UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Faculdade de Ciências e Tecnologia
Câmpus de Presidente Prudente

Bayesian and Classical Inference for Extensions of Geometric Exponential Distribution with Applications in Survival Analysis Under the Presence of the Data Covariated and Randomly Censored

Paulo Roberto de Lima Gianfelice

Academic Supervisor: Prof. Fernando Antonio Moala, Ph.D.

Post-Graduate Program in Applied and Computational Mathematics

Presidente Prudente, February 2020

UNIVERSIDADE ESTADUAL PAULISTA
'JÚLIO DE MESQUITA FILHO'
Faculdade de Ciências e Tecnologia de Presidente Prudente
Post-Graduate Program in Applied and Computational Mathematics

Bayesian and Classical Inference for Extensions of Geometric Exponential Distribution with Applications in Survival Analysis Under the Presence of the Data Covariated and Randomly Censored

Paulo Roberto de Lima Gianfelice

Academic Supervisor: Prof. Fernando Antonio Moala, Ph.D.

Master's Thesis presented to the Post-Graduate Program in Applied and Computational Mathematics of the Faculty of Science and Technology of the São Paulo State University "Júlio de Mesquita Filho", Presidente Prudente campus, in partial fulfillment of the requirements for title of Master in Applied and Computational Mathematics.

Presidente Prudente, February 2020

FICHA CATALOGRÁFICA

Gianfelice, Paulo Roberto de Lima
G433b Bayesian and Classical Inference for Extensions of Geometric Exponential Distribution with Data Covariated and Randomly Applications in Survival Analysis Under the Presence of the Censored/ Paulo Roberto de Lima Gianfelice - Presidente Prudente, 2020
91 p.: il, tabs.

Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia, Presidente Prudente, 2020
Orientador: Fernando Antonio Moala
Inclui bibliografia

1. Censored and covariate data. 2. Maximum likelihood and bayesian estimation. 3. Extensions for Exponential Geometric distribution. 4. Statistical simulation. VI. Título.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Presidente Prudente

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: Bayesian and classical inference for extensions of geometric exponential distribution with applications in survival analysis under the presence of the data covariated and randomly censored

AUTOR: PAULO ROBERTO DE LIMA GIANFELICE
ORIENTADOR: FERNANDO ANTONIO MOALA

Aprovado como parte das exigências para obtenção do Título de Mestre em MATEMÁTICA APLICADA E COMPUTACIONAL, pela Comissão Examinadora:

Prof. Dr. FERNANDO ANTONIO MOALA
Departamento de Estatística / Faculdade de Ciências e Tecnologia de Presidente Prudente

Prof. Dr. LUIZ CARLOS BENINI
Departamento de Estatística / Faculdade de Ciências e Tecnologia de Presidente Prudente

VIDEOCONFERÊNCIA

Prof. Dr. PEDRO LUIZ RAMOS
Pós-Doutorando / Universidade de São Paulo

Presidente Prudente, 26 de fevereiro de 2020

To all the people who supported, taught, instructed and guided me.
In particular, for Olézia Gianfelice is dedicad!

Acknowledgements

Acknowledgements first and foremost, to all gods for my soul, health, wisdom, deliverance from evil, courage and patience, without these items I would not have been not even able even win my self... not even human I would be. It is not possible for man, alone, to leave the most unlikely place, overcome all obstacles, go against all expectations and perform a feat like this, in a place like this.

Thank my parents for conceiving me, as the first and with the greatest love on the 15th day of January 1985. Without their care, examples and deeds I would never have taken this direction, arrived where I am and taken the steps decisions I made.

I will thank you until the end of my life for my three beautiful brothers, they are the primary inspiration and reason for my conquests. I thank my whole family from the bottom of my heart for their understanding, support and encouragement, without their support and advice I would have been lost me in the course of this journey. I especially thank from "Jão", my cousin-friend-brother, and my aunt "Cris", the first for always remembering and believing in me and the second for all the affection, support, help, understanding and value. I seek success in my achievements as an academic, and I will continue to seek in other areas of life, as a way of compensating and honoring them. Their care, guidance and incentives made me a person capable of overcoming anything.

To all my teachers, in particular to Gilcilene Sanchez De Paulo and Manoel Ivanildo Silvestre Bezerra. For the professional inclination and attendance, the first for teaching me to lead the learning of the most important concept in this line of the academy with affection and patience, and the second for believing in my ability, giving me freedom in the construction of my academic works, reading them and mainly trusting my proposals. In particular, I thank Professor Marta, for her Probability and Statistics classes in my 4th year of the Mathematics course opened my eyes and aroused interest in this career.

To my companions in study and unwinding, André Antunes, Bruno Gois, José Vanterler (Zé), Laisson, Leonardo Kengi (Léo), Marcos Viana (Osama) and Rafael Paulino (Pão), who besides being extraordinary people and academics, believed in my capabilities, tolerated me in unpleasant moments of tiredness and frustration. I am grateful to them for the guidance, influence, notes, advice, invitations and doubts taken about n situations of personal and academic life, especially because they are extremely nice people and exemplary academics I feel privileged to have lived with them at this stage.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, case number 88882.441640/2019-01, and FCT/UNESP, through Pos MAC, therefore, a recognition especially to for these institutions by the opportunity, credit, structure and privileges in the development of this work, training and academic career. Through these two institutions I also transfer my gratitude to each of the Brazilians who contributed with culture, kindness, serenity, trust, and work to pay yours taxes, fees and all tax contributions for development and outcome of scientific and technological work , as well as for the training and academic career of scientists/researchers in Brazil.

"E o gládio a erguer, que arrasa e que depreda,
E o olhar que ante a agnomínia não desmaia,
Luta! E é forçoso que ao lutar não caia,
Pois se cair o esmagarão na queda."
(Aleluias, ***Raimundo Correia)***

Abstract

This work presents a study of probabilistic modeling, with applications to survival analysis, based on a probabilistic model called Exponential Geometric (EG), which offers great flexibility for the statistical estimation of its parameters based on samples of life time data complete and censored. In this study, the concepts of estimators and lifetime data are explored under random censorship in two cases of extensions of the EG model: the Extended Geometric Exponential (EEG) and the Generalized Extreme Geometric Exponential (GE2). The work still considers, exclusively for the EEG model, the approach of the presence of covariates indexed in the rate parameter as a second source of variation to add even more flexibility to the model, as well as, exclusively for the GE2 model, a analysis of the convergence, hitherto ignored, it is proposed for its moments. The statistical inference approach is performed for these extensions in order to expose (in the classical context) their maximum likelihood estimators and asymptotic confidence intervals, and (in the bayesian context) their a priori and a posteriori distributions, both cases to estimate their parameters under random censorship, and covariates in the case of EEG. In this work, bayesian estimators are developed with the assumptions that the prioris are vague, follow a Gamma distribution and are independent between the unknown parameters. The results of this work are regarded from a detailed study of statistical simulation applied to compare the estimation procedures approached under the pretext of evaluating these estimators based on the 95% coverage probability, mean square error, mean bias and the mean interval amplitude. At the end of each extension's approach, an application with real data is also presented to highlight the reach and particularities of the extended model addressed.

Keywords: Censored and covariate data, Maximum likelihood and bayesian estimation, Extensions for Exponential Geometric distribution, Statistical simulation.

Resumo

Este trabalho apresenta um estudo de modelagem probabilística, com aplicações à análise de sobrevivência, fundamentado em um modelo probabilístico denominado Exponencial Geométrico (EG), que oferece uma grande flexibilidade para a estimação estatística de seus parâmetros com base em amostras de dados de tempo de vida completos e censurados. Neste estudo são explorados os conceitos de estimadores e dados de tempo de vida sob censuras aleatórias em dois casos de extensões do modelo EG: o Exponencial Geométrico Estendido (EEG) e o Exponencial Geométrico Extremo Generalizado (GE2). O trabalho ainda considera, exclusivamente para o modelo EEG, a abordagem de presença de covariáveis indexadas no parâmetro de taxa como uma segunda fonte de variação para acrescentar ainda mais flexibilidade para o modelo, bem como, exclusivamente para o modelo GE2, uma análise de convergência até então ignorada, é proposta para seus momentos. A abordagem da inferência estatística é realizada para essas extensões no intuito de expor (no contexto clássico) seus estimadores de máxima verossimilhança e intervalos de confiança assintóticos, e (no contexto bayesiano) suas distribuições à priori e posteriori, ambos os casos para estimar seus parâmetros sob as censuras aleatórias, e covariáveis no caso do EEG. Neste trabalho os estimadores bayesianos são desenvolvidos com os pressupostos de que as prioris são vagas, seguem uma distribuição Gama e são independentes entre os parâmetros desconhecidos. Os resultados deste trabalho são resguardados de um estudo detalhado de simulação estatística aplicado para comparar os procedimentos de estimação abordados sob o pretexto de avaliar estes estimadores com base na probabilidade de 95% de cobertura, erro quadrático médio, vício médio e a amplitude intervalar média. Ao final da abordagem de cada extensão é apresentada ainda uma aplicação com dados reais para destacar o alcance e as particularidades do modelo estendido abordado.

Palavras-Chave: Dados censurados e covariados, Estimação de máxima verossimilhança e bayesiano, Extensões para a distribuição Exponencial Geométrica, Simulação estatística.

List of Figures

List of Tables

List of Acronyms

AIC: Akaike Information Criterion
AICC: Akaike Information Criterion Corrected
BIC: Bayesian Information Criterion
DIC: Deviance Information Criterion
E2G: Exponentiated Exponential Geometric.
EG: Exponential Geometric.
CE2G: Complementary Exponentiated Exponential Geometric.
CEG: Complementary Exponential Geometric
EEG: Extended Exponential Geometric.
ELL: Exponentiated Log-Logistic
EM: Expectation Maximization
EW: Weibull Exponentiated
GE2: Generalized Exponential Geometric Extreme
GEG: Generalized Exponential Geometric
GG: Generalized Gamma
HPD: Highest posterior Density
LCEG: Long-Term Complementary Exponential Geometric
LE2G: Long-Term Extended Exponential Geometric
LL: Log-Logistic
MBA: Mean Bias Absolute
MCMC: Monte Carlo Markov Chain
MH: Metropolis-Hastings
MIA: Mean Interval Amplitude
MLE: Maximum Likelihood Estimates
MOEE: Marshall-Olkin Extended Exponential
MOExp: Marshall-Olkin Exponential
MOExpExt: Marshall-Olkin Exponential Extension
MOGE: Marshall-Olkin Generalized Exponential
MSE: Mean Square Errors
TTT: Total Time in Teste
VIF: Variance Inflation Factor

Contents

Chapter 1

Introduction: Probabilistic Models, Genesis and Extended Summary

Statistics can be announced a science inherent to Probability Theory, a concept that in the most refined context is approached in the light of Analysis Mathematics in a branch of study known as Measure Theory and that is in constant discoveries since its origin as scientific knowledge in the first decade of the 19th century (see Laplace [32].

In the same vein, several applications are summarized in the shadow of the Probability Theory that is summarized the data analysis to explain the frequency of occurrence of the natural and/or artificial events and phenomena present in all fields of human knowledge, both in observational studies as in experiments to model randomness and uncertainty in order to estimate or enable the prediction of future observations of these phenomena, since even in the most basic data analysis the processes are still dependent on measures that can be interpreted as counting, size, mass, volume ... or any additive and multiplicative property through the fundamental principle of counting.

In this sense, the Probability Distributions (or Probabilistic Models) until then presented in the literature are intrinsic to the Theory of Probability and Statistics, and are consolidated as the platform for statistical analysis in practically all areas of study, whether in the context of the probabilistic modeling for estimate the probability of occurrence of events, as the same in the statistical inference for the decision making and/or also through statistical modeling under conclusions about a subset of representative values of a universe, the sample of a population.

1.1 Initial Considerations

Since Pierre Simon Laplace, who according to Hájek [21], is considered the precursor of probability theory, a wide range of probability distributions have been proposed to model a multitude of random phenomena, and in recent decades the statistical literature related to Survival Analysis and Reliability Theory, in addition to becoming more complex mathematically and computationally, according to Harris and Albert [23], it is enormously rich in useful results regarding not only analytical results, but also to the proposal of probability distributions due to the modeling of survival and reliability functions that, above all, are probabilistic models for a random variable T positive for a temporal phenomenon, a quantile t for the model, called failure time (or lifetime).

In survival and reliability studies, a given distribution function (or cumulative probability function), commonly denoted by F, is at the heart of the survival or reliability

function, denoted respectively by S and R, which by in turn, with the probability density function (or mass probability function) f derived from F, it is the description of the hazard rate resulting from the effects of time exposures, the hazard rate function (or risk function) commonly denoted by h.

In this context, the common practice in modeling random phenomena through the results of the survival and reliability analysis consists of prior knowledge of the forms of the hazard rate function, since whatever the probability distribution from which a function S derives or R, such functions are strictly decreasing probability functions, while the h function contains five basic forms: constant, increasing, decreasing, unimodal and bathtub.

The richness of the diversity of forms that the hazard rate can take over time is the main feature in the modeling of the functions of lifetime random variables and the more shapes a model captures, the more indicated him it will be to model the failure time.

However, it is very common to observe that the greater the number of θ parameters in the parameter vector of a probability distribution, the greater the number of hazard rate forms captured by its hazard rate function.

Two-parameter models, such as those discussed in the recent studies of Ramos [50], Braguim [5] and Reis [53], respectively derived from the fundamental distributions Gama, Log-Logistic (LL) and Weibull, although limited in their characteristics and unable to show ample flexibility, they present under certain parametric conditions the forms of increasing and decreasing risk, and with the exception of LL, they still capture the constant form. But, as more parameters are inserted in these models, as developed in Stacy et al. [62], Mudholkar and Srivastava [44] and Rosaiah, Kantam and Kumar [56] to propose, respectively, the Generalized Gamma (GG), Weibull Exponentiated (EW) and Exponentiated Log-Logistic (ELL) distributions, more limitations are also inserts in its characteristics and flexibility, but more forms for the hazard rate function are now captured, such as unimodal and bathtub risks in both distributions.

Such searches have been developed over the years, either for a better exploration of the observed phenomena or to obtain distributions with more forms for hazard rate, such as obtaining bimodal risk models as developed in Mendoza, Ortega and Cordeiro [41] and, no less important, the discovery of hazard rate functions with reduced number of parameters but that captures as many forms of hazard rate as possible, as shown in the work of Ramos, Louzada and Ramos [52], where the approached distribution assumes the four basic forms in the parametric conditions of only two parameters.

For family Exponential distribution, for exemple, the obtaining procedures consist, in the most basic approach, as proposed by Gupta and Kundu [20] for a random variable T, in exponentialize the distribution function F of interest in the shape parameter $\pi > 0$ to be added, so that, with the parametric vector $\boldsymbol{\theta}$

$$\Psi(t|\boldsymbol{\theta}, \pi) = [F(t|\boldsymbol{\theta})]^{\pi} \ , \tag{1.1}$$

in the domain $t, \boldsymbol{\theta}, \pi \in \mathbb{R}_{+}^{*}$ so that

$$\psi(t|\boldsymbol{\theta}, \pi) = \frac{d}{d\pi}\Psi(t|\boldsymbol{\theta}, \pi) = \pi[F(t|\boldsymbol{\theta})]^{\pi-1} f(t|\boldsymbol{\theta}) \ . \tag{1.2}$$

Another method, introduced by Marshall and Olkin [39], which in line that the method proposed in Gupta and Kundu [20] by introducing a new shape parameter, is more general because it does not consider restrictions for the shape parameter π, as well as for the random variable X. Generalization is also obtained from a F distribution function as

$$\Psi(x|\boldsymbol{\theta}, \pi) = \frac{\pi f(x|\boldsymbol{\theta})}{1 - \pi F(x|\boldsymbol{\theta})} \ , \tag{1.3}$$

in the domain of $t, \pi \in \mathbb{R}$ and $\boldsymbol{\theta} \in \Theta$, there Θ is parametric space for the parametric vector $\boldsymbol{\theta}$, and so that

$$\psi(x|\boldsymbol{\theta}, \pi) = \frac{\pi F(x|\boldsymbol{\theta})}{[1 - \pi F(x|\boldsymbol{\theta})]^2} \ . \tag{1.4}$$

See that ψ is the distribution function of a new probability distribution, provided that your first derivative exists in relation to π and in the sample space $\Omega \subseteq \mathbb{R}$, for every $\mathbb{X} \in \Omega$ event still occur

$$P(\mathbb{X}) \geq 0 \ , \tag{1.5}$$

$$P\left(\bigcup_{i=1}^{\infty} \mathbb{X}_i\right) = \sum_{i=1}^{\infty} P(\mathbb{X}_i) \ , \tag{1.6}$$

$$P(\Omega) = \int_0^{\infty} \psi(t|\boldsymbol{\theta}, \pi) dt = 1 \ . \tag{1.7}$$

Another widely applied artifice consists of mixing the f distribution with a given φ distribution to obtain a new probability distribution called composite models (or mix models).

In Tahir and Cordeiro [63], the authors scanned the literature on the analysis of survival and reliability in search of the models generated through this technique and cataloged a vast number of distributions composed since that this procedure was used in 1997 by the work of Adamidis and Loukas [3].

The authors in Tahir and Cordeiro [63] still point out that there is not only one procedure for obtaining a composite distribution, and that they follow three categories of composition of distributions that start from the simple generalization called the "first generalization approach" (or G-classes to denote the classes of generalized distributions), actually reaching the procedures properly called composite models in the "second generalization approach" (or composition) and arrive in the current procedures for obtaining composite models in the category called "recent trends in composition".

Above all, the first model composed in the literature, proposed Adamidis and Loukas [3], the Exponential Geometric model (EG), was conceived to model phenomena in function of time with three characteristics: (I) the time minimum t with (II) exponential evolution which (III) marks the n repetitions required until a failure occurs.

To model this phenomenon, the authors considered the auxiliary random variables $X \sim Exp(\lambda)$ and $N \sim Geo(\theta)$ representing, respectively, (II) any time with exponential evolution parameterized in $\lambda > 0$ and (III) the number of Bernoulli attempts required to achieve failure in analysis.

As the objective is to obtain (I) a model based on the minimum t time, the composite model (or baseline distribution as highlighted by Tahir and Cordeiro [63]) should be represented by the random variable defined as $T = \min(\{X_i\}), \forall i = 1, \ldots, n$.

In these conditions, the Exponential Geometric model comes from a population composed of 3 subpopulations, where each of them represents a particular characteristic, that is, for all $\lambda, x \in \mathbb{R}_+^*$, $n \in \mathbb{N}$ and $0 < \theta < 1$:

$$\delta(x|\lambda) = \lambda e^{-\lambda x} \qquad \text{e} \qquad \varepsilon(n|\theta) = \theta(1-\theta)^{n-1} \ . \tag{1.8}$$

As $T = \min(\{X_i\})$ has the distribution to the minimum of the set $\{X_i\}, \forall i = 1, \ldots, n$, due to the $X_i \sim Exp(\lambda)$ being independent and identically distributed, we obtain that $T \sim Exp(n\lambda)$, $\forall\ n\lambda, t \in \mathbb{R}_+^*$, so that

$$\delta_{\min}(t|n\lambda) = n\lambda e^{-n\lambda t} \ . \tag{1.9}$$

Then, if T is the baseline distribution of the composite model between it the discrete random variable N, noting that $e^{-n\lambda t} = e^{-\lambda t}e^{-(n-1)\lambda t}$, we can write

$$g(t|\theta,\lambda) = \sum_{n=1}^{+\infty} \delta(n|\theta)\epsilon(t|n\lambda) = \lambda\theta e^{-\lambda t} \sum_{n=1}^{+\infty} n[(1-\theta)e^{-\lambda t}]^{n-1} \ , \tag{1.10}$$

where $|(1-\gamma)e^{-\lambda t}| < 1$ induces the convergence of the infinite series indexed in n, so that

$$\sum_{n=1}^{+\infty} n[(1-\gamma)e^{-\lambda t}]^{n-1} = \frac{1}{[1-(1-\gamma)e^{-\lambda t}]^2} \ . \tag{1.11}$$

Therefore, the model composed of random variables $T \sim Exp(n\lambda)$ e $N \sim Geo(\theta)$, for $n\lambda, t > 0$ e $0 < \theta < 1$, respectively, has a probability density function given by

$$g(t|\theta,\lambda) = \frac{\lambda\theta e^{-\lambda t}}{[1-(1-\theta)e^{-\lambda t}]^2} \ , \tag{1.12}$$

for all $t, \lambda \in \mathbb{R}_+^*$ e $\theta \in \ ]0,1[$.

Many probabilistic distributions have been proposed as composite models based on the Geometric distribution, in addition, generalizations for these models are acquired and, as such models consider the parameter $\theta \in \ ]0.1[$ inherited from the geometric distribution, extensions for the composite models and their generalizations are also discovered when considering the complement of θ in $\mathbb{R}_+^*$.

As well as defined for the Exponential and Weibull distributions, the models derived from the EG distribution can also be understood as a family formed by the models Extended Geometric Exponential (EEG), Generalized Exponential Geometric (GEG), Complementary Exponential Geometric (CEG), Long-Term Complementary Exponential Geometric (LCEG), Complementary Exponentiated Exponential Geometric (CE2G), Exponentiated Exponential Geometric (E2G), Marshall-Olkin Generalized Exponential (MOGE) and the Generalized Exponential Geometric Extreme (GE2) distributions, proposed respectively by Adamidis, Dimitrakopoulou and Loukas [2], Silva, Barreto and Cordeiro [60], Louzada, Roman and Cancho [37], Louzada et al. [33], Louzada, Cancho and Carpenter [34], Louzada, Marchi and Roman [35], Ristic and Kundu [54] and Ristic and Kundu [55].

The family of distributions for the EG model can also be classified as distributions of two generations of models, the first with 2 parameters and the second with 3, where the first generation starts with the EG model and ends with the EEG model and the second generation starts with the E2G model and ends with the MOGE and GE2 models.

In Ristic and Kundu [54] and Ristic and Kundu [55] it is verified that, although the generalized models MOGE and GE2 (MOGE/GE2) are identified with different names and proposed on different dates, they are obtained under the same conditions, have the same parametric states and have the same expression for the distribution and probability density functions. Both works consider random variables strictly positive and parameterized in α, λ e β defined in $\mathbb{R}_+^*$.

Important extensions or generalizations were, and can be, introduced in the statistical literature, however, the EG distribution was, according to the records of Adamidis and Loukas [3], Gupta and Kundu [20] and Tahir and Cordeiro [63] the first obtained to model rates of decreasing risk under two parameters and started the modern era of probability distribution theory with the main objective of proposing and emphasizing the solution of problems faced by professionals and researchers in practically all areas of knowledge.

1.2 General Objectives

The model EEG, as previously discussed, is the extension of the EG distribution and has the function of probability density given by

$$g(t|\gamma,\lambda) = \frac{\gamma\lambda e^{-\lambda t}}{[1-(1-\gamma)e^{-\lambda t}]^2}\ , \tag{1.13}$$

for all $t, \gamma, \lambda \in \mathbb{R}_+^*$, as shown in Adamidis, Dimitrakopoulou and Loukas [2], where the decreasing and increasing forms for its function of derived hazard rate are presented. Above all, such a model can be applied in situations where the hazard rate shows little or no variation after a considerable period of evolution, as such behavior leads to the adoption of a risk model whose form captures none or the slow increase or decrease in the hazard rate.

That such form for the hazard rate can be captured by the Gamma and Weibull distributions, but in contrast to these cases for EEG, the EEG distribution may be much more appropriate due to its guaranteed flexibility in its basic properties that can be reflected for the computational gain when applied to adjust complete and censored lifetime data.

In Adamidis, Dimitrakopoulou and Loukas [2], the parameter estimation is performed using the classic approach via maximum likelihood estimators, in addition, analytical expressions for the mean and variance of the distribution are presented and the work implies that in that approach to the moments of distribution, such expressions are defined only in the case where $\gamma < 1$. The expectation maximization algorithm (EM) is presented to calculate the estimates of the parameters of this distribution, and in Ramos et al. [51] a bayesian analysis is developed considering complete data with different previous noninformative distributions, but in both works nothing is clarified about the r-th moment in the event that $\gamma > 1$.

Although the distribution is characterized by its mean residual lifetime for all $\gamma \in \mathbb{R}-\{1\}$, as the first objective of this work we will present a study for the moments of the EEG distribution in the general case, specifically when $\gamma \geq 1$.

It is obvious that, since its disclosure, the EEG distribution is already consolidated in the works identified in the analysis of survival and reliability, since under the consideration of censorship, as shown in the works Anwar et al. [4], Mirjalili [42], Dey et al. [11] and Abujarad et al. [1], respectively, with a type-II, progressive type-II, progressive type-I hybrid and censorship on the right approach, it is verified that this distribution is applied in a refined way in relation to the types of censorship considered both in the context of survival and reliability.

However, the approach according to random censorship is not found in the literature, and in none of these lines of application, therefore the second objective of this study is to obtain the estimates for the parameter assuming different estimation methods for the distribution parameters of EEG under randomly censored samples. In this case, the main objective is to study the bayesian estimates considering different noninformative prior distributions and to contrast them with the maximum likelihood estimate.

As the bayesian estimates and marginal posterior densities cannot be obtained in a closed form, we carry out a bayesian analysis for the EEG distribution using Markov Chain Monte Carlo (MCMC) methods (see , Gelfand and Smith [15]; or Chib and Greenberg [8]) to obtain the posterior summaries of interest. Since bayesian analysis remains a new approach proposal for this distribution, as verified by the main works cited here, in the present work we limit ourselves to applying inferences in this context only under non-informative priori.

Numerical integration based on stochastic simulation methods as the MCMC will be

used to simulate samples of the marginal posterior distribution of interest and in particular, we will be using the Metropolis-Hastings (MH) algorithm to obtain the posterior summaries of interest.

As a final proposal for an approach to this distribution, we will consider the presence of covariates in conjunction with random censorship and with the objective of presenting the characteristics of this distribution under the presence of this double influence, in this approach we consider it essential to develop the same approach used for the distribution in the case randomly censored and without covariate, that is, the approach to the moments of the EEG distribution in the presence of random and convariable censorship, as well the obtaining their estimators will follow the same conditions, with the estimators obtained in the cases classic and bayesian in order to be contrasted possible applications.

With similar objectives, in this work we will also consider the distribution that generalizes the second generation of the EG distribution, as presented in the previous section, the MOGE-GE2 distributions, whose probability density function is given by

$$f(t|\alpha,\lambda,\beta) = \frac{\alpha\lambda\beta e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}}{[\beta+(1-\beta)(1-e^{-\lambda t})^{\alpha}]^2}\ , \tag{1.14}$$

for all $t,\alpha,\lambda,\beta \in \mathbb{R}_+^*$.

As it is highlighted in this work that the MOGE/GE2 distributions are the same and are being approached as members of the EG family, we will from here on follow the common notation that the literature presents for the EG distribution family, as discussed in the previous section, we will follow referring to the MOGE/GE2 distributions only as GE2.

From the works of Ristic and Kundu [54] and Ristic and Kundu [55] to the most recent, such as the current ones Khan, Akhtar and Ali Khan [45], Dey et al. [11], Torabi, Bagheri and Mahmoudi [64] and Abujarad et al. [1], there is no significant approach to the moments of the GE2 distribution.

Although in Khan, Akhtar and Ali Khan [45] the authors have announced a study in the light of bayesian inference for Marshall-Olkin family of distributions, which has admirably addressed the moments in the shadows of Laplace Approximation, this is did for Marshall-Olkin Exponential (MOExp) and no reference is given for the GE2 distribution.

In Dey et al. [11], although a study is presented for the model considered through the refined application of type-II progressive censorship, the work considers an approach to the Marshall-Olkin Extended Exponential (MOEE) distribution and is also dont't move of no measure to the GE2 model.

In the work of Torabi, Bagheri and Mahmoudi [64], despite the fact that the authors actually approach the GE2 model, they do so with the parameters strictly limited to $\alpha = 1$ and $\beta \in]0,1[$ for all $\lambda > 0$, reducing the model $GE2(1,\lambda,\theta)$ for the GE distribution. Therefore, the authors present a study that is limited to an application for the GE distribution to complete data, where there is no study for the models of the GE2 distribution in the general case.

In the most recent work, presented by Abujarad et al. [1], although the GE2 distribution proposed by Ristic and Kundu [54] and Ristic and Kundu [55] is indeed explored, the authors invoke this model under the reparametrization $\lambda = \theta^{-1}$ for all $\theta > 0$, and insert a third name for this distribution as Marshall-Olkin Exponentiated Exponential (MOExp-Exp) and develop the approach, together with the MOExp models and the Marshall-Olkin Exponential Extension (MOExpExt), in line with a simulation study for censored data on the right, with application to real censored data. However, they also do not mention anything about the moments of the three models covered.

From what the literature provides, there is a low, or negligible, approach to the GE2

distribution, presenting a reasonable characteristic for this model. Therefore, in this work we bring a Mathematical approach on the r-th moment of this distribution, in order to characterize it as a probability distribution.

In addition, we will present an approach under random censorship for its estimators and, under the same conditions described above on the proposed objectives for the EEG distribution approach, this study will be developed considering the classical and bayesian estimators with vague prioris, with applications to available real data in the literature on survival analysis.

Both in the approach of the EEG distribution and in the GE2, the study of their respective estimators will be developed computationally by the justification that these models are semi-pathological, as we will show in the approach of their respective moments.

Therefore, we consider a simulation study for both models and following the forms of their respective hazard rate functions, that is, we will simulate the EEG in the parametric condition in which its risk is decreasing and increasing, while for the GE2 model the simulation it is performed for cases in which it manifests the decreasing, increasing, unimodal and bath shape for its hazard rate function.

1.3 Dissertation Structure

In detail, this work is organized as that is describe in the sequence.

In the chapter 2, in the section 2.1, the properties of the EEG model are reviewed in the context of survival analysis. A lemma that guarantees obtaining the mean and variance for this model, in the conditions presented by the author in Adamidis, Dimitrakopoulou and Loukas [2] and, consequently, a theorem that reinforces the semi-pathology result for EEG when $\gamma > 1$ são demonstrados, as well as a theorem that guarantees that this model tends to lose memory as time progresses, a result hitherto commented on previous works and, almost always, highlighted geometrically in the approaches for this distribution family.

In the section 2.2, the maximum likelihood estimators for the parameters are presented, as well as a reference to Fisher's information matrix obtained by Kitidamrongsuk et al. [30] in the specific case in which $\gamma < 1$ is described. Subsequently, a bayesian approach to this model is introduced under the application of vague prioris as a result of the product of independent Gamma distributions for its joint composition.

In the subsection 2.2.3, we present the simulation for the EEG model in the presence of random censorship, in 3 particular cases for this category of censorship, with 4 different cases of sample size and two parametric cases to simulate the manifestation of these scenarios with the two forms of the function of risk that this model assumes.

The subsection 2.2.4 describes an application for the EEG model in the presence of censorship for a modeling problem with a set of 137 observations from patients with lung cancer, where the parametric models Gamma, LL and Weibull are taken in competition with to the EEG model for modeling under the classical conditions of the model estimators. A random sample of 40 observations was also considered for modeling, however, under bayesian conditions to highlight a contrast between these two approaches for the model's estimators.

The section 2.3.2 presents the proposed distribution approach in the presence of covariates and censored times, so that a second source of variation is mathematically defined for the density function and, consequently, defined the survival and risk function highlighting that with two sources of variation these three functions geometrically represent surfaces in the space $\mathbb{R}^3$. Still in the section 2.3.2, a theorem is presented to show that the

property of tendency to memory loss also manifests itself in the presence of covariates, however, in this context, the risk function tends for a curve in the $\mathbb{R}^3$.

Following the theorem, the maximum likelihood equation, as well as its system of maximum likelihood equations, for the $p+1$ parameters of the model in the presence of censorship are presented without significant change, maintaining the same properties that in the case considered with a source of variation as shown in the analytical expression presented in the section 2.2, which does not occur in the bayesian approach that is introduced with vague priori and also as the product between the independent Gamma distribution for the parameter γ and the Normal distribution for each of the $p+1$ parameters β_j in the covariate term.

In the subsection 2.3.1, the simulation for the EEG model in the presence of covariates and censored time is also developed in 3 particular cases of random censorship, with 4 cases of different sample size and two parametric cases for simulate the model according to forms for the risk function that remain in the same conditions as in the case of a source of variation, regardless of the composition of the second source in the model.

Subsequently, the subsection 2.3.2 deals with application that seeks to adjust survival and risk models for two groups of patients undergoing different lung cancer treatments. In this application, a group presents proportional risk, which is why the Cox regression model was used to contrast the adjustment implemented by the EEG model in the presence of censored and covariable data.

The application shows that, despite the flexibility that the Cox model offers, although improved with the application of the cubic spline to smooth the function of survival and risk, as we consider in two nodes, the EEG model still presented better results, adjusting efficiently to the empirical risk manifested by the censored data. In the second group, the same performance is observed in contrast to the Weibull model.

In the chapter 3, a generalization for the EEG model is reviewed in the theoretical context with a simulation and an application presented. This chapter has been divided into 2 sections.

In the first, the section 4.1 seeks to review the origin of this three-parameter model, highlighting its origin and consequent functions in the context of survival analysis, where the descending, crescent, unimodal and bathtub shapes are notable as illustrated in the figure 3. 1 displayed.

In the following subsection, in 3.1.1, a significant mathematical approach is presented for the function of r-th moment for the approached distribution, presenting as results three mottos, a proportion and a theorem to highlight a study of the moment of r-th order of this model, considering its possible parametric compositions and its memory loss tendency property inherited from its original distribution.

In the subsection 3.1.2 we will also see, in the classic aspect, the estimators for the parameters of this model and, in the same way, in the subsection 3.1.3, the bayesian approach is considered under vague prioris through the product independent range distributions.

In the sequence, the section 3.2 displays and discusses the results obtained in the simulation study that considered three particular cases of censorship, with five different sample size cases and four parametric cases to simulate and evaluate the estimators of this model approached according to each of the forms of risk function that the model provides.

The sequence application, shown in the section 3.3, again shows the competition between the three-parameter model approached with competing models. The models considered were obtained under the same conditions as the one approached, making the contrast in equality not only due to the parametric quantity and composition of the hazard rate

function shape, but because they have different conditions of flexibility and analysis.

In this work, we try to show that the relevance of the EEG and GE2 models does not consist only of their good adjustments in statistical modeling applications, but that their relevance is highlighted due to the fact that their properties inherited from the Exponential and Geometric distributions, that this relevance consists of the fact the compositions, extensions or generalizations of models derived from Exponential and Geometric distribution. That although most of its analytical results are intractable as a result of the introduction of shape parameters, the flexibility inherited from its source distributions injects highly relevant benefits to the computational approach in real life-time data.

1.4 Other Considerations

The bayesian procedures include several statistical diagnostic tests witch seek to assess the convergence of the markovian chain, this work consider the Heidelberger-Welch stationarity test, known as the heidel diagnostics, which uses the Cramer-von-Mises statistic to test the hypothesis that the chain values provide a stationary distribution.

The diagnosis of heidel is based on the work of Heidelberger and Welch [24], Schruben [59] and heidelberger and Welch [25], and the results of this test with a given significance will shown in the tables by the columns test-stat, p-value, and test result columns, for show that all Markov chains from MCMC processes converged or not.

Since we consider adjustments by the performed using bayesian estimates, to compare these adjustments the most appropriate adjustment measure is the DIC (Deviance Information Criterion) which generalizes the Akaike, AAIC and Schwarz Information Criterion, respectively the AIC, AICC and BIC, to evaluate fitted models with calculated sample estimates via posterior distributions.

In the context of model comparison by the DIC measure, the best-fit model is indicated by penalties that the others suffer through the effective number of parameters and deviance, the statistics commonly indicated by p_D and D, respectivaly, and obtained by summing the squares of the linear regression residues which, according to Collett et al. [9], summarizes to what extent the fit of a current model of interest diverges from a model that is assumed to be a perfect fit to the data.

Suitable for selection problems of bayesian models in which the posterior distributions of the models are obtained via MCMC simulation, when the p_D decreases smaller is the complexity of the model andthe DIC decreases by pointing the best fit, then, the lower p_D and DIC, the better is adjustment. Then, defined

$$\bar{D}(\theta) = E[-2\log(p(y|\theta))] \quad \text{and} \quad D(\hat{\theta}) = -2\log(p(y|\hat{\theta})) \ , \tag{1.15}$$

from calculation of p_D as $p_D = \bar{D}(\theta) - D(\hat{\theta})$, the DIC can be obtained in two ways: according to Spiegelhalter et al. [61] as $DIC = p_D + \bar{D}(\theta)$, or according to Gelman et al. [16] as $DIC = 2p_D + D(\bar{\theta})$.

In the applications from the sections 2.2.4 and 2.3.2, purposing to fit an ideal probabilistic model for the censored lifetimes for the patients according to two chemotherapeutic agents, standard and test, based on the absence and presence of covariates, respectively, a same set of data is considered to discuss possible approaches and results around probabilistic modeling and to point out the benefits and penalties in modeling with data with and without variables.

The data was presented by Prentice [48], record the lifetime of 137 patients with advanced lung cancer and was recorded in the presence of covariates when the patient was taken to the study. Tumors are classified into four groups: squamous (1), small (2), adeno (3) and large (4).

The covariates considered in original study include performance status, a measure of overall medical status for the patient on a scale of 10 out of 10 units from 10 to 90, where 10, 20, and 30 indicate that the patient is fully hospitalized, 40, 50 and 60 indicate a partial confinement in the hospital and 70, 80 and 90 indicate that the patient is able to take care of himself.

Time in months from diagnosis to study initiation, age in years, and any therapy prior to the study (1-yes ou 0-not) were also considered with as covariates in study 2.

In applications, the objective is to model the lifetime of standard and test chemotherapeutic agents and, in the case of the presence of a vector $\boldsymbol{x}$ of covariates, is considers the tumor group as a new covariate. So, denoting as X_1, X_2, X_3, X_4 and X_5 the covariates of the study, where

- X_1: the tumor group with values 1, 2, 3 or 4
- X_2: the medical status with values 10, 20, ..., 80 or 90
- X_3: diagnostic time in months with values $x_{i,3} \in \mathbb{Z}_*^+$
- X_4: the age in years with values $x_{i,4} \in \mathbb{Z}_*^+$
- X_5: if you have had therapy with values 1 or 0

The covariate approach is developed with the covariates indexed in the scale parameter of the model under study. This form the presence of covariates in the probabilistic model integrates into its variability a source of variation with origin in additive effect index parametric that not require the scale parameter coupled in the model as a random variable, but a simple source of variation that integrates effect the covariate in the model.

Therefore, in this approach, the modeling concept considers that an additive model $\nu = \langle \boldsymbol{x}_k, \boldsymbol{\beta} \rangle + \epsilon$ contributes to the probabilistic variation of the model, together with the variable random that represents the lifetime, but as the scale parameter instead of being coupled in the model as a random variable.

The commonly used models that consider the presence of covariates, generally statistical models of regression or joint probability distribution, although powerful, require a delicate or complex approach before, during and after the modeling process.

To circumvent this situation and also find a more flexible model, Cox [10], proposed a semi-parametric model called the proportional risk model that became the most used in the analysis of survival data, under the covariable aspect, due its versatility. The general expression of the Cox model, for $\nu = \langle \boldsymbol{x}_k, \boldsymbol{\beta} \rangle$, is given by

$$h(t|\boldsymbol{x}, \boldsymbol{\beta}) = h_0(t)\phi_{\boldsymbol{\beta}}(\nu) \ , \tag{1.16}$$

on what $h_0(t|\boldsymbol{\beta})$ is the nonparametric component called the hazard rate function and $\phi_{\boldsymbol{\beta}}(\nu)$ is the parametric term, both non-negative on the condition that $\phi_{\boldsymbol{\beta}}(0) = 1$.

