UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

DANIELE CRISTINA TITA GRANZOTTO

TRANSMUTATION MAPS: MODELING, STRUCTURAL PROPERTIES, ESTIMATION AND APPLICATIONS

Thesis submitted to the Departamento de – Des/UFSCar and to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment for the PhD degree in Statistics - Interinstitucional Program of Pos-Graduation in Statistics UFSCar-USP.

Advisor: Prof. Dr. Francisco Louzada Neto

São Carlos
January 2017

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

DANIELE CRISTINA TITA GRANZOTTO

MAPAS DA TRANSMUTAÇÃO: MODELAGEM, PROPRIEDADES ESTRUTURAIS, ESTIMAÇÃO E APLICAÇÕES

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

Orientador: Prof. Dr. Francisco Louzada Neto

São Carlos
Janeiro de 2017

Ficha catalográfica elaborada pelo DePT da Biblioteca Comunitária UFSCar
Processamento Técnico
com os dados fornecidos pelo(a) autor(a)

Granzotto, Daniele Cristina Tita
G765t Transmutation maps : modeling, structural properties, estimation and applications / Daniele Cristina Tita Granzotto. -- São Carlos : UFSCar, 2017.
101 p.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2016.

1. E-extension. 2. Log-logistic model. 3. Weibull model. 4. Cubic transmutation. 5. Quadratic transmutation. I. Título.

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

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado da candidata Daniele Cristina Tita Granzotto, realizada em 05/12/2016:

Prof. Dr. Francisco Louzada Neto
USP

Prof. Dr. Fernando Antonio Moala
UNESP

Prof. Dr. Paulo Henrique Ferreira da Silva
UFBA

Profa. Dra. Teresa Cristina Martins Dias
UFSCar

Profa. Dra. Vera Lucia Damasceno Tomazella
UFSCar

Ao meu esposo e filhos

Agradecimentos

Acredito que todo caminho em que nos aventuramos, nos deparamos com pessoas que nos ajudam a aliviar as angústias,auxiliando e incentivando a seguir sempre em frente. Nestes quatro anos de doutoramento são muitos os que me auxiliaram e apoiaram, me ajudando a ter forças para concluir mais esta importante etapa.

Primeiramente quero agradecer ao meu orientador e amigo Professor Francisco Louzada Neto que, além de compartilhar seus conhecimentos, acreditou no meu trabalhou e me incentivou desde o mestrado.

Tenho muito a agradecer ao meu esposo, Carlos, que colaborou imensamente para a realização deste projeto. Aos meus filhos, Iasmim e João, que apesar da minha ausência demonstraram muita compreensão, sempre me motivando a continuar em busca desta conquista. Ao meu irmão e sua esposa, Marcos e Roberta, que muitas vezes me acudiram e auxiliaram.

Muito obrigado ao professores do Deparamento de Estatística da Universidade Estadual de Maringá, departamento onde me graduei e hoje integro o corpo docente. Aos professores do Programa Interinstitucional de Pós-Graduação em Estatística da Universidades Federal de São Carlos e da Universidade de São Paulo por me acolherem e compartilharem comigo seus conhecimentos.

E enfim, a todos que contribuíram para a realização deste trabalho, seja de forma direta ou indireta, fica registrado aqui, o meu muito obrigado!

Resumo

Inicialmente, usamos os mapas de transmutação quadráticos para compor um novo modelo de probabilidade: a distribuição log-logística transmutada. Mapas de transmutação são uma forma conveniente de construção de novas distribuições, em especial de sobrevivência/confiabilidade, e compreendem a composição funcional da função de distribuição acumulada e da função de distribuição acumulada inversa (quantil) de um outro modelo. Uma descrição detalhada de suas propriedades, tais como, momentos, quantis, estatística de ordem, dentre outras estatísticas, juntamente com o estudo de sobrevivência e métodos de estimação clássico e Bayesiano, também fazem parte deste trabalho. Focando em análise sobrevivência, incluímos no estudo duas situações práticas comumente encontradas: a presença de variáveis regressoras, através do modelo de regressão transmutado log-logístico, e a presença de censura à direita. Em um segundo momento, buscando um modelo mais flexível que o transmutado, apresentamos uma generalização para esta classe de modelos, as distribuições transmutadas de rank cúbico. Usando a metodologia apresentada nesta primeira generalização, dois modelos foram considerados para compor as novas distribuições transmutadas cúbica: os modelos log-logístico e Weibull. Diante de problemas apresentados na classe transmutada de ordens quadrática e cúbica (tal como o espaço paramétrico restrito do parâmetro de transmutação λ), propomos neste trabalho, uma nova família de distribuição. Esta família, a qual chamamos e-transmutada ou e-extendida, é tão simples quanto o modelo transmutado, por incluir um único parâmetro ao modelo base, porém mais flexível do que a classe de modelos transmutados, sendo esta classe um caso particular da família proposta. Além disso, apresenta propriedades importantes, como ortogonalidade entre os parâmetros do modelo base e o parâmetro de e-transmutação, e espaço paramétrico não restrito para o parâmetro de e-transmutação ω, que é definido em toda reta real. Estudos de simulação e aplicações a dados reais foram realizados para todos os modelos e generalizações propostas.

