were studied to compare the immunotherapies BCG (Bacillus Calmette-Guerin) and Corynebacterium parvum for their abilities to prolong remission duration and survival time. It is important to point out that the patients were resected before the treatment began and thus had no evidence of melanoma at the time of the first treatment. The main goal with this type of data is to determine the length of remission and survival and to compare the distributions of remission and survival time in each group. Before comparing the remission and survival distributions, we attempt to determine if the two treatment groups are comparable with respect to prognostic factors.

2.1.1.4 Incidence of Retinopathy

For this example, let us consider a study of the incidence of retinopathy in Oklahoma Indians with NIDDM conducted in the years period 1987-1990 as part of a prospective study of diabetic complications presented in Lee and Wang (2003). In this study, from the 312 patients who were free of retinopathy at initial examination in the 1970s, 228 were found to have developed the eye disease during the 10 to 16-year follow-up period (average follow-up time 12.7 years). Also, twelve potential factors (assessed at time of baseline ex-

amination) were considered to examine their possible relationship to retinopathy (RET): age, gender, duration of diabetes (DUR), fasting plasma glucose (GLU), initial treatment (TRT), systolic (SBP) and diastolicblood pressure (DBP), body mass index (BMI), plasma cholesterol (TC), plasma triglyceride (TG), and presence of macrovascular disease (LVD) or renal disease (RD). The main goal with this type of data is to relate these variables (prognostic factors) to the development of retinopathy.

2.1.1.5 Breast Cancer

As a fifth example, let us consider a study by Sedmak et al. (1989) designed to determine if female breast cancer patients, originally classified as lymph node negative by standard light microscopy (SLM), could be more accurately classified by immunohistochemical (IH) examination of their lymph nodes with an anticytokeratin monoclonal antibody cocktail, identical sections of lymph nodes were sequentially examined by SLM and IH. According to Sedmak et al. (1989), the significance of this study is that 16% of patients with negative axillary lymph nodes, by standard pathological examination, developed recurrent disease within 10 years. In this way, forty-five female breast-cancer patients with negative axillary lymph nodes and a minimum 10-year follow-up were selected from the Ohio State University Hospitals Cancer Registry. Of the 45 patients, 9 were immunoperoxidase positive, and the remaining 36 remained negative.

2.1.1.6 Sexually Transmitted Diseases

A major problem in certain subpopulations is the occurrence of sexually transmitted diseases (STD). Even if one ignores the lethal effects of the Acquired Immune Deficiency Syndrome (AIDS), other sexually transmitted diseases still have a significant impact on the morbidity of the community as for example gonorrhea and chlamydia (see World Health Organization, 2011). According to Klein and Moeschberger (1997) these diseases are of special interest because they are often asymptomatic in the female, and, if left untreated, can lead to complications including sterility. The purpose of this kind of study is to identify those factors which are related to time until reinfection by either gonorrhea or chlamydia, given an initial infection of gonorrhea or chlamydia.

2.1.2 Design a Survival Study

A study that generates survival time data is necessarily longitudinal in nature, in contrast to a cross-sectional study where data is collected at a fixed point of the time line. The time measurements and other data for sampled subjects in a survival study can be collected retrospectively, as in a case-control study, or prospectively (Lilienfeld and Lilienfeld, 1980). A prospective survival study can be an observational cohort study or

a randomized clinical trial. An observational follow-up study allows us to characterize the natural history of the disease in quantitative terms. After we describe the severity of a disease, priorities may then be established for clinical services and public health programs. On the other hand, we usually conduct a trial to modify the natural history of a disease so as to prevent or delay an adverse outcome such as death or disability and to improve the health of an individual patient or the general population. A carefully designed randomized trial can evaluate the effectiveness and side effects of new forms of interventions. Currently, both observational studies and randomized clinical trials are being widely practiced in medicine, and the choice of design depends on the goal of the study, specific disease, and practical limitations (Li and Ma, 2013).

Statisticians can play an important role to protect patient safety and rights in the course of a research study which usually consists of multiple phases. Even though clinical trial may sound like a less frightening term than medical experiment, it essentially places human subjects under a testing environment. Another design aspect relevant to statisticians is the calculation of a proper sample size. The calculation approaches may be based on a required precision of an interval estimate of the population parameter, or a required significance and power for establishing the alternative hypothesis. One additional numeric specification for sample size calculation is how large an effect size we expect to detect. According to Li and Ma (2013), sometimes the limited resources allow only a moderate sample size for the design, and accordingly we may evaluate the power of the study under such a design. Either a sample size calculation or power analysis is necessary at the design stage.

2.1.3 Description of Survival Distributions

Suppose that a patient is randomly selected from the population and denote his time to the failure event as T. According to Li and Ma (2013), it is important to fit a probability distribution for the random variable T in order to answer practical questions such as "how long can a patient with lung cancer survive with a 90% probability?" or "will treatment A be significantly more beneficial to the patients with a chronic mental illness than treatment B?"

In order to answer the questions above, there are two functions which are the great interest in the survival data analysis: the *survival function* and *hazard function*. The survival function could be interpreted and defined as the probability of observing a survival time longer than a fixed value t, that is,

$$S(t) = P(T \geq t) = 1 - F(t) \tag{2.1}$$

where $F(t)$ denoted as the cumulative distribution function is the probability of a randomly selected subject dying before time t. According to Lawless (1982), the survival

function is a non-increasing and left continuous function, that is, $S(0) = 1$ and $S(\infty) = \lim_{t\to\infty} S(t) = 0$.

Another important related function for describing the survival distribution is the hazard function. The hazard function could be defined as the probability of dying at time t given the survival time is no less than t, that is,

$$h(t) = \lim_{\Delta t\to 0} \frac{P(T \in [t, t+\Delta t) \mid T \geq t)}{\Delta t} \tag{2.2}$$

where Δt is an infinitesimal increment of time. In some practical applications, $h(t)$ is also called *instantaneous failure rate* and could be expressed as $h(t) = f(t)/S(t)$ where $f(t)$ is the probability density (or mass, in discrete case) function. The cumulative hazard, $H(t)$, could be also obtained in terms of the survival function as

$$S(t) = \exp\{-H(t)\} \tag{2.3}$$

In summary, knowing one of the four functions, $f(t)$, $S(t)$, $H(t)$ and $h(t)$, allows us to derive the other three. These functions can serve an identical purpose of describing the distribution of survival time. Conventionally, $f(t)$ and $S(t)$ are used to form the likelihood functions for parameter estimation and hypothesis tests, while $h(t)$ is usually used to present a regression model as for example the Cox proportional hazards model (Cox, 1972).

In probability theory, the moments of the probability distributions are usually important summary measures of the distribution characteristics. Mean and variance are familiar examples, and are the first-order and second-order moments, respectively. However, for most survival distributions, it is more relevant to summarize the survival distribution by the moments for the truncated distribution since it is inappropriate to assume a symmetrical distribution. One such useful moment parameter is the mean residual life, which is defined as

$$r(t) = \mathbb{E}(T - t \mid T \geq t) = \frac{\int_t^\infty S(u)du}{S(t)} \tag{2.4}$$

and could be interpreted as the expected lifespan after a subject has survived up to time t. The mean of the overall survival time is simply $r(0)$.

Another important measure in survival analysis is the percentile. The 100^{th} percentile of the survival distribution $F(t)$ is given by

$$F^{-1}(p) = \inf\{t : F(t) > p\} \tag{2.5}$$

The 50^{th} percentile is the median survival time and is interpreted as the time at which half of the subjects in the study population can survive. It offers two advantages over the mean survival time. First, its sample estimate is less affected by extremely large or small values, while the mean estimate may be sensitive to a small number of outliers. Second, we must observe all the deaths in the study population in order to evaluate the sample mean.

2.1.4 Censoring Mechanisms

According to Klein and Moeschberger (1997); Giolo and Colosimo (2006) and Collett (2015), the main characteristic of survival data is the presence of censored or truncated observations. The survival time of a patient is censored when the point of interest was not observed for this patient. In general, this occurs because, for some reason, the patient's follow-up was interrupted or because the patient died of a cause other than the one studied.

In the literature, three common types of censoring schemes are considered: *right censoring*, where all that is known is that the individual is still alive at a given time, *left censoring* when all that is known is that the individual has experienced the event of interest prior to the start of the study, or *interval censoring*, where the only information is that the event occurs within some interval.

The most common type of censoring presented in many studies is the right-censoring scheme. This fact is related to that many studies only have limited funds, and investigators cannot wait until all the subjects develop the event of interest. Mathematically, this censoring scheme could be defined using two random variables, that is, let T be a random variable related to the failure time of a patient and C a random variable independent of T related to the censoring time. In this case, the observed time and censoring scheme are, respectively, given by,

$$t = \min(T, C) \quad \text{and} \quad \delta = \begin{cases} 1, & \text{if } T \leq C \\ 0, & \text{if } T > C \end{cases} \tag{2.6}$$

On other hand, the left-censoring scheme could be also defined using two random variables as the right-censoring scheme. In this case, t is equal to T if the lifetime is observed and δ indicates whether the exact lifetime T is observed ($\delta = 1$) or not ($\delta = 0$). Therefore, for left censoring scheme, the observed data is given by $t = \max(T, C)$.

It is important to point out that not considering the presence of censored observations, classical techniques, such as standard regression analysis with normal errors and usually transformed data, could be used in the analysis of survival data. However, these techniques – usually based on the assumption of normality or nonparametric methods – could be inappropriate in the presence of censored observations or cure rates. In this way, new techniques have been presented in the literature to deal with censored observations or cure rates. The most popular technique is given by the Cox proportional hazards model (Cox, 1972); long-term models (also refereed as cure rate models) (Wienke et al., 2006) and frailty models (Price and Manatunga, 2001). For more details of survival analysis techniques, see some standard existing lifetime books as for example: Lawless (1982); Klein and Moeschberger (1997); Meeker and Escobar (1998); Kalbfleisch and Prentice (2002); Lee and Wang (2003); Collett (2015).

2.1.5 Nonparametric Kaplan-Meier Estimator

One of the basic goals in survival analysis is to estimate the survival function $S(t)$. It could be done in two ways: using parametric or nonparametric approaches. Under a parametric approach, in general, it is assumed a parametric continuous or discrete distribution such as exponential, Weibull, log-normal, log-logistic, Pareto, inverse Gaussian, geometric, to estimate the survival function using usual inference methods (see, for example, Lawless, 1982; Ibrahim et al., 2005 or Collett, 2015). Under a nonparametric approach, the most used existing estimator of the survival function is the Kaplan-Meier estimator (Kaplan and Meier, 1958). It is defined as follows:

Definition 1.1.5.1 (Product-Limit Estimator/Kaplan-Meier Estimator) *Suppose that n individuals have lifetimes represented by random variables $T_i, i = 1, \ldots, n$ which are subject to right censoring and let $C_i, i = 1, \ldots, n$ be the corresponding censoring times. Then the observed data consist of (t_i, δ_i), where $t_i = \min(T_i, C_i)$ and $\delta_i = I(T_i = t_i), i = 1, \ldots, n$ with $I(\cdot)$ defined as the usual indicator function. Suppose that there are $k (k \leq n)$ distinct ordered times $y_1 < \ldots < y_k$ at which death occur and*

$$d_j = \sum_{i=1}^{n} I(t_i = y_j, \delta_j = 1) \tag{2.7}$$

represent the number of deaths at y_j. Then, the product-limit estimator of $S(t)$ is defined as,

$$\hat{S}(t) = \prod_{j:y_j \leq t} \frac{n_j - d_j}{n_j} \tag{2.8}$$

where $n_j = \sum_{i=1}^{n} I(t_j \geq y_j)$ is the number of individuals at risk at y_j, which is the number of individuals alive and uncensored just prior to t_j.

2.1.6 Models Diagnostics

In many practical situations, the investigators need to decide for an appropriated regression model to be fitted by the data. If a model contains too many covariates, many problems may arise, including the following:

- A multicollinearity issue can directly lead to the divergence of the computation algorithm and yield unreasonable estimation results.
- A model with numerous covariates may be more difficult to interpret than a parsimonious model.
- The requirement of estimating too many parameters may decrease the efficiency of estimates and inflate the sampling variability.

2.1.6.1 Cox-Snell Residuals

In regression analysis, model diagnostics is usually conducted using residuals. With the Cox model or a parametric model, one can assess the overall goodness-of-fit of the model using the Cox-Snell residuals (Cox and Snell, 1968; Crowley and Hu, 1977; Collett, 2015) defined by

$$r_{C_i} = -\log\{\widehat{S}_i(t_i)\} \tag{2.9}$$

where $\widehat{S}_i(t)$ is the model-based estimated survival function. When the model fitted to the observed data is overall satisfactory, the Cox-Snell residuals follow an exponential distribution with mean equals to one. This is due to the fact that if T follows a survival distribution $S(t)$, then the random variable $-\log S(T)$ follows a unit exponential distribution. However, if an individual survival time is right-censored, the corresponding Cox-Snell residual $r_{C_i}^{+}$ is smaller than the residual evaluated at an uncensored observation with same value t_i. In this way, Crowley and Hu (1977) proposed the modified Cox-Snell based on the mean and on the median of the unit exponential distribution by assuming that difference between $H(t_i)$ and $H_(t_i^{+})$ also follows the unit exponential distribution. The modified Cox-Snell residuals for censored observations is defined by,

$$r_{C_i}^{+} = 1 - \log\{\widehat{S}_i(t_i)\} \quad \text{or} \quad r_{C_i}^{+} = \log(2) - \log\{\widehat{S}_i(t_i)\} \tag{2.10}$$

The procedure for using Cox—Snell residuals can be summarized as follows:

1. Use any inference procedure and find the estimates of the parameters for the selected theoretical model.

2. Calculate the Cox-Snell residuals $r_{C_i} = -\log\{\widehat{S}_i(t_i)\}, i = 1, \ldots n$ where $\widehat{S}_i(t_i)$ is the estimated survival function with the estimates of Step 1.

3. Apply the Kaplan—Meier method to estimate the survival function $S_R(r_{C_i})$ of the Cox—Snell residuals r_{C_i}'s obtained in Step 2, then using the estimate $\widehat{S}_R(r_{C_i})$, calculate $-\log \widehat{S}_R(r_{C_i})$.

4. Plot r_{C_i} versus $-\log \widehat{S}_R(r_{C_i})$. If the plot is closed to a straight line with unit slope and zero intercept, the fitted model is appropriate.

2.1.6.2 Martingale Residuals

An alternative type of residual that takes censoring into account and is particularly suited for assessing the functional forms of predictor effects is the martingale residual. The i^{th} martingale residual is defined as

$$\widehat{M_i} = \delta_i - c_i \tag{2.11}$$

where c_i is the Cox-Snell residuals. In alternative way, M_i could be expressed, for the i^{th} subject, as

$$M_i = \delta_i + \exp\{\mathbf{x}_i^\top \widehat{\boldsymbol{\beta}}\} \log\{\widehat{S}_0(y_i)\} \tag{2.12}$$

where δ_i is the censoring indicator, y_i is the observed survival time, $\widehat{\boldsymbol{\beta}}$ is the estimated regression coefficient, and $\widehat{S}_0(t)$ is the estimated baseline survival function. The $\widehat{M_i}$ take values in $(-\infty, 1]$ and are always negative for censored observations. In large samples, the martingale residuals are uncorrelated and have expected value equal to zero. However, they are not symmetrically distributed about zero.

2.1.7 Likelihood Function in Presence of Censored Data

As stated previously, the design of survival experiments in presence of censoring needs to be carefully considered when constructing likelihood functions. A critical assumption is that the lifetimes and censoring times are independent. In this section, it is constructed the likelihood function assuming censored data for univariate and bivariate cases.

2.1.7.1 Univariate Case

Let us assume an observation corresponding to an exact event time which provides information on the probability that the event's occurring at this time, which is approximately equal to the density function of T at this time. The likelihood for various types of censoring schemes may be written by incorporating the following components:

D: Exact lifetimes - $f(t)$

R: Right-censored observations - $S(C_r)$

L: Left-censored observations - $1 - S(C_l)$

I: Interval-censored observations - $[S(L) - S(R)]$

Thus, the likelihood function based on a random sample of size n, may be constructed using all components above together as,

$$L \;\propto\; \prod_{i \in D} f(t_i) \prod_{i \in R} S(C_{ri}) \prod_{i \in L} S(C_{li}) \prod_{i \in I} [S(L_i) - S(R_i)] \tag{2.13}$$

For example, suppose that the δ indicates whether the lifetime T is observed ($\delta = 1$) or not ($\delta = 0$) and assume that the observed data is given by the expression

$t_i = \min(T_i, C_i), i = 1, 2, \ldots, n$ (right-censored data scheme), the likelihood function based on a random sample of size n, is given by,

$$L(\theta) = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i} \tag{2.14}$$

where θ is the model vector of parameters. The log-likelihood function is given by

$$\ell(\theta) = \sum_{i=1}^{n} \delta_i \log f(t_i) + \sum_{i=1}^{n} (1 - \delta_i) \log S(t_i) \tag{2.15}$$

On other hand, suppose that the δ indicates whether the lifetime T is observed ($\delta = 1$) or not ($\delta = 0$) and assume that the observed data is given by the expression $t_i = \max(T_i, C_i), i = 1, 2, \ldots, n$ (left-censored data scheme), the likelihood function based on a random sample of size n, is given by,

$$L(\theta) = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [1 - S(t_i)]^{1-\delta_i} \tag{2.16}$$

where θ is the model vector of parameters. The log-likelihood function is given by

$$\ell(\theta) = \sum_{i=1}^{n} \delta_i \log f(t_i) + \sum_{i=1}^{n} (1 - \delta_i) \log(1 - S(t_i)) \tag{2.17}$$

In both cases, the parameters could be estimated in two ways: using the maximum likelihood approach; or using a Bayesian approach. In this thesis, our focus will be on Bayesian methods since it provides less computational instability assuming censoring schemes, more accurate inference results not depending on asymptotically results and most important: the incorporation of prior opinion of experts.

2.1.7.2 Bivariate Case

For the bivariate case, let us assume the right-censoring scheme. In this case, let $(X_{11}, X_{21}), (X_{12}, X_{22}), \ldots, (X_{1n}, X_{2n})$ be a random sample of size n from a bivariate lifetime distribution and define the following indicator variables:

$$\begin{cases} \delta_{1i} = 1 & \text{if} \quad X_{1i} < C_{1i} \quad \text{and 0, for the other part.} \\ \delta_{2i} = 1 & \text{if} \quad X_{2i} < C_{2i} \quad \text{and 0, for the other part.} \end{cases} \tag{2.18}$$

where $i = 1, 2, \ldots, n; (C_{1i}, C_{2i})$ are the right censoring times. In this way, we have four possible situations:

C_1: Both, X_{1i} and X_{2i}, are complete observations ($\delta_{1i} = 1, \delta_{2i} = 1$),

C_2: X_{1i} are complete and X_{2i} are censored ($\delta_{1i} = 1, \delta_{2i} = 0$),

C_3: X_{1i} are censored and X_{2i} are complete ($\delta_{1i} = 0, \delta_{2i} = 1$),

C_4: Both, X_{1i} and X_{2i}, are censored observations ($\delta_{1i} = 0, \delta_{2i} = 0$).

In all cases, the observed data is given by the expressions $t_{1i} = \min(X_{1i}, C_{1i}), i = 1, 2, \ldots, n$ and $t_{2i} = \min(X_{2i}, C_{2i}), i = 1, 2, \ldots, n$. Thus, the likelihood function based on a random sample of size n, is given by,

1. Discrete Case:

$$\begin{aligned} L \;\propto\; & \prod_{i \in C_1} P(T_{1i} = t_{1i}, T_{2i} = t_{2i}) \prod_{i \in C_2} P(T_{1i} = t_{1i}, T_{2i} > t_{2i}) \prod_{i \in C_3} P(T_{1i} > t_{1i}, T_{2i} = t_{2i}) \\ & \times \prod_{i \in C_4} P(T_{1i} > t_{1i}, T_{2i} > t_{2i}) \end{aligned} \tag{2.19}$$

2. Continuous Case:

$$L \;\propto\; \prod_{i \in C_1} f(t_{1i}, t_{2i}) \prod_{i \in C_2} \left[-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} \right] \prod_{i \in C_3} \left[-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} \right] \prod_{i \in C_4} S(t_{1i}, t_{2i}) \tag{2.20}$$

Remark 1.1.7.2.1 *Observe that the number of possible situations (or combinations) for the right-censoring scheme in this case is given by* $\sum_{i=0}^{2} \binom{2}{i}$. *This could be expanded for the multivariate case by* $\sum_{i=0}^{n} \binom{n}{i}$, *where n is the number of lifetimes. In this case, the likelihood function has n components. For example, for three lifetimes, the number of possible situations of the right-censoring scheme is eight and the likelihood has eight components.*

2.1.8 Cure Rate Models

Another common feature of lifetime data is the presence of a cure rate. This situation may occur in different areas, such as in cancer studies where the researchers are interested in the proportion of cured patients and where many individuals may die due to other causes or in the proportion of cured patients in a clinical trial among many other applications. In these situations, we could have a fraction of individuals not expecting the occurrence of the event of interest, that is, these individuals are not at risk ('long-term survivors' or 'cured individuals').

Different approaches have been presented in the literature to model cure rate, especially for univariate lifetime data, for example: Farewell (1982); De Angelis et al. (1999); Cancho and Bolfarine (2001); Price and Manatunga (2001); Yu et al. (2004); Yin and Ibrahim (2005); Lambert et al. (2006); Lu (2010); Othus et al. (2012); Achcar et al.

(2012); Fernandes (2014). In the presence of two lifetimes associated to each unit, that is, bivariate lifetimes, Wienke et al. (2003, 2006) introduced a model for a cure rate in bivariate time-to-event data analysis. For a multivariate situation, a multivariate cure rate model was proposed by Cancho et al. (2016).

In many medical studies, especially related to cancer treatments, an important issue of interest to the medical researchers is the estimation of the fraction of individuals (or patients) in the studied population who never experience the event of interest. These individuals are not at risk with respect to the event of interest and are considered immune, cured, non-susceptible or extremely long-term survivors. Standard survival analysis techniques, for example the Cox proportional hazards (Cox, 1972) model, provides no direct estimation for the cure rate. In this way, it would be appropriate to fit parametric lifetime models which incorporates the cure rate.

According to Vahidpour (2016), the literature introduces two approaches for cure models: the mixture cure rate models, also known as standard cure rate models (see, for example, De Angelis et al., 1999; Tsodikov et al., 2003; Lambert et al., 2006), which have been widely used for modeling survival data in presence of cure rate; and the non-mixture cure rate models which are not so popular (see Achcar et al., 2012; Vahidpour, 2016). The main goal of this thesis is to explore the use of mixture cure rate models in the analysis of bivariate lifetime data assuming continuous or discrete distributions.

2.1.8.1 Mixture and Non-mixture Cure Rate Models

Let us denote by T the event of interest. Following Maller and Zhou (1996), the standard cure rate model (or mixture cure rate model) assuming that the probability of the time-to-event to be greater than a specified time t is given by the survival function,

$$S(t) = \rho + (1-\rho)S_0(t) \tag{2.21}$$

where $\rho \in (0,1)$ is the mixing parameter which represents the proportion of "long-term survivors", "non-susceptible" or "cured patients", and $S_0(t)$ denotes a proper survival function for the non-cured or susceptible group in the population. Observe that if $t \to \infty$, then $S(t) \to \rho$, that is, the survival function has an asymptote at the cure rate ρ. The probability density and the hazard functions corresponding to (2.21) are given, respectively, by,

$$f(t) = (1-\rho)f_0(t) \tag{2.22}$$

and,

$$h(t) = \frac{(1-\rho)f_0(t)}{\rho + (1-\rho)S_0(t)} \tag{2.23}$$

On other hand, an alternative non-mixture model has been proposed in the literature which defines an asymptote for the cumulative hazard and hence for cure rate (see,

Tsodikov et al., 2003). In this case, the survival function for the non-mixture cure rate model is given by,

$$S(t) = \rho^{F_0(t)} = \exp\{\ln(\rho)F_0(t)\} \tag{2.24}$$

where $\rho \in (0,1)$ is the probability of cured patients and $F_0(t) = 1 - S_0(t)$ denotes a proper distribution function for the non-cured or susceptible group in the population. The probability density and the hazard functions corresponding to (2.24) are given, respectively, by,

$$f(t) = -\ln(\rho) f_0(t) \exp\{\ln(\rho)F_0(t)\} \tag{2.25}$$

and,

$$h(t) = -\ln(\rho) f_0(t) \tag{2.26}$$

2.1.8.2 Bivariate Mixture Cure Rate Model

For the bivariate case, let us denote by T_1 and T_2 two lifetimes associated to the same individual. From (2.21), the marginal survival functions for T_1 and T_2 are given, respectively, by,

$$S_1(t) = p_1 + (1-p_1)S_{01}(t) \quad \text{and} \quad S_2(t) = p_2 + (1-p_2)S_{02}(t) \tag{2.27}$$

where p_k (proportion of cured patients) and $S_{0k}(t)$ (survival function for the non-cured patients) are associated to each lifetime $T_k, k = 1, 2$. Defining the indicators variables,

$$V_k = \begin{cases} 1 & \text{if the individual in the } k\text{-th event is susceptible} \\ 0 & \text{if the individual in the } k\text{-th event is cured,} \end{cases} \tag{2.28}$$

where $k = 1, 2$ and $\psi = cov(V_1, V_2)$ such that $0 \leq \psi \leq \min(\rho_1, \rho_2) - \rho_1\rho_2$, we have,

i.) $\phi_{11} = P(V_1 = 1, V_2 = 1) = P(V_1 = 1)P(V_2 = 1) + cov(V_1, V_2) = (1-\rho_1)(1-\rho_2) + \psi$ which indicates the probability that an individual is susceptible to both events;

ii.) $\phi_{10} = P(V_1 = 1, V_2 = 0) = P(V_1 = 1)P(V_2 = 0) - cov(V_1, V_2) = (1-\rho_1)\rho_2 - \psi$ which indicates the probability that an individual is susceptible for the first event but not for the second event;

iii.) $\phi_{01} = P(V_1 = 0, V_2 = 1) = P(V_1 = 0)P(V_2 = 1) - cov(V_1, V_2) = \rho_1(1-\rho_2) - \psi$ which indicates the probability that an individual is susceptible for the second event but not for the first event;

iv.) $\phi_{00} = P(V_1 = 0, V_2 = 0) = P(V_1 = 0)P(V_2 = 0) + cov(V_1, V_2) = \rho_1\rho_2 + \psi$ which indicates the probability that an individual is not susceptible to both events;

Following Wienke et al. (2006), the joint long-term survival function for the bivariate lifetimes T_1 and T_2 is given by,

$$S(t_1, t_2) = \phi_{11} S_0(t_1, t_2) + \phi_{10} S_{10}(t_1) + \phi_{01} S_{20}(t_2) + \phi_{00} \tag{2.29}$$

where $S_0(t_1, t_2)$ is the joint survival function for the susceptible patients; $S_{10}(t_1)$ is the marginal survival function for T_1; $S_{20}(t_2)$ is the marginal survival function for T_2 and $\phi_{11} + \phi_{10} + \phi_{01} + \phi_{00} = 1$.

2.2 An Introduction to Bayesian Inference

In general, the statistical inference is the process of data analysis to deduce properties of a population from a sampled data of that population. According to Ibrahim et al. (2005), the Bayesian paradigm is based on specifying a probability model for the observed data D, given a vector of unknown parameters θ (assuming θ is a random variable) and provides a rational method for updating the new information using the Bayes' rule and prior distributions for the uncertainty about θ. That is, the Bayesian paradigm is the process of fitting a probability model to a set of data and summarizing the result by a probability distribution on the parameters of the model and on unobserved quantities such as predictions for new observations (for more details, see Gelman et al., 1995).

Remark 1.2.1. (Prior Distribution) *A prior distribution, which is supposed to represent what is known about unknown parameters before the data is available, plays an important role in Bayesian analysis. According to Box and Tiao (2011), such a distribution can be used to represent prior knowledge or relative ignorance. In problems of scientific inference we would usually, were it possible, like the data "to speak for themselves". Consequently, it is usually appropriate to conduct the analysis as if a state of relative ignorance existed previously.*

According to Gelman et al. (1995), the process of Bayesian data analysis can be idealized by dividing it into the following three steps:

1. Setting up a *full probability model* - a joint probability distribution for all observable and unobservable quantities in a problem.

2. Conditioning on observed data: calculating and interpreting the appropriate *posterior distribution* - the conditional probability distribution of the unobserved quantities of ultimate interest, given the observed data.

3. Evaluating the fit of the model and the implications of the resulting posterior distribution: how well does the model fit the data, are the substantive conclusions reasonable, and how sensitive are the results to the modeling assumptions in step 1? In response, one can alter or expand the model and repeat the three steps.

2.2.1 Bayes Theorem

In probability theory and statistics, Bayes' theorem is named after Reverend Thomas Bayes, who first provided an equation that allows new evidence to update beliefs in his work *An Essay towards solving a Problem in the Doctrine of Chances* published in 1763 (for more details, see Bayes et al., 1763). In his work, the main objective was, in a simple way, how to infer cause from effects. For that, Bayes proposed the following system: *Initial Belief* $\times$ *New Data* $\rightarrow$ *Improved Belief* which nowadays this system could be interpreted was *Posterior Distribution* $\propto$ *Likelihood Function* $\times$ *Prior Distribution.*

To fix ideas, let $\mathbf{y} = (y_1, \ldots, y_n)$ be a vector of n observations whose probability distribution $\pi(\mathbf{y} \mid \theta)$ depends on the values of a vector $\theta = (\theta_1, \ldots, \theta_k)$ of k parameters. Suppose that θ has a prior distribution given by $\pi(\theta)$. Then, the Bayes' rule is defined by,

$$\pi(\theta \mid \mathbf{y}) = \frac{\pi(\mathbf{y}, \theta)}{\pi(\mathbf{y})} = \frac{\pi(\mathbf{y} \mid \theta)\pi(\theta)}{\pi(\mathbf{y})} \tag{2.30}$$

where it could be seen that the expression $1/\pi(\mathbf{y})$, which is independent from θ, is a normalizing constant of $\pi(\theta, \mathbf{y})$ and $\pi(\theta, \mathbf{y})$ is the posterior distribution which tells us what is known about θ given knowledge of the data (see Berger, 2013).

Definition 1.2.1.1. *For a fixed value of y, the function $L(\theta \mid \mathbf{y})$ provides the likelihood of each possible value of θ while $\pi(\theta)$ is called a priori distribution of θ. These two sources of information, priori and likelihood, are combined leading to the posterior distribution of θ, $\pi(\theta \mid \mathbf{y})$. Thus, the usual form of the Bayes theorem is described by,*

$$\pi(\theta \mid y) \propto L(\theta \mid y)\pi(\theta) \tag{2.31}$$

Definition 1.2.1.2. *From (2.31), the normalizing constant from the posterior distribution given in (2.30) is given by,*

$$\pi(\mathbf{y}) = \int \pi(\mathbf{y}, \theta)d\theta = \int \pi(\mathbf{y} \mid \theta)\pi(\theta)d\theta = \mathbf{E}_\theta[\pi(Y \mid \theta)] \tag{2.32}$$

which is called predictive distribution. Thus,

- *Before observe Y, we could check the adequacy of the priori distribution by predictions via $\pi(\mathbf{y})$.*
- *If the observed Y receives unusual predictive probability then the model must be questioned.*

According to Gelman et al. (1995), a pragmatic rationale for the use of Bayesian methods is the inherent flexibility introduced by their incorporation of multiple levels of randomness and the resultant ability to combine information from different sources,

while incorporating all reasonable sources of uncertainty in inferential summaries. In other words, Bayesian inference are conditional on probability models that invariably contain approximations in their attempt to represent complicated real-world relationships.

2.2.2 Some Family of Prior Distributions

As stated previously, a distribution can be used to represent prior knowledge or relative ignorance which leads to a important choice to use Bayesian methods of inference. In this section, it is illustrated some family of prior distributions which will be assumed in future analysis in this thesis.

2.2.2.1 Informative Prior Distributions

There are two basic interpretations that can be given to prior distributions: *population interpretation* and *knowledge interpretation*. According to Gelman et al. (1995); Box and Tiao (2011), in the population interpretation, the prior distribution represents a population of possible parameter values, from which the θ of current interest has been drawn. In the knowledge interpretation, the guiding principle is that we must express our knowledge (and uncertainty) about θ as if its value could be thought of as a random realization from the prior distribution. For many problems, the prior distribution should include all plausible values of θ, but the distribution need not be realistically concentrated around the true value, because often the information about θ contained in the data will far outweigh any reasonable prior probability specification.

2.2.2.2 Conjugated Prior Distributions

According to Gelman et al. (1995); Ehlers (2007) and Box and Tiao (2011), from the knowledge about θ, we could define a family of densities. In this case, the a priori distribution is represented by a functional form whose parameters must be specified according to this knowledge. These indexing parameters of the a priori distributions family are called hyper-parameters to distinguish them from the parameters of interest.

Definition 1.2.2.2.1. *If $F = \{\pi(x \mid \theta, \theta \in \Theta\}$ is a class of sample distributions, then a class of distributions π is conjugated to F if, for all $\pi(x \mid \theta) \in F$ and $\pi(\theta) \in \pi$, we have $\pi(\theta \mid x)$.*

Definition 1.2.2.2.2. (Conjugated Family) *In order to obtain a family of conjugate distributions, two steps are required. That is,*

1.) *Identifying the π class of distributions for θ such that $L(\theta \mid x)$ is proportional to a class member;*

2.) *Verify that π is closed by sampling, that is, if for all distributions $\pi_1, \pi_2 \in \pi$, there exists a constant k such that $k\pi_1\pi_2 \in \pi$.*

2.2.2.3 Exponential Family Prior Distributions

Probability distributions that belong to an exponential family have natural conjugate prior distributions. In this case, we have the following definition:

Definition 1.2.2.3.1. *The family of distributions with probability density function $\pi(x \mid \theta)$ belongs to the exponential family to a parameter if,*

$$\pi(x \mid \theta) = a(x) \exp[u(x)\phi(\theta) + b(\theta)] \tag{2.33}$$

In this case, the conjugate class is identified as,

$$\pi(\theta) = k(\alpha, \beta) \exp[\alpha\phi(\theta) + \beta b(\theta)] \tag{2.34}$$

and, by Bayes' theorem, it follows that:

$$\pi(\theta \mid x) = k(\alpha + u(x), \beta + 1) \exp\{[\alpha + u(x)]\phi(\theta) + (\beta + 1)b(\theta)\} \tag{2.35}$$

2.2.2.4 Non-informative Prior Distributions

According to Ehlers (2007) and Box and Tiao (2011), the first non-informative prior idea that one can have is to think of all possible values of θ as equally probable. However, this choice of priori may bring some technical difficulties, such as:

i.) If the range of θ is unlimited, then the prior distribution is improper, that is, $\int \pi(\theta) d\theta = \infty$.

ii.) If $\phi = g(\theta)$ is a monotonous nonlinear re-parametrization of θ, then $\pi(\phi)$ is non-uniform since, by transforming variables, we have:

$$\pi(\phi) = \pi(\theta(\phi)) \left| \frac{d\theta}{d\phi} \right| \propto \left| \frac{d\theta}{d\phi} \right| \tag{2.36}$$

Definition 1.2.2.4.1. (Jeffreys, 1946) *Let X be a observation with probability function $\pi(x \mid \theta)$. The non-informative prior distribution has probability density function described by:*

$$\pi(\theta) \propto [I(\theta)]^{1/2} \tag{2.37}$$

and, if θ is a parametric vector, then $P(\theta) \propto |\det I(\theta)|^{1/2}$ where $I(\theta)$ is the expected Fisher information. The equation (2.37) is called Jeffrey's Prior Distribution.

In general, a non-informative prior distribution is obtained by making the scale parameter of the conjugate distribution tend to zero and setting the other parameters in a convenient way (for more details, see Gelman et al., 1995; Zellner, 1996; Ghosh et al., 2007; Box and Tiao, 2011).

2.2.3 Point Estimation Method

In a Bayesian analysis, the posterior distribution, $f(\theta \mid x)$, plays an important role in statistical inferential procedure. There are various ways in which it could be summarized. For example, we can report our findings through point estimates. Basically, *point estimate* is the process of finding an approximate value of some parameter of a population from random samples of the population (Britney and Winkler, 1974; Bernardo and Smith, 2001; Lehmann and Casella, 2006).

Definition 1.2.3.1. (Point Estimate) *Let $X_1, \ldots, X_n$ be a random sample from a distribution with probability (density) function $\pi(x \mid \theta)$ where the value of the parameter θ is unknown. In problems of this type, the value of θ must be estimated from the values observed in the sample.*

If $\theta \in \Theta$, then it is reasonable that the possible values of an estimator $\delta(X)$ must also belong to the parametric space Θ and, in addition, a good estimator is that $\delta(X) - \theta$ tends to zero.

Definition 1.2.3.2. *For each possible value of θ and each possible estimate $a \in \Theta$, let $\eta(a, \theta)$ be the loss function so that, the greater the distance between a and θ, the greater the loss. In this case, the expected minimum loss a posteriori is given by:*

$$\min E(\eta(a, \theta) \mid x) = \min \int \eta(a, \theta) P(\theta \mid x) d\theta \tag{2.38}$$

assuming that $E(\eta(a, \theta) \mid x)$ is finite and that the minimum exists.

Definition 1.2.3.3. *The absolute loss function, defined as $\eta(a, \theta) = |a - \theta|$, introduces penalties that grow linearly with the estimation error, and in this case the associated Bayes estimate is the median of the posterior distribution of θ.*

Definition 1.2.3.4. *To further reduce the effect of large estimation errors, we consider functions that associate a fixed loss with a committed error, regardless of its magnitude. The function that makes this association is called loss 0-1 and is defined by:*

$$\eta(a, \theta) = \begin{cases} 1, & \text{if } |a - \theta| > \epsilon \\ 0, & \text{if } |a - \theta| < \epsilon \end{cases} \tag{2.39}$$

where $\epsilon > 0$. In this case, the Bayes estimator is the mode of the posterior distribution of θ. The mode of the posterior distribution of θ is also called the generalized maximum likelihood estimator.

2.2.4 Interval Estimation Method

The main constraint of the point estimate is that, when we estimate a parameter through a single numerical value, all the information present in the posterior distribution is summarized by means of this number. Some information, such as the coefficient of variation for the posterior mean, the measurement of Fisher's observed information for the posteriori mode, and the quartiles for the posterior median, determine how precise the specification of this number is. From this restriction, it is necessary to use a measure that associates the posterior distribution itself and a point estimate. This measure is called the credibility interval (Box and Tiao, 2011).

Definition 1.2.4.1. (Credibility Interval) *C is a credibility interval of $100(1-\alpha)\%$ for θ if $P(\theta \in C) \geq 1-\alpha$. The credibility intervals are also invariant under 1:1 transformations, that is, if $C = [a, b]$ is a $100(1-\alpha)\%$ credibility interval for θ, then $[\phi(a), \phi(b)]$ is a $100(1-\alpha)\%$ credibility interval for $g(\theta)$.*

Definition 1.2.4.2. (High Posterior Density Interval) *A $100(1-\alpha)\%$ C credibility interval for θ is a high posterior density interval if $C = \{\theta \in \Theta : P(\theta \mid x) \geq k(\alpha)\}$ where $k(\alpha)$ is a constant such that $P(\theta \in C) \geq 1-\alpha$.*

2.2.5 Markov Chains

Markov chains are stochastic models which play an important role in Bayesian analysis and other fields such as biology, finance, and industrial production. Basically, Markov chains are used for modeling how a system moves from one state to another in time based on a conditional probability distribution which assigns a probability to the move into a new state, given the current state of the system (see Gilks et al., 1995; Meyn and Tweedie, 2012; Kijima, 2013).

Definition 1.2.5.1. *A stochastic process is a collection of random variables X_i indexed over some set A, i.e. $\{X_i : i \in A\}$. If A is a discrete set, we have a discrete stochastic process. If A is continuous, we have a continuous stochastic process. Let $X_0, X_1, \cdots$ be a stochastic process.*

Definition 1.2.5.2. (Markov Chain) *A Markov chain is a sequence of random variables such that the next state X_{i+1} depends only on the current state X_i (memoryless property), that is,*

$$P(X_{i+1} = y \mid X_i = x_i, \cdots, X_0 = x_0) = P(X_{i+1} = y \mid X_i = x_i).$$

Remark 1.2.5.1. The probability $P_{xy} = P(X_{i+1} = y | X_i = x)$ is called the *transition kernel*. Note that $\sum_y P_{xy} = 1$. The initial distribution for X_0 determine the distribution for any n-th state.

2.2.6 Gibbs Sampling Algorithm

The Gibbs sampler was developed by Geman and Geman (1984) and could be used to generate specific multivariate distributions (see Gelfand and Smith, 1990; Chib and Greenberg, 1995; Achcar and Leandro, 1998). For example, let $\mathbf{Z}_i = (X_i, Y_i)'$ be a Markov chain, $f(x, y)$ be a given joint density, $f(x \mid y)$ and $f(y \mid x)$ to be conditional densities. In this case, the Gibbs sampling algorithm is given by:

1. Generate $Z_0 = (X_0, Y_0)'$ and set $i = 1$;
2. Generate $X_i \sim f(x_i \mid Y_{i-1} = y_{i-1})$ and $Y_i \sim f(y_i \mid X_i = x_i)$;
3. Set $i = i + 1$ and go to step 2.

2.2.7 Deviance Information Criterion (DIC)

The Deviance Information Criterion (DIC) is a criterion specially useful for selection models under the Bayesian approach where samples of the posterior distribution for the parameters of the model are obtained using MCMC methods. It is similar to AIC Akaike (1974) with two changes: replace the maximum likelihood estimate $\hat{\theta}$ with posterior mean $\hat{\theta}_{Bayes} = \mathbb{E}(\theta \mid \mathbf{y})$ and replace k with a data-based bias correction. The new measure of predictive accuracy, according to Gelman et al. (1995), is,

$$\widehat{\text{elpd}}_{\text{DIC}} = \log p(\mathbf{y} \mid \hat{\theta}_{Bayes}) - p_{\text{DIC}}, \tag{2.40}$$

where p_{DIC} is the effective number of parameters, defined as,

$$p_{\text{DIC}} = 2\left(\log p(\mathbf{y} \mid \hat{\theta}_{Bayes} - \mathbb{E}_{post}(\log p(\mathbf{y} \mid \theta)\right), \tag{2.41}$$

where the expectation in the second term is an average of θ over its posterior distribution. The posterior mean of θ will produce the maximum log predictive density when it happens to be the same as the mode, and negative p_{DIC} can be produced if posterior mean is far from the mode (Gelman et al., 1995; Spiegelhalter et al., 2014).

Finally, the actual quantity called DIC is defined in terms of the deviance rather than the log predictive density. Thus,

$$\text{DIC} = -2\log p(\mathbf{y} \mid \hat{\theta}_{Bayes}) + 2p_{\text{DIC}} \tag{2.42}$$

Smaller values of DIC indicate better models. Note that these values could be negative.

Chapter 3

An Extension of Bivariate Models Using Mixture Cure Rate Models

3.1 Introduction

In many applications of bivariate lifetimes data, time-dependent association measures play a vital role as it is seen in medical recurrent events, engineering component systems, toxicology studies, cancer studies and so on. Many bivariate distributions for continuous random variables are introduced in the literature to be used in data analysis: Gumbel (1960); Freund (1961); Marshall and Olkin (1967a,b); Downton (1970); Hawkes (1972);Block and Basu (1974); Hougaard (1986); Sarkar (1987); Arnold and Strauss (1988); Hanagal (2006); Hanagal and Ahmadi (2008). However, it could be observed in the literature that it is not very common the use of bivariate distributions for survival data assuming discrete data. In fact, few discrete bivariate distributions have been introduced in the literature as the bivariate geometric distribution of Arnold (1975) and Basu and Dhar (1995), but these discrete distributions are still not very popular in the analysis of bivariate lifetime data.

As stated previously, a common situation in lifetime data analysis is the presence of censored observations. This feature in bivariate case provides a great complexity due to the dependence structure between both lifetimes and, in general, it is required the use of parametric models which capture the dependence of the bivariate lifetimes. Many papers related to different parametric distributions are introduced in the literature to analyze bivariate lifetime data in presence of censored data: Gumbel (1960); Freund (1961); Marshall and Olkin (1967a,b); Downton (1970); Hawkes (1972); Block and Basu (1974); Hougaard (1986); Sarkar (1987); Arnold and Strauss (1988); Muraleedharan Nair and Unnikrishnan Nair (1988); Basu and Dhar (1995); Sun and Basu (1995); Dhar (2003); Dhar and Balaji (2006); Hanagal and Ahmadi (2008); Krishna and Pundir (2009); Davarzani et al. (2015).

On other hand, in the presence of a cure fraction, that is, a situation that we could have a fraction of individuals not expecting the occurrence of the event of interest ('individuals are not at risk', 'long-term survivors' or 'cured individuals'), the literature has very few papers related to use of cure rate models in the bivariate lifetime data analysis which is a motivation for the models introduced in this thesis. For the univariate case, different approaches have been presented in the literature to model cure rate: Farewell (1982); De Angelis et al. (1999); Cancho and Bolfarine (2001); Price and Manatunga (2001); Yu et al. (2004); Yin and Ibrahim (2005); Lambert et al. (2006); Lu (2010); Othus et al. (2012); Achcar et al. (2012); Fernandes (2014).

In recent years, the research on the bivariate survival models has grown rapidly, the use of Bayesian methods of inference have been very popular, especially under MCMC (Markov Chain Monte Carlo) simulation techniques, for a parametric estimation in presence of censored data, cure rate or covariates (see for example, Achcar and Leandro, 1998; dos Santos and Achcar, 2011). Based on that, the main goal of this chapter is to explore the performance and introduce new bivariate cure rate models as extensions of some bivariate models presented in the literature as, for example, the bivariate exponential models Gumbel (1960); Freund (1961); Downton (1970); Block and Basu (1974) and bivariate geometric Arnold (1975); Basu and Dhar (1995); Krishna and Pundir (2009) under a Bayesian approach for medical studies where the results could be of great interest for the search of appropriate bivariate lifetime distributions assuming a dependence structure.

3.2 Construction of Bivariate Survival Models

Multivariate survival data are usual in many areas of application. A special characteristic of multivariate data is the presence of a dependence structure among the random variables. This usually occurs when an individual response is reported by a vector related to the occurrence of two or more events of interest, or when different individuals have dependent event times (see Crowder, 2012; de Oliveira et al., 2018; de Oliveira and Achcar, 2018). The multivariate normal distribution is usually assumed for the data analysis considering the original or transformed data. One reason for this parametric model choice is the great flexibility of the multivariate normal distribution in terms of simple mathematical properties and a readily interpretable dependence structure (see, for example, Anderson, 1957; Stein, 1962; Šidák, 1967; Sidák et al., 1968; Baranchik, 1970; Khursheed and Lai Saxena, 1981; Tong, 2012). However, when the underlying process generates skewed data or the occurrence of the event of interest is rare, it is needed to assume non-normal distributions for the data analysis. In this way, the construction and study of skewed distributions is an active recent research field in statistics (see Vaidyanathan et al., 2016).

As more specific examples in multivariate lifetime applications, the times to deterioration level, the times of infection or the times to reaction for a treatment in pairs of lungs, kidneys, eyes or ears of humans could be given as count of the number of days, weeks or months. In this case, the use of a univariate continuous or discrete distribution could lead to inaccurate inference results since the dependence structure is not considered which is a great motivation to introduce new bivariate models, especially discrete. All these applications give motivations to introduce new and more flexible discrete bivariate lifetime distributions where new models are being introduced usually showing their mathematical properties and presenting estimation procedures. In this direction special attention has been given on bivariate geometric distributions and bivariate Poisson distributions (see for example Kocherlakota and Kocherlakota, 1992; Kocherlakota, 1995; Arnold, 1975; Basu and Dhar, 1995; Kumar, 2008; Kemp, 2013; Lee and Cha, 2014; Nekoukhou and Kundu, 2017; Kundu and Nekoukhou, 2018) as alternatives to many bivariate continuous models introduced in the literature (see for example, Block and Basu, 1974; Marshall and Olkin, 1967a,b; Downton, 1970; Freund, 1961; Sarkar, 1987; Arnold and Strauss, 1988; Gumbel, 1960; Hanagal, 2006; Hanagal and Ahmadi, 2008; Hawkes, 1972; Hougaard, 1986; Balakrishnan and Lai, 2009).

In some cases, some of these bivariate lifetime distributions could present the joint probability mass functions or the marginal probability mass functions in a not convenient form to be used in applications, as the two classes of discrete bivariate distributions introduced by Lee and Cha (2014), although the motivation for this model to be quite simple based on the minimum and maximum of two independent non-identical distributed random variables. In this model there are difficulties to compute the estimates of the unknown parameters, and to derive different properties (see also, Nekoukhou and Kundu, 2017). From these considerations, there is a great motivation to introduce new bivariate lifetime distributions with simple mathematical properties and simplifications to get the inferences of interest, the main goal of this study. Another possible justification for the use of the bivariate lifetime discrete distributions is the simplification of the likelihood function in presence of censored data, a common situation in lifetime data applications as in medical or engineering studies, when compared to some existing continuous bivariate lifetime distributions where the likelihood function in presence of censored data usually depends on not closed analytical forms for the joint or marginal survival functions.

In the construction of bivariate probability distributions, especially for the continuous case, the literature presents many different techniques such as: the use of copula functions, mixing and compounding; the use of trivariate reduction; the specification of a conditional and a marginal distribution; the use of a conditionally method; the construction of discrete bivariate distributions with given marginals and correlation; the use of sums and limits of Bernoulli trials; the use of clusters; the construction of finite bivariate distributions via extreme points via convex sets; the use of generalized distributions

methods; the use of canonical correlation coefficients and semi-groups; the use of bivariate distributions generated from weight functions; the use of marginal transformations method; the use of truncation or the use of two-stage failure risks of a two-component parallel system method; Marshall-Olkin methods; stress and shock models; the use of the Morgenstern family proposed by Morgenstern (1956) and the use of the bivariate dependence model proposed by Roy (2004b). According to Kemp and Papageorgiou (1982), the main problem in the construction of bivariate distributions is the impossibility to have a standard set of criteria that can always be applied to produce a unique bivariate distribution obtained from an univariate distribution which could unequivocally be called the bivariate version. For more details about the techniques above, the reader should consult Marshall and Olkin (1967a,b, 1985); Kocherlakota and Kocherlakota (1992); Marshall and Olkin (1997); Lai (2006); Balakrishnan and Lai (2009).

3.2.1 Stress and Shock Models

Shock and stress models are used in reliability to describe different applications. Shocks can refer for example to damage caused to biological organs by illness or environmental causes of damage acting on a technical system; stress can refer for example to two components which were maintained working independently and they had an overall joint maintenance scheduled at a fixed time.

The shock model structure introduced by Marshall and Olkin (1967b) (see also A-hameed and Proschan, 1973) assumes three independent sources of shocks presented in the environment of a system consisting of two components such that: a shock from source 1 destroys the component 1, it occurs at a random time W_1; a shock from source 2 destroys the component 2, it occurs at a random time W_2; a shock from source 3 destroys both components, it occurs at a random time W_3.

Definition 2.2.1.1. *Suppose that the components of a two-component system fail after receiving an overall fatal shock. Independent Poisson processes* $W_1(t, \boldsymbol{\delta}_1), W_2(t, \boldsymbol{\delta}_2), W_3(t, \boldsymbol{\delta}_3)$ *govern the occurrence of fatal shocks. Then, we have:*

- *Events in the process* $W_1(t, \boldsymbol{\delta}_1)$ *are fatal shocks transmitted to component 1.*
- *Events in the process* $W_2(t, \boldsymbol{\delta}_2)$ *are fatal shocks transmitted to component 2.*
- *Events in the process* $W_3(t, \boldsymbol{\delta}_3)$ *are fatal shocks transmitted equally and independently to both components.*

Thus if $X = \min(W_1, W_3)$ *and* $Y = \min(W_2, W_3)$ *denote, respectively, the lifetimes of the first and second components, we have that the probability of the system is working until*

an overall failure (in other words, the joint sf function) is given by,

$$\begin{aligned} P(X > x, Y > y) &= P(\{W_1(x) = 0, W_2(y) = 0, W_3(\max(x, y)) = 0\}) \\ &= P(\{\min(W_1, W_3) > x\}, \{\min(W_2, W_3) > y\}) \\ &= P(\{W_1 > x, W_3 > x\}, \{W_2 > y, W_3 > y\}) \\ &= P(W_1 > x, W_2 > y, W_3 > \max(x, y)). \end{aligned} \tag{3.1}$$

Since the random variables $W_j, (j = 1, 2, 3)$ *are mutually independent, we have,*

$$P(X > x, Y > y) = P(W_1 > x)P(W_2 > y)P(W_3 > z), \tag{3.2}$$

where $z = \max(x, y)$*. For this model, the dependence structure of the random variables* X *and* Y *is related to the common source of shock 3.*

Now, considering a generalization for more than two components series systems, it is considered first an extension of the fatal shock model to a three-component system given in the following definition.

Definition 2.2.1.2. *Let us now consider a three-component system. In this case, we have independent Poisson processes such that:*

- $W_1(t, \boldsymbol{\delta}_1)$*,* $W_2(t, \boldsymbol{\delta}_2)$*,* $W_3(t, \boldsymbol{\delta}_3)$ *govern the occurrence of fatal shocks to components 1, 2, 3, respectively;*
- $W_{12}(t, \boldsymbol{\delta}_{12})$*,* $W_{13}(t, \boldsymbol{\delta}_{13})$*,* $W_{23}(t, \boldsymbol{\delta}_{23})$ *govern the occurrence of fatal shocks to the component pairs 1 and 2, 1 and 3, 2 and 3, respectively;*
- $W_{123}(t, \boldsymbol{\delta}_{123})$ *governs the occurrence of overall fatal shock to components 1, 2, 3.*

Then, $X = \min(W_1, W_{12}, W_{13}, W_{123})$*,* $Y = \min(W_2, W_{12}, W_{23}, W_{123})$ *and* $Z = \min(W_3, W_{13}$*,* W_{23}*,* $W_{123})$*. And, the probability of the system is working until an overall failure (in other words, the trivariate sf function) is given by,*

$$\begin{aligned} P(X > x, Y > y, Z > z) &= P(\{W_1(x) = 0, W_2(y) = 0, W_3(z) = 0, \\ &\quad W_{12}(\max(x, y)) = 0, W_{13}(\max(x, z)) = 0, \\ &\quad W_{23}(\max(y, z)) = 0, W_{123}(\max(x, y, z)) = 0\}) \end{aligned} \tag{3.3}$$

Since the random variables $W_j, W_{ij}, (j = 1, 2, 3; i = 1, 2)$ *and* W_{123} *are mutually independent, we have,*

$$\begin{aligned} P(X > x, Y > y, Z > z) &= P(W_1 > x)P(W_2 > y)P(W_3 > z)P(W_{12} > \max(x, y)) \\ &\times P(W_{13} > \max(x, z))P(W_{23} > \max(y, z)) \\ &\times P(W_{123} > \max(x, y, z)). \end{aligned} \tag{3.4}$$

For the n-dimensional space, similar arguments hold. In this case, the probability of the system is working until an overall failure is given by,

$$\begin{aligned} P(X_1 > x_1, \ldots, X_n > x_n) &= P(W_1 > x_1) \cdot \ldots \cdot P(W_n > x_n) \cdot P(W_{12} > \max(x_1, x_2)) \cdot \\ &\quad \cdot \ldots \cdot P(W_{123} > \max(x_1, x_2, x_3)) \cdot \ldots \cdot \\ &\quad P(W_{12\ldots n} > \max(x_1, \ldots, x_n)) \end{aligned} \tag{3.5}$$

For the stress model structure introduced by Marshall and Olkin (1967a), it is supposed that a system has two components, where each one is subjected to individual independent stresses say U_1 and U_2 and the system has an overall stress U_3 independently transmitted equally to both components. The observed stresses for each one of the two components are $X_1 = \max(U_1, U_3)$ and $X_2 = \max(U_2, U_3)$. Similar arguments to shock model holds for the joint survival function.

3.2.2 Marshall-Olkin Method

Marshall and Olkin (1997) introduced a method to obtain an extended family of distributions including one additional parameter called univariate Marshall-Olkin family having cumulative distribution function $G(x)$ and survival function $\bar{G}(x)$ given, respectively, by,

$$G(x) = \frac{F(x)}{\alpha + \bar{\alpha}F(x)} \quad \text{and} \quad \bar{G}(x) = \frac{\alpha S(x)}{1 - \bar{\alpha}S(x)} \tag{3.6}$$

where $\alpha > 0, \bar{\alpha} = 1 - \alpha$ and $-\infty < x < \infty$. This new family of distributions has an additional parameter α which generalizes the baseline distribution and is related to the dependence structure of two random variables. This family could also be extended to the multivariate case.

Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector with multivariate cumulative and survival functions given, respectively, by $F(x_1, \ldots, x_n)$ and $S(x_1, \ldots, x_n)$, the multivariate Marshall-Olkin family has the cumulative distribution function $G(x_1, \ldots, x_n)$ and the survival function $\bar{G}(x_1, \ldots, x_n)$ given, respectively, by,

$$G(x_1, \ldots, x_n) = \frac{F(x_1, \ldots, x_n)}{\alpha + \bar{\alpha}F(x_1, \ldots, x_n)} \quad \text{and} \quad \bar{G}(x_1, \ldots, x_n) = \frac{\alpha S(x_1, \ldots, x_n)}{1 - \bar{\alpha}S(x_1, \ldots, x_n)} \tag{3.7}$$

where $\alpha > 0, \bar{\alpha} = 1 - \alpha$ and $-\infty < x_i < \infty$. Notice that the functions obtained in (3.7) are more flexible than the functions obtained from the product of independent distributions and provide a great flexibility in the modeling of the dependence structure.

3.3 New Bivariate Probability Models

3.3.1 Bivariate Discrete Generalized Rayleigh Distribution

3.3.1.1 Introduction

In this subsection it is introduced a new bivariate discrete distribution derived from two Rayleigh distributions using a method proposed by Marshall and Olkin where an additional parameter is introduced related to the dependence structure of two discrete random variables X_1 and X_2. It is shown that this new bivariate distribution has good statistical properties and simple mathematical expression for its correlation coefficient. Also, it is presented usual classical and Bayesian estimators for the parameters of the proposed model. A simulation study is carried out in order to evaluate some frequentist properties of the proposed model.

3.3.1.2 Model Description

Let $X_i, i = 1, 2$ be two independent discrete random variables having the Rayleigh distribution (see Roy, 2004a) with parameters $0 < \lambda_i < 1, i = 1, 2$. Since $X_i, i = 1, 2$ are independent, the joint survival function of the bivariate random variable X_1 and X_2 is given by,

$$\mathrm{P}_{(X_1,X_2)}(X_1 > x_1, X_2 > x_2) = \lambda_1^{x_1^2}\lambda_2^{x_2^2}. \tag{3.8}$$

Observe that the joint survival function in (3.8) is restricted to independent lifetimes and it cannot be applied directly assuming dependence structures in bivariate data. In this way, using the Marshall-Olkin survival function given in (3.7), the new proposed joint survival function is given by,

$$\mathrm{P}(X_1 > x_1, X_2 > x_2) = \frac{\alpha\lambda_1^{x_1^2}\lambda_2^{x_2^2}}{1 - \bar{\alpha}\lambda_1^{x_1^2}\lambda_2^{x_2^2}} \tag{3.9}$$

where $\bar{\alpha} = 1 - \alpha$ and $\alpha > 0$. Observe that (3.9) is more flexible than (3.8) and can be applied directly to model the dependence structure for correlated bivariate lifetimes. The joint survival function defined by (3.9) is called the discrete bivariate generalized Rayleigh (DBGR) distribution and it discrete contour is illustrated in Figure 1.