The parametric term is almost always applied in the multiplicative form as

$$\phi_{\boldsymbol{\beta}}(\nu) = e^{\nu}. \tag{1.17}$$

With the same objective, to make the survival data modeling more flexible, the use of spline functions is proposed, since this function also allows the investigation of nonlinear effects and the evaluation of time interactions in the presence the covariates. According to Heinzl and Kaider [26], modeling under the application of spline functions provides two main advantages: no specific functional form needs to be specified and its application is completely computational, excluding applications of algebraic concepts.

In addition, Durrleman and Simon [12] proposes the use of spline functions to detect nonlinear relationships between covariates and Cox model response, and in the same vein are used in Hess [27] and Heinzl, Kaider and Zlabinger [22] to evaluate and identify interactions between time and covariates.

As if the flexibility that Cox's model itself provides does not suffice, using spline functions on the Cox model, it is possible to obtain a parametric form from and which simulate survival times with adjust survival and risk functions with refined precision.

To obtain a flexible model for the hazard rate, Royston and Parmar [57] modeled the logarithm of the risk-risk function as a cubic spline function of the time logarithm as

$$ln[H(t|\boldsymbol{x},\boldsymbol{\beta})] = ln[H_0(t)] + \nu \ , \tag{1.18}$$

where H is the empirical cumulative risk function, $\boldsymbol{x}$ is the vector of covariates, $\boldsymbol{\beta}$ is the vector of parameters and $\nu = \langle \boldsymbol{x}_k, \boldsymbol{\beta} \rangle$ is the additive model.

In this context, according to Lambert and Royston [31], the s spline function is applied over the $ln(t)$ function with the nodes $\kappa_0 = (\kappa_1, \ldots, \kappa_m)$. In this case s can be written as $s[ln(t)|\boldsymbol{a}, \kappa_0]$, with node location κ_0 no coefficiente $\boldsymbol{a} = (a_1, \ldots, a_m)$, and then used for the logarithm of the accumulated baseline risk in the proportional accumulated risk model of the expression (1.16) to obtain (1.18) rewritten as

$$ln[H(t|\boldsymbol{x},\boldsymbol{\beta})] = s[ln(t)|\boldsymbol{a}, \kappa_0] + \nu \ , \tag{1.19}$$

that taken on the scale of survival and hazard rate, provides

$$S(t|\boldsymbol{x},\boldsymbol{\beta}) = e^{-e^{\nu}} \quad \text{and} \quad h(t|\boldsymbol{x},\boldsymbol{\beta}) = \left\{ \frac{d}{dt} s[ln(t)|\boldsymbol{a}, \kappa_0] \right\} e^{\nu} \ , \tag{1.20}$$

because the s spline function is a cubic polynomial and the relations between the functions h and S, in particular with $h(t) = H'(t)$, $S(t) = e^{-H(t)}$ and $H(t) = -ln[S(t)]$ under derivative and integral operators in relation to t.

Thus, according Rutherford, Crowther and Lambert [58] over the Cox models, the spline function can be considered a generalization of Weibull's proportional hazards rates model and is reduced to the Weibull model with only two nodes.

Therefore, in this application a contrast between the EEG model in the presence of covariates with the Cox/Spline$_2$ model, the Cox model under two-node spline function adjustment, it is considered.

In the application 2.3.2 the covariables X_3 and X_4 are integers, the 5 covariables of the problem can be considered categorical, therefore, graphical methods to find proportional risks are not feasible for the application, instead they are considered a statistical test based on Schoenfeld residues to test the overall proportionality for risk functions.

Chapter 2

The Extended Exponential Geometric Distribution

2.1 Definition and Properties of the EEG Distribution

In the previous chapter, the expression (1.12) is obtained as the probability density function (pdf) of a random variable considered as the lifetime of a n identical component system, where the failure occurs when all components leave functioning, this is, if the system's lifetime is $T = \min(\{X_i\}), \forall i = 1, \ldots, n$.

Here, if X the lifetime of the components follows an independent Exponential random variable and the number of components N follows a Geometric distribution, the distribution the composition of the random variables T, X and N provides the model probabilistic $T \sim EG(\theta, \lambda)$, where $0 < \theta < 1$ and $\lambda > 0$.

Then, in the same conditions as before, but considering $T = \max(\{X_i\})$ we'll have to

$$\delta_{\max}(t|n, \lambda) = n\lambda e^{-\lambda t}(1 - e^{-\lambda t})^{n-1} \ , \tag{2.1}$$

is the pdf of the random variable T, whose composed model with the random variable $N \sim Geo(\theta)$ has pdf

$$\psi(t|\theta, \lambda) = \frac{\lambda\theta e^{-\lambda t}}{[\theta + (1 - \theta)e^{-\lambda t}]^2} \ , \tag{2.2}$$

obtained under the same algebraic conditions as the expression (1.12).

Taking $u = \theta + (1 - \theta)e^{-\lambda t}$ with $du = -\lambda(1 - \theta)e^{-\lambda t}dt$, by variable change in the integral of ψ with respect to t, without difficulties it turns out that ψ is a pdf.

Now, taking the reparametrization $\vartheta = \theta^{-1}$ see that $0 < \theta < 1 \Rightarrow \vartheta > 1$ and that under this reparametrization the pdf (2.2) boils down to pdf (1.12) under parameter ϑ. This is, the distribution for random variable T with pdf (2.2) is the complement to the EG model because, with $\vartheta = \theta^{-1} \Rightarrow \theta = \vartheta^{-1}$, (2.2) can be rewritten under the parameter ϑ as

$$\psi(t|\vartheta^{-1}, \lambda) = \frac{\lambda\vartheta e^{-\lambda t}}{[1 - (1 - \vartheta)e^{-\lambda t}]^2} = g(t|\vartheta, \lambda) \ . \tag{2.3}$$

And more, under the parameters $\vartheta > 1$ and $0 < \theta < 1$, we have that g and ψ are well defined and so that

$$g(t|\theta, \lambda) = \psi(t|\theta^{-1}, \lambda) \Leftrightarrow \psi(t|\vartheta^{-1}, \lambda) = g(t|\vartheta, \lambda) \ . \tag{2.4}$$

Thus, noting that g and ψ are both defined in $\vartheta = \theta = 1$, considering the parameter $\gamma \in \mathbb{R}_+^*$, by (1.12) and (2.2) we will have to

$$\frac{\lambda\gamma e^{-\lambda t}}{[1-(1-\gamma)e^{-\lambda t}]^2} = \begin{cases} g(t|\gamma,\lambda), \text{ if } \gamma < 1 \\ \psi(t|\gamma,\lambda), \text{ if } \gamma > 1 \\ \lambda e^{-\lambda t}, \text{ if } \gamma = 1 \end{cases} . \tag{2.5}$$

Therefore, the pdf (2.3) is the extension of the model $EG(\theta,\lambda)$ with pdf (1.12) and is a distribution called Extended Geometric Exponential and denoted by $EEG(\gamma,\lambda)$, or simply by EEG, and was introduced by Adamidis, Dimitrakopoulou and Loukas [2].

Let then T be a random variable that represents the lifetime of any phenomenon and that follows an EEG distribution. In this conditions we have to say that T has pdf given by

$$g(t|\gamma,\lambda) = \frac{\lambda\gamma e^{-\lambda t}}{[1-(1-\gamma)e^{-\lambda t}]^2} , \tag{2.6}$$

for all $t,\gamma,\lambda \in \mathbb{R}_+^*$ where λ is shape and γ is form parameters, respectively.

The distribution function $G(t)$ is given by,

$$G(t|\gamma,\lambda) = \frac{1-e^{-\lambda t}}{1-(1-\gamma)e^{-\lambda t}} . \tag{2.7}$$

The survival and hazard rate functions of EEG distribution are given by

$$S(t|\gamma,\lambda) = \frac{\gamma e^{-\lambda t}}{1-(1-\gamma)e^{-\lambda t}} \text{ and } h(t|\gamma,\lambda) = \frac{\lambda}{1-(1-\gamma)e^{-\lambda t}} , \tag{2.8}$$

respectively.

The quantile function of the EEG distribution, for all $q \in]0,1[$, is given by

$$Q_{EEG}(q|\gamma,\lambda) = -\frac{1}{\lambda}\ln\left[\frac{1-q}{1-(1-\gamma)q}\right] . \tag{2.9}$$

This result is trivial and is obtained by expression (2.7) considering that $G(t|\gamma,\lambda) = q \in]0,1[$, so that $Q_{EEG}(q|\gamma,\lambda) = G^{-1}(q|\gamma,\lambda)$.

Many of the interesting characteristics and features of a distribution can be studied through its moments, such as mean and variance. Expressions for expectation value, variance and the r-th moment on the origin of T can be obtained using the well-known formula

$$E(T^r|\gamma,\lambda) = \int_0^\infty \frac{\lambda\gamma t^r e^{-\lambda t}}{[1-(1-\gamma)e^{-\lambda t}]^2} dt . \tag{2.10}$$

The rth moment provides the most important properties of a probabilistic model, so much so that the characterization of probability distributions, where possible, is indispensable and defined by expression de $E(T^r)$, that is commonly approach in terms of f.d.p. as

$$E(T^r|\alpha,\lambda,\beta) = \int_0^\infty t^r f(t) dt . \tag{2.11}$$

Through this expression, in many cases, obtaining this result is not feasible due to the lack of an elementary primitive for the integrant $t^r f(t)$, mainly for obtaining high order

moments, such as for obtaining asymmetry and kurtosis of model. In the best case, the desired expression is obtained after costly mathematical devices.

However, in cases of models with non-negative random variables a device that facilitates obtaining expressions of order moments $r \geq 1$ is derived in terms of the survival function in parallel with the application of Fubini's theorem, the alternative rth moment given by

$$E(T^r|\alpha,\lambda,\beta) = r\int_0^\infty t^{r-1}S(t)dt\ . \tag{2.12}$$

It is a highly advantageous method when it comes, for example, to moments of a transformed random variable, where it is considerably easier to integrate $x^{r-1}[1-F(x)]$ instead of $t^r f(t)$ (see, for exemple, Hong [28] or Chakraborti, Jardim and Epprecht [7]). Armed with this device, we can proof the result below.

Lemma 1 *For a random variable T with $EEG(\gamma,\lambda)$ distribution, where $\gamma,\lambda > 0$, we have that r-th moment function, for all $r \geq 1$ it's such that*

$$E(T^r|\gamma,\lambda) = \begin{cases} \dfrac{\gamma r!}{\lambda^r}\displaystyle\sum_{k=0}^{\infty}\frac{(1-\gamma)^k}{(k+1)^r}, when\ \gamma \neq 1 \\ \dfrac{r}{\lambda}E(T^{r-1}|1,\lambda),\ when\ \gamma = 1 \end{cases}\ , \tag{2.13}$$

whatever it is $\lambda > 0$.

Proof: Taking

$$E(T^r|\gamma,\lambda) = r\int_0^\infty x^{r-1}S(x)dx\ , \tag{2.14}$$

the expression for order moments $r \geq 1$, derived from the survival function as the application of Fubini's theorem, we have from (2.8) that

$$E(T^r|\gamma,\lambda) = \gamma r\int_0^\infty \frac{t^{r-1}e^{-\lambda t}}{1-(1-\gamma)e^{-\lambda t}}dt\ . \tag{2.15}$$

Since γ and λ are parameters in $E(T^r|\gamma,\lambda)$, $(1-\gamma)e^{-\lambda t}$ varies in t and is dominated by $e^{-\lambda t}$, we have always one t about γ and λ such that $|(1-\gamma)e^{-\lambda t}| < 1$.
In this condition, we can take

$$\frac{1}{1-(1-\gamma)e^{-\lambda t}} = \sum_{k=0}^{\infty}(1-\gamma)^k e^{-\lambda tk}\ , \tag{2.16}$$

that which replaced in (2.15), follow that

$$E(T^r|\gamma,\lambda) = \gamma r\int_0^\infty \frac{t^{r-1}e^{-\lambda t}}{1-(1-\gamma)e^{-\lambda t}}dt = \gamma r\sum_{k=0}^{\infty}(1-\gamma)^k\int_0^\infty t^{r-1}e^{-\lambda t(k+1)}dt\ . \tag{2.17}$$

See also that, with $r, \lambda(k+1), t > 0$ in the integral resulting in (2.17), we also have to

$$\int_0^\infty t^{r-1}e^{-\lambda t(k+1)}dt = \frac{\Gamma(r)}{[\lambda(k+1)]^r} = \frac{(r-1)!}{\lambda^r(k+1)^r}\ . \tag{2.18}$$

Therefore, replacing (2.18) in (2.17), the first equality for the expression (2.13) is obtained, because

$$\gamma r \sum_{k=0}^{\infty}(1-\gamma)^k \int_0^{\infty} t^{r-1}e^{-\lambda t(k+1)}dt = \frac{\gamma r!}{\lambda^r}\sum_{k=0}^{\infty}\frac{(1-\gamma)^k}{(k+1)^r} \ , \tag{2.19}$$

the proof is completed for $\gamma \neq 1$.

Now, note that then $\gamma = 1$ the function $E(T^r|\gamma,\lambda)$ is pathological because in the expression (2.19) result that $E(T^r|\gamma,\lambda) = 0$, but by the expression (2.6), where

$$g(t|\gamma,\lambda) = g(t|1,\lambda) = \lambda e^{-\lambda t} \ , \tag{2.20}$$

it turns out that the random variable $T \sim EEG(1,\lambda)$ is reduced for the Exponential distribution with parameter λ, this is, $T \sim EEG(1,\lambda) = Exp(\lambda)$ and

$$E(T^r|\gamma,\lambda) = E(T^r|1,\lambda) = \int_0^{\infty} t^r g(t|1,\lambda)dt = \int_0^{\infty} t^r \lambda e^{-\lambda t}dt \ , \tag{2.21}$$

such, by integration by parts, with $u = t^r$ and $dv = e^{-\lambda t}dt$, results

$$\int_0^{\infty} t^r \lambda e^{-\lambda t}dt = -t^r e^{-\lambda t}\Big|_0^{\infty} + \frac{r}{\lambda}\int_0^{\infty} t^{r-1}\lambda e^{-\lambda t}dt = \frac{r}{\lambda}E(T^{r-1}|1,\lambda) \ , \tag{2.22}$$

where $-t^r e^{-\lambda t}\Big|_0^{\infty} = 0$ in the rth application of L'Hospital theorem, or because that $e^{-\lambda t}$ grows faster than t^r, which demonstrates the lemma. ■

Now see that, proposed by Erdely [13] (see Jeffrey and Zwillinger [29], Guillera and Sondow [19] and Gradshteyn and Ryzhik [18]), the function Lerch transcendent which converges for any real number for all $|z| < 1$ and $s, a \in \mathbb{R}_+^*$, is given by

$$\Phi(z,s,a) = \sum_{k=0}^{\infty}\frac{z^k}{(k+a)^s} \ , \tag{2.23}$$

where, as one of its special cases, for $a = 1$, it is reduced to a polylogarithmic function given by

$$Li_s(z) = z\Phi(z,s,1) \ . \tag{2.24}$$

that, since $s \geq 0$, any of the following integral representations furnishes the analytic continuation of the polylogarithm beyond the circle of convergence.

With this result we will be able to prove the following theorem:

Theorem 1 *For a random variable T with $EEG(\gamma,\lambda)$ distribution, where $\gamma,\lambda > 0$, we have that r-th moment function, for all $r \geq 1$ it's such that*

$$E(T^r|\gamma,\lambda) = \begin{cases} \dfrac{\gamma r! Li_r(1-\gamma)}{\lambda^r(1-\gamma)}, when\ \gamma < 1 \\ \dfrac{r}{\lambda}E(T^{r-1}|1,\lambda),\ when\ \gamma = 1 \\ Pathological,\ when\ \gamma > 1 \end{cases} \ , \tag{2.25}$$

whatever it is $\lambda > 0$ in $E(T^r|\gamma,\lambda)$.

Proof: By the last term in the equality from the result (2.23), with $z = 1-\gamma$, $s = r$ and $a = 1$, since $0 < \gamma < 1$ we have

$$\sum_{k=0}^{\infty}\frac{z^k}{(k+a)^s} = \sum_{k=0}^{\infty}\frac{(1-\gamma)^k}{(k+1)^r} = \Phi(1-\gamma,r,1) \ , \tag{2.26}$$

defined by a power series in $1-\gamma$.

And more, in this case, by the result (2.24), also we have to $Li_r(1-\gamma) = (1-\gamma)\Phi(1-\gamma, r, 1)$, by which

$$\Phi(1-\gamma, r, 1) = \frac{Li_r(1-\gamma)}{1-\gamma} \ , \tag{2.27}$$

which proves the result (2.25) when $\gamma < 1$ in the first equality.

See also that the case $\gamma = 1$ is proved in the result (2.25) by the lemma (1), because it coincides with the Exponential distribution parameterized in λ. Then, if we show that result (2.25) diverges for all $\gamma > 1$ and $r > 1$, the result is proved.

Considering the general term $a_k = \dfrac{(1-\gamma)^k}{(k+1)^r}$ in the power serie (2.26), there $\phi(k) = (1-\gamma)^k$ is an exponential function of positive k exponent in the base $1-\gamma$ and $\psi(k) = (k+1)^r$ is a polynomial function in k of degree r, see that in a_k the enumerator grows faster than the denominator, that is, the general term function a_k is dominated by $\psi(\gamma)$ such that, as $\gamma > 1$, then $|a_k| > 1$ for all $k \in \mathbb{N}$ and $k \longrightarrow \infty$.

Therefore, the power serie (2.26) diverges alternately and this divergence can be guaranteed by the test of the general term (divergence test) considering that $(1-\gamma)^2 = \delta > 0$, because in this case there will be in the sequence $\{a_k\}$ the subsequence $\{a_{2k}\}$ such that

$$\lim_{k\longrightarrow\infty} a_{2k} = \lim_{k\longrightarrow\infty} \frac{(1-\gamma)^{2k}}{(2k+1)^r} = \lim_{k\longrightarrow\infty} \frac{\delta^k}{(2k+1)^r} \underset{r\ times}{\overset{L'Hospital}{=}} \lim_{k\longrightarrow\infty} \frac{\delta^k [ln(\delta)]^r}{r!} = \infty \ , \tag{2.28}$$

because for all $k \in \mathbb{N}$ and $\gamma > 1$, $(1-\gamma)^{2k} > 0$, which proves that there is no $E(T^r|\gamma, \lambda)$ for all $r \geq 1$ then $\gamma > 1$. ■

More specifically, and strictly, the pathologic case for the EEG distribution is guaranteed showing $\gamma > 1$ e $\lambda > 0$, when $T \sim EEG(\gamma, \lambda)$ then $E(T|\gamma, \lambda) = \infty$.

Note that a given function $\phi(y) = (1+y)^n$ can be written by Taylor expansion as

$$\phi(y) = (1+y)^m = \sum_{n=0}^{\infty} \binom{m}{n} y^n \ , \tag{2.29}$$

a binomial series that, with $y = -\gamma$, independent of the situation at the border of the radius disk $|y| = \gamma > 1$, can be rewritten in terms of the (generalized) binomial coefficients as

$$(1-\gamma)^k = \sum_{n=0}^{\infty} \binom{k}{n} (-\gamma)^n = \sum_{n=0}^{\infty} \frac{(-\gamma)^n k^{\underline{n}}}{n!} = \sum_{n=0}^{\infty} (-1)^n \frac{\gamma^n k^{\underline{n}}}{n!} \ , \tag{2.30}$$

where, for all $k, n \in \mathbb{N}$, $k^{\underline{n}}$ is the falling factorial, so that

$$k^{\underline{n}} = \frac{k!}{(k-n)!} = \begin{cases} 1, \text{ where } n = 0 \\ \displaystyle\prod_{i=0}^{n-1} (k-i), \text{ where } n \geq 1 \text{ and } n \leq k \end{cases} \ , \tag{2.31}$$

As $k, n \in \mathbb{Z}_+$ and $0 \leq n \leq k$, the ith first few falling factorials are such that

$$\begin{aligned} k^{\underline{0}} &= 1 \\ k^{\underline{1}} &= k(k-1) \geq 0 \\ k^{\underline{2}} &= k(k-1)(k-2) \geq 0 \\ &\ddots \\ k^{\underline{i}} &= \prod_{j=0}^{i-1} (k-j) \geq 0 \end{aligned}$$

where the equality $k^{\underline{n}} = 0$ occurs for all $k \leq i$ and inequality $k > 0$ occurs if $n = 0$ or $k > i$, this is, how $k, n \longrightarrow \infty$, exist a k in $\mathbb{Z}_+$ such that for all $k > i$ or $n = 0$, $k^{\underline{n}} \geq 1$ such that the lowest value for $k^{\underline{n}}$ is 1.

In this condition, replacing (2.30) in the power series (2.26) and considering the previous result such that $k^{\underline{n}} \geq 1$, we have

$$\sum_{k=0}^{\infty} \frac{(1-\gamma)^k}{(k+1)^r} = \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \frac{(-\gamma)^n k^{\underline{n}}}{n!(k+1)^r} \geq \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \frac{(-\gamma)^n}{n!(k+1)^r} \ , \tag{2.32}$$

where the smallest term in the inequality is such that

$$\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \frac{(-\gamma)^n}{n!(k+1)^r} = \left[\sum_{n=0}^{\infty} \frac{(-\gamma)^n}{n!}\right]\left[\sum_{k=0}^{\infty} \frac{1}{(k+1)^r}\right] = A_n B_k \ , \tag{2.33}$$

where A_n is a alternating series and B_k is a harmonic series of order r.

See that, $\forall\ n \in \mathbb{N}$, if $n \longrightarrow \infty$ then

$$\begin{aligned}
A_n &= \sum_{n=0}^{\infty} \frac{(-\gamma)^n}{n!} = 1 - \frac{\gamma}{1!} + \frac{\gamma^2}{2!} - \frac{\gamma^3}{3!} + \frac{\gamma^4}{4!} - \frac{\gamma^5}{5!} + \frac{\gamma^6}{6!} - \ldots + \frac{\gamma^{2i}}{2i!} - \frac{\gamma^{2i+1}}{(2i+1)!} + \ldots = \\
&= 1 + \frac{\gamma^2}{2!} + \frac{\gamma^4}{4!} + \frac{\gamma^6}{6!} + \ldots + \frac{\gamma^{2i}}{2i!} + \ldots - \gamma - \frac{\gamma^3}{3!} - \frac{\gamma^5}{5!} - \ldots - \frac{\gamma^{2i+1}}{(2i+1)!} - \ldots = \\
&= 1 + \frac{\gamma^2}{2!} + \frac{\gamma^4}{4!} + \ldots + \frac{\gamma^{2i}}{2i!} + \ldots - \left(\gamma + \frac{\gamma}{1!} + \frac{\gamma^3}{3!} + \frac{\gamma^5}{5!} + \ldots + \frac{\gamma^{2i+1}}{(2i+1)!} + \ldots\right) = \\
&= \sum_{i=0}^{\infty} \frac{\gamma^{2i}}{2i!} - \sum_{i=0}^{\infty} \frac{\gamma^{2i+1}}{(2i+1)!} = cosh(\gamma) - sinh(\gamma) = e^{-\gamma} \ ,
\end{aligned}$$

and also see that the serie B_k is divergent in the order $r = 1$, but converges for any $r \geq 2$.

Thus, it results in (2.33) that

$$A_n B_k = e^{-\gamma} \sum_{k=0}^{\infty} \frac{1}{(k+1)^r} = \begin{cases} \infty, & \text{when } r = 1 \\ M, & \text{when } r \geq 2 \end{cases} \ . \tag{2.34}$$

Therefore, from lemma (1), in the case where $\gamma > 1$, with the result (2.34), as $E(T^r|\gamma, \lambda)$ increases the result (2.34), this is, then $r = 1$ we can assume that there is no order moment $r = 1$, because

$$E(T|\gamma, \lambda) = \frac{\gamma}{\lambda} \sum_{k=0}^{\infty} \frac{(1-\gamma)^k}{k+1} \geq \frac{\gamma}{\lambda} A_n B_k = \infty \ , \tag{2.35}$$

which proves that the random variable $T \sim EEG(\gamma, \lambda)$ is pathological when $\gamma > 1$, because in this case there is no mean and variance for T, which demonstrates the result (2.25) the theorem.

For $1 - \gamma < 1$ in (2.24), for order $r = 1$ and $r = 2$, $Li_r(1-\gamma)$ is given, respectively, by

$$Li_1(1-\gamma) = -ln(1-\gamma) \quad \text{and} \quad Li_2(1-\gamma) = -\int_0^1 \frac{ln[1-(1-\gamma)y]}{y} dy \ ,$$

where Li_2 is a dilogarithm, a special function with analytical results for particular cases of $\gamma < 1$ and which, according to Loxton [38] and Zagier [65], will have $0 < \gamma < 1$ what $0 < Li_2(1-\gamma) < \pi^2/6$.

And more, for any order $r > 1$ in $E(T^r|\gamma, \lambda)$, the power serie the lemma (1) diverge

by divergence test for serie, because we have with r applications the L'Hospital theorem, the general term $a_k = \dfrac{(1-\gamma)^k}{(k+1)^r} \longrightarrow L \neq 0$ when $n \longrightarrow \infty$ and $\gamma > 1$.

Then, by the previous theorem, it is verified that the expected value and the variance presented in the works Adamidis, Dimitrakopoulou and Loukas [2], Kitidamrongsuk et al. [30] and Louzada, Ramos and PerdonÃ¡ [36] can be rewritten with

$$E(T|\gamma,\lambda) = \begin{cases} -\dfrac{\gamma ln(1-\gamma)}{\lambda(1-\gamma)}, \text{ when } \gamma < 1 \\ \dfrac{1}{\lambda}, \text{ when } \gamma = 1 \\ \infty, \text{ when } \gamma > 1 \end{cases} , \tag{2.36}$$

for $r = 1$ and, by divergence test in the power serie the lemma (1) with $\gamma > 1$, when $r = 2$

$$E(T^2|\gamma,\lambda) = \begin{cases} -\dfrac{2\gamma}{\lambda^2(1-\gamma)} \displaystyle\int_0^1 \frac{ln[1-(1-\gamma)y]}{y} dy, \text{ when } \gamma < 1 \\ \dfrac{2}{\lambda^2}, \text{ when } \gamma = 1 \\ \infty, \text{ when } \gamma > 1 \end{cases} . \tag{2.37}$$

These conditions, the mean for $EEG(\gamma,\lambda)$ distribution is trivial by the result (2.36), but when is considered the expression

$$Var(T|\gamma,\lambda) = E(T^2|\gamma,\lambda) - [E(T|\gamma,\lambda)]^2 . \tag{2.38}$$

the variance for $EEG(\gamma,\lambda)$ is given

$$Var(T|\gamma,\lambda) = \begin{cases} -\dfrac{2\gamma}{\lambda^2(1-\gamma)} \left\{ \dfrac{\gamma[ln(1-\gamma)]^2}{2(1-\gamma)} + \displaystyle\int_0^1 \frac{ln[1-(1-\gamma)y]}{y} dy \right\}, \text{ when } \gamma < 1 \\ \dfrac{1}{\lambda^2}, \text{ when } \gamma = 1 \\ \nexists, \text{ when } \gamma > 1 \end{cases} ,$$

that meets the recent work proposed by Louzada, Ramos and Perdoná [36].

In Zagier [65], the author further assures that there are exactly eight values in $[-1;1]$ for which Li_2 has analytical results, so that three are in $]0;1[$, and as $\gamma > 0$ is guaranteed then that when $\gamma < 1$, only three exact values for the expression Li_2 are analytically defined in $\mathbb{R}_+^*$.

Thus, exist and is analytically defined an expected value for the EEG distribution, although this occurs conditionally when $\gamma < 1$, important properties are attributed to this distribution for cases where this parametric condition is acquired.

Exponential and Geometric distributions are characterized by their memory loss property that, in particular, is associated with population parameter as a rate parameter in function of their respective λ and γ fixed parameters, respectively. The following theorem shows that, even though the EEG distribution has no finite mean, this distribution tends towards memory loss as time progresses.

Theorem 2 *Let $h(t|\gamma,\lambda)$ as obtained in (2.8). When $t \longrightarrow \infty$, then $h(t|\gamma,\lambda) \longrightarrow \lambda$, this is, tends to lose memory over time.*

Proof: When $t \longrightarrow \infty$, see that $e^{-\lambda t} \longrightarrow 0$. Then

$$\lim_{t\longrightarrow\infty} h(t|\gamma,\lambda) = \lim_{t\longrightarrow\infty} \frac{\lambda}{1-(1-\gamma)e^{-\lambda t}} = \frac{\lambda}{1-(1-\gamma)(\lim_{t\longrightarrow\infty} e^{-\lambda t})} = \lambda , \tag{2.39}$$

the which proves the theorem. ■

The previous theorem allows us to observe that in the context of modeling this distribution is indicated to model phenomena in which wear, or phenomenon of interest, manifests risk that tends to stabilize over time, ie the risk that an event of interest occurs is influenced until certain time.

Figure 2. 1 in the next page presents different forms for EEG distribution risk functions, considering different values of λ and γ. Moreover to observing the property shown in the theorem, we can also that for any $\lambda > 0$, the hazard rate function (2.8) is decreasing when $\gamma \in \ ]0, 1[$, constant for $\gamma = 1$ and increasing when $\gamma > 1$.

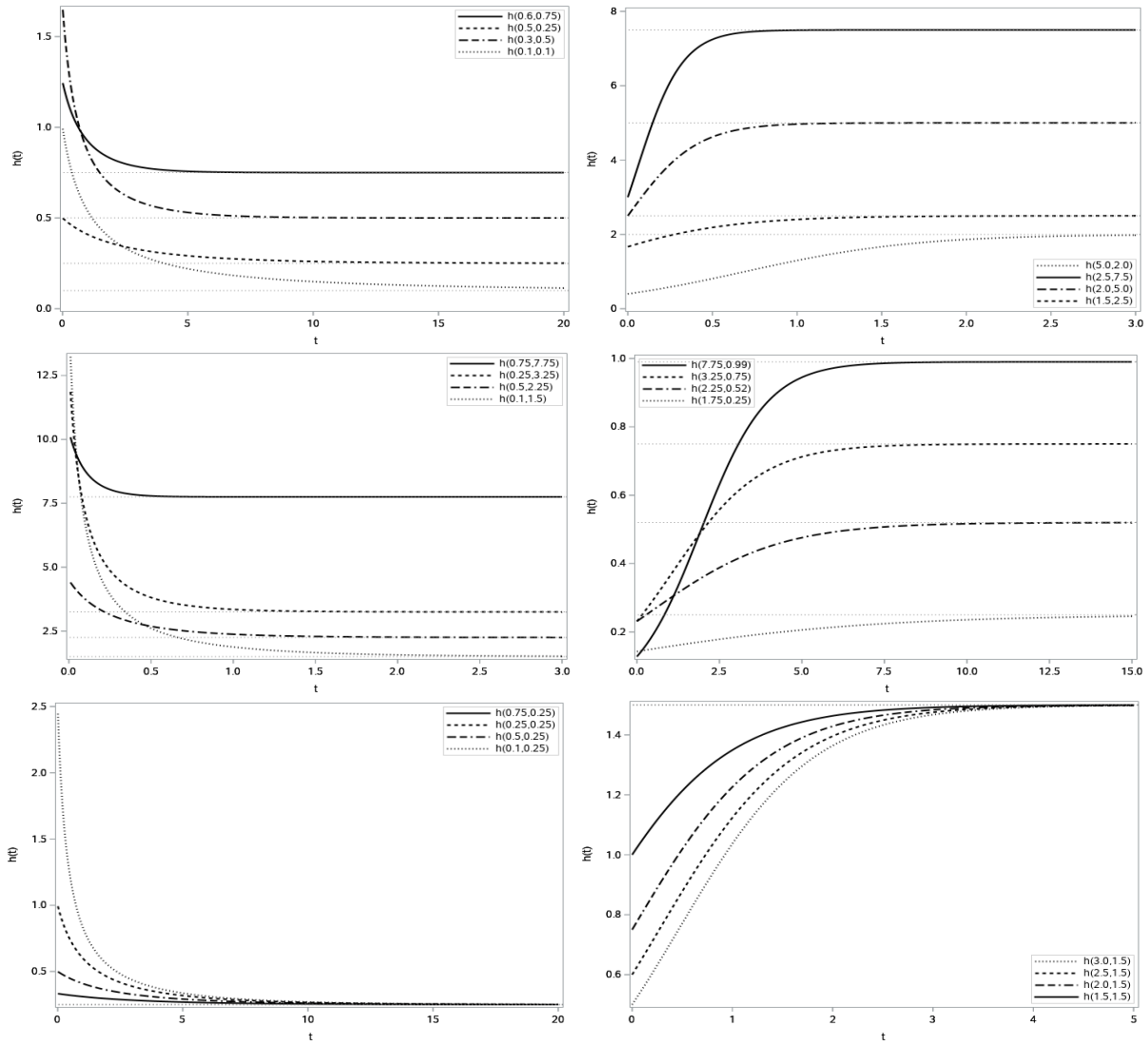

Figure 2. 1 : In panel, the forms for hazard rate functions of the EEG distribution.

2.2 Estimation of Parameters Under Censoring

2.2.1 Maximum Likelihood Estimates

Suppose that $t_1, \ldots, t_n$ are lifetime data with density $f(t)$ and survival function $S(t)$. Assuming that the i-th component could experiment censoring in time C_i, then the data set is (t_i, δ_i), where $t_i = \min\{T_i, C_i\}$ and $\delta_i = 1$ if $T_i \leq C_i$ or $\delta_i = 0$ if $T_i > C_i$. This kind of censorship has random censorship mechanism and type I and II censoring as special schemes and so it is the one used in these study. This way, the likelihood function is given by

$$L(\boldsymbol{\theta}|\boldsymbol{t}, \boldsymbol{\delta}) = \prod_{i=1}^{n} [g(t_i|\boldsymbol{\theta})]^{\delta_i} [S(t_i|\boldsymbol{\theta})]^{1-\delta_i}. \tag{2.40}$$

Now, let $t_1, \ldots, t_n$ be a random sample of EEG distribution then the likelihood function considering data under censoring is given by,

$$L(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta}) = \prod_{i=1}^{n} \frac{\lambda^{\delta_i} \gamma e^{-\lambda t}}{[1-(1-\gamma)e^{-\lambda t}]^{\delta_i+1}} \ . \tag{2.41}$$

The maximum likelihood estimates (MLE) are obtained by maximizing the logarithm of likelihood function given by

$$l(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta}) = r\ln(\lambda) + n\ln(\gamma) - \sum_{i=1}^{n} \left\{ \lambda t_i - (\delta_i + 1) \ln\left[1-(1-\gamma)e^{-\lambda t_i}\right] \right\}, \tag{2.42}$$

where $r = \sum_{i=1}^{n} \delta_i$. From $\dfrac{\partial}{\partial\lambda} l(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta})\Big|_{(\gamma,\lambda)=(\hat{\gamma},\hat{\lambda})} = 0$ and $\dfrac{\partial}{\partial\gamma} l(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta})\Big|_{(\gamma,\lambda)=(\hat{\gamma},\hat{\lambda})} = 0$, we get, respectively, the system of likelihood equations

$$\begin{cases} \dfrac{n}{\hat{\gamma}} - \displaystyle\sum_{i=1}^{n} \frac{(\delta_i+1)e^{-\hat{\lambda}t_i}}{1-(1-\hat{\gamma})e^{-\hat{\lambda}t_i}} = 0 \\ \dfrac{r}{\hat{\lambda}} - \displaystyle\sum_{i=1}^{n} \left[t_i + \frac{t_i(\delta_i+1)(1-\hat{\gamma})e^{-\hat{\lambda}t_i}}{1-(1-\hat{\gamma})e^{-\hat{\lambda}t_i}} \right] = 0 \end{cases}, \tag{2.43}$$

which solutions provide the maximum likelihood estimators of the parameters γ and λ.

Note that the solutions of probability equations cannot be obtained analytically and therefore numerical approaches need to be used in this case.

However, in Kitidamrongsuk et al. [30] is shown that the Fisher information matrix for γ and λ exist, and so the MLE for γ and λ are asymptotically normal and distributed with joint distribution given by,

$$(\hat{\gamma}, \hat{\lambda}) \sim N_2[(\gamma, \lambda), I^{-1}(\gamma, \lambda)] \text{ for } n \to \infty, \tag{2.44}$$

where $I(\gamma, \lambda)$ is the Fisher information matrix with elements given by

$$I_{jk}(\boldsymbol{\theta}) = E\Big[\frac{\partial^2}{\partial\theta_j \partial\theta_k} l(\boldsymbol{\theta}|\boldsymbol{x})\Big], \ j, k = 1, 2 \ , \tag{2.45}$$

where $l(\boldsymbol{\theta}|\boldsymbol{x}) = ln[L(\boldsymbol{\theta}|\boldsymbol{x})]$ and, in case for the EEG distribution, $\boldsymbol{\theta} = (\theta_1, \theta_2) = (\gamma, \lambda)$ whith $l(\boldsymbol{\theta}|\boldsymbol{x}) = l(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta})$ (see subsection 2.1.3 in Kitidamrongsuk et al. [30] for more

detail) which, according to the theorem 1, exists exclusively when $\gamma \leq 1$, and where $\boldsymbol{\delta}$ is a vector with n elements 1, this is, $\boldsymbol{t}, \boldsymbol{\delta} = (1, , \ldots, 1)$, provided that the conditions of regularities necessary for the existence of $I(\gamma, \lambda)$ are satisfied (see Mood, Graybill and Graybill [43] for example).

When all (2.45) it exists, is called the expected information matrix, and its application will be possible if, and only if, it is obtained analytically for all its jk elements, which is not always possible, as possible by the theorem (1) for cases where $\gamma > 1$.