Palavras-Chave: E-extensão, modelo log-logístico, modelo Weibull, transmutação cúbica, transmutação quadrática.

Abstract

Initially, we use the quadratic transmutation maps to compose a new probability model: the transmuted log-logistic distribution. Transmutation maps are a convenient way of constructing new distributions, in particular survival ones. It comprises the functional composition of the cumulative distribution function of one distribution with the inverse cumulative distribution (quantil) function of another. Its comprehensive description of properties, such as moments, quantiles, order statistics etc., along with its survival study and the classical and Bayesian estimation methods, are also part of this work. Focusing on analysis of survival, the study included two practical situations commonly found: the presence of regression variables, through the transmuted log-logistic regression model, and the presence of right censorship. In a second moment, searching for a more flexible model than the transmuted, we present its generalization, the transmuted distributions of cubic rank. Using the methodology presented in this first generalization, two models were considered to compose the new cubic transmuted distributions: the log-logistic and Weibull models. Faced with problems presented in the transmutated class of quadratic and cubic orders (such as the restricted parametric space of the transmutation parameter λ), we propose in this work, a new family of distribution. This family, which we call e-transmuted or e-extended, is as simple as the transmuted model, because it includes a single parameter to the base model, but more flexible than the class of transmuted models, once the transmuted is a particular case of the proposed family. In addition, the nem family presents important properties such as, orthogonality between the baseline model parameters and the e-transmutation parameter, along with unrestricted parametric space for the ω e-transmutation parameter, which is defined on the real line. Simulation studies and real data applications were performed for all proposed models and generalizations.

Keywords: E-extension, log-logistic model, Weibull model, cubic transmutation, quadratic transmutation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Survival analysis researchers are usually concerned with the proposition of new survival probability models. The challenge is to derive statistical survival probability models or simply survival distributions of real world lifetime phenomena that can represent more consistently the random behavior of experimental observations.

In this context, the literature on proposing new survival distributions is rich and growing rapidly. There are many papers extending standard survival distributions designed to serve as statistical survival models for a wide range of real lifetime phenomena, which do not follow any of the standard survival distributions. Interested readers can refer to Ghitany (2001), Marshall and Olkin (2007a), and their references, in which they discuss some common approaches for constructing lifetime distributions.

Generally transmutation maps are a convenient way of constructing new distributions, in particular, survival ones. According to Shaw and Buckley (2009), transmutation maps comprise the functional composition of the cumulative distribution function of one distribution with the inverse cumulative distribution (quantile) function of another. Motivated by the need for parametric families of rich and yet tractable distributions in financial mathematics Shaw and Buckley (2009) used a transmutation map of a non-Gaussian distribution. After that, some studies involving quadratic rank transmutation map could be observed in other application areas such as survival analysis and reliability. Aryal and Tsokos (2009) proposed a generalization of the extreme value distribution by using quadratic rank transmutation maps and applied this new distribution to analyze snow fall data in Midway Airport in the state of Illinois, USA. Furthermore, Aryal and Tsokos (2011) proposed the transmuted Weibull distribution as a generalization of the Weibull probability distribution and its usefulness was illustrated using two published data sets.

In this work, we propose for the first time the transmuted log-logistic model, obtained by considering a quadratic ranking transmutation map. A direct advantage of our modeling is the fact that the log-logistic distribution has a larger range of choices for the shape of the hazard function than Weibull or extreme value distributions do (see Ibrahim *et al.* (2005)

and Bennett (1983)). For instance, the log-logistic distribution can accommodate an unimodal hazard function in detriment of its competitors. The log-logistic distribution is the probability distribution of a random variable whose logarithm has a logistic distribution, Lawless (2011). It is similar in shape to the log-normal distribution but has heavier tails and its cumulative distribution function can be written in closed form, unlike that of the log-normal.