Definition 2.3.1.2.1. *Let* $\mathbf{X} = (X_1, X_2)$ *be a discrete random vector following the joint survival function given by* (3.9) *with parameters* $0 < \lambda_1, \lambda_2 < 1$ *and* $\alpha > 0$. *Defining* $h(z_1, z_2) = \lambda_1^{z_1^2}\lambda_2^{z_2^2}$, $z_1, z_2 \in \mathbb{R}_+$, *the joint probability mass function (pmf) of the* $\mathbf{X}$ *is given by,*

$$P(X_1 = x_1, X_2 = x_2) = \frac{\alpha[1 - \bar{\alpha}^2 h(k_1, k_2)]\,[1 - \bar{\alpha}h(x_1, x_2)]^{-1}\,DR(x_1, \lambda_1)DR(x_2, \lambda_2)}{[1 - \bar{\alpha}\,h(x_1 + 1, x_2)]\,[1 - \bar{\alpha}\,h(x_1, x_2 + 1)]\,[1 - \bar{\alpha}\,h(x_1 + 1, x_2 + 1)]} \tag{3.10}$$

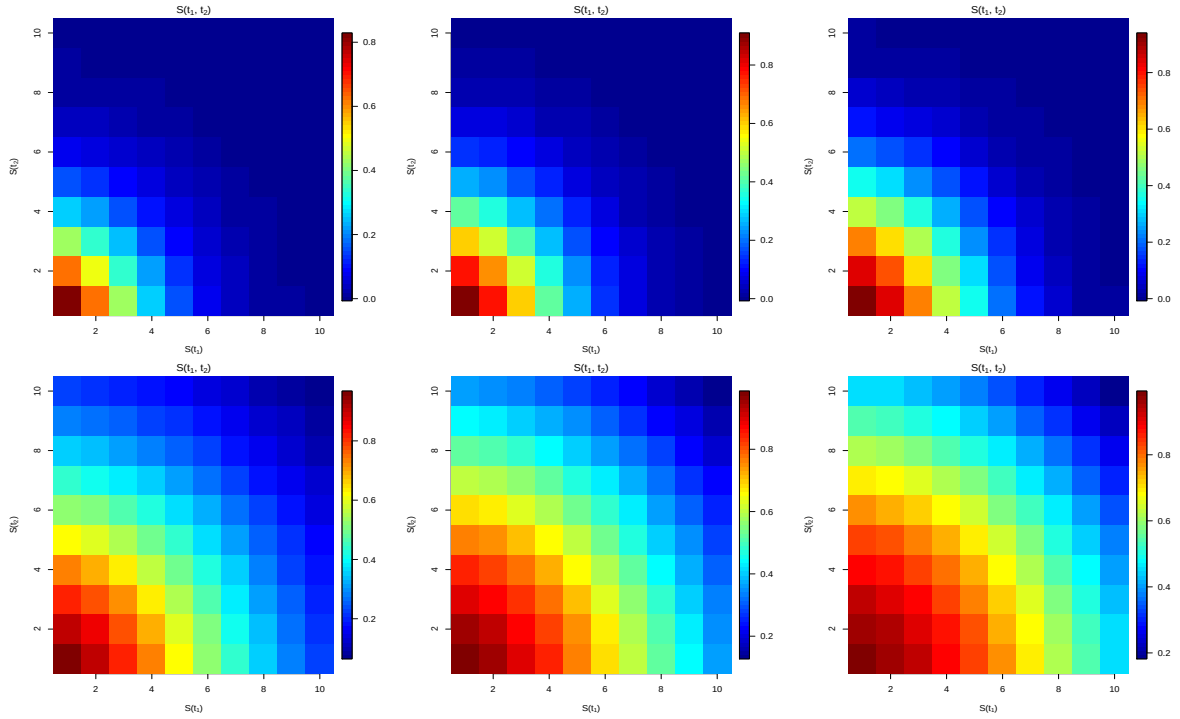

Figure 1 – Discrete contour plots of the joint survival function for DBGR model assuming different parameter values (**Upper-panels:** fixed values given by $\lambda_1 = \lambda_2 = 0.95$, and $\alpha = 0.50 \to 1.00 \to 1.50$. **Lower-panels:** fixed values given by $\lambda_1 = \lambda_2 = 0.99$, and $\alpha = 0.50 \to 1.00 \to 1.50$).

where $k_i = \sqrt{x_i^2 + (x_i+1)^2}, i = 1, 2$; $\bar{\alpha} = 1 - \alpha$ *and* $DR(x_i, \lambda_i) = \lambda_i^{x_i^2} - \lambda_i^{(x_i+1)^2}$, $i = 1, 2$ *denotes a univariate discrete Rayleigh distribution. Observe that* (3.10) *is a proper joint pmf by using the fact that the series expressed by* $\sum_{x_1=0}^{\infty}\sum_{x_2=0}^{\infty} h(x_1, x_2)$ *converges to* $\frac{1}{4}[\vartheta_3(0, \lambda_1) + 1][\vartheta_3(0, \lambda_2) + 1]$ *for* $z_1, z_2 \in \mathbb{R}_+$ *where* $\vartheta_a(x, q)$ *is the Jacobi theta function.*

3.3.1.3 Mathematical Properties

For this distribution, the marginal distributions of X_1 and X_2 are given by discrete generalized Rayleigh (DGR) distributions with corresponding parameters (λ_1, α) and (λ_2, α), respectively. Since the DGR inherits most of the properties of the continuous model in the univariate case, these marginals have great flexibility for the hazard function, given by bathtub, increasing and increasing-decreasing-increasing shapes depending on the parameter values. Many properties of the continuous generalized Rayleigh distribution were studied by MirMostafaee et al. (2017). The marginal survival and marginal pmf functions can be expressed by,

$$\Pr(X_i > x_i) = \frac{\alpha\lambda_i^{x_i^2}}{1 - \bar{\alpha}\lambda_i^{x_i^2}} \quad \text{and} \quad \Pr(X_i = x_i) = \frac{\alpha\left(\lambda_i^{x_i^2} - \lambda_i^{(x_i+1)^2}\right)}{\left(1 - \bar{\alpha}\lambda_i^{(x_i+1)^2}\right)\left(1 - \bar{\alpha}\lambda_i^{x_i^2}\right)}, \; i = 1, 2. \tag{3.11}$$

The expected value and variance for the marginal probability distributions do not have closed forms, however it could be computed using numerical methods from the mean and variance around the origin given by,

$$\mathbb{E}[X_i] = \alpha\left(\sum_{x_i=1}^{\infty}\frac{\lambda_i^{x_i^2}}{1-\bar{\alpha}\lambda_i^{x_i^2}}\right) \quad \text{and} \quad \text{Var}(X_i) = \sum_{x_i=1}^{\infty}\frac{(2x_i^2-1)\alpha\lambda_i^{x_i^2}}{1-\bar{\alpha}\lambda_i^{x_i^2}} - \left(\sum_{x_i=1}^{\infty}\frac{\alpha\lambda_i^{x_i^2}}{1-\bar{\alpha}\lambda_i^{x_i^2}}\right)^2 \tag{3.12}$$

which could be approximated using the *finite series* given by:

$$\sum_{x=1}^{\infty}\frac{\lambda_i^{x^2}}{1-\bar{\alpha}\lambda_i^{x^2}} = \sum_{x=1}^{\infty}\frac{d}{d\alpha}\log(1-\bar{\alpha}\lambda_i^{x^2}) \approx \sum_{j=1}^{M}(-1)^{j+1}(\alpha-1)^{(j-1)^2}\frac{\lambda_i^{j^2}}{1-\lambda_i^{j^2}}. \tag{3.13}$$

Now, let us assume the transformation of the random variables X_1 and X_2 given by $W = \min(X_1, X_2)$. In this case, the cumulative function of W is given by,

$$\Pr(W < w) = 1 - \Pr(W > w) = 1 - \frac{\alpha\lambda_1^{w^2}\lambda_2^{w^2}}{1-\bar{\alpha}\lambda_1^{w^2}\lambda_2^{w^2}} = 1 - \frac{\alpha(\lambda_1\lambda_2)^{w^2}}{1-\bar{\alpha}(\lambda_1\lambda_2)^{w^2}}$$

which implies that the distribution of W is a discrete generalized Rayleigh (DGR) distributions with corresponding parameters $(\lambda_1\lambda_2, \alpha)$. The expected value and variance could also be approximated by the finite series given in (3.13).

On other hand, for the DBGR given by (3.10), the cross factorial moment between X_1 and X_2 is given by,

$$\mu_{X_1,X_2} = \mathbb{E}[X_1X_2] = \sum_{x_1=1}^{\infty}\sum_{x_2=1}^{\infty}\frac{\alpha\lambda_1^{x_1^2}\lambda_2^{x_2^2}}{1-\bar{\alpha}\lambda_1^{x_1^2}\lambda_2^{x_2^2}}. \tag{3.14}$$

which is a monotonic increasing function of λ_1, λ_2 and α since,

$$\frac{\partial\mu_{X_1,X_2}}{\partial\alpha} = \alpha\sum_{x_1=1}^{\infty}\sum_{x_2=1}^{\infty}\frac{(1-\lambda_1^{x_1^2}\lambda_2^{x_2^2})\lambda_1^{x_1^2}\lambda_2^{x_2^2}}{(1-\bar{\alpha}\lambda_1^{x_1^2}\lambda_2^{x_2^2})^2} > 0$$

and,

$$\frac{\partial\mu_{X_1,X_2}}{\partial\lambda_i} = \alpha\sum_{x_1=1}^{\infty}\sum_{x_2=1}^{\infty}\frac{x_i^2\lambda_i^{x_i^2-1}\lambda_j^{x_j^2}}{(1-\bar{\alpha}\lambda_i^{x_i^2}\lambda_j^{x_j^2})^2} > 0$$

for $i, j = 1, 2; i \neq j$. However, the cross factorial also could be approximated using series representation of the logarithmic function. That is, suppose that M is an integer sufficiently large and $|(\alpha-1)\lambda_i x_i^2| < 1, i = 1, 2$, thus we have,

$$\mu_{X_1,X_2} = \sum_{x_1=1}^{\infty}\sum_{x_2=1}^{\infty}\frac{d}{d\alpha}\log(1-\bar{\alpha}\lambda_1^{x_1^2}\lambda_2^{x_2^2}) \approx \sum_{j=1}^{M}(\alpha-1)^{(j-1)^2}\frac{\lambda_1^{j^2}\lambda_2^{j^2}}{(1-\lambda_1^{j^2})(1-\lambda_2^{j^2})}$$

which is a finite series and can be determined when the parameters of the distribution are estimated from a dataset.

Theorem 2.3.1.3.1. (Covariance signal) *Given the function*

$$\Psi(\alpha) = \left(1 - \bar{\alpha}\lambda_1^{x_1^2}\right)\left(1 - \bar{\alpha}\lambda_2^{x_2^2}\right) - \alpha\left(1 - \bar{\alpha}\lambda_1^{x_1^2}\lambda_2^{x_2^2}\right) \tag{3.15}$$

then $\Psi(\alpha) > 0$ *if* $0 < \alpha < 1$, $\Psi(\alpha) = 1$ *if* $\alpha = 0$ *and* $\Psi(\alpha) < 0$ *if* $\alpha > 1$. *That is, the covariance signal only depends on the parameter* α.

Proof. Note that the $\Psi(\alpha)$ is a continuous function on α and its first derivative is given by,

$$\Psi'(\alpha) = -\left(1 - \lambda_1^{x_1^2}\right)\left(1 - \lambda_2^{x_2^2}\right) < 0, \quad \text{for all } \alpha > 0.$$

Note that, since $\Psi(0) = \left(1 - \lambda_1^{x_1^2}\right)\left(1 - \lambda_2^{x_2^2}\right) > 0$ and $\Psi(1) = 0$, then $\Psi(\alpha) > 0$ if $0 < \alpha < 1$. On the other hand, if $\alpha = 1$ then $\Psi(\alpha) = 0$; if $\alpha > 1$ then $\Psi(\alpha) < 0$ and the proof is complete. □

From the function $\Psi(\alpha)$ given in (3.15), the proposed model admits a very flexible correlation coefficient ρ of any sign. In fact it is observed that if $\Psi(\alpha) > 0$, $\rho > 0$ and if $\Psi(\alpha) < 0, \rho < 0$. If $\Psi(\alpha) = 0$ the correlation coefficient is equal to zero. Although there is no closed form for the expressions of the covariance and correlation coefficient this could be computed by taking a large number of terms in the series. In Table 1, it is illustrated a numerical experiment for the covariance and correlation coefficients for some values of α, λ_1 and λ_2 from where it could be seen the flexibility of the correlation coefficient ρ.

Table 1 – Theoretical dependence measures assuming a DBGR distribution.

$(\alpha, \lambda_1, \lambda_2)$	$\mathbb{E}[X_1X_2]$	$\text{cov}(X_1, X_2)$	ρ
(0.2, 0.2, 0.2)	0.0084	0.0061	0.1281
(0.8, 0.2, 0.2)	0.0327	0.0046	0.0310
(0.5, 0.5, 0.5)	0.1777	0.0433	0.0980
(1.0, 0.5, 0.5)	0.3186	0.0000	0.0000
(1.2, 0.5, 0.5)	0.3676	-0.0191	-0.0267
(1.5, 0.5, 0.5)	0.4348	-0.0466	-0.0579
(1.8, 0.8, 0.8)	2.6964	-0.2225	-0.0298
(2.0, 0.8, 0.8)	2.8597	-0.2721	-0.0342

Finally, for the proposed DBGR model, the conditional distribution of X_j given $X_i, i, j = 1, 2$ and $i \neq j$ are given by,

$$\begin{aligned} \text{P}(X_j \mid X_i = x_i) &= \frac{[1 - \bar{\alpha}^2 h(k_i, k_j)]\,[1 - \bar{\alpha} h(x_i, x_j)]^{-1}\,[1 - \bar{\alpha}(DR(x_i, \lambda_i) + \bar{\alpha}\lambda_i^{(x_i+1)^2}\lambda_i^{x_i^2})]}{[1 - \bar{\alpha}\, h(x_i + 1, x_j)]\,[1 - \bar{\alpha}\, h(x_i, x_j + 1)]} \\ &\times \frac{[DR(x_j, \lambda_j)]}{[1 - \bar{\alpha}\, h(x_i + 1, x_j + 1)]} \end{aligned} \tag{3.16}$$

where $DR(x_i, \lambda_i) = \lambda_i^{x_i^2} - \lambda_i^{(x_i+1)^2}$, $i = 1, 2$ denotes a univariate discrete Rayleigh distribution.

Proposition 2.3.1.3.1. (Special cases) *Some especial cases of the DBGR are given by,*

i.) If $\lambda_1 = \lambda_2$*, the pmf of DBGR is symmetric in its arguments, that is,* $P(X_1 = x_1, X_2 = x_2) = P(X_2 = x_2, X_1 = x_1)$ *for all* $x_1, x_2 \in \mathbb{N}$.

ii.) If $\lambda_1 = \lambda_2 = \alpha = \lambda$*, the pmf of DBGR is also symmetric in its arguments and its reduced to a one parameter bivariate discrete distribution with probability mass function given by,*

$$\frac{\lambda^{x_1+x_2+1}[(1-\lambda)^2 - (1-\lambda)^4\lambda^{x_1^2+x_2^2+(x_1+1)^2+(x_2+1)^2}]}{(1-\lambda^{x_1+x_2+1})(1-\lambda^{x_1+x_2+2})(1-\lambda^{x_1+x_2+3})} \tag{3.17}$$

iii.) If $0 < \alpha < 1$ *and* $\lambda_1 = \lambda_2 = \lambda$*, the pmf of DBGR can be rewritten as an infinite mixture of the product of two Rayleigh distributions.*

Proof. (i) and (ii) are trivial. For (iii), the result is obtained by considering the series representation,

$$(1-b)^{-k} = \sum_{j=0}^{\infty} \frac{\Gamma(k+j)}{\Gamma(j+1)\Gamma(k)} b^j,\ |b| < 1,\ k > 0 \tag{3.18}$$

and noticing that $\mathrm{P}(X_1 > x_1, X_2 > x_2) = \dfrac{\alpha\lambda^{x_1^2+x_2^2}}{1-\bar{\alpha}\lambda^{x_1^2+x_2^2}} = \alpha\lambda^{x_1^2+x_2^2}\sum_{j=0}^{\infty}(\bar{\alpha}\lambda^{x_1^2+x_2^2})^j.$ □

Definition 2.3.1.3.2. (Multivariate extension) *Let* $X_i, i = 1, \ldots, n$ *be* n *independent discrete random variables having the Rayleigh distribution with parameters* $0 < \lambda_i < 1, i = 1, \ldots, n$*. Since* $X_i, i = 1, \ldots, n$ *are independent, the multivariate survival and multivariate pmf functions are given, respectively by,*

$$P(X_1 > x_1, \ldots, X_n > x_n) = \frac{\alpha\lambda_1^{x_1^2} \ldots \lambda_n^{x_n^2}}{1-\bar{\alpha}\lambda_1^{x_1^2} \ldots \lambda_n^{x_n^2}} \tag{3.19}$$

and,

$$\begin{aligned} P(X_1 = x_1, \ldots, X_n = x_n) &= \frac{\alpha\prod_{i=1}^{n}\lambda_i^{x_i^2}}{1-\bar{\alpha}\prod_{i=1}^{n}\lambda_i^{x_i^2}} - \frac{\alpha\lambda_1^{x_1^2}\prod_{i=2}^{n}\lambda_i^{(x_i+1)^2}}{1-\bar{\alpha}\lambda_1^{x_1^2}\prod_{i=2}^{n}\lambda_i^{(x_i+1)^2}} + \ldots + \\ &+ \ldots + (-1)^n \frac{\alpha\prod_{i=1}^{n}\lambda_i^{(x_i+1)^2}}{1-\bar{\alpha}\prod_{i=1}^{n}\lambda_i^{(x_i+1)^2}} \end{aligned} \tag{3.20}$$

For the multivariate model, the dependence between the n lifetimes is specified by the parameter α, where if $\alpha = 1$ there is independence between the n lifetimes. The model

adequacy could be checked by comparisons of the fitted marginal survival functions with the empirical estimates of the marginal survival distributions since the marginal survival functions are discrete generalized Rayleigh distributions as well. Moreover, the correlation coefficient for the multivariate case has the same properties of the correlation coefficient of the bivariate model, that is, it could be negative, positive or zero. Since it is needed extensive calculations for the multivariate case, their properties will not be derived here.

3.3.1.4 Inference Methods

Now, let $(X_{11}, X_{21}), (X_{12}, X_{22}), \ldots, (X_{1n}, X_{2n})$ be a random sample of size n from a DBGR distribution. The log-likelihood function $\ell(\lambda_1, \lambda_2, \alpha)$ is given by:

$$\begin{aligned}
\ell(\lambda_1, \lambda_2, \alpha) &= (n_1 + n_2)\log(\alpha) + \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \log\left(1 - \bar{\alpha}^2 \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}\right) \\
&+ \sum_{s=1}^{n_1} \log\left(\lambda_1^{x_{1s}^2} - \lambda_1^{(x_{1s}+1)^2}\right) + \sum_{t=1}^{n_2} \log\left(\lambda_2^{x_{2t}^2} - \lambda_2^{(x_{2t}+1)^2}\right) \\
&- \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \log\left(1 - \bar{\alpha}\lambda_1^{(x_{1s}+1)^2} \lambda_2^{x_{2t}^2}\right) - \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \log\left(1 - \bar{\alpha}\lambda_1^{x_{1s}^2} \lambda_2^{(x_{2t}+1)^2}\right) \\
&- \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \log\left(1 - \bar{\alpha}\lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2}\right) - \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \log\left(1 - \bar{\alpha}\lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}\right).
\end{aligned} \tag{3.21}$$

The normal equations (partial derivatives of the log–likelihood functions equating to zero) obtained from (3.21), are not reproduced here as they cannot be solved explicitly. They must be solved either by numerical methods as, for example, the Newton-Rapshon optimization method or by directly maximizing the log-likelihood function. Since the global maximum of the log–likelihood surface is not guaranteed, different initial values in the parameter space can be considered as a seed point. From the log-likelihood, the first derivatives of $\ell(\lambda_1, \lambda_2, \alpha)$ with respect to λ_1, λ_2 and α are given, respectively by,

i.) $$\begin{aligned}
\frac{\partial \ell}{\partial \lambda_1} &= \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\left(\bar{\alpha}(x_{1s}+1)^2 \lambda_1^{(x_{1s}+1)^2-1} \lambda_2^{(x_{2t}+1)^2}\right)}{\left(1 - \bar{\alpha}\lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}\right)} + \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\left(\bar{\alpha} x_{1s}^2 \lambda_1^{x_{1s}^2-1} \lambda_2^{(x_{2t}+1)^2}\right)}{\left(1 - \bar{\alpha}\lambda_1^{x_{1s}^2} \lambda_2^{(x_{2t}+1)^2}\right)} \\
&+ \sum_{s=1}^{n_1} \frac{\left(x_{1s}^2 \lambda_1^{x_{1s}^2-1} - (x_{1s}+1)^2 \lambda_1^{(x_{1s}+1)^2-1}\right)}{\left(\lambda_1^{x_{1s}^2} - \lambda_1^{(x_{1s}+1)^2}\right)} + \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\left(\bar{\alpha}(x_{1s}+1)^2 \lambda_1^{(x_{1s}+1)^2-1} \lambda_2^{x_{2t}^2}\right)}{\left(1 - \bar{\alpha}\lambda_1^{(x_{1s}+1)^2} \lambda_2^{x_{2t}^2}\right)} \\
&+ \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\left(\bar{\alpha} x_{1s}^2 \lambda_1^{x_{1s}^2-1} \lambda_2^{x_{2t}^2}\right)}{\left(1 - \bar{\alpha}\lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2}\right)} - \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\bar{\alpha}^2 x_{1s}^2 \lambda_1^{x_{1s}^2-1} \lambda_2^{x_{2t}^2} (x_{1s}+1)^2 \lambda_1^{(x_{1s}+1)^2-1} \lambda_2^{(x_{2t}+1)^2}}{\left(1 - \bar{\alpha}^2 \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}\right)}
\end{aligned}$$

ii.) $$\frac{\partial \ell}{\partial \lambda_2} = \sum_{t=1}^{n_2} \frac{\left(x_{2t}^2 \lambda_2^{x_{2t}^2-1} - (x_{2t}+1)^2 \lambda_2^{(x_{2t}+1)^2-1}\right)}{\left(\lambda_2^{x_{2t}^2} - \lambda_2^{(x_{2t}+1)^2}\right)} + \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\left(\bar{\alpha} x_{2t}^2 \lambda_1^{(x_{1s}+1)^2} \lambda_2^{x_{2t}^2-1}\right)}{\left(1 - \bar{\alpha} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{x_{2t}^2}\right)}$$
$$+ \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\left(\bar{\alpha}(x_{2t}+1)^2 \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2-1}\right)}{\left(1 - \bar{\alpha} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}\right)} + \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\left(\bar{\alpha}(x_{2t}+1)^2 \lambda_1^{x_{1s}^2} \lambda_2^{(x_{2t}+1)^2-1}\right)}{\left(1 - \bar{\alpha} \lambda_1^{x_{1s}^2} \lambda_2^{(x_{2t}+1)^2}\right)}$$
$$+ \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\left(\bar{\alpha} x_{2t}^2 \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2-1}\right)}{\left(1 - \bar{\alpha} \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2}\right)} - \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\bar{\alpha}^2 x_{2t}^2 \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2-1} (x_{2t}+1)^2 \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2-1}}{\left(1 - \bar{\alpha}^2 \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}\right)}$$

iii.) $$\frac{\partial \ell}{\partial \alpha} = \frac{n_1+n_2}{\alpha} + \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{2\bar{\alpha} \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}}{\left(1 - \bar{\alpha}^2 \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}\right)}$$
$$+ \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}}{\left(1 - \bar{\alpha} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{(x_{2t}+1)^2}\right)} + \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\lambda_2^{(x_{2t}+1)^2} \lambda_1^{x_{1s}^2}}{\left(1 - \bar{\alpha} \lambda_2^{(x_{2t}+1)^2} \lambda_1^{x_{1s}^2}\right)}$$
$$+ \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\lambda_1^{(x_{1s}+1)^2} \lambda_2^{x_{2t}^2}}{\left(1 - \bar{\alpha} \lambda_1^{(x_{1s}+1)^2} \lambda_2^{x_{2t}^2}\right)} + \sum_{s=1}^{n_1}\sum_{t=1}^{n_2} \frac{\lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2}}{\left(1 - \bar{\alpha} \lambda_1^{x_{1s}^2} \lambda_2^{x_{2t}^2}\right)}$$

Under standard asymptotic maximum likelihood theory, a consistent estimator for the covariance matrix of $(\widehat{\lambda_1}, \widehat{\lambda_2}, \widehat{\alpha})$ is obtained by the inverse of the Fisher information of $(\lambda_1, \lambda_2, \alpha)$, evaluated at $(\lambda_1, \lambda_2, \alpha) = (\widehat{\lambda_1}, \widehat{\lambda_2}, \widehat{\alpha})$. In this case, the Fisher information could be approximate by the second derivatives of the log-likelihood function with respect to λ_1, λ_2 and α locally at the obtained MLE's, that is,

$$I_0(\widehat{\lambda_1}, \widehat{\lambda_2}, \widehat{\alpha}) = \begin{pmatrix} -\frac{\partial^2 \ell}{\partial^2 \lambda_1^2} & -\frac{\partial^2 \ell}{\partial^2 \lambda_1 \lambda_2} & -\frac{\partial^2 \ell}{\partial^2 \lambda_1 \alpha} \\ -\frac{\partial^2 \ell}{\partial^2 \lambda_2 \lambda_1} & -\frac{\partial^2 \ell}{\partial^2 \lambda_2^2} & -\frac{\partial^2 \ell}{\partial^2 \lambda_2 \alpha} \\ -\frac{\partial^2 \ell}{\partial^2 \alpha \lambda_1} & -\frac{\partial^2 \ell}{\partial^2 \alpha \lambda_2} & -\frac{\partial^2 \ell}{\partial^2 \alpha^2} \end{pmatrix}_{\widehat{\lambda_1}, \widehat{\lambda_2}, \widehat{\alpha}} \tag{3.22}$$

Hypothesis testing and confidence intervals for λ_1, λ_2 and α could be obtained by using the asymptotically normality of the MLEs $\widehat{\lambda_1}, \widehat{\lambda_2}$ and $\widehat{\alpha}$ and the observed Fisher information matrix I_0, that is,

$$(\widehat{\lambda_1}, \widehat{\lambda_2}, \widehat{\alpha}) \sim N[(\widehat{\lambda_1}, \widehat{\lambda_2}, \widehat{\alpha}), I_0^{-1}] \tag{3.23}$$

On other hand, in many applications related to lifetime data, it is common the presence of censored data, that could be right, left or interval censoring. In this section, let us assume the presence of right censored data, that is, associated to each lifetime $X_j, j = 1, 2$, there is a fixed censoring time C_j and the data are given by $T_1 = \min(X_1, C_1)$ and

$T_2 = \min(X_2, C_2)$. The likelihood function for the parameters of the DBGR distribution has the dataset classified in four regions:

- C_1: Both, X_{1i} and X_{2i}, are complete observations;
- C_2: X_{1i} are complete and X_{2i} are censored;
- C_3: X_{1i} are censored and X_{2i} are complete;
- C_4: Both, X_{1i} and X_{2i}, are censored observations.

Thus, the likelihood function for λ_1, λ_2 and α based on n bivariate observations $\mathbf{t}_i = (t_{1i}, t_{2i}), i = 1, 2, \ldots, n$ is given by,

$$\begin{aligned} L(\lambda_1, \lambda_2, \alpha) &= \prod_{i \in C_1} \mathrm{P}(T_{1i} = t_{1i}, T_{2i} = t_{2i}) \times \prod_{i \in C_2} \mathrm{P}(T_{1i} = t_{1i}, T_{2i} > t_{2i}) \\ &\times \prod_{i \in C_3} \mathrm{P}(T_{1i} > t_{1i}, T_{2i} = t_{2i}) \times \prod_{i \in C_4} \mathrm{P}(T_{1i} > t_{1i}, T_{2i} > t_{2i}) \end{aligned} \tag{3.24}$$

where $\mathrm{P}(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = \mathrm{P}(T_{1i} > t_{1i} - 1, T_{2i} > t_{2i}) - \mathrm{P}(T_{1i} > t_{1i}, T_{2i} > t_{2i})$ and $\mathrm{P}(T_{1i} > t_{1i}, T_{2i} = t_{2i}) = \mathrm{P}(T_{1i} > t_{1i}, T_{2i} > t_{2i} - 1) - \mathrm{P}(T_{1i} > t_{1i}, T_{2i} > t_{2i})$. Let us also to define the following indicator variables of censoring,

$$\begin{cases} \delta_{1i} = 1 \quad \text{if} \quad X_{1i} < C_{1i} \quad \text{and 0, for the other part.} \\ \delta_{2i} = 1 \quad \text{if} \quad X_{2i} < C_{2i} \quad \text{and 0, for the other part.} \end{cases} \tag{3.25}$$

where $i = 1, 2, \ldots, n; (C_{1i}, C_{2i})$ are the right censoring times. From equation (3.25), it is obtained the following results for the likelihood function:

(a) For $\delta_{1i} = \delta_{2i} = 1$, it is obtained,

$$\mathrm{P}(T_{1i} = t_{1i}, T_{2i} = t_{2i}) = \left\{ \frac{\left[1 - \bar{\alpha}^2 h\left(\sqrt{t_{1i}^2 + (t_{1i}+1)^2}, \sqrt{t_{2i}^2 + (t_{2i}+1)^2}\right)\right]}{\left[1 - \bar{\alpha}\, h(t_{1i}+1, t_{2i})\right]\left[1 - \bar{\alpha}\, h(t_{1i}, t_{2i}+1)\right]} \right.$$
$$\left. \times \frac{\alpha DR(t_{1i}, \lambda_1) DR(t_{2i}, \lambda_2)}{\left[1 - \bar{\alpha}\, h(t_{1i}+1, t_{2i}+1)\right]\left[1 - \bar{\alpha} h(t_{1i}, t_{2i})\right]} \right\}^{\delta_{1i}\delta_{2i}}$$

(b) For $\delta_{1i} = 1, \delta_{2i} = 0$, it is obtained,

$$\mathrm{P}(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = \left\{ \frac{\alpha h(t_{1i}, t_{2i}+1)}{1 - \bar{\alpha} h(t_{1i}, t_{2i}+1)} - \frac{\alpha h(t_{1i}, t_{2i})}{1 - \bar{\alpha} h(t_{1i}, t_{2i})} \right\}^{\delta_{1i}(1-\delta_{2i})}$$

(c) For $\delta_{1i} = 0, \delta_{2i} = 1$, it is obtained,

$$\mathrm{P}(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = \left\{ \frac{\alpha h(t_{1i}+1, t_{2i})}{1 - \bar{\alpha} h(t_{1i}+1, t_{2i})} - \frac{\alpha h(t_{1i}, t_{2i})}{1 - \bar{\alpha} h(t_{1i}, t_{2i})} \right\}^{\delta_{2i}(1-\delta_{1i})}$$

(d) For $\delta_{1i} = 0, \delta_{2i} = 0$, it is obtained,

$$\mathrm{P}(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = \left\{ \frac{\alpha h(t_{1i}, t_{2i})}{1 - \bar{\alpha} h(t_{1i}, t_{2i})} \right\}^{(1-\delta_{2i})(1-\delta_{1i})}$$

3.3.1.5 A Simulation Study

This section reports the results of a simulation study carried out to assess the performance of the MLEs of the DBGR model assuming complete data. The simulation study was performed using the library **maxLik** from the R software and considering the **BFGS** optimization method. To simulate observations from DBGR model, the marginal distribution of X_1 and the conditional distribution of X_2 given X_1 were used following the steps:

- Step 1: Generate $U_1 \sim Uniform(0, 1)$ and $U_2 \sim Uniform(0, 1)$;
- Step 2: Generate a value x_1 of X_1 from the marginal distribution of X_1 using the inverse transformation method;
- Step 3: Generate a value x_2 of X_2 using the inverse transformation method again based on the conditional distribution of X_2 given $X_1 = x_1$;
- Step 4: Return $X = (X_1, X_2)$.

It was performed the simulation study under three scenarios considering the following parameter values assumed for better computational stability: $(\lambda_1, \lambda_2, \alpha) = (\lambda_1,\ 0.90,\ 2.00)$ where $\lambda_1 = 0.40, 0.50, 0.70, 0.80$ for the first scenario; $(\lambda_1, \lambda_2, \alpha) = (0.95, \lambda_2, 2.50)$ where $\lambda_2 = 0.40, 0.50, 0.70, 0.80$ for the second scenario; and $(\lambda_1, \lambda_2, \alpha) = (0.95, 0.97, \alpha)$ where $\alpha = 0.50, 1.00, 1.50, 2.00$ for the third scenario. It was also considered the sample sizes $n = 10, \ldots, 100$, each one involving 10,000 Monte Carlo replications.

For each scenario, the mean of the 10,000 estimated parameter component of the vector of parameters $(\lambda_1, \lambda_2, \alpha)$, the biases and the RMSE were computed using the expressions:

$$BIAS(\widehat{\mathbf{\Psi}}) = \frac{1}{N}\sum_{i=1}^{N}(\widehat{\mathbf{\Psi}}_i - \mathbf{\Psi}),\ RMSE(\widehat{\mathbf{\Psi}}) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widehat{\mathbf{\Psi}}_i - \mathbf{\Psi})^2}$$

where $N = 10,000$ is the number of simulations and $\mathbf{\Psi}$ denotes each parameter λ_1, λ_2 or α. The obtained simulation results for each scenario are illustrated, respectively, in Figures 2, 3 and 4.

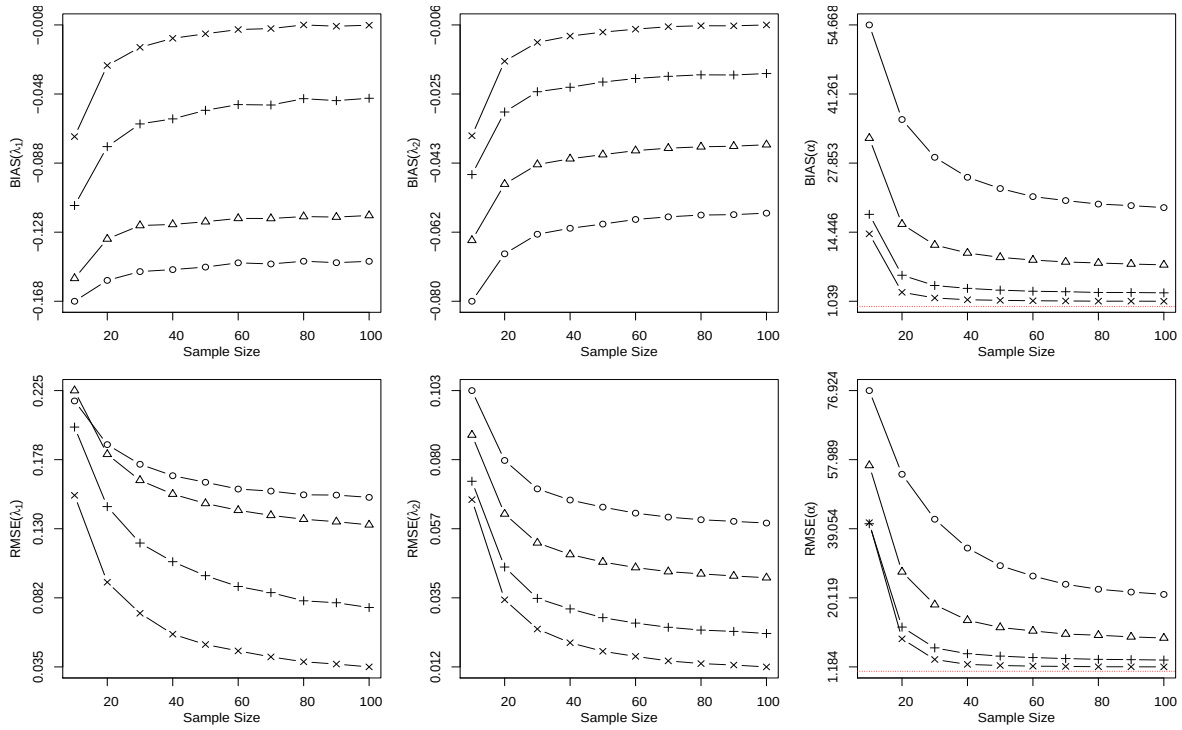

Figure 2 – The biases (upper panels) and RMSEs (lower panels) for the DBGR distribution assuming $(\lambda_1, \lambda_2, \alpha) = (\lambda_1, 0.90, 2.00)$ where $\circ: \lambda_1 = 0.40; \triangle: \lambda_1 = 0.50; +: \lambda_1 = 0.60; \times: \lambda_1 = 0.70$.

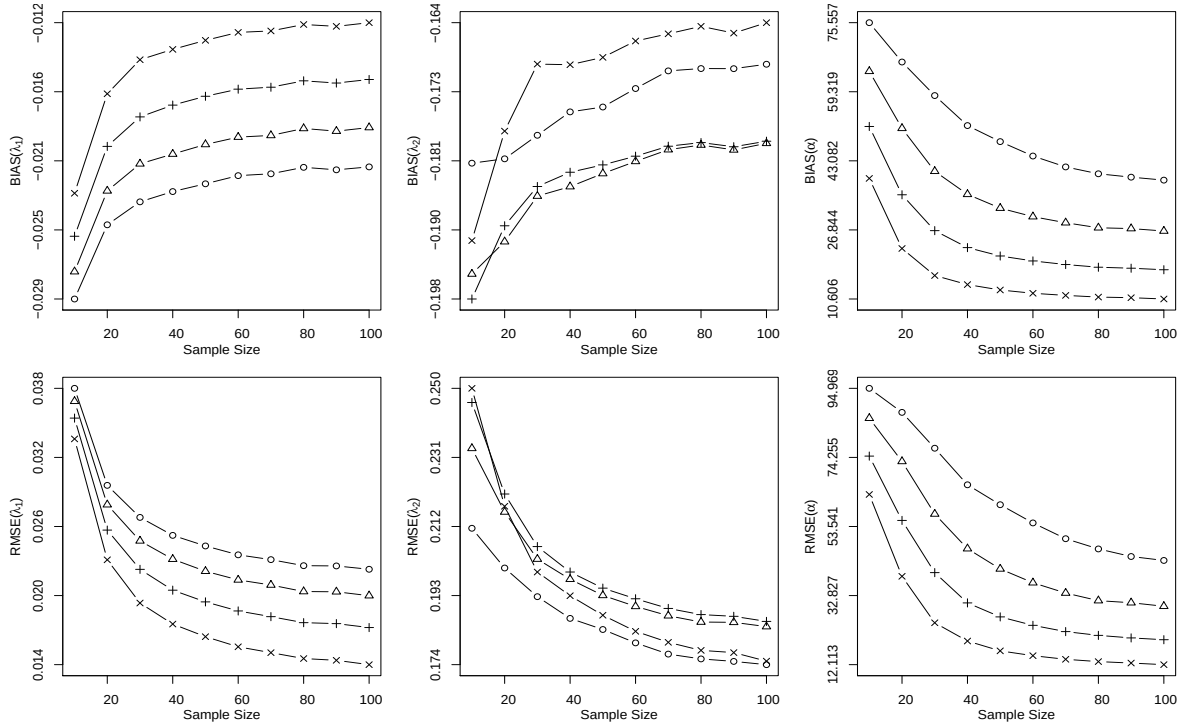

Figure 3 – The biases (upper panels) and RMSEs (lower panels) for the DBGR distribution assuming $(\lambda_1, \lambda_2, \alpha) = (0.95, \lambda_2, 2.50)$ where $\circ: \lambda_2 = 0.35; \triangle: \lambda_2 = 0.40; +: \lambda_2 = 0.45; \times: \lambda_2 = 0.50$.

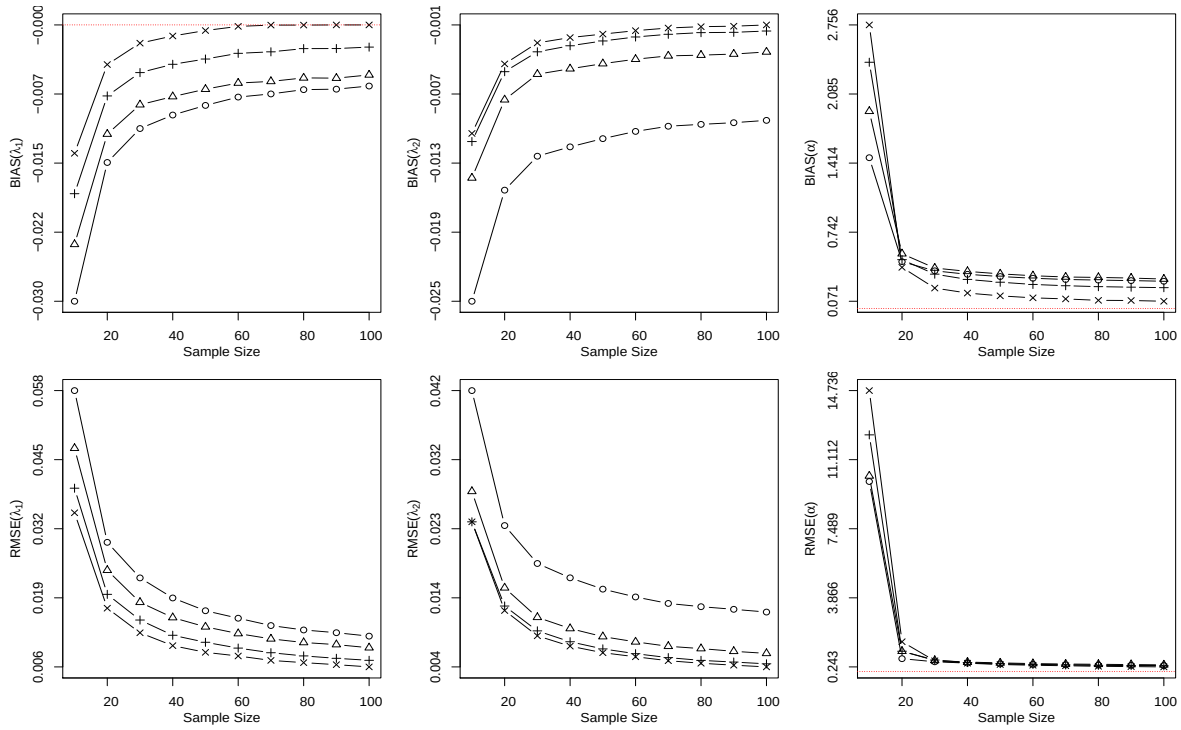

Figure 4 – The biases (upper panels) and RMSEs (lower panels) for the DBGR distribution assuming $(\lambda_1, \lambda_2, \alpha) = (0.95, 0.97, \alpha)$ where $\circ : \alpha = 0.50; \triangle : \alpha = 0.75; + : \alpha = 1.00; \times : \alpha = 1.25$.

From the simulation results illustrated in Figures 2, 3 and 4, it is possible to conclude that,

i.) For all considered scenarios, the biases and RMSEs tends to zero when the sample size increases. The convergence to zero is much faster in the third scenario (λ_1, λ_2 fixed);

ii.) The parameters λ_1 and λ_2 have negative biases for all considered scenarios. The parameter α has a positive bias for all scenarios;

iii.) The parameters λ_1 and λ_2 have small values for the biases and RMSEs; however the parameter α has high values for the biases and RMSEs;

iv.) The smallest values for the biases were obtained for the third scenario; the smallest values for the RMSE were also obtained for the third scenario;

v.) From the simulations results, it is concluded that the the DBGR distribution has better asymptotically non-biased estimation in the third scenario since $E(\lambda_i) \approx \lambda_i, i = 1, 2$ and $E(\alpha) \approx \alpha$ in this scenario; for the other scenarios, it is needed a sample size $n > 100$ to obtain the results $E(\lambda_i) \approx \lambda_i$ and $E(\alpha) \approx \alpha$;

vi.) It is important to point out that the simulation study also could be made using a Bayesian approach with different prior distributions for the parameters of the DBGR distribution.

vii.) Based on these simulation results, it could be concluded that the DBGR distribution could be used as an alternative to other existing discrete bivariate distributions (such as the Basu-Dhar bivariate geometric distribution introduced by Basu and Dhar, 1995) to describe bivariate lifetimes with good accuracy in applications.

3.3.2 A Class of Bivariate Lindley Distributions

3.3.2.1 Introduction

The Lindley probability distribution was introduced in the literature under a Bayesian context (see Lindley, 1958; Ghitany et al., 2008), as a counter example of fiducial statistics. For many years, it has been used in compound processes linked to a Poisson distribution (see Sankaran, 1970). A continuous random variable X is said to have a Lindley distribution if its probability density function (pdf) can be written as,

$$f_L(x) = \frac{\beta^2}{1+\beta}(1+x)\mathrm{e}^{-\beta x}, \tag{3.26}$$

where $x > 0$ and $\beta > 0$ is the scale parameter. The expression (3.26) also could be written as a mixture of two distributions with components given, respectively, by an Exponential density, $f_1(x) = \beta e^{-\beta x}$, and a Gamma density, $f_2(x) = x\beta^2 e^{-\beta x}$. The incidence probabilities of each component in the mixture are given, respectively by $\frac{\beta}{1+\beta}$ and $\frac{1}{1+\beta}$.

A comprehensive discussion on the mathematical properties of the Lindley distribution, such as moments, hazard function, stochastic orderings, parameter estimation, among others is presented by Ghitany et al. (2008). The corresponding survival function (sf) is given by,

$$S_L(x) = \left(1 + \frac{\beta x}{1+\beta}\right)\mathrm{e}^{-\beta x}. \tag{3.27}$$

Another uniparametric distribution usually used in reliability studies with similar form of (3.26) is the well-known Exponential distribution with pdf and sf given, respectively, by,

$$f_E(x) = \beta \mathrm{e}^{-\beta x} \quad \text{and} \quad S_E(x) = \mathrm{e}^{-\beta x}. \tag{3.28}$$

Although the Exponential distribution is very popular in reliability analysis, the Lindley distribution has some mathematical properties more flexible than those of the Exponential distribution, as for example, obtaining inference results for the stress-strength parameter considered by Al-Mutairi et al. (2013) or the non-constant hazard rate which is a great motivation for its use for a better approach when compared to the Exponential distribution.

3.3.2.2 Lindley Models Based on Shock Model

As a first model, let us assume that the random variables W_1 and W_2 have Lindley distributions with parameters β_1 and β_2, respectively, while the random variable W_3 has an Exponential distribution with parameter β_3. In order to investigate the joint sf of the random variables X_1 and X_2, the following theorem presents the joint sf $S(x_1, x_2)$ related to this distribution.

Theorem 3.3.2.2.1. *The joint sf, $S(x_1, x_2)$, of X_1 and X_2 is given by,*

$$S(x_1, x_2) = \frac{(1 + \beta_1 + \beta_1 x_1)(1 + \beta_2 + \beta_2 x_2)}{(1 + \beta_1)(1 + \beta_2)} \exp\{-\beta_1 x_1 - \beta_2 x_2 - \beta_3 z\}, \tag{3.29}$$

where $z = \max(x_1, x_2)$.

Proof. Since the joint sf of X_1 and X_2 is defined as $S(x_1, x_2) = P(X_1 > x_1, X_2 > x_2)$ then,

$$\begin{aligned} S(x_1, x_2) &= P(\{\min(W_1, W_3) > x_1\}, \{\min(W_2, W_3) > x_2\}) \\ &= P(\{W_1 > x_1, W_3 > x_1\}, \{W_2 > x_2, W_3 > x_2\}) \\ &= P(W_1 > x_1, W_2 > x_2, W_3 > \max(x_1, x_2)). \end{aligned} \tag{3.30}$$

As the random variables $W_j, (j = 1, 2, 3)$ are mutually independent, we get,

$$\begin{aligned} S(x_1, x_2) &= P(W_1 > x_1)P(W_2 > x_2)P(W_3 > z) \\ &= S_1(x_1)S_2(x_2)S_3(z), \end{aligned} \tag{3.31}$$

where $z = \max(x_1, x_2)$. Replacing the respective sf of the Lindley and Exponential probability distributions into the above relation, one can obtain the joint sf (3.29). Hence, the proof is complete. □

Let us denote the model defined above as a bivariate Lindley distribution of type I (BL-I). From the plots in Figure 5, we can notice that the BL-I distribution has a continuous part and a singular part which is an expected result since this distribution is an extension of the Marshall-Olkin bivariate Exponential distribution.

From the joint sf (3.29), the marginal sf for $X_j (j = 1, 2)$ of the BL-I distribution are directly obtained by observing that

$$S_{X_j}(x_j) = S_i(x_j)S_3(x_j) = \left(1 + \frac{\beta_j x_j}{1 + \beta_j}\right) \mathrm{e}^{-\beta_j x_j - \beta_3 x_j}, \tag{3.32}$$

and the marginal pdf of $X_j, (j = 1, 2)$ are given by,

$$f_{X_j}(x_j) = \frac{\mathrm{e}^{-(\beta_j + \beta_3)x_j}}{1 + \beta_j} \left[(\beta_j^2 + \beta_3\beta_j)(1 + x_j) + \beta_3\right]. \tag{3.33}$$

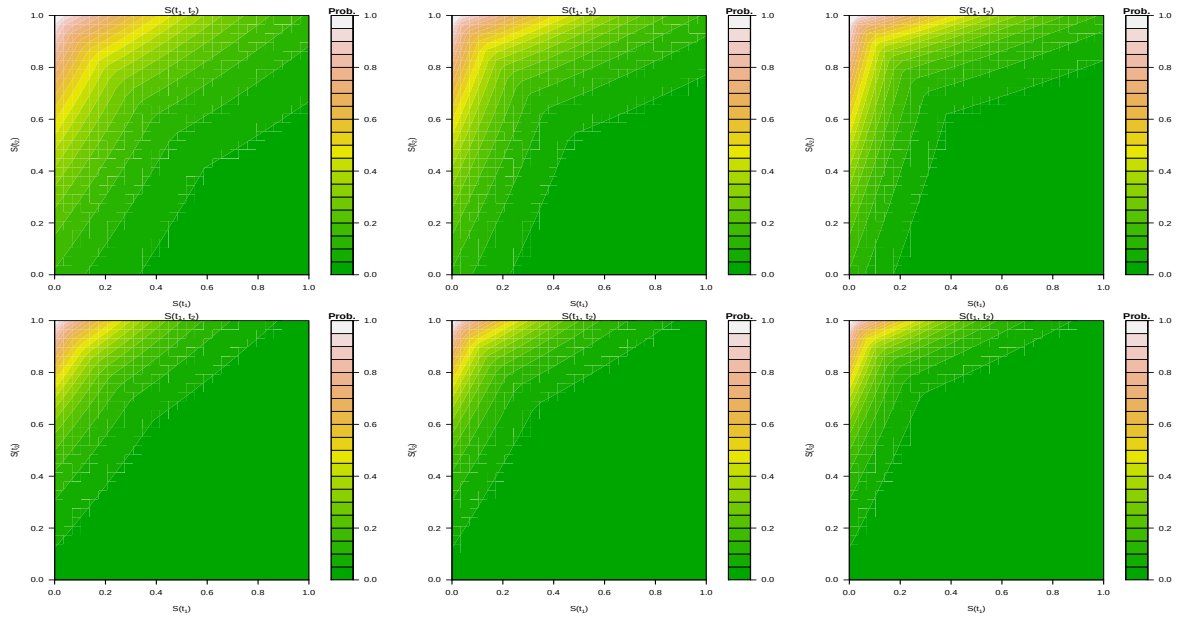

Figure 5 – Contour plots of the joint survival function for BL-I model assuming different parameter values (**Upper-panels:** fixed values given by $\beta_1 = \beta_2 = 0.95$, and $\beta_3 = 0.50 \rightarrow 1.00 \rightarrow 1.50$. **Lower-panels:** fixed values given by $\beta_1 = \beta_2 = 1.50$, and $\beta_3 = 0.50 \rightarrow 1.00 \rightarrow 1.50$).

The joint distribution function $F(x_1, x_2)$ of X_1 and X_2 can be directly obtained from the relationship given by $F(x_1, x_2) = 1 - S_{X_1}(x_1) - S_{X_2}(x_2) + S(x_1, x_2)$, that is,

$$\begin{aligned} F(x_1, x_2) &= 1 - \left(1 + \frac{\beta_1 x_1}{1+\beta_1}\right) \mathrm{e}^{-(\beta_1+\beta_3)x_1} - \left(1 + \frac{\beta_2 x_2}{1+\beta_2}\right) \mathrm{e}^{-(\beta_2+\beta_3)x_2} \\ &+ \left[\frac{(1+\beta_1+\beta_1 x_1)(1+\beta_2+\beta_2 x_2)}{(1+\beta_1)(1+\beta_2)} \exp\{-\beta_1 x_1 - \beta_2 x_2 - \beta_3 z\}\right]. \quad (3.34) \end{aligned}$$

Theorem 3.3.2.2.2. *If the joint sf, $S(x_1, x_2)$, of X_1 and X_2 is given by,*

$$S(x_1, x_2) = \frac{(1+\beta_1+\beta_1 x_1)(1+\beta_2+\beta_2 x_2)}{(1+\beta_1)(1+\beta_2)} \exp\{-\beta_1 x_1 - \beta_2 x_2 - \beta_3 z\}, \quad (3.35)$$

where $z = \max(x_1, x_2)$, then the joint probability density function, $f(x_1, x_2)$ of X_1 and X_2 is given by,

$$f(x_1, x_2) = \begin{cases} f_1(x_1, x_2) = \left[\dfrac{\beta_1(1+x_1)}{1+\beta_1} f_E(x_1, \beta_1+\beta_3) + \dfrac{\beta_3}{1+\beta_1} S_E(x_1, \beta_1+\beta_3)\right] f_L(x_2, \beta_2) \\ \quad \textit{if } x_1 > x_2 \\ f_2(x_1, x_2) = \left[\dfrac{\beta_2(1+x_2)}{1+\beta_2} f_E(x_2, \beta_2+\beta_3) + \dfrac{\beta_3}{1+\beta_2} S_E(x_2, \beta_2+\beta_3)\right] f_L(x_1, \beta_1) \\ \quad \textit{if } x_1 < x_2 \\ f_3(x, x) = \beta_3 e^{-\beta_3 x} \left\{ \left(1 + \dfrac{\beta_1 x}{1+\beta_1}\right) \times \left(1 + \dfrac{\beta_2 x}{1+\beta_2}\right) e^{-(\beta_1+\beta_2)x} \right\} \\ \quad \textit{if } x_1 = x_2 = x, \end{cases}$$

(3.36)

where $f_L(\cdot), f_E(\cdot)$ are, respectively, the pdf of the Lindley and Exponential distributions and $S_E(\cdot)$ is the sf of the Exponential distribution.

Proof. Let us first assume that $x_1 > x_2$. In this case, the joint sf becomes,

$$S(x_1, x_2) = \frac{(1+\beta_1+\beta_1 x_1)(1+\beta_2+\beta_2 x_2)}{(1+\beta_1)(1+\beta_2)} \exp\{-\beta_1 x_1 - \beta_2 x_2 - \beta_3 x_1\}. \quad (3.37)$$

Thus, upon differentiation, one can obtain the expression of the joint probability density function for X_1 and X_2

$$f(x_1, x_2) = \frac{\partial^2 S(x_1, x_2)}{\partial x_1 \partial x_2},$$

to be the $f_1(x_1, x_2)$ given in (3.36). Analogously, it is obtained the expression to be $f_2(x_1, x_2)$ when $x_1 < x_2$. However, $f_3(x, x)$ cannot be derived in a similar way. In this case, it is used the following identity,

$$\begin{aligned} 1 &= \int_0^\infty f_3(x, x)\, dx + \int_0^\infty \int_0^{x_1} f_1(x_1, x_2)\, dx_2 dx_1 \\ &+ \int_0^\infty \int_0^{x_2} f_2(x_1, x_2)\, dx_1 dx_2. \end{aligned} \quad (3.38)$$

In (3.38), the double integrals are given, respectively, by,

$$\begin{aligned} K_1 &= \int_0^\infty \int_0^{x_1} f_1(x_1, x_2)\, dx_2 dx_1 \\ &= \int_0^\infty F_L(x_1, \beta_2) \left[\frac{\beta_1(1+x_1)}{1+\beta_1} f_E(x_1, \beta_1+\beta_3) + \frac{\beta_3}{1+\beta_1} S_E(x_1, \beta_1+\beta_3) \right] dx_1 \\ &= \frac{\beta_1}{(1+\beta_1)} \left[\int_0^\infty f_E(x_1, \beta_1+\beta_3) F_L(x_1, \beta_2)\, dx_1 + \int_0^\infty x_1 f_E(x_1, \beta_1+\beta_3) F_L(x_1, \beta_2)\, dx_1 \right] \\ &+ \frac{\beta_3}{(1+\beta_1)} \int_0^\infty S_E(x_1, \beta_1+\beta_3) F_L(x_1, \beta_2)\, dx_1, \end{aligned}$$

and,

$$\begin{aligned} K_2 &= \int_0^\infty \int_0^{x_1} f_2(x_1, x_2)\, dx_1 dx_2 \\ &= \int_0^\infty F_L(x_2, \beta_1) \left[\frac{\beta_2(1+x_2)}{1+\beta_2} f_E(x_2, \beta_2+\beta_3) + \frac{\beta_3}{1+\beta_2} S_E(x_2, \beta_2+\beta_3) \right] dx_2 \\ &= \frac{\beta_2}{(1+\beta_2)} \left[\int_0^\infty f_E(x_2, \beta_2+\beta_3) F_L(x_2, \beta_1)\, dx_2 + \int_0^\infty x_2 f_E(x_2, \beta_2+\beta_3) F_L(x_2, \beta_1)\, dx_2 \right] \\ &+ \frac{\beta_3}{(1+\beta_2)} \int_0^\infty S_E(x_2, \beta_2+\beta_3) F_L(x_2, \beta_1)\, dx_2. \end{aligned}$$

Using the identity,

$$\int x^m \mathrm{e}^{-\beta x^n}\, dx = -\frac{\Gamma(\gamma, \beta x^n)}{n\beta^\gamma}, \; \gamma = \frac{m+1}{n}, \; \beta \neq 0, \; n \neq 0,$$

and replacing K_1 and K_2 into (3.38) it is obtained,

$$f_3(x,x) = [1+\beta_1+\beta_1 x]\,[1+\beta_2+\beta_2 x]\,\eta \mathrm{e}^{-\beta x},$$

where $\eta = \beta_3/\left[(1+\beta_1)(1+\beta_2)\right]$; $\beta = \beta_1+\beta_2+\beta_3$. Hence, the proof is complete. □

Since it was obtained the expressions for the marginal pdf and the joint pdf for both random variables X_1 and X_2, the conditional probability distributions for the BL-I distribution are directly obtained using the relationship $f(x_j \mid x_k) = f_{X_j,X_k}(x_j,x_k)/f_{X_j}(x_j)$, $j,k = 1,2$ and $j \neq k$. That is, the conditional distributions are given by the following expression:

$$f(x_j \mid x_k) = \begin{cases} f_{(1)}(x_j \mid x_k) = \dfrac{f_L(x_k,\beta_k)(1+\beta_k)\mathrm{e}^{(\beta_k+\beta_3)x_k}}{(\beta_k^2+\beta_3\beta_k)(1+x_k)+\beta_3}\left[\dfrac{\beta_j(1+x_j)}{1+\beta_j} f_E(x_j,\beta_j+\beta_3)\right. \\ \left. + \dfrac{\beta_3}{1+\beta_j} S_E(x_j,\beta_j+\beta_3)\right] \text{ if } x_j > x_k \\ f_{(2)}(x_j \mid x_k) = \dfrac{f_L(x_j,\beta_j)(1+\beta_j)\mathrm{e}^{(\beta_j+\beta_3)x_j}}{(\beta_j^2+\beta_3\beta_j)(1+x_j)+\beta_3}\left[\dfrac{\beta_k(1+x_k)}{1+\beta_k} f_E(x_k,\beta_k+\beta_3)\right. \\ \left. + \dfrac{\beta_3}{1+\beta_k} S_E(x_k,\beta_k+\beta_3)\right] \text{ if } x_j < x_k \\ f_{(3)}(x_j \mid x_k) = \dfrac{\beta_3(1+\beta_k)\mathrm{e}^{\beta_k x_j}}{(\beta_k^2+\beta_3\beta_k)(1+x_j)+\beta_3}\left\{\left(1+\dfrac{\beta_j x_j}{1+\beta_j}\right)\left(1+\dfrac{\beta_k x_j}{1+\beta_k}\right)\right. \\ \left. \times \mathrm{e}^{-(\beta_j+\beta_k)x_j}\right\} \text{ if } x_j = x_k. \end{cases} \tag{3.39}$$

Corollary 3.3.2.2.1. *Assuming that the random variables W_1 and W_2 follow Lindley probability distributions with density (3.26) and parameters β_1 and β_2, respectively, we could assume any continuous probability distribution with positive domain for the latent random variable W_3. The generated bivariate distribution will be a bivariate Lindley distribution of Marshall-Olkin type. For instance, let us assume a Lindley distribution with parameter β_3 for W_3. In this case, the joint sf and joint pdf for the random variables $X_1 = \min(W_1,W_3)$ and $X_2 = \min(W_2,W_3)$ are given, respectively, by,*

$$\begin{aligned} S(x_1,x_2) &= \frac{(1+\beta_1+\beta_1 x_1)(1+\beta_2+\beta_2 x_2)(1+\beta_3+\beta_3 z)}{(1+\beta_1)(1+\beta_2)(1+\beta_3)} \\ &\times \exp\{-\beta_1 x_1 - \beta_2 x_2 - \beta_3 z\}, \end{aligned} \tag{3.40}$$

and,

$$f(x_1,x_2) = \begin{cases} f_1(x_1,x_2) = f_L(x_2,\beta_2)[f_L(x_1,\beta_1)S_L(x_1,\beta_3) + f_L(x_1,\beta_3)S_L(x_1,\beta_1)] \text{ if } x_1 > x_2 \\ f_2(x_1,x_2) = f_L(x_1,\beta_1)[f_L(x_2,\beta_2)S_L(x_2,\beta_3) + f_L(x_2,\beta_3)S_L(x_2,\beta_2)] \text{ if } x_1 < x_2 \\ f_3(x,x) = S_L(x,\beta_1)S_L(x,\beta_2)f_L(x,\beta_3) \text{ if } x_1 = x_2 = x, \end{cases} \tag{3.41}$$

where $z = \max(x_1, x_2)$ *and* $f_L(\cdot)$ *and* $S_L(\cdot)$ *are, respectively, the pdf and the sf of a Lindley distribution. Denote this model as a bivariate Lindley distribution of type II (BL-II).*

3.3.2.3 Lindley Models Based on Stress Model

Now, let us assume that the random variables U_1 and U_2 have Lindley distributions with parameters β_1 and β_2, respectively, while the random variable U_3 has an Exponential distribution with parameter β_3. For this model, the following theorem presents the joint sf $S(x_1, x_2)$ for X_1 and X_2.

Theorem 3.3.2.3.1. *The joint sf,* $S(x_1, x_2)$*, for the random variables* X_1 *and* X_2 *is given by,*

$$S(x_1, x_2) = \frac{(1+\beta_1+\beta_1 x_1)(1+\beta_2+\beta_2 x_2)}{(1+\beta_1)(1+\beta_2)} \exp\{-\beta_1 x_1 - \beta_2 x_2 - \beta_3 z\}, \quad (3.42)$$

where $z = \min(x_1, x_2)$.

Proof. The proof of this theorem is analogous to the proof of Theorem 3.3.2.2.1. □

Let us denote the model with joint sf (3.42) as a bivariate Lindley probability distribution of type III (BL-III). Like the shock model, from Figure 6, it is possible to see that the BL-III distribution also has a continuous part and a singular part which it is also expected since it is also an extension of the Marshall-Olkin bivariate Exponential distribution.

From the joint sf (3.42), the marginal sf for $X_j (j = 1, 2)$ of the BL-II distribution are given directly by,

$$S_{X_j}(x_j) = S_j(x_j) S_3(x_j) = \left(1 + \frac{\beta_j x_j}{1+\beta_j}\right) \mathrm{e}^{-\beta_j x_j - \beta_3 x_j}, \quad (3.43)$$

and the marginal pdf of $X_j, (i = 1, 2)$ are given by,

$$f_{X_j}(x_j) = \frac{\mathrm{e}^{-(\beta_j+\beta_3)x_j}}{1+\beta_j} \left[(\beta_j^2 + \beta_3\beta_j)(1+x_j) + \beta_3\right]. \quad (3.44)$$

The joint cdf $F(x_1, x_2)$ for the random variables X_1 and X_2 can be directly obtained from the relationship given by $F(x_1, x_2) = 1 - S_{X_1}(x_1) - S_{X_2}(x_2) + S(x_1, x_2)$, that is,

$$\begin{aligned} F(x_1, x_2) &= 1 - \left(1 + \frac{\beta_1 x_1}{1+\beta_1}\right) \mathrm{e}^{-(\beta_1+\beta_3)x_1} - \left(1 + \frac{\beta_2 x_2}{1+\beta_2}\right) \mathrm{e}^{-(\beta_2+\beta_3)x_2} \\ &+ \left[\frac{(1+\beta_1+\beta_1 x_1)(1+\beta_2+\beta_2 x_2)}{(1+\beta_1)(1+\beta_2)} \exp\{-\beta_1 x_1 - \beta_2 x_2 - \beta_3 z\}\right] \quad (3.45) \end{aligned}$$

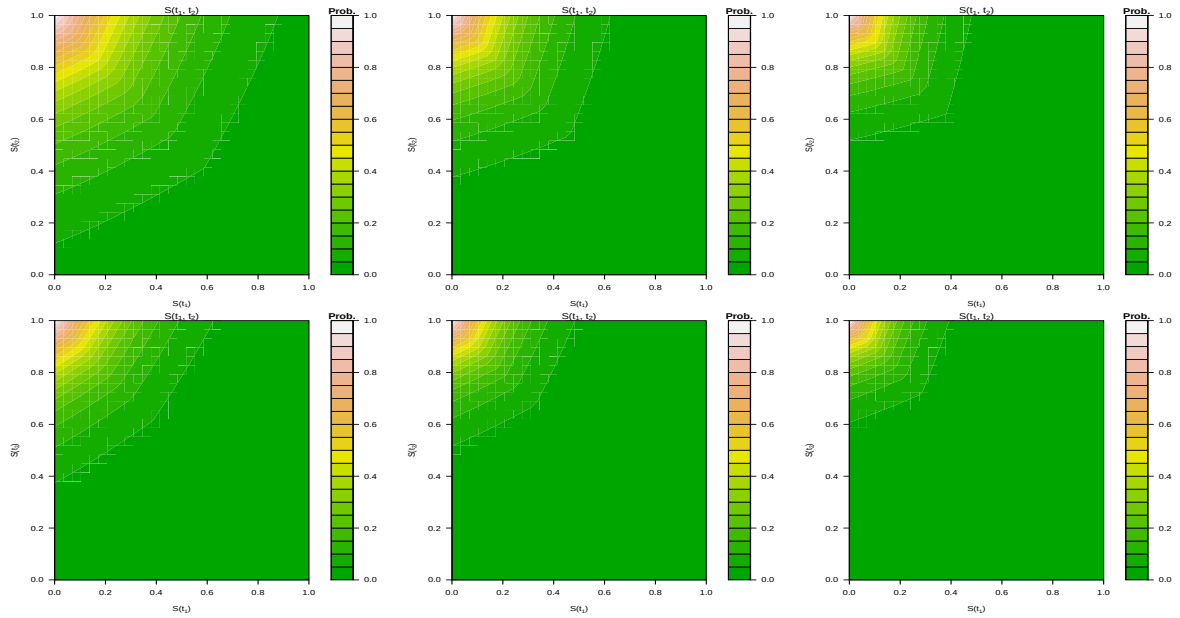

Figure 6 – Contour plots of the joint survival function for BL-III model assuming different parameter values (**Upper-panels:** fixed values given by $\beta_1 = \beta_2 = 0.95$, and $\beta_3 = 0.50 \rightarrow 1.00 \rightarrow 1.50$. **Lower-panels:** fixed values given by $\beta_1 = \beta_2 = 1.50$, and $\beta_3 = 0.50 \rightarrow 1.00 \rightarrow 1.50$).