In cases where (2.45) does not exist, or is not obtained analytically, we can consider that the jkth element is obtained by

$$J_{jk}(\boldsymbol{\theta}) = \frac{\partial^2}{\partial\theta_j \partial\theta_k} l(\boldsymbol{\theta}|\boldsymbol{x}) \ . \tag{2.46}$$

However, there are no records in the literature that in the $I(\boldsymbol{\theta})$ matrix there are, simultaneously, the elements I_{jk} and $J_{j'k'}$, and therefore, in cases where at least one element I_{jk} does not exist, the application of Fisher's observed information matrix is considered, which, according to (2.42) and (2.43), for the maximum likelihood estimator of the EEG model in presence of censorship, taking $l'(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta}) = U(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta})$ and $l''(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\theta}) = U^2(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta})$, respectively the first and second order score function, we will have by (2.42) Fisher's observed information matrix given by

$$J(\hat{\alpha}; \hat{\beta}) = \begin{bmatrix} -U^2(\hat{\lambda}|\hat{\gamma}, \boldsymbol{t}, \boldsymbol{\delta}) & -U^2(\hat{\gamma}, \hat{\lambda}|\boldsymbol{t}, \boldsymbol{\delta}) \\ -U^2(\hat{\lambda}, \hat{\gamma}|\boldsymbol{t}, \boldsymbol{\delta}); \hat{\alpha}) & -U^2(\hat{\gamma}|\hat{\lambda}, \boldsymbol{t}, \boldsymbol{\delta}) \end{bmatrix} \ , \tag{2.47}$$

whose elements, where g is the pdf of the EEG model in the time t_i, will be

$$\begin{cases} -U^2(\hat{\lambda}|\hat{\gamma}, \boldsymbol{t}, \boldsymbol{\delta}) = \dfrac{1}{\hat{\lambda}} \displaystyle\sum_{i=1}^{n} \left[\frac{r}{\hat{\gamma}} + \frac{(\delta_i + 1)(1 - \hat{\gamma})}{\hat{\lambda}} t_i^2 g(t_i|\hat{\gamma}, \hat{\lambda}) \right] \\ -U^2(\hat{\gamma}|\hat{\lambda}, \boldsymbol{t}, \boldsymbol{\delta}) = \dfrac{1}{\hat{\gamma}} \displaystyle\sum_{i=1}^{n} \left[\frac{n}{\hat{\gamma}} + \frac{(\delta_i + 1)}{\hat{\lambda}} e^{-\hat{\lambda} t_i} g(t_i|\hat{\gamma}, \hat{\lambda}) \right] \\ -U^2(\hat{\lambda}, \hat{\gamma}|\boldsymbol{t}, \boldsymbol{\delta}) = -U^2(\hat{\gamma}, \hat{\lambda}|\boldsymbol{t}, \boldsymbol{\delta}) = -\displaystyle\sum_{i=1}^{n} \frac{(\delta_i + 1)}{\hat{\lambda}\hat{\gamma}} t_i g(t_i|\hat{\gamma}, \hat{\lambda}) \end{cases} . \tag{2.48}$$

Thus, we define as the lower limit of the variance of the $\boldsymbol{\theta} = (\gamma, \lambda)$ estimators the statistical inequality given by

$$LI(\boldsymbol{\theta}) = [nJ(\boldsymbol{\theta})] \leq Var(\hat{\boldsymbol{\theta}}) \ , \tag{2.49}$$

called observed inequality of information.

In particular, for the $EEG(\gamma, \lambda)$ model the expected inequality of information will exist if, and only if, it exists in (2.48)

$$E[Tg(T|\gamma, \lambda)] = \int_0^\infty t g^2(t|\gamma, \lambda) dt \ , \tag{2.50}$$

$$E[T^2 g(T|\gamma, \lambda)] = \int_0^\infty t^2 g^2(t|\gamma, \lambda) dt \ , \tag{2.51}$$

$$E[e^{-\lambda T} g(T|\gamma, \lambda)] = \int_0^\infty e^{-\lambda t} g^2(t|\gamma, \lambda) dt \ , \tag{2.52}$$

in $E[U^2(\hat{\lambda}, \hat{\gamma}|\boldsymbol{t}, \boldsymbol{\delta})]$, $E[U^2(\hat{\lambda}|\hat{\gamma}, \boldsymbol{t}, \boldsymbol{\delta})]$ and $E[U^2(\hat{\gamma}|\hat{\lambda}, \boldsymbol{t}, \boldsymbol{\delta})]$, respectively, that by the theorem 1, exists only when $\gamma \leq 1$.

2.2.2 Bayesian Estimation

Now, we carry out a bayesian estimation of the parameters γ and λ. This way, we need to assume some prior distributions for the unknown parameters of the distribution, in view of that the prior distributions considered in this work express little or non information about γ and λ.

This can be obtained by assuming independent Gamma distributions for each parameter resulting in a joint prior distribution given by

$$\pi(\gamma, \lambda) \propto \gamma^{a_1-1}\lambda^{a_2-1}e^{-b_1\gamma-b_2\lambda} \ , \tag{2.53}$$

where a_1, b_1, a_2 and b_2 are known hyperparameters.

In this work, we assume that a_1, a_2, b_1 and b_2 in which (2.53) becomes a flat prior. The joint posterior distribution for γ and λ is proportional to the product of the likelihood function (2.41) and the prior distribution (2.53) resulting in

$$p(\gamma, \lambda|\boldsymbol{t}, \boldsymbol{\delta}) \propto \gamma^{a_1+n-1}\lambda^{a_2-1}e^{-b_1\gamma-b_2\lambda}\prod_{i=1}^{n}\frac{\lambda^{\delta_i}e^{-\lambda t_i}}{\left[1-(1-\gamma)e^{-\lambda t_i}\right]^{\delta_i+1}} \ . \tag{2.54}$$

The full conditional posterior distributions for γ and λ are given as follows:

$$p(\lambda|\boldsymbol{t}, \boldsymbol{\delta}, \gamma) \propto \lambda^{a_2-1}e^{-b_2\lambda}\prod_{i=1}^{n}\frac{\lambda^{\delta_i}e^{-\lambda t_i}}{\left[1-(1-\gamma)e^{-\lambda t_i}\right]^{\delta_i+1}} \ , \tag{2.55}$$

and

$$p(\gamma|\boldsymbol{t}, \boldsymbol{\delta}, \lambda) \propto \ \gamma^{a_1+n-1}e^{-b_1\gamma}\left\{\prod_{i=1}^{n}\left[1-(1-\gamma)e^{-\lambda t_i}\right]^{\delta_i+1}\right\}^{-1} \ . \tag{2.56}$$

These conditional distributions are needed in simulation of parameters of the joint posterior distribution based on Monte Carlo Markov Chain (MCMC) methods.

Since the conditional distributions of λ and γ are not identified, we use the Metropolis-Hastings algorithm (see Gamerman and Lopes [14]) to simulate the quantities of interest.

2.2.3 A Simulation Study in the Presence of Random Censorship

We chose to perform this simulation procedure for $m = 2$ parametric cases given by $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2\} = \{(0.25, 3.25), (3.5, 0.75)\}$, the parameter vectors for the model in the cases which the curve of risk manifests the forms, respectively, increasing and decreasing, as shown in the following image.

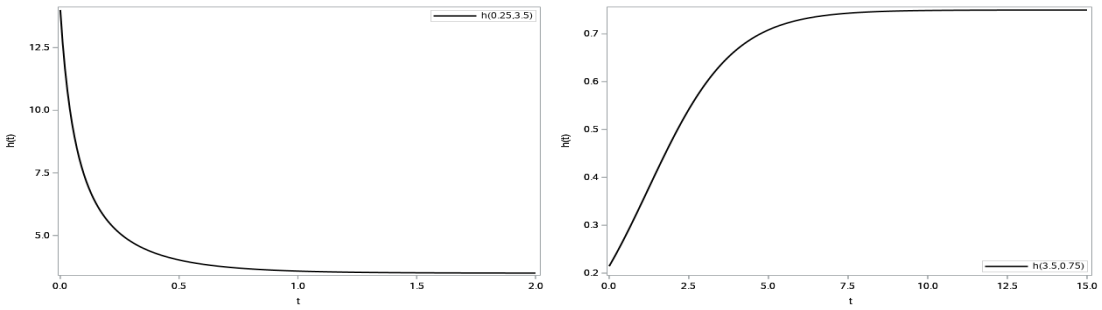

Figure 2. 2 : The hazard rate works for simulations of EEG distribution censored.

In this section, we develop a simulation study used MCMC method whose main objective is to study the efficiency of the MLE method for the distribution $X \sim EEG(x|\gamma, \lambda)$.

For this, the following procedure was computationally implemented

Step 1: Set the values N and n, respectively the number of samples in the simulation and the size of each their, and the values γ and λ in a number of m cases of the parametric vector $\boldsymbol{\theta} = (\gamma, \lambda)$ of the model $EEG(\boldsymbol{\theta}) = EEG(x|\gamma, \lambda)$ censored with a fixed proportion of censorship in N samples, so that $n\tau_\delta$ is the exact number of censorships determined in each sample, where τ_δ is the proportion of censorship.

Step 2: Generate nN values $q \in]0, 1[$, so that $n - n\tau_\delta$ values are complete lifetime and $n\tau_\delta$ values are censored lifetime from each of the N samples of the distribution $X \sim EEG(\boldsymbol{\theta})$ with $x = Q(q)$, according (3.13), and such that $F(x|\boldsymbol{\theta}) \in]0, 1[$.

Step 3: Use the values obtained in step 2 for the $X \sim EEG(\boldsymbol{\theta})$ distribution to calculate in each of the N samples the estimated vector $\hat{\boldsymbol{\theta}} = (\hat{\gamma}, \hat{\lambda})$, this is, for $i = 1, \ldots, N$, get $\hat{\boldsymbol{\theta}}_i$ through MLE of the γ and λ parameters by the MCMC method.

Step 4: Use the N vectors $\hat{\boldsymbol{\theta}} = (\hat{\gamma}, \hat{\lambda})$ and the vector $\boldsymbol{\theta} = (\gamma, \lambda)$ for compute the mean bias absolute (MBA), square root from the mean square errors (MSE) and 95% coverage probability, respectivaly

$$V_{\theta_{jk}} = \frac{1}{N} \sum_{i=1}^{N} |\theta_{jk} - \hat{\theta}_{ijk}| \ , \tag{2.57}$$

$$e_{\theta_{jk}} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\theta_{jk} - \hat{\theta}_{ijk})^2} \ , \tag{2.58}$$

$$\hat{p}_{\theta_{jk}} = \frac{1}{N} \sum_{i=1}^{N} W_{ijk} \ , \tag{2.59}$$

so that, in the i-th sample where, for j-th parameter, with $j = 1, 2, 3$, and k-th case, with $1 \leq k \leq m$

$$W_{ijk} = \begin{cases} 1, \text{ if } \theta_{jk} \in IC_{\{\theta_{jk}, 0.95\}} \\ 0, \text{ otherwise} \end{cases} \ ,$$

and $IC_{\{\theta_{jl}, 0.95\}}$ is the interval of 95% (of credibilaty or confidence) for the parameter θ_{jk}, with $\boldsymbol{\theta}_k = (\theta_{1k}, \theta_{2k}) = (\lambda_k, \gamma_k)$ and $\hat{\boldsymbol{\theta}}_{ik} = (\hat{\theta}_{i1k}, \hat{\theta}_{i2k}) = (\hat{\lambda}_{ik}, \hat{\gamma}_{ik})$, in the i-th sample of the k-th case, respectivaly. Moreover, for the N confidence and credibility intervals obtained, we will also consider the mean interval amplitude (MIA) in each case as

$$\overline{h}_{IC_{\{\theta; 1-\epsilon\}}} = \frac{1}{N} \sum_{i=1}^{N} h_{IC_{\{\theta_i; 1-\epsilon\}}} \ , \tag{2.60}$$

were $h_{IC_{\{\theta_i; 0.95\}}} = 2z_{\frac{0.05}{2}} \hat{\sigma}_{\hat{\theta}_i}$ in the classic case and $h_{IC_{\{\theta_i; 0.95\}}} = \hat{\theta}_i^{(k+[0.95]\eta)} - \hat{\theta}_i^{(k)}$ in the bayesian case, were $\theta_i^{(k)}$ is the $k-$th smallest lower limit and $\theta_i^{(k+[0.95]\eta)}$ is the $[k + (0.95)\eta]-$th smallest upper limit of the ordered set of quantis $\boldsymbol{\theta}_j^* = \{\theta_j^{(1)}; \theta_j^{(2)}; \theta_j^{(3)}; ...; \theta_j^{(\eta)}\}$ from the $j-$th posterior sample for size η.

Repeat steps 2, 3, 4 and 5 for the m cases of $\boldsymbol{\theta}$.

To generate the random data with $\tau_\delta = 0.0, 0.2, 0.4$, the exact proportion of censorships, the sizes $n = 10, 25, 50, 100$ in $N = 500$ samples were used to perform 24 simulation processes by the approach classic and bayesian, totaling 48 processes.

For this approach, MBA and MPE are expected to approach zero as samples of size n increase and your respective this likelihood of empirical coverage p approaches 0.95, the theoretical probability of the confidence and credibility interval which for this simulation process is defined as the most rigorous in the case of credibility, with the Highest Posterior Density (HPD) interval.

The seed used to generate the simulation random values was the $64 - bit$ Windows 10 operating system time, with Intel® Core™ $i5 - 4200U$ CPU @ $1.60GHz$ processor and $8.00GB$ installed RAM. The software used was SAS On Deamand where the implemented code mainly considered the DATA STEP process and the IML and MCMC procedures, all with seed inserted by the STREAMINIT(0) statement.

For each sample of the EEG model, the MCMC process was implemented with 50000 iterations for each of the 500 samples, with 10000 iterations burn-in, and roughing with 4 iterations to average each of the 500 subsequent samples of size $\eta = 10000$. The desired estimate was obtained by averaging these 500 averages and to generate the exact amount of censored data we used the same methods used by Goodman, Li and Tiwari [17].

It is noteworthy that the following results reflect the care taken with the sample autocorrelation, since with burning 20% of the initial iterations for the estimates generated in the Markov chain followed by the thinning of 4 units, the 10000 values generated are uncorrelated throughout the iterations.

The results of the cases described for the simulation are presented in tables (2. 1) and (2. 2) below, where we used the methods describe by Goodman, Li and Tiwari (2006, [17]) for generation the censorships.

Table 2. 1 : Results for the model $EEG(\gamma, \lambda) = EEG(0.25, 3.25)$.

Cases			Classical estimation					Bayesian estimation				
θ	τ_δ	n	$\hat{\theta}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$	$\hat{\mu}_{\hat{\theta}}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$
γ	0	10	1.130	0.978	6.932	0.948	5.707	0.293	0.991	0.018	0.103	0.612
		25	0.453	0.964	0.243	0.277	1.411	0.275	0.980	0.011	0.082	0.588
		50	0.361	0.958	0.071	0.173	0.809	0.282	0.970	0.014	0.088	0.516
		100	0.296	0.970	0.018	0.096	0.486	0.261	0.956	0.010	0.080	0.372
	0.2	10	1.888	0.956	20.726	1.723	11.091	0.283	1.000	0.007	0.066	0.692
		25	0.542	0.976	0.461	0.361	1.923	0.292	0.994	0.011	0.079	0.648
		50	0.417	0.976	0.126	0.228	1.063	0.299	0.988	0.016	0.094	0.585
		100	0.333	0.976	0.036	0.133	0.627	0.277	0.979	0.014	0.087	0.502
	0.4	10	2.497	0.884	36.283	2.350	18.372	0.546	1.000	0.093	0.296	1.227
		25	0.701	0.942	0.937	0.532	2.997	0.765	1.000	0.283	0.515	1.529
		50	0.468	0.986	0.184	0.277	1.445	1.338	1.000	0.751	0.183	1.849
		100	0.390	0.992	0.068	0.186	0.889	1.693	1.000	2.178	1.443	2.031
λ	0	10	7.603	0.974	87.237	5.156	21.783	3.297	0.986	1.114	0.833	6.709
		25	4.403	0.960	9.571	2.130	10.264	3.311	0.984	1.000	0.792	5.985
		50	4.083	0.942	4.509	1.519	6.768	3.385	0.978	1.219	0.852	5.082
		100	3.645	0.958	1.581	0.930	4.554	3.315	0.968	0.849	0.634	3.785
	0.2	10	8.560	0.916	149.707	6.576	26.617	2.986	0.964	0.956	0.802	6.590
		25	4.400	0.960	11.017	2.266	11.225	2.934	0.956	0.887	0.779	5.626
		50	3.783	0.932	4.512	1.561	7.305	2.986	0.942	0.978	0.809	4.941
		100	3.263	0.948	1.553	0.975	4.792	2.263	0.938	1.002	0.977	4.385
	0.4	10	7.820	0.870	109.853	6.159	31.055	0.244	1.000	9.038	3.006	0.545
		25	4.060	0.922	12.817	2.411	12.698	0.272	1.000	8.869	2.978	0.491
		50	3.231	0.962	3.494	1.449	7.832	0.622	1.000	8.341	2.896	0.408
		100	2.902	0.908	1.931	1.137	5.297	0.439	1.000	7.907	2.811	0.361

In the case where $\gamma > 1$ and $\lambda < 1$, for the parametric vector $(3.25, 0.75)$ as shown in the following table, the results are the same, showing that in any parametric case the bayesian approach is preferable over the classical.

Table 2. 2 : Results for the model $EEG(\gamma, \lambda) = EEG(3.5, 0.75)$.

Cases			Classical estimation					Bayesian estimation				
θ	τ_δ	n	$\hat{\theta}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$	$\hat{\mu}_{\hat{\theta}}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$
γ	0	10	13.962	0.948	2127	11.608	86.203	4.120	0.984	3.722	1.394	8.893
		25	5.131	0.936	19.521	2.652	14.825	4.114	0.976	3.072	1.354	7.891
		50	4.446	0.952	6.886	1.713	8.664	4.136	0.968	2.851	1.239	6.417
		100	3.891	0.956	2.120	1.041	5.235	3.862	0.962	1.550	0.918	4.608
	0.2	10	72.309	0.946	857579	69.800	116.44	4.682	0.999	4.127	1.849	10.361
		25	6.436	0.970	46.070	3.686	20.692	4.456	0.988	3.797	1.461	8.843
		50	5.192	0.968	11.435	2.268	11.129	4.277	0.982	3.273	1.400	6.929
		100	4.523	0.984	3.927	1.386	6.672	4.376	0.980	2.520	1.168	5.594
	0.4	10	138.85	0.956	1212328	136.20	512.63	4.469	0.996	2.993	1.346	10.740
		25	9.767	0.972	183.735	6.883	36.721	4.920	0.996	5.475	1.797	10.110
		50	6.645	0.990	32.886	3.564	16.235	5.112	0.990	6.087	1.902	8.859
		100	6.009	0.987	21.291	2.736	7.948	4.891	0.951	6.240	2.311	7.961
λ	0	10	0.968	0.966	0.241	0.353	1.693	0.752	0.978	0.082	0.162	1.204
		25	0.817	0.952	0.061	0.186	0.902	0.729	0.970	0.025	0.126	0.749
		50	0.796	0.936	0.027	0.124	0.613	0.750	0.958	0.016	0.098	0.537
		100	0.770	0.962	0.011	0.083	0.421	0.758	0.964	0.008	0.073	0.393
	0.2	10	0.984	0.982	0.318	0.390	1.853	0.745	0.997	0.111	0.174	1.251
		25	0.825	0.948	0.072	0.198	0.998	0.730	0.960	0.025	0.127	0.749
		50	0.784	0.944	0.029	0.132	0.669	0.719	0.967	0.025	0.133	0.642
		100	0.759	0.944	0.013	0.089	0.459	0.739	0.960	0.010	0.080	0.421
	0.4	10	1.039	0.978	0.567	0.474	2.153	0.664	0.950	0.040	0.166	1.024
		25	0.834	0.960	0.098	0.238	1.132	0.683	0.954	0.027	0.139	0.781
		50	0.767	0.944	0.041	0.153	0.743	0.693	0.948	0.022	0.121	0.602
		100	0.745	0.948	0.007	0.099	0.473	0.738	0.951	0.017	0.082	0.591

The table 2. 1 highlights that the estimates assume discrepant values in all cases of size 10 samples, in both parameters and only in the classic case. In this case, even for $\gamma < 1$ the estimation errors considered are relatively high and, although this is expected, increase dramatically when the censorship ratio increases.

In the bayesian case, we can see that in samples of any size and up to 20% censorship the estimates and estimation errors are reasonable and behave as expected, with errors decreasing as sample size increases. A fact also present in the classic case, however, in the bayesian case convergence occurs faster with estimates close the true value since samples of size 10.

Despite the superiority of the bayesian estimation over the classical one in samples of any size and in the case of 0% and 20% censorship, the results show that in the case of 40% censorship the classical approach is preferable to the bayesian estimator for γ parameter only.

It is noteworthy that in the bayesian case the approach to 40% censorship when $\gamma < 1$ does not follow any expected pattern, with divergence of the true parameter value and therefore increase of the estimation errors, whereas in the classic case it is verified that the As sample size increases, estimates tend to true value and estimation errors decrease.

The table 2. 1 still shows that, although the probability of coverage for the case of 40% of censorship under classical estimation does not tend to the theoretical value of 95%, the estimator in this case is preferable over the bayesian in both parameters, becouse as is observed about the λ parameter the behavior is analogous.

The simulation for the parametric vector $(0.25, 3.5)$ provides the results to evaluate the

behavior of the model estimators in a particular case where $\gamma < 1$ and $\lambda > 1$, and allows us to conclude that the bayesian estimation is preferable the classical in cases where the censorship proportion of the sample is not significant.

Above all, the previous results in table 2. 2 allow us to conclude that as the sample size increases, the values of the estimatives $\hat{\alpha}$ and $\hat{\beta}$ are closer to the true values of the parameters α and β, respectively, and similarly the probability of empirical coverage approaches pre-established confidence level, as expected in applied asymptotic theory. However, the comparetion between the classical and bayesian cases, especially for small samples, shows bayesian estimators are preferable.

Therefore, the results show that the EEG distribution is indeed flexible and can be approached in both classical and bayesian contexts to model censored data from samples of any size.

2.2.4 Aplication for Censored Data for Advanced Lung Cancer

In this application the proposed methodology was implemented to model the life span of patients with advanced lung cancer. The data used refer to the lifetime (in days) of a group of 137 patients observed by the Veterans Administration's Lung Cancer Study Group and reported in a work proposed by Prentice [48] and the modeling considered the support of the classic inference and bayesian to obtain the estimates of the parameters of the survival and risk models intended in the modeling as models derived from the EEG, Gamma, Log Logistica (Log-Log.) and Weibul distributions.

The main objective of this application is to put the EEG probabilistic model in competition with the others, as well as to contrast the estimators of this model according to an application to real data. For this purpose, the 137 observations for the development of classical inference through the MLE (Maximum Likelihood Estimators) obtained from the system solution presented in 2.43 are considered, while the bayesian inference is applied to a random sample of 40 observations among the 137 presented by Prentice [48] and developed with the prioris vague presented in (2.55) and (2.56) to estimate the parameters γ and λ, respectively.

Above all, the modeling is proposed as an adjustment for the data according to a probabilistic model that captures the alternative forms of risk, and as shown in the following figure, the data indicate that the hazard rate for patients is predominantly decreasing, which is confirmed by plotting the empirical risk function based on the Epaneshnikov smoothed kernel function.

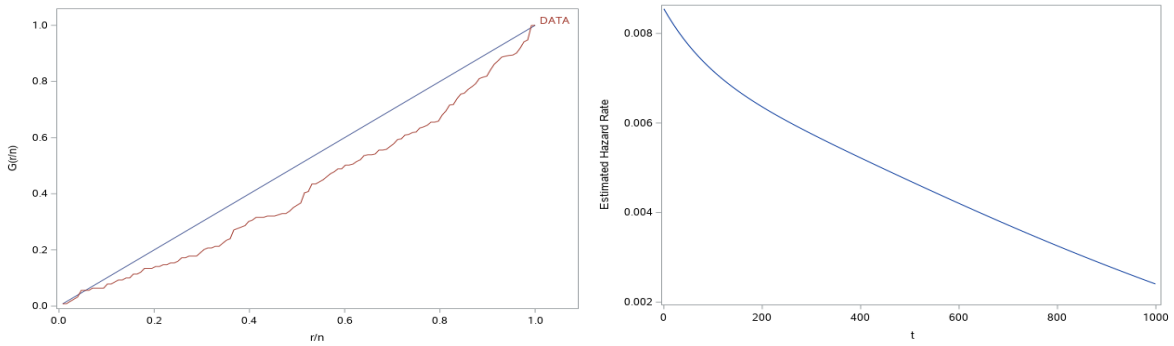

Figure 2. 3 : TTT-plot and empirical hazard rate plot for 137 patients.

One of the most common ways to verify the fit for the survival model is to compare the fitted model with the survival curve obtained by the Kaplan-Meyer nonparametric estimator, and for the problem data, with 6.6% of censorship, the graph of the estimated

Kaplan-Meyer survival curve is shown in the following graph, where + indicates the probability of the censored time.

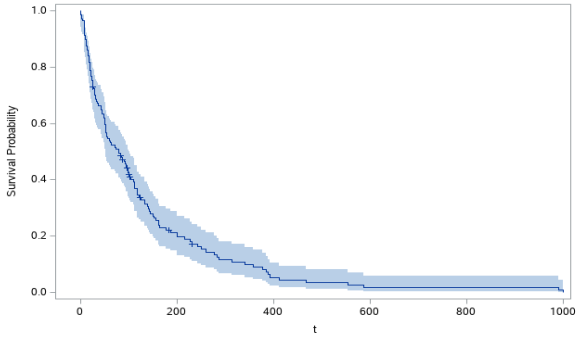

Figure 2. 4 : Empirical survival plot for patients data with lung cancer.

Next, using the maximum likelihood method, the estimates for the parameters of the EEG, Gamma, Log-Log models are presented. and Webull obtained with 5% significance.

The table shows the estimates for the location and scale parameters in column $\hat{\theta}$, where e γ and λ are the location and scale parameters, respectively, indicated by column θ, and with their respective standard errors and asymptotic confidence intervals, respectively in the std.err and $CI_{95\%}(\theta)$ columns, as well as the test statistic and the p-value for the t-Student test, respectively in t-stat and p-value.

Table 2. 3 : Results estimations for the models.

Model	$\boldsymbol{\theta}$	$\hat{\boldsymbol{\theta}}$	std. err.	$IC_{95\%}(\theta)$	test-stat	p-value
EEG	γ_1	0.2611	0.0805	(0.1018; 0.4205)	3.2400	0.0015
	λ_2	0.0038	0.0011	(0.0017; 0.0060)	3.5500	0.0005
Weibull	γ_2	0.8521	0.0570	(0.7393; 0.9649)	14.9400	<0.0001
	λ_2	120.6800	13.0096	(94.9549; 146.4100)	9.2800	<0.0001
Gamma	γ_3	0.8095	0.0860	(0.6394; 0.9795)	9.4100	<0.0001
	λ_3	0.0062	0.0009	(0.0043; 0.0080)	6.7300	<0.0001
LogLog	γ_4	1.2679	0.09249	(1.0850; 1.4508)	13.7100	<0.0001
	λ_4	67.9870	8.0648	(52.0393; 83.9346)	8.4300	<0.0001

The table 2. 5 in the sequence shows the results for the main information criteria for the considered adjustments.

Table 2. 4 : Results for measures of fit.

Criterion	EEG	Weibull	Gamma	Log-Log.
$-2log[L(\theta)]$	1496.0	1496.2	1498.2	1500.5
AIC	1500.0	1500.2	1502.2	1504.5
$AAIC$	1500.1	1500.3	1502.3	1504.6
BIC	1505.9	1506.0	1508.0	1510.4

Given the candidate models for data adjustment, it is conventionally preferred to choose the one that provides the lowest value of the information criterion, in this case, based on the criteria presented in table 2. 5 , it is concluded that the EEG distribution provides the best fit, followed by the Weibull, Gamma and Log-Log. models, respectively, although it is observed that in both criteria a result is shown pointing a narrow difference between the candidate models for the best fit. However, this small difference is observed in the figure 2. 5 in the next page, which shows the graphic result of this adjustment.

Note by figure 2. 5 in the sequence, that the adjusted survival graphs highlight the negligible difference between the 4 values obtained by the information criteria, however,

although under small difference, these values actually point to models less likely to represent the data, which is notable for the graph risk for each model.

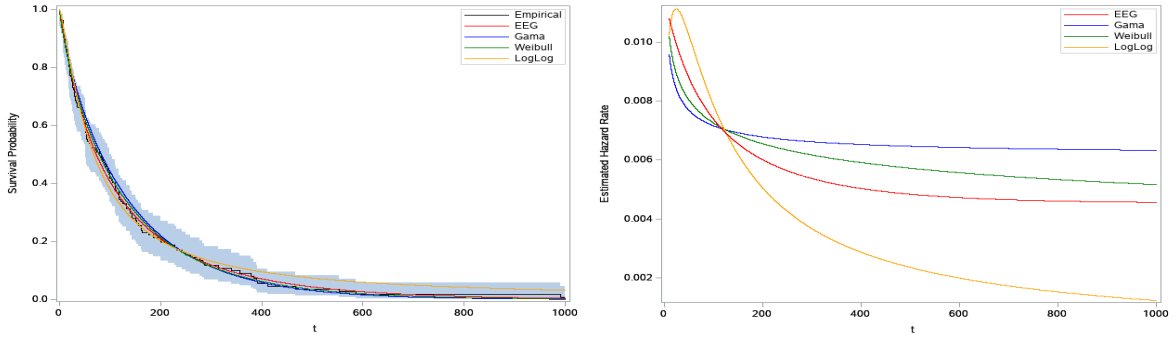

Figure 2. 5 : Survival and hazard rate plot for the fitted models for 137 patients data.

The following table shows the average distance that the survival models adjusted in comparison to the empirical model as the representative of the true model. The comparison is given for the four models as the average of the differences between the estimated value for the corresponding adjusted survival function and the corresponding empirical value obtained via Kaplan-Mayer.

Table 2. 5 : Average distance measures.

EEG	Weibull	Gamma	Log-Log.
0.0130	0.0163	0.0244	0.0231

As said initially to this application, to contrast the previous results, we took a random sample of size 40 from patients with advanced lung cancer.

The table 2. 6 displays the data sampled for the survival times (in days) of 40 patients, where the symbol * indicates the presence of censorship in the times 025*, 103* e 231*, and in the sequence is shows the empirical survival function obtained via Kaplan-Meyer for the sample obtained.

Table 2. 6 : Dataset related to the lifetimes.

1	2	8	8	10	11	12	12	15	16
18	19	20	21	25*	43	44	51	54	56
82	84	90	100	103*	118	126	153	164	177
200	201	231	231*	250	287	340	411	991	999

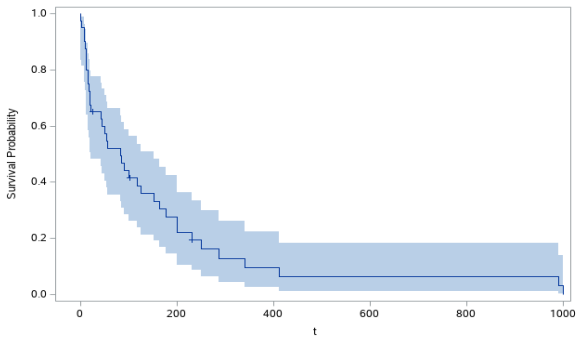

Figure 2. 6 : Empirical survival plot for 40 patients data with lung cancer.

Under the same conditions for the previous adjustment, where it was considered to fit a probabilistic model according to the form of risk that the data presents, the following

figure shows the TTT-plot (total time in test) and the graph for the empirical hazard rate function for the sample of 40 observations.

Under the same conditions for the previous adjustment, where it was considered to fit a probabilistic model according to the form of risk that the data presents, the following figure shows the TTT-plot and the graph for the empirical hazard rate function smoothed for the sample of 40 observations. Note that TTT-plot coincides with the empirical hazard rate function, both of which provide information that survival models must still adjust a decreasing risk model to the sampled data.

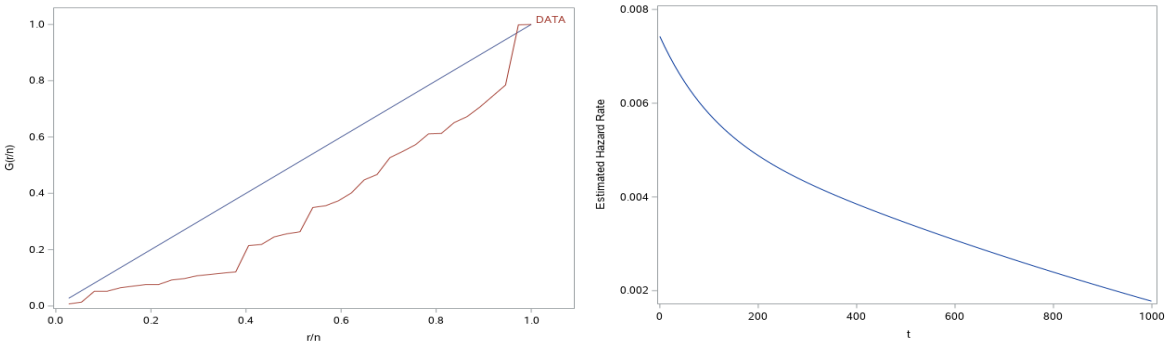

Figure 2. 7 : TTT-plot and empirical hazard rate function graph for 40 patients.

Competition with other probabilistic models is also considered in this case of the adjustment and to maintain the same conditions the same probability models are considered: the Gamma, Lo-log models. and Weibull.

So, based on the table 2. 6 , since data related to lung cancer patients have the censorship mechanism, the equations (2.41) and (2.53) were used to obtain bayesian estimates for the adjustment. Following, the table 2. 7 displays these estimates and the convergence test for the executed MCMC process.

Table 2. 7 : Results for MCMC process with 95% credibility for models.

Cases		Results estimations for the models			Results from convergence test		
Model	$\boldsymbol{\theta}$	$\hat{\boldsymbol{\theta}}$	std.err.	$IC_{95\%}(\theta)$	test-stat	p-value	test outcome
EEG	γ_1	0.1266	0.0803	(0.0048; 0.2840)	0.1432	0.4109	passed
	λ_1	0.00172	0.0009	(0.0002; 0.0036)	0.1523	0.3825	passed
Weibull	γ_2	0.7062	0.0890	(0.5227; 0.8701)	0.1825	0.3041	passed
	λ_2	132.5000	32.8925	(77.2955; 200.3000)	0.1702	0.3335	passed
Log Log.	γ_3	1.0490	0.1391	(0.8001; 1.3237)	0.3251	0.1150	passed
	λ_3	66.1672	17.8906	(36.7660; 105.1000)	0.1097	0.5392	passed
Gamma	γ_4	0.6136	0.1141	(0.4056; 0.8406)	0.0920	0.6257	passed
	λ_4	0.0039	0.0011	(0.0018; 0.0060)	0.0960	0.6046	passed

The table 2. 7 shows the estimates for location and scale parameters in the coluns $\hat{\boldsymbol{\theta}}$, respectivaly indicated for γ and λ by the colun $\boldsymbol{\theta}$, and their respective standard-error and yours 95% credibility intervals in the columns std.err. and $IC_{95\%}(\theta)$, respectivaly.

In graphics panel in the next page, the set formed by the graphics of iteration, autocorrelation, and a posteriori density plots outlines the diagnosis for a parameter of interest.

Geometrically, the iteration plots indicate that the Markov chains converged actually to an estimate the parameters of interest, as indicated in the tests in table 2. 7 , and that all these iterations are independent of each other as shown in the autocorrelation plots.

In the 2. 8 panel, each row represents an estimated model and each column represents a parameter for respective model, the gamma and lambda parameters. Each row contains

the iterations, autocorrelation, and a posteriori density graphs for each parameter in the model respectively for the EEG, Gama, LL and Weibull models.

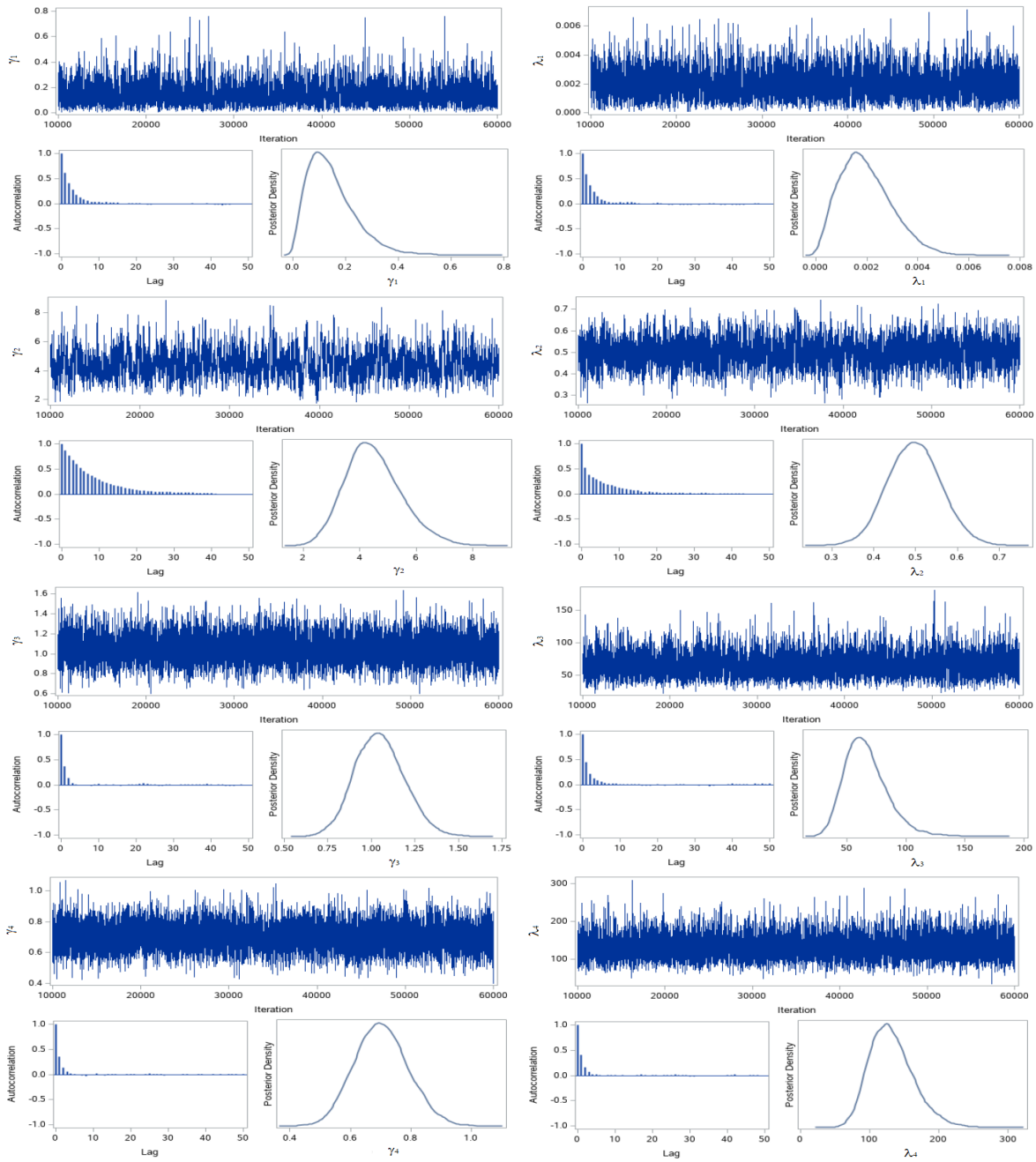

Figure 2. 8 : Diagnostics plot for the convergence the markov chains.

The table 2. 8 below shows the results for the some information criteria for the adjustments made to lung cancer patient data.

Table 2. 8 : Results for measures of fit.

Criterion	EEG	Weibull	Log-Log.	Gamma
p_D	1.289	1.978	1.939	1.908
$D(\hat{\theta})$	438.728	439.510	440.582	439.519
DIC	441.305	443.466	444.459	444.755

Given the set of candidate better models, it is conventionally preferred to choose the one that provides the value lowest DIC. In this case, based on the criteria presented in 2. 8 it is concluded that the EEG distribution provides the best fit.

However, even though the EEG model has the lowest penalty, this is verified by a significantly low difference in relation to the four models confronted. The reader may therefore be tempted to assume that, statistically, the values for these comparisons are the same, which implies considering the average distance shown in table 2. 9 in the sequence, which in turn shows that the adjustment that led to the estimated survival curve closest to the empirical was the curve derived from the Weibull model.

Table 2. 9 : Average distance measures.