Moreover, we propose to apply the transmuted log-logistic model in a Bayesian context. In order to fit this new model, the subjective Bayesian analysis was used. To do so, we used the half-Cauchy prior distribution, cited by various authors (Polson and Scott (2012) and Gelman (2006)), as an alternative prior to an inverse Gamma distribution. Specially, Gelman (2006) made use of this particular prior for variance parameters in hierarchical models, which are our case.

Although several studies involving quadratic rank transmutation maps can be seen in many areas, such as survival analysis and reliability, they do not provide a regression approach and do not consider the presence of censorship. In our work, we particularly consider the presence of right censorship and, in order to fit this new model, we used profile methods, such as pure, adjusted and modified profile; see, for example, Barndorff-Nielsen and McCullagh (1993), Severini (1998) and Ferrari *et al.* (2007). Note that the pure likelihood method cannot estimate the parameters in samples with high percentages of censored points, which is what our application case showed.

A new order of transmuted models was introduced in Chapter 4, the cubic ranking transmutation map. This new order increases the flexibility of transmuted models and is able to analyze more complex data, for example, situations with bimodal hazard rates. In order to exemplify the applicability of a cubic ranking transmutation map, we used two of the best known distributions in reliability fields as a base: Weibull and log-logistic. Several mathematical properties of these new distributions, transmuted Weibull of order 2 and transmuted log-logistic of order 2, were derived.

Although transmutation maps are a convenient way of constructing new distributions, the restricted parametric space of the extra parameter λ may be a problem in some situations. As an alternative to this class of model, we present the regression e-transmuted family of model (or exponential transmuted), which has the property that the extra parameter λ can take any real value. Without the restricted parameter space for λ, this new model continues to present characteristics of a good model: it is simple, more flexible than the transmuted model and continues being interpretable.

Furthermore, an influential analysis was carried out in order to provide an indication of bad model fitting or influential observations; see Ortega *et al.* (2003) and Fachini *et al.* (2008). The bootstrap and Monte Carlo methods were used in order to validate the results presented. Moreover, two quadratic distance measures of goodness-of-fit: Anderson-Darling and Cramér-von Mises are used to verify the quality of the adjustment of the e-transmuted model.

1.1 Objectives

The main objective of this work was to study the transmuted models wich is a convenient way of constructing new distributions. Particularly, we have worked with the log-logistic distribution, which is a continuous probability distribution for a non-negative random variable (it can be used in survival analysis) and presents all functions with closed forms.

A second goal was to generalize the ranking of transmutation map. Thus, we introduced the cubic ranking transmutation map model in the last chapter; however, the kth ranking transmutation map can be seen as a continuation of this work.

Finally, we proposed a new family of distributions to solve the "problem" of the restricted parametric space of λ.

1.2 Organization of chapters

The dissertation was organized as a collection of papers and divided in five chapters besides the Introduction and Conclusion. It should be mentioned that these five chapters were accepted papers for publication, or submitted to specialized journals in the field.

Chapter 2 comprises the paper entitled "The Transmuted Log-Logistic Distribution: Modeling, Inference and Application to a Polled Tabapua Race Time up to First Calving Data", published by Communications in Statistics: Theory and Methods in 2014 and the paper entitled "The transmuted log-logistic regression model: A new model for time up to first calving of cows", published by Statistical Paper in 2015. In these papers, we proposed a new lifetime model called Transmuted Log-Logistic, which is a generalization of the well-known survival log-logistic model. We also presented the main statistical properties as well as their main functions and characteristics in the area of survival analysis. The usefulness of the model was exemplified by using a real dataset.

The third chapter presents the paper entitled "Hierarchical Transmuted Log-Logistic Model: A Subjective Bayesian Analysis", written by us and co-authored by Vera L. D. Tomazella and the paper entitled "Likelihood Based Inference for the Transmuted Log-Logistic Model in the Presence of Right Censored Lifetime", both submitted to specialized journals. There, we covered the transmuted log-logistic model by the Bayesian point of view and include censored observations in the transmuted model.

Chapter 4 presents a new order of ranking transmutation map, order 2, which we called the cubic transmuted model. The study was carried out in partnership with Professor Narayanaswamy Balakrishnan, a senior lecturer at McMaster University and the paper entitled "Cubic Rank Transmuted Distributions: Inferential Issues and Applications".