Theorem 3.3.2.3.2. *From the joint sf (3.42), the joint probability density function, $f(x_1, x_2)$ for the random variables X_1 and X_2 is given by,*

$$f(x_1,x_2) = \begin{cases} f_1(x_1,x_2) = \left[\dfrac{\beta_2(1+x_2)}{1+\beta_2} f_E(x_2, \beta_2+\beta_3) \; + \; \dfrac{\beta_3}{1+\beta_2} S_E(x_2, \beta_2+\beta_3)\right] f_L(x_1,\beta_1) \\ \quad \textit{if } x_1 > x_2 \\ f_2(x_1,x_2) = \left[\dfrac{\beta_1(1+x_1)}{1+\beta_1} f_E(x_1, \beta_1+\beta_3) \; + \; \dfrac{\beta_3}{1+\beta_1} S_E(x_1, \beta_1+\beta_3)\right] f_L(x_2,\beta_2) \\ \quad \textit{if } x_1 < x_2 \\ f_3(x,x) = \beta_3 e^{-\beta_3 x}\left\{\left(1+\dfrac{\beta_1 x}{1+\beta_1}\right) \times \left(1+\dfrac{\beta_2 x}{1+\beta_2}\right) e^{-(\beta_1+\beta_2)x}\right\} \\ \quad \textit{if } x_1 = x_2 = x, \end{cases} \tag{3.46}$$

where $f_L(\cdot), f_E(\cdot)$ and $S_E(\cdot)$ are, respectively, the pdf and sf of the Lindley and Exponential distributions.

Proof. The proof of this theorem is analogous to the proof of Theorem 3.3.2.2.2. □

Since the marginal pdf and the joint pdf of the random variables X_1 and X_2 are known, the conditional probability distributions for the BL-II distribution are directly obtained using the relationship $f(x_j \mid x_k) = f_{X_j,X_k}(x_j, x_k)/f_{X_j}(x_j)$, $j, k = 1, 2$ and $j \neq k$.

That is,

$$f(x_j \mid x_k) = \begin{cases} f_{(1)}(x_j \mid x_k) = \dfrac{f_L(x_j,\beta_j)(1+\beta_j)\mathrm{e}^{(\beta_j+\beta_3)x_j}}{(\beta_j^2+\beta_3\beta_j)(1+x_j)+\beta_3}\left[\dfrac{\beta_k(1+x_k)}{1+\beta_k}f_E(x_k,\beta_k+\beta_3)\right. \\ \left. + \dfrac{\beta_3}{1+\beta_k}S_E(x_k,\beta_k+\beta_3)\right] \text{ if } x_j > x_k \\ f_{(2)}(x_j \mid x_k) = \dfrac{f_L(x_k,\beta_k)(1+\beta_k)\mathrm{e}^{(\beta_k+\beta_3)x_k}}{(\beta_k^2+\beta_3\beta_k)(1+x_k)+\beta_3}\left[\dfrac{\beta_j(1+x_j)}{1+\beta_j}f_E(x_j,\beta_j+\beta_3)\right. \\ \left. + \dfrac{\beta_3}{1+\beta_j}S_E(x_j,\beta_j+\beta_3)\right] \text{ if } x_j < x_k \\ f_{(3)}(x_j \mid x_k) = \dfrac{\beta_3(1+\beta_k)\mathrm{e}^{\beta_k x_j}}{(\beta_k^2+\beta_3\beta_k)(1+x_j)+\beta_3}\left\{\left(1+\dfrac{\beta_j x_j}{1+\beta_j}\right)\left(1+\dfrac{\beta_k x_j}{1+\beta_k}\right)\right. \\ \left. \times \mathrm{e}^{-(\beta_j+\beta_k)x_j}\right\} \text{ if } x_j = x_k. \end{cases} \tag{3.47}$$

Corollary 3.3.2.3.1. *In the same way as it was considered for the shock model, assuming that the the random variables U_1 and U_2 follow a Lindley distribution with parameters β_1 and β_2, it is obtained a new bivariate Lindley distribution that belongs to the new class of bivariate Lindley distributions of Marshall-Olkin type based on a stress model assuming any continuous probability distribution with positive domain for the latent random variable U_3. For instance, let us assume a Lindley distribution with parameter β_3 for U_3. In this case, the joint sf and joint pdf for the random variables $X_1 = \max(U_1, U_3)$ and $X_2 = \max(U_2, U_3)$ are given, respectively, by,*

$$\begin{aligned} S(x_1,x_2) &= \frac{(1+\beta_1+\beta_1x_1)(1+\beta_2+\beta_2x_2)(1+\beta_3+\beta_3z)}{(1+\beta_1)(1+\beta_2)(1+\beta_3)} \\ &\times \exp\{-\beta_1x_1-\beta_2x_2-\beta_3z\}, \end{aligned} \tag{3.48}$$

and,

$$f(x_1,x_2) = \begin{cases} f_1(x_1,x_2) = f_L(x_1,\beta_1)[f_L(x_2,\beta_2)S_L(x_2,\beta_3) + f_L(x_2,\beta_3)S_L(x_2,\beta_2)] \text{ if } x_1 > x_2 \\ f_2(x_1,x_2) = f_L(x_2,\beta_2)[f_L(x_1,\beta_1)S_L(x_1,\beta_3) + f_L(x_1,\beta_3)S_L(x_1,\beta_1)] \text{ if } x_1 < x_2 \\ f_3(x,x) = S_L(x,\beta_1)S_L(x,\beta_2)f_L(x,\beta_3) \text{ if } x_1 = x_2 = x, \end{cases} \tag{3.49}$$

where $z = \min(x_1, x_2)$ and $f_L(\cdot)$ and $S_L(\cdot)$ are the pdf and the sf of a Lindley distribution. Denote this model as a bivariate Lindley distribution of type IV (BL-IV).

3.3.2.4 Correlation Structure

In this section, we present the marginal moments for the proposed models and the correlation structure between X_1 and X_2. Since the process is analogous for all the

proposed models, without loss of generality, only the BL-I model will be considered in this section. In this way, let us assume the BL-I model with marginal distributions given by (3.33). In this case, the expected value and variance for the random variable $X_j, j = 1, 2$ are given, respectively, by

$$\mathbb{E}[X_j] = \frac{\beta_j^2 + (\beta_3 + 2)\beta_j + \beta_3}{(1 + \beta_j)(\beta_j + \beta_3)^2}, \tag{3.50}$$

and,

$$\text{Var}(X_j) = \frac{\beta_j^4 + (2\beta_3 + 4)\beta_j^3}{(1 + \beta_j)^2(\beta_j + \beta_3)^4} + \frac{(\beta_3^2 + 4\beta_3)\beta_j + \beta_3^2}{(1 + \beta_j)^2(\beta_j + \beta_3)^4} + \frac{(\beta_3^2 + 6\beta_3 + 2)\beta_j^2}{(1 + \beta_j)^2(\beta_j + \beta_3)^4}. \tag{3.51}$$

The cross factorial moment between X_1 and X_2 is given by,

$$\begin{aligned}
\mathbb{E}[X_1X_2] &= \frac{1}{\beta^3\beta_{13}^2(1 + \beta_1)(1 + \beta_2)\beta_{23}^2}\Bigg[\{\beta_2^2 + (\beta_3 + 2)\beta_2 + \beta_3\}\beta_1^5 + 3(\beta_2^2 + (\beta_3 + 2)\beta_2 + \beta_3) \\
&\times \left(\beta_2 + \frac{5\beta_3}{3} + \frac{2}{3}\right)\beta_1^4 + \{3\beta_2^4 + (14\beta_3 + 12)\beta_2^3 + (20\beta_3^2 + 42\beta_3 + 12\}\beta_2^2 \\
&+ \{9\beta_3^3 + 40\beta_3^2 + 26\beta_3)\beta_2 + 9\beta_3^3 + 10\beta_3^2\}\beta_1^3 + \{\beta_2^5 + (8\beta_3 + 8)\beta_2^4 + (20\beta_3^2 + 42\beta_3 \\
&+ 12)\beta_2^3 + (20\beta_3^3 + 72\beta_3^2 + 50\beta_3)\beta_2^2 + (7\beta_3^4 + 45\beta_3^3 + 56\beta_3^2)\beta_2 + 7\beta_3^4 + 16\beta_3^3\}\beta_1^2 \\
&+ \{\beta_2^3 + (3\beta_3 + 2)\beta_2^2 + (2\beta_3^2 + 6\beta_3)\beta_2 + 2\beta_3^2\}\beta_{23}((\beta_3 + 2)\beta_2 + \beta_3^2 + 5\beta_3)\beta_1 \\
&+ \{\beta_2^3 + (3\beta_3 + 2)\beta_2^2 + (2\beta_3^2 + 6\beta_3)\beta_2 + 2\beta_3^2\}\beta_3\beta_{23}^2\Bigg],
\end{aligned} \tag{3.52}$$

where $\beta_{13} = \beta_1 + \beta_3$, $\beta_{23} = \beta_2 + \beta_3$ and $\beta = \beta_1 + \beta_2 + \beta_3$. Since the cross factorial moment has a closed form, the covariance and the correlation coefficient between X_1 and X_2 can be directly obtained using the relations:

$$\text{Cov}(X_1, X_2) = \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]$$

and,

$$\rho = \frac{\text{Cov}(X_1, X_2)}{[\text{Var}(X_1)\text{Var}(X_2)]^{1/2}}$$

For BL distributions based on shock model, we could observe that $\mathbb{E}[X_1X_2] > \mathbb{E}[X_1]\mathbb{E}[X_2]$, that is, $0 < \rho < 1$ (positive correlation); and for BL distributions based on stress model, we could observe that $\mathbb{E}[X_1X_2] < \mathbb{E}[X_1]\mathbb{E}[X_2]$, that is, $-1 < \rho < 0$ (negative correlation).

3.3.2.5 Inference Methods

Now, without loss of generality, let us first derive the likelihood for the BL-II distribution (the procedure is analogous for BL-I, BL-II and BL-IV distributions). In this

way, suppose (X_{11}, X_{21}), (X_{12}, X_{22}), ..., (X_{1n}, X_{2n}) is a random sample of size n from a BL-II distribution and define two indicator variables given by $v_{1i} = 1$ if $X_{1i} > X_{2i}$ and 0 otherwise and $v_{2i} = 1$ if $X_{1i} < X_{2i}$ and 0 otherwise for $i = 1, 2, \ldots, n$. In this way, there are three possible situations considering these indicator variables:

$$\begin{aligned}(v_{1i}, v_{2i}) &= (1,0) \quad \text{if} \quad X_{1i} > X_{2i} \\ (v_{1i}, v_{2i}) &= (0,1) \quad \text{if} \quad X_{1i} < X_{2i} \\ (v_{1i}, v_{2i}) &= (0,0) \quad \text{if} \quad X_{1i} = X_{2i}. \end{aligned} \tag{3.53}$$

Thus, from the BL-II pdf and (3.53), the likelihood function for the parameters β_1, β_2 and β_3 assuming complete data is given by,

$$\begin{aligned} L(\boldsymbol{\beta}) &= \prod_{i=1}^{n} [S_L(x_{1i}, \beta_1) S_L(x_{1i}, \beta_2) f_L(x_{1i}, \beta_3)]^{(1-v_{1i})(1-v_{2i})} \prod_{i=1}^{n} f_L(x_{2i}, \beta_2) \big[f_L(x_{1i}, \beta_1) \\ &\times S_L(x_{1i}, \beta_3) + f_L(x_{1i}, \beta_3) S_L(x_{1i}, \beta_1) \big]^{v_{1i}(1-v_{2i})} \prod_{i=1}^{n} f_L(x_{1i}, \beta_1) \big[f_L(x_{2i}, \beta_2) \\ &\times S_L(x_{2i}, \beta_3) f_L(x_{2i}, \beta_3) S_L(x_{2i}, \beta_2) \big]^{v_{2i}(1-v_{1i})}, \end{aligned} \tag{3.54}$$

where $f_L(\cdot)$ and $S_L(\cdot)$ are, respectively, the pdf and sf of the Lindley distribution.

3.3.2.6 A Simulation Study

In this section, it is presented simulated datasets of a connected two-components series system assuming that the systems were put on the life test and the lifetimes of the system were observed until the failures were observed. The simulated datasets were generated from BL distributions considering the sample sizes $n = 20, 50, 100, 150, 300$ and parameter values $\boldsymbol{\beta} = (0.10, 0.20, 0.35)$ using the following algorithm:

1. Generate $S_j \sim Lindley(\beta_j)$ $(j = 1, 2)$ and $S_3 \sim Exponential(\beta_3)$ for BL-I and BL-III distributions. And, for BL-II and BL-IV distributions, generate $S_j \sim Lindley(\beta_j)$ $(j = 1, 2, 3)$.

2. Define $T_1 = \min(S_1, S_3)$ and $T_2 = \min(S_2, S_3)$; return $T = (T_1, T_2)$.

Remark 2.3.2.5.1 *The package* `LindleyR` *(see Mazucheli et al., 2016) is used to generate random values of the Lindley distribution.*

For the analysis of the simulated datasets, it is assumed, as *prior* distributions, approximately noninformative Gamma$(0.01, 0.01)$ *prior* distributions for the parameters β_j $(i = 1, 2)$ and a Gamma *prior* distribution with hyperparameters values obtained from the values of the sample mean $\overline{T}_{\min}$ and $\text{Var}(T_{\min})$ (where $\overline{T}_{\min}$ and $\text{Var}(T_{\min})$ are the sample mean of the data and the sample variance of $T_{\min} = \min(T_1, T_2)$, respectively) for

the parameter β_3. Figure 33 presents the plots of the empirical and the Bayesian estimates for the reliability/survival function $R(t)$ assuming the proposed BL distributions from which we can notice the great accuracy of the obtained Bayesian inferences for the model parameters.

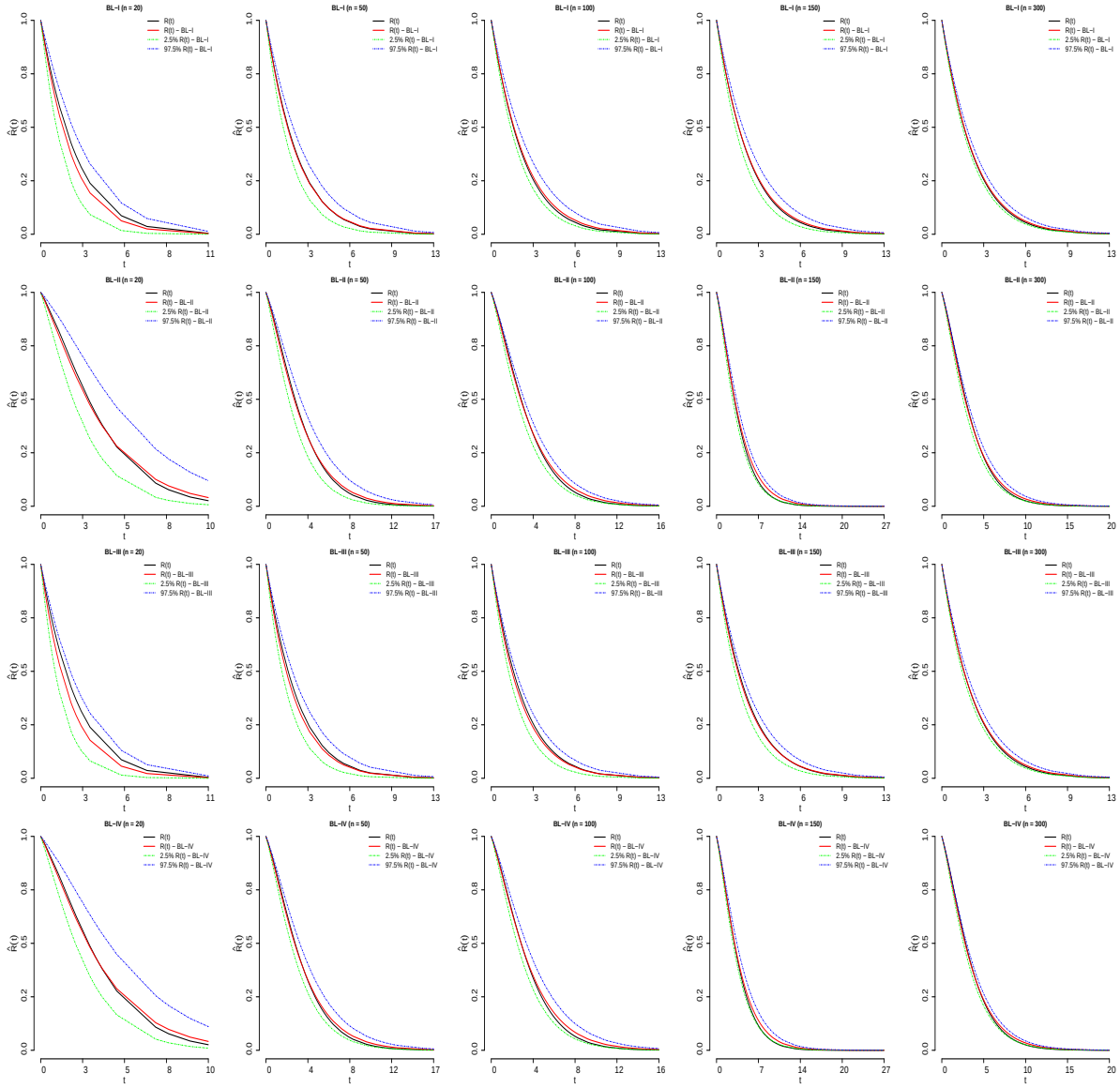

Figure 7 – The mean and 95 percent credible intervals for the reliability/survival function assuming the BL models for each sample size ($n = 20 \rightarrow 300$).

3.4 New Bivariate Cure Rate Models

3.4.1 Gumbel Bivariate Exponential

3.4.1.1 Standard Model

In this section, let us assume the Gumbel bivariate exponential distribution for continuous random variables, denoted by GBE(λ_1,λ_2,θ), introduced by Gumbel (1960) in

reliability analysis. The joint survival function for the lifetimes T_1 and T_2 assuming the GBE model is given by:

$$S_0(t_1, t_2) = \exp\{-\lambda_1 t_1 - \lambda_2 t_2 - \theta\lambda_1\lambda_2 t_1 t_2\} \tag{3.55}$$

where $\lambda_1, \lambda_2 > 0$ and $0 < \theta < 1$. It is important to point out that the GBE model is one of the first models introduced in the literature to model bivariate lifetime data and possible the simplest known exponential bivariate distribution. This model has been applied to many areas, including competing risks, extreme values, failure times, regional analyses of precipitation, and reliability.

From the bivariate survival function described in Equation (3.55), the joint probability density function for the GBE distribution is given by,

$$f_0(t_1, t_2) = [(1-\theta)\lambda_1\lambda_2 + \theta\lambda_1^2\lambda_2 t_1 + \theta\lambda_1\lambda_2^2 t_2 + \theta^2\lambda_1^2\lambda_2^2 t_1 t_2]S(t_1, t_2) \tag{3.56}$$

where the marginal distributions for the random variables T_1 and T_2 are standard exponential univariate distributions with parameters λ_1 and λ_2, respectively. The correlation coefficient between T_1 and T_2 is given by,

$$\rho = 1 - \frac{1}{\theta}\exp\left(\frac{1}{\theta}\right) E_i\left(\frac{1}{\theta}\right) \tag{3.57}$$

where $E_i(x) = \displaystyle\int_{-x}^{\infty} \frac{\exp\{-t\}}{t} dt$ is the exponential integral function. Observe that, the correlation is close to zero when $\theta \to 0$ (the case of independence between T_1 and T_2) and it decreases to -0.40365 as θ increases to 1.

3.4.1.2 Cure Rate Model

Now, let us assume a GBE distribution with parameters λ_1, λ_2 and θ for the bivariate lifetimes T_1 and T_2. From the GBE joint survival function and the equation (2.29), the bivariate mixture cure rate model is given by,

$$\begin{aligned} S(t_1, t_2) &= \phi_{11}\exp\{-\lambda_1 t_1 - \lambda_2 t_2 - \theta\lambda_1\lambda_2 t_1 t_2\} + \phi_{10}\exp\{-\lambda_1 t_1\} \\ &+ \phi_{01}\exp\{-\lambda_2 t_2\} + \phi_{00} \end{aligned} \tag{3.58}$$

where λ_1, λ_2 and $0 < \theta < 1$. The major properties of this model as the joint pdf, conditional distributions and marginal distributions could be directly obtained from the bivariate mixture cure rate model and will not be illustrated here.

For the likelihood-based method of inference for bivariate distributions in presence of censored observations, the contribution of the ith observation for the likelihood function assuming the GBE cure rate model is given by,

i.) If $i \in C_1$,

$$\begin{aligned} f(t_{1i}, t_{2i}) &= \phi_{11}[(1-\theta)\lambda_1\lambda_2 + \theta\lambda_1^2\lambda_2 t_{1i} + \theta\lambda_1\lambda_2^2 t_{2i} + \theta^2\lambda_1^2\lambda_2^2 t_{1i}t_{2i}] \\ &\times \exp\{-\lambda_1 t_{1i} - \lambda_2 t_{2i} - \theta\lambda_1\lambda_2 t_{1i}t_{2i}\} \end{aligned}$$

ii.) If $i \in C_2$,

$$\begin{aligned} -\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} &= \phi_{11}[\lambda_1\lambda_2 t_{2i}\theta + \lambda_1]\exp\{-\lambda_1 t_{1i} - \lambda_2 t_{2i} - \theta\lambda_1\lambda_2 t_{1i}t_{2i}\} \\ &+ \phi_{10}\lambda_1 \exp\{-\lambda_1 t_{1i}\} \end{aligned}$$

iii.) If $i \in C_3$,

$$\begin{aligned} -\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} &= \phi_{11}[\lambda_1\lambda_2 t_{1i}\theta + \lambda_2]\exp\{-\lambda_1 t_{1i} - \lambda_2 t_{2i} - \theta\lambda_1\lambda_2 t_{1i}t_{2i}\} \\ &+ \phi_{01}\lambda_2 \exp\{-\lambda_2 t_{2i}\} \end{aligned}$$

iv.) If $i \in C_4$,

$$\begin{aligned} S(t_{1i}, t_{2i}) &= \phi_{11}\exp\{-\lambda_1 t_{1i} - \lambda_2 t_{2i} - \theta\lambda_1\lambda_2 t_{1i}t_{2i}\} + \phi_{10}\exp\{-\lambda_1 t_{1i}\} \\ &+ \phi_{01}\exp\{-\lambda_2 t_{2i}\} + \phi_{00} \end{aligned}$$

In Figure 8, it is illustrated the contour plots for the survival function expressed in Equation (3.58) considering different parameter values. It is important to point out that the contour plots could be useful for establishing relations of the values of the marginal distributions of the joint survival function, especially in presence of cure rate.

3.4.2 Block and Basu Bivariate Exponential

3.4.2.1 Standard Model

In this section, it is assumed the Block and Basu bivariate exponential distribution also defined for continuous random variables, denoted by BBBE$(\lambda_1, \lambda_2, \lambda_3)$, introduced by Block and Basu (1974). The joint survival function for the BBBE distribution for the random variables T_1 and T_2 is given by,

$$S_0(t_1, t_2) = \begin{cases} \dfrac{\lambda}{\lambda_{12}}\exp\{-\lambda_1 t_1 - \lambda_{23} t_2\} - \dfrac{\lambda_3}{\lambda_{12}}\exp\{-\lambda t_2\} & \text{if} \quad t_1 < t_2 \\ \dfrac{\lambda}{\lambda_{12}}\exp\{-\lambda_{13} t_1 - \lambda_2 t_2\} - \dfrac{\lambda_3}{\lambda_{12}}\exp\{-\lambda t_1\} & \text{if} \quad t_1 \geq t_2 \end{cases} \tag{3.59}$$

where $\lambda = \lambda_1 + \lambda_2 + \lambda_3, \lambda_{12} = \lambda_1 + \lambda_2, \lambda_{13} = \lambda_1 + \lambda_3, \lambda_{23} = \lambda_2 + \lambda_3$ and $\lambda_1, \lambda_2, \lambda_3 > 0$. From the bivariate survival function described in Equation (3.59), the joint probability

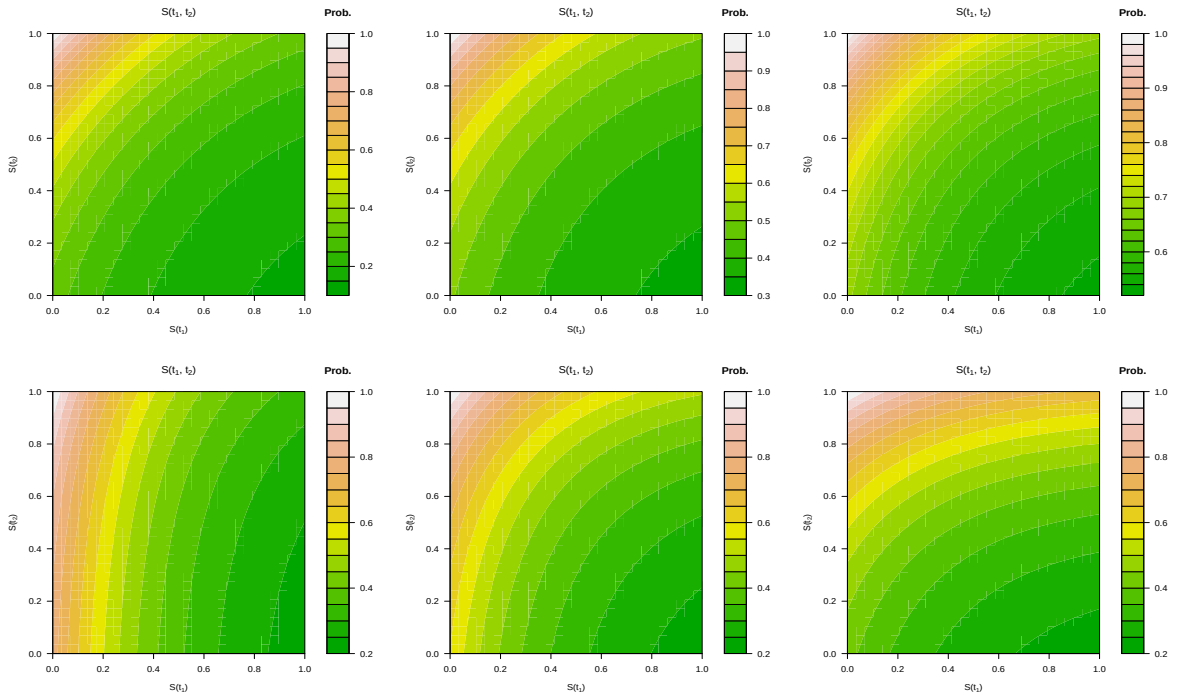

Figure 8 – Contour plots of the joint survival function in presence of cure rate for GBE model assuming different parameter values (**Upper-panels:** fixed values given by $\lambda_1 = \lambda_2 = 0.75, \theta = 0.10, \phi_{10} = \phi_{01} = 0.15$, and $\phi_{11} = (0.60, 0.40, 0.20), \phi_{00} = (0.10, 0.30, 0.50)$. **Lower-panels:** fixed values given by $\lambda_1 = \lambda_2 = 0.75, \theta = 0.10, \phi_{11} = \phi_{00} = 0.15$, and $\phi_{10} = (0.60, 0.40, 0.20), \phi_{01} = (0.10, 0.30, 0.50)$).

density function for the BBBE distribution is given by,

$$f_0(t_1, t_2) = \begin{cases} \dfrac{\lambda\lambda_1\lambda_{23}}{\lambda_{12}} \exp\{-\lambda_1 t_1 - \lambda_{23} t_2\} & \text{if} \quad t_1 < t_2 \\ \dfrac{\lambda\lambda_2\lambda_{13}}{\lambda_{12}} \exp\{-\lambda_{13} t_1 - \lambda_2 t_2\} & \text{if} \quad t_1 \geq t_2. \end{cases} \tag{3.60}$$

Remark 2.4.2.1.1. *Observe that the marginal distributions of the joint distribution defined by Equation (3.60) are not exponentials; however, if $\lambda_3 = 0$, T_1 and T_2 are independent exponential distributions with parameters λ_1 and λ_2, respectively. In other way, note that the marginal distributions for the lifetimes T_1 and T_2 could be also written as mixtures of two exponential distributions (see, Kundu and Gupta, 2010) with densities given, respectively, by,*

$$\begin{aligned} f_{01}(t_1) &= \frac{\lambda}{\lambda_{12}} f_E(t_1, \lambda_{13}) - \frac{\lambda_3}{\lambda_{12}} f_E(t_1, \lambda) \\ f_{02}(t_2) &= \frac{\lambda}{\lambda_{12}} f_E(t_2, \lambda_{23}) - \frac{\lambda_3}{\lambda_{12}} f_E(t_2, \lambda) \end{aligned} \tag{3.61}$$

where $f_E(t, \alpha) = \alpha \exp\{-\alpha t\}$ denotes an exponential density. In this case, the marginal

survival functions are respectively given by,

$$\begin{aligned} S_{01}(t_1) &= \frac{\lambda}{\lambda_{12}} S_E(t_1, \lambda_{13}) - \frac{\lambda_3}{\lambda_{12}} S_E(t_1, \lambda) \\ S_{02}(t_2) &= \frac{\lambda}{\lambda_{12}} S_E(t_2, \lambda_{23}) - \frac{\lambda_3}{\lambda_{12}} S_E(t_2, \lambda) \end{aligned} \tag{3.62}$$

where $S_E(t, \alpha) = \exp\{-\alpha t\}$ *denotes an exponential survival function.*

The correlation coefficient between the lifetimes T_1 and T_2 is given by,

$$\rho = \frac{\lambda\lambda_3\lambda_{12}(1 + \lambda_1^2 + \lambda_2^2)}{\{K(\lambda_2, \lambda_1)\}^{1/2}\{K(\lambda_1, \lambda_2)\}^{1/2}}. \tag{3.63}$$

where $K(\lambda_i, \lambda_j) = \lambda^2\lambda_{12}^2 + \lambda_i\lambda_3(2\lambda_j\lambda + \lambda_i\lambda_3), i \neq j = 1, 2$. The correlation is limited to $0 < \rho \leq 1$, that is, this distribution is useful for bivariate lifetimes with positive correlation.

3.4.2.2 Cure Rate Model

The mixture cure rate model for bivariate lifetimes T_1 and T_2 assuming the BBBE distribution is given by,

$$S(t_1, t_2) = \begin{cases} \phi_{11}\left[\frac{\lambda}{\lambda_{12}}\exp\{-\lambda_1 t_1 - \lambda_{23} t_2\} - \frac{\lambda_3}{\lambda_{12}}\exp\{-\lambda t_2\}\right] + \phi_{10}\left[\frac{\lambda}{\lambda_{12}} S_E(t_1, \lambda_{13}) - \frac{\lambda_3}{\lambda_{12}} S_E(t_1, \lambda)\right] \\ + \phi_{01}\left[\frac{\lambda}{\lambda_{12}} S_E(t_2, \lambda_{23}) - \frac{\lambda_3}{\lambda_{12}} S_E(t_2, \lambda)\right] + \phi_{00} \quad \text{if} \quad t_1 < t_2 \\ \phi_{11}\left[\frac{\lambda}{\lambda_{12}}\exp\{-\lambda_{13} t_1 - \lambda_2 t_2\} - \frac{\lambda_3}{\lambda_{12}}\exp\{-\lambda t_1\}\right] + \phi_{10}\left[\frac{\lambda}{\lambda_{12}} S_E(t_1, \lambda_{13}) - \frac{\lambda_3}{\lambda_{12}} S_E(t_1, \lambda)\right] \\ + \phi_{01}\left[\frac{\lambda}{\lambda_{12}} S_E(t_2, \lambda_{23}) - \frac{\lambda_3}{\lambda_{12}} S_E(t_2, \lambda)\right] + \phi_{00} \quad \text{if} \quad t_1 \geq t_2 \end{cases} \tag{3.64}$$

where $\lambda = \lambda_1 + \lambda_2 + \lambda_3, \lambda_{12} = \lambda_1 + \lambda_2, \lambda_{13} = \lambda_1 + \lambda_3, \lambda_{23} = \lambda_2 + \lambda_3$; $S_E(t, \alpha)$ is the exponential survival function; and $\phi_{11}, \phi_{10}, \phi_{01}, \phi_{00}$ are defined in Equation (2.29). Thus, the contribution of the ith observation for the likelihood function assuming the BBBE cure rate model is given by,

i.) If $i \in C_1$,

$$f(t_{1i}, t_{2i}) = \begin{cases} \phi_{11}\frac{\lambda\lambda_1\lambda_{23}}{\lambda_{12}}\exp\{-\lambda_1 t_{1i} - \lambda_{23} t_{2i}\} & \text{if} \quad t_{1i} < t_{2i} \\ \phi_{11}\frac{\lambda\lambda_2\lambda_{13}}{\lambda_{12}}\exp\{-\lambda_{13} t_{1i} - \lambda_2 t_{2i}\} & \text{if} \quad t_{1i} \geq t_{2i} \end{cases}$$

ii.) If $i \in C_2$,

$$-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} = \begin{cases} \phi_{11}\left[\frac{\lambda\lambda_1}{\lambda_{12}}\exp\{-\lambda_1 t_{1i} - \lambda_{23} t_{2i}\}\right] + \phi_{10}\left[\frac{\lambda}{\lambda_{12}} f_E(t_{1i}, \lambda_{13})\right. \\ \left. - \frac{\lambda_3}{\lambda_{12}} f_E(t_{1i}, \lambda)\right] \quad \text{if} \quad t_{1i} < t_{2i} \\ \phi_{11}\left[\frac{\lambda\lambda_{13}}{\lambda_{12}}\exp\{-\lambda_{13} t_{1i} - \lambda_2 t_{2i}\} - \frac{\lambda\lambda_3}{\lambda_{12}}\exp\{-\lambda t_{1i}\}\right] \\ + \phi_{10}\left[\frac{\lambda}{\lambda_{12}} f_E(t_{1i}, \lambda_{13}) - \frac{\lambda_3}{\lambda_{12}} f_E(t_{1i}, \lambda)\right] \quad \text{if} \quad t_{1i} \geq t_{2i} \end{cases}$$

iii.) If $i \in C_3$,

$$-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} = \begin{cases} \phi_{11}\left[\frac{\lambda\lambda_{23}}{\lambda_{12}}\exp\{-\lambda_1 t_{1i} - \lambda_{23} t_{2i}\} - \frac{\lambda\lambda_3}{\lambda_{12}}\exp\{-\lambda t_{2i}\}\right] \\ + \phi_{01}\left[\frac{\lambda}{\lambda_{12}} f_E(t_{2i}, \lambda_{23}) - \frac{\lambda_3}{\lambda_{12}} f_E(t_{1i}, \lambda)\right] \quad \text{if} \quad t_{1i} < t_{2i} \\ \phi_{11}\left[\frac{\lambda\lambda_2}{\lambda_{12}}\exp\{-\lambda_{13} t_{1i} - \lambda_2 t_{2i}\}\right] + \phi_{01}\left[\frac{\lambda}{\lambda_{12}} f_E(t_{2i}, \lambda_{23})\right. \\ \left. - \frac{\lambda_3}{\lambda_{12}} f_E(t_{2i}, \lambda)\right] \quad \text{if} \quad t_{1i} \geq t_{2i} \end{cases}$$

iv.) If $i \in C_4$,

$$\begin{aligned} S(t_{1i}, t_{2i}) &= \phi_{10}\left[\frac{\lambda}{\lambda_{12}} S_E(t_{1i}, \lambda_{13}) - \frac{\lambda_3}{\lambda_{12}} S_E(t_{1i}, \lambda)\right] + \phi_{11} S_0(t_{1i}, t_{2i}) + \phi_{00} \\ &+ \phi_{01}\left[\frac{\lambda}{\lambda_{12}} S_E(t_{2i}, \lambda_{23}) - \frac{\lambda_3}{\lambda_{12}} S_E(t_{2i}, \lambda)\right] \end{aligned}$$

where $f_E(t, \alpha)$ and $S_E(t, \alpha)$ are the exponential density and survival function, respectively; and $S_0(t_{1i}, t_{2i})$ is given in Equation (3.59).

In Figure 9, it is illustrated the contour plots for the survival function expressed in Equation (3.64) for different parameter values. Observe that the behavior expressed in the contour plots for the BBBE cure rate model is quite similar to the GBE cure rate model assuming the same proportions for the parameters of incidence of cure rate and similar value for the parameters. This remark is important since both models have different mathematical expressions which we can see that the GBE model is simpler than the BBBE model in mathematical expressions.

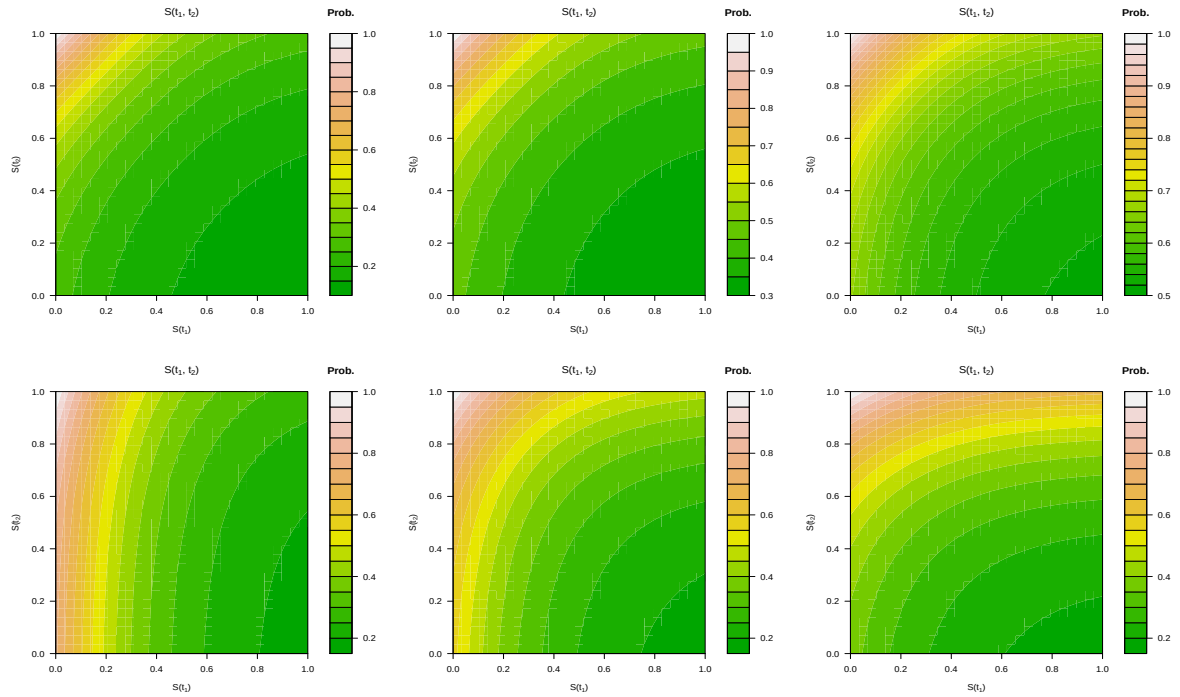

Figure 9 – Contour plots of the joint survival function in presence of cure rate for BBBE model assuming different parameter values (**Upper-panels:** fixed values given by $\lambda_1 = \lambda_2 = 0.75, \lambda_3 = 0.40, \phi_{10} = \phi_{01} = 0.15$, and $\phi_{11} = (0.60, 0.40, 0.20), \phi_{00} = (0.10, 0.30, 0.50)$. **Lower-panels:** fixed values given by $\lambda_1 = \lambda_2 = 0.75, \lambda_3 = 0.40, \phi_{11} = \phi_{00} = 0.15$, and $\phi_{10} = (0.60, 0.40, 0.20), \phi_{01} = (0.10, 0.30, 0.50)$).

3.4.3 Marshall-Olkin Bivariate Exponential

3.4.3.1 Standard Model

In this section, it is assumed the Marshall-Olkin bivariate exponential distribution, denoted by MOBE$(\lambda_1, \lambda_2, \lambda_{12})$, introduced by Marshall and Olkin (1967b) in reliability analysis. The joint survival function for the lifetimes T_1 and T_2 assuming the MOBE distribution is given by,

$$S_0(t_1, t_2) = \exp\{-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12} \max(t_1, t_2)\} \tag{3.65}$$

where $\lambda_1, \lambda_2, \lambda_{12} > 0$. From the bivariate survival function described in Equation (3.65), the joint probability density function for the MOBE distribution is given by,

$$f_0(t_1, t_2) = \begin{cases} \lambda_2(\lambda_1 + \lambda_{12})S(t_1, t_2) & \text{if} \quad t_1 > t_2 \\ \lambda_1(\lambda_2 + \lambda_{12})S(t_1, t_2) & \text{if} \quad t_1 < t_2 \\ \lambda_{12}S(t_1, t_1) & \text{if} \quad t_1 = t_2 \end{cases} \tag{3.66}$$

where the marginal distributions for the random variables T_1 and T_2 are univariate exponential distributions with survival functions given by,

$$S_{01}(t_1) = \exp\{-(\lambda_1 + \lambda_{12})t_1\} \quad \text{and} \quad S_{02}(t_2) = \exp\{-(\lambda_2 + \lambda_{12})t_2\} \tag{3.67}$$

The correlation coefficient between the lifetimes T_1 and T_2 is given by,

$$\rho = \frac{\lambda_{12}}{\lambda_1 + \lambda_2 + \lambda_{12}}. \tag{3.68}$$

Observe that, the correlation is limited to $0 < \rho < 1$, that is, this distribution is useful for bivariate lifetimes with positive correlation.

3.4.3.2 Cure Rate Model

The mixture cure rate model for bivariate lifetimes T_1 and T_2 assuming the MOBE distribution is given by,

$$S(t_1,t_2) = \begin{cases} \phi_{11}\exp\{-(\lambda_1+\lambda_{12})t_1 - \lambda_2 t_2\} + \phi_{10}\exp\{-(\lambda_1+\lambda_{12})t_1\} + \phi_{01}\exp\{-(\lambda_2+\lambda_{12})t_2\} + \phi_{00} \\ \text{if} \quad t_1 > t_2 \\ \phi_{11}\exp\{-(\lambda_2+\lambda_{12})t_2 - \lambda_1 t_1\} + \phi_{10}\exp\{-(\lambda_1+\lambda_{12})t_1\} + \phi_{01}\exp\{-(\lambda_2+\lambda_{12})t_2\} + \phi_{00} \\ \text{if} \quad t_1 < t_2 \\ \phi_{11}\exp\{-(\lambda_1+\lambda_{12}+\lambda_2)t_1\} + \phi_{10}\exp\{-(\lambda_1+\lambda_{12})t_1\} + \phi_{01}\exp\{-(\lambda_2+\lambda_{12})t_1\} + \phi_{00} \\ \text{if} \quad t_1 = t_2 \end{cases} \tag{3.69}$$

where $\lambda_1, \lambda_2, \lambda_{12} > 0$; and $\phi_{11}, \phi_{10}, \phi_{01}, \phi_{00}$ are defined in Equation (2.29). Thus, the contribution of the ith observation for the likelihood function assuming the MOBE cure rate model is given by,

i.) If $i \in C_1$,

$$f(t_{1i},t_{2i}) = \begin{cases} \phi_{11}\lambda_2(\lambda_1+\lambda_{12})S_0(t_{1i},t_{2i}) & \text{if} \quad t_{1i} > t_{2i} \\ \phi_{11}\lambda_1(\lambda_2+\lambda_{12})S_0(t_{1i},t_{2i}) & \text{if} \quad t_{1i} < t_{2i} \\ \phi_{11}\lambda_{12}S(t_{1i},t_{2i}) & \text{if} \quad t_{1i} = t_{2i} \end{cases}$$

ii.) If $i \in C_2$,

$$-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{1i}} = \begin{cases} \phi_{11}(\lambda_1+\lambda_{12})\exp\{-(\lambda_1+\lambda_{12})t_{1i} - \lambda_2 t_{2i}\} + \phi_{10}(\lambda_1+\lambda_{12})\exp\{-(\lambda_1+\lambda_{12})t_{1i}\} \\ \text{if} \quad t_{1i} > t_{2i} \\ \phi_{11}\lambda_1\exp\{-(\lambda_2+\lambda_{12})t_{2i} - \lambda_1 t_{1i}\} + \phi_{10}(\lambda_1+\lambda_{12})\exp\{-(\lambda_1+\lambda_{12})t_{1i}\} \\ \text{if} \quad t_{1i} < t_{2i} \\ \phi_{11}(\lambda_1+\lambda_2+\lambda_{12})\exp\{-(\lambda_1+\lambda_2+\lambda_{12})t_{1i}\} + \phi_{10}(\lambda_1+\lambda_{12})\exp\{-(\lambda_1+\lambda_{12})t_{1i}\} \\ +\phi_{01}(\lambda_2+\lambda_{12})\exp\{-(\lambda_2+\lambda_{12})t_{1i}\} \quad \text{if} \quad t_{1i} = t_{2i} \end{cases}$$

iii.) If $i \in C_3$,

$$-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{2i}} = \begin{cases} \phi_{11}\lambda_2\exp\{-(\lambda_1+\lambda_{12})t_{1i} - \lambda_2 t_{2i}\} + \phi_{01}(\lambda_2+\lambda_{12})\exp\{-(\lambda_2+\lambda_{12})t_{2i}\} \\ \text{if} \quad t_{1i} > t_{2i} \\ \phi_{11}(\lambda_2+\lambda_{12})\exp\{-(\lambda_1+\lambda_{12})t_{2i} - \lambda_1 t_{1i}\} + \phi_{01}(\lambda_2+\lambda_{12})\exp\{-(\lambda_2+\lambda_{12})t_{2i}\} \\ \text{if} \quad t_{1i} < t_{2i} \\ 0 \quad \text{if} \quad t_{1i} = t_{2i} \end{cases}$$

iv.) If $i \in C_4$,

$$S(t_{1i}, t_{2i}) = \begin{cases} \phi_{11}\exp\{-(\lambda_1+\lambda_{12})t_{1i} - \lambda_2 t_{2i}\} + \phi_{10}\exp\{-(\lambda_1+\lambda_{12})t_{1i}\} + \phi_{01}\exp\{-(\lambda_2+\lambda_{12})t_{2i}\} \\ + \phi_{00} \quad \text{if} \quad t_{1i} > t_{2i} \\ \phi_{11}\exp\{-(\lambda_2+\lambda_{12})t_{2i} - \lambda_1 t_{1i}\} + \phi_{10}\exp\{-(\lambda_1+\lambda_{12})t_{1i}\} + \phi_{01}\exp\{-(\lambda_2+\lambda_{12})t_{2i}\} \\ + \phi_{00} \quad \text{if} \quad t_{1i} < t_{2i} \\ \phi_{11}\exp\{-(\lambda_1+\lambda_2+\lambda_{12})t_{1i}\} + \phi_{10}\exp\{-(\lambda_1+\lambda_{12})t_{1i}\} + \phi_{01}\exp\{-(\lambda_2+\lambda_{12})t_{1i}\} \\ + \phi_{00} \quad \text{if} \quad t_{1i} = t_{2i} \end{cases}$$

In Figure 10, it is illustrated the contour plots for the survival function expressed in Equation (3.69) for different parameter values. Observe that the behavior of the MOBE cure rate model has a singular part which is expected since its survival function depends on a $\max(t_1, t_2)$.

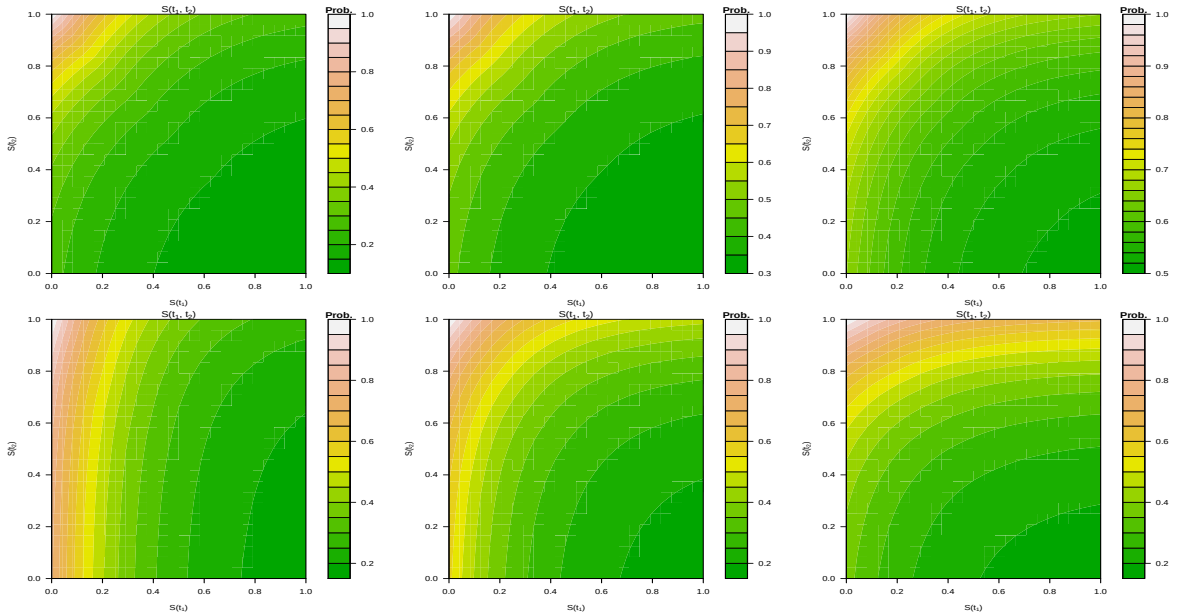

Figure 10 – Contour plots of the joint survival function in presence of cure rate for MOBE model assuming different parameter values (**Upper-panels:** fixed values given by $\lambda_1 = \lambda_2 = 0.75, \lambda_{12} = 0.40, \phi_{10} = \phi_{01} = 0.15$, and $\phi_{11} = (0.60, 0.40, 0.20), \phi_{00} = (0.10, 0.30, 0.50)$. **Lower-panels:** fixed values given by $\lambda_1 = \lambda_2 = 0.75, \lambda_{12} = 0.40, \phi_{11} = \phi_{00} = 0.15$, and $\phi_{10} = (0.60, 0.40, 0.20), \phi_{01} = (0.10, 0.30, 0.50)$).

3.4.4 Bivariate Lindley

In this section, it is assumed the first bivariate Lindley distribution based on a shock model with exponential distribution for W_3 and denoted by BLI$(\beta_1, \beta_2, \beta_3)$; and the third bivariate Lindley distribution based a on stress model with exponential distribution for U_3 and denoted by BLIII$(\beta_1, \beta_2, \beta_3)$ introduced in Section 2.3.

3.4.4.1 Cure Rate Model Based on Shock Model

The mixture cure rate model for bivariate lifetimes T_1 and T_2 assuming the BLI distribution is given by,

$$S(t_1,t_2)=\begin{cases}\phi_{11}\left[\frac{(1+\beta_1+\beta_1t_1)(1+\beta_2+\beta_2t_2)}{(1+\beta_1)(1+\beta_2)}\exp\{-(\beta_1+\beta_3)t_1\}\right]+\phi_{10}\left[\left(1+\frac{\beta_1t_1}{1+\beta_1}\right)\mathrm{e}^{-\beta_1t_1-\beta_3t_1}\right]\\ +\phi_{01}\left[\left(1+\frac{\beta_2t_2}{1+\beta_2}\right)\mathrm{e}^{-\beta_2t_2-\beta_3t_2}\right]+\phi_{00}\quad \text{if}\quad t_1>t_2\\ \phi_{11}\left[\frac{(1+\beta_1+\beta_1t_1)(1+\beta_2+\beta_2t_2)}{(1+\beta_1)(1+\beta_2)}\exp\{-(\beta_2+\beta_3)t_1\}\right]+\phi_{10}\left[\left(1+\frac{\beta_1t_1}{1+\beta_1}\right)\mathrm{e}^{-\beta_1t_1-\beta_3t_1}\right]\\ +\phi_{01}\left[\left(1+\frac{\beta_2t_2}{1+\beta_2}\right)\mathrm{e}^{-\beta_2t_2-\beta_3t_2}\right]+\phi_{00}\quad \text{if}\quad t_1<t_2\\ \phi_{11}\left[\frac{(1+\beta_1+\beta_1t_1)(1+\beta_2+\beta_2t_1)}{(1+\beta_1)(1+\beta_2)}\exp\{-(\beta_1+\beta_2+\beta_3)t_1\}\right]+\phi_{10}\left[\left(1+\frac{\beta_1t_1}{1+\beta_1}\right)\mathrm{e}^{-\beta_1t_1-\beta_3t_1}\right]\\ +\phi_{01}\left[\left(1+\frac{\beta_2t_1}{1+\beta_2}\right)\mathrm{e}^{-\beta_2t_1-\beta_3t_1}\right]+\phi_{00}\quad \text{if}\quad t_1=t_2\end{cases} \tag{3.70}$$

where $\beta_1,\beta_2,\beta_3>0$; and $\phi_{11},\phi_{10},\phi_{01},\phi_{00}$ are defined in Equation (2.29). Thus, the contribution of the ith observation for the likelihood function assuming the BLI cure rate model could be direct obtained assuming the joint pdf:

$$f(t_1,t_2)=\begin{cases}\phi_{11}\left[\frac{\beta_1(1+t_1)}{1+\beta_1}f_E(t_1,\beta_1+\beta_3)+\frac{\beta_3}{1+\beta_1}S_E(t_1,\beta_1+\beta_3)\right]f_L(t_2,\beta_2)\text{ if }t_1>t_2\\ \phi_{11}\left[\frac{\beta_2(1+t_2)}{1+\beta_2}f_E(t_2,\beta_2+\beta_3)+\frac{\beta_3}{1+\beta_2}S_E(t_2,\beta_2+\beta_3)\right]f_L(t_1,\beta_1)\text{ if }t_1<t_2\\ \phi_{11}\beta_3\mathrm{e}^{-\beta_3t}\left\{\left(1+\frac{\beta_1t}{1+\beta_1}\right)\left(1+\frac{\beta_2t}{1+\beta_2}\right)\mathrm{e}^{-(\beta_1+\beta_2)t}\right\}\text{ if }t_1=t_2=t,\end{cases}$$

and the likelihood function based on a random sample of size n, given by,

$$L\ \propto\ \prod_{i\in C_1}f(t_{1i},t_{2i})\prod_{i\in C_2}\left[-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{1i}}\right]\prod_{i\in C_3}\left[-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{2i}}\right]\prod_{i\in C_4}S(t_{1i},t_{2i})$$

3.4.4.2 Cure Rate Model Based on Stress Model

The mixture cure rate model for bivariate lifetimes T_1 and T_2 assuming the BLIII distribution is given by,

$$S(t_1,t_2)=\begin{cases}\phi_{11}\left[\frac{(1+\beta_1+\beta_1t_1)(1+\beta_2+\beta_2t_2)}{(1+\beta_1)(1+\beta_2)}\exp\{-(\beta_1+\beta_3)t_1\}\right]+\phi_{10}\left[\left(1+\frac{\beta_1t_1}{1+\beta_1}\right)\mathrm{e}^{-\beta_1t_1-\beta_3t_1}\right]\\ +\phi_{01}\left[\left(1+\frac{\beta_2t_2}{1+\beta_2}\right)\mathrm{e}^{-\beta_2t_2-\beta_3t_2}\right]+\phi_{00}\quad \text{if}\quad t_1<t_2\\ \phi_{11}\left[\frac{(1+\beta_1+\beta_1t_1)(1+\beta_2+\beta_2t_2)}{(1+\beta_1)(1+\beta_2)}\exp\{-(\beta_2+\beta_3)t_1\}\right]+\phi_{10}\left[\left(1+\frac{\beta_1t_1}{1+\beta_1}\right)\mathrm{e}^{-\beta_1t_1-\beta_3t_1}\right]\\ +\phi_{01}\left[\left(1+\frac{\beta_2t_2}{1+\beta_2}\right)\mathrm{e}^{-\beta_2t_2-\beta_3t_2}\right]+\phi_{00}\quad \text{if}\quad t_1>t_2\\ \phi_{11}\left[\frac{(1+\beta_1+\beta_1t_1)(1+\beta_2+\beta_2t_1)}{(1+\beta_1)(1+\beta_2)}\exp\{-(\beta_1+\beta_2+\beta_3)t_1\}\right]+\phi_{10}\left[\left(1+\frac{\beta_1t_1}{1+\beta_1}\right)\mathrm{e}^{-\beta_1t_1-\beta_3t_1}\right]\\ +\phi_{01}\left[\left(1+\frac{\beta_2t_1}{1+\beta_2}\right)\mathrm{e}^{-\beta_2t_1-\beta_3t_1}\right]+\phi_{00}\quad \text{if}\quad t_1=t_2\end{cases} \tag{3.71}$$

where $\beta_1, \beta_2, \beta_3 > 0$; and $\phi_{11}, \phi_{10}, \phi_{01}, \phi_{00}$ are defined in Equation (2.29). Thus, in the same way as for the BLI cure rate model, the contribution of the ith observation for the likelihood function assuming the BLIII cure rate model could be direct obtained assuming the joint pdf:

$$f(t_1,t_2) = \begin{cases} \phi_{11}\left[\dfrac{\beta_1(1+t_1)}{1+\beta_1} f_E(t_1,\beta_1+\beta_3) + \dfrac{\beta_3}{1+\beta_1} S_E(t_1,\beta_1+\beta_3)\right] f_L(t_2,\beta_2) \text{ if } t_1 < t_2 \\ \phi_{11}\left[\dfrac{\beta_2(1+t_2)}{1+\beta_2} f_E(t_2,\beta_2+\beta_3) + \dfrac{\beta_3}{1+\beta_2} S_E(t_2,\beta_2+\beta_3)\right] f_L(t_1,\beta_1) \text{ if } t_1 > t_2 \\ \phi_{11}\beta_3 \mathrm{e}^{-\beta_3 t}\left\{\left(1+\dfrac{\beta_1 t}{1+\beta_1}\right)\left(1+\dfrac{\beta_2 t}{1+\beta_2}\right)\mathrm{e}^{-(\beta_1+\beta_2)t}\right\} \text{ if } t_1 = t_2 = t, \end{cases}$$

and the likelihood function based on a random sample of size n, given by,

$$L \propto \prod_{i\in C_1} f(t_{1i},t_{2i}) \prod_{i\in C_2}\left[-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{1i}}\right] \prod_{i\in C_3}\left[-\frac{\partial S(t_{1i},t_{2i})}{\partial t_{2i}}\right] \prod_{i\in C_4} S(t_{1i},t_{2i})$$

3.4.5 Bivariate Geometric Distribution Type II

3.4.5.1 Standard Model

In this section, it is assumed the bivariate geometric distribution type II, denoted by BG-Type II(θ_1,θ_2,θ_{12}) proposed by Xiaoling et al. (2012) in reliability analysis. The joint survival function for the lifetimes T_1 and T_2 assuming the BG-Type II distribution is given by,

$$P_0(T_1 > t_1, T_2 > t_2) = \theta_1^{t_1}\theta_2^{t_2}\theta_{12}^{\min(t_1,t_2)} \tag{3.72}$$

where the parameters are $0 < \theta_1\theta_2 < 1$, $0 < \theta_{12} \leq 1$ and satisfy $1-\theta_1-\theta_2+\theta_1\theta_2\theta_{12} \geq 0$. From (3.72), the bivariate probability mass function of the BG-Type II for two discrete random variables T_1 and T_2, is given by,

$$P_0(T_1 = t_1, T_2 = t_2) = \begin{cases} \theta_1^{t_1-1}\theta_2{}^{t_2-1}\theta_{12}^{t_1-1}(1-\theta_2)(1-\theta_1\theta_{12}) & \text{if} \quad t_1 < t_2 \\ \theta_1^{t_1-1}\theta_2{}^{t_2-1}\theta_{12}^{t_2-1}(1-\theta_1)(1-\theta_2\theta_{12}) & \text{if} \quad t_1 > t_2 \\ (\theta_1\theta_2\theta_{12})^{t_1-1}(1-\theta_1\theta_2-\theta_1\theta_2+\theta_1\theta_2\theta_{12}) & \text{if} \quad t_1 = t_2 \end{cases} \tag{3.73}$$

where $0 < \theta_1\theta_2 < 1$ and $0 < \theta_{12} \leq 1$. The marginal probability mass functions of the BG-Type II model are, respectively, standard geometric distributions starting at one with parameters θ_1 and θ_2.

Now, let (T_1, T_2) be a bivariate discrete random vector with a BG-Type II distribution. The correlation coefficient for the BG-Type II distribution is given by,

$$\rho = -\frac{(1-\theta_{12})(\theta_1\theta_2)^{1/2}}{1-\theta_1\theta_2\theta_{12}}. \tag{3.74}$$

Note that, the correlation is equals to zero when $\theta_{12} = 1$; otherwise $0 \leq \rho \leq 1$, that is, this distribution is useful for bivariate lifetimes with positive correlation.

3.4.5.2 Cure Rate Model

Now, let us assume a BG-Type II distribution with parameters θ_1, θ_2 and θ_{12} for the bivariate lifetimes T_1 and T_2. From the BG-Type II joint survival function and the equation (2.29), the bivariate mixture cure rate model is given by,

$$P(T_1 > t_1, T_2 > t_2) = \phi_{11}\left[\theta_1^{t_1}\theta_2^{t_2}\theta_{12}^{\min(t_1,t_2)}\right] + \phi_{10}\theta_1^{t_1} + \phi_{01}\theta_2^{t_2} + \phi_{00} \tag{3.75}$$

In Figure 11, it is illustrated the contour plots for the survival function expressed in Equation (3.75) for different parameter values.