EEG	Weibull	Log-Log.	Gamma
0.0329	0.0327	0.0353	0.0404

Situations like this lead to comparisons by graphical routes, which eventually, due to the survival curve plotted with the 95% confidence bounds, does not allow considering the best fit to the data between the adjusted models, as long as some curve is plotted outside confidence bounds. In addition, if we consider the risk curve, provided that any of them is contrary to the previous information provided by the TTT-plot and the empirical risk smoothed, as shown in figure 2. 7 in the sequence, no significant conclusion can be considered.

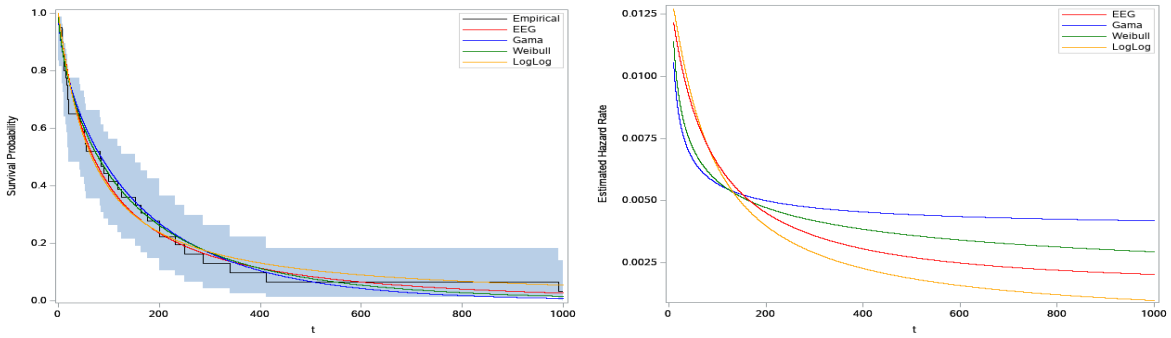

Figure 2. 9 : Survival and hazard rate plot for the fitted models.

Then, by geometric means, it is possible to verify at most which of the adjusted models are not appropriate so that estimates are taken through it. See in the image that the adjusted survival functions are entirely contained in the confidence region and that the risk functions are all decreasing.

In the image 2. 7 , on the left, if the adjusted survival graph is considered as a tiebreaker between the EEG and Weibull models, noting that the largest distance from the Kaplan-Meyer curve for the EEG and Log-Log models is highlighted in most of the observed time, this allows the interpretation that the EEG model presents a lower adjustment in relation to the Weibull for the survival data shown in the table 2. 8 .

On the right, in the image 2. 7 , the adjusted risk graphs highlight only that the estimated models capture the form of risk expected by the data.

In the following section we will consider a mechanism that, inserted or omitted from the modeling and model selection process, situations such as that obtained in the sample of 40 observations can be easily circumvented.

2.3 The EEG Distribution in the Presence of Covariates

The Exponential distribution can be alternatively parameterized according to its probability density function, and when a random variable T with such a distribution is interpreted as the length of time a mechanical or biological system survives, we say that $\lambda > 0$

is the scale parameter of the distribution and is the inverse of the rate parameter $\phi > 0$.

In this case we have to

$$\lambda = \frac{1}{\phi} \, . \tag{2.61}$$

Furthermore, to study the lifetime data, it is important to consider the relationship of lifetime with other factors. From this, let us assume the presence of a vector $\boldsymbol{x} = (x_1, \ldots, x_p)$ of independent covariates associated to lifetime T with exponential distribution and scale parameter λ in the presence the covariates as follows

$$\lambda(\boldsymbol{x}) = \frac{1}{\phi(\boldsymbol{x})} \, . \tag{2.62}$$

Since there are a multitude of shapes for covariates to relate to in a model, then let's assume that ϕ is a linearizable nonlinear model defined as

$$\phi(\boldsymbol{x}) = e^{\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle} \, , \tag{2.63}$$

were $\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle$ is an internal product between $\boldsymbol{x}$ and $\boldsymbol{\beta}$, this is, $\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle = \boldsymbol{x}\boldsymbol{\beta}' = y \in \mathbb{R}$.

In this condition, let us assume the presence of a vector $\boldsymbol{x} = (x_1, \ldots, x_p)$ of covariates associated to lifetime T under the EEG distribution (2.6) with the scale parameter λ depending on the covariates as follows

$$\lambda(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle} \, , \tag{2.64}$$

where $\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle \in \mathbb{R}$, with $\boldsymbol{x} = (1, x_1, \ldots, x_p)$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)$ is the coefficients vector with $\beta_j \in \mathbb{R}$ and fixed for all $j = 0, \ldots, p$.

Note that, how the vectors $\boldsymbol{x}$ and $\boldsymbol{\beta}$ are of dimension $p + 1$, we have that $\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle = y$ is a additive effect, that reasonably describes relationships between various explanatory variables of a given process, and how the collection of variables that make up $\boldsymbol{x}$ are not observable, $\lambda(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}$ becomes a curve nonlinear composed of multiple variates that confers randomness with considerable influence in the variability of the probabilistic model.

Formally, since $\lambda(\boldsymbol{x})$ varies in $\boldsymbol{x} \in \mathbb{R}^{p+1}$ for any $\boldsymbol{\beta} \in \mathbb{R}^{p+1}$ fixed, we can define it as a function $\lambda_{\boldsymbol{\beta}}$ as follows.

Definition 1 *Let $\boldsymbol{\beta} \in \mathbb{R}^{p+1}$ a parameter vector fixed for any $\boldsymbol{x} \in \mathbb{R}^{p+1}$. If we define a relation $\lambda_{\boldsymbol{\beta}}$ as*

$$\begin{aligned} \lambda_{\boldsymbol{\beta}} : \mathbb{R}^{p+1} &\longrightarrow \quad \mathbb{R}_+^* \\ \boldsymbol{x} \quad &\longmapsto \lambda_{\boldsymbol{\beta}}(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle} \end{aligned}$$

then it is easy to see that $\lambda_{\boldsymbol{\beta}}$ define a function for $\boldsymbol{x}$ with fixed $\boldsymbol{\beta}$.

See also that $Im\, \lambda_{\boldsymbol{\beta}} = \mathbb{R}_+^*$ and in this case we can take a subset $\mathbb{E}_{\boldsymbol{\beta}} \subseteq Im\, \lambda_{\boldsymbol{\beta}}$ defined as

$$\mathbb{E}_{\boldsymbol{\beta}} = \{\lambda_{\boldsymbol{\beta}}(\boldsymbol{x}) \in Im\, \lambda_{\boldsymbol{\beta}};\, \forall\, \boldsymbol{x} \in \mathbb{R}^{p+1}\} \, , \tag{2.65}$$

and such that $\forall\, t \in \mathbb{R}_+^*$, it is also defined that

Definition 2 *Let $\lambda_{\boldsymbol{\beta}} : \mathbb{R}^{p+1} \longrightarrow \mathbb{R}_+^*$ be a function, as in definition 1, for any $\boldsymbol{x} \in \mathbb{R}^{p+1}$ and a fixed $\boldsymbol{\beta}$. Then we define*

$$\begin{aligned} g : \mathbb{R}_+^* \times \mathbb{E}_{\boldsymbol{\beta}} &\longrightarrow \quad]0, 1[\\ (t, \lambda_{\boldsymbol{\beta}}(\boldsymbol{x})) &\longmapsto g(t|\boldsymbol{x}, \boldsymbol{\beta}, \gamma) = \frac{\gamma \exp\left(-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle - te^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}\right)}{\left[1 - (1 - \gamma) \exp\left(-te^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}\right)\right]^2} \, , \end{aligned} \tag{2.66}$$

$$
\begin{aligned}
G:\ \mathbb{R}_+^* \times \mathbb{E}_{\boldsymbol{\beta}} &\longrightarrow\ ]0,1[\\
(t, \lambda_{\boldsymbol{\beta}}(\boldsymbol{x})) &\longmapsto G(t|\boldsymbol{x},\boldsymbol{\beta},\gamma) = \frac{1-\exp(-te^{-\langle \boldsymbol{x},\boldsymbol{\beta}\rangle})}{1-(1-\gamma)\exp(-te^{-\langle \boldsymbol{x},\boldsymbol{\beta}\rangle})}\ ,
\end{aligned} \tag{2.67}
$$

$$
\begin{aligned}
S:\ \mathbb{R}_+^* \times \mathbb{E}_{\boldsymbol{\beta}} &\longrightarrow\ ]0,1[\\
(t, \lambda_{\boldsymbol{\beta}}(\boldsymbol{x})) &\longmapsto S(t|\boldsymbol{x},\boldsymbol{\beta},\gamma) = \frac{\gamma\exp\left(-te^{-\langle \boldsymbol{x},\boldsymbol{\beta}\rangle}\right)}{1-(1-\gamma)\exp\left(-te^{-\langle \boldsymbol{x},\boldsymbol{\beta}\rangle}\right)}\ ,
\end{aligned} \tag{2.68}
$$

$$
\begin{aligned}
h:\ \mathbb{R}_+^* \times \mathbb{E}_{\boldsymbol{\beta}} &\longrightarrow\ \mathbb{R}_+^* \\
(t, \lambda_{\boldsymbol{\beta}}(\boldsymbol{x})) &\longmapsto h(t|\boldsymbol{x},\boldsymbol{\beta},\gamma) = \frac{e^{-\langle \boldsymbol{x},\boldsymbol{\beta}\rangle}}{1-(1-\gamma)\exp(-te^{-\langle \boldsymbol{x},\boldsymbol{\beta}\rangle})}\ ,
\end{aligned} \tag{2.69}
$$

are, respectivaly, the density, distribution, survival and hazard rate functions of the EEG distribution in the presence the covariates.

Following figure 2. 10 shows density graphs for two parametric cases for γ and $\boldsymbol{\beta}$.

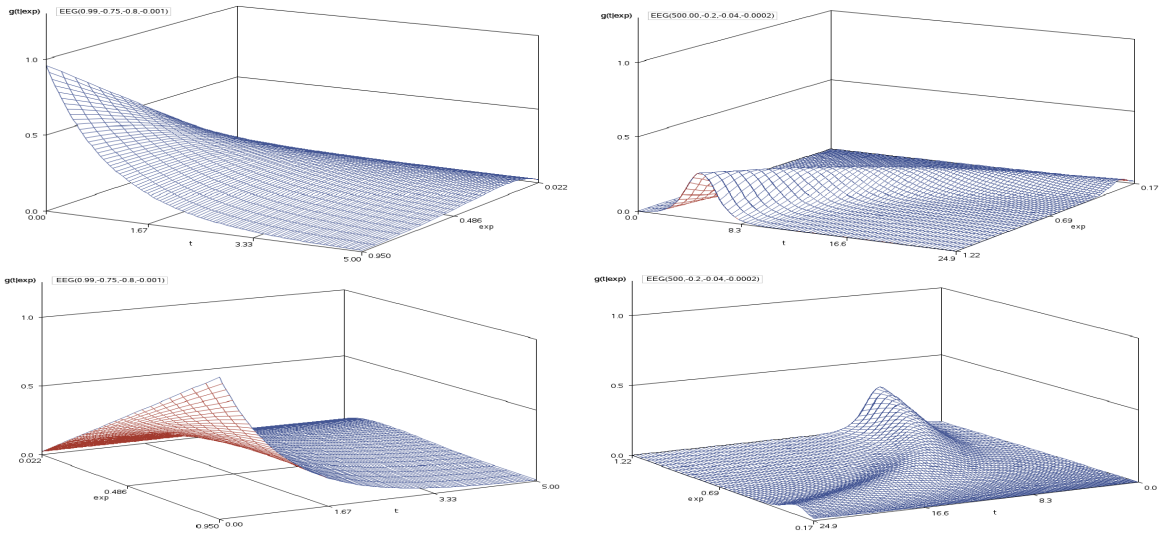

Figure 2. 10 : In the panel, the density function for EEG distribution.

The figure 2. 10 draught that the density function of the EEG distribution in the presence of covariates has two sources of variability: lifetime t and the set of multiple variates that constitute the additive effect index $\lambda_{\boldsymbol{\beta}}(\boldsymbol{x}) = e^{-\langle \boldsymbol{x},\boldsymbol{\beta}\rangle}$.

Empirically, for all $i = 1,\ldots,n$, since $\lambda_{\boldsymbol{\beta}}(\boldsymbol{x}_i) = e^{-\langle \boldsymbol{x}_i,\boldsymbol{\beta}\rangle} = \lambda_i$ is dependent for the vector $\boldsymbol{x}_i$ and is related to lifetime t_i, in the sample of size n of lifetimes t, $\boldsymbol{\lambda}$ is a vector related for $\boldsymbol{t}$ vector, both of size n, that is, $\boldsymbol{x}_i \mapsto \lambda_i$ no necessarily with $t_i \mapsto \lambda_i$ but, for each λ_i exist a t_i related in the sample.

Then, for $i = 1,\ldots,n$, observations from sample, we have then n vectors $\boldsymbol{x}$ compose the lines from a matrix $\boldsymbol{X}$, and if the n elements $\hat{\nu}_i = \langle \boldsymbol{x}_i, \hat{\boldsymbol{\beta}}\rangle$ form an $\hat{\boldsymbol{\nu}} = (\hat{\nu}_1, \hat{\nu}_2, \ldots, \hat{\nu}_n)'$ vector, the matrix $\boldsymbol{X}$ is such that

$$
\hat{\boldsymbol{\nu}} = \begin{bmatrix} \hat{\nu}_1 \\ \hat{\nu}_2 \\ \vdots \\ \hat{\nu}_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \ldots & x_{1p} \\ 1 & x_{21} & \ldots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \ldots & x_{np} \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{bmatrix} = \begin{bmatrix} \langle \boldsymbol{x_1}, \hat{\boldsymbol{\beta}}\rangle \\ \langle \boldsymbol{x_2}, \hat{\boldsymbol{\beta}}\rangle \\ \vdots \\ \langle \boldsymbol{x_n}, \hat{\boldsymbol{\beta}}\rangle \end{bmatrix} = \boldsymbol{X}\hat{\boldsymbol{\beta}}'\ . \tag{2.70}
$$

See then that each observation of the function $\lambda_{\boldsymbol{\beta}}$ of the definition 1 is a constant, which although varying by $\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle$, is not considered as an observation of a random variable because it enters the model as a particular parameter for its respective observation, and therefore, even if the functions in (2) are defined in $\mathbb{R}_+^* \times \mathbb{E}_{\boldsymbol{\beta}}$ the expressions derived from the density (2.66) follow the same analytical form presented in the previous section.

In the sequence, the figure 2. 11 shows survival graphs for two parametric cases for γ and $\boldsymbol{\beta}$.

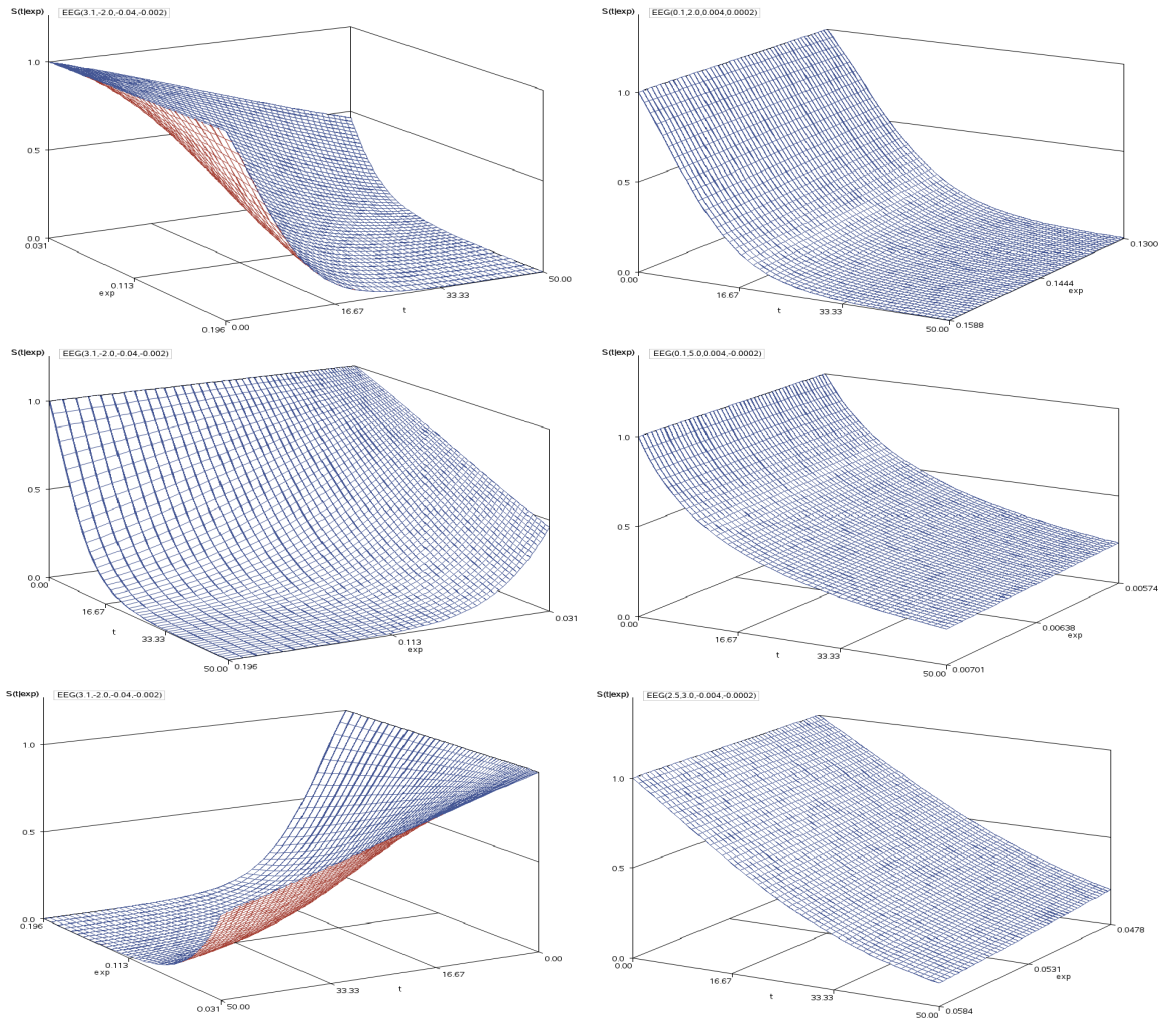

Figure 2. 11 : Survival plot for only model in colun left and diferents model in right.

Theorem 3 *Let $h(t|\boldsymbol{x}, \boldsymbol{\beta}, \gamma)$ the hazard rate functions as obtained in (2.69). For each $\boldsymbol{x}$ fixed, when $t \longrightarrow \infty$ then $h(t|\boldsymbol{x}, \boldsymbol{\beta}, \gamma) \longrightarrow \lambda_{\boldsymbol{\beta}}(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}$, this is, tends to lose memory over time and converge for $\lambda_{\boldsymbol{\beta}}(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle} \in \mathbb{R}_+^*$.*

Proof: When $t \longrightarrow \infty$, as $e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle} \in \mathbb{R}_+^*$ is a constant, see that $\exp(-te^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}) \longrightarrow 0$. Then

$$\lim_{t \longrightarrow \infty} h(t|\boldsymbol{x}, \boldsymbol{\beta}, \gamma) = \lim_{t \longrightarrow \infty} \frac{e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}}{1 - (1 - \gamma) \exp(-te^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle})} = \frac{e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}}{1 - (1 - \gamma)0} = e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle} \ , \quad (2.71)$$

the which proves the theorem. ■

Note that $\boldsymbol{x}$ is a vector of size $p + 1$ associated with a time t in the sample of size n. Then, when $h(t|\boldsymbol{x}, \boldsymbol{\beta}, \gamma) \longrightarrow \lambda_{\boldsymbol{\beta}}(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}$, for each vectors $\boldsymbol{x}$ associated with one t of

the sample, we have that geometrically hazard rate functions h tends for a curve in $\mathbb{R}^2$, whose domain is time t, that which is perfectly reasonable becouse in the the presence of covariates the density function g can be understood as a function of $\mathbb{R}^2_+ \setminus \{(0,0)\}$ in $]0,1[$, that is, $(t, \lambda_{\boldsymbol{\beta}}(\boldsymbol{x})) \mapsto]0,1[$.

The hazard rate plots show that the criteria for the form of the hazard rate function, established by the γ parameter, are still met in the presence of covariates. The figure 2. 12 in the sequence shows the hazards rates graphs for two cases for γ and $\boldsymbol{\beta}$.

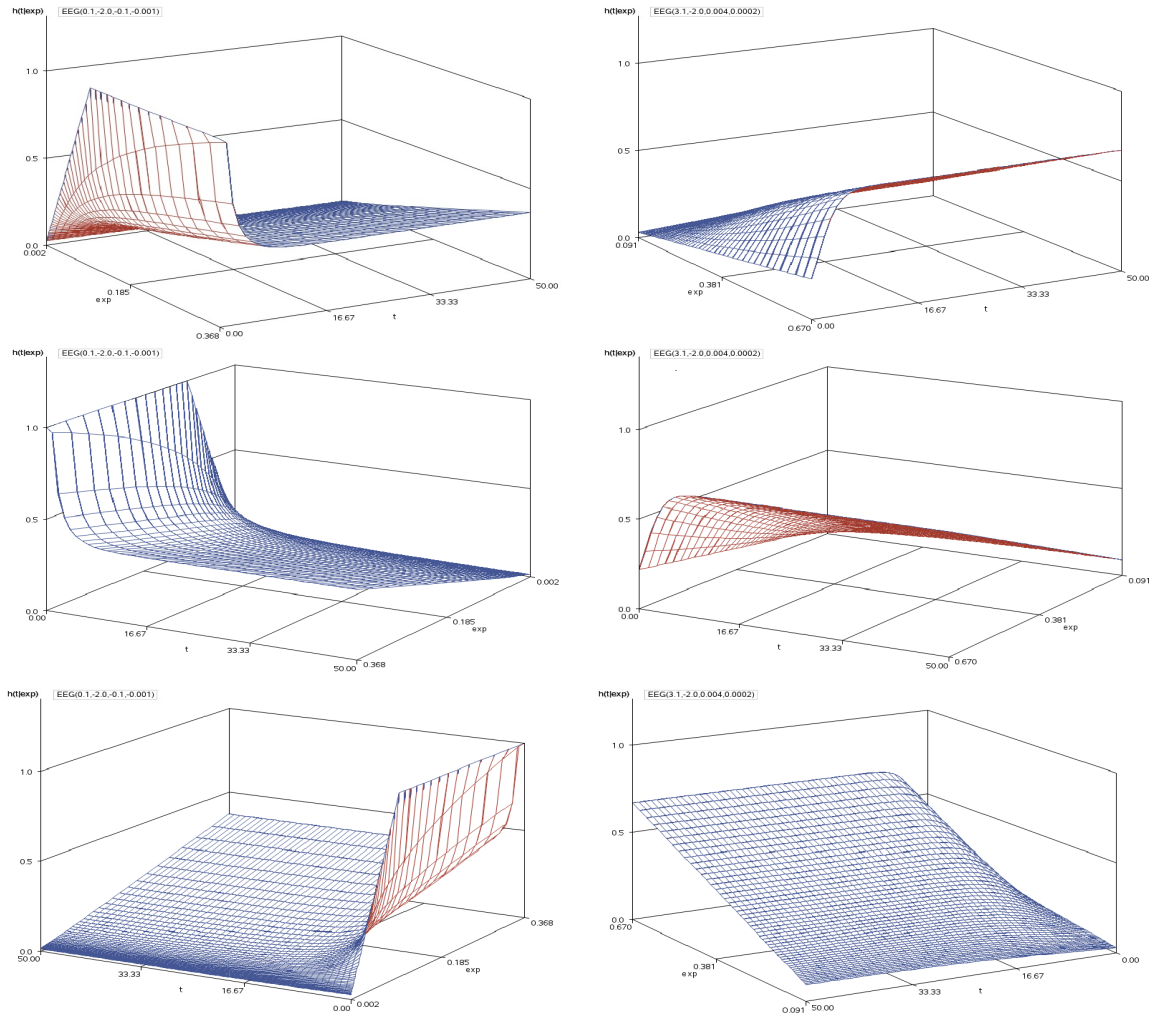

Figure 2. 12 : Hazards rates plot for only model in colun left and in colun left.

Note from figure that the nonlinear function $\lambda_{\boldsymbol{\beta}}$, that constitutes the additive effect index of the model, does not influence the shape of the hazard rate function, independent of the number of covariates present in $\lambda_{\boldsymbol{\beta}}$, that is, the number of additional parameters in the model through $\boldsymbol{\beta}$ vector.

As presented in the (1) theorem, the rth moment for the EEG distribution in the presence of covariates can be then written as

$$E(T^r|\gamma,\boldsymbol{x},\boldsymbol{\beta}) = \begin{cases} \dfrac{\gamma r! Li_r(1-\gamma)}{(1-\gamma)e^{-r\langle \boldsymbol{x},\boldsymbol{\beta}\rangle}}, \text{when } \gamma < 1 \\ \dfrac{r}{e^{-\langle \boldsymbol{x},\boldsymbol{\beta}\rangle}} E(T^{r-1}|1,\boldsymbol{x},\boldsymbol{\beta}), \text{ when } \gamma = 1 \\ Pathological, \text{ when } \gamma > 1 \end{cases} . \tag{2.72}$$

While in the absence of covariables the moment r th provides a global measure for the random variable, since $\mu_r = E(T^r|\gamma,\boldsymbol{x},\boldsymbol{\beta})$ for each of the n observations, in the presence

of covariables this measure can be summarized as a local measure due to the sample of size n, whose information can be summarized as a position measure for the vector $\vec{\mu}_r$, such as median (M_d) and geometric mean (M_g) as central trend statistics and the norm of the vector $\vec{\mu}_r$, which mathematically can be interpreted as a length, that is, the distance from its end point to the origin.

In the case where $r = 1$, we will have for example that $\vec{\mu}_1 = (\mu_{1,1}, \ldots, \mu_{n,1})$, then

$$M_d(\vec{\mu}_1) = \begin{cases} \dfrac{\mu_{\frac{n}{2},1} + \mu_{\frac{n}{2},1}}{2}, \text{ if } n \text{ par} \\ \mu_{\lceil \frac{n}{2} \rceil,1}, \text{ if } n \text{ impar} \end{cases}, \; M_g(\vec{\mu}_1) = \sqrt[n]{\prod_{i=1}^{n} \mu_{i,1}} \;\text{ and }\; \|\vec{\mu}_1\|_2 = \sqrt{\sum_{i=1}^{n} \mu_{i,1}^2} \; . \quad (2.73)$$

Thus, when it exists, the rth moment for the EEG distribution in the presence of covariates can be interpreted as the r th moment for the i th observation among the n observations in the sample. Furthermore, in this condition there will be a vector $\vec{\mu}_r$ with n elements where so that each element will be a value $E(T^r|\gamma, \boldsymbol{x}, \boldsymbol{\beta})$ under parameter $\lambda_{\boldsymbol{\beta}}(\boldsymbol{x}_i)$.

Thus, considering the inverse of the equation (2.67), a sample of n quantiles of the EEG distribution in the presence of covariates is given by the quantile function

$$Q_{EEG}(q|\boldsymbol{X}, \boldsymbol{\beta}, \gamma) = -\frac{1}{e^{-\langle \boldsymbol{x}, \boldsymbol{\beta} \rangle}} \ln \left[\frac{1-q}{1-(1-\gamma)q} \right] \; , \quad (2.74)$$

which, less than the covariant term, is analogous to the (2.9) function.

Furthermore, assuming survival data in the presence of covariates and censored data, the likelihood function for the parameter vector $(\gamma, \boldsymbol{\beta})$ is given by

$$L(\gamma, \boldsymbol{\beta}|\boldsymbol{t}, \boldsymbol{X}, \boldsymbol{\delta}) = \prod_{i=1}^{n} \frac{\gamma e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta} \rangle}}{\left[1-(1-\gamma)\exp\left(-t_i e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta} \rangle}\right)\right]^{\delta_i + 1}} \; . \quad (2.75)$$

The maximum likelihood estimates are obtained by maximizing the logarithm of likelihood function given by

$$\begin{aligned} l(\gamma, \boldsymbol{\beta}|\boldsymbol{t}, \boldsymbol{X}, \boldsymbol{\delta}) = & \, n \ln(\gamma) - \\ & - \sum_{i=1}^{n} \left\{ (\delta_i + 1) \ln\left[1-(1-\gamma)\exp\left(-t_i e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta} \rangle}\right)\right] + t_i e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta} \rangle} \right\} \; . \end{aligned} \quad (2.76)$$

From $\dfrac{\partial}{\partial \lambda} l(\gamma, \boldsymbol{\beta}|\boldsymbol{t}, \boldsymbol{X}, \boldsymbol{\delta})\Big|_{(\gamma,\boldsymbol{\beta})=(\hat{\gamma},\hat{\boldsymbol{\beta}})} = 0$ and $\dfrac{\partial}{\partial \beta_j} l(\gamma, \boldsymbol{\beta}|\boldsymbol{t}, \boldsymbol{X}, \boldsymbol{\delta})\Big|_{(\gamma,\boldsymbol{\beta})=(\hat{\gamma},\hat{\boldsymbol{\beta}})} = 0$ for $j = 0, 1, \ldots, p$, we get the system of likelihood equations

$$\begin{cases} \dfrac{n}{\hat{\gamma}} - \displaystyle\sum_{i=1}^{n} \frac{(1+\delta_i)\exp(-t_i e^{-\langle \boldsymbol{x_i}, \hat{\boldsymbol{\beta}} \rangle})}{1-(1-\hat{\gamma})\exp(-t_i e^{-\langle \boldsymbol{x_i}, \hat{\boldsymbol{\beta}} \rangle})} = 0 \\ \displaystyle\sum_{i=1}^{n} t_i x_i \left[\frac{(1-\hat{\gamma})(2+\delta_i) t_i e^{-\langle \boldsymbol{x_i}, \hat{\boldsymbol{\beta}} \rangle}}{1-(1-\hat{\gamma})\exp(-t_i e^{-\langle \boldsymbol{x_i}, \hat{\boldsymbol{\beta}} \rangle})} \right] = 0 \end{cases} , \quad (2.77)$$

whose solutions provide the maximum likelihood estimators of the parameters vectors $(\gamma, \boldsymbol{\beta}) = (\gamma, \beta_0, \ldots, \beta_p)$, and its second-order score functions provide Fisher's observed information matrix analogous to (2.47).

By considering the bayesian approach it is assumed as prior distributions Gamma for γ and Multivariate Normal distributions for $\boldsymbol{\beta}$ vector, with $\gamma \in \mathbb{R}_+^*$ and $\beta_j \in \mathbb{R}$, such that

$$\gamma \sim \Gamma(a, b) \quad \text{and} \quad \boldsymbol{\beta} \sim \mathrm{N_p}(\boldsymbol{\mu}, \mathrm{Sigma}) \ , \tag{2.78}$$

by which

$$\pi(\gamma, \boldsymbol{\beta}) \propto \gamma^{a-1} \exp\left(-b\gamma - \frac{1}{2}\sum_{j=0}^{p} \frac{\beta_j^2}{\sigma_j^2}\right) \ , \tag{2.79}$$

where a, b and σ_j^2 for $j = 0, 1, \ldots, p$ are known hyperparameters. In this work we assume that a, b and σ_j^2 such that (2.75) becomes a flat prior.

The joint posterior distribution for γ and $\boldsymbol{\beta}$ is proportional to the product of the likelihood function (2.75) and the prior distribution (2.79), resulting in

$$p(\gamma, \boldsymbol{\beta}|\boldsymbol{t}, \boldsymbol{X}, \boldsymbol{\delta}) \propto \frac{\gamma^{a-1}}{e^{b\gamma}} \exp\left(-\frac{1}{2}\sum_{j=0}^{p} \frac{\beta_j^2}{\sigma_j^2}\right) \left\{\prod_{i=1}^{n} \frac{\exp\left(\delta_i \langle \boldsymbol{x_i}, \boldsymbol{\beta}\rangle - t_i e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta}\rangle}\right)}{\left[1 - (1-\gamma)\exp\left(-t_i e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta}\rangle}\right)\right]^{\delta_i + 1}}\right\} \ .$$

The full conditional posterior distributions for γ and $\boldsymbol{\beta}$, respectively, are given as follows:

$$p(\gamma|\boldsymbol{t}, \boldsymbol{X}, \boldsymbol{\delta}, \boldsymbol{\beta}) \propto \prod_{i=1}^{n} \frac{\gamma^{a-1} e^{-b\gamma}}{\left[1 - (1-\gamma)\exp\left(-t_i e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta}\rangle}\right)\right]^{\delta_i + 1}} \ , \tag{2.80}$$

and for all $j = 0, 1, \ldots, p$

$$p(\beta_j|\boldsymbol{t}, \boldsymbol{X}, \boldsymbol{\delta}, \lambda) \propto \exp\left(-\frac{\beta_j^2}{2\sigma_j^2}\right) \prod_{i=1}^{n} \frac{\exp\left(\delta_i \langle \boldsymbol{x_i}, \boldsymbol{\beta}\rangle - t_i e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta}\rangle}\right)}{\left[1 - (1-\gamma)\exp\left(-t_i e^{-\langle \boldsymbol{x_i}, \boldsymbol{\beta}\rangle}\right)\right]^{\delta_i + 1}} \ . \tag{2.81}$$

Since the conditional distributions of γ and $\boldsymbol{\beta}$ are not easily identified, we use the Metropolis-Hastings algorithm (see Gamerman and Lopes [14]) to simulate the quantities of interest.

2.3.1 A Simulation Study with Censorship and Covariates

We have that N samples of size n from a EEG distribution in the presence of covariates and with rate τ_δ under randomly censored data in each their, can be randomly generated according to following steps:

Step 1: Fix a value for $\gamma > 0$ and the $p + 1$ values $\beta_i \in \mathbb{R}$, with $i = 0, 1, \ldots, p$, for form the vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)$.

Step 2: Set a numbers of N, n and m, respectively, the number of samples, the size of each their and cases of the parametric vector $(\gamma, \beta_0, \ldots, \beta_p)$ in the simulation.

Step 3: Set the τ_δ proportion of censorships for each of the N simulation samples so that $n\tau_\delta$ is the exact number of censorships determined in each sample, where δ_i is the i-th observation censored in a censorships vector, so that

$$\delta_i = \begin{cases} 1, & \text{if } t_i \text{ is time failure} \\ 0, & \text{otherwise} \end{cases} \ . \tag{2.82}$$

Step 4: Generate p random variables X_j with size n, in the wich $X_j \sim \mathbb{P}(\cdot|\boldsymbol{\theta})$, $j = 1, \ldots, p$, were $\mathbb{P}$ is a particular j-th probability distribution and θ are its parameters, also previously set.

Step 5: Use the values obtained in step 2 for generate the n vectors $\boldsymbol{x}_k$, $k = 1, \ldots, n$, for with $\boldsymbol{\beta}$ defined in the step 1, calculate os n values $\langle \boldsymbol{x}_k, \boldsymbol{\beta} \rangle$, were

$$\boldsymbol{x}_k = (1, x_{k,1}, x_{k,2}, \ldots, x_{k,p}) \quad \text{and} \quad \boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \ldots, \beta_p) \ . \tag{2.83}$$

Step 6: Generate the nN values from $q \in \ ]0, 1[$ and build each of the N samples of the n size EEG distribution, where each $t_k = Q(q_k | \boldsymbol{x}_k, \boldsymbol{\beta}, \gamma)$, $k = 1, \ldots, n$, according to with (2.74), is such that $F(t_k | \boldsymbol{x}_k, \boldsymbol{\beta}, \gamma) = q_k \in \ ]0, 1[$, that is

$$t_k = \frac{1}{e^{-\langle \boldsymbol{x}_k, \boldsymbol{\beta} \rangle}} \ln \left[\frac{1 - q_k}{1 - (1 - \gamma) q_k} \right] \ . \tag{2.84}$$

Step 7: Generate N random vectors of size n with $n\tau_\delta$ elements $\delta = 0$ and $n(1 - \tau_\delta)$ elements $\delta = 1$ to form N bivariate vectors with n elements (t, δ).

Step 8: For $k = 1, \ldots, n$, when $\delta_k = 0$, generate $q_k^* \sim Unif(0, t_k)$ and recalculate t_k as

$$t_k^* = \frac{1}{e^{-\langle \boldsymbol{x}_k, \boldsymbol{\beta} \rangle}} \ln \left[\frac{1 - q_k^*}{1 - (1 - \gamma) q_k^*} \right] \ , \tag{2.85}$$

and t_k^* will be a censored lifetime in the simulation

In this condition, each of the n vectors $\upsilon_k = (t_k, \delta_k, 1, x_{k,1}, x_{k,2}, \ldots, x_{k,p})$ represents an element in the simulation for the random variable EEG distribution in the presence of covariates for the case for a predefined parametric vector $(\gamma, \beta_0, \ldots, \beta_p)$.

For generate the under randomly censored data, we utilize the same methods used by Goodman, Li and Tiwari (2006, [17]. In the next page, the table 2. 10 and 2. 11 exibe the results for a simulation, according to the classical and bayesian approach, the models EEG for the parametric vector $(\gamma, \boldsymbol{\beta}) = (0.50, 2.00, -0.25, 1.75, -0.75)$ and $(\gamma, \boldsymbol{\beta}) = (2.75, -0.50, 1.00, -0.75, 1.25)$, respectivaly, for the decreasing and increasing hazard rate with covariates.

The results in the next pages summarize a simulation study developed on 500 samples of sizes 10, 25, 50 and 100, in cases for τ_δ%, the censorship proportion, set at 0% (complete data), 20% and 40%. The theoretical confidence level, fixed for the classic case, and the level of credibility for the bayesian case, was 95%.

In the classic case, the maximum likelihood estimates were calculated based on the expression (2.76) for each of the 500 simulated samples, and in the bayesian case the estimates were obtained from the posteriori defined for the parameters γ and each β_j respectively through the expressions (2.80) and (2.80).

The computational process was implemented in SAS language under the procedures NLMIXED and MCMC, respectively, for the clasic and bayesian approaches, respectively, both aided by the SQL and IML procedures for the extraction and calculation of interest statistics in the simulated process, such the estimates for the parameters of interest and asymptotic confidence intervals and HPD, for the calculation of the estimation error measurements presented in sub section 2.2.3 by the expressions (2.57), (2.58), (2.59) and (2.60).

The conclusion drawn from these results, as follows, boils down to the fact that the EEG distribution in the presence of censorship and covariates still remains a flexible model, which can be approached in either the classical or bayesian context to model sample data from any size and proportion of censorship.

In tables 2. 10 and 2. 11 in the sequence we can see that the MSE for all parameters generally decrease to 0 with the increase of n, as well as their MBA and MIA.

Table 2. 10 : Results for the model $EEG(0.50, 2.00, -0.25, 1.75, -0.75)$.