Finally, in Chapter 5, we present a new family of distribution called e-transmuted. The study was carried out in partnership with Professor Karim Anaya Izquierdo, a lecturer at University of Bath and the paper entitled “Goodness-of-Fit Testing to the Regression e-Transmuted Family of Distribution” was presented as a result. This study was done in a sandwich stage, in England.

Finally, in the last section, we present some conclusions and propose further researcher topics. Thi is followed by the bibliography and appendix sections.

Chapter 2

The transmuted log-logistic model

In this Chapter, we introduce the generalization of the log-logistic model by using a convenient way of constructing new distributions, i.e, the rank transmutation map methodology. Transmutation maps comprise the functional composition of the cumulative distribution function with the inverse cumulative distribution (quantile) function. The new three parameters model is called transmuted log-logistic model and its main properties and statistics are presented in this Chapter as well as the model construction.

2.1 The model

Let X be a nonnegative random variable denoting the lifetime of an individual in some population. The random variable X is said to be log-logistically distributed, with scale parameter μ and shape parameter β, if its probability density function (pdf) is given by

$$g\left(x\right) = \frac{e^{\mu}\beta x^{\beta-1}}{\left(1+e^{\mu}x^{\beta}\right)^{2}}, \tag{2.1}$$

where $\beta > 0$ and $-\infty < \mu < +\infty$ (note that the shape parameter β determines the tail behavior of the distribution). The corresponding cumulative distribution function is given by

$$G\left(x\right) = \frac{e^{\mu}x^{\beta}}{\left(1+e^{\mu}x^{\beta}\right)}. \tag{2.2}$$

According to Shaw and Buckley (2009), a ranking quadratic transmutation map has the following simple form:

$$F(x) = (1+\lambda)G(x) - \lambda G^{2}(x), \tag{2.3}$$

for $\mid \lambda \mid \leq 1$, where $G(x)$ is the cumulative distribution function of the baseline distribution. Observe that at $\lambda = 0$, we just obtain the baseline cumulative distribution function. Following this idea, several authors have considered extensions of some common survival distributions; see Aryal and Tsokos (2009) and Ibrahim *et al.* (2005).

The construction of the ranking quadratic transmutation map considered here is simple and intuitive. Let X_1 and X_2 be independent and identically distributed nonnegative random variables with distribution $G(x)$. Then, consider

$$\begin{cases} Y \stackrel{d}{=} \min(X_1, X_2), & \text{with probability } \pi, \\ Y \stackrel{d}{=} \max(X_1, X_2), & \text{with probability } 1-\pi, \end{cases}$$

where $0 \leq \pi \leq 1$. The distribution of Y is evidently

$$F_Y(x) = \pi Pr(\min(X_1, X_2) \leq x) + (1-\pi) Pr(\max(X_1, X_2) \leq x).$$

We know that $F_{(\min)}(x) = 1 - [1 - G(x)]^n$ and $F_{(\max)}(x) = [G(x)]^n$; see, for example, Arnold *et al.* (1992). Hence,

$$\begin{aligned} F_Y(x) &= \pi\left[1 - (1 - G(x))^2\right] + (1-\pi) G^2(x) \\ &= 2\pi G(x) + (1 - 2\pi) G^2(x). \end{aligned} \tag{2.4}$$

If we take $2\pi = \lambda$, the distribution in (2.4) is the well-known ranking quadratic transmutation (or simply the transmuted distribution) as presented in (2.3).

Following this idea, several authors have considered extensions from usual survival distributions; see for example Aryal and Tsokos (2009) and Ibrahim *et al.* (2005).

Hence, from (2.2) and (2.3), the cdf of a transmuted generalized log-logistic distribution is

$$F(x) = \frac{e^{\mu} x^{\beta}}{(1 + e^{\mu} x^{\beta})^2}\left(1 + e^{\mu} x^{\beta} + \lambda\right). \tag{2.5}$$

The corresponding pdf is given by

$$f(x) = \frac{e^{\mu} \beta x^{\beta-1}\left[\left(1 + e^{\mu} x^{\beta}\right) - \lambda\left(e^{\mu} x^{\beta} - 1\right)\right]}{(1 + e^{\mu} x^{\beta})^3}. \tag{2.6}$$

Note that the transmuted log-logistic distribution is an extended model to analyze more complex data and it generalizes some important distributions in reliability analysis. The log-logistic distribution is clearly a special case for $\lambda = 0$. Some of the possible shapes of the transmuted log-logistic distribution were illustrated by Figure 2.1 for selected values of parameters λ and β and for $\mu = 1$. The λ is responsible for introducing skewness into the log-logistic distribution. This is in full agreement with Shaw and Buckley (2009), who pointed out that the introduction of skewness into a distribution is a direct effect of the transmutation maps.