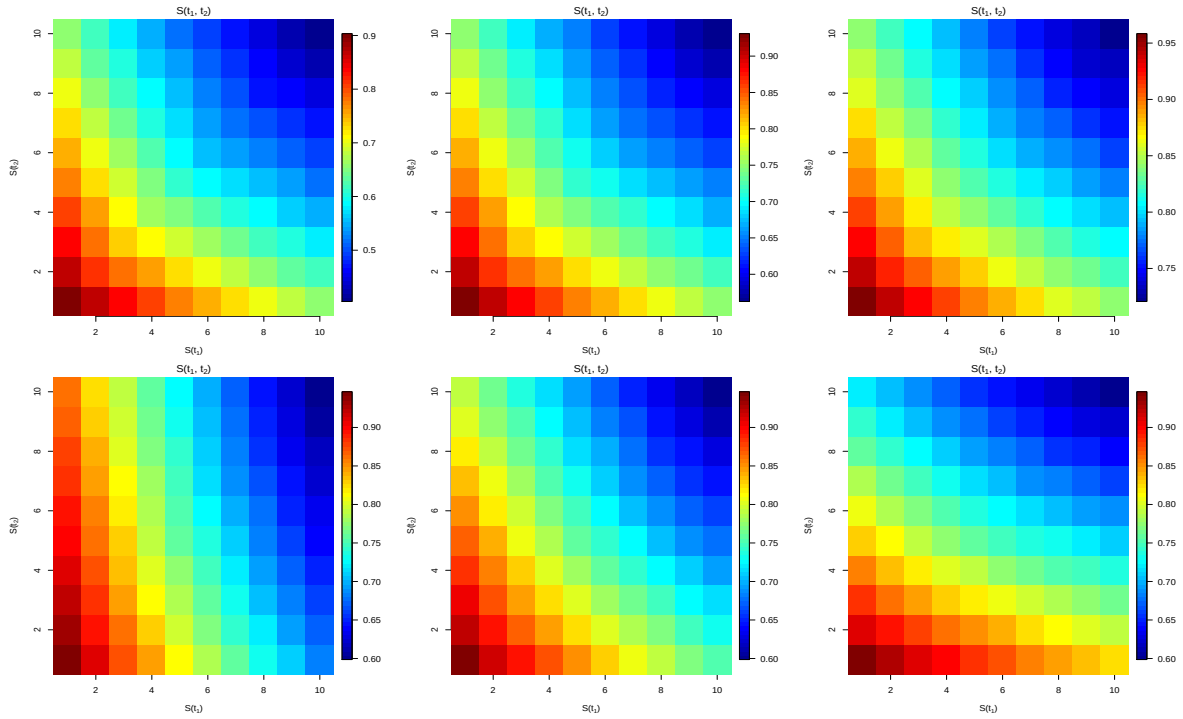

Figure 11 – Contour plots of the joint survival function in presence of cure rate for BG-Type II model assuming different parameter values (**Upper-panels:** fixed values given by $\theta_1 = \theta_2 = \theta_{12} = 0.95, \phi_{10} = \phi_{01} = 0.15$, and $\phi_{11} = (0.60, 0.40, 0.20), \phi_{00} = (0.10, 0.30, 0.50)$. **Lower-panels:** fixed values given by $\theta_1 = \theta_2 = \theta_{12} = 0.95, \phi_{11} = \phi_{00} = 0.15$, and $\phi_{10} = (0.60, 0.40, 0.20), \phi_{01} = (0.10, 0.30, 0.50)$).

For this discrete model, the contribution of the ith unit for the likelihood function is given by,

i.) If $i \in C_1$,

$$P(T_{1i} = t_{1i}, T_{2i} = t_{2i}) = \phi_{11}\left[\theta_1^{t_{1i}-1}\theta_2^{t_{2i}-1}\theta_{12}^{z_{1i}} - \theta_1^{t_{1i}}\theta_2^{t_{2i}-1}\theta_{12}^{z_{2i}} - \theta_1^{t_{1i}-1}\theta_2^{t_{2i}}\theta_{12}^{z_{3i}} + \theta_1^{t_{1i}}\theta_2^{t_{2i}}\theta_{12}^{z_{4i}}\right]$$

ii.) If $i \in C_2$,

$$P(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = \phi_{11}\left[\theta_1^{t_{1i}-1}\theta_2^{t_{2i}}\theta_{12}^{z_{3i}} - \theta_1^{t_{1i}}\theta_2^{t_{2i}}\theta_{12}^{z_{4i}}\right] + \phi_{10}\theta_1^{t_{1i}-1}(1-\theta_1)$$

iii.) If $i \in C_3$,

$$P(T_{1i} > t_{1i}, T_{2i} = t_{2i}) = \phi_{11}\left[\theta_1^{t_{1i}}\theta_2^{t_{2i}-1}\theta_{12}^{z_{2i}} - \theta_1^{t_{1i}}\theta_2^{t_{2i}}\theta_{12}^{z_{4i}}\right] + \phi_{01}\theta_2^{t_{2i}-1}(1-\theta_2)$$

iv.) If $i \in C_4$,

$$P(T_{1i} > t_{1i}, T_{2i} > t_{2i}) = \phi_{11}\left[\theta_1^{t_{1i}}\theta_2^{t_{2i}}\theta_{12}^{z_{4i}}\right] + \phi_{10}\theta_1^{t_{1i}} + \phi_{01}\theta_2^{t_{2i}} + \phi_{00}$$

where $z_1 = \min(t_1 - 1, t_2 - 1), z_2 = \min(t_1, t_2 - 1), z_3 = \min(t_1 - 1, t_2)$ e $z_4 = \min(t_1, t_2)$.

3.4.6 Basu-Dhar Bivariate Geometric

3.4.6.1 Standard Model

In this section, it is assumed the Basu-Dhar bivariate geometric distribution defined for discrete random variables, denoted by BDBG(θ_1,θ_2,θ_{12}), proposed by Basu and Dhar (1995) in reliability analysis. This distribution was also studied by Achcar et al. (2016a); de Oliveira and Achcar (2018) in the presence of covariates and censored data. The survival function is given by,

$$P_0(T_1 > t_1, T_2 > t_2) = \theta_1^{t_1}\theta_2^{t_2}\theta_{12}^{\max(t_1,t_2)} \tag{3.76}$$

where $0 < \theta_1, \theta_2 < 1$ and $0 < \theta_{12} \leq 1$. From (3.76), the bivariate probability mass function of the Basu-Dhar bivariate geometric distribution, for two discrete random variables T_1 and T_2, is given by,

$$P_0(T_1 = t_1, T_2 = t_2) = \begin{cases} \theta_1^{t_1-1}(\theta_2\theta_{12})^{t_2-1}q_1(1-\theta_2\theta_{12}) & \text{if} \quad t_1 < t_2 \\ \theta_2^{t_2-1}(\theta_1\theta_{12})^{t_1-1}q_2(1-\theta_1\theta_{12}) & \text{if} \quad t_1 > t_2 \\ (\theta_1\theta_2\theta_{12})^{t_1-1}(1-\theta_1\theta_{12}-\theta_2\theta_{12}+\theta_1\theta_2\theta_{12}) & \text{if} \quad t_1 = t_2 \end{cases} \tag{3.77}$$

where $0 < \theta_1 < 1$, $0 < \theta_2 < 1$, $0 < \theta_{12} \leq 1$, $q_1 = 1-\theta_1$ and $q_2 = 1-\theta_2$. The marginal probability mass functions of the BDBG model are respectively, given by standard geometric distributions starting at one, that is,

$$\begin{aligned} P_{01}(T_1 = t_1) &= (\theta_1\theta_{12})^{t_1-1}(1-\theta_1\theta_{12}) \\ P_{02}(T_2 = t_2) &= (\theta_2\theta_{12})^{t_2-1}(1-\theta_2\theta_{12}). \end{aligned} \tag{3.78}$$

and marginal survival functions given, respectively by,

$$\begin{aligned} P_{01}(T_1 > t_1) &= (\theta_1\theta_{12})^{t_1-1} \\ P_{02}(T_2 > t_2) &= (\theta_2\theta_{12})^{t_2-1} \end{aligned} \tag{3.79}$$

Now, let (T_1, T_2) be a bivariate discrete random vector with a BDBG distribution. Following Li and Dhar (2013), the product moment of T_1 and T_2 is derived as follows:

$$\mathbb{E}[T_1 T_2] = \frac{1 - \theta_1\theta_2\theta_{12}^2}{(1-\theta_1\theta_{12})(1-\theta_2\theta_{12})(1-\theta_1\theta_2\theta_{12})} \tag{3.80}$$

From (3.80), the correlation coefficient for the BDBG distribution is given by,

$$\rho = \frac{(1-\theta_{12})(\theta_1\theta_2)^{1/2}}{1-\theta_1\theta_2\theta_{12}}. \tag{3.81}$$

Note that, the correlation is equals to zero when $\theta_{12} = 1$; otherwise $0 \leq \rho \leq 1$, that is, this distribution is useful for bivariate lifetimes with positive correlation.

3.4.6.2 Cure Rate Model

Now, let us assume a BDBG distribution with parameters θ_1, θ_2 and θ_{12} for the bivariate lifetimes T_1 and T_2. From the BDBG joint survival function and the equation (2.29), the bivariate mixture cure rate model is given by,

$$P(T_1 > t_1, T_2 > t_2) = \phi_{11}\left[\theta_1^{t_1}\theta_2^{t_2}\theta_{12}^{\max(t_1,t_2)}\right] + \phi_{10}(\theta_1\theta_{12})^{t_1} + \phi_{01}(\theta_2\theta_{12})^{t_2} + \phi_{00} \tag{3.82}$$

In Figure 12, it is illustrated the contour plots for the survival function expressed in Equation (3.82) for different parameter values.

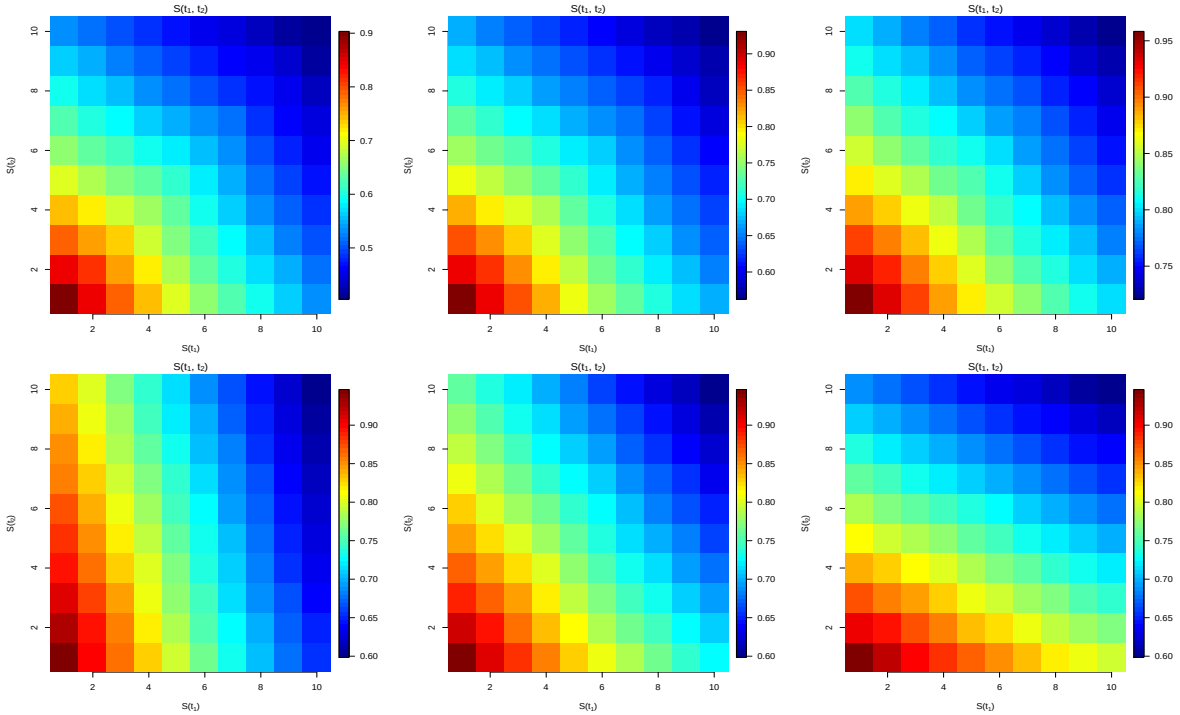

Figure 12 – Contour plots of the joint survival function in presence of cure rate for BDBG model assuming different parameter values (**Upper-panels:** fixed values given by $\theta_1 = \theta_2 = \theta_{12} = 0.95, \phi_{10} = \phi_{01} = 0.15$, and $\phi_{11} = (0.60, 0.40, 0.20), \phi_{00} = (0.10, 0.30, 0.50)$. **Lower-panels:** fixed values given by $\theta_1 = \theta_2 = \theta_{12} = 0.95, \phi_{11} = \phi_{00} = 0.15$, and $\phi_{10} = (0.60, 0.40, 0.20), \phi_{01} = (0.10, 0.30, 0.50)$).

For this discrete model, the contribution of the ith unit for the likelihood function is given by,

i.) If $i \in C_1$,

$$P(T_{1i} = t_{1i}, T_{2i} = t_{2i}) = \phi_{11}\left[\theta_1^{t_{1i}-1}\theta_2^{t_{2i}-1}\theta_{12}^{z_{1i}} - \theta_1^{t_{1i}}\theta_2^{t_{2i}-1}\theta_{12}^{z_{2i}} - \theta_1^{t_{1i}-1}\theta_2^{t_{2i}}\theta_{12}^{z_{3i}} + \theta_1^{t_{1i}}\theta_2^{t_{2i}}\theta_{12}^{z_{4i}}\right]$$

ii.) If $i \in C_2$,

$$P(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = \phi_{11}\left[\theta_1^{t_{1i}-1}\theta_2^{t_{2i}}\theta_{12}^{z_{3i}} - \theta_1^{t_{1i}}\theta_2^{t_{2i}}\theta_{12}^{z_{4i}}\right] + \phi_{10}(\theta_1\theta_{12})^{t_{1i}-1}(1-\theta_1\theta_{12})$$

iii.) If $i \in C_3$,

$$P(T_{1i} > t_{1i}, T_{2i} = t_{2i}) = \phi_{11}\left[\theta_1^{t_{1i}}\theta_2^{t_{2i}-1}\theta_{12}^{z_{2i}} - \theta_1^{t_{1i}}\theta_2^{t_{2i}}\theta_{12}^{z_{4i}}\right] + \phi_{01}(\theta_2\theta_{12})^{t_{2i}-1}(1-\theta_2\theta_{12})$$

iv.) If $i \in C_4$,

$$P(T_{1i} > t_{1i}, T_{2i} > t_{2i}) = \phi_{11}\left[\theta_1^{t_{1i}}\theta_2^{t_{2i}}\theta_{12}^{z_{4i}}\right] + \phi_{10}(\theta_1\theta_{12})^{t_{1i}} + \phi_{01}(\theta_2\theta_{12})^{t_{2i}} + \phi_{00}$$

where $z_1 = \max(t_1 - 1, t_2 - 1), z_2 = \max(t_1, t_2 - 1), z_3 = \max(t_1 - 1, t_2)$ e $z_4 = \max(t_1, t_2)$.

3.4.7 Arnold Bivariate Geometric

3.4.7.1 Standard Model

In this section it is assumed a bivariate geometric distribution also for discrete lifetime data denoted as ABG(θ_1, θ_2) introduced by Arnold (1975). The ABG distribution has joint survival function given by,

$$P_0(T_1 > t_1, T_2 > t_2) = \begin{cases} (1-\theta_1-\theta_2)^{t_1}(1-\theta_2)^{t_2-t_1} & \text{if} \quad t_1 < t_2 \\ (1-\theta_1-\theta_2)^{t_2}(1-\theta_1)^{t_1-t_2} & \text{if} \quad t_1 > t_2 \\ 0 \quad \text{if} \quad t_1 = t_2 \end{cases} \tag{3.83}$$

where $0 < \theta_1, \theta_2 < 1$. For this model, marginal probability mass functions are respectively, given by standard geometric distributions starting at one, that is,

$$\begin{aligned} P_{01}(T_1 = t_1) &= (1-\theta_1)^{t_1-1}\theta_1 \\ P_{02}(T_2 = t_2) &= (1-\theta_2)^{t_2-1}\theta_2 \end{aligned} \tag{3.84}$$

and marginal survival functions, given respectively by,

$$\begin{aligned} P_{01}(T_1 > t_1) &= (1-\theta_1)^{t_1-1} \\ P_{02}(T_2 > t_2) &= (1-\theta_2)^{t_2-1} \end{aligned} \tag{3.85}$$

The correlation coefficient between T_1 and T_2, assuming the ABG model, is given by,

$$\rho = -\frac{\theta_1\theta_2}{(\theta_1+\theta_2)[(1-\theta_1)(1-\theta_2)]^{1/2}} \tag{3.86}$$

Note that, the correlation is equals to zero when $\theta_1 = 0$ or $\theta_2 = 0$; otherwise $-1 \leq \rho < 0$, that is, this distribution is useful for bivariate lifetimes with negative correlation.

3.4.7.2 Cure Rate Model

For likelihood-based inference methods, assuming the Arnold bivariate geometric distribution with joint survival function given by (3.83) in the presence of censored data and cure rate, the contribution of the ith unit for the likelihood function is given by,

i.) If $i \in C_1$,

$$P(T_{1i} = t_{1i}, T_{2i} = t_{2i}) = \begin{cases} \phi_{11}\theta_1\theta_2(1-\theta_1-\theta_2)^{t_{1i}-1}(1-\theta_2)^{t_{2i}-t_{1i}-1} & \text{if} \quad t_{1i} < t_{2i} \\ \phi_{11}\theta_1\theta_2(1-\theta_1-\theta_2)^{t_{2i}-1}(1-\theta_1)^{t_{1i}-t_{2i}-1} & \text{if} \quad t_{1i} > t_{2i} \end{cases}$$

ii.) If $i \in C_2$,

$$P(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = \begin{cases} \phi_{11}\theta_1(1-\theta_1-\theta_2)^{t_{1i}-1}(1-\theta_2)^{t_{2i}-t_{1i}-1} \\ + \ \phi_{10}P_{01}(T_{1i} = t_{1i}) \quad t_{1i} < t_{2i} \\ \phi_{11}\theta_1(1-\theta_1-\theta_2)^{t_{2i}}(1-\theta_1)^{t_{1i}-t_{2i}-1} \\ + \ \phi_{10}P_{01}(T_{1i} = t_{1i}) \quad t_{1i} > t_{2i} \end{cases}$$

where $P_{01}(T_{1i} = t_{1i}) = (1-\theta_1)^{t_{1i}-1}\theta_1$.

iii.) If $i \in C_3$,

$$P(T_{1i} > t_{1i}, T_{2i} = t_{2i}) = \begin{cases} \phi_{11}\theta_2(1-\theta_1-\theta_2)^{t_{1i}}(1-\theta_2)^{t_{2i}-t_{1i}-1} \\ + \ \phi_{01}P_{02}(T_{2i} = t_{2i}) \quad t_{1i} < t_{2i} \\ \phi_{11}\theta_2(1-\theta_1-\theta_2)^{t_{2i}-1}(1-\theta_1)^{t_{1i}-t_{2i}} \\ + \ \phi_{01}P_{02}(T_{2i} = t_{2i}) \quad t_{1i} > t_{2i} \end{cases}$$

where $P_{02}(T_{2i} = t_{2i}) = (1-\theta_2)^{t_{2i}-1}\theta_2$.

iv.) If $i \in C_4$,

$$\begin{aligned} P(T_{1i} > t_{1i}, T_{2i} > t_{2i}) &= \phi_{11}P_0(T_{1i} > t_{1i}, T_{2i} > t_{2i}) + \phi_{10}P_{01}(T_{1i} > t_{1i}) \\ &+ \phi_{01}P_{02}(T_{2i} > t_{2i}) + \phi_{00} \end{aligned}$$

where $P_0(T_{1i} > t_{1i}, T_{2i} > t_{2i})$ is given in Equation (2.27); $P_{01}(T_{1i} > t_{1i})$ and $P_{02}(T_{2i} > t_{2i})$ are given in Equation (2.29).

In Figure 13, it is illustrated the contour plot of survival function probabilities of T_1 and T_2 assuming different parameter values and cure rate for the ABG distribution. Observe that, in the case $T_1 = T_2$, the joint survival function values assume only the marginal survival values since the joint survival probability is zero at $T_1 = T_2$.

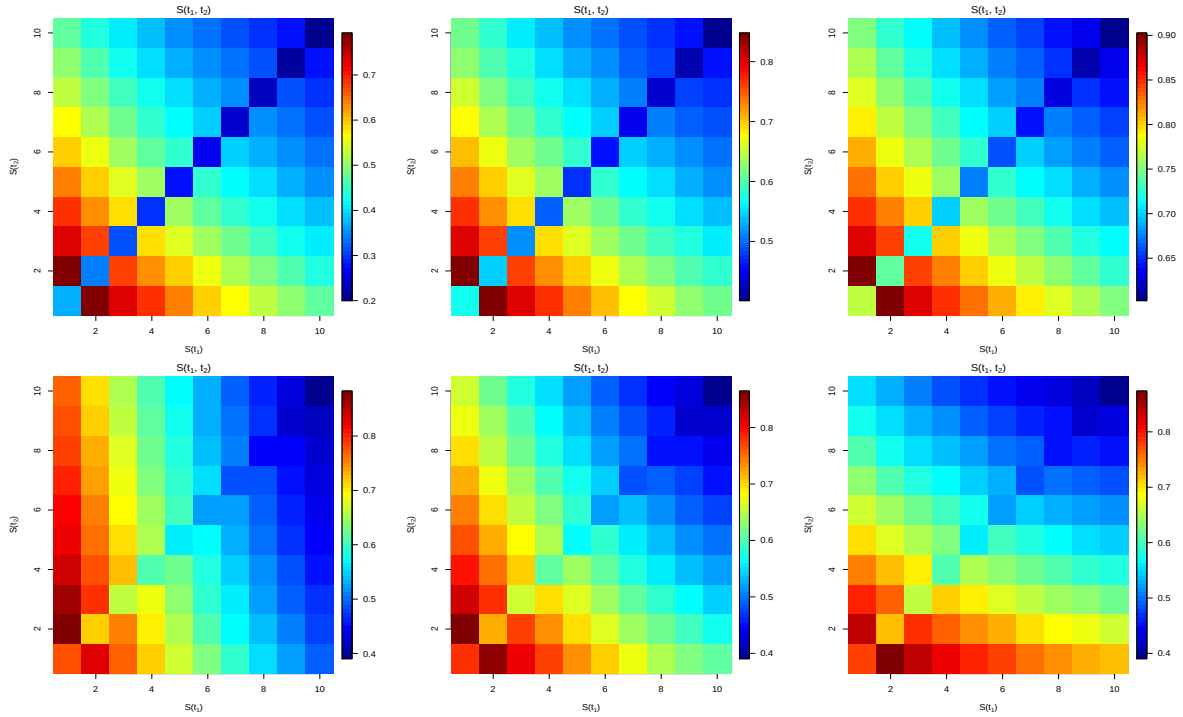

Figure 13 – Contour plots of the joint survival function in presence of cure rate for ABG model assuming different parameter values (**Upper-panels:** fixed values given by $\theta_1 = \theta_2 = 0.10, \phi_{10} = \phi_{01} = 0.15$, and $\phi_{11} = (0.60, 0.40, 0.20), \phi_{00} = (0.10, 0.30, 0.50)$. **Lower-panels:** fixed values given by $\theta_1 = \theta_2 = 0.10, \phi_{11} = \phi_{00} = 0.15$, and $\phi_{10} = (0.60, 0.40, 0.20), \phi_{01} = (0.10, 0.30, 0.50)$).

3.4.8 Bivariate Discrete Generalized Rayleigh Cure Rate Model

In this section, it is assumed the bivariate discrete generalized Rayleigh model denoted by DBGR$(\lambda_1, \lambda_2, \alpha)$ introduced in Section 2.3. The mixture cure rate model for bivariate lifetimes T_1 and T_2 assuming the DBGR distribution is given by,

$$\mathrm{P}(T_1 > t_1, T_2 > t_2) = \phi_{11}\left[\frac{\alpha\lambda_1^{t_1^2}\lambda_2^{t_2^2}}{1-\bar{\alpha}\lambda_1^{t_1^2}\lambda_2^{t_2^2}}\right] + \phi_{10}\left[\frac{\alpha\lambda_1^{t_1^2}}{1-\bar{\alpha}\lambda_1^{t_1^2}}\right] + \phi_{01}\left[\frac{\alpha\lambda_2^{t_2^2}}{1-\bar{\alpha}\lambda_2^{t_2^2}}\right] + \phi_{00} \quad (3.87)$$

where $\bar{\alpha} = 1 - \alpha$. Thus, the contribution of the i*th* observation for the likelihood function assuming the DBGR cure rate model could be obtained directly from its marginal distributions and the likelihood function based on a random sample of size n, is given by,

$$\begin{aligned} L \;\propto\; & \prod_{i\in C_1} P(T_{1i} = t_{1i}, T_{2i} = t_{2i}) \prod_{i\in C_2} P(T_{1i} = t_{1i}, T_{2i} > t_{2i}) \prod_{i\in C_3} P(T_{1i} > t_{1i}, T_{2i} = t_{2i}) \\ & \times \prod_{i\in C_4} P(T_{1i} > t_{1i}, T_{2i} > t_{2i}) \end{aligned}$$

3.5 Summary of the Proposed Models

In Table 2, it is presented a summary for the proposed models including the correlation coefficients and their ranges which is an important aspect to consider when deciding which model is appropriate for the data analysis (see de Oliveira et al., 2018).

Table 2 – Summary of the proposed models and their correlation coefficients.

Model	Type of Model	Parameters	Correlation Coefficient (ρ)	Correlation Range
GBE	Continuous	$\lambda_1, \lambda_2 > 0$ $0 < \theta < 1$	$1 - \frac{1}{\lambda_3} \exp\left(\frac{1}{\theta}\right) E_i\left(\frac{1}{\theta}\right)$	$0 < \rho \leq 1$
BBBE	Continuous	$\lambda_1, \lambda_2, \lambda_3 > 0$	Equation (3.63)	$0 < \rho \leq 1$
MOBE	Continuous	$\lambda_1, \lambda_2, \lambda_{12} > 0$	$\frac{\lambda_{12}}{\lambda_1 + \lambda_2 + \lambda_{12}}$	$0 < \rho < 1$
BLI	Continuous	$\beta_1, \beta_2, \beta_3 > 0$	Not illustrated here	$0 < \rho \leq 1$
BLIII	Continuous	$\beta_1, \beta_2, \beta_3 > 0$	Not illustrated here	$-1 \leq \rho < 0$
ABG	Discrete	$0 < \theta_1, \theta_2 < 1$	$-\frac{\theta_1\theta_2}{(\theta_1 + \theta_2)[(1-\theta_1)(1-\theta_2)]^{1/2}}$	$-1 \leq \rho < 0$
BDBG	Discrete	$0 < \theta_1, \theta_2 < 1$ $0 < \theta_{12} \leq 1$	$\frac{(1-\theta_{12})(\theta_1\theta_2)^{1/2}}{1-\theta_1\theta_2\theta_{12}}$	$0 \leq \rho \leq 1$
BG-Type II	Discrete	$0 < \theta_1, \theta_2 < 1$ $0 < \theta_{12} \leq 1$	$-\frac{(1-\theta_{12})(\theta_1\theta_2)^{1/2}}{1-\theta_1\theta_2\theta_{12}}$	$-1 \leq \rho \leq 0$
DBGR	Discrete	$0 < \lambda_1, \lambda_2 < 1$ $\alpha > 0$	Not illustrated here	$-1 \leq \rho \leq 1$

From Table 2, it could be seen that the discrete models have simpler equations for the correlation coefficients when compared to the continuous models (except for MOBE model). However, analyzing the correlation range, it could be concluded, except for BL-III model, that the continuous models have positive correlation and, in discrete case, only BDBG has positive correlation. However, in terms of flexibility, the DBGR model has positive, null and negative correlation which is the most flexible model among the nine proposed models, but its correlation coefficient depends on a infinite sum due the cross factorial moment presented in Section 2.3 which is one disadvantage of this model. Moreover, the models that have only negative correlation range could be more limited in lifetime data applications due to its negative range which could be not appropriated for the data analysis. Finally, in terms of mathematical expressions, the MOBE model is the most flexible model among the nine models, however, this model has a singular part which could be a problem in the estimation process. On the other hand, the BDBG model also has simpler expression for the correlation coefficient and no singular part which is a great advantage to model lifetime data with positive correlation.

Chapter 4

Modeling Long-Term Survivors in Medical Studies Under Bayesian Approach

4.1 Introduction

Most of multivariate lifetime data are derived from studies that most of the researchers are interested related to the waiting times until the ocurrence of an event of interest. As a special case, clinical trial studies where the patients are followed-up during a fixed period of time to assess the lifetimes efficacy of a new treatment (Vahidpour, 2016). In general, the event of interest in such studies could be death, cure, time to a heart attack, remission time, reaction time for a treatment, deterioration time of a organ, or adverse reaction; and the follow-up time for the study may range from few weeks to many years. In the literature, this kind of data is called *time-to-event data*. However, in some situations, the event of interest may not occur for some individuals, even after a very long period of follow-up time. In those cases, the standard survival models cannot accurately describe the behavior of all individuals. According to Vahidpour (2016), cure rate models could be useful to be fitted by time-to-event data with long term survivors.

The use of cure rate models could be an useful tool that could provide invaluable information about the clinical trial (see, for example, Fleming and Lin, 2000; Singh and Mukhopadhyay, 2011; Ghadimi et al., 2011). Most of the lifetime data from those studies are often modeled using the standard nonparametric proportional hazards model introduced by Cox (1972) which estimates the covariate effects as the log hazard ratios or parametric regression models as log-additive regression models assuming standard Weibull, log-normal or gamma probability distributions. However, the Cox proportional hazards model, provides no direct estimation for the cure rate. Thus, as an alternative to Cox's models, fully parametric models such as the models introduced in Chapter 3 could be used offering a gain that may not be obtained under the Cox's model (see, Ghadimi et al., 2011).

4.2 Study 1: Pelvic Sarcoma Data

4.2.1 Introduction

According to Yazdanbod et al. (2004) and Ghadimi et al. (2011), cancer is known as one of the major causes leading to many disorders, death, and disabilities worldwide. According to Zali et al. (2005), it has been increasingly affecting the human population during the past decades so that considerable amount of health care resources have been allocated to lessen its side effects. In 2012, according to World Cancer Research Foundation International (WCRF), there were an estimated 14.1 million cancer cases around the world, of these 7.4 million cases were in men and 6.7 million in women(Ferlay et al., 2013). Moreover, it is an extremely heterogeneous disease. In this way, genetic differences between people lead to differences in susceptibility and cancers arising from the same tissue can be stratified into subtypes of the disease based on differences in genomic measurements (Curtis et al., 2012). Some examples related to cancer morbidity among the years could be seen in Wingo et al. (1996); Siegel et al. (2013, 2014, 2016); Facina (2014).

The main goal of our study is to investigate long-term lifetimes and risk factors with a special application related to the lifetimes of patients receiving a treatment for pelvic sarcomas (part of the results of this study, were already published; see de Oliveira et al., 2019). According to Sugarbaker (2004), sarcomas are unusual but not rare malignancies and account for only 1% of adult solid tumors. These sarcomas appear most frequently between the fourth and sixth decades of life with a 2:1 male/female ratio and could arise anywhere in the body with the lower extremity being the most common site. Incidence rates are as follows: lower extremities (46%), upper extremities (13%), retroperitoneum, pelvis and visceral (12%), truncal sarcomas (19%), and head and neck sarcomas (9%).

An important characteristic of the natural history of retroperitoneal and pelvic sidewall sarcomas is how it differs significantly from the more common abdominal and pelvic adenocarcinomas and from visceral sarcoma. These differences are important when planning a treatment since the treatment of pelvic sarcomas are surgically difficult due to the anatomic proximity of the pelvis to many neurovascular structures and the urinary and intestinal tracts with poor oncological outcomes and high complications rate (see, for example, Figure 14).

4.2.2 Material and Methods

To accomplish our main goal, it is considered a retrospective cohort study by Puchner et al. (2017) which consists in 147 consecutive cases with surgical treatment of a sarcoma of the pelvis observed between the years 1980 and 2012. The records included 68

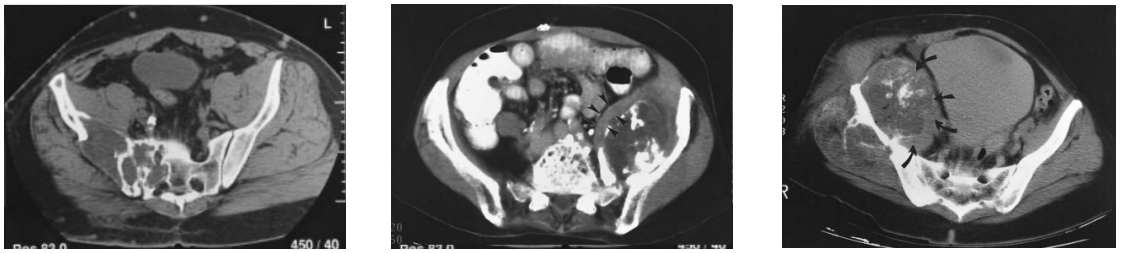

Figure 14 – **Left panel:** High-grade chondrosarcoma of the right sacrum, ilium, and periacetabular region, encasing the ipsilateral sacral foramina. **Middle panel:** The iliacus muscle is "pushed" by a growing bone sarcoma and serves as a barrier to direct extension of the tumor to the pelvic viscera. **Right panel:** High-grade sarcoma of the left ilium "pushing" the iliacus muscle (arrows) towards the midline (**Source:** Bickels et al., 2012)

.

males (46%) and 79 females (54%) with an average age of 38 ± 20 years at time of surgery. Also, the diagnosis was based on conclusive clinical and imaging findings and was always confirmed by biopsy and histological analysis. As prognostic and oncological factors, there are the patient sex, tumor grade, radiotherapy, chemotherapy, among many others in the data set. The statistical analysis was conducted assuming a univariate Weibull distribution under three responses (survival time, infection time and metastasis time) considering the mixture and non-mixture approaches.

4.2.2.1 Discrete Weibull Cure Rate Models

In this section, it is assumed that the probability distribution for the lifetimes of the suscetible population follows a discrete Weibull (DW) distribution introduced by Nakagawa and Osaki (1975), which can be considered as a discrete analogue of the continuous Weibull distribution. The probability mass function (p.m.f.) of a DW distribution is defined by

$$\mathrm{P}\left(T = t \mid \phi, \beta\right) = \phi^{t^{\beta}} - \phi^{(t+1)^{\beta}}, \qquad t \in \mathbb{N}_0 = \{0, 1, 2, \ldots\} \tag{4.1}$$

and its corresponding survival function is given by,

$$S(t \mid \phi, \beta) = \Pr\left(T > t \mid \phi, \beta\right) = \phi^{(t+1)^{\beta}} \tag{4.2}$$

where $\beta > 0$ and $0 < \phi < 1$. Note that, when $\beta = 1$, the DW distribution reduces to the geometric distribution and when $\beta = 2$, it reduces to the Rayleigh distribution introduced by Roy (2004a). This model has been applied to many areas, including competing risks, extreme values, failure times, regional analyses of precipitation, and reliability (see, for example, Khan et al., 1989; Kulasekera, 1994; Roy, 2002; Murthy et al., 2004; Englehardt and Li, 2011; Almalki and Nadarajah, 2014; Brunello and Nakano, 2015).

Table 3 – Baseline characteristics and operative data (available in Puchner et al., 2017).

Variable	**n (%)**
Sex	
Male	68 (46%)
Female	79 (54%)
Histology	
Chondrosarcoma	54 (37%)
Ewing's sarcoma/PNET	37 (25%)
Osteosarcoma	32 (22%)
Leiomyosarcoma	4 (3%)
Sarcoma-Not other specified	4 (3%)
Hemangiopericytoma	3 (2%)
Others	13 (9%)
Grading	
G3	101 (69%)
G2	38 (26%)
G1	8 (5%)
Age at time of surgery (Years; SD)	38 ± 20
Size (cm^3; SD)	1023 ± 1848
Location	
Ileum	110 (75%)
Ischium	9 (6%)
Pubis	28 (19%)
Periacetabular involvement	67 (46%)
Type of surgery	
Resection without reconstruction	46 (31%)
Endoprosthetic reconstruction	47 (32%)
Biological reconstruction	21 (14%)
Internal hemipelvectomy and transposition of the hip	14 (10%)
External hemipelvectomy	19 (13%)
Type of resection	
Type I	27 (18%)
Type III	14 (10%)
Type I/II	19 (13%)
Type I/IV	10 (7%)
Type I/II/IV	5 (3%)
Type II/III	25 (17%)
Type I/II/III	33 (22%)
Type I/II/III/IV	14 (10%)

Using the proposed methodology introduced in Chapter 3 for the univariate case, the mixture and non-mixture cure fraction model for the lifetime T assuming the DW distribution are given, respectively, by,

$$S(t \mid \phi, \beta, \rho) = \Pr(T > t \mid \phi, \beta, \rho) = \rho + (1-\rho)\,\phi^{(t+1)^\beta} \tag{4.3}$$

and,

$$S(t \mid \phi, \beta, \rho) = \Pr(T > t \mid \phi, \beta, \rho) = \exp\left\{\ln(\rho)\,\left[1 - \phi^{(t+1)^\beta}\right]\right\}, \tag{4.4}$$

where $\rho \in (0,1)$ is the cure fraction parameter.

Remark 3.4.2.1.1. *Since the DW model has no closed form for the expected value and variance, the subsequently mixture and non-mixture cure fraction models also have no closed form for their expected values and variances. However, the expected values and the variances could be obtained using numerical methods directly from the definition of the r-th moment given by*

$$\mathbb{E}(T^r) = \sum_{k=0}^{\infty} k^r \left\{\phi^{\log\left[1+\left(\frac{k}{\theta}\right)^\alpha\right]} - \phi^{\log\left[1+\left(\frac{k+1}{\theta}\right)^\alpha\right]}\right\}$$

where, in particular, for $r = 1$, we have $\mathbb{E}(T) = \sum_{k=1}^{\infty} \phi^{\log\left[1+\left(\frac{k}{\theta}\right)^\alpha\right]}$ and for $r = 2$ we have $\mathbb{E}(T^2) = \sum_{k=1}^{\infty} (2k-1)\,\phi^{\log\left[1+\left(\frac{k}{\theta}\right)^\alpha\right]}$.

4.2.2.2 Inference and Residuals

To get the inferences of interest, let us consider the situation when the lifetime, T_i, is not completely observed and may be subject to right censoring. Let C_i be the censoring time for the ith individual. From a sample of size n, it is observed $T_i = \min\{T_i, C_i\}$ and $\delta_i = I(T_i < C_i)$, where $\delta_i = 1$ if T_i is an observed lifetime and $\delta_i = 0$ if it is right censored lifetime. In this case, the log-likelihood function considering the DW distribution with pmf defined in (4.1), can be written as

$$\ell(\beta, \phi \mid \boldsymbol{t}, \boldsymbol{\delta}) = \sum_{i=1}^{n} \delta_i \log\left[\phi^{t^\beta} - \phi^{(t+1)^\beta}\right] + \sum_{i=1}^{n} (1-\delta_i) \log\left[\phi^{(t+1)^\beta}\right] \tag{4.5}$$

where $\boldsymbol{t} = (t_1, \ldots, t_n)^\top$ and $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)^\top$. Assuming the DW mixture model the log-likelihood function is expressed as

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{t}, \boldsymbol{\delta}) = r\,\ln(1-\rho) + \sum_{i=1}^{n} \delta_i \ln\left[\phi^{t_i^\beta} - \phi^{(t_i+1)^\beta}\right] + \sum_{i=1}^{n} (1-\delta_i) \ln\left[\rho + (1-\rho)\,\phi^{(t_i+1)^\beta}\right] \tag{4.6}$$

where $\boldsymbol{\theta} = (\beta, \phi, \rho)^\top$ and $r = \sum_{i=1}^{n} \delta_i$ is the number of uncensored observations. Additionally, considering the discrete Weibull non-mixture model, the log-likelihood function is given by,

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{t}, \boldsymbol{\delta}) = r \, \ln(-\ln \rho) + \sum_{i=1}^{n} \delta_i \, \ln\left[\phi^{t_i^\beta} - \phi^{(t_i+1)^\beta}\right] + (\ln \rho) \sum_{i=1}^{n} 1 - \phi^{(t_i+1)^\beta}. \tag{4.7}$$

The maximum likelihood estimates (MLEs) $\widehat{\boldsymbol{\theta}}$ for the unknown parameters in the vector parameter $\boldsymbol{\theta}$ are obtained by maximizing the log-likelihood functions defined in Equations (4.5), (4.6) and (4.7) using standard optimization methods, such as Newton-Raphson and quasi-Newton. In this study, the MLEs were obtained by the quasi-Newton method available in the `SAS/NLMIXED` procedure (SAS, 2010). Under suitable regularity conditions (see Lehmann and Casella, 1998), the asymptotic distribution of the maximum likelihood estimator $\widehat{\boldsymbol{\theta}}$ is a multivariate Normal distribution with mean $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}\left(\widehat{\boldsymbol{\theta}}\right)$, which can be consistently estimated by the inverse of the observed Fisher information matrix given by

$$\widehat{\boldsymbol{\Sigma}}\left(\widehat{\boldsymbol{\theta}}\right) = \left[-\frac{\partial\, \ell(\boldsymbol{\theta} \mid \boldsymbol{t}, \boldsymbol{\delta})}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}^\top}\right]^{-1} \tag{4.8}$$

evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. The required second derivatives are computed numerically using the `SAS/NLMIXED` procedure.

For regression analysis, it is proposed to relate the parameters ϕ and ρ of the mixture and non-mixture models to the vectors of explanatory variables $\boldsymbol{x}_i$ and $\boldsymbol{z}_i$, respectively. Thus, it is assumed the the following link functions

$$\log(-\log(\phi_i)) = \boldsymbol{x}_i^\top \boldsymbol{\alpha} \qquad \text{and} \qquad \log\left(\frac{\rho_i}{1-\rho_i}\right) = \boldsymbol{z}_i^\top \boldsymbol{\delta} \tag{4.9}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$ denote the vectors of unknown regression parameters. It is noteworthy, that the log-log link function in ϕ is motivated by the analytical formula for the quantile function of the DW model (see Klakattawi et al., 2018), which facilitates the interpretation of the coefficients. According to Klakattawi et al. (2018) the regression parameters $\boldsymbol{\alpha}$ can be interpreted in relation to the logarithm of the median.

Now, since the response is discrete, in the evaluation and study of departures from the model assumptions we propose the use of the randomized quantile residuals introduced by Dunn and Smyth (1996), which are defined as follows,

$$\widehat{r}_i = \Phi^{-1}(u_i), \qquad i = 1, \ldots, n \tag{4.10}$$

where $\Phi(\cdot)$ is the standard normal distribution function and u_i is a random value from the uniform distribution on the interval

$$u_i = \begin{cases} \left[F(t_i - 1 \mid \widehat{\boldsymbol{\theta}}), F(t_i \mid \widehat{\boldsymbol{\theta}})\right], & \text{for } \delta_i = 1 \\ \left[F(t_i \mid \widehat{\boldsymbol{\theta}}), 1\right], & \text{for } \delta_i = 0 \end{cases} \tag{4.11}$$

where $F(t_i \mid \widehat{\boldsymbol{\theta}})$ is the cumulative distribution function of mixture and non-mixture DW models. Apart from the variability due to the estimates of the parameters these residuals have standard normal distribution if the proposed model is is correctly specified (Dunn and Smyth, 1996).

Hence, to check if the model assumption is adequate we can examine the half-normal plots with simulated envelope proposed by Atkinson (1981). The simulated envelope can be construct as follows:

i.) fit the model and generate a sample set of n independent observations using the estimated parameters of the fitted model;

ii.) fit the model from the generated sample, calculate the absolute values of the residuals and arrange them in order;

iii.) repeat steps (i) and (ii) B number of times;

iv.) consider the n sets of the B ordered statistics of the residuals, then for each set calculate the quantile $\gamma/2$, the median and the quantile $1-\gamma/2$;

v.) plot these values and the ordered residuals of the original sample set versus the expected order statistics of a half-normal distribution, which is approximated as

$$\Phi^{-1}\left(\frac{i+n-0.125}{2\,n+0.5}\right).$$

According to Atkinson (1981) if the model was correctly specified then no more than $\gamma \times 100\%$ of the observations are expected to appear outside the envelope bands. Additionally, if a large proportion of the observations lies outside the envelope, thus one has evidence against the adequacy of the fitted model.

Finally, to discriminate the proposed models we can consider the Akaike Information Criterion (AIC) (Akaike, 1974) which is given by $\text{AIC} = -2\,\ell\left(\widehat{\boldsymbol{\theta}}\right) + 2\,p$, where p is the number of model parameters. Among all fitted models, the one with the smallest value is commonly considered as the better model to describe the data (Rohde, 2014).

4.2.3 Results

First of all, it was assumed discrete mixture and non-mixture Weibull models for the PS dataset without considering the presence of covariates under a Classical Approach. The corresponding inference results are given in Table 4. From these results, it is observed that all models, on computational aspects, did not show instability and the estimation method converged successful. As expected, the smallest AIC values were obtained assuming the mixture and non-mixture models.

Table 4 – Inference results for the three lifetimes for the PS dataset: survival, infection and metastasis times.

Model	Par.	Survival Time MLE	Survival Time 95% C.I.	AIC	Infection Time MLE	Infection Time 95% C.I.	AIC	Metastasis Time MLE	Metastasis Time 95% C.I.	AIC
Weibull	μ	19.6126	(8.5066, 30.7186)	761.5373	26.0097	(7.8812, 44.1382)	311.0863	37.1204	(8.8439, 65.3968)	423.8983
	β	0.6181	(0.4921, 0.7442)		0.4562	(0.3025, 0.6098)		0.6229	(0.4521, 0.7936)	
Mixture Weibull	μ	23.9595	(6.7375, 41.1815)	751.9783	7.7367	(1.4697, 14.0038)	309.0108	31.7240	(1.0969, 62.3511)	411.5414
	β	0.9199	(0.7088, 1.1310)		0.6420	(0.4156, 0.8683)		1.0576	(0.7822, 1.3330)	
	ρ	0.3663	(0.2593, 0.4734)		0.7256	(0.6009, 0.8503)		0.6261	(0.5156, 0.7367)	
Non-Mixture Weibull	μ	41.3519	(10.2851, 72.4186)	754.2051	9.1000	(1.3374, 16.8625)	309.0300	43.4083	(-0.4776, 87.2941)	411.3725
	β	0.9714	(0.7394, 1.2035)		0.6609	(0.4277, 0.8941)		1.1071	(0.8196, 1.3946)	
	ρ	0.3632	(0.2455, 0.4809)		0.7255	(0.5993, 0.8517)		0.6246	(0.5133, 0.7359)	

Moreover, the estimated proportion of cure fraction for the survival times, that is, the proportions of people in population who will not die due to sarcoma pelvic, are given by 36.63% and 36.32% considering the mixture and non-mixture models, respectively, while from the Kaplan-Meier estimate it is approximately 37.40%. Regarding the estimated proportion of cure fraction for the time to infection it was verified that the estimated proportions of cure fraction are given by 72.56% and 72.55% for the mixture and non-mixture models, respectively, whereas from the Kaplan-Meier estimate the cure fraction is 72.80%. In addition, for the time to metastasis the estimated proportions of cure fraction are given by 62.61% and 62.46% assuming the mixture and non-mixture models, respectively, and it is 62.20% from the Kaplan-Meier estimate. In addition, Figure 15 illustrate the plots of theoretical fitted survival functions based on the maximum likelihood estimates along with the empirical survival functions in presence of censored data using the Kaplan-Meier method. It is observed that the DW mixture and non-mixture models have better fit than the DW distribution with no cure rate and also captured the cure rate with great accuracy.

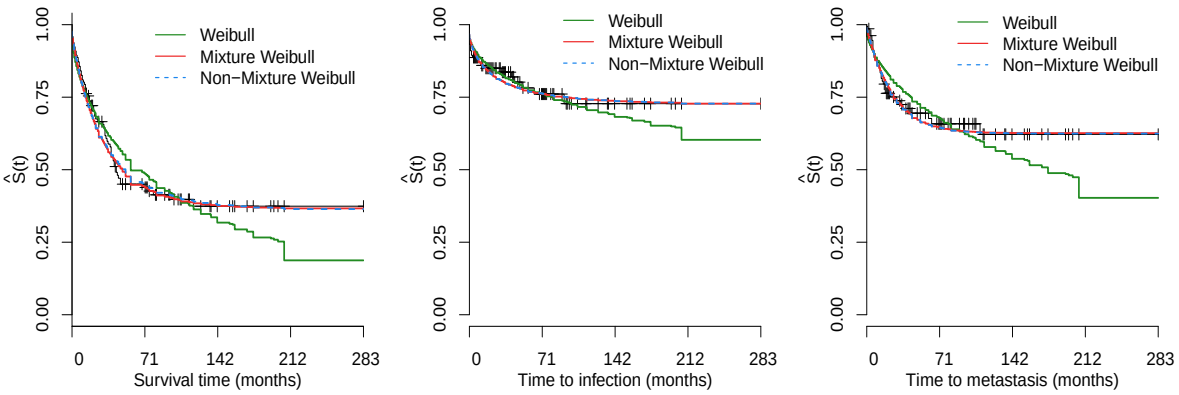

Figure 15 – Kaplan-Meier curves and the estimated survival functions for three times: Survival, infection and metastasis time.

Finally, it is compared in Figure 16 the empirical estimates based on the Kaplan-Meier versus the corresponding predicted values obtained from the parametric models.

From these plots, it is observed that the predicted values obtained from the mixture and non-mixture models are closer to the empirical values than those obtained from the standard discrete Weibull. Nevertheless, for the time to infection it seems that no one of the models have a satisfactory fit.

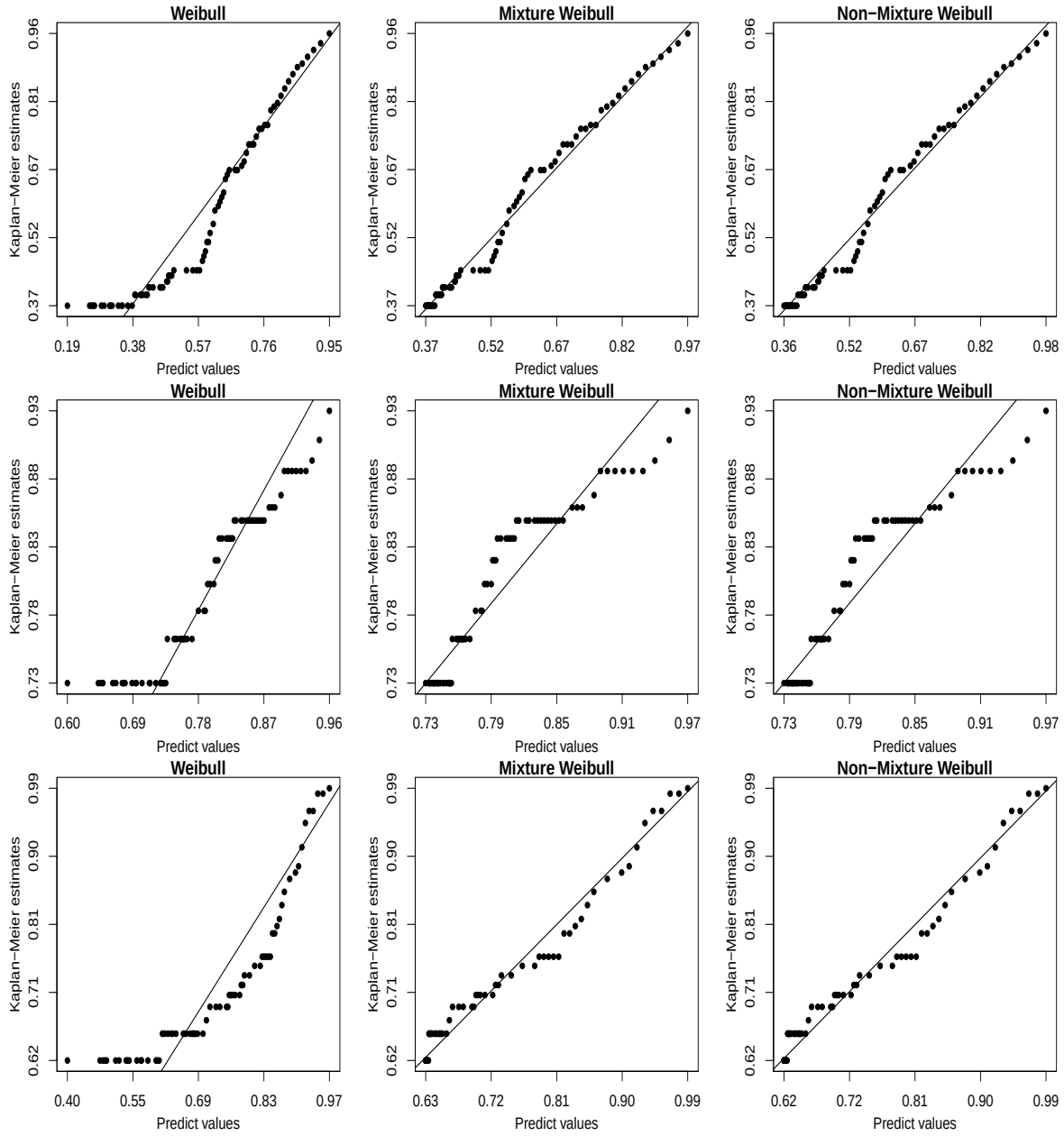

Figure 16 – Plots of the Kaplan-Meier estimates for the survival function versus the respective predicted values obtained from the parametric models for the three lifetimes. (**Upper Panel:** Survival time. **Middle Panel:** Infection time. **Lower Panel:** Metastasis time.)

Suppose now that the medical interest is investigate possible prognostic/diagnose/influence factors for pelvic sarcoma. In this case, under a regression model approach, let us consider the following covariates associated to each patient:

→ age40$_i$: age of the patients at start of follow-up, classified as less than 40 years (age40$_i = 0$) versus greater or equal to 40 years (age40$_i = 1$);

→ sex$_i$: patient gender, classified as male (sex$_i = 0$) versus female (sex$_i = 0$);

→ chemo$_i$: whether the patient received chemotherapy (chemo$_i = 1$) or not (chemo$_i = 0$);

→ radio$_i$: whether the patient received radiotherapy (radio$_i = 1$) or not (radio$_i = 0$);

→ grade$_i$: tumor grade of the patient classified as G3 (grade$_i = 1$) versus G2 (grade$_i = 0$);

→ log2volume$_i$: tumor volume (log2-transformed).

In our analysis considering regression models, it was assumed that the covariates affect the probability of being cured and the parameter ϕ, considering the following regression structures:

$$\begin{aligned} \log\left(\frac{\rho_i}{1-\rho_i}\right) &= \delta_0 + \delta_1\,\text{age40}_i + \delta_2\,\text{sex}_i + \delta_3\,\text{chemo}_i \\ &+ \delta_4\,\text{radio}_i + \delta_5\,\text{grade}_i + \delta_6\,\text{log2volume}_i \end{aligned} \tag{4.12}$$

and,

$$\begin{aligned} \log\left(-\log(\phi_i)\right) &= \alpha_0 + \alpha_1\,\text{age40}_i + \alpha_2\,\text{sex}_i + \alpha_3\,\text{chemo}_i \\ &+ \alpha_4\,\text{radio}_i + \alpha_5\,\text{grade}_i + \alpha_6\,\text{log2volume}_i \end{aligned} \tag{4.13}$$

for $i = 1, \ldots, 147$. The inference results for the survival time, time to infection and time to metastasis are reported in Tables 5, 6 and 7, respectively.

From the obtained results of Table 5 it was observed that for both models (mixture and non-mixture) the covariates grade (δ_5) and tumor volume (δ_6) affect the overall survival to death since the zero value is not included in the 95% confidence intervals. The significance of these covariates indicate that the probability of being "cured", that is, not dying due to sarcoma pelvic, depends on patient's tumor grade and volume. Patients who have tumor grade G3 seem to have lower chance of not dying compared to patients with tumor grade G2. In addition, we can observe that as the tumor volume increases the odds of not dying due to pelvic sarcoma decreases, since the parameter δ_6 associated to the tumor volume has a negative sign. A difference exist regarding the significance of the covariate age40 (δ_1) and radio (δ_4), because they are statistically not significants in the mixture model. For the DW mixture model only the covariate tumor volume (α_6) affects the median of the survival time, whereas in the non-mixture model the covariates age40 (α_1) and radio (α_4) are also significant at the 5% significance level.

Table 5 – Inference results for the discrete Weibull cure rate regression models — Survival time.

Model	Parameter	Mixture			Non-Mixture		
		MLE	SE	95% C.I.	MLE	SE	95% C.I.
Reg. Model (3.14)	δ_0	6.2878	2.1468	(2.0431, 10.5325)	7.0469	1.8754	(3.3389, 10.7549)
	δ_1	-1.6036	0.8449	(-3.2741, 0.0669)	-1.8573	0.7275	(-3.2956, -0.4189)
	δ_2	-0.3822	0.4918	(-1.3546, 0.5901)	-0.2539	0.5196	(-1.2813, 0.7734)
	δ_3	0.1498	0.6348	(-1.1054, 1.4050)	-0.0420	0.6864	(-1.3991, 1.3151)
	δ_4	-1.4408	0.7506	(-2.9248, 0.0433)	-1.5720	0.6847	(-2.9258, -0.2181)
	δ_5	-2.2784	0.7974	(-3.8551, -0.7018)	-2.4059	0.7795	(-3.9471, -0.8648)
	δ_6	-0.4533	0.1567	(-0.7632, -0.1434)	-0.5113	0.1483	(-0.8046, -0.2181)
Reg. Model (3.15)	α_0	-0.4324	2.1326	(-4.6489, 3.7841)	1.3765	1.3332	(-1.2594, 4.0124)
	α_1	-0.7533	0.7378	(-2.2121, 0.7054)	-1.4037	0.5125	(-2.4171, -0.3904)
	α_2	-0.5068	0.3380	(-1.1750, 0.1614)	-0.5165	0.3768	(-1.2615, 0.2286)
	α_3	-0.6700	0.4044	(-1.4695, 0.1296)	-0.9122	0.5152	(-1.9309, 0.1065)
	α_4	-0.7493	0.6431	(-2.0208, 0.5222)	-1.2694	0.4968	(-2.2516, -0.2872)
	α_5	0.2029	0.6441	(-1.0705, 1.4764)	-0.4409	0.5908	(-1.6090, 0.7272)
	α_6	-0.2206	0.1108	(-0.4397, -0.0015)	-0.3697	0.1065	(-0.5802, -0.1591)
	β	1.1115	0.1793	(0.7569, 1.4661)	1.1596	0.1375	(0.8877, 1.4315)

MLE: maximum likelihood estimates; SE: standard error; 95% C.I.: 95% confidence interval.

Figure 17 presents the half-normal with simulated envelope considering the randomized quantile residuals. It is observed, for both models, that all points lie inside the envelopes, suggesting that there is no serious violation of the model assumptions. Additionally, it is noteworthy that the non-mixture model fits the data better than the mixture, since the observed residuals are closer to the median line.

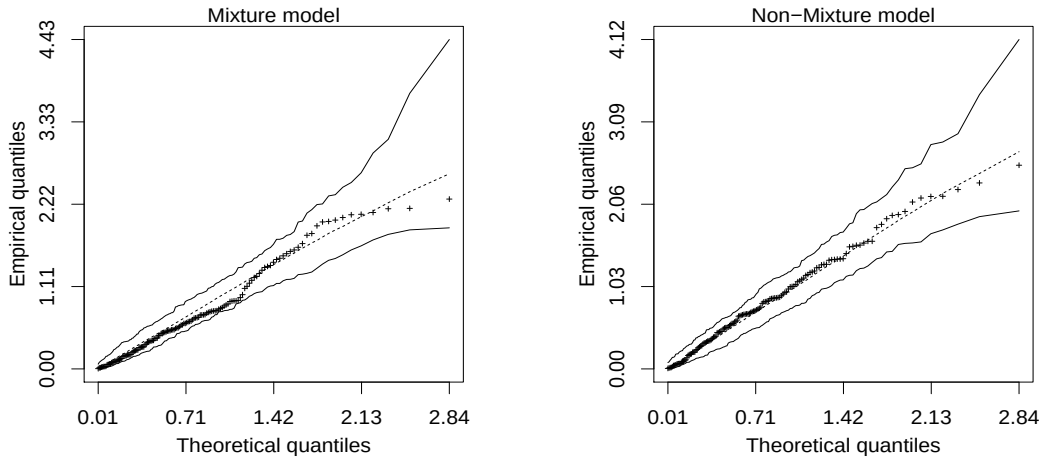

Figure 17 – Half-normal plot with simulated envelope for the randomized quantile residuals — Survival time.

In respect to the time to infection the results from Table 6 shows that all covariates do not affect the probability of being cured assuming the mixture model, since the 95% confidence intervals include the zero value. On the other hand, the covariate age40 (δ_1) was only statistically significant under the non-mixture model, although the upper limit of the 95% confidence interval for δ_1 assuming the non-mixture model is very close to zero, indicating some significance of the covariate age40.

Table 6 – Inference results for the discrete Weibull cure rate regression models - Time to infection.

Model	Parameter	Mixture			Non-Mixture		
		MLE	SE	95% C.I.	MLE	SE	95% C.I.
Reg. Model (3.14)	δ_0	-16.4294	11.3838	(-38.9385, 6.0798)	-9.5600	7.3180	(-23.9030, 4.7830)
	δ_1	-3.6810	1.9659	(-7.5682, 0.2062)	-3.2156	1.6364	(-6.4229, -0.0082)
	δ_2	8.0530	4.6309	(-1.1036, 17.2096)	4.7289	2.6811	(-0.5260, 9.9837)
	δ_3	7.6546	4.7980	(-1.8324, 17.1417)	4.7702	2.9249	(-0.9625, 10.5030)
	δ_4	-3.9429	2.3855	(-8.6598, 0.7740)	-3.1499	1.7926	(-6.6632, 0.3635)
	δ_5	-3.9574	2.7146	(-9.3250, 1.4102)	-2.1437	1.8444	(-5.7586, 1.4712)
	δ_6	1.1875	0.7627	(-0.3206, 2.6956)	0.7942	0.5411	(-0.2664, 1.8548)
Reg. Model (3.15)	α_0	-8.5747	1.9547	(-12.4398, -4.7095)	-11.9181	3.0884	(-17.9713, -5.8650)
	α_1	-0.9091	0.6660	(-2.2261, 0.4079)	-2.2399	1.0368	(-4.2719, -0.2079)
	α_2	1.9000	0.8354	(0.2483, 3.5518)	3.2018	0.9450	(1.3496, 5.0540)
	α_3	2.3371	0.8708	(0.6154, 4.0589)	3.6894	1.0830	(1.5668, 5.8120)
	α_4	-1.4132	0.7004	(-2.7981, -0.0284)	-2.5335	1.0909	(-4.6717, -0.3954)
	α_5	-0.7874	0.7070	(-2.1855, 0.6106)	-1.2127	1.0060	(-3.1844, 0.7590)
	α_6	0.6456	0.1777	(0.2942, 0.9969)	0.9284	0.2783	(0.3830, 1.4738)
	β	0.3737	0.0895	(0.1968, 0.5507)	0.4387	0.1170	(0.2095, 0.6680)

MLE: maximum likelihood estimates; SE: standard error; 95% C.I.: 95% confidence interval.

A visual inspection of the half-normal plots given in Figure 18 suggests that although all points lie inside the envelope, indicating that there is no serious violation of the model assumptions, it is also observed that the DW mixture model does not fit the data as well as the DW non-mixture model, since most of the observed residuals of the mixture model are near to the boundary of the envelope.