Cases			Classical estimation					Bayesian estimation				
θ	τ_δ	n	$\hat{\theta}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$	$\hat{\mu}_{\hat{\theta}}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$
γ	0	10	12.676	0.976	2238.95	12.232	86.025	0.567	0.996	0.043	0.168	1.447
		25	1.377	0.974	2.377	0.950	4.058	0.590	0.986	0.068	0.201	1.261
		50	0.781	0.972	0.291	0.365	1.615	0.554	0.954	0.053	0.171	0.978
		100	0.645	0.964	0.091	0.215	0.940	0.531	0.948	0.027	0.134	0.659
	20	10	125.44	0.972	1497303	124.99	392.73	0.583	0.998	0.043	0.167	1.513
		25	1.914	0.990	6.300	1.472	6.284	0.631	0.992	0.074	0.206	1.402
		50	0.961	0.992	0.560	0.522	2.215	0.646	0.992	0.077	0.210	1.189
		100	0.742	0.982	0.161	0.290	1.216	0.604	0.982	0.053	0.172	0.899
	40	10	590.87	0.926	6389379	590.41	1130.19	0.576	0.998	0.034	0.151	1.538
		25	3.339	0.980	47.503	2.887	13.175	0.666	0.998	0.079	0.215	1.545
		50	1.327	0.986	1.802	0.887	3.510	0.674	0.996	0.097	0.233	1.354
		100	0.936	0.992	0.418	0.479	1.776	0.684	0.992	0.097	0.236	1.128
β_0	0	10	0.985	0.744	13.775	3.148	6.561	2.039	0.998	16.430	4.039	1.453
		25	1.498	0.832	13.423	3.498	3.785	2.013	0.990	16.317	4.013	1.266
		50	1.745	0.912	14.566	3.745	2.677	2.017	0.986	16.337	4.017	0.982
		100	1.885	0.918	15.330	3.885	1.826	2.014	0.971	16.315	4.009	0.714
	20	10	0.771	0.696	14.950	3.219	7.047	2.079	1.000	16.750	4.079	1.520
		25	1.691	0.858	15.102	3.691	4.172	2.137	0.998	17.306	4.137	1.407
		50	2.373	0.898	24.436	4.873	2.913	2.461	0.986	24.834	4.961	1.193
		100	1.946	0.942	15.870	3.946	2.049	2.112	0.972	17.075	4.112	0.902
	40	10	0.754	0.568	22.850	3.350	277.46	2.086	1.000	16.797	4.086	1.545
		25	1.686	0.822	15.756	0.650	3.687	4.542	0.990	17.958	4.218	1.551
		50	1.977	0.900	16.861	3.977	3.578	2.258	0.982	18.337	4.258	1.358
		100	2.045	0.948	16.728	4.045	2.349	2.244	0.976	18.172	4.244	1.132
β_1	0	10	−0.238	0.794	1.501	0.973	3.211	−0.225	0.998	0.435	0.541	2.644
		25	−0.238	0.906	0.533	0.600	1.900	−0.237	0.968	0.391	0.524	1.811
		50	−0.266	0.918	0.399	0.539	1.329	−0.260	0.960	0.359	0.524	1.307
		100	−0.242	0.940	0.302	0.496	0.925	−0.244	0.951	0.289	0.500	0.983
	20	10	−0.275	0.754	1.800	1.063	3.327	−0.241	0.996	0.466	0.554	2.772
		25	−0.233	0.886	0.597	0.638	2.048	−0.209	0.968	0.391	0.523	1.937
		50	−0.268	0.916	0.422	0.555	1.441	−0.233	0.964	0.342	0.506	1.398
		100	−0.253	0.942	0.318	0.507	1.006	−0.245	0.960	0.300	0.498	1.003
	40	10	−0.415	0.648	9.272	1.610	307.66	−0.241	1.000	0.457	0.562	2.942
		25	−0.258	0.894	0.646	0.650	2.209	−0.223	0.984	0.400	0.520	2.120
		50	−0.260	0.922	0.457	0.565	1.587	−0.235	0.964	0.364	0.517	1.564
		100	−0.277	0.934	0.368	0.534	1.117	−0.260	0.960	0.334	0.514	1.114
β_2	0	10	1.740	0.816	12.526	3.490	1.786	1.747	0.986	12.342	3.497	1.757
		25	1.759	0.910	12.387	3.509	0.998	1.760	0.960	12.380	3.510	1.036
		50	1.738	0.922	12.204	3.488	0.682	1.738	0.940	12.200	3.488	0.701
		100	1.750	0.938	12.268	3.500	0.472	1.751	0.948	12.307	3.479	0.452
	20	10	1.755	0.728	12.945	3.510	1.926	1.772	0.976	12.571	3.522	1.904
		25	1.777	0.920	12.530	3.527	1.095	1.773	0.968	12.477	3.523	1.137
		50	1.740	0.936	12.219	3.500	0.741	1.738	0.956	12.200	3.488	0.755
		100	1.759	0.918	12.329	3.508	0.515	1.758	0.934	12.324	3.508	0.528
	40	10	1.808	0.678	13.592	3.563	1.960	1.775	0.980	12.622	3.525	2.070
		25	1.739	0.900	12.308	3.489	1.194	1.741	0.966	12.278	3.491	1.278
		50	1.746	0.934	12.284	3.498	0.828	1.746	0.966	12.267	3.496	0.864
		100	1.748	0.938	12.257	3.498	0.570	1.747	0.942	12.255	3.497	0.588
β_3	0	10	−0.748	0.810	2.396	1.499	1.131	−0.713	1.000	2.159	1.463	0.803
		25	−0.731	0.902	2.236	1.483	0.631	−0.726	0.988	2.190	1.475	0.543
		50	−0.737	0.924	2.224	1.487	0.435	−0.731	0.972	2.203	1.481	0.399
		100	−0.751	0.948	2.258	1.501	0.297	−0.753	0.962	2.212	1.502	0.270
	20	10	−0.733	0.748	2.446	1.492	1.196	−0.686	0.996	2.083	1.436	0.864
		25	−0.772	0.890	2.353	1.522	0.674	−0.727	0.988	2.194	1.477	0.585
		50	−0.758	0.934	2.291	1.508	0.467	−0.716	0.962	2.158	1.466	0.427
		100	−0.749	0.952	2.254	1.500	0.323	−0.742	0.972	2.232	1.492	0.308
	40	10	−0.696	0.692	2.426	1.464	1.202	−0.625	0.992	1.915	1.375	0.927
		25	−0.761	0.886	2.339	1.512	0.736	−0.702	0.994	2.123	1.452	0.648
		50	−0.762	0.922	2.308	1.512	0.523	−0.728	0.980	2.194	1.478	0.479
		100	−0.749	0.944	2.256	1.499	0.359	−0.735	0.966	2.211	1.485	0.342

Table 2. 11 : Results for the model $EEG(2.75, -0.50, 1.00, -0.75, 1.25)$.

Cases			Classical estimation					Bayesian estimation				
θ	τ_δ	n	$\hat{\theta}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$	$\hat{\mu}_{\hat{\theta}}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$
γ	0	10	117.60	0.976	916583	115.14	364.34	3.354	0.992	1.755	1.045	8.168
		25	6.167	0.982	51.458	3.808	18.444	3.430	0.988	2.432	1.177	6.733
		50	3.929	0.961	6.976	1.704	7.699	3.197	0.960	1.937	1.024	5.051
		100	3.294	0.954	2.113	0.998	4.430	3.061	0.938	1.242	0.804	3.693
	20	10	452.16	0.954	4352995	449.60	1108.18	3.520	0.996	2.049	1.136	8.732
		25	8.484	0.992	274.75	5.987	29.201	3.672	0.994	2.715	1.268	7.529
		50	4.715	0.982	13.573	2.288	10.205	3.508	0.980	2.458	1.139	5.866
		100	5.107	0.990	15.210	2.598	11.060	3.576	0.950	2.153	1.084	4.584
	40	10	1383.43	0.884	15383476	1380.77	1992.76	3.481	1.000	1.640	1.016	8.894
		25	13.787	0.992	746.96	11.243	56.689	3.857	0.994	3.242	1.389	8.295
		50	6.863	0.992	39.261	4.257	17.070	4.187	0.990	4.314	1.609	7.442
		100	5.006	0.998	10.419	2.335	8.406	4.099	0.942	3.694	1.449	5.747
β_0	0	10	−1.125	0.775	5.218	1.917	4.602	−0.769	1.000	1.788	1.269	8.174
		25	−0.770	0.895	2.086	1.286	2.444	−0.637	0.980	1.489	1.137	6.737
		50	−0.602	0.923	1.431	1.105	1.713	−0.559	0.968	1.258	1.059	5.055
		100	−0.542	0.944	1.176	1.042	1.180	−0.519	0.952	1.111	1.019	3.695
	20	10	−1.333	0.703	6.296	2.077	4.659	−0.749	1.000	1.733	1.249	8.738
		25	−0.776	0.882	2.162	1.304	2.676	−0.619	0.996	1.434	1.119	7.533
		50	−0.626	0.900	1.527	1.129	1.871	−0.553	0.972	1.257	1.053	5.869
		100	−0.638	0.899	1.534	1.141	1.824	−0.507	0.952	1.102	1.007	4.587
	40	10	−1.093	0.702	20.384	2.668	157.78	−0.701	0.998	1.610	1.203	8.900
		25	−0.803	0.860	2.539	1.363	3.096	−0.576	0.994	1.353	1.077	8.299
		50	−0.584	0.916	1.481	1.099	2.055	−0.478	0.982	1.105	0.980	7.446
		100	−0.507	0.934	1.162	1.009	1.419	−0.446	0.958	0.999	0.946	5.750
β_1	0	10	1.019	0.825	4.659	2.035	2.237	1.063	0.992	4.434	2.063	2.120
		25	1.012	0.915	4.175	2.012	1.263	1.034	0.962	4.230	2.034	1.286
		50	1.004	0.925	4.073	2.004	0.881	1.018	0.942	4.123	2.018	0.889
		100	0.988	0.938	3.978	1.988	0.613	0.995	0.942	4.007	1.995	0.612
	20	10	1.024	0.757	5.347	2.051	110.80	1.033	0.990	4.321	2.033	2.297
		25	1.000	0.919	4.140	2.000	1.356	1.032	0.976	4.228	2.032	1.411
		50	1.016	0.951	4.129	2.016	0.954	1.033	0.964	4.188	2.033	0.975
		100	0.996	0.897	4.059	1.996	0.933	0.996	0.976	4.009	1.996	0.662
	40	10	0.917	0.712	14.427	2.417	284.75	1.047	0.996	4.377	2.047	2.523
		25	1.012	0.881	4.262	2.012	1.539	1.050	0.970	4.330	2.050	1.623
		50	1.018	0.918	4.153	2.018	1.029	1.039	0.962	4.223	2.039	1.085
		100	0.994	0.932	4.015	1.994	0.727	1.008	0.938	4.067	2.008	0.747
β_2	0	10	−0.766	0.831	2.499	1.518	1.277	−0.769	0.974	2.392	1.519	1.332
		25	−0.758	0.887	2.315	1.508	0.671	−0.759	0.924	2.312	1.509	0.714
		50	−0.746	0.941	2.254	1.496	0.459	−0.747	0.954	2.257	1.497	0.472
		100	−0.747	0.954	2.248	1.497	0.310	−0.749	0.952	2.252	1.499	0.313
	20	10	−0.774	0.757	2.635	1.537	1.361	−0.775	0.970	2.441	1.525	1.474
		25	−0.745	0.890	2.279	1.495	0.719	−0.746	0.948	2.276	1.496	0.786
		50	−0.751	0.912	2.270	1.501	0.495	−0.752	0.948	2.274	1.502	0.519
		100	−0.742	0.913	2.244	1.492	0.485	−0.746	0.946	2.247	1.496	0.340
	40	10	−0.717	0.712	2.725	1.540	1.422	−0.759	0.976	2.346	1.486	1.688
		25	−0.768	0.910	2.361	1.518	0.807	−0.766	0.976	2.344	1.516	0.911
		50	−0.745	0.930	2.256	1.495	0.533	−0.744	0.946	2.253	1.494	0.574
		100	−0.743	0.924	2.240	1.493	0.369	−0.746	0.938	2.248	1.496	0.384
β_3	0	10	1.254	0.855	6.355	2.504	0.812	1.319	0.988	6.611	2.569	0.641
		25	1.256	0.917	6.295	2.506	0.414	1.279	0.978	6.403	2.529	0.395
		50	1.251	0.933	6.263	2.501	0.288	1.265	0.946	6.332	2.515	0.279
		100	1.250	0.952	6.255	2.500	0.196	1.256	0.950	6.284	2.506	0.192
	20	10	1.270	0.785	6.450	2.520	0.818	1.335	0.990	6.697	2.585	0.692
		25	1.254	0.907	6.287	2.504	0.452	1.288	0.976	6.450	2.538	0.433
		50	1.258	0.939	6.295	2.508	0.312	1.274	0.954	6.377	2.524	0.305
		100	1.254	0.929	6.276	2.504	0.304	1.258	0.954	6.293	2.508	0.208
	40	10	1.245	0.702	6.828	2.543	0.909	1.367	0.982	6.868	2.617	0.773
		25	1.260	0.902	6.325	2.510	0.515	1.308	0.972	6.553	2.558	0.497
		50	1.245	0.937	6.233	2.495	0.340	1.270	0.970	6.354	2.520	0.338
		100	1.247	0.928	6.241	2.497	0.232	1.258	0.946	6.291	2.508	0.232

In the classic case, the random effects exerted on the response are realizations of a probability distribution across each parameter of the model with covariates. To ensure that the effects of covariates are fixed as well as the accuracy of the estimates and that this random effect remains in the particular $k-$th parameter, without the influence of the others p parameter by the estimation process, the method quadrature-adapted of Gauss-Hermite was considered to approximate the integral of the likelihood given by expression (2.76) in relation to the $k-$th estimated parameter. In addition, to perform the optimization of solutions for expression (2.76), the procedure was also implemented with the Quasi-Newton method with 300 iterations.

In the bayesian case the estimates were obtained by applying the MCMC method (Monte Carlo via Markovian Chains) under the Metropolis algorithm proposed with random walk around the chain the test values. In this case, a random perturbation initiated by ϵ around 0 is performed such that the proposed function, the density $q(\theta^*|\theta_k) = g(|\theta^* - \theta_k|)$ is dependent θ_k and be symmetrical around this value.

In this simulation the method Metropolis algorithm proposed with random walk was preferred over Metropilis-Hasting because it produces a super adaptation to the proposed density q, which drastically impacts the likelihood of simulation coverage. The MCMC process was implemented based on the Metropolis algorithm with Normal distribution for the proposed function to generate 50000 iterations and take 10000 initial discards with 4 element roughing to ensure independence, totaling 10000 independent elements resampled to the posterior sample.

As discussed for the simulation developed in the 2.2.3 sub-section, here we also observe bad estimates and therefore high estimation errors in the case of samples with 40% censorship of any size. However, it is observed in this case that as the sample size increases, the estimates tend to their real values, and thus the MSE, MBA and MIA tend to 0 and the probability of empirical coverage tends to 0.95.

2.3.2 Aplication Data for Advanced Lung Cancer with Covariates

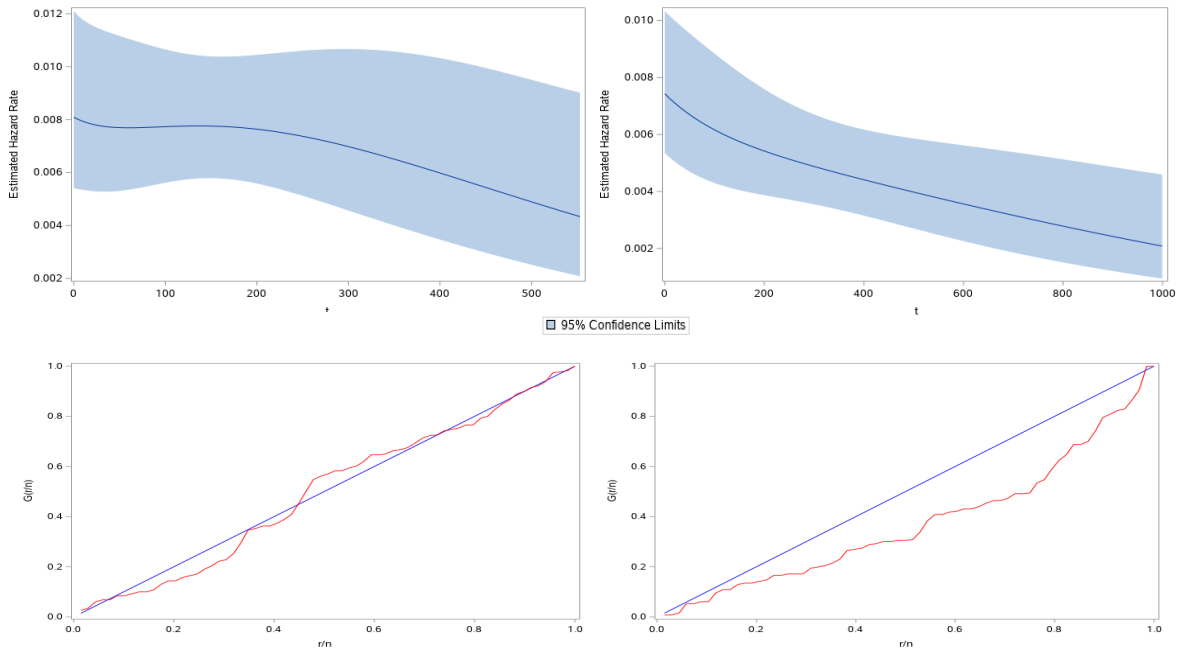

Figure 2. 13 : The empirical hazard rate function plot and TTT-plot for two agents.

For the data this application, the previous panel displays the empirical hazard rate function graph and TTT-plot for the standard agent and for the test agent, respectively,

in the left column and the right.

In the displayed panel, in the 95% confidence band display the hazards rates estimated by Epanechnikov kernel-smoothed hazard rate function is exibed, and below, in the red line on the anti-diagonal axis, the TTT-plot for both chemotherapeutic agents under effects the covariate data presented by Prantice [48].

Of particular interest are the effects of therapy on the tumor cell and the EEG model is proposed because, as we show in 2. 13 , the data manifest the decreasing hazard form that can be captured by the model-derived hazard rate function EEG.

Based on Schoenfeld residues, the table 2. 12 in the sequence presents a proportionality test for the chemotherapeutic agents of this application, where only standard agent can be treated as proportional hazard rate under 5% significance test.

The table presents the results of the proportionality test for chemotherapeutic agents in three steps: (1°) obtaining the Schoenfeld residues, (2°) the classification of survival time by the Schoenfeld residue and (3°) the correlation test between the classified survival times and the Schoenfeld residues.

Table 2. 12 : Results for proportionality test.

Chemotherapy	test-stat	p-value	test outcome for 5%
Standard	8.9810	0.1098	passed
Test	38.4444	<0.0001	not passed

However, the non-zero slope in a generalized linear regression of time-scaled Schoenfeld residues is an indication of violation of the proportional risk, or more conservatively, assumption when the correlation of Schoenfeld residues with survival are significated, the assumption of proportionality cannot be assumed.

When it comes to multiple variables, it is important to check the condition in which these covariates correlate, since if the variables are highly correlated, the inferences under the model may be erroneous or unreliable, and therefore all estimates for the parameters in the probabilistic model are unreliable.

If there is no relationship between them, we say they are orthogonal and since we have a set of 5 covariates in the problem, a multicollinearity diagnosis to quantify the linear relationship between one covariable and the others is indispensable.

Multicollinearity attributes a problem in model fit and we can diagnose it by using the variance inflation factor (VIF). Table 2. 13 presents the VIF of each covariate of the model.

Table 2. 13 : VIF for covariables by agents.

Chemotherapy	x_1	x_2	x_3	x_4	x_5
Standard	1.0300	1.0523	1.3134	1.0287	1.2645
Test	1.1224	1.0816	1.2520	1.0477	1.2264

In the literature, VIF is indicative of multicollinearity problems if VIF is greater than 10. However, some authors designate that the maximum level for VIF is 5.

Above all, the results of table 2. 13 show that in both chemotherapeutic agents, each of the covariates manifests the $VIF < 1.32$, indicating that there is no multicollinearity problem in the covariates.

In the next page, the maximum likelihood method, the estimates for the parameters of the EEG and Cox-Splines$_2$ models obtained with 5% of significance, are shown. Note that parameters κ_1 and κ_2 will be contained unconditionally in the model, because through the spline function, are parameters for node.

In the sequence, the information criterions Akaike (AIC), Akaike corrected (AICC) and the Schwarz Bayesian (BIC) are show.

Table 2. 14 : Results estimations for the models.

Model	$\boldsymbol{\theta}$	$\hat{\boldsymbol{\theta}}$	std. err.	$IC_{95\%}(\theta)$	test-stat	p-value
Cox-Splines$_2$	κ_1	-3.3954	1.1092	(-5.4695; -1.1213)	-	-
	κ_2	1.0148	0.0964	(0.8258 ; 1.2038)	-	-
	β_1	-0.0252	0.1151	(-0.2509; 0.2005)	0.0500	0.8267
	β_2	-0.0252	0.0090	(-0.0429; -0.0075)	7.7900	0.0052
	β_3	-0.0023	0.0200	(-0.0415; 0.0368)	0.0100	0.9071
	β_4	-0.0012	0.0125	(-0.0258; 0.0233)	0.0100	0.9211
	β_5	0.3345	0.3101	(-0.2733; 0.9422)	1.1600	0.2808
EEG	γ	0.7359	0.3375	(0.0627; 1.4091)	2.1800	0.0326
	β_0	3.2190	1.0556	(1.1132; 5.3248)	3.0500	0.0032
	β_1	0.0270	0.1257	(-0.2238; 0.2779)	0.2200	0.8303
	β_2	0.0270	0.0095	(0.0080; 0.0461)	2.8300	0.0060
	β_3	0.0017	0.0207	(-0.0396; 0.0430)	0.0800	0.9337
	β_4	0.0023	0.0134	(-0.0243; 0.0290)	0.1800	0.8610
	β_5	-0.3563	0.3298	(-1.0142; 0.3016)	-1.0800	0.2837

The calculated criteria and presented in the table 2. 15 allow us to observe that, although the differences are not significant between the two models, the EEG model presents the values of the minor information criteria and therefore the EEG model is indicated as the preferred model for modeling covariate lifetime data.

Table 2. 15 : Information criterions.

Model	-2Log	AIC	AICC	BIC
Cox-Splines$_2$	734.9	750.9	753.2	768.6
EEG	734.4	748.4	750.2	764.0

Then, through the estimated parameters for the EEG model, presented in table 2. 14 , we can also describe the algebraic expression for the nonlinear function adjusted $\lambda_{\boldsymbol{\beta}}$ as

$$\lambda_{\hat{\boldsymbol{\beta}}}(\boldsymbol{x}) = \exp[-3.219 - 0.027(x_1 + x_2) - 0.001x_3 - 0.0023x_4 + 0.3563x_5] \tag{2.86}$$

and is defined $\lambda_{\hat{\boldsymbol{\beta}}}$, we get

Table 2. 16 : Statistics for Distribution $\lambda_{\hat{\boldsymbol{\beta}}}$.

Minimum	Maximum	Median	Quartile Range	n
0.0027	0.0258	0.0067	0.0062	69

the frequency distribution for $\lambda_{\hat{\boldsymbol{\beta}}}$ based in the sample data for the X_j covariates, through the values $x_{i,j}$, for $i = 1, \ldots, 69$ and $j = 1, \ldots, 5$.

The following image shows the geometric representation for the previous frequency distribution.

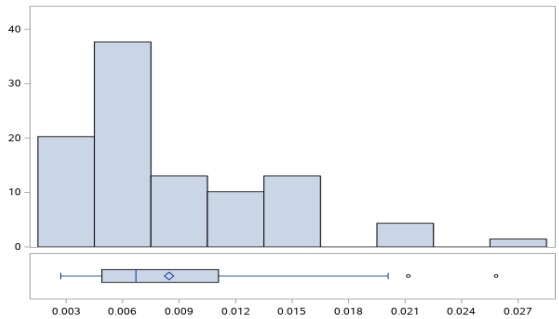

Figure 2. 14 : Frequency distribution for $\lambda_{\hat{\boldsymbol{\beta}}}$ by the sample of covariates.

In this condition, the comparison turns to prediction errors measures, such as the mean bias absolute (MBA) and mean square error (MSE) with presented in 2.57 and 2.58, respectivaly, but defined here for the two models as the differences weighted by number of terms between the estimated for respective fitted survival function value and or corresponding value empirical obtained through of the Kaplan-Mayer. The results are shown in the following table.

Table 2. 17 : Precision.

Model	MBA	MSE
Cox-Splines$_2$	0.0002	0.0138
EEG	0.0155	0.0985

From the results of the previous table, see that while the values obtained are both low because the calculation values are decimal numbers less than 1, the Cox-Splines$_2$ model actually prevails over the EEG in this numeric comparison.

However, in the survival analysis, the interest is also in the adjustment of the hazard rate function and since Cox's proportional hazard rate model directly models the hazard rate function, it is essential still to verify whether the shape of the hazard rate function is appropriate to the form suggested by the empirical data, as shown in the figure 2. 13 .

The graphical comparison for fitting the Cox-Splines$_2$ model with the empirical model is shown in figure 2. 15 below and outlines how well this model fit. The describe how it fits the empirical survival model and the its shape of hazard rate function.

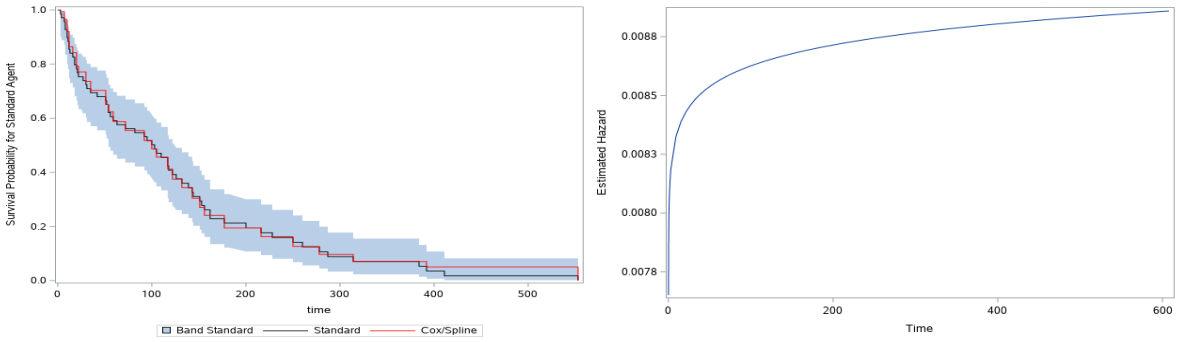

Figure 2. 15 : The adjusted survival and hazard rate functions by Cox-Splines$_2$.

As a consequence of this distribution, we are able, in addition to plotting the hazard rate surface for the EEG model in the presence of covariates, to obtain the optimal curve to represent the hazard rate function geometrically for the data of patients with advanced lung cancer under the influence of standard chemotherapeutic agent.

The following image shows the hazard rate surface according for the $\lambda_{\hat{\beta}}$ frequency by the sample of covariates and the optimal curve obtained with the median displayed in table 2. 16 .

Without difficulty, it turns out that the estimated risk curve captures the true form of the corresponding risk function shown in figure 2. 13 , and the conclusion that the EEG model is preferable in this case is immediate, although the great difference between the survival curves presented by the table 2. 17 .

Now, analogous to what has been developed so far for the standard chemotherapeutic agent, what follows are the adjustments to the hazard rate functions of the EEG and Weibull model for the patients data treated with the test chemotherapeutic agent.

As shown in table 2. 12 , these data do not manifest proportional hazards rates, so in this fit considered a parametric model to contrast the fit of the EEG model.

However, in this case the estimates for the parameters of the two models were obtained by the MCMC method. Table 2. 18 following shows the estimates obtained, the standard

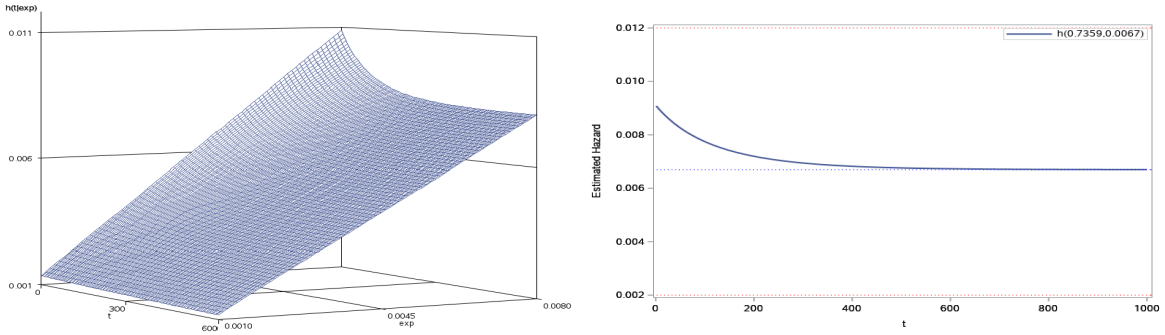

Figure 2. 16 : The hazard rate surface and the hazard rate optima curve adjusted.

errors, and the HPD interval with credibility 95% for each of the model estimates in the presence of covariates.

In the same table, with a significance level of 5%, the stationarity test results for the MCMC process convergence performed for this problem are presented for each parameter in table 2. 18 . The data in the table show that convergence was achieved for each parameter under chain markovian of 50000 iterations, 10000 burning and thinning for 4 unit.

Table 2. 18 : Results for MCMC process with 5% credibility for models.

Cases		Results estimations for the models			Results from convergence test		
Model	$\boldsymbol{\theta}$	$\hat{\boldsymbol{\theta}}$	std. err.	$IC_{95\%}(\theta)$	test-stat	p-value	test outcome
EEG	γ	0.8931	0.3923	(0.2406; 1.6822)	0.0495	0.8796	passed
	β_0	2.1204	0.6876	(0.7478; 3.4154)	0.0783	0.7018	passed
	β_1	-0.2251	0.1067	(-0.4289; -0.0081)	0.2983	0.1366	passed
	β_2	0.0371	0.0061	(0.0250; 0.0490)	0.3340	0.1087	passed
	β_3	-0.0055	0.0107	(-0.0256; 0.0161)	0.0919	0.6259	passed
	β_4	0.0125	0.0105	(-0.0074; 0.0341)	0.0903	0.6348	passed
	β_5	0.5890	0.3260	(-0.0321; 1.2335)	0.3626	0.0909	passed
Weibull	μ	1.0364	0.1028	(0.8438; 1.2459)	0.2937	0.1408	passed
	β_0	2.2451	0.6820	(0.8910; 3.5413)	0.1147	0.5174	passed
	β_1	-0.2323	0.1010	(-0.4432; -0.0476)	0.0578	0.8277	passed
	β_2	0.0366	0.0058	(0.0254; 0.0480)	0.0187	0.9981	passed
	β_3	-0.0060	0.0104	(-0.0257; 0.0149)	0.2284	0.2186	passed
	β_4	0.0113	0.0101	(-0.00915; 0.0305)	0.0945	0.6123	passed
	β_5	0.5868	0.3172	(-0.0280; 1.2059)	0.1372	0.4310	passed

Based on these estimates, replacing them in the theoretical model, we then have the estimated models EEG and Weibull adjusted for the test chemotherapeutic agent data.

The table 2. 19 in the sequence shows the results for the information criteria for the bayesian models adjust resulting.

Table 2. 19 : Measures for fit.

Model	p_D	$\bar{D}(\theta)$	$D(\hat{\theta})$	DIC
EEG	5.66	711.30	705.64	716.96
Weibull	6.40	711.04	704.64	717.44

In the subsection 2.2.4 (page 37) a brief discussion is made about these measures, and analogously, it is also noted here that no relevant information about the two adjusted models is acquired. Although estimates indicate that an adjustment for the EEG model would be more appropriate because it manifests the smallest p_D and DIC, the difference of one unit in these measures does not allow affirm, statistically, to state which is the most appropriate fit for the data between the two models contrasted.

As we saw in the previous case of this application, it is possible to obtain estimates that allow an excellent adjustment to the survival model but that diverge from the true risk model. Therefore, due to the tie of the DICs for the adjustment of the EEG and Weibull models, we will consider comparing the shape of the hazard rate function for these models through their respective risk surface and the risk curve under the optimal value for the source covariate of variation.

In this case, denoting by $\langle \boldsymbol{x}, \boldsymbol{\beta_1} \rangle$ and $\langle \boldsymbol{x}, \boldsymbol{\beta_2} \rangle$ the covariate variation source of the EEG and Weibull model, respectively, we have to the exponentialized estimates for their values are obtained through the following functions

$$\langle \boldsymbol{x}, \hat{\boldsymbol{\beta}}_1 \rangle = 2.1204 - 0.2251x_1 + 0.0371x_2 - 0.0055x_3 + 0.0125x_4 + 0.5890x_5 \ , \quad (2.87)$$

$$\langle \boldsymbol{x}, \hat{\boldsymbol{\beta}}_2 \rangle = 2.2451 - 0.2323x_1 + 0.0366x_2 - 0.0060x_3 + 0.0113x_4 + 0.5868x_5 \ , \quad (2.88)$$

whose exponentials through $\lambda_{\hat{\boldsymbol{\beta}}_1}(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, \hat{\boldsymbol{\beta}}_1 \rangle}$ and $\lambda_{\hat{\boldsymbol{\beta}}_2}(\boldsymbol{x}) = e^{-\langle \boldsymbol{x}, \hat{\boldsymbol{\beta}}_2 \rangle}$, respectivaly, provide the distribution

Table 2. 20 : Statistics for Distribution $\lambda_{\hat{\boldsymbol{\beta}}}$.

Model	Minimum	Maximum	Median	Quartile Range	n
EEG	0.0016	0.0621	0.0091	0.0140	69
Weibull	0.0016	0.0605	0.0091	0.0139	69

In the sequence, the figure shows the frequency distribution for $\lambda_{\hat{\boldsymbol{\beta}}_1}$ and $\lambda_{\hat{\boldsymbol{\beta}}_2}$ adjusted, in the left with asymmetry by the EEG model and on the right forming an absolutely descending curve by the Weibull model, both by the covariate sample.

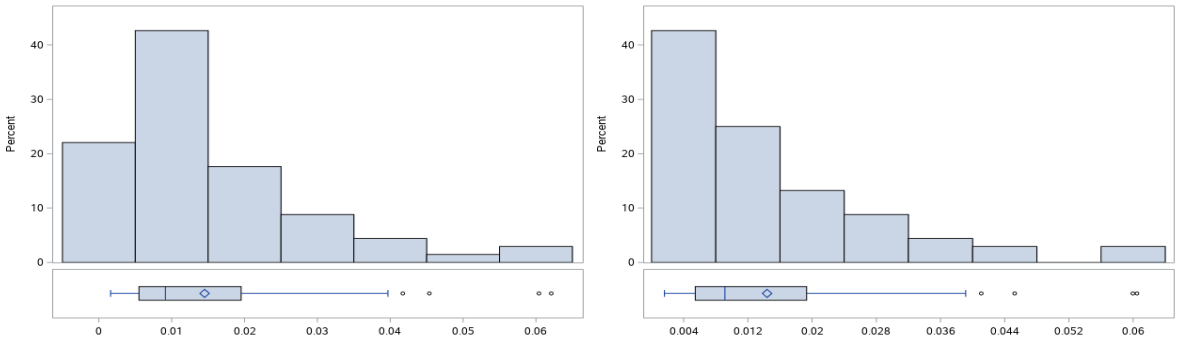

Figure 2. 17 : Frequency distribution for $\lambda_{\hat{\boldsymbol{\beta}}_1}$ and $\lambda_{\hat{\boldsymbol{\beta}}_2}$ by the sample of covariates.

Then, based on this distributions, the surface and optimal hazard rate curve for the fitted models can be evaluated geometrically witch follows and we have so ensure that the adjustment made by the estimated model EEG is the also most appropriate in these contrast for patient data on the treatment of the test chemotherapeutic agent.

Therefore, on the data from study presented by Prentice [48], the adjustment developed under the EEG model is more efficient than the cox-spline$_2$ semiparametric model for the standard chemotherapeutic agent and also for the Weibull parametric model for the test chemotherapeutic agent.

Follow, to apply this adjustment, it is considered a scenario in which patients under these treatments are at time domain hazard rate with limit set at 228 and 242, respectively in the case of treatment chemotherapeutic under standard and testing agents. The empirical hazard rate rates by the discrete kernel of Epanechnikov are shown in the following image under the range of 95% confidence.

The image in next page shows the previously estimated models in the your respective cases. They are properly adjusted in this time domain as shown in the red curves.

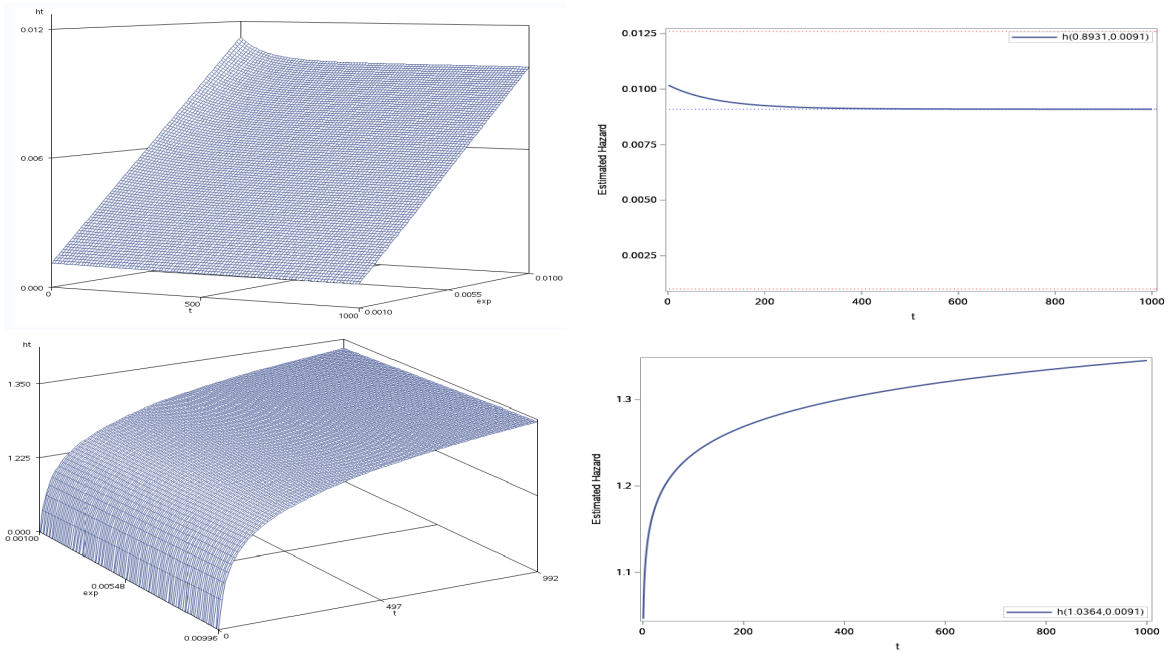

Figure 2. 18 : The hazard rate surface and the hazard rate optima curve adjusted.

In the sequence, in their respective scenarios the graphs show that, in the case of the standard agent the risk curve remains sufficient over the risk rate data, as well as in the case of the test agent an oscillation around the median curve can be considered in search of the ideal curve.