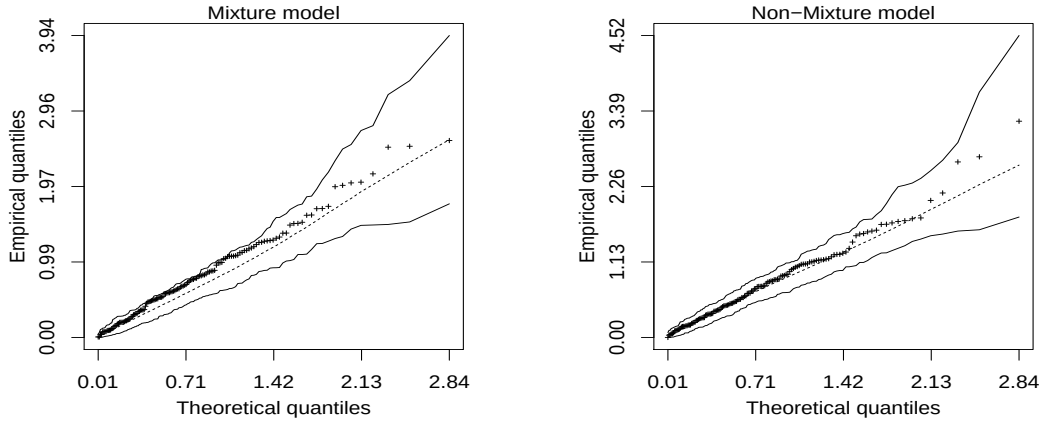

Figure 18 – Half-normal plot with simulated envelope for the randomized quantile residuals - Time to infection.

The results of Table 7 related to the metastasis times indicate that the covariate chemotherapy (δ_3) and tumor volume are statistically significant for both the mixture and non-mixture models. The significance of these covariates reveal that the probability of being "cured", depends on the patient receiving or not receiving chemotherapy, which means that individuals who received the chemotherapy seem to have lower chance of not present metastasis compared to individuals who do not received chemotherapy. Additionally, we expected that as the tumor volume increases the odds of not presented metastasis decreases, since the parameter δ_6 has negative sign.

Table 7 – Inference results for the discrete Weibull cure rate regression models — Time to metastasis.

Model	Parameter	Mixture			Non-Mixture		
		MLE	SE	95% C.I.	MLE	SE	95% C.I.
Reg. Model (3.14)	δ_0	7.1233	1.9217	(3.3570, 10.8897)	6.9469	1.8971	(3.2288, 10.6651)
	δ_1	-0.3116	0.6474	(-1.5805, 0.9573)	-0.2062	0.6569	(-1.4938, 1.0813)
	δ_2	-0.6215	0.5210	(-1.6425, 0.3996)	-0.6769	0.5173	(-1.6908, 0.3370)
	δ_3	-1.8591	0.7426	(-3.3146, -0.4036)	-1.8487	0.7351	(-3.2894, -0.4079)
	δ_4	0.3352	0.6439	(-0.9269, 1.5973)	0.4549	0.6303	(-0.7805, 1.6904)
	δ_5	-1.1753	0.9068	(-2.9526, 0.6021)	-1.1484	0.8935	(-2.8996, 0.6028)
	δ_6	-0.4700	0.1669	(-0.7971, -0.1429)	-0.4570	0.1645	(-0.7795, -0.1345)
Reg. Model (3.15)	α_0	-4.8403	1.5214	(-7.8222, -1.8584)	-4.2341	1.7083	(-7.5823, -0.8858)
	α_1	-0.2126	0.6624	(-1.5109, 1.0858)	-0.0527	0.7592	(-1.5407, 1.4353)
	α_2	-0.3770	0.4503	(-1.2596, 0.5057)	-0.6333	0.5045	(-1.6221, 0.3554)
	α_3	-2.5830	0.6866	(-3.9287, -1.2373)	-2.9090	0.7692	(-4.4166, -1.4014)
	α_4	0.9750	0.5980	(-0.1970, 2.1471)	1.2403	0.6482	(-0.0301, 2.5108)
	α_5	3.6399	0.9562	(1.7659, 5.5140)	3.6976	1.0326	(1.6736, 5.7215)
	α_6	-0.1645	0.1219	(-0.4035, 0.0745)	-0.2751	0.1368	(-0.5433, -0.0069)
	β	1.7503	0.2417	(1.2766, 2.2239)	1.8718	0.2582	(1.3658, 2.3778)

MLE: maximum likelihood estimates; SE: standard error; 95% C.I.: 95% confidence interval.

The simulated envelope plots of the mixture and non-mixture models corresponding to the metastasis time are shown in Figure 19. For both models, we can see that most of the observed randomized quantile residuals are within the simulated envelope, showing no evidence of violated model assumptions.

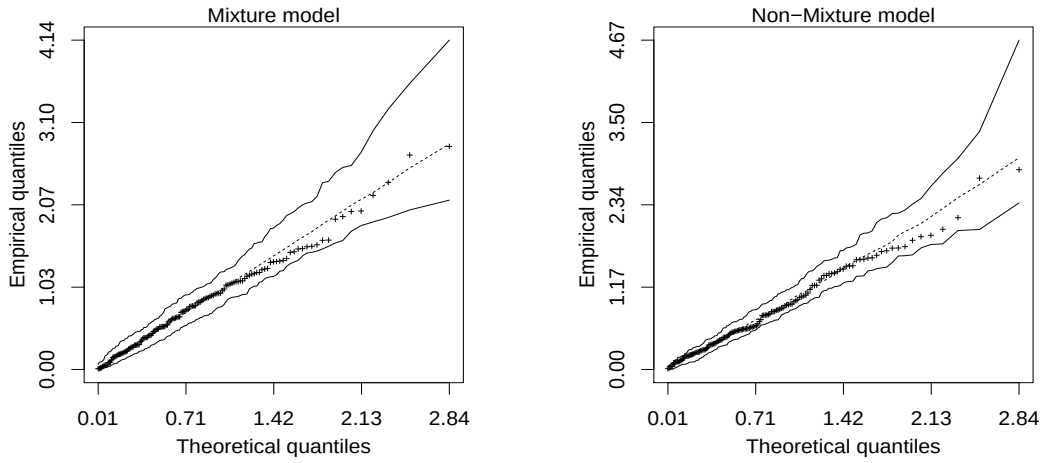

Figure 19 – Half-normal plot with simulated envelope for the quantile residuals - Time to metastasis.

Table 8 reports the AIC values considering the regression model and the model not including the presence of covariates. From these results it is observed that the mixture and non-mixture models have similar AIC values. Regardless the response variable it is verified that the model with covariates presented the smallest values of AIC. However, for the time to infection the AIC values indicated no improvement on the fit to the data, as expected, since for this time the covariates were not significant.

Table 8 – Values of model discrimination criteria (AIC) for mixture and non-mixture models.

Time	Mixture		Non-Mixture	
	Without covariates	With covariates	Without covariates	With covariates
Survival	751.9783	694.6915	754.2051	697.4294
Infection	309.0108	267.6405	309.0300	270.0043
Metastasis	411.5414	370.2743	411.3725	370.9714

4.2.4 Discussion and Remarks

The main goal of this study was the introduction of mixture and non-mixture cure fraction models assuming a discrete Weibull distribution in place of standard existing continuous lifetime distributions, with special application to the statistical analysis of a dataset related to long-term oncological treatment outcomes of resection of pelvic sarcomas. The obtained results of this study show many advantages for the use of discrete cure fraction models in terms of great accuracy for the obtained point and interval inferences, great computational simplicity to get the inferences of interest under the classical approach and simple interpretations for the parameters of the models which is an important point in medical applications. It was also observed that using regression models, the identification of important covariates was easily obtained with good accuracy assuming the DW model due to the best simplicity of the likelihood function assuming the DW model when compared to standard continuous Weibull cure fraction models commonly used in the analysis of lifetime data in presence of cure fraction, censored data and covariates. Finally, the general framework for the computer codes of the proposed modeling approach is presented in Appendix A at the end of this thesis which could be carried out using the OpenBugs or R softwares.

4.3 Study 2: Tobacco-Stained Fingers Data

4.3.1 Introduction

According to World Health Organization (2018), tobacco is one of the biggest public health threats the world has ever faced, killing more than 7 million people a year. In addition, around 80% of the 1.1 billion smokers worldwide live in low- and middle-income countries, where the burden of tobacco-related illness and death tt is too high. In recent decades, many studies published in the literature show the health risks for the use of tobacco, as the studies of Wynder and Graham (1988); Peto et al. (1992); Anthony et al. (1994); Giovino et al. (1995); Organization et al. (1997); Fiore (2000); Initiative (2004); of Health and Services (2006); Chapman (2007); Organization and for International Tobacco Control (2008); Tobacco et al. (2008); John et al. (2013, 2015a);

Food et al. (2016). Moreover, according to the World Health Organization (2018), many studies show that few people understand the specific health risks of tobacco use and among smokers who are aware of the dangers of tobacco, where most of the smokers want to quit. In this case, counseling and medication can more than double the chance that a smoker who tries to quit will succeed.

According to John et al. (2013), although cigarette smoking is by far the most common risk factor for chronic obstructive pulmonary disease, it is strongly associated with cardiovascular disease, and is responsible for 30% of all cancer deaths. In a study by Ezzati and Lopez (2004); Team (2011), some additional factors for tobacco-related disease are involved in determining each individual's susceptibility as well identifying smokers at a higher risk of developing tobacco disease. In addition, according to John et al. (2013), smoking has been linked to many skin conditions for more than 150 years and evidence exists that skin involvement might be a conspicuous marker of other tobacco-related disease. Tobacco stains on fingers are commonly seen among smokers where behavioral and environmental factors may be important in the stain development (see Thakurdas, 1963; Mitchell and Dahlgren, 1986; Hafezi et al., 2001; John et al., 2013, 2015a). For example, yellow-stained fingers, often observed in psychiatric units and among young drug users (autopsy series), support a possible association between those conditions and tar deposition as illustrated in Figure 20.

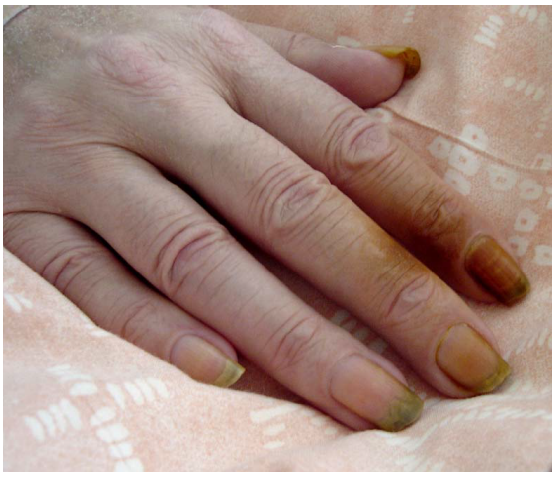

Figure 20 – Tar staining seen in a 70-year-old male smoker (**Source:** John et al., 2013)

.

4.3.2 Material and Methods

In our present study, it is assumed a real dataset introduced by John et al. (2015b). This data set reports a retrospective cohort study in a population of 143 smokers screened between March 2006 and January 2010 in a 180-bed community hospital in La Chaux-de-Fonds, Switzerland (see Table 9). The interested reader should consult John et al. (2013) and John et al. (2015a,b) for further details about this data set.

Table 9 – TSF data structure (partial): A clue for smoking-related disease (full structure and data available in John et al., 2015a,b).

Variable	Description
N	*Patient group: tobacco-stain or controls*
Sex	*Patient sex: male or female*
Age	*Patient age*
Time to Death (in Days)	*Survival time until death*
Death	*0 = alive, 1 = dead*
Readmission Binary	*0 = no readmission, 1 = readmission*
Time to Readmission (Days)	*Time before the first readmission*
Readmission Causes	*Reason for readmission: psychiatric, alcohol, tobacco, cardiopulmonary, pulmonary and cancer*
Follow-up	*Follow-up time*
Weight	*Patient weight*
Hypertension	*0 = no, 1 = yes*
Harmful Alcohol Use	*0 = no, 1 = yes*
Depression	*0 = no, 1 = yes*
Tobacco Years	*Years of tobacco use*
Pack Per Day	*Number of packs of tobacco per day*
Stain Past	*0 = no, 1 = yes*

For the statistical analysis, it is considered as lifetime T_1 the time before the first readmission which was censored in case of death before the closure date; and as lifetime T_2 the survival time of the patient. For T_1 there are 105 censored observations and 38 not censored observations; and for T_2 there are 69 censored observations and 74 not censored observations. Both times were measured in a discrete way (number of months). The non-parametric estimators for the means obtained from Kaplan-Meier estimators Kaplan and Meier (1958) are given, respectively, by 31.58 months for T_1 and 68.49 months for T_2. Moreover, it is assumed the proposed cure rate models described in Chapter 3. In the case of BL models, only the BL-I was considered here. For a Bayesian approach, it is assumed Beta prior distributions with hyperparameter values $(\alpha, \beta) = (1, 1)$ for the discrete model parameters; Gamma prior distributions with hyperparameter values $(\alpha, \beta) = (0.001, 0.001)$ for the continuous model parameters and also Beta prior distributions for the incidence parameters $\phi_{11}, \phi_{10}, \phi_{01}$ and ϕ_{00} with hyperparameter values $(\alpha_i, \beta_i) = (1, 1), i = 1, 2, 3, 4$. Note that with these values for the hyperparameters, it is assumed approximately non-informative prior distributions for the parameters of the proposed models. In Table 10, it is presented the posterior summaries of interest.

In Table 11, it is presented the Monte Carlo estimates of DIC for each model. From the results in Table 11, using DIC as the discrimination criteria, it could be concluded that the BL-I model has a good fit among the continuous models and the BDBG model has a good fit among the discrete models to the TSF data set in presence of cure rate

Table 10 – Posterior summaries for the proposed models in presence of cure rate for the TSF data set.

Continuous Models				Discrete Models			
Model	Par.	Post. Mean (S.D.)	95% Cred. Interval	Model	Par.	Post. Mean (S.D.)	95% Cred. Interval
BBBE	λ_1	0.0390 (0.0143)	(0.0090, 0.0614)	ABG	θ_1	0.0530 (0.0057)	(0.0428, 0.0649)
	λ_2	0.0175 (0.0089)	(0.0022, 0.0342)		θ_2	0.0279 (0.0052)	(0.0188, 0.0387)
	λ_3	0.0167 (0.0161)	(0.0001, 0.0480)		ϕ_{00}	0.1402 (0.0302)	(0.0890, 0.2063)
	ϕ_{00}	0.1356 (0.0307)	(0.0803, 0.2024)		ϕ_{01}	0.0124 (0.0127)	(0.0005, 0.0476)
	ϕ_{01}	0.0126 (0.0121)	(0.0004, 0.0503)		ϕ_{10}	0.3250 (0.0539)	(0.2100, 0.4225)
	ϕ_{10}	0.3327 (0.0531)	(0.2304, 0.4283)		ϕ_{11}	0.5224 (0.0556)	(0.4274, 0.6421)
	ϕ_{11}	0.5191 (0.0541)	(0.4233, 0.6170)				
GBE	λ_1	0.0521 (0.0056)	(0.0423, 0.0641)	BDBG	θ_1	0.9501 (0.0053)	(0.9392, 0.9595)
	λ_2	0.0260 (0.0052)	(0.0168, 0.0363)		θ_2	0.9765 (0.0054)	(0.9651, 0.9867)
	θ	0.0020 (0.0105)	(0.0001, 0.0172)		θ_{12}	0.9955 (0.0025)	(0.9900, 0.9992)
	ϕ_{00}	0.1371 (0.0312)	(0.0752, 0.2034)		ϕ_{00}	0.1380 (0.0302)	(0.0853, 0.1997)
	ϕ_{01}	0.0133 (0.0131)	(0.0003, 0.0493)		ϕ_{01}	0.0147 (0.0143)	(0.0004, 0.0526)
	ϕ_{10}	0.3151 (0.0524)	(0.2102, 0.4175)		ϕ_{10}	0.3236 (0.0540)	(0.1986, 0.4189)
	ϕ_{11}	0.5345 (0.0564)	(0.4368, 0.6472)		ϕ_{11}	0.5237 (0.0554)	(0.4246, 0.6459)
MOBE	λ_1	0.0521 (0.0062)	(0.0413, 0.0650)	BG-Type II	θ_1	0.9672 (0.0068)	(0.9529, 0.9801)
	λ_2	0.0002 (0.0009)	(0.0001, 0.0024)		θ_2	0.9627 (0.0040)	(0.9543, 0.9700)
	λ_{12}	0.0122 (0.0038)	(0.0069, 0.0212)		θ_{12}	0.9686 (0.0081)	(0.9518, 0.9826)
	ϕ_{00}	0.1474 (0.0355)	(0.0876, 0.2191)		ϕ_{00}	0.1002 (0.0416)	(0.0073, 0.1789)
	ϕ_{01}	0.0242 (0.0188)	(0.0019, 0.0728)		ϕ_{01}	0.0035 (0.0033)	(0.0001, 0.0122)
	ϕ_{10}	0.1785 (0.1010)	(0.0086, 0.3715)		ϕ_{10}	0.4355 (0.0526)	(0.3398, 0.5454)
	ϕ_{11}	0.6499 (0.1076)	(0.4560, 0.8384)		ϕ_{11}	0.4608 (0.0387)	(0.3859, 0.5394)
BL-I	β_1	0.1009 (0.0084)	(0.0857, 0.1185)	DBGR	α	0.0277 (0.0144)	(0.0063, 0.0608)
	β_2	0.0084 (0.0067)	(0.0006, 0.0243)		λ_1	0.9997 (0.0002)	(0.9993, 0.9999)
	β_3	0.0158 (0.0040)	(0.0094, 0.0244)		λ_2	0.9999 (0.0001)	(0.9997, 0.9999)
	ϕ_{00}	0.1604 (0.0315)	(0.1035, 0.2173)		ϕ_{00}	0.1453 (0.0297)	(0.0876, 0.2036)
	ϕ_{01}	0.0198 (0.0167)	(0.0018, 0.0607)		ϕ_{01}	0.0085 (0.0085)	(0.0002, 0.0321)
	ϕ_{10}	0.2809 (0.0722)	(0.1119, 0.4027)		ϕ_{10}	0.3420 (0.0415)	(0.2631, 0.4307)
	ϕ_{11}	0.5389 (0.0733)	(0.4281, 0.7009)		ϕ_{11}	0.5042 (0.0420)	(0.4159, 0.5793)

parameters. Overall, the BDBG is the best model to be fitted for the TSF data set. However, the DIC criterion in medical studies could not be a good indicator to model discrimination/selection.

Table 11 – DIC values for each model for the DRS data set assuming the presence of cure rate parameters.

Continuous Models		Discrete Models	
Model	DIC value	Model	DIC value
GBE	1703.563	ABG	1692.545
BBBE	1703.171	BDBG	1692.223
MOBE	1736.251	BG-Type II	1702.339
BL-I	1696.983	DBGR	1756.366

In this way, it also considered the marginal Kaplan-Meier and survival probability plots as a discrimination criteria in contrast to DIC criteria. Thus, by observing the Kaplan-Meier curves and the Kaplan-Meier estimates versus the predicted value plots

(see, Figures 21, 22, 23 and 24) for both times T_1 and T_2, we could conclude that the proposed models captured the cure rate with good accurate; the estimated credibility bonds incorporate the non-parametric Kaplan-Meier curve and the results are quite similar indicating that these models could be an alternative to describe and predict the time before the first readmission and the survival time.

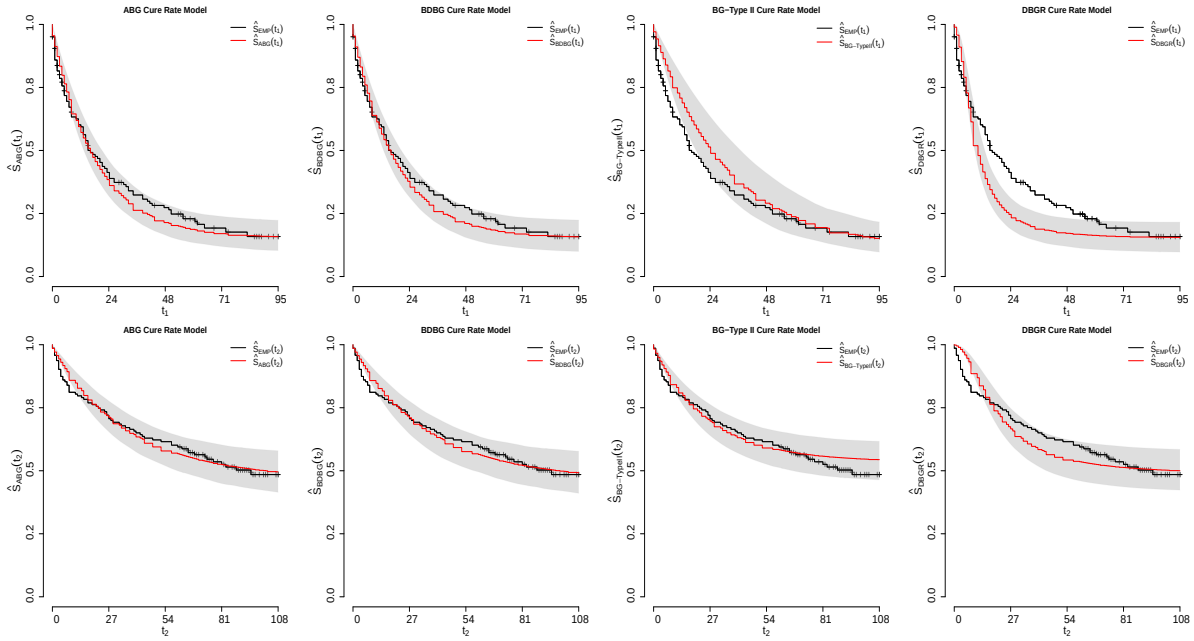

Figure 21 – Kaplan-Meier estimators versus discrete marginal fitted survival functions for T_1 (upper panels) and T_2 (lower panels) for TSF dataset.

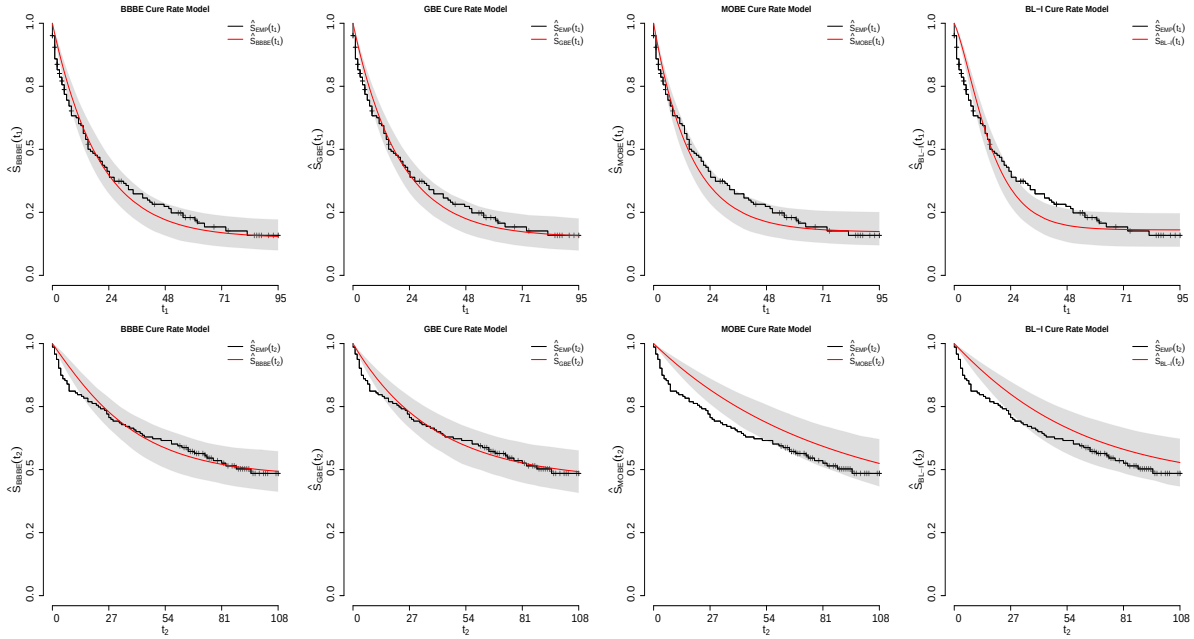

Figure 22 – Kaplan-Meier estimators versus continuous marginal fitted survival functions for T_1 (upper panels) and T_2 (lower panels) for TSF dataset.

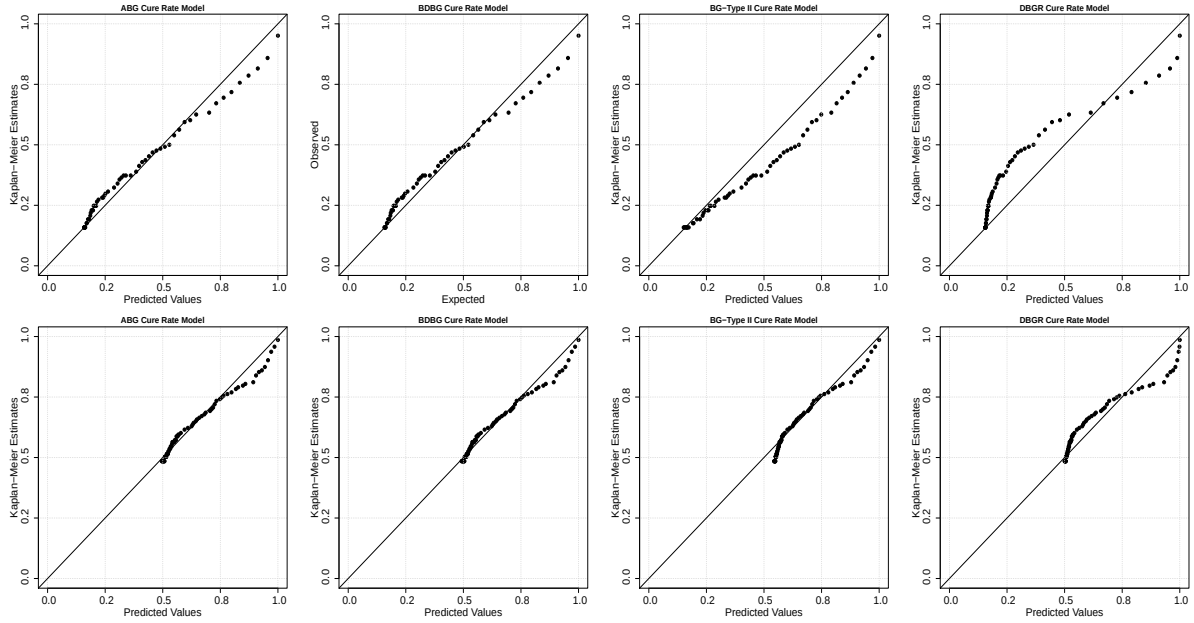

Figure 23 – Plots of the Kaplan-Meier estimates for the survival function versus the respective predicted values obtained from the proposed discrete models for T_1 (upper panels) and T_2 (lower panels).

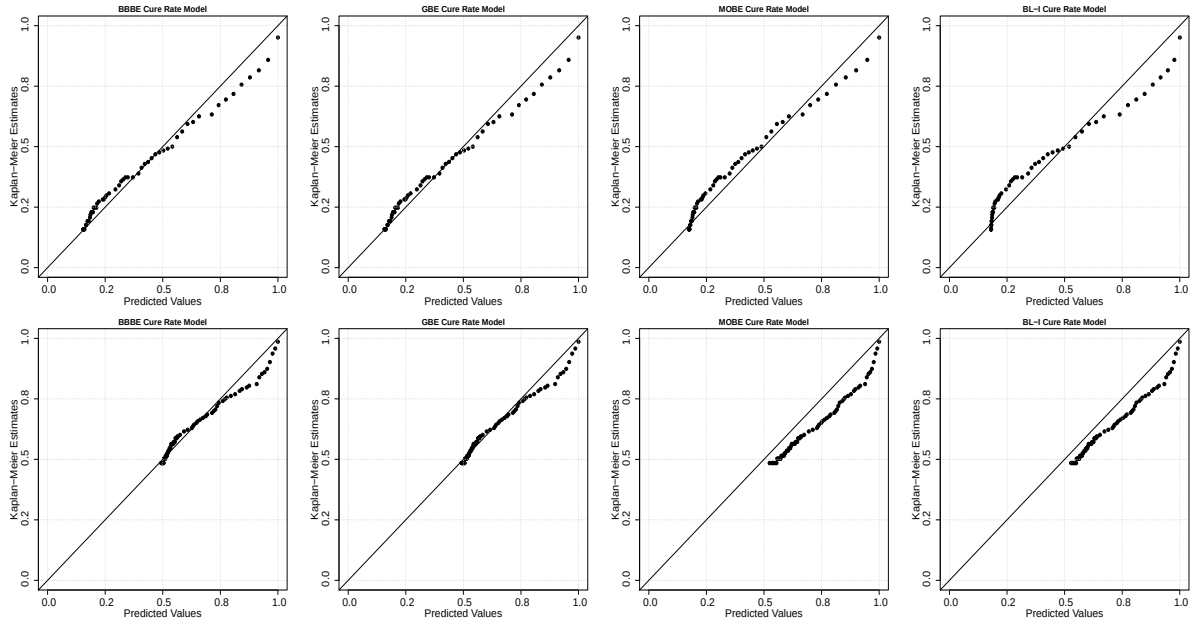

Figure 24 – Plots of the Kaplan-Meier estimates for the survival function versus the respective predicted values obtained from the proposed continuous models for T_1 (upper panels) and T_2 (lower panels).

Finally, the estimated proportions of non-susceptible patients for the time before the first readmission (T_1) and for the survival time (T_2), are given in Table 12. The estimated proportions of non-susceptible individuals obtained from the non-parametric Kaplan-Meier estimator are given, respectively by 15.80% for time T_1 and 48.48% for time T_2. Therefore, from the obtained inference results of this application, it could be

concluded that the DBGR model has better fit for the cure rate under times T_1 and T_2. However, as stated in Kaplan-Meier plots, the DBGR model could be worse to predict survival probabilities if we look the times before the plateau that indicates the cure rate incidence. In this case, the second best fitted model, the BDBG model, could provide a better fit instead of the DBGR model.

Table 12 – The estimated proportions of non-susceptible patients for the TSF data set.

Continuous Models			Discrete Models		
Model	T_1 $(\phi_{00}+\phi_{01})$	T_2 $(\phi_{00}+\phi_{10})$	Model	T_1 $(\phi_{00}+\phi_{01})$	T_2 $(\phi_{00}+\phi_{10})$
GBE	15.05%	45.22%	AGB	15.26%	46.52%
BBBE	14.82%	46.83%	BDBG	15.27%	46.16%
MOBE	17.16%	32.59%	BG-Type II	10.37%	53.57%
BL-I	18.02%	44.13%	DBGR	15.38%	48.73%

4.3.3 Discussion and Remarks

Our analysis of the TSF dataset illustrated a new way to predict survival and readmission probabilities for the tobacco-stained fingers study. Based on the analysis, the findings of the present study demonstrated, under a nonparametric approach, that 15.80% of the patients are non-susceptible to the readmission event while 48.48% of the patients are long-term survivors. In contrast, under a parametric approach, the BDBG model demonstrated that 15.27% of the patients are non-susceptible to readmission event while 46.16% of the patients are long-term survivors. The DBGR model provided better approximation for the cure rate (15.38% of the patients are non-susceptible to readmission event while 48.73% of the patients are long-term survivors), however this model could be worse to predict survival probabilities if we look the time before the plateau that indicates the cure rate incidence. In general, the proposed models has a good accuracy for the estimation of the long-term survivors. Finally, in terms of computational aspects, the convergence was faster for discrete models even using non-informative prior distributions. The continuous models showed some instabilities and higher number of MCMC iterations to get the convergence. The general framework for the computer codes of the proposed modeling approach is presented in Appendix A at the end of this thesis which could be carried out using the OpenBugs or R softwares.

4.4 Study 3: Diabetic Retinopathy

4.4.1 Introduction

According to Retina Labs (2018) and Mayo Clinic (2018), the diabetic retinopathy disease is a chronic progressive, potentially sight-threatening disease of the retinal microvasculature associated with the prolonged hyperglycaemia. It is caused by damage

to the blood vessels of the light-sensitive tissue at the back of the eye called the retina, which process light and vision for the brain. Over time, diabetes damages the blood vessels in the retina which could cause the retinal tissue to swell, resulting in blurred vision (see Figure 25).

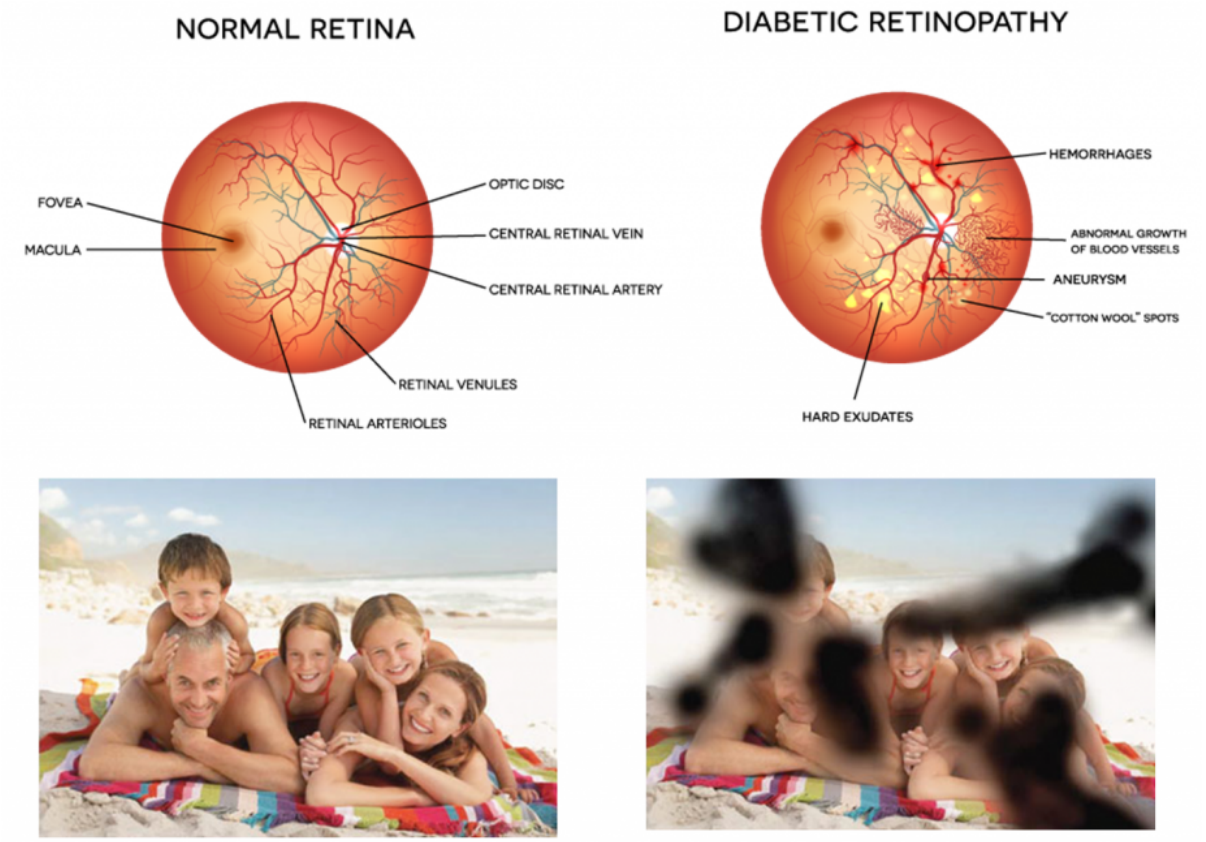

Figure 25 – An illustration of normal retina and diabetic retinopathy retina (**Source:** Retina Labs, 2018).

For *The National Eye Institute* (Institute, 2018), diabetic retinopathy disease may progress through four stages:

1. **Mild nonproliferative retinopathy:** Small areas of balloon-like swelling in the retina's tiny blood vessels, called microaneurysms, occur at this stage.

2. **Moderate nonproliferative retinopathy:** As the disease progresses, blood vessels that nourish the retina may swell and distort.

3. **Severe nonproliferative retinopathy:** Many more blood vessels are blocked, depriving blood supply to areas of the retina. These areas secrete growth factors that signal the retina to grow new blood vessels.

4. **Proliferative diabetic retinopathy (PDR):** At this advanced stage, growth factors secreted by the retina trigger the proliferation of new blood vessels, which grow along the inside surface of the retina and into the vitreous gel.

According to *The National Eye Institute* (Institute, 2018), vision lost to diabetic retinopathy is sometimes irreversible. However, early detection and treatment can reduce the risk of blindness by 95 percent. It is important to point out that diabetic retinopathy often lacks early symptoms, people diagnosticated with diabetes should get a comprehensive dilated eye exam at least once a year. The treatment for diabetic retinopathy is often delayed until it starts to progress to PDR, or when diabetic macular edema occurs (DME). As reported by *The National Eye Institute* research (Institute, 2018), DME can be treated with several therapies that may be used alone or in combination:

- **Anti-VEGF Injection Therapy:** Anti-VEGF drugs are injected into the vitreous gel to block a protein called vascular endothelial growth factor (VEGF), which can stimulate abnormal blood vessels to grow and leak fluid. Blocking VEGF can reverse abnormal blood vessel growth and decrease fluid in the retina.
- **Focal/grid macular laser surgery:** In focal/grid macular laser surgery, a few to hundreds of small laser burns are made to leaking blood vessels in areas of edema near the center of the macula. Laser burns for DME slow the leakage of fluid, reducing swelling in the retina.
- **Corticosteroids:** Corticosteroids, either injected or implanted into the eye, may be used alone or in combination with other drugs or laser surgery to treat DME. The Ozurdex (dexamethasone) implant is for short-term use, while the Iluvien (fluocinolone acetonide) implant is longer lasting. Both are biodegradable and release a sustained dose of corticosteroids to suppress DME. Corticosteroid use in the eye increases the risk of cataract and glaucoma.

4.4.2 Material and Methods

In order to illustrate the proposed methodology in Chapter 3, it is considered a data set introduced by Huster et al. (1989) with 197 patients where 50% of the patients where classified by the authors as "high-risk" for diabetic retinopathy. Each patient had one eye randomized to laser treatment and the other eye received no treatment. For each eye, the event of interest was the time from beginning of the treatment to the time when visual acuity dropped below 2/200 (call it "blindness"). There was a built-in lag time of approximately six months (visits were every three months). Survival times in this data set are, therefore, the actual times to blindness in months, minus the minimum possible time to event (6.5 months). Censoring was caused by death, dropout or end of the study. This data set will be denoted as DRS and the data structure is presented in Table 13.

Table 13 – A trial of laser coagulation as a treatment to delay diabetic retinopathy data structure (available in R software).

Variable	Description
ID	*Subject ID*
Laser	*Type of laser used: xenon or argon*
Eye	*Which eye was treated: right or left*
Age	*Age at diagnosis of diabetes*
Type	*Type of diabetes: juvenile or adult*
Trt	*0 = control eye, 1 = treated eye*
Futime	*Time to loss of vision or last follow-up*
Status	*0 = censored, 1 = loss of vision in this eye*
Risk	*A risk score for the eye*

For the statistical analysis, it is assumed as lifetimes the time to blindness for the eye randomized to laser treatment (T_1), with 143 censored observations and 54 not censored observations, and the time to blindness for the eye randomized that not received the treatment (T_2), with 96 censored observations and 101 non-censored observations. Moreover, in Figure 26, it is presented the plots of the nonparametric Kaplan and Meier (1958) estimator of the survival functions for both times where could be seen the incidence of a cure rate; the nonparametric estimators for the means obtained from Kaplan-Meier estimators are given, respectively, by 53.72 months for T_1 and 43.52 months for T_2.

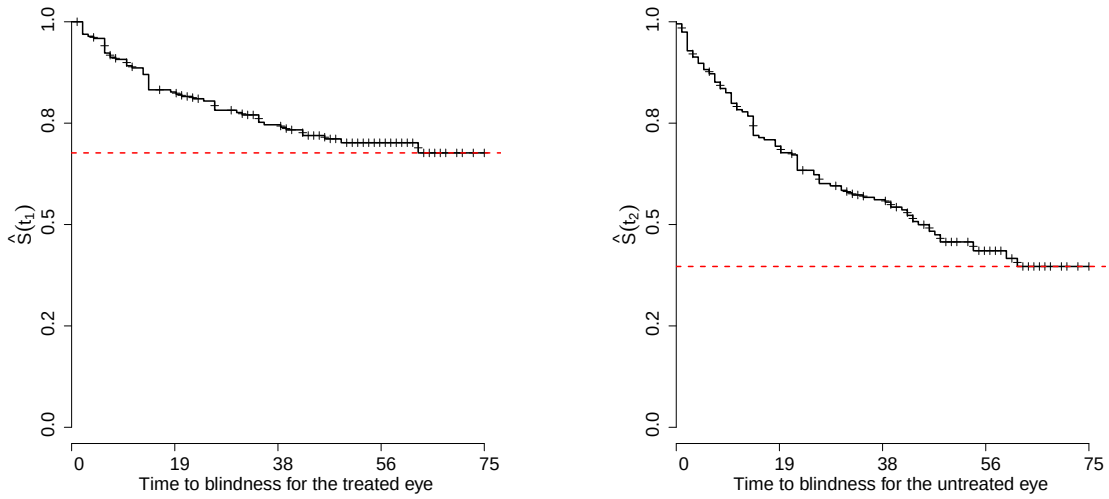

Figure 26 – Kaplan-Meier estimators for the survival functions of T_1 (time to blindness for the treated eye) and T_2 (time to blindness for the untreated eye).

4.4.2.1 Statistical Analysis in Presence of Cure Rate

As a first modeling approach for the DRS data set, it is assumed the proposed cure rate models described in Chapter 3. In the case of BL models, only the BL-I was considered here. For a Bayesian approach, it is assumed Beta prior distributions with

hyperparameter values $(\alpha, \beta) = (1, 1)$ for the discrete model parameters, Gamma prior distributions with hyperparameter values $(\alpha, \beta) = (0.001, 0.001)$ for the continuous model parameters and also Beta prior distributions for the incidence parameters $\phi_{11}, \phi_{10}, \phi_{01}$ and ϕ_{00} with hyperparameter values $(\alpha_i, \beta_i) = (1, 1), i = 1, 2, 3, 4$. Note that with these values for the hyperparameters, it is assumed approximately non-informative prior distributions for the parameters of the proposed models. In Table 14, it is presented the posterior summaries of interest.

Table 14 – Posterior summaries for the proposed models in presence of cure rate for the DRS data set.

Continuous Models				Discrete Models			
Model	Par.	Post. Mean (S.D.)	95% Cred. Interval	Model	Par.	Post. Mean (S.D.)	95% Cred. Interval
GBE	λ_1	0.0298 (0.0101)	(0.0124, 0.0484)	ABG	θ_1	0.0359 (0.0093)	(0.0179, 0.0544)
	λ_2	0.0320 (0.0057)	(0.0221, 0.0438)		θ_2	0.0333 (0.0059)	(0.0227, 0.0456)
	θ	0.0010 (0.0041)	(0.0001, 0.0142)		ϕ_{11}	0.3162 (0.0614)	(0.2179, 0.4646)
	ϕ_{11}	0.3489 (0.0903)	(0.2325, 0.5916)		ϕ_{10}	0.0478 (0.0289)	(0.0033, 0.1113)
	ϕ_{10}	0.0515 (0.0343)	(0.0031, 0.1193)		ϕ_{01}	0.3490 (0.0675)	(0.2039, 0.4776)
	ϕ_{01}	0.3278 (0.0892)	(0.0772, 0.4633)		ϕ_{00}	0.2870 (0.0499)	(0.1826, 0.3798)
	ϕ_{00}	0.2719 (0.0526)	(0.1614, 0.3647)				
BBBE	λ_1	0.0007 (0.0013)	(0.0001, 0.0044)	BDBG	θ_1	0.9724 (0.0084)	(0.9549, 0.9878)
	λ_2	0.0013 (0.0023)	(0.0001, 0.0080)		θ_2	0.9711 (0.0057)	(0.9596, 0.9813)
	λ_3	0.0485 (0.0085)	(0.0317, 0.0629)		θ_{12}	0.9937 (0.0029)	(0.9871, 0.9982)
	ϕ_{11}	0.3140 (0.0634)	(0.2126, 0.4496)		ϕ_{11}	0.3155 (0.0645)	(0.2158, 0.4667)
	ϕ_{10}	0.0802 (0.0281)	(0.0219, 0.1365)		ϕ_{10}	0.0598 (0.0312)	(0.0066, 0.1277)
	ϕ_{01}	0.3293 (0.0465)	(0.2501, 0.4226)		ϕ_{01}	0.3388 (0.0676)	(0.1955, 0.4632)
	ϕ_{00}	0.2765 (0.0551)	(0.1548, 0.3782)		ϕ_{00}	0.2858 (0.0507)	(0.1794, 0.3809)
MOBE	λ_1	0.0202 (0.0076)	(0.0015, 0.0341)	BG-Type II	θ_1	0.9546 (0.0064)	(0.9418, 0.9662)
	λ_2	0.0278 (0.0058)	(0.0178, 0.0393)		θ_2	0.9619 (0.0044)	(0.9525, 0.9697)
	λ_{12}	0.0054 (0.0026)	(0.0015, 0.0108)		θ_{12}	0.9463 (0.0119)	(0.9215, 0.9674)
	ϕ_{11}	0.3696 (0.0889)	(0.2415, 0.5718)		ϕ_{11}	0.1742 (0.0220)	(0.1327, 0.2193)
	ϕ_{10}	0.0722 (0.0438)	(0.0079, 0.1787)		ϕ_{10}	0.0893 (0.0154)	(0.0625, 0.1215)
	ϕ_{01}	0.2751 (0.0860)	(0.0616, 0.4142)		ϕ_{01}	0.4216 (0.0372)	(0.3511, 0.4939)
	ϕ_{00}	0.2831 (0.0562)	(0.1695, 0.3743)		ϕ_{00}	0.3150 (0.0383)	(0.2442, 0.3933)
BL-I	β_1	0.0716 (0.0117)	(0.0488, 0.0947)	DBGR	α	0.2119 (0.1069)	(0.0655, 0.4796)
	β_2	0.0755 (0.0086)	(0.0606, 0.0913)		λ_1	0.9997 (0.0001)	(0.9994, 0.9999)
	β_3	0.0111 (0.0033)	(0.0057, 0.0174)		λ_2	0.9997 (0.0001)	(0.9994, 0.9999)
	ϕ_{11}	0.2455 (0.0409)	(0.0387, 0.3134)		ϕ_{11}	0.4166 (0.0829)	(0.2752, 0.5866)
	ϕ_{10}	0.0789 (0.0436)	(0.0416, 0.1339)		ϕ_{10}	0.0251 (0.0210)	(0.0009, 0.0750)
	ϕ_{01}	0.3151 (0.0245)	(0.2399, 0.4105)		ϕ_{01}	0.2538 (0.0603)	(0.1139, 0.3637)
	ϕ_{00}	0.3605 (0.0387)	(0.2771, 0.4488)		ϕ_{00}	0.3045 (0.0474)	(0.2095, 0.3944)

In Table 15, it is presented the Monte Carlo estimates of DIC for each model. From the results in Table 15, using DIC as the discrimination criteria, it could be concluded that the BL-I model has a good fit among the continuous models and the BDBG model has a good fit among the discrete models to the DRS data set in presence of cure rate

parameters. Overall, the BL-I is the best model to be fitted for DRS data set. However, the DIC criteria in medical studies might not be a good indicator to model discrimination/selection.

Table 15 – DIC values for each model for the DRS data set assuming the presence of cure rate parameters.

Continuous Models		Discrete Models	
Model	DIC value	Model	DIC value
GBE	1666.162	ABG	1657.612
BBBE	1660.827	BDBG	1649.907
MOBE	1674.111	BG-Type II	1671.265
BL-I	1595.778	DBGR	1685.621

In this way, it also considered the marginal Kaplan-Meier and survival probability plots as a discrimination criteria in contrast to DIC criteria. In this way, looking at the Kaplan-Meier plots (see, Figures 27 and 28) for both times T_1 and T_2, it is observed that the assumed models captured the cure rate with good accuracy; the estimated credibility bonds incorporate the nonparametric Kaplan-Meier curve and the results are quite similar meaning that these models could be an alternative to describe and predict the time to blindness for the treated or untreated eye for the patients diagnosed with diabetic retinophaty considered in this study. In addition, in Figures 29 and 30, it is illustrated the probability plots of the Kaplan-Meier estimates for the survival functions versus the respective predicted values from the proposed models for both times T_1 and T_2 under a Bayesian approach. Again, it is observed that the assumed models have a good accuracy to predict the survival probability of the time to blindness for the treated and untreated eyes. In conclusion, although the models are quite similar in the results, the discrete models have better fits for the DRS once they have more parsimony in terms of computational aspects (no terms with exponential function).

Finally, the estimated proportions of non-susceptible patients for the time to blindness for the treated eye (T_1) and for the time to blindness for the untreated eye (T_2), are given in Table 16. Under a nonparametric approach, the estimated proportions of non-susceptible individuals for the time T_1 and T_2, are, respectively, 67.1% and 39.4% considering the Kaplan-Meier nonparametric estimator. Therefore, from the results of Table 16, it could be concluded that the BL-I model has better fit for cure rate under time T_1 and the BG-Type II model has better fit for cure rate under time T_2.

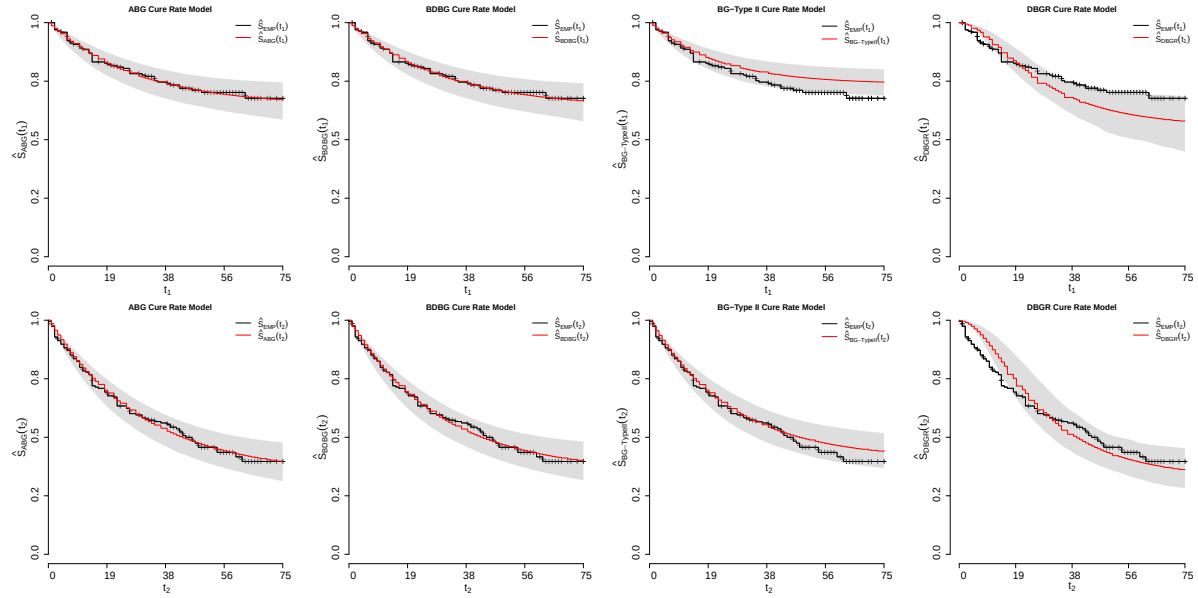

Figure 27 – Kaplan-Meier estimators versus Bayesian estimated survival functions for the marginal survival functions assuming discrete models for T_1 (upper panels) and T_2 (lower panels) in presence of cure rate for the DRS data set.

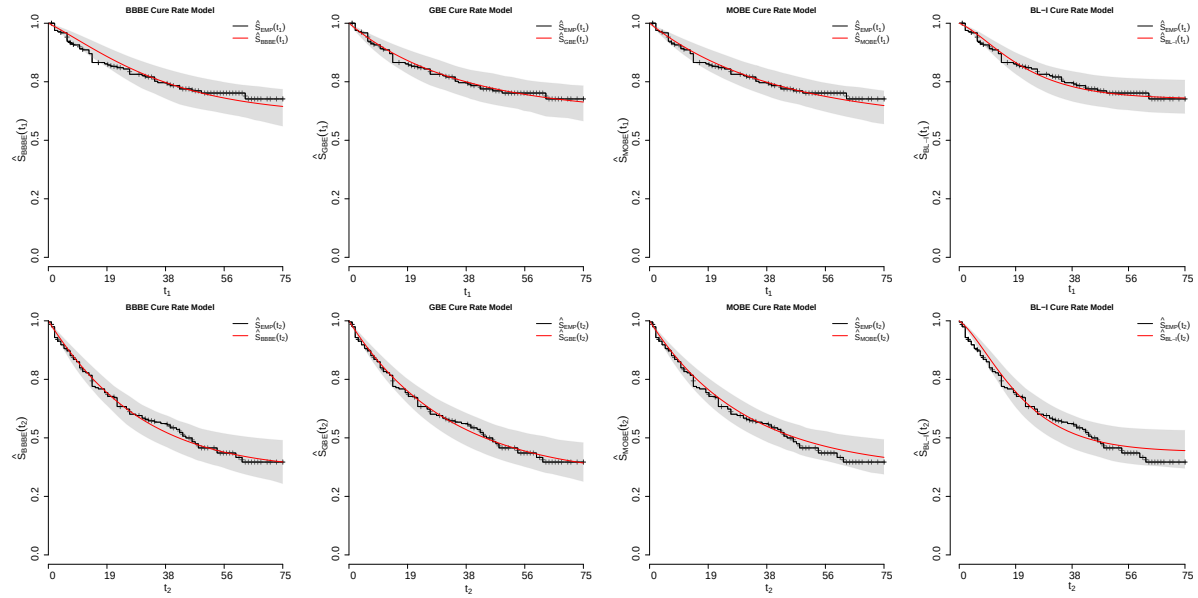

Figure 28 – Kaplan-Meier estimators versus Bayesian estimated survival functions for the marginal survival functions assuming continuous models for T_1 (left panel) and T_2 (right panel) in presence of cure rate for the DRS data set.

Table 16 – The estimated proportions of non-susceptible patients for the DRS data set.

Continuous Models			Discrete Models		
Model	T_1 $(\phi_{00}+\phi_{01})$	T_2 $(\phi_{00}+\phi_{10})$	Model	T_1 $(\phi_{00}+\phi_{01})$	T_2 $(\phi_{00}+\phi_{10})$
GBE	59.97%	32.34%	AGB	63.60%	33.48%
BBBE	60.58%	35.67%	BDBG	62.46%	34.56%
MOBE	55.82%	35.53%	BG-Type II	73.66%	40.43%
BL-I	67.56%	43.94%	DBGR	55.83%	32.96%

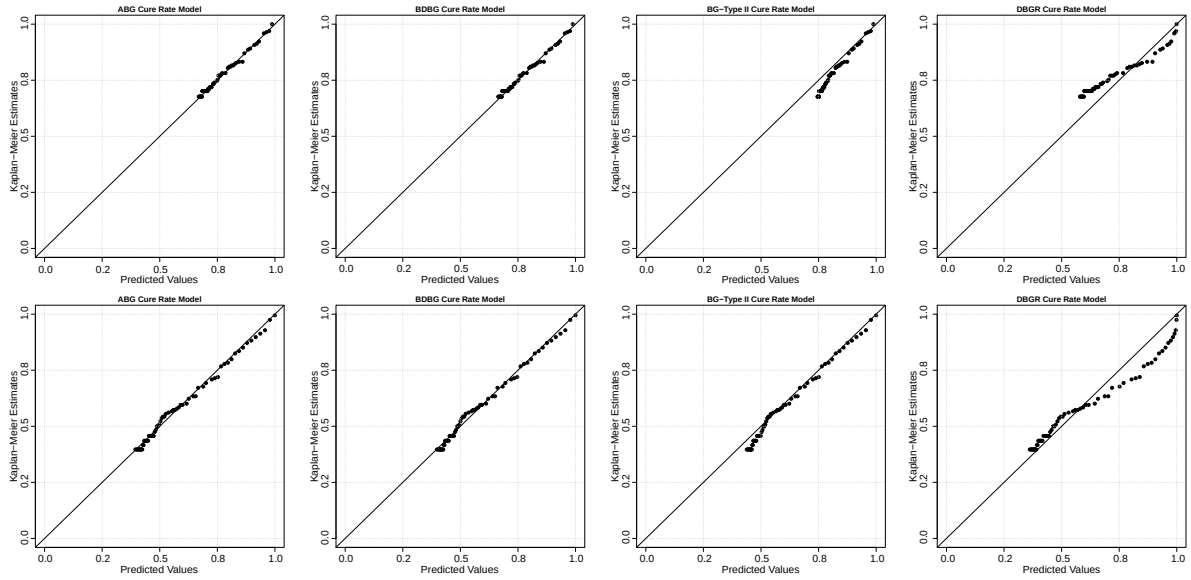

Figure 29 – Plots of the Kaplan-Meier estimates for the survival function versus the respective predicted values obtained from the proposed discrete models for T_1 (upper panels) and T_2 (lower panels) for the DRS data set.

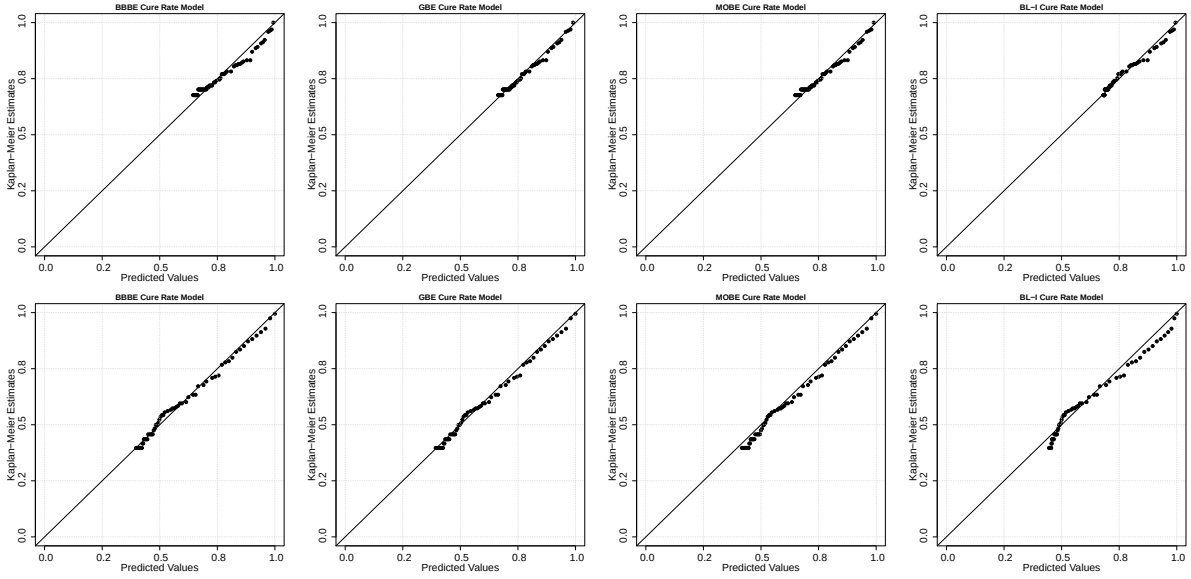

Figure 30 – Plots of the Kaplan-Meier estimates for the survival function versus the respective predicted values obtained from the proposed continuous models for T_1 (upper panels) and T_2 (lower panels) for the DRS data set.

4.4.2.2 Statistical Analysis in the Presence of Risk Factors

Suppose now that the medical interest is investigate possible prognostic/diagnose/influence factors for diabetic retinopathy disease. In our dataset considered in this study, we have four covariates of interest: *type of laser, age at diagnosis of diabetes, type of diabetes* and *risk score*. However, this study is limited to investigate only one of them, *type of diabetes*, which is the majority issue for diabetic retinophaty disease.

Let X_1 denoting the covariate type of diabetes categorized by an indicator variable where the indicator 0 is for juvenile (age at diagnosis < 20) and the indicator 1 is for adults. In this case, it is assumed two regression models for the parameters of the proposed models under a Bayesian approach: the logistic regression model (discrete case) and the linear regression model (continuous case). The assumed logistic regression model in the discrete case is given by,

$$\text{Logit Model: } \begin{cases} \text{logit}(\gamma_{1i}) = \omega_{10} + \omega_{11} X_{1i} \\ \text{logit}(\gamma_{2i}) = \omega_{20} + \omega_{21} X_{1i} \end{cases} \tag{4.14}$$

where $\text{logit}(\alpha) = \log\left(\dfrac{\alpha}{1-\alpha}\right)$ and X_1 refers to age status (adult or juvenile); the assumed linear regression model in the continuous case is given by:

$$\text{Linear Model: } \begin{cases} \eta_{1i} = \gamma_1 \exp(\omega_{10} + \omega_{11} X_{1i}) \\ \eta_{2i} = \gamma_2 \exp(\omega_{20} + \omega_{21} X_{1i}) \end{cases} \tag{4.15}$$

For both regression models, let us assume the same prior distributions considered in the analysis not considering the presence of covariates for the model and incidence parameters. For both regression models given in (4.14) and (4.15), let us assume approximately normal non-informative prior distributions $N(0, 100)$ for the regression parameters $\omega_{j0}, j = 1, 2$ and $\omega_{j1}, j = 1, 2$; and Gamma$(0.001, 0.001)$ prior distributions for the parameters $\gamma_j, j = 1, 2$ in the continuous case.

The posterior summaries of interest are presented in Table 17. From the results of Table 17, it is observed that zero is included in the 95% credible intervals for ω_{21} for all the models, that is, the lifetime T_2 is not affected by the covariate X_1 denoting the type of diabetes. In this way, we conclude that only T_1 is affected (the credible intervals do not include zero, except for BBBE and DBGR models) by the covariate denoting the type of diabetes where 0 = juvenile (age at diagnosis < 20) and 1 = for adults.

Moreover, in Table 18, it is presented the Monte Carlo estimates of DIC for each model under the regression approach. From the results in Table 18 it is possible to conclude that the DIC values in presence of covariates are very similar to the DIC values considering only the cure rate, that is, the models not considering the presence of *type of diabetes* are better fitted to the DRS data set based on DIC criteria and parsimony. However, in the regression approach there is more information for the researcher, especially in medical studies where the main interest is the prognostic factors as well the cure rate. In Figures 31 and 32, it is also illustrated the Kaplan-Meier plots in presence of covariate *type of diabetes* from where it is concluded that that the proposed models has a quite good accuracy (except for BBBE and DBGR models) to predict the time to blindness for each type of diabetes.

Table 17 – Posterior summaries for the proposed models considering the type of diabetes for the DRS data set.

Continuous Models				Discrete Models			
Model	Par.	Post. Mean (S.D.)	95% Cred. Interval	Model	Par.	Post. Mean (S.D.)	95% Cred. Interval
GBE	γ_1	1.4276 (2.5039)	(0.0024, 9.7564)	ABG	ω_{10}	-3.1807 (0.3170)	(-3.9122, -2.6603)
	γ_2	0.2661 (0.5125)	(0.0001, 1.7654)		ω_{11}	-0.8803 (0.4078)	(-1.6901, -0.0864)
	ω_{10}	-1.9890 (2.3400)	(-5.6767, 2.6334)		ω_{20}	-3.6196 (0.2323)	(-4.0568, -3.1627)
	ω_{11}	-0.8151 (0.3385)	(-1.4625, -0.1998)		ω_{21}	0.3500 (0.2582)	(-0.1641, 0.8378)
	ω_{20}	1.1710 (3.7345)	(-4.2742, 8.0235)		ϕ_{11}	0.3642 (0.0699)	(0.2513, 0.5269)
	ω_{21}	0.3499 (0.2475)	(-0.1966, 0.7686)		ϕ_{10}	0.0407 (0.0286)	(0.0020, 0.1058)
	θ	0.0004 (0.0020)	(0.0001, 0.0043)		ϕ_{01}	0.3218 (0.0765)	(0.1588, 0.4649)
	ϕ_{11}	0.3943 (0.0796)	(0.2722, 0.5925)		ϕ_{00}	0.2732 (0.0514)	(0.1671, 0.3680)
	ϕ_{10}	0.0384 (0.0304)	(0.0013, 0.1099)				
	ϕ_{01}	0.3128 (0.0877)	(0.1059, 0.4501)				
	ϕ_{00}	0.2544 (0.0505)	(0.1480, 0.3428)				
BBBE	γ_1	0.0779 (0.2004)	(0.0001, 0.7416)	BDBG	ω_{10}	3.3840 (0.3508)	(2.8140, 4.2069)
	γ_2	0.4321 (1.1860)	(0.0001, 5.1015)		ω_{11}	1.2202 (0.5361)	(0.2417, 2.3442)
	ω_{10}	-5.2060 (3.4029)	(-12.4762, 0.3028)		ω_{20}	3.7675 (0.2855)	(3.2224, 4.3353)
	ω_{11}	-1.3148 (4.2635)	(-9.3880, 6.3233)		ω_{21}	-0.3623 (0.3066)	(-0.9727, 0.2290)
	ω_{20}	-6.6493 (4.5510)	(-15.3456, 0.7470)		θ_{12}	0.9945 (0.0025)	(0.9888, 0.9983)
	ω_{21}	-0.0902 (4.2219)	(-8.1723, 7.2838)		ϕ_{11}	0.3643 (0.0727)	(0.2500, 0.5381)
	λ_3	0.0453 (0.0083)	(0.0292, 0.0607)		ϕ_{10}	0.0541 (0.0335)	(0.0038, 0.1279)
	ϕ_{11}	0.3447 (0.0765)	(0.2472, 0.5555)		ϕ_{01}	0.3124 (0.0751)	(0.1472, 0.4550)
	ϕ_{10}	0.0733 (0.0267)	(0.0266, 0.1301)		ϕ_{00}	0.2692 (0.0531)	(0.1582, 0.3672)
	ϕ_{01}	0.3282 (0.0538)	(0.2198, 0.4262)				
	ϕ_{00}	0.2539 (0.0603)	(0.1130, 0.3654)				
MOBE	γ_1	0.0211 (0.0770)	(0.0001, 0.2859)	BG-Type II	ω_{10}	3.2945 (0.3515)	(2.7289, 4.1324)
	γ_2	0.1942 (0.2149)	(0.0001, 0.7299)		ω_{11}	0.8763 (0.3837)	(0.1250, 1.6324)
	ω_{10}	5.4122 (4.2759)	(-2.3573, 13.1010)		ω_{20}	3.6678 (0.2465)	(3.1656, 4.1410)
	ω_{11}	-1.0697 (0.5524)	(-2.3453, -0.1603)		ω_{21}	-0.3601 (0.2577)	(-0.8537, 0.1575)
	ω_{20}	-0.5578 (2.7158)	(-3.5710, 5.3107)		θ_{12}	0.9981 (0.0016)	(0.9949, 1.0000)
	ω_{21}	0.3772 (0.3038)	(-0.1948, 0.8964)		ϕ_{11}	0.3626 (0.0752)	(0.2461, 0.5552)
	λ_{12}	0.0054 (0.0025)	(0.0018, 0.0110)		ϕ_{10}	0.0500 (0.0335)	(0.0028, 0.1250)
	ϕ_{11}	0.3868 (0.0868)	(0.2591, 0.6188)		ϕ_{01}	0.3280 (0.0809)	(0.1456, 0.4739)
	ϕ_{10}	0.0605 (0.0380)	(0.0045, 0.1501)		ϕ_{00}	0.2594 (0.0551)	(0.1424, 0.3576)
	ϕ_{01}	0.2754 (0.0850)	(0.0658, 0.4139)				
	ϕ_{00}	0.2773 (0.0531)	(0.1697, 0.3770)				
BL-I	γ_1	1.2603 (3.0638)	(0.0001, 11.2275)	DBGR	ω_{10}	8.5882 (0.5270)	(7.5524, 9.7509)
	γ_2	0.0562 (0.0859)	(0.0001, 0.2975)		ω_{11}	0.2678 (0.3873)	(-0.5152, 0.9921)
	ω_{10}	7.1414 (9.4409)	(-4.9411, 28.4038)		ω_{20}	8.7585 (0.5100)	(7.8421, 9.8073)
	ω_{11}	-0.8906 (0.4113)	(-1.7344, -0.1129)		ω_{21}	-0.1176 (0.4214)	(-0.9561, 0.7791)
	ω_{20}	1.8732 (2.3053)	(-1.5180, 6.7436)		α	0.1489 (0.0677)	(0.0403, 0.2941)
	ω_{21}	0.1920 (0.2317)	(-0.2512, 0.6658)		ϕ_{11}	0.4736 (0.0964)	(0.3238, 0.7214)
	ω_3	0.0104 (0.0034)	(0.0049, 0.0179)		ϕ_{10}	0.0179 (0.0168)	(0.0007, 0.0626)
	ϕ_{11}	0.2721 (0.0416)	(0.1957, 0.3569)		ϕ_{01}	0.2358 (0.0729)	(0.0613, 0.3543)
	ϕ_{10}	0.0831 (0.0266)	(0.0372, 0.1388)		ϕ_{00}	0.2726 (0.0506)	(0.1617, 0.3637)
	ϕ_{01}	0.2932 (0.0444)	(0.2086, 0.3796)				
	ϕ_{00}	0.3516 (0.0412)	(0.2719, 0.4311)				

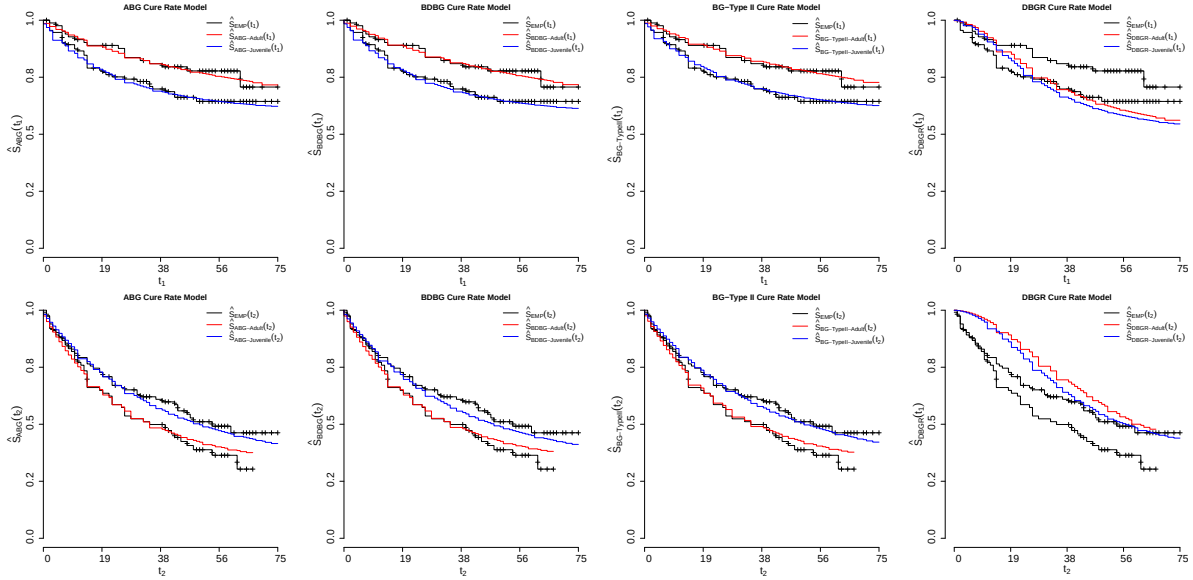

Figure 31 – Kaplan-Meier estimators versus Bayesian estimated survival functions for the marginal survival functions assuming discrete models for T_1 (upper panels) and T_2 (lower panels) assuming the type of diabetes for the DRS data set.

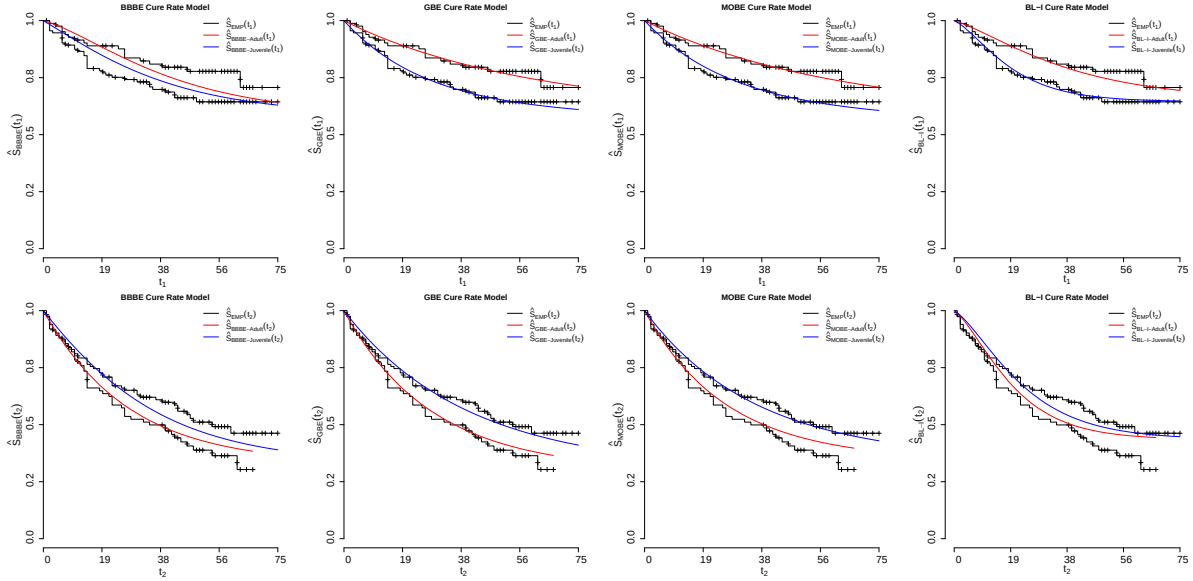

Figure 32 – Kaplan-Meier estimators versus Bayesian estimated survival functions for the marginal survival functions assuming continuous models for T_1 (upper panels) and T_2 (lower panels) assuming the type of diabetes for the DRS data set.

Table 18 – DIC values for each model for the DRS data set under regression approach.

Continuous Models		Discrete Models	
Model	DIC value	Model	DIC value
GBE	1659.695	ABG	1654.468
BBBE	1662.653	BDBG	1646.271
MOBE	1669.756	BG-Type II	1666.157
BL-I	1593.050	DBGR	1675.782

4.4.3 Discussion and Remarks

Our analysis of DRS dataset illustrated a new way to predict and identify some important in the risk factors for diabetic retinopathy disease. Based on the analysis, the findings of the present study demonstrated, under a nonparametric approach, that 67.1% of the patients which the eye was treated are non-susceptible to blindness and 39.4% of the patients for untreated eye. In contrast, under a parametric approach, the BL-I model demonstrated that 67.56% of the patients which the eye was treated are non-susceptible to blindness and the BG-Type II model demonstrated that 40.43% of the patients are non-susceptible to blindness for untreated eye. That is, the proposed models has a good accuracy for the estimation of the long-term survivors. It is important to point out that the proposed continuous models demonstrated similar accuracy for the estimation of the long-term survivors compared to discrete models. However, these models also demonstrated some computational difficulties to get inferences for the parameters of interest. This fact may be related to the use of non-informative prior distributions for the MCMC simulation algorithm. In this case, the results could be improved using very informative prior distributions as well more iterations to get convergence of the MCMC simulation algorithm.

On other hand, related to the risk factors and according to the regression results, only the patients with treated eye had the time to blindness affected by the type of diabetes assuming the proposed models where it could be seen that the probability of the patients which the eye was treated that are non-susceptible to blindness is higher for patients that were prognosticated with diabetes at adult phase as compared to the patients that were prognosticated with diabetes at juvenile phase (age at diagnosis $<$ 20). As a future work, we should investigate another factors as for example the age at diagnosis that could be an important factor clinically.

In conclusion, the results emerging from this study reinforce the fact that the search of appropriate bivariate lifetime distributions could be extremely difficult (see BBBE model, for example) depending on the correlation structure of the lifetime data. However, the proposed methodology could be very useful in the medical data analysis where the interest is the estimation of the fraction of patients in the studied population who never experience the event of interest. The results could be also extended to other cross-over trials in clinical research; reliability analysis in engineering; risk analysis in economics; among many others areas. Finally, the general framework for the computer codes of the proposed modeling approach is presented in Appendix A at the end of this thesis which could be carried out using the OpenBugs software (Spiegelhalter et al., 2007) or R2jags (Su and Yajima, 2015) library from R software.

Chapter 5

Extension to Other Fields of Study

5.1 Reliability Analysis

5.1.1 Introduction

A series system is a component configuration usually assumed in engineering studies, such that, if any one of the system components fails, the entire system fails. Associated to each system component there is a response given by a random variable that could be binary (fail/no fail) or denoting its lifetime (a positive value). In this way, the estimation of the reliability is obtained using an inference approach based on probabilistic models given by the probability of failure (response: fail/no fail) or by the probability $P(T > t)$, where T denotes the lifetime of the component or system and t is a fixed value (see, for example, Jensen and Bard, 2003).

Studies of series system reliability typically assume that the lifetimes of each component are independent. This assumption, in general, is not reasonable in many practical engineering situations, since it is possible that two components assembled into a system structure may share the same load, may be subject to the same set of stresses or under the same environment, which could lead to similar performances. In this way, when estimating the reliability of 2-component systems, it is important to consider statistical models that support the presence of some dependence structure between the lifetimes of the components, since it may affect the evaluation of the full reliability of the system. Different bivariate lifetime models could be assumed for correlated lifetimes of a series system (see, for example, Gumbel, 1960; Freund, 1961; Marshall and Olkin, 1967a,b; Downton, 1970; Hawkes, 1972; Block and Basu, 1974; Sarkar, 1987; Arnold and Strauss, 1988).

The system's reliability is usually derived as the product of the reliabilities for a fixed time $t = \min(t_i)$ $(i = 1, \ldots, n)$. Under the independence assumption (product of the reliabilities), considering as a special case a two-component series systems, it is assumed two random variables T_1 and T_2 belonging to the same distributional family, but indexed

by different parameters, typically in a univariate form representing the lifetime series system, $t = \min(t_1, t_2)$. Any lifetime distribution could be assumed for each component of the series system. For example, T_1 and T_2 could be assumed as following an Exponential or a Weibull distribution with different parameters. A mixture of these distributions (T_1 follows a Weibull distribution and T_2 follows an Exponential) or random field discretization can also be considered (see Burr, 1968, 1973; Singh and Billinton, 1977; Blanchard et al., 1990; Chao and Fu, 1991; Mori and Ellingwood, 1993; Hulting and Robinson, 1994; Zhang and Horigome, 2001; Rausand and Arnljot, 2004; Kołowrocki, 2008; Eryilmaz and Tank, 2012; Hu and Mahadevan, 2015; Oliveira and Achcar, 2019). However, it is possible that there is a deterioration process destroying both components at the same time. In this case, the dependence assumption is essential requiring the use of a bivariate distribution since the dependence is related to the deterioration process.

The novelty of the present study is the introduction of a time-dependent structure among the deterioration process using BL distributions derived from two Lindley distributions for a 2-component series system since it could provide better accuracy for the estimation of the full system reliability instead of assuming independent univariate lifetime distributions for each component. Moreover, since it is a great challenge obtaining the system's reliability function in terms of component reliabilities, the use of a probability distribution, especially a multivariate one, can be useful to get warranty periods and the system failure rate.

5.1.2 Inference Methods

Usually, in studies related to reliability of 2-component series systems, it is observed only the times of failure of the system and the indication of which component had failed. In other words, one has the information on the complete lifetime for one component, but for the other one, there is only partial information of its lifetime, since it possibly continues to work. In this case, it is observed only the minimum lifetime between T_1 and T_2, that is, it is observed $T = \min(T_1, T_2)$. Let us define an indicator variable:

$$\delta = \begin{cases} 1 & \text{if} \quad T_1 < T_2 \\ 0 & \text{if} \quad T_1 \geq T_2. \end{cases} \tag{5.1}$$

One can notice that, if $\delta_i = 1$, the individual contribution for the likelihood function assuming bivariate lifetimes is given by $P(T_{1i} = t_{1i}, T_{2i} > t_{2i}) = -\partial S(t_{1i}, t_{2i})/\partial t_{1i}$ and, if $\delta_i = 0$, the contribution is $P(T_{1i} > t_{1i}, T_{2i} = t_{2i}) = -\partial S(t_{1i}, t_{2i})/\partial t_{2i}$. For the proposed models, assuming a random sample of size n of a 2-component series system with lifetimes T_1 and T_2, the likelihood function (see, e.g. Lawless, 1982) is given by

$$\mathcal{L}(\boldsymbol{\beta}) = \prod_{i=1}^{n} \left[-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} \right]^{\delta_i} \left[-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} \right]^{1-\delta_i}. \tag{5.2}$$

5.1.3 Data Applications

In this section, some applications associated to reliability and lifetime data are considered for illustrative purposes of the usefulness of the proposed BL models. In the first application, it is considered 50 simulated simple computer series systems consisting of a processor and a memory. The computer works if both system components are working correctly. Suppose that there is a latent deterioration process occurring in the system. During a relatively short time (in hours), the deterioration progresses rapidly and makes the system susceptible to shocks in which a random fatal shock can destroy the first component, the second or both components. Since the fatal shock can simultaneously destroy both components, the independence assumption could be not appropriate.

The second application is related to engine winding reliability (the dataset was introduced in a data file of the Minitab software (Minitab, 1991) where a reliability engineer studied the failure rates of engine windings of turbine assemblies to determine the times at which the windings fail. At high temperatures, the windings could decompose too fast.

5.1.3.1 Computer Series System Simulated Data (CS)

Let us consider an industrial design previously described which relates to the simple computer series systems consisting of a processor and a memory, that is, two-component series systems. Since the system has a deteriorating process, suppose that a shock from source one destroys the processor (component one) occurring at a random time W_1 such that $W_1 \sim Lindley(0.5)$; a shock from source two destroys the memory (component two) occurring at a random time W_2 such that $W_2 \sim Lindley(0.5)$; and a shock from source three destroys simultaneously the processor and the memory at a random time W_3 such that $W_3 \sim Lindley(0.5)$.

In this industrial design, the lifetime of the processor (in hours) is given by $T_1 = \min(W_1, W_3)$ and the lifetime of the memory (in hours) is given by $T_2 = \min(W_2, W_3)$. However, our interest is in the lifetime of the entire system. In this case, the lifetime of the entire system is given by $T = \min(T_1, T_2)$ and the reliability of the system for the time $t = \min(t_1, t_2)$ is given by $R(t) = S(t, t)$ due to the common source of shock. For a statistical analysis under a Bayesian approach, it is assumed the proposed models BL models. As *prior* distributions, it is assumed approximately noninformative Gamma$(0.01, 0.01)$ *prior* distributions for the parameters β_j $(j = 1, 2)$ and an approximately informative Gamma *prior* distribution with hyperparameters values $n\overline{T_{\min}}$ and $n\text{Var}(T_{\min})$ (where $\overline{T_{\min}}$ and $\text{Var}(T_{\min})$ are the sample mean and the sample variance of $T_{\min} = \min(T_1, T_2)$, respectively; and n is the sample size) for β_3 to avoid computational instability. The *posterior* summaries of interest for the assumed models are presented in Table 20.

Table 19 – Fifty simulated computer series system of two components.

Sys	Processor Lifetime	Memory Lifetime	Sys	Processor Lifetime	Memory Lifetime	Sys	Processor Lifetime	Memory Lifetime	Sys	Processor Lifetime	Memory Lifetime	Sys	Processor Lifetime	Memory Lifetime
1	1.9292	3.9291	11	1.9386	4.0043	21	1.1739	3.3857	31	0.1181	0.0884	41	0.6270	1.7289
2	3.6621	0.0026	12	2.1000	2.0513	22	1.3482	1.9705	32	5.0533	2.3238	42	0.7947	0.7947
3	3.9608	0.8323	13	0.9867	0.9867	23	3.0935	3.0935	33	1.6465	2.0197	43	0.5079	5.3535
4	2.3504	3.3364	14	0.1837	0.1837	24	2.1396	2.1548	34	0.9096	0.6214	44	2.5913	2.5913
5	1.0833	3.3059	15	1.3989	4.1268	25	1.3288	0.9689	35	1.7494	2.3643	45	2.5372	2.4923
6	2.8414	1.8438	16	2.3757	2.7953	26	0.1115	0.1115	36	0.1058	0.1058	46	1.1917	0.0801
7	0.3309	0.3309	17	3.5202	1.4095	27	0.8503	2.8578	37	0.4593	0.4593	47	1.5254	4.4088
8	2.9884	1.5961	18	2.3364	0.1624	28	0.1955	0.1955	38	0.9938	1.7689	48	1.0986	1.0986
9	0.5784	1.8795	19	0.8584	1.9556	29	0.4614	0.8584	39	5.7561	0.3212	49	1.0051	1.0051
10	0.5520	0.5520	20	4.3435	1.0001	30	3.3887	1.9796	40	6.2950	1.0495	50	1.3640	1.3640

Table 20 – *Posterior* summaries assuming BL models for the CS data.

Shock bivariate Lindley models						Stress bivariate Lindley models					
Model	Par.	Post. Mean	Std. Dev.	95% Cred. Int.	DIC	Model	Par.	Post. Mean	Std. Dev.	95% Cred. Int.	DIC
BL-I	β_1	0.4847	0.0777	0.3545, 0.6428)	336.90	BL-III	β_1	0.0027	0.0096	(0.0001, 0.0391)	387.80
	β_2	0.4723	0.0784	(0.3227, 0.6578)			β_2	0.0046	0.0172	(0.0001, 0.0694)	
	β_3	0.4182	0.0525	(0.3355, 0.5376)			β_3	0.8982	0.0585	(0.7783, 0.9954)	
BL-II	β_1	0.4758	0.0733	(0.3415, 0.6268)	294.60	BL-IV	β_1	0.0012	0.0088	(0.0001, 0.0108)	316.20
	β_2	0.4737	0.0792	(0.3310, 0.6416)			β_2	0.0092	0.0255	(0.0001, 0.0937)	
	β_3	0.4978	0.0610	(0.3869, 0.6298)			β_3	1.0149	0.0736	(0.8741, 1.1598)	

From the results of Table 20, using DIC as discrimination criterion, it is observed that the BL-II model is better fitted to the data which is expected since the generated data follows the construction structure of the model BL-II. However, all models have closer DIC values which lead us to the conclusion that the use of BL distributions could be a good alternative to analyse reliability system data since it was obtained accurate inference results (small values for standard deviations and good length for the 95% credible intervals). However, the main goal of the study is to estimate the reliability for each considered series systems. For instance, using the estimates of the parameters given in Table 20, it is presented in Table 21 the Bayesian estimates of the reliabilities as well the 95% credibility intervals for the reliability of the of the entire system for the first ten systems given in Table 19. For comparative purposes, the product of two Lindley distributions is also considered (independence assumption).

From the results presented in Table 21, it could be concluded that the BL-II model is the best model to predict the reliability of the entire system with good accuracy. This result is expected due to the system structure and it is important to point out that the dependence assumption, in this case, is crucial for the data analysis. In fact, under independence assumption, the reliability of many systems has a poor prediction and the credible intervals do not contain the true value of the system reliability; on the other hand,

all proposed models can capture the true reliability of the entire system by the credibility intervals range. In many practical situations which a common source of shock for both components, the entire series system reliability could be predicted using BL distributions which are the novelty of this study.

Table 21 – The estimated reliability and the 95%credibility intervals of the first ten simulated computer series system, assuming BL models and independent Lindley models; and the true reliabilities of the systems for the CS data given in Table 19.

Sys	R_{BL-I}	R_{BL-II}	R_{BL-III}	R_{BL-IV}	R_L	R_T
1	0.1879 (0.1274, 0.2487)	0.2663 (0.1810, 0.3622)	0.1777 (0.1464, 0.2228)	0.2795 (0.2167, 0.3511)	0.3476 (0.2527, 0.4546)	0.2456
2	0.9981 (0.9977, 0.9984)	0.9988 (0.9984, 0.9991)	0.9977 (0.9974, 0.9980)	0.9987 (0.9984, 0.9989)	0.9990 (0.9987, 0.9993)	0.9987
3	0.5128 (0.4403, 0.5769)	0.6144 (0.5312, 0.6936)	0.4739 (0.4365, 0.5232)	0.6098 (0.5505, 0.6702)	0.6747 (0.5970, 0.7505)	0.5982
4	0.1247 (0.0766, 0.1758)	0.1858 (0.1141, 0.2722)	0.1220 (0.0962, 0.1605)	0.2025 (0.1476, 0.2680)	0.2620 (0.1752, 0.3652)	0.1670
5	0.4116 (0.3360, 0.4804)	0.5155 (0.4235, 0.6064)	0.3785 (0.3400, 0.4303)	0.5151 (0.4496, 0.5834)	0.5865 (0.4974, 0.6761)	0.4966
6	0.2039 (0.1410, 0.2665)	0.2858 (0.1983, 0.3830)	0.1918 (0.1594, 0.2381)	0.2980 (0.2339, 0.3705)	0.3675 (0.2716, 0.4745)	0.2649
7	0.7769 (0.7344, 0.8133)	0.8408 (0.7977, 0.8792)	0.7429 (0.7193, 0.7729)	0.8336 (0.8022, 0.8642)	0.8685 (0.8305, 0.9033)	0.8332
8	0.2577 (0.1884, 0.3246)	0.3492 (0.2565, 0.4483)	0.2392 (0.2040, 0.2887)	0.3579 (0.2910, 0.4316)	0.4306 (0.3334, 0.5355)	0.3281
9	0.6357 (0.5737, 0.6894)	0.7251 (0.6582, 0.7866)	0.5950 (0.5621, 0.6375)	0.7178 (0.6698, 0.7655)	0.7707 (0.7101, 0.8278)	0.7127
10	0.6497 (0.5894, 0.7020)	0.7372 (0.6724, 0.7964)	0.6092 (0.5771, 0.6507)	0.7297 (0.6832, 0.7758)	0.7809 (0.7226, 0.8358)	0.7252

R_L: product of two independent Lindley distribution; R_T: True reliability of the entire system.

5.1.3.2 Engine Winding Reliability Dataset (EW)

The proposed BL models can also be useful for other reliability applications not related to series systems. In this subsection, for example, it is considered an application related to engine winding reliability. The main goal of this study is the estimation of the failure time of engine windings of the turbine. The dataset consists of two lifetimes where the lifetime T_1 denotes the time to failure of windings that was exposed to a temperature of 80 degrees Celsius; C_1 indicates whether the unit failed or survived at 80 degrees Celsius: 1 = an actual failure, or 0 = the censored unit was removed from the test before it failed; and T_2 denotes the time to failure of windings that was exposed to a temperature of 100 degrees Celsius; C_2 indicates whether the unit failed or survived at 100 degrees Celsius: 1 = an actual failure, or 0 = the censored unit was removed from the test before it failed.

For a Bayesian analysis, it is assumed the proposed BL models described in Chapter 3 and approximately noninformative Gamma$(0.001, 0.001)$ *prior* distributions were assumed for the parameters $\beta_j\,(j = 1, 2)$ and an approximately noninformative U$(0, 1)$ for β_3 to avoid computational instability. The *posterior* summaries of interest for the assumed models as well for the bivariate Gumbel Exponential Gumbel (1960) and the bivariate Marshall-Olkin Exponential Marshall and Olkin (1967a,b) are presented in Table 22. Table 22 also present the Monte Carlo estimates of DIC. Using DIC discrimination criterion, it is observed that model BL-II model is better fitted to the data assuming the shock model structure and the BL-IV model is better fitted to the data assuming the stress model structure. The obtained results lead us to the conclusion that the use of BL distributions could be a good alternative to analyse bivariate reliability data since it was obtained accurate inference results (small values for standard deviations and good length for the 95% credible intervals).

Table 22 – *Posterior* summaries assuming BL models for the EW dataset.

Shock bivariate Lindley models						Stress bivariate Lindley models					
Model	Par.	Post. Mean	Std. Dev.	95% Cred. Int.	DIC	Model	Par.	Post. Mean	Std. Dev.	95% Cred. Int.	DIC
BL-I	β_1	0.0281	0.0034	(0.0217, 0.0345)	652.80	BL-III	β_1	0.0256	0.0031	(0.0197, 0.0315)	642.10
	β_2	0.0419	0.0047	(0.0335, 0.0496)			β_2	0.0240	0.0066	(0.0117, 0.0384)	
	β_3	0.0009	0.0008	(0.0001, 0.0032)			β_3	0.0174	0.0055	(0.0076, 0.0280)	
BL-II	β_1	0.0281	0.0035	(0.0212, 0.0347)	647.70	BL-IV	β_1	0.0253	0.0037	(0.0189, 0.0332)	628.10
	β_2	0.0424	0.0048	(0.0337, 0.0525)			β_2	0.0192	0.0062	(0.0095, 0.0330)	
	β_3	0.0048	0.0037	(0.0002, 0.0125)			β_3	0.0446	0.0070	(0.0314, 0.0595)	
Gumbel bivariate Exponential model (Gamma(0.01,0.01) *priors*)						Marshall-Olkin bivariate Exponential model (Gamma(0.01,0.01) *priors*)					
GBE	β_1	0.0138	0.0025	(0.0093, 0.0188)	676.20	MOBE	β_1	0.0132	0.0028	(0.0086, 0.0193)	677.30
	β_2	0.0188	0.0032	(0.0128, 0.0254)			β_2	0.0211	0.0036	(0.0150, 0.0285)	
	β_3	0.4374	0.2609	(0.4122, 0.9653)			β_3	0.0011	0.0010	(0.0001, 0.0034)	

5.1.4 Concluding Remarks

From the different model structure approaches considered in this study to get inferences of interest for a two-component series system reliability (see CS data application), it is possible to conclude that the use of the new class of BL distributions introduced in this paper leads to accurate inference results for the parameters of interest assuming the system lifetime data as univariate or bivariate (despite the presence of censored data) even using noninformative *priors* for the parameters of the model under a Bayesian approach. For example, considering the EW dataset and the LC real dataset applications (despite the nature of both applications), the use of noninformative *prior* provided great accurate inference results (small values for the standard deviations) for each BL model.

On the other hand, in the CS application, it is important to point out that the dependence structure is a crucial assumption as well the use of an informative *prior* distribution with hyperparameters based on sample mean and variance to get good inference

accuracy for the $R(t)$ of the entire system which is a great indicator that the BL models can be applied in practical industrial situations where the series systems fails when a common source of shock destroy both components. Moreover, based on the $R(t)$ estimates, it is possible to conclude that the use of any one of the proposed BL models may be a good option instead using the product of two Lindley distributions considered under the independence assumption. In addition, other probability distributions could be assumed for the random variables W_3 and U_3 related to common shock and stress considering the new class of BL models proposed in this thesis, however this issue is outside the scope of this thesis and will be the subject of future research.

5.2 Multivariate Analysis

5.2.1 Introduction

In this section, the defining properties of the multivariate geometric distribution are based on models in which a n-component system fails according to the occurrences of fatal shocks to each one of the components or for all of the components. The first approach related to this idea introduced in the literature was proposed by Marshall and Olkin (1967b) from where the authors introduced a multivariate exponential distribution. The results for the multivariate geometric distribution are analogous to the results obtained using the multivariate exponential distribution introduced in the literature by Arnold (1975).

5.2.2 Model Structure

Suppose that the components of a two-component system fail after receiving an overall fatal shock. Independent Poisson processes $U_1(t,\theta_1), U_2(t,\theta_2), U_{12}(t,\theta_{12})$ govern the occurrence of fatal shocks. Events in the process $U_1(t,\theta_1)$ are fatal shocks transmitted to component 1, events in the process $U_2(t,\theta_2)$ are fatal shocks transmitted to component 2, and events in the process $U_{12}(t,\theta_{12})$ are fatal shocks transmitted equaly and independently to both components. Thus if $X = \min(U_1, U_{12})$ and $Y = \min(U_2, U_{12})$ denote, respectively, the lifes of the first and second components, we have,

- Assume that the probability of transmitting a fatal shock to the component 1 is equal to $1-\theta_1$; observe that the event $X > x$ occurs if and only if there were no fatal shocks until $X = x$, that is, $P(X > x) = \theta_1^x$.
- Assume that the probability of transmitting a fatal shock to the component 2 is equal to $1-\theta_2$; observe that the event $Y > y$ occurs if and only if there were no fatal shocks until $Y = y$, that is, $P(Y > y) = \theta_2^y$.

In this case, the probability of the system is working until an overall failure is given by,

$$\mathrm{P}(X > x, Y > y) = \theta_1^x \theta_2^y \theta_{12}^{\max(x,y)} \tag{5.3}$$

The probability model given by (5.3) is known in the literature as the Basu-Dhar bivariate geometric distribution introduced by Basu and Dhar (1995). Inferences and some computational aspects for this distribution under a Bayesian approach in presence of censoring and covariates are introduced by Achcar et al. (2016a); de Oliveira and Achcar (2018). An implementation of this distribution in R software is given by the library BivGeo introduced by de Oliveira et al. (2018).

Now, considering a generalization for more than two components series systems, it is considered first an extension of the fatal shock model to a three-component system (Oliveira et al., 2019). Let the independent Poisson processes $U_1(t, \theta_1)$, $U_2(t, \theta_2)$, $U_3(t, \theta_3)$ govern the occurrence of fatal shocks to components 1, 2, 3, respectively; $U_{12}(t, \theta_{12})$, $U_{13}(t, \theta_{13})$, $U_{23}(t, \theta_{23})$ govern the occurrence of fatal shocks to the component pairs 1 and 2, 1 and 3, 2 and 3, respectively; and $U_{123}(t, \theta_{123})$ governs the occurrence of overall fatal shock to components 1, 2, 3. If $X = \max(U_1, U_{12}, U_{13}, U_{123})$, $Y = \max(U_2, U_{12}, U_{23}, U_{123})$ and $Z = \max(U_3, U_{13}, U_{23}, U_{123})$ denote, respectively, the life lengths of the first, second, and third components, thus the probability of the system is working until an overall failure is given by,

$$\begin{aligned} \mathrm{P}(X > x, Y > y, Z > z) &= \mathrm{P}\{U_1(x) = 0, U_2(y) = 0, U_3(z) = 0, U_{12}(\max(x, y)) = 0, \\ &\quad U_{13}(\max(x, z)) = 0, U_{23}(\max(y, z)) = 0, \\ &\quad U_{123}(\max(x, y, z)) = 0\} \\ &= \theta_1^x \theta_2^y \theta_3^z \theta_{12}^{\max(x,y)} \theta_{13}^{\max(x,z)} \theta_{23}^{\max(y,z)} \theta_{123}^{\max(x,y,z)}. \end{aligned} \tag{5.4}$$

Similar arguments yield the n-dimensional geometric distribution given by,

$$\mathrm{P}(X_1 > x_1, \ldots, X_n > x_n) = \prod_{i=1}^{n} \theta_i^{x_i} \cdot \prod_{i=1<j}^{n} \theta_{ij}^{\max(x_i, x_j)} \cdot \prod_{i=1<j<k}^{n} \theta_{ijk}^{\max(x_i, x_j, x_j)} \cdot \ldots \cdot \theta_{12\ldots n}^{\max(x_1, x_2, \ldots, x_n)}, \tag{5.5}$$

where $0 < \theta_i < 1, i = 1, \ldots, n$ and $0 < \theta_{ij}, \ldots, \theta_{12\ldots n} \leq 1, i = 1, \ldots, n; j = 2, \ldots, n; i < j$.

Proposition 4.2.2.1. *Let S denote the set of vectors $(s_1, \ldots, s_n)$ where each $s_j = 0$ or 1 but $(s_1, \ldots, s_n) \neq (0, \ldots, 0)$. For any vector $s \in S$, $\max(x_i, s_i)$ is the minimum of the $x_i's$ for which $s_i = 1$. Thus,*

$$P(X_1 > x_1, \ldots, X_n > x_n) = \prod_{s \in S}^{n} \theta_s^{\max(x_i, s_i)}. \tag{5.6}$$

Proof. It is obtained directly from equation (5.5). □

Proposition 4.2.2.2. *The multivariate probability mass function for the random variables $X_i, i = 1, \ldots, n$ is given by,*

$$\begin{aligned}
\mathrm{P}(X_1 = x_1, \ldots, X_n = x_n) &= \prod_{i=1}^{n} \theta_i^{x_i - 1} \prod_{i=1<j}^{n} \theta_{ij}^{\max(x_i - 1, x_j - 1)} \\
&\times \prod_{i=1<j<l}^{n} \theta_{ijl}^{\max(x_i - 1, x_j - 1, x_l - 1)} \ldots \theta_{12\ldots n}^{\max(x_1 - 1, x_2 - 1, \ldots, x_n - 1)} \\
&- \theta_1^{x_1} \prod_{i=2}^{n} \theta_i^{x_i - 1} \prod_{j=2}^{n} \theta_{1j}^{\max(x_1, x_j - 1)} \\
&\times \prod_{i=2<j}^{n} \theta_{ij}^{\max(x_i - 1, x_j - 1)} \ldots \theta_{12\ldots n}^{\max(x_1, x_2 - 1, \ldots, x_n - 1)} \\
&\vdots \\
&+ (-1)^n \prod_{i=1}^{n} \theta_i^{x_i} \prod_{i=1<j}^{n} \theta_{ij}^{\max(x_i, x_j)} \\
&\times \prod_{i=1<j<l}^{n} \theta_{ijl}^{\max(x_i, x_j, x_l)} \ldots \theta_{12\ldots n}^{\max(x_1, x_2, \ldots, x_n)}. \qquad (5.7)
\end{aligned}$$

Proof. From (5.5) and the equation below,

$$\begin{aligned}
\mathrm{P}(X_1 = x_1, \ldots, X_n = x_n) &= \mathrm{P}(X_1 > x_1 - 1, \ldots, X_n > x_n - 1) \\
&- \mathrm{P}(X_1 > x_1, \ldots, X_n > x_n - 1) - \\
&- \ldots \mathrm{P}(X_1 > x_1 - 1, \ldots, X_n > x_n) + \ldots + \\
&+ (-1)^n \mathrm{P}(X_1 > x_1, \ldots, X_n > x_n)
\end{aligned}$$

the proof is completed. □

We define the distribution given by (5.7) as the *multivariate geometric distribution* (abbreviated MVG).

Proposition 4.2.2.3. *The $(n-1)$-dimensional marginals (hence k-dimensional marginals, $k = 1, 2, \ldots, n-1$) are MVG. In particular, the two-dimensional marginal distributions are bivariate geometric distributions, so the one-dimensional marginal distributions are geometric distributions.*

Proof. For the case $k = 1$, that is, one-dimensional marginals, we have for a random variable $X_i, i = 1, \ldots, n$ that the marginal survival function can be expressed as,

$$\mathrm{P}(X_i > x_i) = \theta_i^{x_i} \cdot \prod_{i<j}^{n} \theta_{ij}^{x_i} \cdot \prod_{i<j<k}^{n} \theta_{ijk}^{x_i} \cdot \ldots \cdot \theta_{12\ldots n}^{x_i} \qquad (5.8)$$

where $i = 1, \ldots, n; j = 2, \ldots, n; k = 3, \ldots, n; i < j < k$. From the survival function in (5.8), the marginal distribution of X_i is given by,

$$\mathrm{P}(X_i = x_i) = \left(\theta_i^{x_i} \cdot \prod_{i<j}^{n} \theta_{ij} \cdot \prod_{i<j<k}^{n} \theta_{ijk} \cdot \ldots \cdot \theta_{12\ldots n}\right)^{x_i} \left(1 - \theta_i^{x_i} \cdot \prod_{i<j}^{n} \theta_{ij} \cdot \prod_{i<j<k}^{n} \theta_{ijk} \cdot \ldots \cdot \theta_{12\ldots n}\right) \tag{5.9}$$

which corresponds to a one-dimensional geometric distribution with mean and variance given, respectively, by

$$\mathbb{E}(X_i) = \frac{1}{\left(1 - \theta_i^{x_i} \cdot \prod_{i<j}^{n} \theta_{ij} \cdot \prod_{i<j<k}^{n} \theta_{ijk} \cdot \ldots \cdot \theta_{12\ldots n}\right)}$$

and

$$\mathrm{Var}(X_i) = \frac{\left(\theta_i^{x_i} \cdot \prod_{i<j}^{n} \theta_{ij} \cdot \prod_{i<j<k}^{n} \theta_{ijk} \cdot \ldots \cdot \theta_{12\ldots n}\right)}{\left(1 - \theta_i^{x_i} \cdot \prod_{i<j}^{n} \theta_{ij} \cdot \prod_{i<j<k}^{n} \theta_{ijk} \cdot \ldots \cdot \theta_{12\ldots n}\right)^2}$$

Since similar arguments yield for the k-dimensional marginals, $k = 1, 2, \ldots, n-1$, the proof is completed.

□

Proposition 4.2.2.4. *Assuming the transformation of the random variables $T_i, i = 1, \ldots, n$ given by $T = \min(T_1, \ldots, T_n)$, the cumulative distribution function for T is given by:*

$$P(T \leq t) = 1 - P(T > t) = 1 - (\theta_1\theta_2 \ldots \theta_{12}\theta_{13} \ldots \theta_{12\ldots n-1}\theta_{12\ldots n})^t. \tag{5.10}$$

which implies that $T \sim Geo(\theta_1\theta_2 \ldots \theta_{12}\theta_{13} \ldots \theta_{12\ldots n-1}\theta_{12\ldots n})$ with mean $\mathbb{E}(T) = 1/(1 - \theta_1\theta_2 \ldots \theta_{12} \ldots \theta_{12\ldots n-1}\theta_{12\ldots n})$ and variance $Var(T) = \theta_1\theta_2 \ldots \theta_{12}\theta_{13} \ldots \theta_{12\ldots n-1}\theta_{12\ldots n}/(1 - \theta_1\theta_2 \ldots \theta_{12}\theta_{13} \ldots \theta_{12\ldots n-1}\theta_{12\ldots n})^2$.

Proof. The proof is directly obtained in the same way as the proof of Proposition 3. □

5.2.3 Likelihood Function

Let us assume a random sample of size n of a n-component series system with lifetimes $T_i, i = 1, \ldots, n$ related to the n-components. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{12}, \ldots, \theta_{12\ldots n-1}, \theta_{12\ldots n})$ assuming complete data is

given by,

$$\begin{aligned}
\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\log\left(\prod_{i=1}^{n}\theta_i^{x_i-1}\prod_{i=1<j}^{n}\theta_{ij}^{\max(x_i-1,x_j-1)}\right.\\
&\times \prod_{i=1<j<l}^{n}\theta_{ijl}^{\max(x_i-1,x_j-1,x_l-1)}\ldots\theta_{12\ldots n}^{\max(x_1-1,x_2-1,\ldots,x_n-1)}\\
&- \theta_1^{x_1}\prod_{i=2}^{n}\theta_i^{x_i-1}\prod_{j=2}^{n}\theta_{1j}^{\max(x_1,x_j-1)}\prod_{i=2<j}^{n}\theta_{ij}^{\max(x_i-1,x_j-1)}\ldots\theta_{12\ldots n}^{\max(x_1,x_2-1,\ldots,x_n-1)}\\
&\vdots\\
&+ \left.(-1)^n\prod_{i=1}^{n}\theta_i^{x_i}\prod_{i=1<j}^{n}\theta_{ij}^{\max(x_i,x_j)}\prod_{i=1<j<l}^{n}\theta_{ijl}^{\max(x_i,x_j,x_l)}\ldots\theta_{12\ldots n}^{\max(x_1,x_2,\ldots,x_n)}\right).
\end{aligned} \tag{5.11}$$

Observe that the log-likelihood expressed in (5.11) has no compact form which implies that the MLEs and the observed information Fisher's matrix should be obtained using standard numeric optimization algorithms such the Newton-Raphson or the Nelder-Mead methods. However, in this study, inferences for the system parameters are obtained from Bayesian methods obtained using MCMC (Markov Chain Monte Carlo) methods (see Gelfand and Smith, 1990; Chib and Greenberg, 1995) based on the squared error loss function, $\eta(\beta, a) = (\beta - a)^2$ and assuming independent beta prior distributions for each parameter in the vector parameters $\boldsymbol{\theta}$ since each component of $\boldsymbol{\theta}$ is restricted to the interval $(0, 1)$.

5.2.4 A Numerical Illustration

In this section, without loss of generality, it is presented simulated datasets of 2-components/3-components or 4-components series system as an application for the proposed methodology. The results holds in a similar way for the n-components series system. In this way, it is assumed that n series systems of 2-components/3-components or 4-components were put on the life test and the lifetimes of the systems are observed until the occurrence of failure. In the statistical analysis, it is assumed a Bayesian analysis considering the likelihood function assuming the multivariate lifetimes from a MVG distribution from the 2-components/3-components and 4-components series system. In this case, the dependence assumption are considered. In practical engineering studies, we usually only have the univariate data and the independence assumption.

The simulated datasets were generated from a MVG distribution considering the sample sizes $n = 20, 50, 100, 150, 300$ and parameters values given in Table 23. The generation algorithm for the n-components series system is given by the steps:

- Generate $U_i \sim Geo(1 - \theta_i), i = 1, \ldots, n$.

- Generate $U_{ij} \sim Geo(1-\theta_{ij}), i = 1, \ldots, n; j = 2, \ldots, n$ and $i < j$.
- Generate $U_{ijk} \sim Geo(1-\theta_{ijk}), i = 1, \ldots, n; j = 2, \ldots, n; k = 3, \ldots, n$ and $i < j < k$.

$$\vdots$$

- Generate $U_{12\ldots n} \sim Geo(1-\theta_{12\ldots n})$.
- Define $T_1 = \min(U_1, U_{1j}, U_{1jk} \ldots, U_{12\ldots n})$, $T_2 = \min(U_2, U_{12}, U_{12k}, \ldots, U_{2j}, U_{2jk}, \ldots, U_{12\ldots n}), \ldots, T_n = \min(U_n, U_{1n}, U_{2n}, \ldots, U_{12n}, \ldots, U_{12\ldots n})$.
- Return $\mathbf{T} = (T_1, \ldots, T_n)$.

Table 23 – True parameter values for the considered series system for the simulated datasets.

	Components		
Param.	n = 2	n = 3	n = 4
θ_1	0.90	0.90	0.90
θ_2	0.90	0.90	0.90
θ_3	—	0.90	0.90
θ_4	—	—	0.90
θ_{12}	0.95	0.95	0.95
θ_{13}	—	0.95	0.95
θ_{14}	—	—	0.95
θ_{23}	—	0.95	0.95
θ_{24}	—	—	0.95
θ_{34}	—	—	0.95
θ_{123}	—	0.97	0.97
θ_{124}	—	—	0.97
θ_{134}	—	—	0.97
θ_{234}	—	—	0.97
θ_{1234}	—	—	0.98

As a statistical analysis of the simulated datasets, it is assumed the observed lifetimes as multivariate lifetimes and a MVG distribution described in Section 5.2.2. In this analysis, it is presented the Bayesian Monte Carlo estimators for the reliability function $R(t)$ for 2-components/3-components/4-components series system. The posterior summaries of interest will be not presented here due to the number of parameters of the considered series system. However, the results are summarized in Figures 33 and 34 in which it is presented the plots of the empirical estimated reliability functions and the Bayesian estimates for the reliability functions R(t) and the probability plots for the considered model assuming the observed data as multivariate lifetimes. In Figure 33, it is also presented the 95% credible intervals for the estimated reliability functions assuming the MVG distribution.

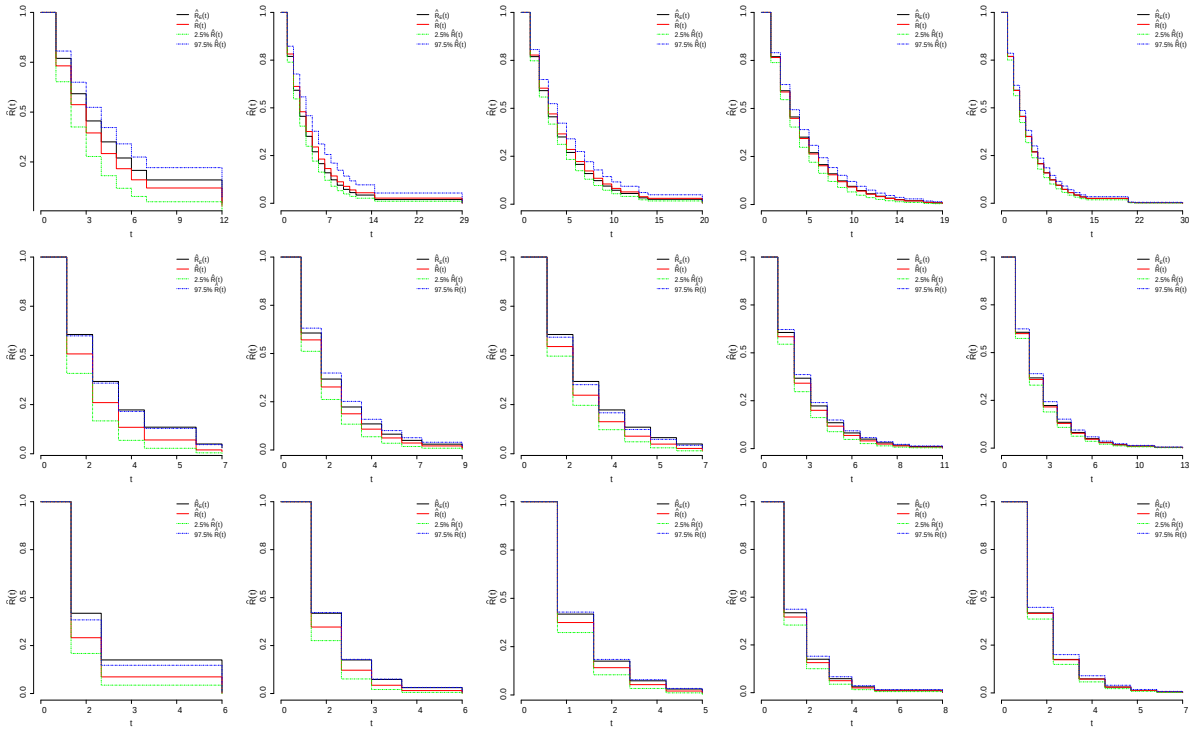

Figure 33 – The mean and 95 percent credible intervals for the reliability functions for the MVG model assuming the multivariate lifetimes of each considered series system (top to bottom: 2-components → 4-components) for each sample size (left to right: $n = 20 \rightarrow 300$).

Based on this simulation data, the reliability function for the entire system can also be estimated. For the specified time, $t = 1$, the true reliability value is obtained in each case by $R(1) = 0.7695$ (2-components), $R(1) = 0.6063$ (3-components) and $R(1) = 0.4184$ (4-components). The estimated Bayesian estimators based on the simulated Gibbs samples for $R(1)$ are presented in Table 24 for each sample size assuming the MVG model and assuming the independence structure.

Table 24 – Bayes estimators for $R(1)$ for each simulated dataset for each series system under dependence and independence assumption.

Dependence Assumption				Independence Assumption			
Sample	2-comp.	3-comp.	4-comp.	Sample	2-comp.	3-comp.	4-comp.
20	0.7319	0.5076	0.2912	20	0.7008	0.4387	0.2026
50	0.7822	0.5707	0.3466	50	0.7410	0.4813	0.1762
100	0.7776	0.5450	0.3748	100	0.7413	0.4597	0.2236
150	0.7652	0.5841	0.3960	150	0.7237	0.4700	0.2132
300	0.7687	0.5994	0.4161	300	0.7281	0.4927	0.2227

From the obtained results of Table 24, it is observed that for large sample sizes the Bayesian estimators for the reliability function at $t = 1$ are more accurate, especially under the dependence assumption. In this case, there is a great indication that using the multivariate lifetimes of the n-component system it is possible to get inference results

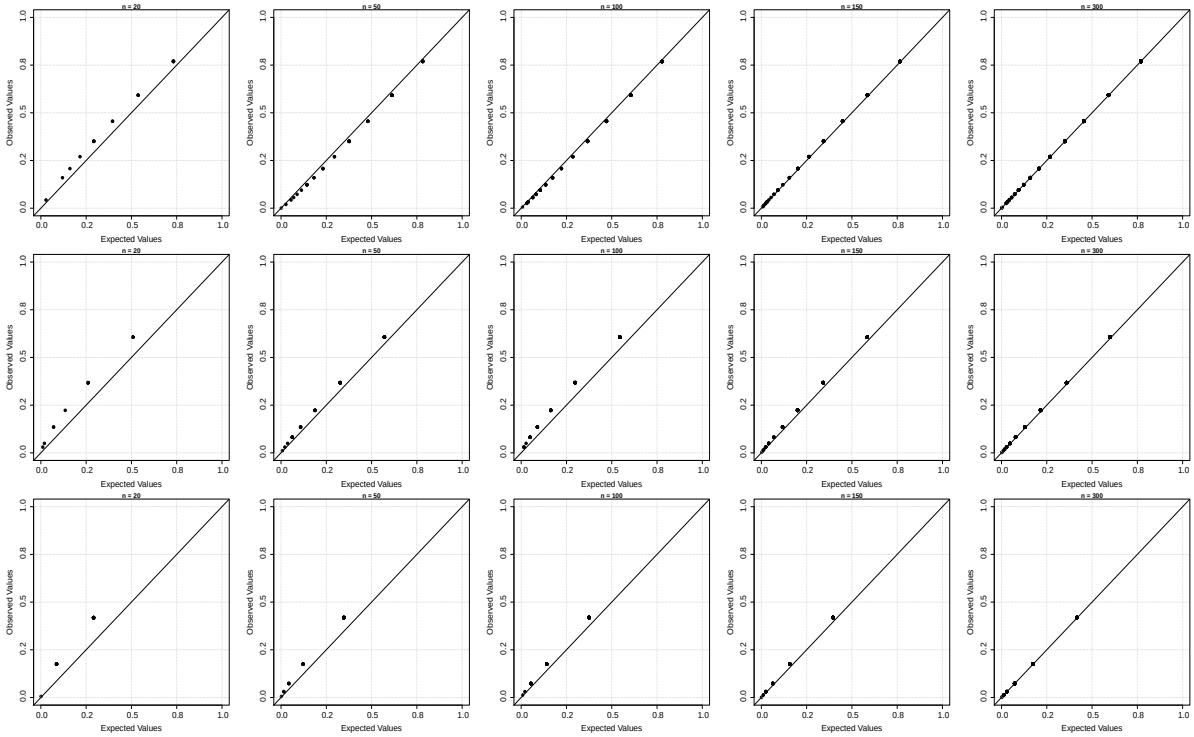

Figure 34 – PP-plots for the reliability functions for the MVG model assuming the multivariate lifetimes of each considered series system (top to bottom: 2-components → 4-components) for each sample size (left to right: $n = 20 \rightarrow 300$).

with great accuracy assuming the MVG distribution instead of the product of geometric distributions (independence assumption). In Figure 35, it is presented the posterior distribution for the reliability function at $t = 1$ estimated assuming the MVG distribution from where it is possible to conclude that the posterior distribution converges asymptotically to the normal distribution as expected.

5.2.5 Applications to Real Data

In order to illustrate our proposed methodology, in this section it is presented two real dataset applications assuming a special case, named trivariate geometric distribution, of the discrete MVG distribution introduced in this study. Classical and Bayesian methods were used to get the inferences of interest. Under a Classical approach the **BFGS** optimization method, available in the **maxLik** library of R software, was considered to obtain the maximum likelihood estimates (MLEs), whereas under a Bayesian approach the **R2jags** library from the R software was used to obtain the posterior summaries of interest.

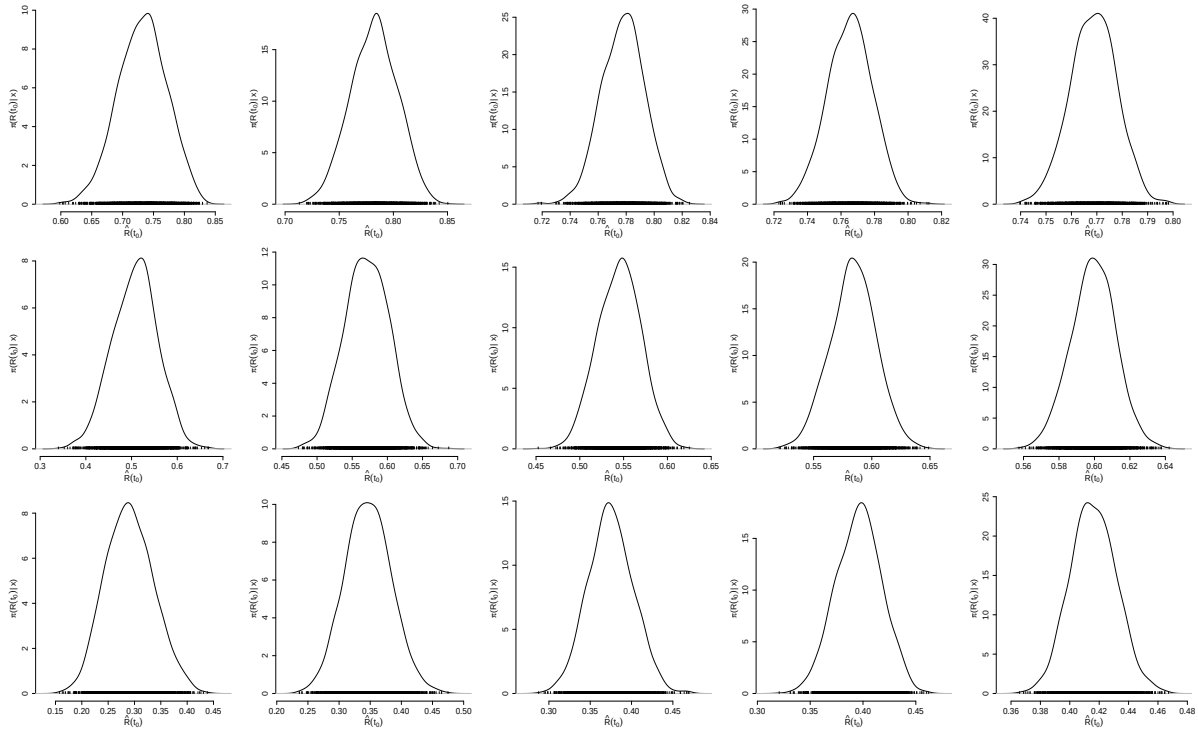

Figure 35 – Posterior density function for the reliability function at $t = 1$ for the MVG model assuming the multivariate lifetimes of each considered series system (top to bottom: 2-components $\rightarrow$ 4-components) for each sample size (left to right: $n = 20 \rightarrow 300$).

5.2.5.1 Automatic Blood Pressure Measuring Machine

As a first application, let us assume a real data set introduced by Bland and Altman (1999). This data set consists of systolic blood pressure measurements made simultaneously by two observers (J and R) and an automatic blood pressure measuring machine (S), each making three observations in quick succession (dataset in Table 25).

Table 25 – Automatic blood pressure measuring machine three measures.

X_1: Measure 1	122	121	95	127	140	139	122	130	119	126	107	123	131	123	127
	142	104	117	139	143	181	149	173	160	158	139	153	138	228	190
	103	131	131	126	121	97	116	215	141	153	113	109	145	192	112
	152	141	206	151	112	162	117	119	136	112	120	117	194	167	173
	228	77	154	154	145	200	188	149	136	128	204	184	163	93	178
	202	162	227	133	202	158	124	114	137	121					
X_2: Measure 2	128	127	94	127	131	142	112	129	122	113	113	125	129	126	119
	133	116	113	127	155	170	156	170	155	152	144	150	144	228	183
	99	131	123	129	114	94	121	201	133	143	107	105	102	178	116
	144	141	188	147	125	165	118	131	116	115	118	118	191	160	161
	218	89	156	155	154	180	147	217	132	125	222	187	160	88	181
	199	166	227	127	190	121	149	118	135	123					
X_3: Measure 3	124	128	98	135	124	136	112	135	122	111	111	125	122	114	126
	137	115	112	113	133	166	140	154	170	154	141	154	131	226	184
	106	124	124	125	125	96	127	207	146	138	102	97	137	171	116
	147	137	166	136	124	189	109	124	113	104	132	115	196	161	154
	189	101	141	148	166	179	139	192	133	142	224	192	152	88	181
	195	148	219	126	213	134	137	126	134	128					

For the statistical analysis, it is assumed as lifetimes the first measure X_1, the second measure X_2 and the third measure X_3. Since the dataset do not include censored observations, the MLE's were calculated using the **maxLik** function from the R software using the moments estimates as initial values for the optimization BFGS method. In addition, under a Bayesian approach, it was assumed Beta(1,1) prior distributions for the parameters θ_1, θ_2, θ_3, θ_{12}, θ_{13}, θ_{23} and θ_{123}. Monte Carlo Bayesian estimates were obtained using MCMC (Markov Chain Monte Carlo) simulation methods (see Table 26). Convergence of the MCMC algorithm was verified from traceplots of the simulated samples.

Table 26 – Maximum likelihood and Bayesian estimates for the automatic blood pressure measuring machine dataset.

Parameter	Classical approach		Bayesian approach	
	MLE (S.E)	95% Conf. Int.	Bayes (S.D.)	95% Cred. Int.
θ_1	0.9931 (0.0010)	(0.9911, 0.9952)	0.9948 (0.0008)	(0.9929, 0.9960)
θ_2	0.9952 (0.0007)	(0.9938, 0.9968)	0.9953 (0.0007)	(0.9939, 0.9964)
θ_3	0.9934 (0.0009)	(0.9917, 0.9951)	0.9942 (0.0008)	(0.9926, 0.9957)
θ_{12}	0.9989 (0.0008)	(0.9973, 1.0000)	0.9985 (0.0008)	(0.9971, 0.9997)
θ_{13}	1.0000 (0.0011)	(0.9977, 1.0000)	0.9996 (0.0003)	(0.9986, 1.0000)
θ_{23}	0.9983 (0.0008)	(0.9968, 1.0000)	0.9984 (0.0007)	(0.9969, 0.9994)
θ_{123}	0.9982 (0.0010)	(0.9962, 1.0000)	0.9987 (0.0007)	(0.9970, 0.9999)

From the results of Table 26, despite the results assuming both inference approaches are quite similar to estimate the model parameters, the obtained Bayesian inference results are in general better when compared to the obtained classical inference results. Figure 36 presents the plots of the marginal posterior densities from which it could be seen the convergence of the proposed model.