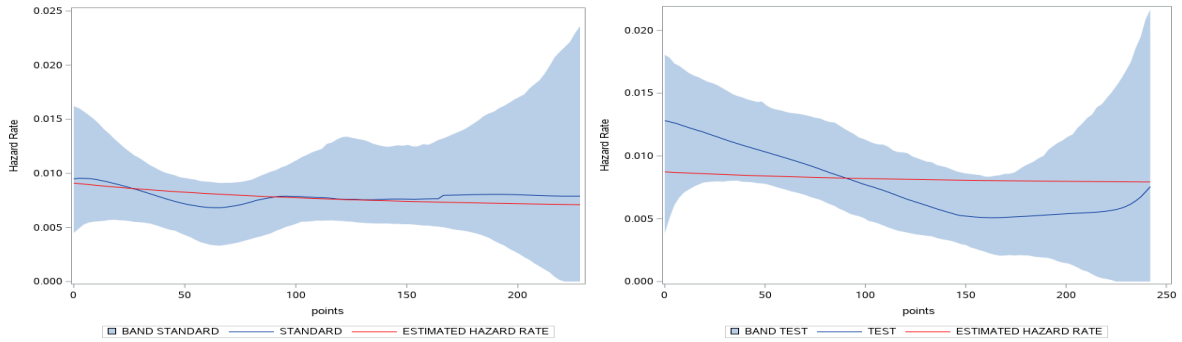

Figure 2. 19 : Adjustment for the maximum time ten patients remain in hazard rate.

Since under two sources of variation the risk model generates a risk surface, it is reasonable to consider taking an oscillation around the median as long as this oscillation remains within the limits of the sample distribution, which is not possible when Only one source of variation is modeled.

Chapter 3

The Generalized Exponential Geometric Extreme Distribution

3.1 The Generalization of the EG Distribution

Proposed by Gupta and Kundu [20] a model called Exponential Exponential (EE) was obtained as

$$\psi(x|\alpha,\lambda) = \alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1} \ , \tag{3.1}$$

where $\alpha, \lambda, x > 0$, and as developed for the composite models EG and EEG, considering the distribution for $T = \min(\{X_i\}), 1 < i < n$, with $X \sim EE(\alpha,\lambda)$ e $N \sim Geo(\theta)$, we have pdf

$$\delta_{\min}(t|n,\alpha,\lambda) = n\alpha\lambda e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}[1-(1-e^{-\lambda t})^{\alpha}]^{n-1} \ , \tag{3.2}$$

whence it results, $p(n|\theta) = \theta(1-\theta)^{n-1}$, the pdf

$$\varphi(t|\alpha,\lambda,\theta) = \frac{\alpha\lambda\theta e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}}{[\theta+(1-\theta)(1-e^{-\lambda t})^{\alpha}]^2} \ . \tag{3.3}$$

The function 3.3 is the pdf of the random variable X with distribution E2G, with $\alpha, \lambda, x > 0$ and $0 < \theta < 1$, and was proposed by Louzada, Marchi and Roman [35].

Also taking the distribution to $T = \max(\{X_i\}), 1 < i < n$, still with $X \sim EE(\alpha,\lambda)$ e $N \sim Geo(\theta)$, pdf is obtained

$$\delta_{\max}(t|n,\alpha,\lambda) = n\alpha\lambda e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}(1-e^{-\lambda t})^{\alpha(n-1)} \ , \tag{3.4}$$

with which, as a composition model, one obtains

$$\phi(t|\alpha,\lambda,\theta) = \frac{\alpha\lambda\theta e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}}{[1-(1-\theta)(1-e^{-\lambda t})^{\alpha}]^2} \ , \tag{3.5}$$

the pdf of the CE2G model proposed by Louzada, Marchi and Carpenter [34].

Analogously to the process of obtaining EEG models, assuming reparametrization $\beta = \theta^{-1}$, from fdp (3.3) e (3.5), both with $\alpha, \lambda, x > 0$ and $0 < \theta < 1$, in particular there are $0 < \theta < 1 \Rightarrow \beta > 1$ whereby we obtain the GE2 model, whose pdf is

$$f(t|\alpha,\lambda,\beta) = \frac{\alpha\lambda\beta e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}}{[\beta+(1-\beta)(1-e^{-\lambda t})^{\alpha}]^2} \ , \tag{3.6}$$

as proposed in Ristic and Kundu [54] and Ristic and Kundu [55].

Note that replacing $\theta = \beta^{-1}$, has in (3.5) the $\phi(t|\alpha, \lambda, \theta) = \phi(t|\alpha, \lambda, \beta^{-1})$, whereby through some algebraic manipulations results that

$$\frac{\alpha\lambda\beta^{-1}e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}}{[1-(1-\beta^{-1})(1-e^{-\lambda t})^{\alpha}]^2} = \frac{\alpha\lambda\beta e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}}{[\beta+(1-\beta)(1-e^{-\lambda t})^{\alpha}]^2} \ , \tag{3.7}$$

that is, it turns out that $\phi(t|\alpha, \lambda, \theta) = \varphi(t|\alpha, \lambda, \beta)$ under the parameter $\beta > 1$, the pdf (3.6) of the GE2 model as unification of the both distribution, (3.5) and (3.4).

Consequently, by reparametrization, the pdf of the CE2G distribution is the same as the E2G distribution. The main, and hitherto unnoticed, joint contribution of Louzada, Marchi e Roman [35] and Louzada, Marchi and Carpenter [34] was the obtaining of distributions whose pdf coincide through the reparametrization $\beta = \theta^{-1}$.

Therefore, unifying both distribution, (3.3) and (3.7), and noting that both reduce the Exponential distribution when $\alpha = \beta = 1$, the random variable T is said to have an Exponential Geometric Extreme Distribution (GE2) if its probability density function is given by

$$f(t|\alpha, \lambda, \beta) = \frac{\alpha\lambda\beta e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}}{[\beta+(1-\beta)(1-e^{-\lambda t})^{\alpha}]^2} \ , \tag{3.8}$$

were $\alpha, \lambda, \beta > 0$, which contributes to the natural interpretation of the GE2 distribution.

If $\alpha = 1$ and $\beta \neq 1$, the pdf (3.8) reduces to EEG distribution Adamidis, Dimitrakopoulou and Loukas [2], it is important to point out that the credit and the main motivation for introducing this distribution goes to the Louzada, Roman and Cancho [37] and Louzada, Marchi and Roman [35]. Moreover, following those authors we obtain a significant account of mathematical properties of the GE2 distribution such as moments, rth moment of the ith order statistic and some entropy measures.

Beyond this, if $\alpha = \beta = 1$, the pdf (3.8) reduces to Exponential distribution. The distribution function for the $GE2(\alpha, \lambda, \beta)$ distribution is given by,

$$F(x|\alpha, \lambda, \beta) = \frac{(1-e^{-\lambda x})^{\alpha}}{\beta+(1-\beta)(1-e^{-\lambda x})^{\alpha}} \ . \tag{3.9}$$

The survival, hazard functions of the $GE2(\alpha, \lambda, \beta)$ distribution is given by

$$S(t|\alpha, \lambda, \beta) = \frac{\beta-\beta(1-e^{-\lambda t})^{\alpha}}{\beta+(1-\beta)(1-e^{-\lambda t})^{\alpha}} \ , \tag{3.10}$$

and

$$h(t|\alpha, \lambda, \beta) = \frac{\alpha\lambda e^{-\lambda t}(1-e^{-\lambda t})^{\alpha-1}}{\beta+[1-2\beta-(1-\beta)(1-e^{-\lambda t})^{\alpha}](1-e^{-\lambda t})^{\alpha}} \ . \tag{3.11}$$

An important characteristic from the hazard function for the GE2 model is your forms, that take the form constantly, decreasing, crescent, bathtub and unimodal. In the next page, the figure 1 present different forms for the density and hazard functions for the EEG distribution considering different values of α, λ e β in the first column charts and with λ fixed to different values of α and β in the second column charts.

For all $\lambda > 0$ whith $\alpha = \beta = 1$, result that

$$h(t|1, \lambda, 1) = \lambda \ . \tag{3.12}$$

By the figure 3. 1 in the sequence, note that for values the $\alpha \in \ ]0, 1[$ and $\beta \in \ ]1, \infty[$, the risk function takes the form of the bathtub, and is such that when the parameters α,

λ and β increase, the curve valley (the minimum point $h(t|\alpha, \lambda, \beta)$) decreases and the risk function tends to form increasing to any λ with $\alpha \geq 1$.

When $\alpha \in]1, \infty[$ and $\beta \in]0, 1[$, the risk function takes the unimodal form and the measure that the parameters α, λ e β are decrease, the curve crest (the maximum point for $h(t|\alpha, \lambda, \beta)$ function) are increase, from which the risk function tends to decrease, reaching such a shape for any λ value with $\alpha < 1$.

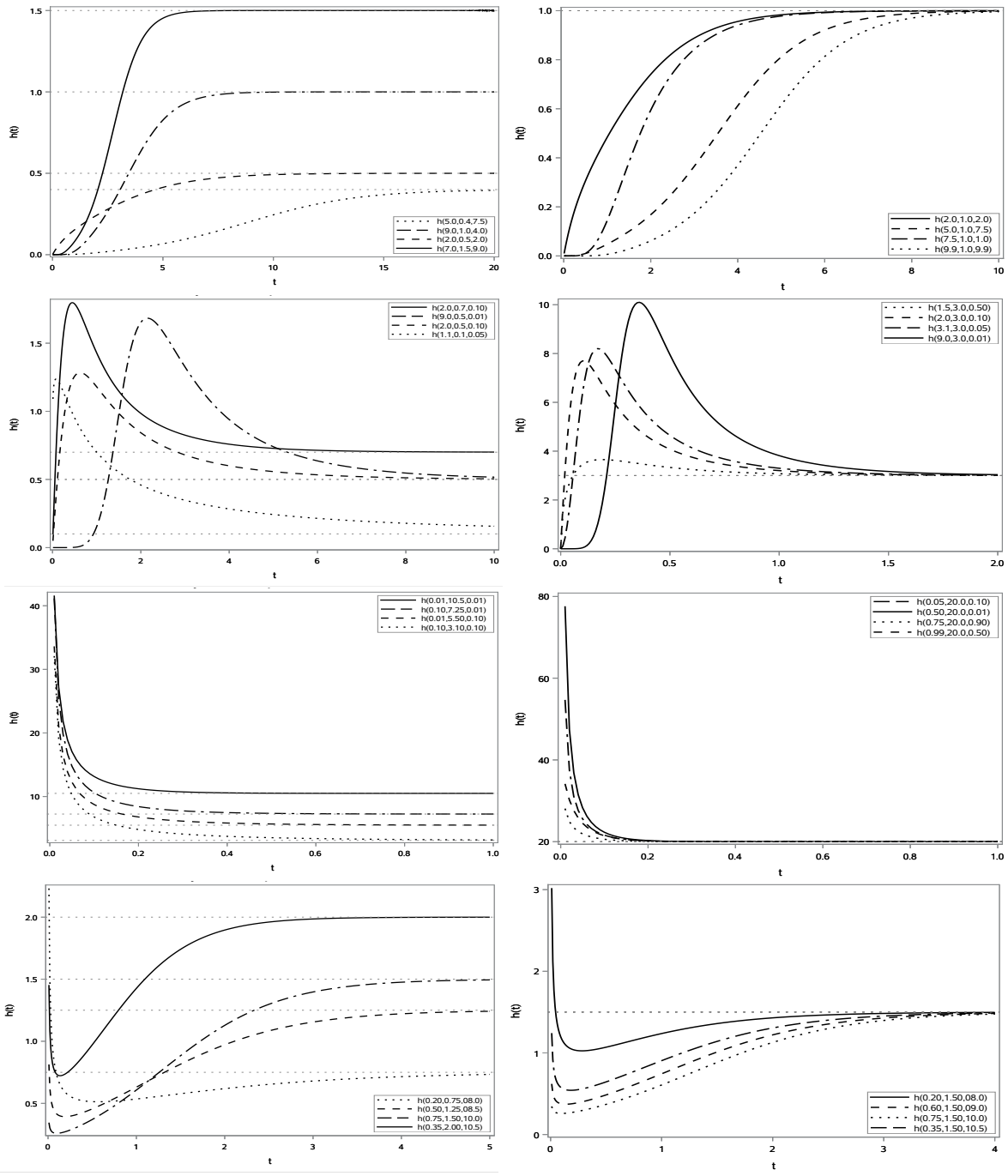

Figure 3. 1 : In panel, the forms for hazard functions of the GE2 distribution.

Note that as the λ parameter increases the maximum value of the risk rate in the time interval it also increases, that the highlights a scalar behavior in the risk function, that is, λ represents a scale parameter in the model. Similarly, the measure that α and β vary, the risk function changes shape to a fixed λ highlighting a shape characteristic for these parameters.

In this condition, as in cases where $\lambda \in]0,1[$ we have that λ acts as a time rate in the GE2 model, for this, it is always necessary to check whether λ is a rate parameter or scale in the distribution.

Moreover, the figure 3. 1 also it shows that when $t \longrightarrow \infty$, $h(t|\alpha,\lambda,\beta) \longrightarrow \lambda$, this is, these result reflects that the risk function of the GE2 model tends to lose memory or not wear out as time goes by. In the following section will be formally presented the demonstration of this result in the form of a theorem.

For research in reliability analysis, for example, where almost always in durability studies it is intended to obtain small quantiles of the time that report premature failures, under the conditions described above, the GE2 distribution risk function emerges as a great alternative when you seek information that a random phenomenon, which we consider to have survived for the while t, won't have your probability of surviving altered.

As in the risk model $h(t|\alpha,\lambda,\beta)$, when $T \sim GE2(\alpha,\lambda,\beta)$ and from time t_n, the probability of the phenomenon under study to fail does not depend on how long it has been running and there is no aging or greater likelihood of failure in this period of the operation then the most relevant information about premature failure risk is contained in the time range $(0, t_n$, the ideal failure times to be modeled for $GE2$.

The quantile function of the GE2 distribution, for $q \in]0,1[$, is given by

$$Q(q|\alpha,\lambda,\beta) = -\frac{1}{\lambda}\ln\left\{1-\left[\frac{\beta q}{1-q(1-\beta)}\right]^{1/\alpha}\right\}, \tag{3.13}$$

were $\alpha,\lambda,\beta \in \mathbb{R}_+^*$.

See that, the qth quantile of the CE2G distribution is given by

$$Q_{CE2G}(q|\alpha,\lambda,\theta) = -\frac{1}{\lambda}\ln\left\{1-\left[\frac{q}{q+(1-q)\theta}\right]^{1/\alpha}\right\}. \tag{3.14}$$

Under parametrization $\theta \in]0,1[$, but using a change in the parametrization $\theta = \beta^{-1}$ in (3.14) were $\beta > 1$, and under parametrization $\beta > 1$ the pth quantile of the Q_{CE2G} can be rewritten as

$$Q_{CE2G}(q|\alpha,\lambda,\beta^{-1}) = -\frac{1}{\lambda}\ln\left[1-\left(\frac{\beta q}{1-q(1-\beta)}\right)^{1/\alpha}\right] = Q(q|\alpha,\lambda,\beta)\ , \tag{3.15}$$

were $\alpha,\lambda \in \mathbb{R}_+^*$ and $\beta \in\ ]1,\infty[$. When $\theta \in\ ]0,1[$, see occour $Q(q|\alpha,\lambda,\beta) = Q_{CE2G}(q|\alpha,\lambda,\beta)$.

The GE2 distribution arises naturally in competing risks scenarios, in which the random variable has distribution of $T = \min(\{X1,\ldots,X_M\})$ were M is a random variable with Geometrical distribution and X_i are assumed to be independent and identically distributed according to a EE distribution.

Still in the competing risks scenarios, other motivation for considering the GE2 distribution lies on the fact the parameter has a biological interpretation.

According Ristic and Kundu [55], since the GE2 is a probability distribution generated by composing a random variable M with geometric distribution with a random variable X distributed according to an EE distribution, then GE2 represents the minimum time of $X_1,\ldots,X_M$ when $0<\beta<1$, the random time of $X_1,\ldots,X_M$ for $\beta = 1$ and the maximum time of $X_1,\ldots,X_M$ when $\beta > 1$, this is, if $T \sim GE2(\alpha,\lambda,\beta)$, for all $\alpha,\lambda,\beta \in \mathbb{R}_+^*$

$$T = \begin{cases} \min\{X1,\ldots,X_M\}, \text{when } 0<\beta<1 \\ \text{random}\{X1,\ldots,X_M\}, \text{ when } \beta = 1 \\ \max\{X1,\ldots,X_M\}, \text{ when } \beta > 1 \end{cases}. \tag{3.16}$$

3.1.1 General Properties

Many of the interesting characteristics and features of a distribution can be studied through its moments, such as mean and variance. Expressions for expectation value, variance and the r-th moment on the origin of X can be obtained using the well-known formula

$$E(X^r|\alpha,\lambda,\beta) = \int_0^\infty \frac{\alpha\lambda\beta x^r e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1}}{[\beta+(1-\beta)(1-e^{-\lambda x})^\alpha]^2}dx \ . \tag{3.17}$$

The rth moment provides the most important properties of a probabilistic model, so much so that the characterization of probability distributions, where possible, is indispensable and defined by expression de $E(X^r)$, that is commonly approach in terms of f.d.p. as

$$E(X^r|\alpha,\lambda,\beta) = \int_0^\infty x^r f(x)dx \ . \tag{3.18}$$

Through this expression, in many cases, obtaining this result is not feasible due to the lack of an elementary primitive for the integrant $x^r f(x)$, mainly for obtaining high order moments, such as for obtaining asymmetry and kurtosis of model. In the best case, the desired expression is obtained after costly mathematical devices.

However, in cases of models with non-negative random variables a device that facilitates obtaining expressions of order moments $r \geq 1$ is derived in terms of the survival function in parallel with the application of Fubini's theorem, the alternative rth moment given by

$$E(X^r|\alpha,\lambda,\beta) = r\int_0^\infty x^{r-1}S(x)dx \ . \tag{3.19}$$

It is a highly advantageous method when it comes, for example, to moments of a transformed random variable, where it is considerably easier to integrate $x^{r-1}[1-F(x)]$ instead of $x^r f(x)$ (see, for exemple, Hong [28] or Chakraborti, Jardim and Epprecht [7]).

A general expression for rth ordinary moment $\mu'_r = E(X^r)$ with density function given by distribution $X \sim GE2(\alpha,\lambda,\beta)$ variable, can be obtained analytically the as follows.

At first, let's also consider that, when $Y \sim E2G(\alpha,\lambda,\theta)$ according to Louzada, Marchi and Roman [35], when $\alpha > 0$, $\lambda > 0$ and $0 < \beta < 1$ results that

$$E(Y^r|\alpha,\lambda,\theta) = \frac{\theta r!}{\lambda^r}\sum_{k,l,m,n=0}^{\infty}(-1)^{k+l}\frac{(1-\theta)^n(n-k+2)^{\underline{k}}(\alpha k+m-l+1)^{\underline{l}}}{k!(l+1)^r l!} \ . \tag{3.20}$$

In Equation (3.20), see that $(n-k+2)^{\underline{k}}$ and $(\alpha k+m-l+1)^{\underline{l}}$ are the decreasing factor power polynomials also known as "Pochhammer symbols" proposed for Pochhammer [47] (see, for exemple, Olver [46] or Qi, Shi and Liu [49]), and as $k,l,m,n \in \mathbb{N}$ with $\alpha \in \mathbb{R}$, can be rewritten in terms of fractions as

$$(n-k+2)^{\underline{k}} = \frac{(n-k+2)!}{(n-2k+2)!} = \begin{cases} \prod_{i=1}^{k-1} n-i+1, \text{if } k \geq 1 \\ 1, \text{if } k = 0 \end{cases} \ , \tag{3.21}$$

and

$$(\alpha k+m-l+1)^{\underline{l}} = \frac{\Gamma(\alpha k+m-l+2)}{\Gamma(\alpha k+m-2l+2)} = \begin{cases} \prod_{j=1}^{l-1} \alpha k+m-j+1, \text{ if } l \geq 1 \\ 1, \text{ if } l = 0 \end{cases} \ . \tag{3.22}$$

Then, when $X \sim GE2(\alpha, \lambda, \beta)$ with $\alpha, \lambda > 0$ and $0 < \beta < 1$, $X = Y \sim GE2(\alpha, \lambda, \beta) = GE2(\alpha, \lambda, \theta)$, this is, for $X = Y \sim E2G(\alpha, \lambda, \beta)$ (model proposed by [35]), with $0 < \beta = \theta < 1$ and we can ensure the following lemma

Lemma 2 *For the random variable X with $GE2(\alpha, \lambda, \beta)$ distribution, when $\alpha, \lambda > 0$, with $\alpha \neq 1$ and $0 < \beta < 1$, we have that r-th moment function not exists, this is*

$$E(X^r|\alpha, \lambda, \beta < 1) = r \int_0^\infty x^{r-1} S(x|\alpha, \lambda, \beta < 1) dx \longrightarrow \infty \ . \tag{3.23}$$

Proof: It is immediate that, when the particular case $\alpha = 1$ occurs, $GE2(\alpha, \lambda, \beta) = EEG(\lambda, \beta)$ and the demonstration considers two cases about the expression (3.20): when $n - k + 2$ e $\alpha k + m - l + 1$ are positive or negative in (3.21) and (3.22).

When they are positive, their equality in (3.21) and (3.22) is satisfied, but when they are negative $(n - k + 2)^{\underline{k}}$ e $(\alpha k + m - l + 1)^{\underline{l}}$ are reduced to

$$(n - k + 2)^{\overline{k}} = (-1)^k \frac{(-n + k - 2)!}{(-n + 2k - 2)!} \ , \tag{3.24}$$

and

$$(\alpha k + m - l + 1)^{\overline{l}} = (-1)^l \frac{\Gamma(-\alpha k - m + l - 2)}{\Gamma(-\alpha k - m + 2l - 2)} \ . \tag{3.25}$$

And more, see that (i) $\dfrac{(n - k + 2)!}{(n - 2k + 2)!} \geq 1$ and (ii) $\dfrac{\Gamma(\alpha k + m - l + 2)}{\Gamma(\alpha k + m - 2l + 2)} \geq 1$ (according to the exposed on page 28, paragraph 7), and assuming (3.20) rewrite as

$$E(X^r|\alpha, \lambda, \beta) = \frac{\beta r!}{\lambda^r} \sum_{k,l,m,n=0}^{\infty} \frac{(-1)^{k+l}(1 - \beta)^n (n - k + 2)! \Gamma(\alpha k + m - l + 2)}{k!(l + 1)^r l! (n - 2k + 2)! \Gamma(\alpha k + m - 2l + 2)} \ , \tag{3.26}$$

taking a general term $A_{k,l,m,n}$ such that

$$A_{k,l,m,n} = \frac{(1 - \beta)^n (n - k + 2)! \Gamma(\alpha k + m - l + 2)}{k!(l + 1)^r l! (n - 2k + 2)! \Gamma(\alpha k + m - 2l + 2)} \ , \tag{3.27}$$

and with the aid of a term $B_{k,l,m,n} = (-1)^{k+l} A_{k,l,m,n}$, such that

$$\begin{cases} A_{k,l,m,n} = \dfrac{(1 - \beta)^n (n - k + 2)! \Gamma(\alpha k + m - l + 2)}{k!(l + 1)^r l! (n - 2k + 2)! \Gamma(\alpha k + m - 2l + 2)}, \text{ if (3.21), (3.22)} \\ B_{k,l,m,n} = \dfrac{(-1)^{k+l}(1 - \beta)^n (-n + k - 2)! \Gamma(-\alpha k - m + l - 2)}{k!(l + 1)^r l! (-n + 2k - 2)! \Gamma(-\alpha k - m + 2l - 2)}, \text{ if (3.24), (3.25)} \end{cases} ,$$

we can assume (3.20) rewrite as

$$\begin{aligned} E(X^r|\alpha, \lambda, \beta) &= \frac{\beta r!}{\lambda^r} \sum_{k,l,m,n=0}^{\infty} (-1)^{k+l} \frac{(1 - \beta)^n (n - k + 2)! \Gamma(\alpha k + m - l + 2)}{k!(l + 1)^r l! (n - 2k + 2)! \Gamma(\alpha k + m - 2l + 2)} = \\ &= \begin{cases} \dfrac{\beta r!}{\lambda^r} \displaystyle\sum_{k,l,m,n=0}^{\infty} (-1)^{k+l} A_{k,l,m,n}, \text{ if } n - k + 2, \alpha k + m - l + 1 > 0 \\ \dfrac{\beta r!}{\lambda^r} \displaystyle\sum_{k,l,m,n=0}^{\infty} (-1)^{k+l} B_{k,l,m,n}, \text{ if } n - k + 2, \alpha k + m - l + 1 < 0 \end{cases} , \end{aligned}$$

so it turns out that

$$E(X^r|\alpha,\lambda,\beta) = \begin{cases} \dfrac{\beta r!}{\lambda^r}\displaystyle\sum_{k,l,m,n=0}^{\infty}(-1)^{k+l}A_{k,l,m,n}, & \text{if } n-k+2, \alpha k+m-l+1>0 \\ \dfrac{\beta r!}{\lambda^r}\displaystyle\sum_{k,l,m,n=0}^{\infty}A_{k,l,m,n}, & \text{if } n-k+2, \alpha k+m-l+1<0 \end{cases} .$$

So, by i and ii, (3.27) it's such that

$$\begin{aligned} A_{k,l,m,n} &= \left[\frac{(1-\beta)^n}{k!(l+1)^r l!}\right]\left[\frac{(n-k+2)!}{(n-2k+2)!}\right]\left[\frac{\Gamma(\alpha k+m-l+2)}{\Gamma(\alpha k+m-2l+2)}\right] = \\ &= a_{(k,l,n)}a_{(n,k)}a_{(k,m,l)} \geq a_{(k,l,n)}(1)(1) = a_{(k,l)}a_n = A^*_{k,l,n} . \end{aligned} \tag{3.28}$$

this is, when $n-k+2>0$ and $\alpha k+m-l+1>0$, it turns out that $A_{k,l,m,n} \geq A^*_{k,l,n}$, but when $n-k+2<0$ and $\alpha k+m-l+1<0$, as $\left|(-n+k-2)^{\bar{l}}\right| \geq 1$ and $\left|(-\alpha k-m+l-2)^{\bar{l}}\right| \geq 1$, it turns out that $A_{k,l,m,n} \geq (-1)^{k+l}A^*_{k,l,n}$, where

$$A^*_{k,l,n} = a_{(k,l)}a_n = \frac{1}{k!(l+1)^r l!}(1-\beta)^n = \frac{(1-\beta)^n}{k!(l+1)^r l!} . \tag{3.29}$$

Then, in the case that $n-k+2>0$ and $\alpha k+m-l+1>0$

$$E(X^r|\alpha,\lambda,\beta) = \frac{\beta r!}{\lambda^r}\sum_{k,l,m,n=0}^{\infty}(-1)^{k+l}A_{k,l,m,n} \geq \frac{\beta r!}{\lambda^r}\sum_{k,l,m,n=0}^{\infty}(-1)^{k+l}A^*_{k,l,m,n} , \tag{3.30}$$

where

$$\sum_{k,l,m,n=0}^{\infty}(-1)^{k+l}A^*_{k,l,m,n} = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{(-1)^{k+l}}{k!(l+1)^r l!}\left[\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}(1-\beta)^n\right] , \tag{3.31}$$

and, in the case that $n-k+2<0$ and $\alpha k+m-l+1<0$, as $(-1)^{2(k+l)}=1$, occurs

$$E(X^r|\alpha,\lambda,\beta) = \frac{\beta r!}{\lambda^r}\sum_{k,l,m,n=0}^{\infty}A_{k,l,m,n} \geq \frac{\beta r!}{\lambda^r}\sum_{k,l,m,n=0}^{\infty}A^*_{k,l,m,n} = \frac{\beta r!}{\lambda^r}\sum_{k,l,m,n=0}^{\infty}\frac{(1-\beta)^n}{k!(l+1)^r l!} ,$$

where

$$\sum_{k,l,m,n=0}^{\infty}\frac{(1-\beta)^n}{k!(l+1)^r l!} = \sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{1}{k!(l+1)^r l!}\left[\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}(1-\beta)^n\right] . \tag{3.32}$$

In this condition, as $0<\beta<1$, in the double sum indexed in m and n, at follows that

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}(1-\beta)^n = \lim_{m\longrightarrow\infty}\sum_{i=0}^{m}\frac{1}{\beta} = \frac{1}{\beta}\left(\lim_{m\longrightarrow\infty}\sum_{i=0}^{m}(1)\right) = \infty , \tag{3.33}$$

this is, with the double sum indexed at m and n diverging, the double sum indexed at k and l also diverges, since it is the infinite sum of the infinite sum of infinite terms of the double sum indexed at m e n. Therefore, in (3.26) it is concluded

$$E(X^r|\alpha,\lambda,\beta) \longrightarrow \infty , \tag{3.34}$$

which proof the lemma. ■

Lemma 3 *For the random variable X with $GE2(\alpha, \lambda, \beta)$ distribution, when $\alpha, \lambda > 0$ and $\beta = 1$, we have that r-th moment function exists, is given for*

$$E(X^r|\alpha, \lambda, \beta = 1) = \frac{r!}{\lambda^r} \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(\alpha)^{\underline{n}}}{n! n^r} \ , \tag{3.35}$$

and such that

$$\frac{r!}{\lambda^r} \sum_{n=1}^{\infty} (-1)^{n-1} \frac{(\alpha)^{\underline{n}}}{n! n^r} \leq \frac{r!}{\lambda^r} e^r \ , \tag{3.36}$$

where e is the Euler's number.

Proof: In Equation (3.17), when $\beta = 1$, see that

$$E(X^r|\alpha, \lambda, \beta = 1) = \int_0^{\infty} \alpha \lambda x^r e^{-\lambda x} (1 - e^{-\lambda x})^{\alpha - 1} dx \ , \tag{3.37}$$

where

$$\begin{aligned} (1 - e^{-\lambda x})^{\alpha - 1} &= [1 + (-e^{-\lambda x})]^{\alpha - 1} = \sum_{k=0}^{\infty} \binom{\alpha - 1}{k} (-e^{-\lambda x})^k = \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{(\alpha - 1)^{\underline{k}}}{k!} e^{-k\lambda x} \ , \end{aligned} \tag{3.38}$$

a converging series, since if $x > 0$ then $0 < e^{-k\lambda x} < 1$, that replaced in (3.37) provides

$$E(X^r|\alpha, \lambda, \beta = 1) = \alpha \lambda \sum_{k=0}^{\infty} (-1)^k \frac{(\alpha - 1)^{\underline{k}}}{k!} \int_0^{\infty} x^{(r+1)-1} e^{-(k+1)\lambda x} dx \ . \tag{3.39}$$

Moreover, if $\alpha \in \mathbb{R}^*$ result, similarly to (3.22), that $(\alpha - 1)^{\underline{k}} = \dfrac{\Gamma(\alpha)}{\Gamma(\alpha - k)}$ and if $\lambda \in \mathbb{R}^*$, for all $k \in \mathbb{N}$, result that $(k + 1)\lambda \in \mathbb{R}^*$ by which considering the probability function of the Gamma distribution, we have $\displaystyle\int_0^{\infty} x^{(r+1)-1} e^{-(k+1)\lambda x} dx = \frac{\Gamma(r + 1)}{[(k + 1)\lambda]^{r+1}}$. Then, in (3.39) result that

$$\begin{aligned} E(X^r|\alpha, \lambda, \beta = 1) &= \alpha \lambda \sum_{k=0}^{\infty} (-1)^k \frac{\Gamma(\alpha)\Gamma(r + 1)}{k! \Gamma(\alpha - k)[(k + 1)\lambda]^{r+1}} = \\ &= \frac{r!}{\lambda^r} \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\alpha \Gamma(\alpha)}{n! n^r \Gamma(\alpha - n + 1)} = \frac{r!}{\lambda^r} \sum_{n=1}^{\infty} (-1)^{n-1} a_n \ . \end{aligned} \tag{3.40}$$

Now, see that $\alpha\Gamma(\alpha) = \Gamma(\alpha + 1)$, $\dfrac{1}{n! n^r} \leq \dfrac{r^n}{n!}$ and $\Gamma(\alpha - n + 1) = (\alpha)^{\underline{n+1}} \Gamma(\alpha + 1)$, where $(\alpha)^{\underline{n+1}} \geq 1$, taking $a_n = \dfrac{\Gamma(\alpha + 1)}{n! n^r \Gamma(\alpha - n + 1)}$, we have

$$a_n = \frac{\Gamma(\alpha + 1)}{n! n^r \Gamma(\alpha - n + 1)} = \frac{r^n \Gamma(\alpha + 1)}{n! (\alpha)^{\underline{n+1}} \Gamma(\alpha + 1)} = \frac{r^n}{n! (\alpha)^{\underline{n+1}}} \leq \frac{r^n}{n!} \ , \tag{3.41}$$

and if for all n we have $(-1)^{n-1} \dfrac{1}{n! n^r \Gamma(\alpha - n + 1)} \leq \dfrac{1}{n! n^r \Gamma(\alpha - n + 1)}$, applaying (3.41) in (3.40), result

$$E(X^r|\alpha, \lambda, \beta = 1) = \frac{r!}{\lambda^r} \sum_{n=1}^{\infty} (-1)^{n-1} \frac{\Gamma(\alpha + 1)}{n! n^r \Gamma(\alpha - n + 1)} \leq \frac{r!}{\lambda^r} \sum_{n=1}^{\infty} \frac{r^n}{n!} = \frac{r!}{\lambda^r} e^r \ ,$$

this is, as $\dfrac{\Gamma(\alpha+1)}{\Gamma(\alpha-n+1)} = (\alpha)^{\underline{n}}$, the proof is completed. ■

Moreover, let's consider the polylogarithm with order 1 for a variable y given by

$$Li_1(y) = -\ln(1-y) = \sum_{n=1}^{\infty}\frac{y^n}{n} < \sum_{n=1}^{\infty} y^n = \sum_{n=1}^{\infty}\frac{y^n}{n^0} = \frac{y}{1-y} = Li_0(y)\ , \tag{3.42}$$

when $0 < y < 1$ so we can ensure the following lemma.

Lemma 4 *For the random variable X with $GE2(\alpha,\lambda,\beta)$ distribution, when $\alpha,\lambda > 0$ and $\beta > 1$, we have that r-th moment function exists and is such that*

$$E(X^r|\alpha,\lambda,\beta) \leq \frac{r!}{\lambda^r}e^r\ . \tag{3.43}$$

Proof: From Equation (3.17), taking $\beta + (1-\beta)(1-e^{-\lambda x})^{\alpha} = u$ we have rewrite her

$$E(X^r|\alpha,\lambda,\beta) = \frac{\beta}{\lambda^r(1-\beta)}\int_{\beta}^{1}\frac{1}{u^2}\left\{-\ln\left[1-\left(\frac{u-\beta}{1-\beta}\right)^{1/\alpha}\right]\right\}^r du\ , \tag{3.44}$$

and, taking $y = \left(\dfrac{u-\beta}{1-\beta}\right)^{1/\alpha}$ results

$$E(X^r|\alpha,\lambda,\beta) = \frac{\alpha\beta}{\lambda^r}\int_0^1 \frac{y^{\alpha-1}[-\ln(1-y)]^r}{[\beta+(1-\beta)y^{\alpha}]^2}dy\ , \tag{3.45}$$

where $0 < y < 1$.

From equation (3.42) and by Maclaurin series, we have that

$$-\ln(1-y) < \frac{y}{1-y} < \frac{1}{1-y}\ . \tag{3.46}$$

Then, by inequality (3.46), assuming $r \geq 0$, the integrate in (3.45) is such that

$$\int_0^1 \frac{y^{\alpha-1}[-log(1-y)]^r}{[\beta+(1-\beta)y^{\alpha}]^2}dy < \int_0^1 \frac{y^{\alpha-1}(1-y)^{-r}}{[\beta+(1-\beta)y^{\alpha}]^2}dy\ , \tag{3.47}$$

where in the denominator of the integrand in the second integral, taking

$$[\beta+(1-\beta)y^{\alpha}]^2 = \beta^2[1-(1-\beta^{-1})y^{\alpha}]^2\ ,$$

as $\beta, y > 0$ such that $0 < (1-\beta^{-1})y^{\alpha} < 1$, results in the geométric series

$$\frac{1}{[1-(1-\beta^{-1})y^{\alpha}]^2} = \sum_{l=1}^{\infty} l[(1-\beta^{-1})y^{\alpha}]^{l-1}\ , \tag{3.48}$$

and in the binomial series

$$(1-y)^{-r} = [1+(-y)]^{(-r)} = \sum_{k=0}^{\infty}\binom{-r}{k}(-y)^k = \sum_{k=0}^{\infty}\frac{[(-1)r]^{\underline{k}}}{k!}[(-1)y]^k = \sum_{k=0}^{\infty}\frac{(r)^{\underline{k}}}{k!}y^k,$$

both converging and such that in the secund member of the inequality (3.47), we obtain

$$\begin{aligned}\int_0^1 \frac{y^{\alpha-1}(1-y)^{-r}}{[\beta+(1-\beta)y^\alpha]^2}dy &= \frac{1}{\beta^2}\int_0^1 y^{\alpha-1}\left[\sum_{k=0}^{\infty}\frac{(r)^{\underline{k}}}{k!}y^k\right]\left[\sum_{l=1}^{\infty} l[(1-\beta^{-1})y^\alpha]^{l-1}\right]dy = \\ &= \frac{1}{\beta^2}\left[\sum_{k=0}^{\infty}\sum_{l=1}^{\infty}\frac{(r)^{\underline{k}}}{k!}l(1-\beta^{-1})^{l-1}\right]\int_0^1 y^{\alpha l+k-1}dy = \\ &= \frac{1}{\beta^2}\sum_{k=0}^{\infty}\sum_{l=1}^{\infty}\frac{(r)^{\underline{k}}l(1-\beta^{-1})^{l-1}}{k!(\alpha l+k)} \ .\end{aligned}$$

See that

$$\frac{1}{\beta^2}\sum_{k=0}^{\infty}\sum_{l=1}^{\infty}\frac{(r)^{\underline{k}}l(1-\beta^{-1})^{l-1}}{k!(\alpha l+k)} < \frac{1}{\beta^2}\sum_{k=0}^{\infty}\frac{(r)^{\underline{k}}}{k!}\sum_{l=1}^{\infty}\frac{l(1-\beta^{-1})^{l-1}}{\alpha} \ , \tag{3.49}$$

and if for all $\beta \geq 0$ result $0 < 1-\beta^{-1} < 1$ and if $r \in \mathbb{N}$, according to (3.21), we have $(r)^{\underline{k}} = \dfrac{r!}{(r-k)!} \leq r!r^k$, replacing (3.49) in the inequation (3.47), it is concluded by equation (3.45) that

$$E(X^r|\alpha,\lambda,\beta) \leq \frac{\alpha\beta}{\lambda^r}\left[\frac{1}{\alpha\beta^2}\sum_{k=0}^{\infty}\frac{(r)^{\underline{k}}}{k!}\sum_{l=1}^{\infty}l(1-\beta^{-1})^{l-1}\right] = \frac{r!}{\beta\lambda^r}\sum_{k=0}^{\infty}\frac{r^k}{k!} = \frac{r!}{\lambda^r}e^r \ ,$$

which proof the lemma. ■

Note that, when $r = 0$, in the lemma (4) it follows that

$$\mu'_0 = E(X^0|\alpha,\lambda,\beta) = \frac{0!}{\lambda^0}e^0 = \int_0^\infty x^0 f(x|\alpha,\lambda,\beta)dx = \int_0^\infty f(x|\alpha,\lambda,\beta)dx = 1 \ ,$$

this is, the function of the r-th moment coincides with the probability density function when $r = 0$.

Similarly, when $r = 1$ and $r = 2$ we have

$$\mu'_1 \leq \frac{e}{\lambda}, \quad \mu'_2 \leq \frac{2e^2}{\lambda^2} \quad \text{and} \quad Var(X|\alpha,\lambda,\beta) \leq \frac{e^2}{\lambda^2} \ ,$$

which will be proved in the following proportion.