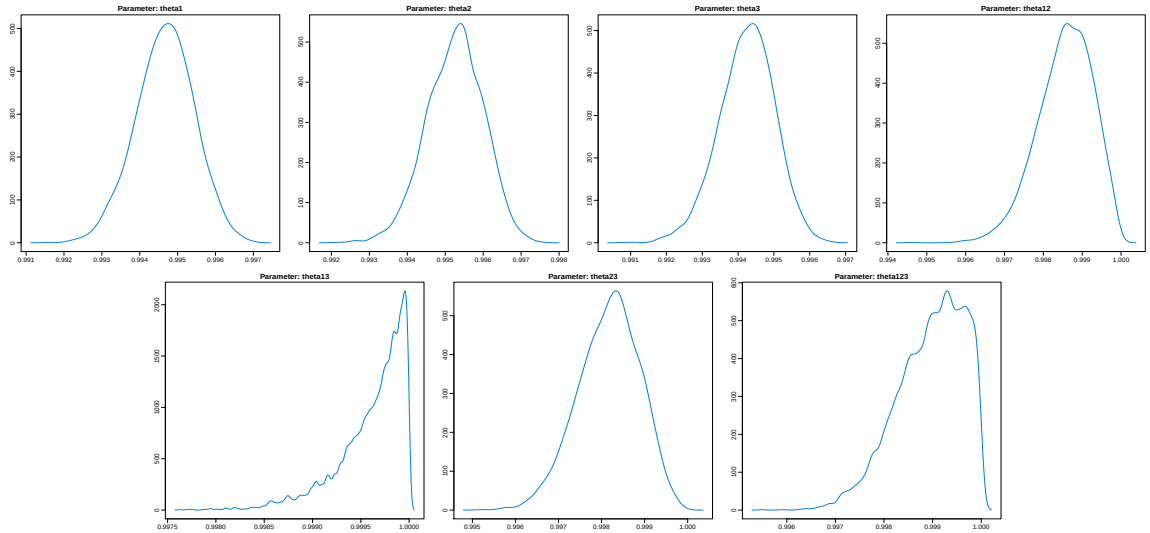

Figure 36 – Posterior density plots for each parameter assuming the MVG distribution for the automatic blood pressure measuring machine dataset.

5.2.5.2 Rats Litter-Matched Study of the Tumorigenesis

As a second application, let us assume a real data set introduced by Mantel et al. (1977). This dataset reports the results of a litter-matched study of the tumorigenesis of a drug and have the presence of censored times. In the experiment, female rats were chosen from fifty distinct litters, and one rat of each litter was randomly selected and given the drug. For each litter, two rats were selected as controls receiving placebo. Possible dependences between litter mates in their times to development of tumors may be due to common genetic backgrounds shared by siblings which is an indication for the use of a trivariate parametric model. The dataset is given in Table 27.

Table 27 – 50 litters of rats on tumorigenesis of a drug study

Litter	T_1	T_2	T_3	δ_1	δ_2	δ_3	Litter	T_1	T_2	T_3	δ_1	δ_2	δ_3	Litter	T_1	T_2	T_3	δ_1	δ_2	δ_3
1	101	104	49	0	0	1	18	104	104	74	0	0	0	35	45	104	79	1	0	0
2	104	104	102	0	0	0	19	81	104	69	0	0	0	36	94	104	104	1	0	0
3	104	104	104	0	0	0	20	67	104	68	1	0	1	37	104	104	104	0	0	0
4	77	97	79	0	0	0	21	104	104	104	0	0	0	38	104	101	94	0	1	0
5	89	104	104	0	0	0	22	104	104	104	0	0	0	39	76	84	78	0	1	1
6	88	104	96	1	0	1	23	104	83	40	0	0	1	40	80	80	76	1	1	0
7	104	94	77	1	0	1	24	87	104	104	0	0	0	41	72	104	95	1	0	0
8	96	104	104	1	0	0	25	104	104	104	0	0	0	42	73	104	66	1	0	1
9	82	104	77	0	0	0	26	89	104	104	0	0	0	43	92	104	102	1	0	1
10	70	104	77	1	0	0	27	78	104	104	0	0	0	44	104	98	78	0	0	0
11	89	91	90	1	0	0	28	104	81	64	0	1	1	45	55	104	104	0	0	0
12	91	92	70	0	0	0	29	86	94	55	1	0	1	46	49	83	77	0	0	0
13	39	50	45	1	1	0	30	34	104	54	1	0	1	47	89	104	104	1	0	0
14	103	91	69	1	0	0	31	76	87	74	0	0	0	48	88	99	79	0	0	0
15	93	104	103	0	0	0	32	103	84	73	1	1	1	49	103	104	91	1	0	0
16	85	104	72	0	0	0	33	102	104	80	1	0	0	50	104	104	79	0	0	1
17	104	104	63	0	0	0	34	80	104	73	1	0	0							

T_1 : treated rat, T_2, T_3 : control rats; $\delta_1, \delta_2, \delta_3$: censoring (1: non-censored, 0: censored)

Due to the great complexity of the likelihood function for $\boldsymbol{\theta}$ in the censored case, only a Bayesian approach was used to get inferences of interest assuming TG distribution. Thus, under a Bayesian analysis, it was assumed Beta(1,1) prior distributions for the parameters $\theta_1, \theta_2, \theta_3, \theta_{12}, \theta_{13}, \theta_{23}$ and θ_{123}, and the use of MCMC simulation methods to get the Monte Carlo Bayesian estimates. The obtained posterior summaries of interest are presented in Table 28 and Figure 37 presents the plots of the marginal posterior densities for each parameter.

Table 28 – Posterior summaries for the litters of rats on tumorigenesis dataset.

Parameter	Posterior Mean	Standard Deviation	95% Credibility Interval
θ_1	0.9955	0.0011	(0.9932, 0.9973)
θ_2	0.9994	0.0005	(0.9982, 0.9999)
θ_3	0.9968	0.0009	(0.9948, 0.9983)
θ_{12}	0.9996	0.0004	(0.9986, 1.0000)
θ_{13}	0.9995	0.0005	(0.9982, 1.0000)
θ_{23}	0.9994	0.0005	(0.9982, 1.0000)
θ_{123}	0.9996	0.0004	(0.9987, 1.0000)

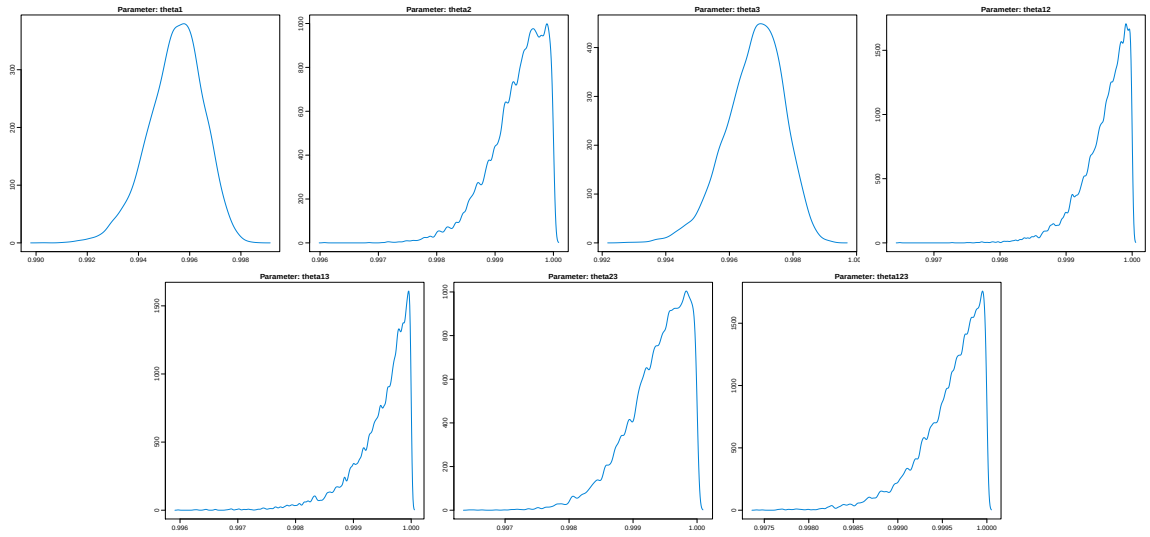

Figure 37 – Posterior density plots for each parameter assuming the TG distribution for the litters of rats on tumorigenesis dataset.

5.2.6 Concluding Remarks

In this study, it was introduced a new multivariate distribution obtained from a generalization of the bivariate Basu-Dhar geometric distribution introduced by Basu and Dhar (1995). Mathematical properties, inferences under classical and Bayesian approaches were introduced for the multivariate lifetimes in presence or not of right censored data. It was observed in this study that the mathematical expression for the likelihood function is relatively simple not requiring sophisticated computational expertize to get the inferences of interest especially under a Bayesian approach. Moreover, in all applications presented in this study, it was observed that the classical inference approach, in general, is not an appropriate inference method to get accurate inferences of interest for the MVG distribution due to the great computational instabilities. However, in censored case, the MVG showed some instability in the estimates of some parameters under a Bayesian approach using the MCMC simulation methodology. This fact may be related to the number of simulated samples, the final Gibbs sample or the presence autocorrelation of the Markov chains. Despite this fact, the MVG distribution could be a good alternative to be fitted to multivariate lifetime datasets with great accuracy, especially for complete data.

Chapter 6

General Conclusion

The search of appropriate probability distributions for data analysis still is a great problem in most studies, especially in medical area due the complexity of the data structure. In this thesis, it was presented some techniques to model survival medical data as well reliability data common in the engineering scenario. In a first approach, a bivariate discrete generalized Rayleigh (DBGR) distribution was proposed as a new model, obtained using the Marshall and Olkin (1997) method to add a new parameter to the survival function of the discrete Rayleigh distribution proposed by Roy (2004a) in order to have a more flexible joint survival function as an alternative to existing discrete models as the popular bivariate geometric distributions introduced by Arnold (see Arnold, 1975) and by Basu-Dhar (see Basu and Dhar, 1995). Some properties of this new distribution were also discussed in this study and an extension to multivariate case was provided. An extensive simulation study was performed to verify the effectiveness of the maximum likelihood estimation method assuming different fixed values for the parameters of the model and different sample sizes. The results obtained from Monte Carlo studies showed that the biases and RMSEs of the DBGR distribution are asymptotically non-biased and tends to zero when the sample size increases even assuming negative values for $\lambda_i, i = 1, 2$ in some scenarios. The model showed small computational costs to get the inference results as compared to many existing bivariate parametric lifetime distributions introduced in the literature or bivariate models derived from copula functions for continuous bivariate lifetime data (see, for example, Achcar et al., 2016b).

In a second approach, it was proposed a class of bivariate Lindley distributions based on stress and shock models. This class of models worked very well when applied to data of medical studies considered here even if they are more appropriate to reliability analysis. In fact, from the different model structure approaches considered in this study to get inferences of interest for a two-component series system reliability (see the computer system data application), it is possible to conclude that the use of the new class of BL distributions introduced in this thesis leads to accurate inference results for the parameters of interest assuming the system lifetime data as univariate or bivariate (de-

spite the presence of censored data) even using noninformative *priors* for the parameters of the model under a Bayesian approach. In special case, considering the EW dataset applications, the use of noninformative *prior* provided accurate inference results (small values for standard deviation) for each BL model. In addition, it is important to point out that the dependence structure is a crucial assumption as well the use of an informative *prior* distribution with hyperparameters based on sample mean and variance to get good inference accuracy for the $R(t)$ of the entire system which is a great indicator that the BL models can be applied in practical industrial situations where the series systems fails when a common source of shock destroy both components. Moreover, based on the $R(t)$ estimates, it is possible to conclude that the use of any of the proposed BL models may be a good alternative in the data analysis instead of the common use of product of two Lindley distributions considered under the independence assumption. Finally, based on the simulation study, it could be seen that the proposed models provide a better estimation for the reliability of the entire system when there are many two-component series systems put on test for a determined product as for example, computer systems depending on the processor and memory to work. This result could be of great interest of engineering studies or other areas of interest.

Finally, in a third approach, the extension of some models using cure rate models was presented as an alternative to the traditional models as the Cox model in the analysis of time-to-event data. The obtained results of this study showed many advantages for the use of cure fraction models in terms of great accuracy for the obtained point and interval inferences, great computational simplicity to get the inferences of interest under classical and Bayesian approaches and with simple interpretations for the parameters of the models which is an important point in medical applications.

In conclusion, the results emerging from this study reinforce the fact that the search of appropriate bivariate lifetime distributions could be extremely difficult depending on the correlation structure of the lifetime data. However, the proposed methodology could be very useful in the medical data analysis where the interest is the estimation of the fraction of patients in the studied population who never experience the event of interest. In addition, the identification of important covariates was also easily obtained assuming the proposed models even using non-informative priors for the parameters of the model, under a Bayesian approach. The results could be also extended to other cross-over trials in clinical research; reliability analysis in engineering; risk analysis in economics; among many others areas. For reproducible research, the general framework for the computer codes of the proposed modeling approach is presented in Appendix A at the end of this thesis which could be carried out using the OpenBugs software (Spiegelhalter et al., 2007) or R2jags (Su and Yajima, 2015) library from R software.

Appendix A. Models Programming Codes

```
## Discrete Weibull - Mixture Cure Rate Model

dDWEIBULLmix   <- function (y, mu, sigma, nu, log = FALSE)
{
  stopifnot(y > 0, mu > 0, mu < 1, sigma > 0, nu > 0, nu < 1)
  if(log)
  {
    pdf <- log(1 - nu) + log(mu ^ (y ^ sigma) - mu ^ ((y + 1) ^ sigma))
  }
  else
  {
    pdf <- (1 - nu) * (mu ^ (y ^ sigma) - mu ^ ((y + 1) ^ sigma))
  }
  return(pdf)
}

pDWEIBULLmix   <- function (q, mu, sigma, nu, lower.tail = TRUE, log.p =
    FALSE)
{
  stopifnot(q > 0, mu > 0, mu < 1, sigma > 0, nu > 0, nu < 1)
  if(lower.tail)
  {
    cdf <- 1 - (nu + (1 - nu) * (mu ^ (q + 1) ^ sigma))
  }
  else
  {
    cdf <- nu + (1 - nu) * (mu ^ (q + 1) ^ sigma)
  }
  if(log.p) return(log(cdf)) else return(cdf)
}
```

```
## Discrete Weibull - Non-mixture Cure Rate Model

dDWEIBULLnmix   <- function (y, mu, sigma, nu, log = FALSE)
{
  stopifnot(y > 0, mu > 0, mu < 1, sigma > 0, nu > 0, nu < 1)
  if(log)
  {
    pdf <- log(-log(nu)) + (y + 1) ^ sigma * log(mu) + sigma * log1p(y) +
        log(sigma) - log1p(y) + log(-log(mu)) + log(nu) - log(nu) * mu ^ (y
        + 1) ^ sigma
  }
  else
  {
    pdf <- log(nu) * mu ^ (y + 1) ^ sigma * (y + 1) ^ sigma * sigma / (y +
        1) * log(mu) * exp(log(nu) * (1 - mu ^ (y + 1) ^ sigma))
  }
  return(pdf)
}

pDWEIBULLnmix   <- function (q, mu, sigma, nu, lower.tail = TRUE, log.p =
    FALSE)
{
  stopifnot(q > 0, mu > 0, mu < 1, sigma > 0, nu > 0, nu < 1)
  if(lower.tail)
  {
    cdf <- 1 - exp(log(nu) * (1 - (mu ^ (q + 1) ^ sigma)))
  }
  else
  {
    cdf <- exp(log(nu) * (1 - (mu ^ (q + 1) ^ sigma)))
  }
  if(log.p) return(log(cdf)) else return(cdf)
}

## Block and Basu Bivariate Exponential Distribution

model.jags.bbe    <- function()
{
  lambda12        <- lambda1 + lambda2
  lambda13        <- lambda1 + lambda3
  lambda23        <- lambda2 + lambda3
  lambda          <- lambda1 + lambda2 + lambda3

  eta             <- g*(min(rho1,rho2) - rho1*rho2)
  phi01           <- rho1*(1-rho2) - eta
  phi10           <- (1-rho1)*rho2 - eta
  phi11           <- (1-rho1)*(1-rho2) + eta
  phi00           <- rho1 * rho2 + eta
```

```
alpha1          <- lambda * lambda1/lambda12
alpha2          <- lambda * lambda13/lambda12
alpha3          <- lambda3 * lambda/lambda12
alpha4          <- lambda * lambda23/lambda12
alpha5          <- lambda * lambda2/lambda12
alpha6          <- lambda3/lambda12
alpha7          <- lambda/lambda12

for (i in 1:n)
{
  phi[i]        <- -log(L[i])
  zeros[i]~dpois(phi[i])

  aux1[i]       <- exp(-lambda1 * t1[i] - lambda23 * t2[i])
  aux2[i]       <- exp(-lambda13 * t1[i] - lambda2 * t2[i])

  F1[i]         <- phi11 * lambda23 * alpha1 * aux1[i]
  F2[i]         <- phi11 * lambda2  * alpha2 * aux2[i]

  aux3[i]       <- lambda13 * exp(-lambda13 * t1[i])
  aux4[i]       <- lambda * exp(-lambda * t1[i])
  aux5[i]       <- alpha7 * aux3[i] - alpha6 * aux4[i]
  aux6[i]       <- alpha2 * aux2[i] - alpha3 * exp(-lambda * t1[i])

  dS1t1[i]      <- phi11 * alpha1 * aux1[i] + phi10 * aux5[i]
  dS2t1[i]      <- phi11 * aux6[i] + phi10 * aux5[i]

  aux7[i]       <- alpha4 * aux1[i] - alpha3 * exp(-lambda * t2[i])
  aux8[i]       <- lambda23 * exp(-lambda23 * t2[i])
  aux9[i]       <- lambda * exp(-lambda * t2[i])
  aux10[i]      <- alpha7 * aux8[i] - alpha6 * aux9[i]

  dS1t2[i]      <- phi11 * aux7[i] + phi01 * aux10[i]
  dS2t2[i]      <- phi11 * alpha5 * aux2[i] + phi01 * aux10[i]

  aux33[i]      <- exp(-lambda13 * t1[i])
  aux44[i]      <- exp(-lambda * t1[i])
  aux55[i]      <- alpha7 * aux33[i] - alpha6 * aux44[i]

  aux88[i]      <- exp(-lambda23 * t2[i])
  aux99[i]      <- exp(-lambda * t2[i])
  aux100[i]     <- alpha7 * aux88[i] - alpha6 * aux99[i]

  aux11[i]      <- alpha7 * aux1[i] - alpha6 * aux99[i]
  aux12[i]      <- alpha7 * aux2[i] - alpha6 * aux44[i]
```

```
    S1[i]          <- phi11 * aux11[i] + phi10 * aux55[i] + phi01 * aux100[i
        ] + phi00
    S2[i]          <- phi11 * aux12[i] + phi10 * aux55[i] + phi01 * aux100[i
        ] + phi00

    ## Log-Likehoood
    L1[i]          <- v.bbe[i] * c1[i] * c2[i] * log(F1[i]) + (1 - v.bbe[i])
         * c1[i] * c2[i] * log(F2[i])
    L2[i]          <- v.bbe[i] * c1[i] * (1-c2[i]) * log(dS1t1[i]) + (1 - v.
        bbe[i]) * c1[i] * (1 - c2[i]) * log(dS2t1[i])
    L3[i]          <- v.bbe[i] * (1 - c1[i]) * c2[i] * log(dS1t2[i]) + (1-v.
        bbe[i]) * (1 - c1[i]) * c2[i] * log(dS2t2[i])
    L4[i]          <- v.bbe[i] * (1 - c1[i]) * (1 - c2[i]) * log(S1[i]) + (1
         - v.bbe[i]) * (1 - c1[i]) * (1 - c2[i]) * log(S2[i])

    L[i]           <- exp(L1[i] + L2[i] + L3[i] + L4[i])
  }

## Priors
lambda1~dgamma(0.001,0.001)
lambda2~dgamma(0.001,0.001)
lambda3~dgamma(0.001,0.001)
rho1~dbeta(1,1)
rho2~dbeta(1,1)
g~dbeta(1,1)
}

##  Gumbel Bivariate Exponential Distribution

model.jags.gum     <- function()
{
  eta              <- g*(min(rho1,rho2) - rho1*rho2)
  phi01            <- rho1*(1-rho2) - eta
  phi10            <- (1-rho1)*rho2 - eta
  phi11            <- (1-rho1)*(1-rho2) + eta
  phi00            <- rho1 * rho2 + eta

  eta1             <- (1 - theta) * lambda1 * lambda2
  eta2             <- theta * lambda1^2 * lambda2
  eta3             <- theta * lambda2^2 * lambda1
  eta4             <- theta^2 * lambda1^2 * lambda2^2
  eta5             <- lambda1 * lambda2 * theta

  for (i in 1:n)
  {
    phi[i]         <-  -log(L[i])
    zeros[i]~dpois(phi[i])
```

```
    S[i]          <- exp(-(lambda1 * t1[i] + lambda2 * t2[i] + eta5 * t1[i]
        * t2[i]))
    S1[i]         <- phi11 * S[i] + phi10 * exp(-lambda1 * t1[i]) + phi01 *
        exp(-lambda2 * t2[i]) + phi00

    F[i]          <- phi11 * (eta1 + eta2 * t1[i] + eta3 * t2[i] + eta4 *
       t1[i] * t2[i]) * S[i]

    dSt1[i]       <- phi11 * (lambda1*lambda2*t2[i]*theta+lambda1) * S[i] +
        phi10 * lambda1 * exp(-lambda1 * t1[i])

    dSt2[i]       <- phi11 * (lambda1*lambda2*t1[i]*theta+lambda2) * S[i] +
        phi01 * lambda2 * exp(-lambda2 * t2[i])

    ## Log-Likehoood
    L1[i]         <- c1[i] * c2[i] * log(F[i])
    L2[i]         <- c1[i] * (1-c2[i]) * log(dSt1[i])
    L3[i]         <- (1 - c1[i]) * c2[i] * log(dSt2[i])
    L4[i]         <- (1 - c1[i]) * (1 - c2[i]) * log(S1[i])
    L[i]          <- exp(L1[i] + L2[i] + L3[i] + L4[i])
  }

## Priors
lambda1~dgamma(0.001,0.001)
lambda2~dgamma(0.001,0.001)
theta~dgamma(0.001,0.001)
rho1~dbeta(1,1)
rho2~dbeta(1,1)
g~dbeta(1,1)
}

## Marshall-Olkin Bivariate Exponential Distribution

model.jags.mobe   <- function()
{
  eta             <- g*(min(rho1,rho2) - rho1*rho2)
  phi01           <- rho1*(1-rho2) - eta
  phi10           <- (1-rho1)*rho2 - eta
  phi11           <- (1-rho1)*(1-rho2) + eta
  phi00           <- rho1 * rho2 + eta

  for (i in 1:n)
  {
    phi[i]        <- -log(L[i])
    zeros[i]~dpois(phi[i])
```

```
S1[i]          <- exp(-(lambda1 + lambda12) * t1[i] - lambda2 * t2[i])
S2[i]          <- exp(-(lambda2 + lambda12) * t2[i] - lambda1 * t1[i])
S3[i]          <- exp(-(lambda1 + lambda2 + lambda12) * t1[i])

F1[i]          <- (lambda1 + lambda12) * exp(-(lambda1 + lambda12) * t1[
    i])
F2[i]          <- (lambda2 + lambda12) * exp(-(lambda2 + lambda12) * t2[
    i])
F3[i]          <- (lambda2 + lambda12) * exp(-(lambda2 + lambda12) * t1[
    i])

SS1[i]         <- exp(-(lambda1 + lambda12) * t1[i])
SS2[i]         <- exp(-(lambda2 + lambda12) * t2[i])
SS3[i]         <- exp(-(lambda2 + lambda12) * t1[i])

A1[i]          <- phi11*lambda2 * (lambda1 + lambda12) * S1[i]
A2[i]          <- phi11*lambda1 * (lambda2 + lambda12) * S2[i]
A3[i]          <- phi11*lambda12 * S3[i]

B1[i]          <- phi11*(lambda1 + lambda12) * S1[i] + phi10*F1[i]
B2[i]          <- phi11*lambda1 * S2[i] + phi10*F1[i]
B3[i]          <- phi11*(lambda1+lambda2+lambda12)*S3[i] + phi10*F1[i] +
     phi01*F3[i]

C1[i]          <- phi11*lambda2 * S1[i] + phi01*F2[i]
C2[i]          <- phi11*(lambda2 + lambda12) * S2[i] + phi01*F2[i]

D1[i]          <- phi11*S1[i] + phi10*SS1[i] + phi01*SS2[i] + phi00
D2[i]          <- phi11*S2[i] + phi10*SS1[i] + phi01*SS2[i] + phi00
D3[i]          <- phi11*S3[i] + phi10*SS1[i] + phi01*SS3[i] + phi00

## Log-Likehoood
L1[i]          <- v1.mobe[i]*(1 - v2.mobe[i]) * c1[i] * c2[i] * log(A1[i
    ])
L2[i]          <- v2.mobe[i]*(1 - v1.mobe[i]) * c1[i] * c2[i] * log(A2[i
    ])
L3[i]          <- (1 - v1.mobe[i])*(1 - v2.mobe[i]) * c1[i] * c2[i] *
    log(A3[i])
L4[i]          <- L1[i] + L2[i] + L3[i]

L5[i]          <- v1.mobe[i]*(1 - v2.mobe[i]) * c1[i] * (1 - c2[i]) *
    log(B1[i])
L6[i]          <- v2.mobe[i]*(1 - v1.mobe[i]) * c1[i] * (1 - c2[i]) *
    log(B2[i])
L7[i]          <- (1 - v1.mobe[i])*(1 - v2.mobe[i]) * c1[i] * (1 - c2[i
    ]) * log(B3[i])
L8[i]          <- L5[i] + L6[i] + L7[i]
```

```
    L9[i]          <- v1.mobe[i]*(1 - v2.mobe[i]) * (1 - c1[i]) * c2[i] *
        log(C1[i])
    L10[i]         <- v2.mobe[i]*(1 - v1.mobe[i]) * (1 - c1[i]) * c2[i] *
        log(C2[i])
    L11[i]         <- L9[i] + L10[i]

    L12[i]         <- v1.mobe[i]*(1 - v2.mobe[i]) * (1 - c1[i]) * (1 - c2[i
        ]) * log(D1[i])
    L13[i]         <- v2.mobe[i]*(1 - v1.mobe[i]) * (1 - c1[i]) * (1 - c2[i
        ]) * log(D2[i])
    L14[i]         <- (1 - v1.mobe[i])*(1 - v2.mobe[i]) * (1 - c1[i]) * (1 -
         c2[i]) * log(D3[i])
    L15[i]         <- L12[i] + L13[i] + L14[i]

    L[i]           <- exp(L4[i] + L8[i] + L11[i] + L15[i])
  }

## Priors
lambda1~dgamma(0.001,0.001)
lambda2~dgamma(0.001,0.001)
lambda12~dgamma(0.001,0.001)
rho1~dbeta(1,1)
rho2~dbeta(1,1)
g~dbeta(1,1)
}

## Bivariate Lindley Distribution Type I - BL-I

model.jags.bl1     <- function()
{
  eta              <- g*(min(rho1,rho2) - rho1*rho2)
  phi01            <- rho1*(1-rho2) - eta
  phi10            <- (1-rho1)*rho2 - eta
  phi11            <- (1-rho1)*(1-rho2) + eta
  phi00            <- rho1 * rho2 + eta

  for (i in 1:n)
  {
    phi[i]         <- -log(L[i])
    zeros[i]~dpois(phi[i])

    fl1[i]         <- beta1^2/(1 + beta1) * (1 + t1[i]) * exp(-beta1 * t1[i
        ])
    fl2[i]         <- beta2^2/(1 + beta2) * (1 + t2[i]) * exp(-beta2 * t2[i
        ])
    fl3[i]         <- beta3^2/(1 + beta3) * (1 + t1[i]) * exp(-beta3 * t1[i
```

```
    ])
fl4[i]          <- beta3^2/(1 + beta3) * (1 + t2[i]) * exp(-beta3 * t2[i
    ])
fl5[i]          <- beta2^2/(1 + beta2) * (1 + t1[i]) * exp(-beta2 * t1[i
    ])


fe1[i]          <- (beta1 + beta3) * exp(-(beta1 + beta3) * t1[i])
fe2[i]          <- (beta2 + beta3) * exp(-(beta2 + beta3) * t2[i])
fe3[i]          <- beta3 * exp(-beta3 * t1[i])
fe4[i]          <- beta3 * exp(-beta3 * t2[i])

se1[i]          <- exp(-(beta1 + beta3) * t1[i])
se2[i]          <- exp(-(beta2 + beta3) * t2[i])
se3[i]          <- exp(-(beta3) * t1[i])
se4[i]          <- exp(-(beta3) * t2[i])

sl1[i]          <- (1 + beta1 + beta1 * t1[i])/(1 + beta1) * exp(-beta1 *
     t1[i])
sl2[i]          <- (1 + beta2 + beta2 * t2[i])/(1 + beta2) * exp(-beta2 *
     t2[i])
sl3[i]          <- (1 + beta2 + beta2 * t1[i])/(1 + beta2) * exp(-beta2 *
     t1[i])

M1[i]           <- (1 + (beta1*t1[i])/(1 + beta1)) * exp(-beta1*t1[i] -
    beta3*t1[i])
M2[i]           <- (1 + (beta2*t2[i])/(1 + beta2)) * exp(-beta2*t2[i] -
    beta3*t2[i])

A1[i]           <-  phi11 * (fl2[i] * (beta1 * (1 + t1[i])/(1 + beta1) *
    fe1[i] + beta3/(1 + beta1) * se1[i]))
A2[i]           <-  phi11 * (fl1[i] * (beta2 * (1 + t2[i])/(1 + beta2) *
    fe2[i] + beta3/(1 + beta2) * se2[i]))
A3[i]           <-  phi11 * (fe3[i] * sl1[i] * sl3[i])

B1[i]           <- phi11 * (fl1[i] * sl2[i] * se3[i] + sl1[i] * sl2[i] *
    fe3[i]) + phi10 * (fl1[i] * se3[i] + fe3[i] * sl1[i])
B2[i]           <- phi11 * (fl1[i] * sl2[i] * se4[i]) + phi10 * (fl1[i] *
     se3[i] + fe3[i] * sl1[i])
B3[i]           <- phi11 * (fl1[i] * sl3[i] * se1[i] + sl1[i] * (fl3[i] *
     se1[i] + fe1[i] * sl3[i])) + phi10 * (fl1[i] * se3[i] + fe3[i] *
    sl1[i]) + phi01 * (fl5[i] * se3[i] + fe3[i] * sl3[i])

C1[i]           <- phi11 * (sl1[i] * fl2[i] * se3[i]) + phi01 * (fl2[i] *
     se4[i] + fe4[i] * sl2[i])
C2[i]           <- phi11 * (fl2[i] * sl1[i] * se4[i] + sl1[i] * sl2[i] *
    fe4[i]) + phi01 * (fl2[i] * se4[i] + fe4[i] * sl2[i])
```

```
    D1[i]          <- phi11 * sl1[i] * sl2[i] * se3[i] + phi10 * M1[i] +
        phi01 * M2[i] + phi00
    D2[i]          <- phi11 * sl1[i] * sl2[i] * se4[i] + phi10 * M1[i] +
        phi01 * M2[i] + phi00
    D3[i]          <- phi11 * sl1[i] * sl3[i] * se3[i] + phi10 * M1[i] +
        phi01 * M2[i] + phi00

    ## Log-Likehoood
    P1[i]          <- v1.bl1[i]*(1 - v2.bl1[i]) * c1[i] * c2[i] * log(A1[i])
         + v2.bl1[i]*(1 - v1.bl1[i]) * c1[i] * c2[i] * log(A2[i]) + (1 - v1.
        bl1[i])*(1 - v2.bl1[i]) * c1[i] * c2[i] * log(A3[i])
    P2[i]          <- v1.bl1[i]*(1 - v2.bl1[i]) * c1[i] * (1 - c2[i]) * log
        ((B1[i])) + v2.bl1[i]*(1 - v1.bl1[i]) * c1[i] * (1 - c2[i]) * log((
        B2[i])) + (1 - v1.bl1[i])*(1 - v2.bl1[i]) * c1[i] * (1 - c2[i]) *
        log((B3[i]))
    P3[i]          <- v1.bl1[i]*(1 - v2.bl1[i]) * (1 - c1[i]) * c2[i] * log
        ((C1[i])) + v2.bl1[i]*(1 - v1.bl1[i]) * (1 - c1[i]) * c2[i] * log((
        C2[i]))
    P4[i]          <- v1.bl1[i]*(1 - v2.bl1[i]) * (1 - c1[i]) * (1 - c2[i])
        * log((D1[i])) + v2.bl1[i]*(1 - v1.bl1[i]) * (1 - c1[i]) * (1 - c2[i
        ]) * log((D2[i])) + (1 - v1.bl1[i])*(1 - v2.bl1[i]) * (1 - c1[i]) *
        (1 - c2[i]) * log((D3[i]))

    L[i]           <- exp(P1[i] + P2[i] + P3[i] + P4[i])
  }

## Priors
beta1~dgamma(0.001,0.001)
beta2~dgamma(0.001,0.001)
beta3~dgamma(0.001,0.001)
rho1~dbeta(1,1)
rho2~dbeta(1,1)
g~dbeta(1,1)
}

## Arnold Bivariate Geometric Distribution

model.jags.arn <- function()
{
  gamma1         <- 1 - theta1 - theta2
  gamma2         <- 1 - theta1
  gamma3         <- 1 - theta2

  eta            <- g*(min(rho1,rho2) - rho1*rho2)
  phi01          <- rho1*(1-rho2) - eta
  phi10          <- (1-rho1)*rho2 - eta
  phi11          <- (1-rho1)*(1-rho2) + eta
```

```
phi00          <- rho1 * rho2 + eta

for (i in 1:n)
{
  phi[i]       <- -log(L[i])
  zeros[i]~dpois(phi[i])

  a1[i]        <- gamma1^(t1[i] - 1)
  a2[i]        <- gamma3^(t2[i] - t1[i] - 1)
  a3[i]        <- gamma1^(t2[i] - 1)
  a4[i]        <- gamma2^(t1[i] - t2[i] - 1)

  P1[i]        <- phi11 * theta1 * theta2 * a1[i] * a2[i]
  P2[i]        <- phi11 * theta1 * theta2 * a3[i] * a4[i]

  a5[i]        <- gamma1^(t2[i])
  a6[i]        <- gamma2^(t1[i] - t2[i] - 1)

  S1[i]        <- phi11 * theta1 * a1[i] * a2[i] + phi10 * theta1 * (1 -
     theta1)^(t1[i] - 1)
  S2[i]        <- phi11 * theta1 * a5[i] * a6[i] + phi10 * theta1 * (1 -
     theta1)^(t1[i] - 1)

  a7[i]        <- gamma1^t1[i]
  a8[i]        <- gamma3^(t2[i] - t1[i] - 1)
  a9[i]        <- gamma2^(t1[i] - t2[i])

  R1[i]        <- phi11 * theta2 * a8[i] * a7[i] + phi01 * theta2 * (1 -
     theta2)^(t2[i] - 1)
  R2[i]        <- phi11 * theta2 * a9[i] * a3[i] + phi01 * theta2 * (1 -
     theta2)^(t2[i] - 1)

  a10[i]       <- gamma3^(t2[i] - t1[i])

  U1[i]        <- phi11 * a10[i] * a7[i] + phi10 * (1 - theta1)^t1[i] +
     phi01 * (1 - theta2)^t2[i] + phi00
  U2[i]        <- phi11 * a9[i]  * a5[i] + phi10 * (1 - theta1)^t1[i] +
     phi01 * (1 - theta2)^t2[i] + phi00

  ## Log-Likelihood
  L1[i]        <- v.arn[i] * c1[i] * c2[i] * log(P1[i]) + (1 - v.arn[i]) *
      c1[i] * c2[i] * log(P2[i])
  L2[i]        <- v.arn[i] * c1[i] * (1 - c2[i]) * log(S1[i]) + (1 - v.arn
     [i]) * c1[i] * (1 - c2[i]) * log(S2[i])
  L3[i]        <- v.arn[i] * (1 - c1[i]) * c2[i] * log(R1[i]) + (1 - v.arn
     [i]) * (1 - c1[i]) * c2[i] * log(R2[i])
  L4[i]        <- v.arn[i] * (1 - c1[i]) * (1 - c2[i]) * log(U1[i]) + (1 -
```

```
        v.arn[i])*(1 - c1[i])*(1 - c2[i]) * log(U2[i])
    L[i]          <- exp(L1[i] + L2[i] + L3[i] + L4[i])
  }

## Priors
theta1~dbeta(1,1)
theta2~dbeta(1,1)
rho1~dbeta(1,1)
rho2~dbeta(1,1)
g~dbeta(1,1)
}

## Basu-Dhar Bivariate Geometric Distribution

model.jags.bd   <- function()
{
  eta           <- g*(min(rho1,rho2) - rho1*rho2)
  phi01         <- rho1*(1-rho2) - eta
  phi10         <- (1-rho1)*rho2 - eta
  phi11         <- (1-rho1)*(1-rho2) + eta
  phi00         <- rho1 * rho2 + eta

  for (i in 1:n)
  {
    phi[i]      <- -log(L[i])
    zeros[i]~dpois(phi[i])

    z1[i]       <- max(t1[i] - 1,t2[i] - 1)
    z2[i]       <- max(t1[i],t2[i] - 1)
    z3[i]       <- max(t1[i] - 1,t2[i])
    z4[i]       <- max(t1[i],t2[i])

    A[i]        <- p1^(t1[i] - 1) * p2^(t2[i] - 1) * p12^(z1[i])
    B[i]        <- p1^(t1[i])     * p2^(t2[i] - 1) * p12^(z2[i])
    C[i]        <- p1^(t1[i] - 1) * p2^(t2[i])     * p12^(z3[i])
    D[i]        <- p1^(t1[i])     * p2^(t2[i])     * p12^(z4[i])

    A1[i]       <- (A[i] - B[i]  - C[i] + D[i])  * phi11
    B1[i]       <- (C[i] - D[i]) * phi11 + phi10 * (p1 * p12)^(t1[i] - 1) *
        (1 - p1*p12)
    C1[i]       <- (B[i] - D[i]) * phi11 + phi01 * (p2 * p12)^(t2[i] - 1) *
        (1 - p2*p12)
    D1[i]       <- phi11 * D[i]  + phi10 * (p1 * p12)^(t1[i]) + phi01 * (p2
        * p12)^(t2[i]) + phi00

    ## Log-Likehoood
    E[i]        <- c1[i] * c2[i] * log(A1[i])
```

```
    K[i]          <- c1[i] * (1 - c2[i]) * log(B1[i])
    G[i]          <- c2[i] * (1 - c1[i]) * log(C1[i])
    H[i]          <- (1 - c1[i]) * (1 - c2[i]) * log(D1[i])
    L[i]          <- exp(E[i] + K[i] + G[i] + H[i])
  }

## Priors
p1~dbeta(1,1)
p2~dbeta(1,1)
p12~dbeta(1,1)
rho1~dbeta(1,1)
rho2~dbeta(1,1)
g~dbeta(1,1)
}

##  Bivariate Geometric Distribution Type-II

model.jags.bgII <- function()
{
  eta           <- g*(min(rho1,rho2) - rho1*rho2)
  phi01         <- rho1*(1-rho2) - eta
  phi10         <- (1-rho1)*rho2 - eta
  phi11         <- (1-rho1)*(1-rho2) + eta
  phi00         <- rho1 * rho2 + eta

  for (i in 1:n)
  {
    phi[i]      <- -log(L[i])
    zeros[i]~dpois(phi[i])

    z1[i]       <- min(t1[i] + 1,t2[i] + 1)
    z2[i]       <- min(t1[i],t2[i] + 1)
    z3[i]       <- min(t1[i] + 1,t2[i])
    z4[i]       <- min(t1[i],t2[i])

    A[i]        <- p1^(t1[i] + 1) * p2^(t2[i] + 1) * p12^(z1[i])
    B[i]        <- p1^(t1[i])     * p2^(t2[i] + 1) * p12^(z2[i])
    C[i]        <- p1^(t1[i] + 1) * p2^(t2[i])     * p12^(z3[i])
    D[i]        <- p1^(t1[i])     * p2^(t2[i])     * p12^(z4[i])

    A1[i]       <- (A[i] - B[i] - C[i] + D[i]) * phi11
    B1[i]       <- (C[i] - D[i]) * phi11 + phi10 * (p1)^(t1[i] - 1) * (1 -
        p1)
    C1[i]       <- (B[i] - D[i]) * phi11 + phi01 * (p2)^(t2[i] - 1) * (1 -
        p2)
    D1[i]       <- phi11 * D[i] + phi10 * (p1)^(t1[i]) + phi01 * (p2)^(t2[i
        ]) + phi00
```

```
    ## Log-Likehoood
    E[i]          <- c1[i] * c2[i] * log(abs(A1[i]))
    K[i]          <- c1[i] * (1 - c2[i]) * log(abs(B1[i]))
    G[i]          <- c2[i] * (1 - c1[i]) * log(abs(C1[i]))
    H[i]          <- (1 - c1[i]) * (1 - c2[i]) * log(D1[i])
    L[i]          <- exp(E[i] + K[i] + G[i] + H[i])
  }

## Priors
p1~dbeta(1,1)
p2~dbeta(1,1)
p12~dbeta(1,1)
rho1~dbeta(1,1)
rho2~dbeta(1,1)
g~dbeta(1,1)
}

## Discrete Bivariate Generalized Rayleigh

model.jags.dgr <- function()
{
  zeta          <- g*(min(rho1,rho2) - rho1*rho2)
  phi01         <- rho1*(1-rho2) - zeta
  phi10         <- (1-rho1)*rho2 - zeta
  phi11         <- (1-rho1)*(1-rho2) + zeta
  phi00         <- rho1 * rho2 + zeta

  eta           <- 1 - alpha

  for (i in 1:n)
  {
    phi[i]      <- -log(L[i])
    zeros[i]~dpois(phi[i])

    A1[i]       <- lambda1^(t1[i]^2 + (t1[i] + 1)^2)
    A2[i]       <- lambda2^(t2[i]^2 + (t2[i] + 1)^2)
    A3[i]       <- lambda1^((t1[i] + 1)^2)
    A4[i]       <- lambda2^((t2[i] + 1)^2)
    A5[i]       <- lambda1^(t1[i]^2)
    A6[i]       <- lambda2^(t2[i]^2)

    # Log-Likelihood
    L1[i]       <- 1 - eta^2 * A1[i] * A2[i]
    L2[i]       <- A5[i] - A3[i]
    L3[i]       <- A6[i] - A4[i]
    L4[i]       <- 1 - eta * A5[i] * A6[i]
```

```
    L5[i]       <-  1 - eta * A3[i] * A6[i]
    L6[i]       <-  1 - eta * A5[i] * A4[i]
    L7[i]       <-  1 - eta * A3[i] * A4[i]
    L8[i]       <-  alpha * A5[i] * A4[i]
    L9[i]       <-  alpha * A3[i] * A6[i]
    L10[i]      <-  alpha * A5[i] * A6[i]

    LL1[i]      <-  ((L1[i] * L2[i] * L3[i] * alpha) / (L4[i] * L5[i] * L6[
        i] * L7[i])) * phi11
    LL2[i]      <-  phi11 * (L8[i]/L6[i] - L10[i]/L4[i]) + phi10 * ((alpha
        * (A5[i] - A3[i]))/((1 - eta * A3[i]) * (1 - eta * A5[i])))
    LL3[i]      <-  phi11 * (L9[i]/L5[i] - L10[i]/L4[i]) + phi01 * ((alpha
        * (A6[i] - A4[i]))/((1 - eta * A4[i]) * (1 - eta * A6[i])))
    LL4[i]      <-  L10[i]/L4[i] * phi11 + phi10 * ((alpha * A5[i])/(1 -
        eta * A5[i])) + phi01 * ((alpha * A6[i])/(1 - eta * A6[i])) + phi00

    C1[i]       <-  c1[i] * c2[i] * log((LL1[i]))
    C2[i]       <-  c1[i] * (1 - c2[i]) * log((LL2[i]))
    C3[i]       <-  (1 - c1[i]) * c2[i] * log((LL3[i]))
    C4[i]       <-  (1 - c1[i]) * (1 - c2[i]) * log((LL4[i]))

    L[i]        <-  exp(C1[i] + C2[i] + C3[i] + C4[i])
  }

# Priors
lambda1~dbeta(1,1)
lambda2~dbeta(1,1)
alpha~dunif(0,1)
rho1~dbeta(1,1)
rho2~dbeta(1,1)
g~dbeta(1,1)
}

# Trivariate Geometric Distribution - Special Case of MVG Distribution

model.jags.tg   <-      function()
{
    for (i in 1:n)
    {
        phi[i]  <-      -log(L[i])
        zeros[i]~dpois(phi[i])

        # Max between ti's
        b1[i]   <-      max(t1[i], t2[i])
        b2[i]   <-      max(t1[i], t3[i])
        b3[i]   <-      max(t2[i], t3[i])
        b4[i]   <-      max(t1[i] - 1, t2[i])
        b5[i]   <-      max(t1[i] - 1, t3[i])
```

```
b6[i]   <-      max(t2[i] - 1, t3[i])
b7[i]   <-      max(t1[i], t2[i] - 1)
b8[i]   <-      max(t1[i], t3[i] - 1)
b9[i]   <-      max(t2[i], t3[i] - 1)
b10[i]  <-      max(t1[i] - 1, t2[i] - 1)
b11[i]  <-      max(t1[i] - 1, t3[i] - 1)
b12[i]  <-      max(t2[i] - 1, t3[i] - 1)
b13[i]  <-      max(t1[i], max(t2[i], t3[i]))
b14[i]  <-      max(t1[i] - 1, max(t2[i], t3[i]))
b15[i]  <-      max(t1[i] - 1, max(t2[i] - 1, t3[i]))
b16[i]  <-      max(t1[i] - 1, max(t2[i], t3[i] - 1))
b17[i]  <-      max(t1[i] - 1, max(t2[i] - 1, t3[i] - 1))
b18[i]  <-      max(t1[i], max(t2[i] - 1, t3[i]))
b19[i]  <-      max(t1[i], max(t2[i] - 1, t3[i] - 1))
b20[i]  <-      max(t1[i], max(t2[i], t3[i] - 1))

# Likelihood auxiliary parts
P1[i]   <-      p1^(t1[i] - 1)  * p2^(t2[i] - 1) * p3^(t3[i] - 1)
P2[i]   <-      p1^(t1[i] - 1)  * p2^(t2[i] - 1) * p3^(t3[i])
P3[i]   <-      p1^(t1[i] - 1)  * p2^(t2[i])     * p3^(t3[i] - 1)
P4[i]   <-      p1^(t1[i])      * p2^(t2[i] - 1) * p3^(t3[i] - 1)
P5[i]   <-      p1^(t1[i] - 1)  * p2^(t2[i])     * p3^(t3[i])
P6[i]   <-      p1^(t1[i])      * p2^(t2[i] - 1) * p3^(t3[i])
P7[i]   <-      p1^(t1[i])      * p2^(t2[i])     * p3^(t3[i] - 1)
P8[i]   <-      p1^(t1[i])      * p2^(t2[i])     * p3^(t3[i])

Q1[i]   <-      p12^(b10[i]) * p13^(b11[i]) * p23^(b12[i]) * p123^(
    b17[i])
Q2[i]   <-      p12^(b10[i]) * p13^(b5[i])  * p23^(b6[i])  * p123^(
    b15[i])
Q3[i]   <-      p12^(b4[i])  * p13^(b11[i]) * p23^(b9[i])  * p123^(
    b16[i])
Q4[i]   <-      p12^(b7[i])  * p13^(b8[i])  * p23^(b12[i]) * p123^(
    b19[i])
Q5[i]   <-      p12^(b4[i])  * p13^(b5[i])  * p23^(b3[i])  * p123^(
    b14[i])
Q6[i]   <-      p12^(b7[i])  * p13^(b2[i])  * p23^(b6[i])  * p123^(
    b18[i])
Q7[i]   <-      p12^(b1[i])  * p13^(b8[i])  * p23^(b9[i])  * p123^(
    b20[i])
Q8[i]   <-      p12^(b1[i])  * p13^(b2[i])  * p23^(b3[i])  * p123^(
    b13[i])

# Likelihood final
K1[i]   <-      P1[i] * Q1[i]
K2[i]   <-      P2[i] * Q2[i]
K3[i]   <-      P3[i] * Q3[i]
```

```
        K4[i]   <-      P4[i] * Q4[i]
        K5[i]   <-      P5[i] * Q5[i]
        K6[i]   <-      P6[i] * Q6[i]
        K7[i]   <-      P7[i] * Q7[i]
        K8[i]   <-      P8[i] * Q8[i]

        L1[i]   <-      K1[i]
        L2[i]   <-      K2[i] + K3[i] + K4[i]
        L3[i]   <-      K5[i] + K6[i] + K7[i]
        L4[i]   <-      K8[i]
        L5[i]   <-      L1[i] - L2[i] + L3[i] - L4[i]

        L[i]    <-      exp(log(L5[i]))
    }

# Priors
p1~dbeta(1,1)
p2~dbeta(1,1)
p3~dbeta(1,1)
p12~dbeta(1,1)
p13~dbeta(1,1)
p23~dbeta(1,1)
p123~dbeta(1,1)
}

## MVG Distribution - 2-Components Serie System

model.jags.mvg2 <-  function()
{
    for (i in 1:n)
    {
        phi[i] <- -log(L[i])
        zeros[i]~dpois(phi[i])

        z1[i]   <- max(t1[i] - 1, t2[i] - 1)
        z2[i]   <- max(t1[i], t2[i] - 1)
        z3[i]   <- max(t1[i] - 1, t2[i])
        z4[i]   <- max(t1[i], t2[i])

        A[i]    <- pow(theta1, t1[i] - 1)   * pow(theta2, t2[i] - 1) * pow(
            theta12, z1[i])
        B[i]    <- pow(theta1, t1[i])   * pow(theta2, t2[i] - 1) * pow(
            theta12, z2[i])
        C[i]    <- pow(theta1, t1[i] - 1)   * pow(theta2, t2[i])     * pow(
            theta12, z3[i])
        D[i]    <- pow(theta1, t1[i])   * pow(theta2, t2[i])     * pow(
            theta12, z4[i])
```

```
        E[i]    <- log(A[i] - B[i] - C[i] + D[i])

        L[i]    <- exp(E[i])

        t[i]    <- min(t1[i], t2[i])
        Sys[i]  <- theta1^t[i] * theta2^t[i] * theta12^t[i]
    }
# Priors
theta1~dbeta(1,1)
theta2~dbeta(1,1)
theta12~dbeta(1,1)
}

## MVG Distribution - 3-Components Serie System

model.jags.mvg3 <-  function()
{
    for (i in 1:n)
    {
        phi[i]  <- -log(L[i])
        zeros[i]~dpois(phi[i])

        # Min between ti's

        b1[i]   <-  max(t1[i], t2[i])
        b2[i]   <-  max(t1[i], t3[i])
        b3[i]   <-  max(t2[i], t3[i])
        b4[i]   <-  max(t1[i] - 1, t2[i])
        b5[i]   <-  max(t1[i] - 1, t3[i])
        b6[i]   <-  max(t2[i] - 1, t3[i])
        b7[i]   <-  max(t1[i], t2[i] - 1)
        b8[i]   <-  max(t1[i], t3[i] - 1)
        b9[i]   <-  max(t2[i], t3[i] - 1)
        b10[i]  <-  max(t1[i] - 1, t2[i] - 1)
        b11[i]  <-  max(t1[i] - 1, t3[i] - 1)
        b12[i]  <-  max(t2[i] - 1, t3[i] - 1)
        b13[i]  <-  max(t1[i], max(t2[i], t3[i]))
        b14[i]  <-  max(t1[i] - 1, max(t2[i], t3[i]))
        b15[i]  <-  max(t1[i] - 1, max(t2[i] - 1, t3[i]))
        b16[i]  <-  max(t1[i] - 1, max(t2[i], t3[i] - 1))
        b17[i]  <-  max(t1[i] - 1, max(t2[i] - 1, t3[i] - 1))
        b18[i]  <-  max(t1[i], max(t2[i] - 1, t3[i]))
        b19[i]  <-  max(t1[i], max(t2[i] - 1, t3[i] - 1))
        b20[i]  <-  max(t1[i], max(t2[i], t3[i] - 1))

        # Likelihood Auxiliary Parts
```

```
P1[i]   <-  theta1^(t1[i] - 1)  * theta2^(t2[i] - 1) * theta3^(t3[i
    ] - 1)
P2[i]   <-  theta1^(t1[i] - 1)  * theta2^(t2[i] - 1) * theta3^(t3[i
    ])
P3[i]   <-  theta1^(t1[i] - 1)  * theta2^(t2[i])     * theta3^(t3[i
    ] - 1)
P4[i]   <-  theta1^(t1[i])      * theta2^(t2[i] - 1) * theta3^(t3[i
    ] - 1)
P5[i]   <-  theta1^(t1[i] - 1)  * theta2^(t2[i])     * theta3^(t3[i
    ])
P6[i]   <-  theta1^(t1[i])      * theta2^(t2[i] - 1) * theta3^(t3[i
    ])
P7[i]   <-  theta1^(t1[i])      * theta2^(t2[i])     * theta3^(t3[i
    ] - 1)
P8[i]   <-  theta1^(t1[i])      * theta2^(t2[i])     * theta3^(t3[i
    ])

Q1[i]   <-  theta12^(b10[i]) * theta13^(b11[i]) * theta23^(b12[i])
    * theta123^(b17[i])
Q2[i]   <-  theta12^(b10[i]) * theta13^(b5[i])  * theta23^(b6[i])
    * theta123^(b15[i])
Q3[i]   <-  theta12^(b4[i])  * theta13^(b11[i]) * theta23^(b9[i])
    * theta123^(b16[i])
Q4[i]   <-  theta12^(b7[i])  * theta13^(b8[i])  * theta23^(b12[i])
    * theta123^(b19[i])
Q5[i]   <-  theta12^(b4[i])  * theta13^(b5[i])  * theta23^(b3[i])
    * theta123^(b14[i])
Q6[i]   <-  theta12^(b7[i])  * theta13^(b2[i])  * theta23^(b6[i])
    * theta123^(b18[i])
Q7[i]   <-  theta12^(b1[i])  * theta13^(b8[i])  * theta23^(b9[i])
    * theta123^(b20[i])
Q8[i]   <-  theta12^(b1[i])  * theta13^(b2[i])  * theta23^(b3[i])
    * theta123^(b13[i])

K1[i]   <-  P1[i] * Q1[i]
K2[i]   <-  P2[i] * Q2[i]
K3[i]   <-  P3[i] * Q3[i]
K4[i]   <-  P4[i] * Q4[i]
K5[i]   <-  P5[i] * Q5[i]
K6[i]   <-  P6[i] * Q6[i]
K7[i]   <-  P7[i] * Q7[i]
K8[i]   <-  P8[i] * Q8[i]

L1[i]   <-  K1[i]
L2[i]   <-  K2[i] + K3[i] + K4[i]
L3[i]   <-  K5[i] + K6[i] + K7[i]
L4[i]   <-  K8[i]
```

```
        L5[i]   <-  abs(L1[i] - L2[i] + L3[i] - L4[i])

        # Likelihood Final

        L[i]    <-  exp(log(L5[i]))

        # Series System

        t[i]    <-  min(t1[i], t2[i], t3[i])
        Sys[i]  <-  (theta1 * theta2 * theta3 * theta12 * theta13 * theta23
            * theta123)^t[i]
    }

# Priors
theta1~dbeta(1,1)
theta2~dbeta(1,1)
theta3~dbeta(1,1)
theta12~dbeta(1,1)
theta13~dbeta(1,1)
theta23~dbeta(1,1)
theta123~dbeta(1,1)
}

## MVG Distribution - 4-Components Serie System

model.jags.mvg4 <-  function()
{
    for (i in 1:n)
    {
        phi[i]  <-  -log(L[i])
        zeros[i]~dpois(phi[i])

        # Max between ti's

        b1.1[i]     <-  max(t1[i] - 1, t2[i] - 1)
        b2.1[i]     <-  max(t1[i] - 1, t3[i] - 1)
        b3.1[i]     <-  max(t1[i] - 1, t4[i] - 1)
        b4.1[i]     <-  max(t2[i] - 1, t3[i] - 1)
        b5.1[i]     <-  max(t2[i] - 1, t4[i] - 1)
        b6.1[i]     <-  max(t3[i] - 1, t4[i] - 1)
        b7.1[i]     <-  max(t1[i] - 1, t2[i] - 1, t3[i] - 1)
        b8.1[i]     <-  max(t1[i] - 1, t2[i] - 1, t4[i] - 1)
        b9.1[i]     <-  max(t1[i] - 1, t3[i] - 1, t4[i] - 1)
        b10.1[i]    <-  max(t2[i] - 1, t3[i] - 1, t4[i] - 1)
        b11.1[i]    <-  max(t1[i] - 1, t2[i] - 1, t3[i] - 1, t4[i] - 1)

        b1.2[i]     <-  max(t1[i], t2[i] - 1)
```

```
b2.2[i]    <- max(t1[i], t3[i] - 1)
b3.2[i]    <- max(t1[i], t4[i] - 1)
b4.2[i]    <- max(t2[i] - 1, t3[i] - 1)
b5.2[i]    <- max(t2[i] - 1, t4[i] - 1)
b6.2[i]    <- max(t3[i] - 1, t4[i] - 1)
b7.2[i]    <- max(t1[i], t2[i] - 1, t3[i] - 1)
b8.2[i]    <- max(t1[i], t2[i] - 1, t4[i] - 1)
b9.2[i]    <- max(t1[i], t3[i] - 1, t4[i] - 1)
b10.2[i]   <- max(t2[i] - 1, t3[i] - 1, t4[i] - 1)
b11.2[i]   <- max(t1[i], t2[i] - 1, t3[i] - 1, t4[i] - 1)

b1.3[i]    <- max(t1[i] - 1, t2[i])
b2.3[i]    <- max(t1[i] - 1, t3[i] - 1)
b3.3[i]    <- max(t1[i] - 1, t4[i] - 1)
b4.3[i]    <- max(t2[i], t3[i] - 1)
b5.3[i]    <- max(t2[i], t4[i] - 1)
b6.3[i]    <- max(t3[i] - 1, t4[i] - 1)
b7.3[i]    <- max(t1[i] - 1, t2[i], t3[i] - 1)
b8.3[i]    <- max(t1[i] - 1, t2[i], t4[i] - 1)
b9.3[i]    <- max(t1[i] - 1, t3[i] - 1, t4[i] - 1)
b10.3[i]   <- max(t2[i], t3[i] - 1, t4[i] - 1)
b11.3[i]   <- max(t1[i] - 1, t2[i], t3[i] - 1, t4[i] - 1)

b1.4[i]    <- max(t1[i] - 1, t2[i] - 1)
b2.4[i]    <- max(t1[i] - 1, t3[i])
b3.4[i]    <- max(t1[i] - 1, t4[i] - 1)
b4.4[i]    <- max(t2[i] - 1, t3[i])
b5.4[i]    <- max(t2[i] - 1, t4[i] - 1)
b6.4[i]    <- max(t3[i], t4[i] - 1)
b7.4[i]    <- max(t1[i] - 1, t2[i] - 1, t3[i])
b8.4[i]    <- max(t1[i] - 1, t2[i] - 1, t4[i] - 1)
b9.4[i]    <- max(t1[i] - 1, t3[i], t4[i] - 1)
b10.4[i]   <- max(t2[i] - 1, t3[i], t4[i] - 1)
b11.4[i]   <- max(t1[i] - 1, t2[i] - 1, t3[i], t4[i] - 1)

b1.5[i]    <- max(t1[i] - 1, t2[i] - 1)
b2.5[i]    <- max(t1[i] - 1, t3[i] - 1)
b3.5[i]    <- max(t1[i] - 1, t4[i])
b4.5[i]    <- max(t2[i] - 1, t3[i] - 1)
b5.5[i]    <- max(t2[i] - 1, t4[i])
b6.5[i]    <- max(t3[i] - 1, t4[i])
b7.5[i]    <- max(t1[i] - 1, t2[i] - 1, t3[i] - 1)
b8.5[i]    <- max(t1[i] - 1, t2[i] - 1, t4[i])
b9.5[i]    <- max(t1[i] - 1, t3[i] - 1, t4[i])
b10.5[i]   <- max(t2[i] - 1, t3[i] - 1, t4[i])
b11.5[i]   <- max(t1[i] - 1, t2[i] - 1, t3[i] - 1, t4[i])
```

```
b1.6[i]    <- max(t1[i], t2[i])
b2.6[i]    <- max(t1[i], t3[i] - 1)
b3.6[i]    <- max(t1[i], t4[i] - 1)
b4.6[i]    <- max(t2[i], t3[i] - 1)
b5.6[i]    <- max(t2[i], t4[i] - 1)
b6.6[i]    <- max(t3[i] - 1, t4[i] - 1)
b7.6[i]    <- max(t1[i], t2[i], t3[i] - 1)
b8.6[i]    <- max(t1[i], t2[i], t4[i] - 1)
b9.6[i]    <- max(t1[i], t3[i] - 1, t4[i] - 1)
b10.6[i]   <- max(t2[i], t3[i] - 1, t4[i] - 1)
b11.6[i]   <- max(t1[i], t2[i], t3[i] - 1, t4[i] - 1)

b1.7[i]    <- max(t1[i], t2[i] - 1)
b2.7[i]    <- max(t1[i], t3[i])
b3.7[i]    <- max(t1[i], t4[i] - 1)
b4.7[i]    <- max(t2[i] - 1, t3[i])
b5.7[i]    <- max(t2[i] - 1, t4[i] - 1)
b6.7[i]    <- max(t3[i], t4[i] - 1)
b7.7[i]    <- max(t1[i], t2[i] - 1, t3[i])
b8.7[i]    <- max(t1[i], t2[i] - 1, t4[i] - 1)
b9.7[i]    <- max(t1[i], t3[i], t4[i] - 1)
b10.7[i]   <- max(t2[i] - 1, t3[i], t4[i] - 1)
b11.7[i]   <- max(t1[i], t2[i] - 1, t3[i], t4[i] - 1)

b1.8[i]    <- max(t1[i], t2[i] - 1)
b2.8[i]    <- max(t1[i], t3[i] - 1)
b3.8[i]    <- max(t1[i], t4[i])
b4.8[i]    <- max(t2[i] - 1, t3[i] - 1)
b5.8[i]    <- max(t2[i] - 1, t4[i])
b6.8[i]    <- max(t3[i] - 1, t4[i])
b7.8[i]    <- max(t1[i], t2[i] - 1, t3[i] - 1)
b8.8[i]    <- max(t1[i], t2[i] - 1, t4[i])
b9.8[i]    <- max(t1[i], t3[i] - 1, t4[i])
b10.8[i]   <- max(t2[i] - 1, t3[i] - 1, t4[i])
b11.8[i]   <- max(t1[i], t2[i] - 1, t3[i] - 1, t4[i])

b1.9[i]    <- max(t1[i] - 1, t2[i])
b2.9[i]    <- max(t1[i] - 1, t3[i])
b3.9[i]    <- max(t1[i] - 1, t4[i] - 1)
b4.9[i]    <- max(t2[i], t3[i])
b5.9[i]    <- max(t2[i], t4[i] - 1)
b6.9[i]    <- max(t3[i], t4[i] - 1)
b7.9[i]    <- max(t1[i] - 1, t2[i], t3[i])
b8.9[i]    <- max(t1[i] - 1, t2[i], t4[i] - 1)
b9.9[i]    <- max(t1[i] - 1, t3[i], t4[i] - 1)
b10.9[i]   <- max(t2[i], t3[i], t4[i] - 1)
b11.9[i]   <- max(t1[i] - 1, t2[i], t3[i], t4[i] - 1)
```

```
b1.10[i]    <-  max(t1[i] - 1, t2[i])
b2.10[i]    <-  max(t1[i] - 1, t3[i] - 1)
b3.10[i]    <-  max(t1[i] - 1, t4[i])
b4.10[i]    <-  max(t2[i], t3[i] - 1)
b5.10[i]    <-  max(t2[i], t4[i])
b6.10[i]    <-  max(t3[i] - 1, t4[i])
b7.10[i]    <-  max(t1[i] - 1, t2[i], t3[i] - 1)
b8.10[i]    <-  max(t1[i] - 1, t2[i], t4[i])
b9.10[i]    <-  max(t1[i] - 1, t3[i] - 1, t4[i])
b10.10[i]   <-  max(t2[i], t3[i] - 1, t4[i])
b11.10[i]   <-  max(t1[i] - 1, t2[i], t3[i] - 1, t4[i])

b1.11[i]    <-  max(t1[i] - 1, t2[i] - 1)
b2.11[i]    <-  max(t1[i] - 1, t3[i])
b3.11[i]    <-  max(t1[i] - 1, t4[i])
b4.11[i]    <-  max(t2[i] - 1, t3[i])
b5.11[i]    <-  max(t2[i] - 1, t4[i])
b6.11[i]    <-  max(t3[i], t4[i])
b7.11[i]    <-  max(t1[i] - 1, t2[i] - 1, t3[i])
b8.11[i]    <-  max(t1[i] - 1, t2[i] - 1, t4[i])
b9.11[i]    <-  max(t1[i] - 1, t3[i], t4[i])
b10.11[i]   <-  max(t2[i] - 1, t3[i], t4[i])
b11.11[i]   <-  max(t1[i] - 1, t2[i] - 1, t3[i], t4[i])

b1.12[i]    <-  max(t1[i], t2[i])
b2.12[i]    <-  max(t1[i], t3[i])
b3.12[i]    <-  max(t1[i], t4[i] - 1)
b4.12[i]    <-  max(t2[i], t3[i])
b5.12[i]    <-  max(t2[i], t4[i] - 1)
b6.12[i]    <-  max(t3[i], t4[i] - 1)
b7.12[i]    <-  max(t1[i], t2[i], t3[i])
b8.12[i]    <-  max(t1[i], t2[i], t4[i] - 1)
b9.12[i]    <-  max(t1[i], t3[i], t4[i] - 1)
b10.12[i]   <-  max(t2[i], t3[i], t4[i] - 1)
b11.12[i]   <-  max(t1[i], t2[i], t3[i], t4[i] - 1)

b1.13[i]    <-  max(t1[i], t2[i])
b2.13[i]    <-  max(t1[i], t3[i] - 1)
b3.13[i]    <-  max(t1[i], t4[i])
b4.13[i]    <-  max(t2[i], t3[i] - 1)
b5.13[i]    <-  max(t2[i], t4[i])
b6.13[i]    <-  max(t3[i] - 1, t4[i])
b7.13[i]    <-  max(t1[i], t2[i], t3[i] - 1)
b8.13[i]    <-  max(t1[i], t2[i], t4[i])
b9.13[i]    <-  max(t1[i], t3[i] - 1, t4[i])
b10.13[i]   <-  max(t2[i], t3[i] - 1, t4[i])
```

```
b11.13[i]   <-  max(t1[i], t2[i], t3[i] - 1, t4[i])

b1.14[i]    <-  max(t1[i], t2[i] - 1)
b2.14[i]    <-  max(t1[i], t3[i])
b3.14[i]    <-  max(t1[i], t4[i])
b4.14[i]    <-  max(t2[i] - 1, t3[i])
b5.14[i]    <-  max(t2[i] - 1, t4[i])
b6.14[i]    <-  max(t3[i], t4[i])
b7.14[i]    <-  max(t1[i], t2[i] - 1, t3[i])
b8.14[i]    <-  max(t1[i], t2[i] - 1, t4[i])
b9.14[i]    <-  max(t1[i], t3[i], t4[i])
b10.14[i]   <-  max(t2[i] - 1, t3[i], t4[i])
b11.14[i]   <-  max(t1[i], t2[i] - 1, t3[i], t4[i])

b1.15[i]    <-  max(t1[i] - 1, t2[i])
b2.15[i]    <-  max(t1[i] - 1, t3[i])
b3.15[i]    <-  max(t1[i] - 1, t4[i])
b4.15[i]    <-  max(t2[i], t3[i])
b5.15[i]    <-  max(t2[i], t4[i])
b6.15[i]    <-  max(t3[i], t4[i])
b7.15[i]    <-  max(t1[i] - 1, t2[i], t3[i])
b8.15[i]    <-  max(t1[i] - 1, t2[i], t4[i])
b9.15[i]    <-  max(t1[i] - 1, t3[i], t4[i])
b10.15[i]   <-  max(t2[i], t3[i], t4[i])
b11.15[i]   <-  max(t1[i] - 1, t2[i], t3[i], t4[i])

b1.16[i]    <-  max(t1[i], t2[i])
b2.16[i]    <-  max(t1[i], t3[i])
b3.16[i]    <-  max(t1[i], t4[i])
b4.16[i]    <-  max(t2[i], t3[i])
b5.16[i]    <-  max(t2[i], t4[i])
b6.16[i]    <-  max(t3[i], t4[i])
b7.16[i]    <-  max(t1[i], t2[i], t3[i])
b8.16[i]    <-  max(t1[i], t2[i], t4[i])
b9.16[i]    <-  max(t1[i], t3[i], t4[i])
b10.16[i]   <-  max(t2[i], t3[i], t4[i])
b11.16[i]   <-  max(t1[i], t2[i], t3[i], t4[i])

# Likelihood Auxiliary Parts (Part 1)

P1[i]   <-  theta1^(t1[i] - 1)  * theta2^(t2[i] - 1) * theta3^(t3[i
    ] - 1) * theta4^(t4[i] - 1)
P2[i]   <-  theta1^(t1[i])      * theta2^(t2[i] - 1) * theta3^(t3[i
    ] - 1) * theta4^(t4[i] - 1)
P3[i]   <-  theta1^(t1[i] - 1)  * theta2^(t2[i])     * theta3^(t3[i
    ] - 1) * theta4^(t4[i] - 1)
P4[i]   <-  theta1^(t1[i] - 1)  * theta2^(t2[i] - 1) * theta3^(t3[i
```

```
    ])     * theta4^(t4[i] - 1)
P5[i]   <- theta1^(t1[i] - 1)  * theta2^(t2[i] - 1) * theta3^(t3[i
    ] - 1) * theta4^(t4[i])
P6[i]   <- theta1^(t1[i])      * theta2^(t2[i])     * theta3^(t3[i
    ] - 1) * theta4^(t4[i] - 1)
P7[i]   <- theta1^(t1[i])      * theta2^(t2[i] - 1) * theta3^(t3[i
    ])     * theta4^(t4[i] - 1)
P8[i]   <- theta1^(t1[i])      * theta2^(t2[i] - 1) * theta3^(t3[i
    ] - 1) * theta4^(t4[i])
P9[i]   <- theta1^(t1[i] - 1)  * theta2^(t2[i])     * theta3^(t3[i
    ])     * theta4^(t4[i] - 1)
P10[i]  <- theta1^(t1[i] - 1)  * theta2^(t2[i])     * theta3^(t3[i
    ] - 1) * theta4^(t4[i])
P11[i]  <- theta1^(t1[i] - 1)  * theta2^(t2[i] - 1) * theta3^(t3[i
    ])     * theta4^(t4[i])
P12[i]  <- theta1^(t1[i])      * theta2^(t2[i])     * theta3^(t3[i
    ])     * theta4^(t4[i] - 1)
P13[i]  <- theta1^(t1[i])      * theta2^(t2[i])     * theta3^(t3[i
    ] - 1) * theta4^(t4[i])
P14[i]  <- theta1^(t1[i])      * theta2^(t2[i] - 1) * theta3^(t3[i
    ])     * theta4^(t4[i])
P15[i]  <- theta1^(t1[i] - 1)  * theta2^(t2[i])     * theta3^(t3[i
    ])     * theta4^(t4[i])
P16[i]  <- theta1^(t1[i])      * theta2^(t2[i])     * theta3^(t3[i
    ])     * theta4^(t4[i])

# Likelihood Auxiliary Parts (Part 2)

Q1[i]   <- theta12^(b1.1[i]) * theta13^(b2.1[i]) * theta14^(b3.1[i
    ]) * theta23^(b4.1[i]) * theta24^(b5.1[i]) * theta34^(b6.1[i])
Q2[i]   <- theta12^(b1.2[i]) * theta13^(b2.2[i]) * theta14^(b3.2[i
    ]) * theta23^(b4.2[i]) * theta24^(b5.2[i]) * theta34^(b6.2[i])
Q3[i]   <- theta12^(b1.3[i]) * theta13^(b2.3[i]) * theta14^(b3.3[i
    ]) * theta23^(b4.3[i]) * theta24^(b5.3[i]) * theta34^(b6.3[i])
Q4[i]   <- theta12^(b1.4[i]) * theta13^(b2.4[i]) * theta14^(b3.4[i
    ]) * theta23^(b4.4[i]) * theta24^(b5.4[i]) * theta34^(b6.4[i])
Q5[i]   <- theta12^(b1.5[i]) * theta13^(b2.5[i]) * theta14^(b3.5[i
    ]) * theta23^(b4.5[i]) * theta24^(b5.5[i]) * theta34^(b6.5[i])
Q6[i]   <- theta12^(b1.6[i]) * theta13^(b2.6[i]) * theta14^(b3.6[i
    ]) * theta23^(b4.6[i]) * theta24^(b5.6[i]) * theta34^(b6.6[i])
Q7[i]   <- theta12^(b1.7[i]) * theta13^(b2.7[i]) * theta14^(b3.7[i
    ]) * theta23^(b4.7[i]) * theta24^(b5.7[i]) * theta34^(b6.7[i])
Q8[i]   <- theta12^(b1.8[i]) * theta13^(b2.8[i]) * theta14^(b3.8[i
    ]) * theta23^(b4.8[i]) * theta24^(b5.8[i]) * theta34^(b6.8[i])
Q9[i]   <- theta12^(b1.9[i]) * theta13^(b2.9[i]) * theta14^(b3.9[i
    ]) * theta23^(b4.9[i]) * theta24^(b5.9[i]) * theta34^(b6.9[i])
Q10[i]  <- theta12^(b1.10[i]) * theta13^(b2.10[i]) * theta14^(b3
```

```
    .10[i]) * theta23^(b4.10[i]) * theta24^(b5.10[i]) * theta34^(b6
    .10[i])
Q11[i] <- theta12^(b1.11[i]) * theta13^(b2.11[i]) * theta14^(b3
    .11[i]) * theta23^(b4.11[i]) * theta24^(b5.11[i]) * theta34^(b6
    .11[i])
Q12[i] <- theta12^(b1.12[i]) * theta13^(b2.12[i]) * theta14^(b3
    .12[i]) * theta23^(b4.12[i]) * theta24^(b5.12[i]) * theta34^(b6
    .12[i])
Q13[i] <- theta12^(b1.13[i]) * theta13^(b2.13[i]) * theta14^(b3
    .13[i]) * theta23^(b4.13[i]) * theta24^(b5.13[i]) * theta34^(b6
    .13[i])
Q14[i] <- theta12^(b1.14[i]) * theta13^(b2.14[i]) * theta14^(b3
    .14[i]) * theta23^(b4.14[i]) * theta24^(b5.14[i]) * theta34^(b6
    .14[i])
Q15[i] <- theta12^(b1.15[i]) * theta13^(b2.15[i]) * theta14^(b3
    .15[i]) * theta23^(b4.15[i]) * theta24^(b5.15[i]) * theta34^(b6
    .15[i])
Q16[i] <- theta12^(b1.16[i]) * theta13^(b2.16[i]) * theta14^(b3
    .16[i]) * theta23^(b4.16[i]) * theta24^(b5.16[i]) * theta34^(b6
    .16[i])

# Likelihood Auxiliary Parts (Part 3)

R1[i] <- theta123^(b7.1[i]) * theta124^(b8.1[i]) * theta134^(b9
    .1[i]) * theta234^(b10.1[i]) * theta1234^(b11.1[i])
R2[i] <- theta123^(b7.2[i]) * theta124^(b8.2[i]) * theta134^(b9
    .2[i]) * theta234^(b10.2[i]) * theta1234^(b11.2[i])
R3[i] <- theta123^(b7.3[i]) * theta124^(b8.3[i]) * theta134^(b9
    .3[i]) * theta234^(b10.3[i]) * theta1234^(b11.3[i])
R4[i] <- theta123^(b7.4[i]) * theta124^(b8.4[i]) * theta134^(b9
    .4[i]) * theta234^(b10.4[i]) * theta1234^(b11.4[i])
R5[i] <- theta123^(b7.5[i]) * theta124^(b8.5[i]) * theta134^(b9
    .5[i]) * theta234^(b10.5[i]) * theta1234^(b11.5[i])
R6[i] <- theta123^(b7.6[i]) * theta124^(b8.6[i]) * theta134^(b9
    .6[i]) * theta234^(b10.6[i]) * theta1234^(b11.6[i])
R7[i] <- theta123^(b7.7[i]) * theta124^(b8.7[i]) * theta134^(b9
    .7[i]) * theta234^(b10.7[i]) * theta1234^(b11.7[i])
R8[i] <- theta123^(b7.8[i]) * theta124^(b8.8[i]) * theta134^(b9
    .8[i]) * theta234^(b10.8[i]) * theta1234^(b11.8[i])
R9[i] <- theta123^(b7.9[i]) * theta124^(b8.9[i]) * theta134^(b9
    .9[i]) * theta234^(b10.9[i]) * theta1234^(b11.9[i])
R10[i] <- theta123^(b7.10[i]) * theta124^(b8.10[i]) * theta134^(
    b9.10[i]) * theta234^(b10.10[i]) * theta1234^(b11.10[i])
R11[i] <- theta123^(b7.11[i]) * theta124^(b8.11[i]) * theta134^(
    b9.11[i]) * theta234^(b10.11[i]) * theta1234^(b11.11[i])
R12[i] <- theta123^(b7.12[i]) * theta124^(b8.12[i]) * theta134^(
    b9.12[i]) * theta234^(b10.12[i]) * theta1234^(b11.12[i])
```

```
    R13[i] <- theta123^(b7.13[i]) * theta124^(b8.13[i]) * theta134^(
        b9.13[i]) * theta234^(b10.13[i]) * theta1234^(b11.13[i])
    R14[i] <- theta123^(b7.14[i]) * theta124^(b8.14[i]) * theta134^(
        b9.14[i]) * theta234^(b10.14[i]) * theta1234^(b11.14[i])
    R15[i] <- theta123^(b7.15[i]) * theta124^(b8.15[i]) * theta134^(
        b9.15[i]) * theta234^(b10.15[i]) * theta1234^(b11.15[i])
    R16[i] <- theta123^(b7.16[i]) * theta124^(b8.16[i]) * theta134^(
        b9.16[i]) * theta234^(b10.16[i]) * theta1234^(b11.16[i])

    # Likelihood Auxiliary Parts (Part 4)

    K1[i]  <- P1[i]  * Q1[i]  * R1[i]
    K2[i]  <- P2[i]  * Q2[i]  * R2[i]
    K3[i]  <- P3[i]  * Q3[i]  * R3[i]
    K4[i]  <- P4[i]  * Q4[i]  * R4[i]
    K5[i]  <- P5[i]  * Q5[i]  * R5[i]
    K6[i]  <- P6[i]  * Q6[i]  * R6[i]
    K7[i]  <- P7[i]  * Q7[i]  * R7[i]
    K8[i]  <- P8[i]  * Q8[i]  * R8[i]
    K9[i]  <- P9[i]  * Q9[i]  * R9[i]
    K10[i] <- P10[i] * Q10[i] * R10[i]
    K11[i] <- P11[i] * Q11[i] * R11[i]
    K12[i] <- P12[i] * Q12[i] * R12[i]
    K13[i] <- P13[i] * Q13[i] * R13[i]
    K14[i] <- P14[i] * Q14[i] * R14[i]
    K15[i] <- P15[i] * Q15[i] * R15[i]
    K16[i] <- P16[i] * Q16[i] * R16[i]

    L1[i]  <- K1[i]
    L2[i]  <- K2[i]  + K3[i]  + K4[i]  + K5[i]
    L3[i]  <- K6[i]  + K7[i]  + K8[i]  + K9[i] + K10[i] + K11[i]
    L4[i]  <- K12[i] + K13[i] + K14[i] + K15[i]
    L5[i]  <- K16[i]
    L6[i]  <- abs(L1[i] - L2[i] + L3[i] - L4[i] + L5[i])

    # Likelihood Final

    L[i]   <- exp(log(L6[i]))

    # Series System

    t[i]   <- min(t1[i], t2[i], t3[i], t4[i])
    Sys[i] <- (theta1 * theta2 * theta3 * theta4 * theta12 * theta13
        * theta14 * theta23 * theta24 * theta34 * theta123 * theta124 *
        theta134 * theta234 * theta1234)^t[i]
}
```

```
# Priors
theta1~dbeta(1,1)
theta2~dbeta(1,1)
theta3~dbeta(1,1)
theta4~dbeta(1,1)
theta12~dbeta(1,1)
theta13~dbeta(1,1)
theta14~dbeta(1,1)
theta23~dbeta(1,1)
theta24~dbeta(1,1)
theta34~dbeta(1,1)
theta123~dbeta(1,1)
theta124~dbeta(1,1)
theta134~dbeta(1,1)
theta234~dbeta(1,1)
theta1234~dbeta(1,1)
}
```