As the result of previous lemma (4) that guarantees that $E(X^r|\alpha,\lambda,\beta)$ is limited, the general expression of the r-th moment of the random variable $X \sim GE2(\alpha,\lambda,\beta)$ is given by

Proposition 1 *For the random variable X with $GE2(\alpha,\lambda,\beta)$ distribution, we have that r-th moment function is given by*

$$\mu'_r = \frac{r!}{\lambda^r}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}(-1)^m\frac{(1-\beta^{-1})^l[(\alpha l+k)^{\underline{m}} - (\alpha l+\alpha+k)^{\underline{m}}]}{(m+1)m!} \ . \tag{3.50}$$

Proof: From Equation (3.18), taking in the results (3.10) the variable change $y = 1-e^{-\lambda x}$, we have in $E(X^r|\alpha,\lambda,\beta) = \mu'_r$ that

$$r\int_0^\infty x^{r-1}\frac{[1-(1-e^{-\lambda x})^\alpha]}{1-(1-\beta^{-1})(1-e^{-\lambda x})^\alpha}dx = \frac{r}{\lambda^r}\int_0^1\frac{[-ln(1-y)]^{r-1}(1-y^\alpha)}{(1-y)[1-(1-\beta^{-1})y^\alpha]}dy \ , \tag{3.51}$$

is this, rewritten with

$$\mu'_r = \frac{r}{\lambda^r}\int_0^1 [-ln(1-y)]^{r-1}(1-y^\alpha)\left(\frac{1}{1-y}\right)\left[\frac{1}{1-(1-\beta^{-1})y^\alpha}\right]dy \ , \tag{3.52}$$

where $0 < \beta^{-1}, y < 1$ and considering the geometric series

$$\frac{1}{1-y} = \sum_{k=0}^{\infty} y^k \quad \text{and} \quad \frac{1}{1-(1-\beta^{-1})y^\alpha} = \sum_{l=0}^{\infty}[(1-\beta^{-1})y^\alpha]^l \ , \tag{3.53}$$

it turns out that

$$\mu'_r = \frac{r}{\lambda^r}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(1-\beta^{-1})^l\int_0^1 [-ln(1-y)]^{r-1}(y^{\alpha l+k} - y^{\alpha l+\alpha+k})dy \ , \tag{3.54}$$

and still, taking $z = -ln(1-y)$ we have in the μ'_r that

$$\mu'_r = \frac{r}{\lambda^r}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}(1-\beta^{-1})^l\int_0^{\infty} z^{r-1}e^{-z}[(1-e^{-z})^{\alpha l+k} - (1-e^{-z})^{\alpha l+\alpha+k}]dz \ . \tag{3.55}$$

See that $1 - e^{-z} = 1 + (-e^{-z})$ and so that

$$[1+(-e^{-z})]^w = \sum_{m=0}^{\infty}\binom{w}{m}(-e^{-z})^m = \sum_{m=0}^{\infty}(-1)^m\frac{(w)^{\underline{m}}}{m!}e^{-mz} \ . \tag{3.56}$$

Then, with proper adjustment in terms of subtraction, (3.55) can be rewritten as

$$\mu'_r = \frac{r}{\lambda^r}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}(-1)^m\frac{(1-\beta^{-1})^l[(\alpha l+k)^{\underline{m}} - (\alpha l+\alpha+k)^{\underline{m}}]}{m!}\int_0^{\infty} z^{r-1}e^{-(m+1)z}dz \ ,$$

Now, considering Gamma distribution, where $\int_0^{\infty} z^{r-1}e^{-(m+1)z}dz = \frac{\Gamma(r)}{(m+1)!}$, noting that $r\Gamma(r) = r!$ concludes

$$\mu'_r = \frac{r!}{\lambda^r}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\sum_{m=0}^{\infty}(-1)^m\frac{(1-\beta^{-1})^l[(\alpha l+k)^{\underline{m}} - (\alpha l+\alpha+k)^{\underline{m}}]}{(m+1)!m!} \ , \tag{3.57}$$

what completes the proof. ■

An infinite series by itself can bring several difficulties in obtaining its sum, since it is not always possible to define a value to which a series converges ... when it converges.

The previous proposition states that the exact calculation of the expected value of the GE2 distribution can only be obtained by the infinite sum of the elements contained in each $k \times l \times m$ matrix, or by the iterated series corresponding to the sum infinity of the rows, columns, and level of the expression $E(X^r|\alpha,\lambda,\beta)$ as

$$\lim_{k\longrightarrow\infty}\lim_{l\longrightarrow\infty}\lim_{m\longrightarrow\infty}(-1)^m\frac{(1-\beta^{-1})^l[(\alpha l+k)^{\underline{m}} - (\alpha l+\alpha+k)^{\underline{m}}]}{(m+1)!m!} \ , \tag{3.58}$$

which can be calculated by means of Pringsheim's theorem (see Bromwich [6]), that lays down criteria for exchanging infinity limit operators.

Theorem 4 *Let $h(t|\alpha,\lambda,\beta)$ as obtained in (3.11). When $t \longrightarrow \infty$, then $h(t|\alpha,\lambda,\beta) \longrightarrow \lambda$, this is, tends to lose memory over time.*

Proof: Applying in (3.11) the change of the variable $w = e^{-\lambda t}$, when $t \longrightarrow \infty$ is easy watched that $w \longrightarrow 0$ and

$$\lim_{t\longrightarrow\infty} h(t|\alpha,\lambda,\beta) = \lim_{w\longrightarrow 0} \frac{\alpha\lambda w(1-w)^{\alpha-1}}{\beta + [1-2\beta-(1-\beta)(1-w)^{\alpha}](1-w)^{\alpha}} \ , \tag{3.59}$$

an indeterminacy of type $\frac{0}{0}$. So, by L'Hospital's rule, it results out that

$$\lim_{t\longrightarrow\infty} h(t|\alpha,\lambda,\beta) = \lim_{w\longrightarrow 0} \frac{-\alpha\lambda(\alpha w-1)(1-w)^{\alpha-2}}{\alpha(1-w)^{\alpha-1}2\beta[(1-w)^{\alpha}-1]-2(1-w)^{\alpha}+1} = \lambda \ , \tag{3.60}$$

the which proves the theorem. ■

3.1.2 Maximum Likelihood Estimation

Let $t_1,\ldots,t_n$ be a random sample of GE2 distribution, that is, $T \sim GE2(\alpha,\lambda,\beta)$, then the likelihood function is given by,

$$L(\alpha,\lambda,\beta|\boldsymbol{t}) = \frac{(\alpha\lambda\beta)^n}{\exp\left(\lambda\sum_{i=1}^{n} t_i\right)} \prod_{i=1}^{n} \frac{(1-e^{-\lambda t_i})^{\alpha-1}}{[\beta+(1-\beta)(1-e^{-\lambda t_i})^{\alpha}]^2} \ . \tag{3.61}$$

The logarithm of likelihood function is given by

$$\begin{aligned} l(\alpha,\lambda,\beta|\boldsymbol{t}) = & \, n[log(\alpha)+log(\lambda)+log(\beta)] + \sum_{i=1}^{n}\{(\alpha-1)log(1-e^{-\lambda t_i}) - \\ & - 2log[\beta+(1-\beta)(1-e^{-\lambda t_i})^{\alpha}] - \lambda t_i\} \ , \end{aligned} \tag{3.62}$$

with system of likelihood equations given by

$$\begin{cases} \dfrac{n}{\hat{\alpha}} + \displaystyle\sum_{i=1}^{n} log(1-e^{-\hat{\lambda}t_i}) \left[\frac{\hat{\beta}-(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}}{\hat{\beta}+(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}}\right] = 0 \\ \dfrac{n}{\hat{\lambda}} - \displaystyle\sum_{i=1}^{n} \left[\frac{\hat{\lambda}(\hat{\alpha}-1)e^{-\hat{\lambda}t_i}}{1-e^{-\hat{\lambda}t_i}} - \frac{2\hat{\alpha}\hat{\lambda}(1-\hat{\beta})e^{-\hat{\lambda}t_i}(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}-1}}{\hat{\beta}+(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}} - t_i\right] = 0 \\ \dfrac{n}{\hat{\beta}} - 2\displaystyle\sum_{i=1}^{n} \frac{1-(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}}{\hat{\beta}+(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}} = 0 \end{cases} . \tag{3.63}$$

Now, by the expression (2.40), when $t_1,\cdots,t_n$ be a random sample of GE2 in the presence of the censure, your distribution then the likelihood function is given by,

$$L(\alpha,\lambda,\beta|\boldsymbol{t},\boldsymbol{\delta}) = \frac{(\alpha\lambda)^r\beta^n}{\exp\left(\lambda\sum_{i=1}^{n}\delta_i t_i\right)} \prod_{i=1}^{n} \frac{[1-(1-e^{-\lambda t_i})^{\alpha}]^{1-\delta_i}(1-e^{-\lambda t_i})^{\delta_i(\alpha-1)}}{[\beta+(1-\beta)(1-e^{-\lambda t_i})^{\alpha}]^{1+\delta_i}} \ , \tag{3.64}$$

and logarithm of likelihood function in the presence of the censure given by

$$\begin{aligned} l(\alpha,\lambda,\beta|\boldsymbol{t},\boldsymbol{\delta}) = & \, r\ln(\alpha\lambda) + n\ln(\beta) + \sum_{i=1}^{n}\big\{(1+\delta_i)\ln[\beta+(1-\beta)(1-e^{-\lambda t_i})^{\alpha}] \ + \\ & + \delta_i(\alpha-1)\ln(1-e^{-\lambda t_i}) - (1-\delta_i)\ln[1-(1-e^{-\lambda t_i})^{\alpha}] - \lambda\delta_i t_i\big\} \ , \end{aligned} \tag{3.65}$$

whose system of likelihood equations given by

$$\begin{cases} \dfrac{r}{\hat{\alpha}} + \displaystyle\sum_{i=1}^{n} \ln(1-e^{-\hat{\lambda}t_i}) \left[\delta_i - \dfrac{(1-\delta_i)(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}}{1-(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}} - \right. \\ \qquad \left. - \dfrac{(1+\delta_i)(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}}{\hat{\beta}+(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}} \right] = 0 \\ \dfrac{r}{\hat{\lambda}} + \displaystyle\sum_{i=1}^{n} t_i e^{-\hat{\lambda}t_i} \left[\dfrac{\delta_i(\hat{\alpha}-1)}{1-e^{-\hat{\lambda}t_i}} - \dfrac{\hat{\alpha}(1-\delta_i)(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}-1}}{1-(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}} - \right. \\ \qquad \left. - \dfrac{\hat{\alpha}(1+\delta_i)(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}-1}}{\hat{\beta}+(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}} \right] - \displaystyle\sum_{i=1}^{n} \delta_i t_i = 0 \\ \dfrac{n}{\hat{\beta}} - \displaystyle\sum_{i=1}^{n} \dfrac{(1-\delta_i)[1-(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}]}{\hat{\beta}+(1-\hat{\beta})(1-e^{-\hat{\lambda}t_i})^{\hat{\alpha}}} = 0 \end{cases}, \tag{3.66}$$

whose solutions provide the maximum likelihood estimators of the parameters α, λ e β.

3.1.3 Bayesian Analysis Approach

As developed in the previous chapter for the EEG model with censorship, with censorship for the model in the presence multiple variableand the cure rate model, here we also consider the approach bayesian for the estimators of this model and so we need to assume some prior distributions for the unknown parameters for development the inference.

Then, as assumed in the previous cases, a priori distributions for the parameters of the GE2 model correspond to the non-informative context, this is, it is considered an non-informative a priori distribution that satisfies the conditions α, λ, > 0 and it is sufficient consider the Gamma distribution for each of the parameters since it is defined continuous in $\mathbb{R}$ and tends to a Normal distribution as the size of its sample increases.

For this, we assume that the parameters are independent with the following prior distributions

$$\alpha \sim \Gamma(a_1, b_1), \quad \lambda \sim \Gamma(a_2, b_2) \quad \text{and} \quad \beta \sim \Gamma(a_3, b_3)\ . \tag{3.67}$$

Thus, under the bayesian approach, both the marginal distribution and the a priori distribution considered are not symmetric. The previous distributions expressing little or no information on α, λ and β can be obtained assuming independent Gamma distribution for each parameter, generate the probability distribution given by

$$\pi(\alpha, \lambda, \beta) \propto \alpha^{a_1-1}\lambda^{a_2-1}\beta^{a_3-1}e^{-b_1\alpha-b_2\lambda-b_3\beta}\ , \tag{3.68}$$

where a_1, a_2, a_3, b_1, b_2 and b_3 are known hyperparameters.

The joint posterior distribution for α, λ and β is proportional to the product of the likelihood function (**??**) and the prior distribution (3.68). Then, the expression (3.63) taking

$$\Delta(\alpha, \lambda, \beta|\boldsymbol{t}, \boldsymbol{\delta}) = \prod_{i=1}^{n} \frac{[1-(1-e^{-\lambda t_i})^{\alpha}]^{1-\delta_i}(1-e^{-\lambda t_i})^{\delta_i(\alpha-1)}}{[\beta+(1-\beta)(1-e^{-\lambda t_i})^{\alpha}]^{1+\delta_i}}\ , \tag{3.69}$$

it follows that

$$p(\alpha, \lambda, \beta|\boldsymbol{t}, \boldsymbol{\delta}) \propto \frac{\alpha^{r+a_1-1}\lambda^{r+a_2-1}\beta^{n+a_3-1}\Delta(\alpha, \lambda, \beta|\boldsymbol{t}, \boldsymbol{\delta})}{\exp\left[b_1\alpha + \lambda\left(b_2 + \displaystyle\sum_{i=1}^{n}\delta_i t_i\right) + b_3\beta\right]}\ . \tag{3.70}$$

The full conditional posterior distributions for α, λ and β are given as follows:

$$p(\alpha|\boldsymbol{t},\boldsymbol{\delta},\lambda,\beta) \propto \alpha^{r+a_1-1}e^{-b_1\alpha}\Delta(\alpha,\lambda,\beta|\boldsymbol{t},\boldsymbol{\delta}) \ , \tag{3.71}$$

$$p(\lambda|\boldsymbol{t},\boldsymbol{\delta},\alpha,\beta) \propto \frac{\lambda^{r+a_2-1}\Delta(\alpha,\lambda,\beta|\boldsymbol{t},\boldsymbol{\delta})}{\exp\left[\lambda\left(b_2+\sum_{i=1}^{r}\delta_i t_i\right)\right]} \ , \tag{3.72}$$

and

$$p(\lambda|\boldsymbol{t},\boldsymbol{\delta},\alpha,\beta) \propto \beta^{n+a_3-1}e^{-b_3\beta}\Delta(\alpha,\lambda,\beta|\boldsymbol{t},\boldsymbol{\delta}) \ . \tag{3.73}$$

These conditional distributions are needed in simulation of parameters of the joint posterior distribution based on MCMC methods.

Since the conditional distributions of α, λ and β are not easily identified, we use the Metropolis-Hastings algorithm (see for example, Gelfand and Smith [15] or Chib and Greenberg [8] to obtain the posterior summaries of interest.

3.2 A Simulation Study for the Model Randomly Censored

In this section, we develop a simulation study used MCMC method whose main objective is to study the efficiency of the MLE method for the distribution $X \sim GE2(x|\alpha,\lambda,\beta)$.

For this, the following procedure was computationally implemented

Step 1: Set the values N and n, respectively the number of samples in the simulation and the size of each their, and the values α, λ and β in a number of m cases of the parametric vector $\boldsymbol{\theta} = (\alpha,\lambda,\beta)$ of the model $GE2(\boldsymbol{\theta}) = GE2(x|\alpha,\lambda,\beta)$ censored with a fixed proportion of censorship in N samples, so that $n\tau_\delta$ is the exact number of censorships determined in each sample, where τ_δ is the proportion of censorship.

Step 2: Generate nN values $q \in]0,1[$ and n values from each of the N samples of the distribution $X \sim GE2(\boldsymbol{\theta})$ with $x = Q(q)$, according (3.13), and such that $F(x|\boldsymbol{\theta}) \in]0,1[$.

Step 3: Use the values obtained in step 2 for the $X \sim GE2(\boldsymbol{\theta})$ distribution to calculate in each of the N samples the estimated vector $\hat{\boldsymbol{\theta}} = (\hat{\alpha},\hat{\lambda},\hat{\beta})$, this is, for $i = 1,\ldots,N$, get $\hat{\boldsymbol{\theta}}_i$ through MLE of the α, λ and β parameters by the MCMC method.

Step 4: Use the N vectors $\hat{\boldsymbol{\theta}} = (\hat{\alpha},\hat{\lambda},\hat{\beta})$ and the vector $\boldsymbol{\theta} = (\alpha,\lambda,\beta)$ for compute for j-th parameter, with $j = 1,2,3$ in the k-th case, with $1 \leq k \leq m$, the mean bias absolute (MBA), square root from the mean square errors (RMSE) and 95% coverage probability, respectivaly

$$V_{\theta_{jk}} = \frac{1}{N}\sum_{i=1}^{N}|\theta_{jk} - \hat{\theta}_{ijk}| \ , \tag{3.74}$$

$$e_{\theta_{jk}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\theta_{jk} - \hat{\theta}_{ijk})^2} \ , \tag{3.75}$$

$$\hat{p}_{\theta_{jk}} = \frac{1}{N}\sum_{i=1}^{N} W_{ijk} \ , \tag{3.76}$$

so that, in the i-th sample where

$$W_{ijk} = \begin{cases} 1, \text{ if } \theta_{jk} \in IC_{\{\theta_{jk},0.95\}} \\ 0, \text{ otherwise} \end{cases} ,$$

and $IC_{\{\theta_{jl},0.95\}}$ is the interval of 95% for the parameter θ_{jk}, with $\boldsymbol{\theta}_k = (\theta_{1k}, \theta_{2k}) = (\lambda_k, \gamma_k)$ and $\hat{\boldsymbol{\theta}}_{ik} = (\hat{\theta}_{i1k}, \hat{\theta}_{i2k}) = (\hat{\lambda}_{ik}, \hat{\gamma}_{ik})$, in the i-th sample of the k-th case, respectivaly. Moreover, for the N confidence and credibility intervals obtained, we will also consider the mean interval amplitude (MIA) in each case as

$$\overline{h}_{IC_{\{\theta;1-\epsilon\}}} = \frac{1}{N}\sum_{i=1}^{N} h_{IC_{\{\theta_i;1-\epsilon\}}} , \tag{3.77}$$

were $h_{IC_{\{\theta_i;0.95\}}} = 2z_{\frac{0.05}{2}}\hat{\sigma}_{\hat{\theta}_i}$ in the classic case and $h_{IC_{\{\theta_i;0.95\}}} = \hat{\theta}_i^{(k+[0.95]\eta)} - \hat{\theta}_i^{(k)}$ in the bayesian case, were $\theta_i^{(k)}$ is the k-th smallest lower limit and $\theta_i^{(k+[0.95]\eta)}$ is the $[k+(0.95)\eta]$-th smallest upper limit of the ordered set of quantis $\boldsymbol{\theta}_j^* = \{\theta_j^{(1)}; \theta_j^{(2)}; \theta_j^{(3)}; ...; \theta_j^{(\eta)}\}$ from the j-th posterior sample for size η.

Repeat steps 2, 3, 4 and 5 for the m cases of $\boldsymbol{\theta}$.

We chose to perform this simulation procedure for $m = 4$ parametric cases given by $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3, \boldsymbol{\theta}_4\} = \{(0.75, 0.5, 3.5), (1.0, 2.0, 0.8), (3.0, 1.0, 0.1), (5.0, 1.5, 2.0)\}$, parameter vectors for the model in the cases which the curve of risk manifests the forms, respectively, increasing, unimodal, decreasing and bathtub, as shown in the following image.

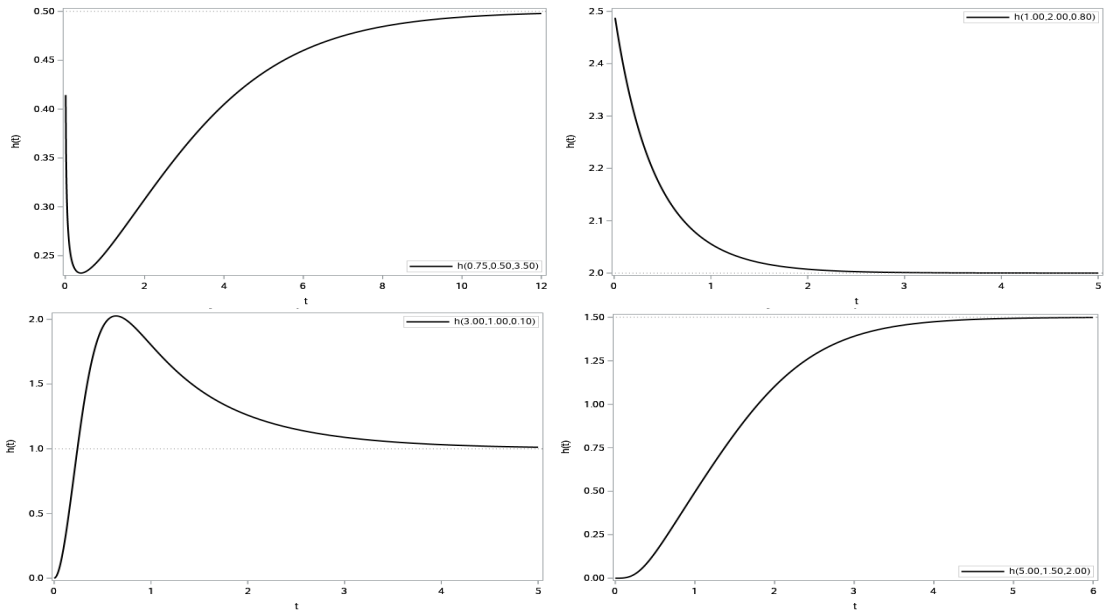

Figure 3. 2 : The hazard works for simulations of GE2 distribution censored.

For the random data generated with the exact proportion of censorships of $\tau_\delta = 0.0, 0.2, 0.4$ in 4 case parametric with sizes n = 10, 25, 50, 100, 200, by the processes approach classic and bayesian, were necessary a total of 120 processes.

For this approach, MBAs, RMSEs and MIAs are expected to be close to zero as the samples of size n increases and that probability of empirical coverage p approaches $P = 0.95$, the probability of the theoretical credibility range which, for this simulation

process is defined as the most rigorous, the Highest Posterior Density (HPD) interval.

The results are presented in the following tables, were the statistics considered for evaluation criteria are presented in the Classical estimation column for the classical approach via MLE and in bayesian estimation for the bayesian approach via MCMC.

Table 3. 1 : Results for the model $GE2(\alpha, \lambda, \beta) = GE2(0.75, 0.5, 3.5)$.

Cases			Classical estimation					Bayesian estimation				
θ	τ_δ	n	$\hat{\theta}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$	$\hat{\mu}_{\hat{\theta}}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$
α	0	10	1.697	0.996	3.396	1.145	6.604	0.775	0.993	0.184	0.153	1.387
		25	1.005	0.957	0.666	0.466	2.365	0.805	0.994	0.207	0.163	1.183
		50	0.879	0.949	0.430	0.313	1.550	0.804	0.991	0.207	0.155	0.995
		100	0.802	0.946	0.276	0.214	1.054	0.778	0.982	0.173	0.135	0.795
		200	0.786	0.939	0.449	0.155	0.740	0.771	0.965	0.479	0.115	0.627
	0.2	10	2.015	0.989	3.686	1.464	8.932	0.787	0.988	0.184	0.150	1.419
		25	1.146	0.936	0.890	0.594	2.783	0.856	0.992	0.247	0.192	1.272
		50	0.943	0.936	0.515	0.364	1.744	0.855	0.995	0.237	0.175	1.070
		100	0.865	0.941	0.341	0.254	1.192	0.841	0.986	0.212	0.157	0.870
		200	0.813	0.943	0.109	0.174	0.819	0.889	0.970	0.167	0.126	0.673
	0.4	10	2.770	0.989	8.095	2.196	15.490	0.800	0.990	0.179	0.148	1.460
		25	1.221	0.939	1.032	0.671	3.199	0.890	0.993	0.261	0.203	1.340
		50	1.015	0.926	0.619	0.435	1.988	0.913	0.997	0.281	0.214	1.150
		100	0.908	0.931	0.404	0.293	1.350	0.897	0.984	0.255	0.189	0.941
		200	0.858	0.922	0.281	0.211	0.940	0.869	0.961	0.212	0.158	0.741
λ	0	10	0.588	0.987	0.300	0.214	1.346	0.473	0.965	0.130	0.104	1.387
		25	0.516	0.968	0.164	0.127	0.688	0.473	0.961	0.100	0.083	1.183
		50	0.509	0.964	0.114	0.088	0.459	0.484	0.972	0.077	0.065	0.996
		100	0.506	0.961	0.077	0.063	0.314	0.491	0.968	0.063	0.051	0.795
		200	0.501	0.947	0.055	0.043	0.218	0.495	0.955	0.045	0.038	0.627
	0.2	10	0.560	0.987	0.339	0.234	1.479	0.436	0.933	0.148	0.119	1.419
		25	0.510	0.967	0.184	0.143	0.770	0.452	0.938	0.114	0.093	1.272
		50	0.500	0.968	0.122	0.099	0.512	0.465	0.957	0.089	0.073	1.070
		100	0.494	0.952	0.083	0.069	0.347	0.473	0.962	0.071	0.057	0.870
		200	0.491	0.939	0.063	0.051	0.240	0.480	0.947	0.055	0.045	0.673
	0.4	10	0.579	0.980	0.371	0.273	1.749	0.396	0.898	0.179	0.148	1.460
		25	0.499	0.957	0.219	0.173	0.881	0.415	0.899	0.261	0.203	1.340
		50	0.479	0.955	0.145	0.115	0.581	0.429	0.913	0.109	0.092	1.155
		100	0.477	0.934	0.100	0.082	0.392	0.444	0.910	0.089	0.073	0.941
		200	0.476	0.923	0.077	0.062	0.271	0.456	0.902	0.071	0.059	0.741
β	0	10	11.036	0.789	41.581	10.260	228.300	3.418	1.000	0.690	0.558	1.392
		25	5.758	0.824	9.610	4.395	39.010	3.577	0.999	0.935	0.745	1.186
		50	4.569	0.861	4.874	2.735	18.130	3.718	0.994	1.107	0.868	1.000
		100	4.136	0.901	2.949	1.880	10.750	3.855	0.983	1.233	0.971	0.800
		200	3.781	0.910	1.769	1.275	6.680	3.827	0.970	1.197	0.927	0.629
	0.2	10	17.480	0.787	106.789	16.650	279.100	3.445	1.000	0.629	0.502	1.419
		25	7.435	0.834	28.888	6.039	78.110	3.683	0.999	0.920	0.730	1.277
		50	5.399	0.871	7.028	3.493	24.390	3.855	0.998	1.156	0.908	1.076
		100	4.655	0.918	3.839	2.341	13.430	4.023	0.988	1.350	1.041	0.875
		200	4.319	0.929	2.311	1.640	8.417	4.144	0.908	1.414	1.106	0.676
	0.4	10	20.650	0.773	132.864	19.910	327.800	3.437	1.000	0.538	0.432	1.465
		25	10.285	0.848	23.831	8.866	104.300	3.761	1.000	0.894	0.701	1.340
		50	6.955	0.891	11.576	5.006	37.820	3.976	0.997	1.204	0.929	1.161
		100	5.656	0.923	5.415	3.205	18.640	4.256	0.993	1.491	1.163	0.947
		200	5.224	0.945	3.693	2.422	11.670	4.518	0.983	1.762	1.373	0.743

The above table shows, above all, that between the n sizes considered, the highest measurements point to high RMSE, MBA and MIA values for 10 samples in the classic case and that such measures are compensated for the sample size increases given the

asymptotically expected convergences in the case of estimation via MLE. In the bayesian case, besides the measurements of erros are smaller, this is common in any sample.

In the bayesian case, although the same statistics presented low values already in samples of size 10, it is verified that the estimator of the parameter β, in this case, overestimates the estimates in samples with high proportions of censorship. In addition, it is found that the coverage probability for the 3 estimators converge more slowly to the theoretical value compared to the MLEs.

An analogous behavior is observed in the parametric case shown in table below.

Table 3. 2 : Results for the model $GE2(\alpha, \lambda, \beta) = GE2(1.0, 2.0, 0.8)$.

Cases			Classical estimation					Bayesian estimation				
θ	τ_δ	n	$\hat{\theta}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$	$\hat{\mu}_{\hat{\theta}}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$
α	0	10	1.312	0.991	0.980	0.604	3.792	1.212	0.988	0.472	0.338	1.846
		25	1.090	0.949	0.453	0.343	1.765	1.104	0.978	0.318	0.236	1.253
		50	1.051	0.955	0.303	0.235	1.185	1.060	0.976	0.232	0.175	0.953
		100	1.027	0.945	0.214	0.170	0.812	1.030	0.966	0.176	0.138	0.722
		200	1.013	0.958	0.141	0.115	0.572	1.010	0.967	0.126	0.102	0.534
	0.2	10	1.494	0.986	2.403	0.790	5.139	1.273	0.987	0.532	0.383	1.995
		25	1.112	0.966	0.492	0.362	1.953	1.147	0.988	0.363	0.250	1.338
		50	1.056	0.949	0.329	0.255	1.293	1.088	0.976	0.251	0.190	1.006
		100	1.035	0.945	0.228	0.179	0.898	1.054	0.966	0.187	0.142	0.766
		200	1.019	0.951	0.158	0.127	0.625	1.030	0.963	0.138	0.110	0.571
	0.4	10	1.915	0.977	5.748	1.197	7.807	1.342	0.991	0.598	0.442	2.184
		25	1.176	0.971	0.576	0.414	2.207	1.210	0.989	0.410	0.294	1.466
		50	1.099	0.935	0.378	0.291	1.450	1.145	0.977	0.300	0.222	1.093
		100	1.057	0.947	0.257	0.202	1.001	1.095	0.977	0.214	0.163	0.822
		200	1.034	0.947	0.179	0.141	0.706	1.057	0.961	0.158	0.122	0.622
λ	0	10	2.901	0.983	2.803	1.447	8.104	2.141	0.979	0.735	0.561	1.846
		25	2.267	0.963	1.057	0.783	4.140	2.023	0.975	0.553	0.432	1.255
		50	2.128	0.957	0.713	0.546	2.809	2.008	0.976	0.474	0.373	0.955
		100	2.071	0.959	0.489	0.382	1.931	2.010	0.973	0.383	0.303	0.723
		200	2.017	0.959	0.339	0.268	1.334	1.990	0.970	0.298	0.238	0.536
	0.2	10	3.036	0.977	2.517	1.637	9.262	2.034	0.969	0.680	0.540	1.995
		25	2.228	0.965	1.179	0.866	4.589	1.903	0.967	0.563	0.447	1.338
		50	2.056	0.957	0.776	0.608	3.084	1.878	0.959	0.487	0.394	1.007
		100	1.982	0.952	0.522	0.409	2.102	1.883	0.963	0.390	0.320	0.767
		200	1.936	0.949	0.363	0.294	1.458	1.886	0.959	0.321	0.262	0.572
	0.4	10	2.989	0.948	2.928	1.802	10.430	1.839	0.957	0.659	0.539	2.186
		25	2.049	0.939	1.241	0.951	5.134	1.708	0.919	0.590	0.491	1.466
		50	1.909	0.955	0.851	0.669	3.446	1.690	0.925	0.532	0.448	1.093
		100	1.851	0.940	0.611	0.491	2.355	1.702	0.910	0.482	0.403	0.824
		200	1.816	0.931	0.434	0.349	1.624	1.724	0.905	0.410	0.343	0.623
β	0	10	3.417	0.903	10.089	3.073	43.500	0.917	1.000	0.263	0.202	1.847
		25	1.529	0.883	2.488	1.119	7.893	0.944	0.995	0.370	0.285	1.255
		50	1.070	0.901	0.973	0.607	3.569	0.933	0.981	0.395	0.302	0.955
		100	0.937	0.927	0.600	0.414	2.172	0.920	0.975	0.383	0.293	0.723
		200	0.855	0.927	0.389	0.270	1.392	0.877	0.957	0.311	0.235	0.536
	0.2	10	4.865	0.906	27.788	4.461	64.830	0.945	1.000	0.249	0.199	1.997
		25	1.882	0.904	3.175	1.447	10.840	0.980	1.000	0.365	0.282	1.340
		50	1.350	0.922	1.568	0.860	5.048	1.003	0.994	0.439	0.335	1.007
		100	1.114	0.942	0.829	0.536	2.873	1.012	0.983	0.447	0.338	0.767
		200	1.005	0.957	0.527	0.369	1.822	0.991	0.977	0.407	0.304	0.572
	0.4	10	6.840	0.887	32.496	6.480	156.000	0.963	1.000	0.243	0.195	2.186
		25	2.314	0.899	4.438	1.910	15.400	0.998	0.999	0.348	0.271	1.466
		50	1.660	0.934	2.931	1.170	7.330	1.041	0.996	0.440	0.335	1.092
		100	1.346	0.962	1.204	0.763	4.000	1.086	0.995	0.510	0.383	0.824
		200	1.220	0.977	0.762	0.540	2.550	1.124	0.989	0.513	0.396	0.623

Tables 3. 1 and 3. 2 show that the advantage of bayesian estimators over MLE consists mainly in applications with small size sample and whe more severe interval estimates, since the values for The MIA in this approach is significantly lower when the true value of the parameter is greater than 1. In general, bayesian estimates are more accurate as observed in the calculated RMSEs.

The tables 3. 3 and 3. 4 in the sequence reflect the behavior of the EEG model parameter estimators described so far and show that, both via MLE and MCMC, the provided estimates show have the same properties.

Table 3. 3 : Results for the model $GE2(\alpha, \lambda, \beta) = GE2(3.0, 1.0, 0.1)$.

Cases			Classical estimation					Bayesian estimation				
θ	τ_δ	n	$\hat{\theta}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$	$\hat{\mu}_{\hat{\theta}}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$
		10	4.892	0.985	11.362	2.614	17.790	3.492	0.979	1.219	0.889	4.673
		25	3.258	0.983	1.155	0.795	4.559	3.210	0.973	0.760	0.564	2.950
	0	50	3.089	0.971	0.640	0.484	2.613	3.092	0.967	0.465	0.402	2.071
		100	3.022	0.959	0.419	0.325	1.677	3.024	0.956	0.367	0.287	1.456
		200	3.001	0.931	0.291	0.226	1.109	2.988	0.950	0.263	0.205	1.031
		10	5.766	0.985	14.446	3.500	26.240	3.517	0.978	1.260	0.909	4.978
		25	5.354	0.655	2.802	2.381	5.943	3.190	0.967	0.816	0.607	3.142
α	0.2	50	3.083	0.971	0.701	0.520	2.940	3.079	0.964	0.556	0.430	2.222
		100	2.982	0.950	0.442	0.344	1.801	2.982	0.949	0.387	0.305	1.547
		200	2.966	0.923	0.305	0.242	1.201	2.955	0.939	0.276	0.221	1.106
		10	9.664	0.982	42.415	7.439	63.470	3.601	0.989	1.368	1.001	5.435
		25	3.351	0.962	1.998	1.063	6.261	3.199	0.963	0.902	0.655	3.437
	0.4	50	3.095	0.952	0.880	0.619	3.370	3.072	0.954	0.627	0.478	2.444
		100	2.934	0.938	0.497	0.388	2.013	2.929	0.933	0.425	0.338	1.679
		200	2.892	0.918	0.339	0.277	1.323	2.878	0.920	0.315	0.259	1.189
		10	2.542	0.940	2.437	1.664	6.959	1.061	0.987	0.316	0.254	4.675
		25	1.743	0.875	1.331	0.952	3.867	1.024	0.991	0.245	0.195	2.951
	0	50	1.387	0.880	0.872	0.645	2.666	0.994	0.991	0.205	0.164	2.072
		100	1.186	0.881	0.615	0.460	1.977	0.986	0.991	0.197	0.157	1.467
		200	1.076	0.913	0.418	0.320	1.451	0.974	0.981	0.197	0.158	1.032
		10	2.545	0.967	2.475	1.683	7.763	1.009	0.985	0.305	0.247	4.978
		25	1.469	0.627	0.562	0.495	1.163	0.946	0.979	0.255	0.209	3.143
λ	0.2	50	1.318	0.912	0.821	0.604	2.999	0.924	0.987	0.207	0.167	2.223
		100	1.097	0.905	0.577	0.439	2.108	0.898	0.981	0.202	0.166	1.547
		200	1.009	0.914	0.419	0.325	1.555	0.897	0.969	0.202	0.168	1.106
		10	2.814	0.976	2.985	1.982	9.307	0.957	0.977	0.310	0.255	5.440
		25	1.706	0.916	1.156	0.987	4.670	0.872	0.963	0.277	0.229	3.438
	0.4	50	1.322	0.904	0.959	0.682	3.289	0.848	0.959	0.255	0.214	2.446
		100	1.053	0.917	0.589	0.451	2.317	0.812	0.959	0.247	0.211	1.680
		200	0.932	0.930	0.413	0.328	1.722	0.802	0.941	0.245	0.215	1.190
		10	2.343	0.984	8.331	2.266	31.080	0.110	1.000	0.071	0.017	4.673
		25	0.748	0.939	1.786	0.681	4.539	0.114	1.000	0.032	0.022	2.951
	0	50	0.362	0.911	0.725	0.301	1.625	0.115	0.998	0.032	0.028	2.071
		100	0.217	0.908	0.297	0.158	0.776	0.116	0.990	0.045	0.034	1.456
		200	0.144	0.917	0.122	0.082	0.416	0.112	0.979	0.045	0.035	1.031
		10	2.455	0.986	8.133	2.375	33.670	0.111	1.000	0.071	0.017	4.980
		25	2.765	0.992	3.905	2.667	9.429	0.113	0.999	0.032	0.020	3.142
β	0.2	50	0.405	0.917	1.032	0.343	2.017	0.116	0.999	0.032	0.025	2.220
		100	0.223	0.911	0.338	0.167	0.884	0.114	1.000	0.032	0.029	1.550
		200	0.157	0.918	0.155	0.098	0.505	0.114	0.988	0.045	0.033	1.106
		10	7.119	0.988	52.010	7.049	110.700	0.112	1.000	0.032	0.018	5.435
		25	1.547	0.956	5.952	1.478	13.490	0.116	1.000	0.032	0.021	3.437
	0.4	50	0.574	0.936	1.480	0.514	3.170	0.117	0.999	0.032	0.023	2.444
		100	0.273	0.929	0.424	0.215	1.218	0.117	0.999	0.032	0.027	1.680
		200	0.175	0.939	0.187	0.114	0.653	0.116	0.999	0.032	0.030	1.189

Table 3. 4 : Results for the model $GE2(\alpha, \lambda, \beta) = GE2(5.0, 1.5, 2.0)$.