Appendix B. Articles and R Package

Published Articles:

- de Oliveira, R. P., & Achcar, J. A. (2018). Basu-Dhar's bivariate geometric distribution in presence of censored data and covariates: some computational aspects. Electronic Journal of Applied Statistical Analysis, 11(1), 108-136.
- de Oliveira, R. P., Achcar, J. A., Peralta, D., & Mazucheli, J. (2018). Discrete and continuous bivariate lifetime models in presence of cure rate: a comparative study under Bayesian approach. Journal of Applied Statistics, 1-19.
- de Oliveira, R. P., & Achcar, J. A. (2019). Use of Basu-Dhar bivariate geometric distribution in the analysis of the reliability of component series systems. International Journal of Quality & Reliability Management, 36(4), 569-586.
- de Oliveira, R. P., Menezes, A. F., Mazucheli, J., & Achcar, J. A. (2019). Mixture and nonmixture cure fraction models assuming discrete lifetimes: Application to a pelvic sarcoma dataset. Biometrical Journal.

Published R Package:

- de Oliveira, R. P., & Achcar, J. A. (2019). BivGeo: Basu-Dhar Bivariate Geometric Distribution. R package version 2.0.1. URL: https://CRAN.R-project.org/package=BivGeo.

Accepted Articles:

- de Oliveira, R. P., Mazucheli, J., & Achcar, J. A. (2019). A generalization of Basu-Dhar's bivariate geometric distribution to the trivariate case. Communications in Statistics: Simulation and Computation. Accepted in 10 July 2019.

Electronic Journal of Applied Statistical Analysis
Vol. 11, Issue 01, April 2018, 108-136
DOI: 10.1285/i20705948v11n1p108

Basu-Dhar's bivariate geometric distribution in presence of censored data and covariates: some computational aspects

Ricardo Puziol de Oliveira*and Jorge Alberto Achcar

University of São Paulo, Department of Social Medicine, Medical School
Ribeirão Preto, S.P., Brazil

Published: 26 April 2018

Some computational aspects to obtain classical and Bayesian inferences for the Basu and Dhar (1995) bivariate geometric distribution in presence of censored data and covariates are discussed in this paper. The posterior summaries of interest are obtained using standard existing MCMC (Markov Chain Monte Carlo) simulation methods available in popular free softwares as the OpenBugs software and the R software. Numerical illustrations are introduced considering simulated and real datasets showing that the use of discrete bivariate distributions may be a good alternative to the use of continuous bivariate distributions, in many areas of application.

keywords: Basu-Dhar distribution, Bayesian estimates, censored data, covariates, maximum likelihood estimates.

*Corresponding author: rpuziol.oliveira@gmail.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

JOURNAL OF APPLIED STATISTICS
https://doi.org/10.1080/02664763.2018.1495701

Taylor & Francis
Taylor & Francis Group

Discrete and continuous bivariate lifetime models in presence of cure rate: a comparative study under Bayesian approach

Ricardo Puziol de Oliveira [a], Jorge Alberto Achcar [a], Danielle Peralta [b] and Josmar Mazucheli [b]

[a]Medical School, University of São Paulo, Ribeirão Preto, Brazil; [b]Department of Statistics, Maringá State University, Brazil

ABSTRACT

The modeling and analysis of lifetime data in which the main end-points are the times when an event of interest occurs is of great interest in medical studies. In these studies, it is common that two or more lifetimes associated with the same unit such as the times to deterioration levels or the times to reaction to a treatment in pairs of organs like lungs, kidneys, eyes or ears. In medical applications, it is also possible that a cure rate is present and needed to be modeled with lifetime data with long-term survivors. This paper presented a comparative study under a Bayesian approach among some existing continuous and discrete bivariate distributions such as the bivariate exponential distributions and the bivariate geometric distributions in presence of cure rate, censored data and covariates. In presence of lifetimes related to cured patients, it is assumed standard mixture cure rate models in the data analysis. The posterior summaries of interest are obtained using Markov Chain Monte Carlo methods. To illustrate the proposed methodology two real medical data sets are considered.

ARTICLE HISTORY
Received 24 November 2017
Accepted 22 June 2018

KEYWORDS
Bivariate lifetimes; Bayesian analysis; cure rate models; diabetic retinopathy; medical studies; survival analysis; tobacco

1. Introduction

In many medical studies, especially related to cancer treatments, an important issue of interest to the medical researchers is the estimation of the fraction of individuals (or patients) in the studied population who never experience the event of interest. These individuals are not at risk with respect to the event of interest and they are considered immune, cured, non-susceptible or extremely long-term survivors. Different approaches have been presented in the literature to model cure rate, especially for univariate lifetime data (see, e.g. [2,9,13,22,23,25,26,35,36]). In the presence of two lifetimes associated to each unit, that is, bivariate lifetimes, Wienke *et al.* [33,34] introduced a model for a cure rate in bivariate time-to-event data analysis.

The cure rate models became popular since the standard survival analysis techniques, for example the Cox proportional hazards [11] model, provide no direct estimation for the

CONTACT Ricardo Puziol de Oliveira rpuziol.oliveira@gmail.com Medical School, University of São Paulo, Ribeirão Preto, Brazil

This article was originally published with errors. This version has been corrected. Please see Corrigendum (http://dx.doi.org/10.1080/02664763.2018.1505203).

Use of Basu-Dhar bivariate geometric distribution in the analysis of the reliability of component series systems

Ricardo Puziol Oliveira and Jorge Alberto Achcar
Faculty of Medicine of Ribeirao Preto, University of Sao Paulo, Ribeirao Preto, Brazil

Received 10 May 2018
Revised 10 May 2018
Accepted 26 July 2018

Abstract

Purpose – The purpose of this paper is to provide a new method to estimate the reliability of series system by using a discrete bivariate distribution. This problem is of great interest in industrial and engineering applications.

Design/methodology/approach – The authors considered the Basu–Dhar bivariate geometric distribution and a Bayesian approach with application to a simulated data set and an engineering data set.

Findings – From the obtained results of this study, the authors observe that the discrete Basu–Dhar bivariate probability distribution could be a good alternative in the analysis of series system structures with accurate inference results for the reliability of the system under a Bayesian approach.

Originality/value – System reliability studies usually assume independent lifetimes for the components (series, parallel or complex system structures) in the estimation of the reliability of the system. This assumption in general is not reasonable in many engineering applications, since it is possible that the presence of some dependence structure between the lifetimes of the components could affect the evaluation of the reliability of the system.

Keywords Engineering, Reliability analysis, Basu–Dhar, Bayesian analysis, Dependent lifetimes, Series systems

Paper type Research paper

1. Introduction

A series system is a component configuration usually assumed in engineering studies, such that, if any one of the system components fails, the entire system fails. Associated to each system component, there is a response given by a random variable that could be binary (fail/no fail) or denoting its lifetime (a positive value). Independence is usually assumed among these random variables, an assumption that is not always realistic in applications. Figure 1 illustrates a series system (see, e.g. Romeu, 2004).

According to Jensen and Bard (2003), the estimation of the reliability is obtained using an inference approach based on probabilistic models. The reliability could be given by the probability of fail (response: fail/no fail) or by the probability $P(T > t)$, where T denotes the lifetime of the component or system and t is a fixed value. Many different probability models are introduced in the literature to capture the dependence among the components (see, e.g. Henley and Kumamoto, 1981; Kolowrocki, 2008; Singh and Billinton, 1977; Chao and Fu, 1991; Hulting and Robinson, 1994; Rausand and Arnljot, 2004; Usher, 1996; Eryilmaz, 2011; Achcar and Moala, 2015; Mukherjee and Saran, 1984; Tang *et al.*, 2013; Eryilmaz and Tank, 2012; Aven and Jensen, 1999). Many physical and non-physical systems (e.g. bridges, car engines, air-conditioning systems, biological and ecological systems, chains of command in civilian or military organizations, quality control systems in manufacturing plants, among many others) may be viewed as assemblies of many interacting components (series systems, parallel systems, k-out-of m systems or complex systems).

The main goal of this paper is to explore the performance of the Basu–Dhar bivariate geometric distribution (see Basu and Dhar, 1995) in the estimation of the reliability of

International Journal of Quality & Reliability Management

0265-671X
DOI 10.1108/IJQRM-05-2018-0115

Received: 14 August 2018 | Revised: 7 January 2019 | Accepted: 4 February 2019

DOI: 10.1002/bimj.201800030

RESEARCH PAPER

Biometrical Journal

Mixture and nonmixture cure fraction models assuming discrete lifetimes: Application to a pelvic sarcoma dataset

Ricardo Puziol de Oliveira[1] | André F. B. Menezes[2] | Josmar Mazucheli[2] | Jorge A. Achcar[1]

[1]Medical School, Universidade de São Paulo, Ribeirão Preto, SP, Brasil

[2]Department of Statistics, Universidade Estadual de Maringá, Paraná, PR, Brasil

Correspondence
Ricardo Puziol de Oliveira, Medical School, Universidade de São Paulo, Ribeirão Preto, SP, Brasil.
Email: rpuziol.oliveira@gmail.com

Abstract

Different cure fraction models have been used in the analysis of lifetime data in presence of cured patients. This paper considers mixture and nonmixture models based on discrete Weibull distribution to model recurrent event data in presence of a cure fraction. The novelty of this study is the use of a discrete lifetime distribution in place of usual existing continuous lifetime distributions for lifetime data in presence of cured fraction, censored data, and covariates. In the verification of the fit of the proposed model it is proposed the use of randomized quantile residuals. An extensive simulation study is considered to evaluate the properties of the estimates of the parameters related to the proposed model. As an illustration of the proposed methodology, it is considered an application considering a medical dataset related to lifetimes in a retrospective cohort study conducted by Puchner et al. (2017) that consists of 147 consecutive cases with surgical treatment of a sarcoma of the pelvis between the years of 1980 and 2012.

KEYWORDS

cure fraction, discrete Weibull distribution, long-term survivors, pelvic sarcomas, survival analysis

1 | INTRODUCTION

In many medical studies, an issue of great interest in medical research is the estimation of the fraction of patients in the studied population who never experience the event of interest. These patients are not at risk with respect to the event of interest and are considered immune, cured, nonsusceptible, or extremely long-term survivors. Standard survival analysis techniques, as for example, the Cox proportional hazards (Cox, 1972) model, provide no direct estimation for the cure fraction that is a motivation for the use of mixture and nonmixture cure fraction models.

According to Vahidpour (2016), in the literature it is presented different models to be fitted by data in presence of cure fraction with great emphasis on the mixture cure fraction models, also known as standard cure fraction models (see, e.g., De Angelis, Capocaccia, Hakulinen, Soderman, & Verdecchia, 1999; Lambert, Thompson, Weston, & Dickman, 2006), which have been widely used for modeling survival data in presence of cure fraction and the nonmixture cure fraction models that are not very popular (see Achcar, Coelho-Barros, & Mazucheli, 2012; Tsodikov, Ibrahim, & Yakovlev, 2003; Vahidpour, 2016). Different approaches have been presented in the literature to model cure fraction for univariate lifetime data: (see, e.g., Achcar et al., 2012; Cancho & Bolfarine, 2001; De Angelis et al., 1999; Farewell, 1982; Lambert et al., 2006; Lu, 2010; Othus et al., 2012; Price & Manatunga, 2001; Yin & Ibrahim, 2005; Yu, Tiwari, Cronin, & Feuer, 2004, among many other studies).

Let us denote by T a positive random variable related to the the time until the event. Following Maller and Zhou (1996), the standard fraction model (or mixture cure fraction model) assumes that the probability of the time-to-event to be greater than a specified time t (the survival function) is given by

Package 'BivGeo'

January 3, 2019

Type Package

Title Basu-Dhar Bivariate Geometric Distribution

Version 2.0.1

Author Ricardo Puziol de Oliveira and Jorge Alberto Achcar

Maintainer Ricardo Puziol de Oliveira <rpuziol.oliveira@gmail.com>

Description Computes the joint probability mass function (pmf), the joint cumulative function (cdf), the joint survival function (sf), the correlation coefficient, the covariance, the cross-factorial moment and generate random deviates for the Basu-Dhar bivariate geometric distribution as well the joint probability mass, cumulative and survival function assuming the presence of a cure fraction given by the standard bivariate mixture cure fraction model. The package also computes the estimators based on the method of moments.

Depends R (>= 3.0.2)

Imports stats

URL For more details, see de Oliveira et. al (2018) <doi:10.1285/i20705948v11n1p108>.

RoxygenNote 6.1.1

Encoding UTF-8

NeedsCompilation no

LazyData TRUE

License GPL (>= 2)

Repository CRAN

Date/Publication 2019-01-03 15:10:16 UTC

R topics documented:

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2019.1643881

A generalization of Basu-Dhar's bivariate geometric distribution to the trivariate case

Q2 Ricardo Puziol de Oliveira[a], Josmar Mazucheli[b], and Jorge Alberto Achcar[a]

[a]Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil; [b]Department of Statistics, Maringá State University, Maringá, PR, Brazil

ABSTRACT

In this paper, it is introduced a new parametric distribution to be used in multivariate lifetime data analysis as an alternative for the use of some existing multivariate parametric models as the popular multivariate normal distribution the most widely used model assumed in the analysis of continuous multivariate data analysis. Although the normal multivariate distribution has univariate marginal normal probability distributions and simple interpretations for all their parameters, it may not be well fitted by many data sets, especially in survival data applications, usually considering logarithm transformed data. In addition, in many cases the use of parametric multivariate discrete models could be more appropriate for the data analysis. In this paper, it is introduced a generalization of the bivariate Basu-Dhar geometric distribution to a trivariate case applied to count data. Some properties of this trivariate geometric distribution, including their marginal probability distributions, the order statistics distributions, the probability generating function and some simulation studies are presented. It is also presented some discussion on an extension of the trivariate case for the multivariate case. Classical and Bayesian inferences are presented assuming censored or uncensored observations. To illustrate the proposed methodology, two applications with real lifetime data are considered as examples.

ARTICLE HISTORY
Received 29 March 2018
Accepted 10 July 2019

KEYWORDS
Basu-Dhar bivariate distribution; Bayesian estimation; Maximum likelihood estimation; Monte Carlo simulations; Right-censoring; Survival analysis; Trivariate geometric distribution

1. Introduction

In many applications of lifetime data analysis, usually there is the presence of two or more lifetimes associated to each unit, as for example in medical recurrent events. In these situations, especially in medical or engineering applications, the researchers usually assume independent lifetimes using standard parametric probability distributions like the exponential, Weibull or log-normal probability distributions or non-parametric set-ups like the product-limit Kaplan and Meier (1958) non-parametric estimator for the survival function or the proportional hazards regression model proposed by Cox (1972) in presence of censored data and covariates, but in general, the lifetime of one component can influence the lifetimes of the other components. In this way, it is needed statistical models which capture the dependence structure among the lifetimes associated to each unit where it is common the use of a continuous random variable modeling approach.

CONTACT Ricardo Puziol de Oliveira rpuziol.oliveira@gmail.com Medical School, University of São Paulo, Ribeirão Preto, SP, Brazil.

Q1 Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lssp.

Bibliography

M. A-hameed and F. Proschan. Nonstationary shock models. *Stochastic Processes and their Applications*, 1(4):383–404, 1973.

J. Achcar, N. Davarzani, and R. Souza. Basu–Dhar bivariate geometric distribution in the presence of covariates and censored data: a bayesian approach. *Journal of Applied Statistics*, 43(9):1636–1648, 2016a.

J. A. Achcar and R. A. Leandro. Use of Markov Chain Monte Carlo methods in a Bayesian analysis of the Block and Basu bivariate exponential distribution. *Annals of the Institute of Statistical Mathematics*, 50(3):403–416, 1998.

J. A. Achcar, E. A. Coelho-Barros, and J. Mazucheli. Cure fraction models using mixture and non-mixture models. *Tatra Mountains Mathematical Publications*, 51(1):1–9, 2012.

J. A. Achcar, E. Z. Martinez, and J. R. T. Cuevas. Bivariate lifetime modeling using copula functions in presence of mixture and non-mixture cure fraction models, censored data and covariates. *Model Assisted Statistics and Applications*, 11(4):261–276, 2016b.

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

H. Akaike. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer, 1998.

D. Al-Mutairi, M. Ghitany, and D. Kundu. Inferences on stress-strength reliability from lindley distributions. *Communications in Statistics-Theory and Methods*, 42(8):1443–1463, 2013.

S. J. Almalki and S. Nadarajah. Modifications of the Weibull distribution: A review. *Reliability Engineering & System Safety*, 124:32–55, 2014.

T. W. Anderson. Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *Journal of the american Statistical Association*, 52(278):200–203, 1957.

J. C. Anthony, L. A. Warner, and R. C. Kessler. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: basic findings from the National Comorbidity Survey. *Experimental and clinical psychopharmacology*, 2(3):244, 1994.

B. C. Arnold. A characterization of the exponential distribution by multivariate geometric compounding. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 164–173, 1975.

B. C. Arnold and D. Strauss. Bivariate distributions with exponential conditionals. *Journal of the American Statistical Association*, 83(402):522–527, 1988.

A. C. Atkinson. Two graphical displays for outlying and influential observations in regression. *Biometrika*, 68(1):13–20, 1981.

N. Balakrishnan and C. Lai. *Continuous Bivariate Distributions*. Springer, New York, 2009.

A. J. Baranchik. A family of minimax estimators of the mean of a multivariate normal distribution. *The Annals of Mathematical Statistics*, pages 642–645, 1970.

A. P. Basu and S. Dhar. Bivariate geometric distribution. *Journal of Applied Statistical Science*, 2(1):33–44, 1995.

T. Bayes, R. Price, and J. Canton. An essay towards solving a problem in the doctrine of chances. 1763.

J. O. Berger. *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 2013.

J. M. Bernardo and A. F. Smith. Bayesian theory, 2001.

J. Bickels, J. C. Wittig, et al. *Operative techniques in orthopaedic surgical oncology*. Lippincott Williams & Wilkins, 2012.

B. S. Blanchard, W. J. Fabrycky, and W. J. Fabrycky. *Systems engineering and analysis*, volume 4. Prentice Hall Englewood Cliffs, NJ, 1990.

J. M. Bland and D. G. Altman. Measuring agreement in method comparison studies. *Statistical methods in medical research*, 8(2):135–160, 1999.

H. W. Block and A. Basu. A continuous bivariate exponential extension. *Journal of the American Statistical Association*, 69(348):1031–1037, 1974.

H. Bolfarine and M. C. Sandoval. *Introdução à inferência estatística*, volume 2. SBM, 2001.

G. E. Box and G. C. Tiao. *Bayesian inference in statistical analysis*, volume 40. John Wiley & Sons, 2011.

R. R. Britney and R. L. Winkler. Bayesian point estimation and prediction. *Annals of the Institute of Statistical Mathematics*, 26(1):15–34, 1974.

G. Brunello and E. Nakano. Inference Bayesian in the discrete Weibull model in data with presence of censorship (in portuguese). *TEMA (São Carlos)*, 16(2):97–110, 2015.

I. W. Burr. On a general system of distributions. III. The sample range. *J. Amer. Statist. Assoc.*, 63, 1968.

I. W. Burr. Parameters for a general system of distributions to match a grid of α_3 and α_4. *Comm. Statist.*, 2, 1973.

V. G. Cancho and H. Bolfarine. Modeling the presence of immunes by using the exponentiated-Weibull model. *Journal of Applied Statistics*, 28(6):659–671, 2001.

V. G. Cancho, D. K. Dey, and F. Louzada. Unified multivariate survival model with a surviving fraction: an application to a Brazilian customer churn data. *Journal of Applied Statistics*, 43(3):572–584, 2016.

G. Casella and R. L. Berger. *Statistical inference*, volume 2. Duxbury Pacific Grove, CA, 2002.

M. Chao and J. C. Fu. The reliability of a large series system under Markov structure. *Advances in Applied Probability*, 23(4):894–908, 1991.

S. Chapman. *Public health advocacy and tobacco control: making smoking history*. Blackwell Pub., 2007.

S. Chib and E. Greenberg. Understanding the Metropolis-Hastings algorithm. *The american statistician*, 49(4):327–335, 1995.

M. Clinic. Diabetic retinopathy, 2018. URL https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/symptoms-causes/syc-20371611. Accessed: 2018-08-14.

D. Collett. *Modelling survival data in medical research*. CRC press, 2015.

G. M. Cordeiro. Introdução à teoria assintótica. *22º Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro*, 1999.

D. R. Cox. Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, 34:187–220, 1972.

D. R. Cox and E. J. Snell. A general definition of residuals (with discussion). *Journal of the Royal Statistical Society, Series B*, 30:248–275, 1968.

M. J. Crowder. *Multivariate survival analysis and competing risks*. CRC Press, 2012.

J. Crowley and M. Hu. Covariance analysis of heart transplant survival data. *Journal of the American Statistical Association*, 72:27–36, 1977.

C. Curtis, S. P. Shah, S.-F. Chin, G. Turashvili, O. M. Rueda, M. J. Dunning, D. Speed, A. G. Lynch, S. Samarajiwa, Y. Yuan, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, 486(7403):346, 2012.

N. Davarzani, J. A. Acgcar, E. N. Smirno, and R. Peeters. Bivariate lifetime geometric distribution in presence of cure fractions. *Journal of Data Science*, 13(4):755–770, 2015.

R. De Angelis, R. Capocaccia, T. Hakulinen, B. Soderman, and A. Verdecchia. Mixture models for cancer survival analysis: application to population-based data with covariates. *Statistics in medicine*, 18(4):441–454, 1999.

R. P. de Oliveira and J. A. Achcar. Basu-Dhar's bivariate geometric distribution in presence of censored data and covariates: some computational aspects. *Electronic Journal of Applied Statistical Analysis*, 11(1):108–136, 2018.

R. P. de Oliveira and J. A. Achcar. *BivGeo: Basu-Dhar Bivariate Geometric Distribution*, 2019. URL https://CRAN.R-project.org/package=BivGeo. R package version 2.0.1.

R. P. de Oliveira, J. A. Achcar, D. Peralta, and J. Mazucheli. Discrete and continuous bivariate lifetime models in presence of cure rate: a comparative study under Bayesian approach. *Journal of Applied Statistics*, pages 1–19, 2018.

R. P. de Oliveira, A. F. Menezes, J. Mazucheli, and J. A. Achcar. Mixture and nonmixture cure fraction models assuming discrete lifetimes: Application to a pelvic sarcoma dataset. *Biometrical Journal*, 2019.

M. H. DeGroot et al. *Probability and statistics*. 1986.

S. K. Dhar. Modeling with a bivariate geometric distribution. *Advances on Methodological and Applied Aspects of Probability and Statistics*, 1:101–109, 2003.

S. K. Dhar and S. Balaji. On the characterization of a bivariate geometric distribution. *Communications in Statistics—Theory and Methods*, 35(5):759–765, 2006.

A. Donaldson, L. Razak, J. Li, D. Fisher, and P. Tambyah. Carbapenems and subsequent multiresistant bloodstream infection: Does treatment duration matter? *International Journal of Antimicrobial Agents*, 34:246–251, 2009.

C. A. dos Santos and J. A. Achcar. A bayesian analysis for the Block and Basu bivariate exponential distribution in the presence of covariates and censored data. *Journal of Applied Statistics*, 38(10):2213–2223, 2011.

F. Downton. Bivariate exponential distributions in reliability theory. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 408–417, 1970.

P. K. Dunn and G. K. Smyth. Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5(3):236–244, 1996.

R. S. Ehlers. Introdução à inferência Bayesiana. 2007.

J. D. Englehardt and R. Li. The discrete Weibull distribution: An alternative for correlated counts with confirmation for microbial counts in water. *Risk Analysis*, 31(3): 370–381, 2011.

S. Eryilmaz and F. Tank. On reliability analysis of a two-dependent-unit series system with a standby unit. *Applied Mathematics and Computation*, 218(15):7792–7797, 2012.

M. Ezzati and A. D. Lopez. Regional, disease specific patterns of smoking-attributable mortality in 2000. *Tobacco control*, 13(4):388–395, 2004.

T. Facina. Estimativa 2014–Incidência de câncer no Brasil. *Rev Bras Cancerol*, 60(1): 63–64, 2014.

V. T. Farewell. The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, pages 1041–1046, 1982.

W. Feller. *An introduction to probability theory and its applications*. John Wiley & Sons London-New York-Sydney-Toronto, 1968.

J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. Parkin, D. Forman, and F. Bray. Cancer incidence and mortality worldwide: IARC Cancer Base no. 11. International Agency for Research on Cancer, France, 2013.

L. M. Fernandes. Inferência Bayesiana em modelos discretos com fração de cura. 2014.

M. C. Fiore. Us public health service clinical practice guideline: treating tobacco use and dependence. *Respiratory care*, 45(10):1200–1262, 2000.

T. R. Fleming and D. Lin. Survival analysis in clinical trials: past developments and future directions. *Biometrics*, 56(4):971–983, 2000.

Food, H. Drug Administration, et al. Deeming tobacco products to be subject to the Federal Food, Drug, and Cosmetic Act, as amended by the Family Smoking Prevention and Tobacco Control Act; restrictions on the sale and distribution of tobacco products and required warning statements for tobacco products. *Federal Register*, 81(90):28973, 2016.

J. E. Freund. A bivariate extension of the exponential distribution. *Journal of the American Statistical Association*, 56(296):971–977, 1961.

A. E. Gelfand and A. F. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409, 1990.

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian data analysis*. Chapman and Hall/CRC, 1995.

S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741, 1984.

M. Ghadimi, M. Mahmoodi, K. Mohammad, H. Zeraati, M. Rasouli, and M. Sheikhfathollahi. Family history of the cancer on the survival of the patients with gastrointestinal cancer in northern Iran, using frailty models. *BMC gastroenterology*, 11(1):104, 2011.

M. Ghitany, B. Atieh, and S. Nadarajah. Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78(4):493 – 506, 2008.

J. K. Ghosh, M. Delampady, and T. Samanta. *An introduction to Bayesian analysis: theory and methods*. Springer Science & Business Media, 2007.

W. R. Gilks, S. Richardson, and D. Spiegelhalter. *Markov Chain Monte Carlo in practice*. CRC press, 1995.

S. R. Giolo and E. A. Colosimo. Análise de sobrevivência aplicada. 2006.

G. A. Giovino, J. E. Henningfield, S. L. Tomar, L. G. Escobedo, and J. Slade. Epidemiology of tobacco use and dependence. *Epidemiologic reviews*, 17(1):48–65, 1995.

E. J. Gumbel. Bivariate exponential distributions. *Journal of the American Statistical Association*, 55(292):698–707, 1960.

M. Hafezi, M. Bohnert, W. Weinmann, and S. Pollak. Prevalence of nicotine consumption in drug deaths. *Forensic science international*, 119(3):284–289, 2001.

D. Hanagal and K. Ahmadi. Estimation of parameters by EM algorithm in bivariate exponential distribution based on censored samples. *Econ. Quality Control*, 23(2):257–66, 2008.

D. D. Hanagal. Bivariate Weibull regression model based on censored samples. *Statistical Papers*, 47(1):137–147, 2006.

A. G. Hawkes. A bivariate exponential distribution with applications to reliability. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 129–131, 1972.

P. Hougaard. A class of multivariate failure time distributions. *Biometrika*, pages 671–678, 1986.

P. Hougaard. *Analysis of multivariate survival data*. Springer Science & Business Media, 2012.

Z. Hu and S. Mahadevan. Time-dependent system reliability analysis using random field discretization. *Journal of Mechanical Design*, 137(10):101404, 2015.

F. L. Hulting and J. A. Robinson. The reliability of a series system of repairable subsystems: a Bayesian approach. *Naval Research Logistics (NRL)*, 41(4):483–506, 1994.

W. J. Huster, R. Brookmeyer, and S. G. Self. Modelling paired survival data with covariates. *Biometrics*, pages 145–156, 1989.

J. G. Ibrahim, M.-H. Chen, and D. Sinha. *Bayesian survival analysis*. Wiley Online Library, 2005.

T. F. Initiative. Why is tobacco a public health priority. *World Health Organization*, pages 27–12, 2004.

S. Institute. *SAS/STAT user's guide: version 6*, volume 2. SAS Institute, 1990.

T. N. E. Institute. Facts about diabetic eye disease, 2018. URL https://nei.nih.gov/health/diabetic/retinopathy. Accessed: 2018-08-14.

H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. A*, 186(1007):453–461, 1946.

P. A. Jensen and J. F. Bard. *Operations research models and methods*, volume 1. John Wiley & Sons Incorporated, 2003.

G. John, S. Pasche, N. Rothen, A. Charmoy, C. Delhumeau-Cartier, and D. Genné. Tobacco-stained fingers: a clue for smoking-related disease or harmful alcohol use? a case–control study. *BMJ open*, 3(11):e003304, 2013.

G. John, C. Louis, A. Berner, and D. Genné. Tobacco stained fingers and its association with death and hospital admission: A retrospective cohort study. *PloS one*, 10(9): e0138211, 2015a.

G. John, C. Louis, A. Berner, and D. Genné. Data from: Tobacco stained fingers and its association with death and hospital admission: a retrospective cohort study, 2015b. URL https://doi.org/10.5061/dryad.4478v.

N. L. Johnson, A. W. Kemp, and S. Kotz. *Univariate discrete distributions*, volume 444. John Wiley & Sons, 2005.

J. D. Kalbfleisch and R. L. Prentice. *The Statistical Analysis of Failure Time Data*. Wiley, New York, NY, 2nd edition, 2002.

E. L. Kaplan and P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

A. W. Kemp. New discrete Appell and Humbert distributions with relevance to bivariate accident data. *Journal of multivariate analysis*, 113:2–6, 2013.

C. Kemp and H. Papageorgiou. Bivariate Hermite distributions. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 269–280, 1982.

M. A. Khan, A. Khalique, and A. Abouammoh. On estimating parameters in a discrete Weibull distribution. *Reliability, IEEE Transactions on*, 38(3):348–350, 1989.

A. Khursheed and K. Lai Saxena. Positive dependence in multivariate distributions. *Communications in Statistics-Theory and Methods*, 10(12):1183–1196, 1981.

M. Kijima. *Markov processes for stochastic modeling*. Springer, 2013.

H. S. Klakattawi, V. Vinciotti, and K. Yu. A simple and adaptive dispersion regression model for count data. *Entropy*, 20(2):1–15, 2018. ISSN 1099-4300.

J. P. Klein and M. L. Moeschberger. *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag, New York, 1997.

S. Kocherlakota. Discrete bivariate weighted distributions under multiplicative weight function. *Communications in statistics-theory and methods*, 24(2):533–551, 1995.

S. Kocherlakota and K. Kocherlakota. *Bivariate discrete distributions*. Wiley Online Library, 1992.

A. N. Kolmogorov. Foundations of the theory of probability. 1950.

K. Kołowrocki. *Reliability of large systems*. Wiley Online Library, 2008.

S. Kotz, N. L. Johnson, and N. Balakrishnan. Continuous univariate distributions, 2000.

H. Krishna and P. S. Pundir. A bivariate geometric distribution with applications to reliability. *Communications in Statistics—Theory and Methods*, 38(7):1079–1093, 2009.

K. B. Kulasekera. Approximate MLE's of the parameters of a discrete Weibull distribution with type I censored data. *Microelectronics Reliability*, 34(7):1185–1188, 1994.

C. S. Kumar. A unified approach to bivariate discrete distributions. *Metrika*, 67(1): 113–123, 2008.

D. Kundu. Geometric skew normal distribution. *Sankhya B*, 76(2):167–189, 2014.

D. Kundu and A. K. Dey. Estimating the parameters of the Marshall–Olkin bivariate Weibull distribution by EM algorithm. *Computational Statistics & Data Analysis*, 53 (4):956–965, 2009.

D. Kundu and R. D. Gupta. A class of absolutely continuous bivariate distributions. *Statistical Methodology*, 7(4):464–477, 2010.

D. Kundu and V. Nekoukhou. Univariate and bivariate geometric discrete generalized exponential distributions. *Journal of Statistical Theory and Practice*, 12(3):595–614, 2018.

R. Labs. What is diabetic retinopathy?, 2018. URL https://www.retina-labs.com/understanding-diabetic-retinopathy/. Accessed: 2018-08-14.

C.-D. Lai. Constructions of discrete bivariate distributions. In *Advances in Distribution Theory, Order Statistics, and Inference*, pages 29–58. Springer, 2006.

P. C. Lambert, J. R. Thompson, C. L. Weston, and P. W. Dickman. Estimating and modeling the cure fraction in population-based cancer survival analysis. *Biostatistics*, 8(3):576–594, 2006.

J. F. Lawless. *Statistical models and methods for lifetime data*. John Wiley & Sons, 1982.

E. T. Lee and J. W. Wang. *Statistical methods for survival data analysis*. Wiley Series in Probability and Statistics. Hoboken, NJ, third edition, 2003.

H. Lee and J. H. Cha. On construction of general classes of bivariate distributions. *Journal of Multivariate Analysis*, 127:151–159, 2014.

E. J. Lehmann and G. Casella. *Theory of Point Estimation*. Springer Verlag, 1998.

E. L. Lehmann and G. Casella. *Theory of point estimation*. Springer Science & Business Media, 2006.

J. Li and S. K. Dhar. Modeling with bivariate geometric distributions. *Communications in Statistics-Theory and Methods*, 42(2):252–266, 2013.

J. Li and S. Ma. *Survival analysis in medicine and genetics*. CRC Press, 2013.

A. Lilienfeld and D. Lilienfeld. *Foundations of Epidemiology*. Oxford University Press, New York, 1980.

E. Lim, S. Zhang, J. Li, W. Yap, T. Howe, B. Tan, Y. Lee, D. Wong, K. Khoo, K. Seto, L. Tan, T. Agasthian, H. Koong, T. John, C. Tan, M. Caleb, A. Chang, A Ng, and P. Tan. Using whole genome amplification (WGA) of low-volume biopsies to assess the prognostic role of EGFR, KRAS, p53, and CMET mutations in advanced-stage nonsmall cell lung cancer (NSCLC). *Journal of Thoracic Oncology*, 4:12–21, 2009.

D. V. Lindley. Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20:102–107, 1958.

W. Lu. Efficient estimation for an accelerated failure time model with a cure fraction. *Statistica Sinica*, 20:661, 2010.

M. N. Magalhães and A. C. P. de Lima. *Noções de probabilidade e estatística*. IME-USP São Paulo:, 2000.

R. A. Maller and X. Zhou. *Survival analysis with long-term survivors*. Wiley New York, 1996.

N. Mantel, N. R. Bohidar, and J. L. Ciminera. Mantel-Haenszel analyses of litter-matched time-to-response data, with modifications for recovery of interlitter information. *Cancer Research*, 37(11):3863–3868, 1977.

A. W. Marshall and I. Olkin. A generalized bivariate exponential distribution. *Journal of Applied Probability*, 4(02):291–302, 1967a.

A. W. Marshall and I. Olkin. A multivariate exponential distribution. *Journal of the American Statistical Association*, 62(317):30–44, 1967b.

A. W. Marshall and I. Olkin. A family of bivariate distributions generated by the bivariate Bernoulli distribution. *Journal of the American Statistical Association*, 80(390):332–338, 1985.

A. W. Marshall and I. Olkin. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84 (3):641–652, 1997.

J. Mazucheli, L. B. Fernandes, and R. P. de Oliveira. *LindleyR: The Lindley Distribution and Its Modifications*, 2016. URL https://CRAN.R-project.org/package=LindleyR. R package version 1.1.0.

W. Q. Meeker and L. A. Escobar. *Statistical Methods for Reliability Data*. John Wiley & Sons, New York, 1998.

S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.

R. B. Millar. *Maximum likelihood estimation and inference: with examples in R, SAS and ADMB*, volume 111. John Wiley & Sons, 2011.

C. Minitab. Minitab reference manual (release 7.1). *State Coll., PA16801, USA*, 1991.

S. MirMostafaee, M. Mahdizadeh, and A. J. Lemonte. The marshall–olkin extended generalized rayleigh distribution: Properties and applications. *Communications in Statistics-Theory and Methods*, 46(2):653–671, 2017.

J. E. Mitchell and L. A. Dahlgren. Prevalence of smoking among psychiatric outpatients. *Am J psychiatry*, 143(8):993–997, 1986.

D. Morgenstern. Einfache beispiele zweidimensionaler verteilungen. *Mitteilingsblatt fur Mathematische Statistik*, 8:234–235, 1956.

Y. Mori and B. R. Ellingwood. Time-dependent system reliability analysis by adaptive importance sampling. *Structural safety*, 12(1):59–73, 1993.

K. Muraleedharan Nair and N. Unnikrishnan Nair. On characterizing the bivariate exponential and geometric distributions. *Annals of the institute of Statistical Mathematics*, 40(2):267–271, 1988.

D. P. Murthy, M. Xie, and R. Jiang. *Weibull models*, volume 505. John Wiley & Sons, 2004.

T. Nakagawa and S. Osaki. The discrete Weibull distribution. *IEEE Transactions on Reliability*, 24(5):300–301, 1975.

V. Nekoukhou and D. Kundu. Bivariate discrete generalized exponential distribution. *Statistics*, 51(5):1143–1158, 2017.

G. W. Oehlert. A note on the delta method. *The American Statistician*, 46(1):27–29, 1992.

U. D. of Health and H. Services. The health consequences of involuntary exposure to tobacco smoke: a report of the surgeon general. *Coordinating Center for Health Promotion, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health*, 2006.

R. Oliveira, J. Mazucheli, and J. Achcar. A generalization of basu-dhar's bivariate geometric distribution to the trivariate case. *Communications in Statistics - Simulation and Computation*, 2019. doi: 10.1080/03610918.2019.1643881.

R. P. Oliveira and J. A. Achcar. Use of basu-dhar bivariate geometric distribution in the analysis of the reliability of component series systems. *International Journal of Quality & Reliability Management*, 36(4):569–586, 2019.

W. H. Organization. Prevalence and incidence of selected sexually transmitted infections, Chlamydia Trachomatis, Neisseria Gonorrhoeae, Syphilis and Trichomonas Vaginalis: methods and results used by WHO to generate 2005 estimates. Technical report, Geneva: World Health Organization, 2011.

W. H. Organization. Tobacco, 2018. URL http://www.who.int/news-room/fact-sheets/detail/tobacco. Accessed: 2018-08-14.

W. H. Organization and R. for International Tobacco Control. *WHO report on the global tobacco epidemic, 2008: the MPOWER package*. World Health Organization, 2008.

W. H. Organization et al. Tobacco or health: a global status report. 1997.

M. Othus, B. Barlogie, M. L. LeBlanc, and J. J. Crowley. Cure models as a useful statistical tool for analyzing survival. *Clinical Cancer Research*, 18(14):3731–3736, 2012.

A. Papoulis and S. U. Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

Y. Pawitan. *In all likelihood: statistical modelling and inference using likelihood*. Oxford University Press, 2001.

D. H. Peter and A. Hoff. A first course in Bayesian statistical methods, 2009.

R. Peto, J. Boreham, A. D. Lopez, M. Thun, and C. Heath. Mortality from tobacco in developed countries: indirect estimation from national vital statistics. *The Lancet*, 339 (8804):1268–1278, 1992.

D. L. Price and A. K. Manatunga. Modelling survival data with a cured fraction using frailty models. *Statistics in medicine*, 20(9-10):1515–1527, 2001.

S. E. Puchner, P. T. Funovics, C. Böhler, A. Kaider, C. Stihsen, G. M. Hobusch, J. Panotopoulos, and R. Windhager. Oncological and surgical outcome after treatment of pelvic sarcomas. *PloS one*, 12(2):e0172203, 2017.

M. Rausand and H. Arnljot. *System reliability theory: models, statistical methods, and applications*, volume 396. John Wiley & Sons, 2004.

C. A. Rohde. *Introductory Statistical Inference with the Likelihood Function*. Springer-Verlag, New York, 2014.

J. L. Romeu. Understanding series and parallel systems reliability. *Selected Topics in Assurance Related Technologies (START), Department of Defense Reliability Analysis Center (DoD RAC)*, 11(5), 2004.

D. Roy. Discretization of continuous distributions with an application to stress-strength reliability. *Bulletin of the Calcutta Statistical Association*, 52(205-208):295–314, 2002.

D. Roy. Discrete Rayleigh distribution. *IEEE Transactions on Reliability*, 53(2):255–260, 2004a.

D. Roy. Bivariate models from univariate life distributions: a characterization cum modeling approach. *Naval Research Logistics (NRL)*, 51(5):741–754, 2004b.

M. Sankaran. The discrete Poisson-Lindley distribution. *Biometrics*, 26:145–149, 1970.

S. K. Sarkar. A continuous bivariate exponential distribution. *Journal of the American Statistical Association*, 82(398):667–675, 1987.

SAS. *The NLMIXED Procedure, SAS/STAT® User's Guide, Version 9.4*. Cary, NC: SAS Institute Inc., 2010.

D. Sedmak, T. Meineke, D. Knechtges, and J. Anderson. Prognostic significance of cytokeratin-positive breast cancer metastases. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc*, 2(5):516–520, 1989.

Z. Šidák. Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, 62(318):626–633, 1967.

Z. Sidák et al. On multivariate normal probabilities of rectangles: their dependence on correlations. *The Annals of Mathematical Statistics*, 39(5):1425–1434, 1968.

R. Siegel, D. Naishadham, and A. Jemal. Cancer statistics. *CA: a cancer journal for clinicians*, 63(1):11–30, 2013.

R. Siegel, J. Ma, Z. Zou, and A. Jemal. Cancer statistics. *CA: a cancer journal for clinicians*, 64(1):9–29, 2014.

R. L. Siegel, K. D. Miller, and A. Jemal. Cancer statistics. *CA: a cancer journal for clinicians*, 66(1):7–30, 2016.

C. Singh and R. Billinton. *System reliability, modelling and evaluation*, volume 769. Hutchinson London, 1977.

R. Singh and K. Mukhopadhyay. Survival analysis in clinical trials: Basics and must know areas. *Perspectives in clinical research*, 2(4):145, 2011.

D. Spiegelhalter, A. Thomas, N. Best, and D. Lunn. OpenBUGS user manual, version 3.0. 2. *MRC Biostatistics Unit, Cambridge*, 2007.

D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van Der Linde. The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(3):485–493, 2014.

V. Sreeja and P. Sankaran. *Regression models for bivariate survival data.* PhD thesis, Cochin University of Science and Technology, 2008.

C. M. Stein. Confidence sets for the mean of a multivariate normal distribution. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 265–296, 1962.

Y.-S. Su and M. Yajima. *R2jags: Using R to Run 'JAGS'*, 2015. URL https://CRAN.R-project.org/package=R2jags. R package version 0.5-7.

P. Sugarbaker. Management of abdominopelvic sarcoma. *Musculoskeletal Cancer Surgery*, pages 147–163, 2004.

K. Sun and A. P. Basu. A characterization of a bivariate geometric distribution. *Statistics & probability letters*, 23(4):307–311, 1995.

N. L. S. T. R. Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*, 365(5):395–409, 2011.

H. Thakurdas. Smoking and psychiatric patients. *British Medical Journal*, 2(5353):388, 1963.

T. C. P. G. T. Tobacco et al. A clinical practice guideline for treating tobacco use and dependence: 2008 update: a us public health service report. *American journal of preventive medicine*, 35(2):158, 2008.

Y. L. Tong. *The multivariate normal distribution.* Springer Science & Business Media, 2012.

A. Tsodikov, J. Ibrahim, and A. Yakovlev. Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association*, 98(464):1063–1078, 2003.

G. Tutz, M. Schmid, et al. *Modeling discrete time-to-event data*. Springer, 2016.

M. Vahidpour. *Cure Rate Models*. PhD thesis, École Polytechnique de Montréal, 2016.

V. Vaidyanathan, A. Sharon Varghese, et al. Morgenstern type bivariate Lindley distribution. *Statistics, Optimization & Information Computing*, 4(2):132–146, 2016.

H. M. Walker and J. Lev. Statistical inference. 1953.

A. Wienke, P. Lichtenstein, and A. I. Yashin. A bivariate frailty model with a cure fraction for modeling familial correlations in diseases. *Biometrics*, 59(4):1178–1183, 2003.

A. Wienke, I. Locatelli, and A. I. Yashin. The modelling of a cure fraction in bivariate time-to-event data. *Austrian Journal of Statistics*, 35(1):67–76, 2006.

P. A. Wingo, S. Bolden, T. Tong, S. L. Parker, L. M. Martin, and C. W. Heath Jr. Cancer statistics for african americans. *CA: a cancer journal for clinicians*, 46(2): 113–125, 1996.

E. L. Wynder and E. A. Graham. Tobacco smoking as a possible etiologic factor in bronchogenic carcinoma: A study of six hundred and eighty-four proved cases1. *The Challenge of Epidemiology: Issues and Selected Readings*, 505:458, 1988.

X. Xiaoling, G. Beiqing, and W. Ronghua. The properties of Bivariate Geometric Distribution-Type II. In *2012 Second International Conference on Business Computing and Global Informatization*, pages 584–587, Oct 2012. doi: 10.1109/BCGIN.2012.158.

A. Yazdanbod, M. S. Naseri, and R. Malekzadeh. Upper gastrointestinal cancer in Ardabil, North-West of Iran: A review. 2004.

G. Yin and J. G. Ibrahim. Cure rate models: a unified approach. *Canadian Journal of Statistics*, 33(4):559–570, 2005.

B. Yu, R. C. Tiwari, K. A. Cronin, and E. J. Feuer. Cure fraction estimation from the mixture cure models for grouped survival data. *Statistics in medicine*, 23(11):1733–1747, 2004.

M. Zali et al. Indices related to gastric cancer in Tehran and seven city provinces in the years 1999 to 2002. *J Islamic Azad Univ Med*, 15(1):15–8, 2005.

A. Zellner. *An introduction to Bayesian inference in econometrics*. 1996.

T. Zhang and M. Horigome. Availability and reliability of system with dependent components and time-varying failure and repair rates. *IEEE Transactions on reliability*, 50 (2):151–158, 2001.