Cases			Classical estimation					Bayesian estimation				
θ	τ_δ	n	$\hat{\theta}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$	$\hat{\mu}_{\hat{\theta}}$	p_θ	e_θ	V_θ	$\overline{h}_{IC}$
α	0	10	13.085	0.983	12.845	9.436	74.440	5.120	0.967	1.520	1.234	8.933
		25	6.128	0.985	3.846	2.446	14.500	5.331	0.974	1.563	1.206	7.164
		50	5.440	0.973	2.200	1.570	8.271	5.316	0.963	1.399	1.050	5.625
		100	5.151	0.966	1.358	1.037	5.394	5.185	0.974	1.035	0.790	4.250
		200	5.047	0.957	0.910	0.716	3.701	5.093	0.967	0.761	0.596	3.197
	0.2	10	23.442	0.984	122.783	19.700	145.200	5.120	0.977	1.480	1.208	9.189
		25	6.787	0.989	6.393	3.045	18.090	5.404	0.980	1.630	1.242	7.547
		50	5.565	0.980	2.438	1.643	9.318	5.366	0.977	1.430	1.048	5.921
		100	5.354	0.967	1.561	1.175	5.943	5.361	0.975	1.162	0.876	4.568
		200	5.167	0.959	1.030	0.798	4.013	5.215	0.968	0.855	0.656	3.392
	0.4	10	51.166	0.966	364.692	47.420	220.000	5.001	0.972	1.382	1.140	9.312
		25	7.303	0.974	7.764	3.597	22.540	5.355	0.971	1.618	1.271	7.836
		50	5.874	0.973	3.172	2.005	11.270	5.469	0.977	1.561	1.174	6.382
		100	5.393	0.964	1.737	1.292	6.621	5.384	0.969	1.257	0.943	4.851
		200	5.282	0.955	1.182	0.908	4.418	5.325	0.965	0.960	0.731	3.638
λ	0	10	1.761	0.909	0.861	0.625	3.770	1.197	0.930	0.373	0.324	8.934
		25	1.568	0.963	0.518	0.405	2.144	1.272	0.951	0.298	0.254	7.165
		50	1.508	0.950	0.387	0.296	1.492	1.324	0.963	0.255	0.209	5.626
		100	1.507	0.953	0.263	0.204	1.041	1.382	0.963	0.205	0.166	4.251
		200	1.507	0.957	0.182	0.144	0.721	1.433	0.962	0.155	0.124	3.197
	0.2	10	1.803	0.981	0.990	0.688	4.279	1.137	0.908	0.422	0.376	9.190
		25	1.559	0.965	0.555	0.430	2.371	1.204	0.937	0.345	0.304	7.548
		50	1.481	0.952	0.432	0.336	1.649	1.255	0.931	0.307	0.264	5.922
		100	1.469	0.955	0.311	0.232	1.163	1.321	0.943	0.245	0.205	4.568
		200	1.472	0.957	0.207	0.163	0.811	1.377	0.940	0.195	0.158	3.393
	0.4	10	1.879	0.971	1.206	0.793	4.859	1.064	0.871	0.484	0.443	9.313
		25	1.498	0.959	0.626	0.485	2.711	1.112	0.870	0.429	0.392	7.836
		50	1.468	0.940	0.485	0.374	1.901	1.177	0.889	0.367	0.328	6.383
		100	1.418	0.947	0.369	0.279	1.329	1.228	0.887	0.320	0.281	4.851
		200	1.416	0.955	0.262	0.199	0.940	1.290	0.897	0.263	0.222	3.639
β	0	10	13.762	0.832	80.592	13.180	271.500	1.323	0.999	0.711	0.677	8.935
		25	4.513	0.853	11.524	3.643	36.320	1.463	0.978	0.672	0.575	7.165
		50	3.058	0.855	4.219	2.098	13.890	1.597	0.960	0.678	0.563	5.626
		100	2.561	0.904	2.257	1.354	7.827	1.783	0.951	0.662	0.546	4.253
		200	2.302	0.921	1.296	0.917	4.824	1.940	0.949	0.656	0.532	3.199
	0.2	10	12.993	0.815	70.725	12.400	235.500	1.298	0.999	0.727	0.702	9.191
		25	5.288	0.858	11.546	4.446	45.030	1.422	0.991	0.676	0.596	7.5504
		50	3.637	0.864	6.498	2.681	19.110	1.555	0.971	0.672	0.564	5.924
		100	2.765	0.895	2.955	1.630	9.429	1.720	0.957	0.677	0.562	4.570
		200	2.442	0.924	1.601	1.070	5.698	1.917	0.950	0.683	0.550	3.394
	0.4	10	21.772	0.819	134.205	21.160	294.400	1.281	1.000	0.734	0.718	9.314
		25	6.733	0.839	20.635	6.033	74.910	1.353	0.993	0.715	0.652	7.837
		50	4.838	0.878	12.708	3.843	31.790	1.508	0.980	0.447	0.572	6.384
		100	3.223	0.899	4.056	2.086	12.570	1.663	0.965	0.669	0.562	4.853
		200	2.631	0.919	2.012	1.292	7.059	1.855	0.949	0.685	0.553	3.641

As the significance level for the confidence intervals and credibility in each sample was 5%, the probability of theoretical coverage for each of the parametric case combinations, sample size, and censorship ratio was set at 0.95, for each of the parameters, ie $P_\theta = 0.95$ is the nominal or expected coverage probability in all cases.

However, the determining factor to be taken into consideration, which is the evaluation p_θ proportion of empirical coverage obtained, in the sense of what is the acceptable level of variation around 0.95 as well as your precision. Therefore, in this simulation study all p_θ statistics are considered acceptable as long as it is observed that, as the values for n

increases, p_θ approaches 0.95.

Note that, in the descriptions of simulationof this study it is not considered in which approach, classical or bayesian, p_θ is closer to 0.95, but whether this statistic converges to its nominal value. The question of proximity for the expected value for the displayed statistics is considered by the error measures, as MBA, RMSE and MIA presented, respectively, by 3.75, 3.74, and 3.77.

For each of the 1500 GE2 samples, the MCMC process was implemented with 50000 iterations for each of the later 1500 samples, with 10000 burn-in and roughing every 4 iterations to average each of the 1500 later samples of the size 10000 and the desired estimate by averaging these averages. To generate the random censored data, we utilize the same methods used by Goodman, Li and Tiwari [17].

The seed used to generate the simulation random values was the $64-bit$ Windows 10 operating system time, with Intel® Core^{TM} $i5-4200U$ CPU @ $1.60GHz$ processor and $8.00GB$ installed RAM. The software used was SAS On Deamand where the implemented code mainly considered the DATA STEP process and the IML and MCMC procedures, all with seed inserted by the STREAMINIT(0) statement.

3.3 Application for Measures of Recurrence Censored Times

In this application, with a sample 38 observations, the main objective is to fit a probabilistic model for the recurrence times to infection at point of insertion of the catheter for kidney patients users portable dialysis equipment. The data was presented by Mcgilchrist and Aisbett [40], record the lifetime of 38 patients and for each patient two times of infection recurrence are observed: the first and second time of occurrence of the event of interest.

Since the times recorded refer to the time of return of the infection, one of the objectives of this modeling is to predict the probability of recurrence (return) of the infection due to the insertion of the catheter. In this case, in the probabilistic context, predicting the probability of recurrence of the infection is the same as predicting the probability for the event that is occurrence of infection at the catheter insertion point.

One of the ways to obtain this measure is through a survival analysis, which can empirically be taken as an estimate given by the survival curve obtained by the Kaplan-Meyer non-parametric estimator for both cases of recurrence. The graphs of the estimated survival curves the Kaplan-Meyer are shown in the graph below.

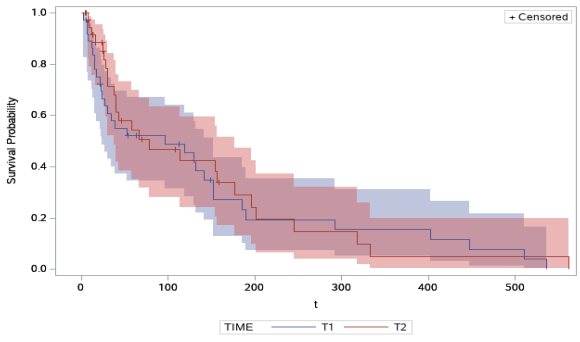

Figure 3. 3 : Product-limit survival estimate plot for two cases for recorrences.

The table 3. 5 in the sequence consists of the approximate chi-square statistics (χ^2), degrees of freedom (DF), and p-values ($Pr > \chi^2$) for the Log-Rank, Wilcoxon, and

likelihood ratio tests. All this three tests show strong evidence of that not exist difference among the survival curves for the two cases of recurrence (p-values > 0.3).

Table 3. 5 : Test of equality.

Test	χ^2	DF	p-values
Log-Rank	0.2678	1	0.6048
Wilcoxon	0.9882	1	0.3202
-2Log(LR)	0.0058	1	0.9394

However, since the recurrence time is an observable continuous random variable and dependent on a population with characteristic structures and conditions, nonparametric approaches using estimates such as Kaplan-Meyer, limit not only the quality of the intended forecast, but also the inference that can be made about the study population.

In these conditions, for the intended modeling, an adjustment is proposed for the data with a probabilistic model that captures alternative forms of risk, as shown in the figure below, the data indicate a predominantly decreasing form of risk, however, leaves evidence that it can also take the form bathtub and unimodal, respectively.

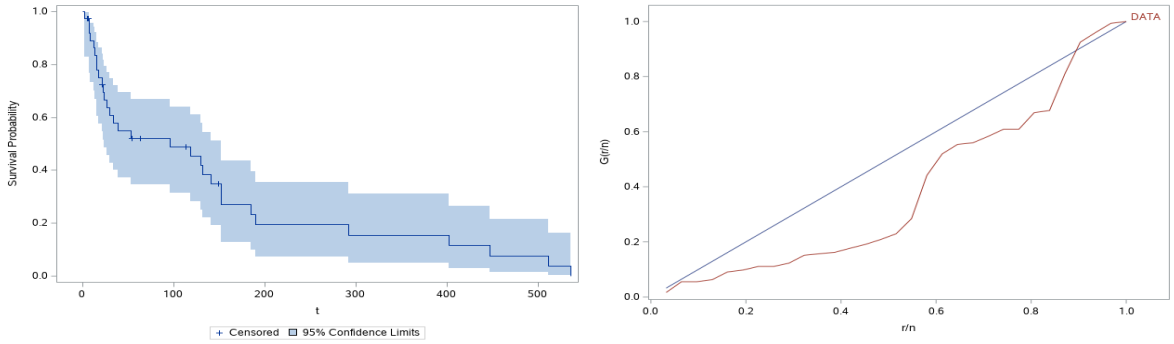

Figure 3. 4 : Product-limit survival estimate plot for first case for recorrence.

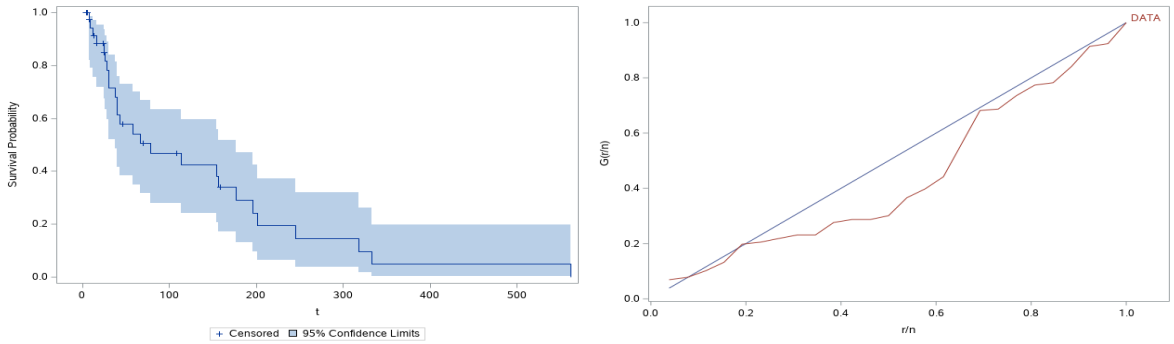

Figure 3. 5 : Product-limit survival estimate graph for second case for recorrence.

In addition, it is also of particular interest to estimate the risk rate of recurrence of the infection, since a risk model also attributes relevant information about the future health status of patients submitted to catheter insertion, but according to the equality of the curves Kaplan-Meyer, the best scenario would be to assume that the risk models for the two cases of recurrence are also the equals.

However, the empirical risk function always results in an increasing function and, consequently, it is assumed that the failure rate increases as time increases.

That way, in the context of modeling via continuous survival models, although the survival function implies in the form of the risk function, and decreases monotonously, although the probability of recurrence of the infection tends to zero, the risk of infection

may actually be reduced, but in three ways: or the risk may be reduced as time increases , or it can reduce up to a certain time after which the risk shows an increase, or it can show reduction after a period of time when the risk has been increasing.

It is worth mentioning that, although the probability decreases as time passes, the hazard of recurrence of infection does not always decrease as a patient is subject to different factors that are not summarize in the cure time.

Therefore, the GE2 model is considered as a competitor for the proposed adjustment, since both forms of risk presented in 3. 4 and 3. 5 can be captured by the risk function derived from this model.

In addition to the GE2 model, the Generalized Gamma (GG), Weibull Exponentiated (EW) and Exponentiated Log-Logistics (ELL) distributions, which also present the same forms for the risk function.

And more, based on the 38 observations presented in Mcgilchrist and Aisbett [40], time recurrence data still follow event type 1, if "infection occurs", or 0, if "censorship" occurs, highlighting that censored observations were recorded in the course of the study without the failure of interest event occurrering, "infection occurs". In this condition, the data have the mechanism of random censorship, and the equations (3.71), (3.72) and (3.73) were used to obtain bayesian estimates for this adjustment.

In the sequence, the table 3. 6 shows these estimates and the convergence test for the MCMC process performed. The table display the estimates for location and scale parameters in the coluns $\hat{\boldsymbol{\theta}}$, where α and β are location parameters and λ is the scale parameters, indicated by the column $\boldsymbol{\theta}$, and whith their respective standard-error and the HPD intervals with 95% confidence, in the columns std. err. and $IC_{95\%}(\theta)$, respectivaly.

Table 3. 6 : Results for MCMC process for models for first case for recorrence.

Cases		Results estimations for the models			Results from convergence test		
Model	$\boldsymbol{\theta}$	$\hat{\boldsymbol{\theta}}$	std. err.	$HPDI_{95\%}(\theta)$	test-stat	p-value	test outcome
GE2	α_1	0.9315	0.1934	(0.5930, 1.3145)	0.1601	0.3601	passed
	λ_1	0.0038	0.0017	(0.0004, 0.0067)	0.0720	0.7391	passed
	β_1	0.4154	0.2861	(0.0197, 0.9939)	0.4219	0.0633	passed
GG	α_2	0.1692	0.0866	(0.0836, 0.3390)	0.0532	0.8566	passed
	λ_2	4.2455	0.2718	(3.7092, 4.7542)	0.1714	0.3305	passed
	β_2	1.5158	0.2002	(1.1627, 1.9163)	0.1119	0.5295	passed
EW	α_3	6.6238	5.3099	(0.2060, 17.5152)	0.2900	0.1443	passed
	λ_3	0.3246	0.0556	(0.2211, 0.4335)	0.0604	0.8115	passed
	β_3	7.0170	3.1090	(2.7192, 13.8023)	0.2674	0.1675	passed
ELL	α_4	24.6154	16.3566	(2.1027, 56.4252)	0.0981	0.5940	passed
	λ_4	0.9042	0.1554	(0.6153, 1.2060)	0.0767	0.7114	passed
	β_4	2.3182	1.1238	(0.7308, 4.6215)	0.0890	0.6414	passed

The table 3. 7 in the sequence shows the results for some information criteria for the adjustments made by the recurrence time data for the first case.

Table 3. 7 : Results for measures of fit.

Criterion	GE2	GG	EW	ELL
AIC	369.7	413.1	369.8	372.3
AICC	370.4	413.8	370.5	373.0
BIC	374.6	418.0	374.7	377.2
DIC	368.3	368.6	377.7	368.6

The figure in the next page shows the plot for result of this adjustment, and analogous to what has been developed so far for the first case for recorrence, in the sequence follows

the adjustments to the hazard functions of the models GE2, GG, Weibull and ELL for the patients data for second case for recorrence.

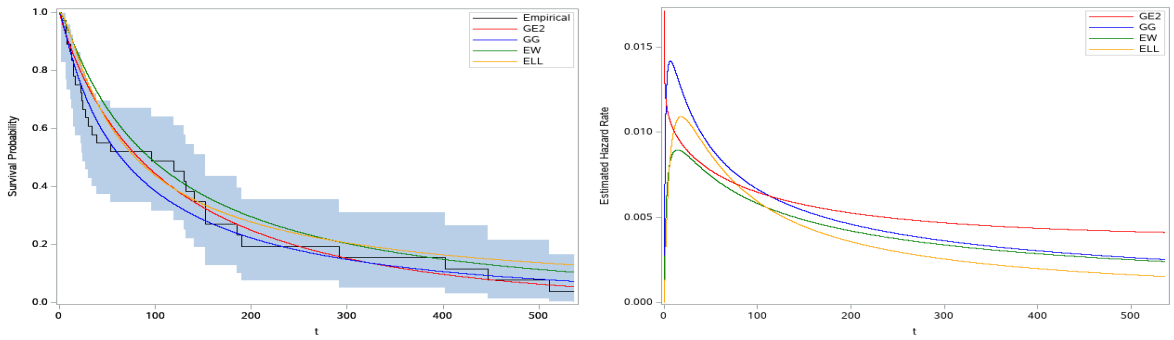

Figure 3. 6 : Survival and hazard graphs for the fitted models for first case for recorrence.

Given the candidate data-fitting models, it is conventionally preferred to choose the one that provides the value lowest information criterion, in this case, based on the criteria presented in table 3. 7 , it is concluded that the GE2 distribution provides the best adjust.

Table 3. 8 following shows the estimates obtained, the standard errors, and the HPD interval with credibility 95% for each of in the model estimates.

In the same table, with a significance level of 5%, the stationarity test results for the MCMC process convergence performed for this problem are presented for each parameter. The data in the table show that convergence was achieved.

Table 3. 8 : Results for MCMC process for models for second case for recorrence.

Cases		Results estimations for the models			Results from convergence test		
Model	$\boldsymbol{\theta}$	$\hat{\boldsymbol{\theta}}$	std. err.	$HPDI_{95\%}(\theta)$	test-stat	p-value	test outcome
GE2	α_1	1.3422	0.2784	(0.8336, 1.9247)	0.3450	0.1015	passed
	λ_1	0.0048	0.0022	(0.0009, 0.0092)	0.3888	0.0773	passed
	β_1	0.3504	0.2843	(0.0097, 0.8891)	0.2121	0.2452	passed
GG	α_2	0.1645	0.0784	(0.0837, 0.3239)	0.2126	0.2444	passed
	λ_2	4.4183	0.2291	(3.9349, 4.8332)	0.0695	0.7548	passed
	β_2	1.2023	0.1652	(0.9204, 1.5542)	0.0413	0.9261	passed
EW	α_3	103.7	57.6383	(13.9215, 220.0000)	0.3550	0.0953	passed
	λ_3	0.8504	0.2841	(0.4359, 1.4685)	0.4508	0.0532	passed
	β_3	1.6983	0.9914	(0.3435, 3.5922)	0.2831	0.1510	passed
ELL	α_4	23.0030	15.4581	(1.9036, 55.8937)	0.1002	0.5839	passed
	λ_4	1.0786	0.1916	(0.6944, 1.4422)	0.1604	0.3593	passed
	β_4	3.5417	2.0586	(0.9169, 7.5370)	0.1054	0.5590	passed

The table 3. 9 in the sequence shows the results for the information criteria for the bayesian models adjust resulting.

Table 3. 9 : Results for measures of fit.

Criterion	GE2	GG	EW	ELL
AIC	311.6	364.8	312.6	312.6
AICC	312.3	365.5	313.3	313.3
BIC	316.5	369.7	317.5	317.5
DIC	306.5	308.8	308.1	308.1

Table 3. 8 previous shows the estimates obtained, the standard errors, and the HPD interval with credibility 95% for each of in the model estimates.

Based on these estimates, replacing them in the theoretical model, we then have the models GE2, GG, EW and ELL adjusted for second case for recorrence time. Then, based on this results, the figure in the sequence shows the graphic result the adjustment.

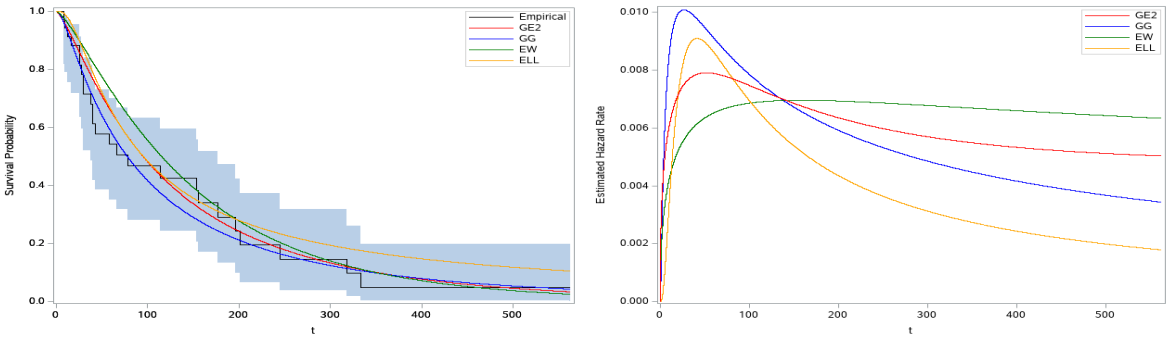

Figure 3. 7 : Survival and hazard plots for the fitted models for second case for recorrence.

Then, we have so ensure that the adjustment made by the estimated model GE2 is most appropriate in these contrast for patient data for first and second case for recorrence.

In the table in follows, are showns the means distance that the survival models fitted in comparison for empirical model as the representative for the true model. The comparison turns for the four models as the differences weighted by number of terms between the estimated for respective fitted survival function value and or corresponding value empirical obtained through of the Kaplan-Mayer.

Table 3. 10 : Average distance measures.

Recorrence	GE2	GG	EW	ELL
First case	0.0547	0.0436	0.0775	0.0770
Second case	0.0461	0.0378	0.0788	0.0673

Chapter 4

Final Results and Conclusions

4.1 Final Results and Conclusions

The present work does not present a chapter for the theoretical framework or basic concepts used for the proposals made, but sought, when a given concept was used, to refer to it and describe its relevance with subtlety.

To compensate for this absence, in the introduction to this work, the most relevant concepts are described through four subsections. In particular, a bibliographic review that highlights the genesis of the models studied, mainly for the distribution from which they originate, was presented in the subsection 1.1.

The subsections 1.2 and 1.3 presented a detailed summary of the intended results in the present study and described the structure of how the work was developed, pointing out what and how the approaches were carried out. In the subsection 1.4 we sought to give a more detailed focus to the concepts that are less relevant to the study's proposals, but which were applied in parallel to the proposed objectives for obtaining the results and even for the applications considered.

Under valuable mathematical concepts, where theoretical points from different lines of study were approached such as those of Infinite Series, Special Functions and the Mathematical Analysis itself which somehow absorbs these fields, the properties of the models were demonstrated, proposing them exclusively to support in processes from the resolution in real problens with these models.

However, the work sought to present the approach of these problems according to the assumed properties and, in some cases, demonstrated for the models approached here. The property of as constitutes the forms of hazard rate, derived from the density and survival functions of the models, conditioned to the scale parameter inherited from the Geometric distribution, was the main one.

It has also been seen, and demonstrated, that the extensions for the Geometric Exponential distribution provide hazard rate function that inherits the memory loss property of the Exponential and Geometric functions that generated them, as well as the relationships that their respective parameters have for a given form from the hazard rate function, making clear the parametric conditions under which they are generated.

Particularly, regarding the considered EEG models, varieting in time only and/or in the time with presence the covariables, it was clearly highlighted that, in both cases approached, whatever the value of the $\lambda > 0$ parameter assumed, the hazard function derived from this model has a decreasing behavior for cases where $\gamma < 1$ and increasing when $\gamma > 1$, where λ is the scale or rate parameter of the model that, becouse by construction

is inherited from the Exponential distribution, and γ is its shape parameter. When $\gamma = 1$ the EEG model is reduced to the Geometric Exponential model.

Similarly, regarding the GE2 model, it was clarified that for any assumed $\lambda > 0$ value, when $\alpha < 1$ the resulting hazard rate function takes the decreasing form for any $\beta < 1$ and bathtub form for any $\beta > 1$, but in the case where $\alpha > 1$, the hazard rate is unimodal when $\beta < 1$ and increasing if $\beta > 1$.

Still on the hazard rate forms of the GE2 model, in the cases where (1) $\alpha = \beta = 1$, which (2) $\beta = 1$ for any $\alpha > 0$ with $\alpha \neq 1$, and that (3) $\alpha = 1$ for any $\beta > 0$ with $\alpha \neq 1$, the GE2 model is reduced to the Geometric Exponential model. In (1) and (2) this is immediate by substitution, as can be seen in [3], and in (3) this is guaranteed in [35].

Under the parametric conditions evidenced for the EEG and GE2 models, the computational approach is simplified, since in the bayesian case, for example, an appropriate initial kick is needed to calculate model estimates via MCMC, and knowing a favorable region for this kick, the markovian process reaches its convergence faster.

Therefore, knowing the hazard rate form for a given data set and, consequently, where the scale or rate parameter from model EEG converges, the initial kick to the γ estimate is simplified because of the region for this parameter to be previously known. The same way that is for the kicks for the estimates of α and β in the GE2 model, in the both cases the kicking λ are made according to the inherited memory loss property.

Also worth highlighting from the conclusions obtained about the properties demonstrated, the fact of the EEG and GE2 distributions are semi pathological in consequence from the yours parameters γ and β, respectivaly. Depending of the parametric condition that is considered for γ or β, exist or not a mean and variance for distribution EEG or GE2, and in this constatation we prove the parametric conditions under which we can guarantee that the expected value exists for the EEG and GE2 distribution, obtaining it in terms of infinite series.

The work shows that, equipped with these propertie, the use r-th moment it is eventually impracticable, mainly for distribution GE2 by virtue your expression, and even though it exists under certain parametric conditions the expected value and variance for EEG and GE2, presuppositions for important statistical concepts, such as expected Fisher's matrix, the Central Limit Theorem and the moments of order r are not achieved under in the complementary cases.

Based on elaborate computational resources to support the approaches considered in the study of these models, the work considers the outline of the geometric shapes of some functions, such as the confidence bands presented in the applications and in particular for the density, survival and hazard surface considered for the EEG model in the presence of covariables. Due to these computational resources, as a consequence of the obtained properties, the simulation study was concluded as follows.

The elaboration of the statistical simulation under the classical and bayesian approach via MLE and MCMC, respectively, required delicate computational processes for the extraction and calculation of statistics relevant to the results evaluation criteria. As described in chapter 1, this work was developed focusing on the study area of survival analysis and in these simulations were considered at least 4 samples sizes and 3 cases of censorships, one of them being the case for complete-time data.

As shown, we present the results of 3 simulation studies developed for (1) EEG model in the presence of censorship, (2) for the EEG model in the presence of censorship and multiple variables and (3) for the GE2 model in the presence of censorship. The 3 case studies were chosen according to the forms of the hazard function derived from the respective simulated models.

All 3 results were satisfactory, although with specific features. It is worth mentioning

the slow convergence of the probability of empirical coverage in the case of the simulation (3) described in the subsection 3.2, where both bayesian and classical estimators showed equal conditions and properties over their estimates in any parametric case, censorship ratio and samples larger than 20 observations. Particularly, in small samples the bayesian estimators were superior in all the considered criteria.

In the subsection 2.2.3, the case of simulation (1) shows that, despite the high quality of the estimates obtained by the MLE, this quality was attributed to large samples and, in cases of low censorship, those lower than $tau = 0.20$. For these estimators, it was common to observe reasonable situations with relevant sample sizes and censorship ratios, such as $n = 50$ and $\tau = 0.2$, where estimates attributed high estimation errors, pointing to an estimated inefficient.

In this cases, the simulation study highlights that bayesian estimators are significantly superior to the classical one. Although it shows high estimation errors for high censorship ratios, it is still significantly better than MLE in these situations.

When the same model was simulated, also in the presence of censorship but considering multiple variables, as developed in simulation (2), the efficiency of the bayesian and classical estimators are analogous, and the results highlight that both the MLE and MCMC methods are satisfactorily accurates in the model in the presence of censorship and multiple variables, even with high proportion of censorship, showing that in both inference cases the presence of covariables in the model attributes high flexibility and excellent results in relation to their estimators.

With regard to the simulations developed, it is evident that while MLE is preferable for its speed in large samples, MCMC is preferable for its accuracy in small sample sizes, but what is noteworthy is that the presence of multiple variables compensates the penalty that estimates suffer by high censorship ratios, as shown in simulations (1) and (2).

Above all, the 3 studies allow to conclude that the estimators of these models are appropriate in the considered sample conditions and in view of their particular properties, as the biparametric EEG model in the presence of censorship and in the presence of covariables with $p + 1$ parameters in the presence of censorship or the GE2 model as a 3 parameter model and under censorship.

Focused in the EEG parametric model, the work presents the applications for this model in the presence of censorship and under censorship in the presence from covariable, respectively in the sub sections 2.2.4 and 2.3.2 . In these applications it is shown that this model is preferable in the contrasts performed, which was possible to conclude with its obtained properties, as described in 2.3.2 where, even though enhancing the specific competing model for a case of the proportional hazard rate, the EEG model presents a superior adjustment, because besides capturing the empirical hazard form presented by the data, it fits the empirical hazard with interpolation and adjustment satisfactory in the confidence range considered.

Similarly, in the case of the tri-parametric model, the application in the subsection 3.3 highlights that the GE2 model is also preferable over its competitors, that although they present excellent adjustments and capture the form of the risk rate function manifested empirically, all the information criteria presented, as well as the average distance between the survival function adjusted with the Kaplan-Meyer empirical, are the best because they are the smallest among all the models considered, which highlights the efficiency of the GE2 model over the others.

Bibliography

[1] M. H. AbuJarad and et al. Bayesian reliability analysis of marshall and olkin model. *Annals of Data Science*, pages 1–29, 2019.

[2] K. Adamidis, T. Dimitrakopoulou, and S. Loukas. On an extension of the exponential-geometric distribution. *Statistics & probability letters*, 73(3):259–269, 2005.

[3] K. Adamidis and S. Loukas. A lifetime distribution with decreasing failure rate. *Statistics & Probability Letters*, 39(1):35–42, 1998.

[4] S. Anwar, S. Shahab, and A. U. Islam. Accelerated life testing design using geometric process for marshall-olkin extended exponential distribution with type ii censored data. *International Journal of Scientific Engineering and Technology*, 3(5):538–542, 2014.

[5] S. D. Braguim. Modelo hierárquico log-logístico: uma aplicação no estudo da linezolida. Master's thesis, Universidade Estadual de Maringá, 2015.

[6] T. J. I. Bromwich. *An Introduction To The Theory Of Infinite Series.* Macmillan and Company, 1908.

[7] S. Chakraborti, F. Jardim, and E. Epprecht. Higher-order moments using the survival function: The alternative expectation formula. *The American Statistician*, 73(2):191–194, 2019.

[8] S. Chib and E. Greenberg. Hierarchical analysis of sur models with extensions to correlated serial errors and time-varying parameter models. *Journal of Econometrics*, 68(2):339–360, 1995.

[9] D. Collett. *Modelling Survival Data in Medical Research.* Chapman and Hall/CRC, 2015.

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

[11] S. Dey and et al. Estimation and prediction of marshall–olkin extended exponential distribution under progressively type-ii censored data. *Journal of Statistical Computation and Simulation*, 88(12):2287–2308, 2018.

[12] S. Durrleman and R. Simon. Flexible regression models with cubic splines. *Statistics in medicine*, 8(5):551–561, 1989.

[13] A. Erdélyi. *Higher Transcendental Functions.* McGraw-Hill, 1953.

[14] D. Gamerman and H. F. Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference.* Chapman and Hall/CRC, 2006.

[15] A. E. Gelfand and A. F. M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409, 1990.

[16] A. Gelman and et al. *Bayesian Data Analysis*. Chapman and Hall/CRC, 2013.

[17] M. S. Goodman, Y. Li, and R. C. Tiwari. Survival analysis with change point hazard functions. 2006.

[18] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, 2014.

[19] J. Guillera and J. Sondow. Double integrals and infinite products for some classical constants via analytic continuations of lerchś transcendent. *The Ramanujan Journal*, 16(3):247–270, 2008.

[20] R. D. Gupta and D. Kundu. Exponentiated exponential family: an alternative to gamma and weibull distributions. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, 43(1):117–130, 2001.

[21] A. Hájek. Interpretations of probability. 2002.

[22] H. Harald, A. Kaider, and G. Zlabinger. Assessing interactions of binary time-dependent covariates with time in cox proportional hazards regression models using cubic spline functions. *Statistics in medicine*, 15(23):2589–2601, 1996.

[23] E. K. Harris and A. Albert. *Survivorship Analysis for Clinical Studies*, volume 114. CRC Press, 1990.

[24] P. Heidelberger and P. D. Welch. A spectral method for confidence interval generation and run length control in simulations. *Commun. ACM*, 24(4):233–245, 1981.

[25] P. Heidelberger and P. D. Welch. Simulation run length control in the presence of an initial transient. *Operations Research*, 31(6):1109–1144, 1983.

[26] H. Heinzl and A. Kaider. Gaining more flexibility in cox proportional hazards regression models with cubic spline functions. *Computer methods and programs in biomedicine*, 54(3):201–208, 1997.

[27] K. R. Hess. Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. *Statistics in medicine*, 13(10):1045–1062, 1994.

[28] L. Hong. A remark on the alternative expectation formula. *The American Statistician*, 66(4):232–233, 2012.

[29] A. Jeffrey and D. Zwillinger. *Table of Integrals, Series, and Products*. Elsevier, 2007.

[30] P. Kitidamrongsuk. Discriminating between the extended exponential geometric distribution and the gamma distribution. *Unpublished Doctoral Dissertation). Assumption University of Thailand*, 2010.

[31] P. C. Lambert and P. Royston. Further development of flexible parametric models for survival analysis. *The Stata Journal*, 9(2):265–290, 2009.

[32] P. S. Laplace. *Philosophical Essay on Probabilities: Translated from the Fifth French Edition of 1825*, volume 13. Springer Science & Business Media, 1998.

[33] F. Louzada and et al. A new long-term lifetime distribution induced by a latent complementary risk framework. *Journal of Applied Statistics*, 39(10):2209–2222, 2012.

[34] F. Louzada, V. A. A. Marchi, and J. Carpenter. The complementary exponentiated exponential geometric lifetime distribution. *Journal of Probability and Statistics*, 2013, 2013.

[35] F. Louzada, V. A. A. Marchi, and M. Roman. The exponentiated exponential-geometric distribution: a distribution with decreasing, increasing and unimodal failure rate. *Statistics*, 48(1):167–181, 2014.

[36] F. Louzada, P. L. Ramos, and G. S. C. Perdoná. Different estimation procedures for the parameters of the extended exponential geometric distribution for medical data. *Computational and Mathematical Methods in Medicine*, 2016, 2016.

[37] F. Louzada, M. Roman, and V. G. Cancho. The complementary exponential geometric distribution: Model, properties, and a comparison with its counterpart. *Computational Statistics & Data Analysis*, 55(8):2516–2524, 2011.

[38] J. Loxton. Special values of the dilogarithm function. *Acta Arithmetica*, 43(2):155–166, 1984.

[39] 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.

[40] C. A. McGilchrist and C. W. Aisbett. Regression with frailty in survival analysis. *Biometrics*, pages 461–466, 1991.

[41] N. V. R. Mendoza, E. M. M. Ortega, and G. M. Cordeiro. The exponentiated-log-logistic geometric distribution: Dual activation. *Communications in Statistics-Theory and Methods*, 45(13):3838–3859, 2016.

[42] M. Mirjalili, H. Torabi, and H. Nadeb. Stress-strength reliability of exponential distribution based on type-i progressively hybrid censored samples. *Journal of Statistical Research of Iran JSRI*, 13(1):89–105, 2016.

[43] M. Mood, D. C. Boes, and F. A. Graybill. *Introduction to the Theory of Statistics*. McGraw-Hill, 1974.

[44] G. S Mudholkar and D. K. Srivastava. Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, 42(2):299–302, 1993.

[45] K. Najrullah, A. Tanwir, and A. K. Atha. Bayesian analysis of marshall and olkin family of distributions. 8(7):18692–18699, 2017.

[46] P. J. Olver. *Classical Invariant Theory*, volume 44. Cambridge University Press, 1999.

[47] L Pochhammer. Über hypergeometrische funktionen höherer ordnungen. *Crelles J*, 71:316–352, 1870.

[48] R. L. Prentice. Exponential survivals with censoring and explanatory variables. *Biometrika*, 60(2):279–288, 1973.

[49] F. Qi, X. T. Shi, and F. F. Liu. Several identities involving the falling and rising factorials and the cauchy, lah, and stirling numbers. *Acta Universitatis Sapientiae, Mathematica*, 8(2):282–297, 2016.

[50] P. L. Ramos. Aspectos computacionais para inferência na distribuição gama generalizada. 2014.

[51] P. L. Ramos and et al. Bayesian analysis of the generalized gamma distribution using non-informative priors. *Statistics*, 51(4):824–843, 2017.

[52] P. L. Ramos, F. Louzada, and E. Ramos. Posterior properties of the nakagamim distribution using noninformative priors and applications in reliability. *IEEE Transactions on reliability*, 67(1):105–117, 2017.

[53] T. C. S. Reis. Extensões da distribuição weibull aplicadas na análise de séries climatológicas. 2017.

[54] M. M. Ristić and D. Kundu. Marshall-olkin generalized exponential distribution. *Metron*, 73(3):317–333, 2015.

[55] M. M. Ristić and D. Kundu. Generalized exponential geometric extreme distribution. *Journal of Statistical Theory and Practice*, 10(1):179–201, 2016.

[56] K. Rosaiah, R. R. L. Kantam, and S. Kumar. Reliability test plans for exponentiated log-logistic distribution. *Economic Quality Control*, 21(2):279–289, 2006.

[57] P. Royston and M. K. B. Parmar. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in medicine*, 21(15):2175–2197, 2002.

[58] M. J. Rutherford, M. J. Crowther, and P. C. Lambert. The use of restricted cubic splines to approximate complex hazard functions in the analysis of time-to-event data: a simulation study. *Journal of Statistical Computation and Simulation*, 85(4):777–793, 2015.

[59] L. W. Schruben. Detecting initialization bias in simulation output. *Operations Research*, 30(3):569–590, 1982.

[60] R. B. Silva, W. B. Souza, and G. M. Cordeiro. A new distribution with decreasing, increasing and upside-down bathtub failure rate. *Computational Statistics & Data Analysis*, 54(4):935–944, 2010.

[61] D. J. Spiegelhalter and et al. Bayesian measures of model complexity and fit. *Journal of the royal statistical society: Series b (statistical methodology)*, 64(4):583–639, 2002.

[62] E. W. Stacy. A generalization of the gamma distribution. *The Annals of mathematical statistics*, 33(3):1187–1192, 1962.

[63] M. H. Tahir and G. M. Cordeiro. Compounding of distributions: a survey and new generalized classes. *Journal of Statistical Distributions and Applications*, 3(1):13, 2016.

[64] H. Torabi, F. L. Bagheri, and E. Mahmoudi. Estimation of parameters for the marshall–olkin generalized exponential distribution based on complete data. *Mathematics and Computers in Simulation*, 146:177–185, 2018.

[65] D. Zagier. The dilogarithm function. In *Frontiers in number theory, physics, and geometry II*, pages 3–65. Springer, 2007.