The following results show that the random variable X has a probability density func-

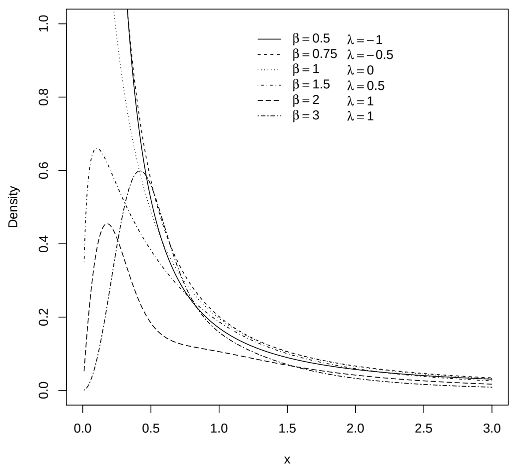


Figure 2.1: *pdf of transmuted log-logistic distribution for $\mu = 1$ and different values of λ and β.*

tion given by (2.6).

Proposition 2.1.1 *Let X be a non-negative random variable with log-logistic distribution and the quadratic transmutation is given, respectively, by Equations (2.1) and (2.3), then X is distributed according to a transmuted log-logistic with probability function given by (2.6).*

Proof According to Mood *et al.* (1974), any function $f(.)$ with domain the real line and conterdomain $[0, \infty)$ is defined as a probability density function if and only if

(i) $f(x) \geq 0$ for all x and;

(ii) $\int_{-\infty}^{\infty} f(x)dx = 1$.

Therefore, the first property is satisfied for all $x > 0$. The second property is shown below:

$$\int_0^{+\infty} f(x)dx = \int_0^{+\infty} e^{\mu}\beta x^{\beta-1}\frac{\left[\left(1+e^{\mu}x^{\beta}\right) - \lambda\left(e^{\mu}x^{\beta}-1\right)\right]}{\left(1+e^{\mu}x^{\beta}\right)^3}dx =$$
$$\int_0^{+\infty} e^{\mu}\beta x^{\beta-1}(1+e^{\mu}x^{\beta})^{-3}dx + \int_0^{+\infty} e^{2\mu}\beta x^{2\beta-1}(1+e^{\mu}x^{\beta})^{-3}dx -$$
$$\int_0^{+\infty} \lambda e^{2\mu}\beta x^{2\beta-1}(1+e^{\mu}x^{\beta})^{-3}dx + \int_0^{+\infty} \lambda e^{\mu}\beta x^{\beta-1}(1+e^{\mu}x^{\beta})^{-3}dx$$

Replacing $y = x^{\beta}e^{\mu}$, $x = e^{-\mu/\beta}y^{1/\beta}$ and $dy = e^{\mu}\beta x^{\beta-1}dx$ in the equation above, it follows that

$$\int_0^{+\infty} f(x)dx = (\lambda+1)\int_0^{+\infty}(1+y)^{-3}dy + (1-\lambda)\int_0^{+\infty} y(1+y)^{-3}dy.$$

We know that

$$B(\upsilon,\omega)=\int_0^{+\infty}\frac{z^{\omega-1}}{(1+z)^{\upsilon+\omega}} \tag{2.7}$$

in terms of Gamma function, we have

$$B(\upsilon,\omega)=\frac{\Gamma(\upsilon)\Gamma(\omega)}{\Gamma(\upsilon+\omega)}=\frac{(\upsilon-1)!(\omega-1)!}{(\upsilon+\omega-1)!}.$$

Then, the integrate is given by

$$(\lambda+1)\int_0^{+\infty}(1+y)^{-3}dy=\frac{\lambda+1}{2} \tag{2.8}$$

and

$$(1-\lambda)\int_0^{+\infty}y(1+y)^{-3}dy=\frac{1-\lambda}{2}. \tag{2.9}$$

Finally, adding the result of (2.8) with (2.9), we have

$$\frac{\lambda+1}{2}+\frac{1-\lambda}{2}=1.$$

Therefore, Equation (2.6) is a density function.

2.2 Properties of the model

2.2.1 Moments and quantiles

In this section, we shall present the moments and quantiles for the transmuted log-logistic distribution. The r^{th} order moments of a transmuted log-logistic random variable X is given by

$$\begin{aligned}
E\left(X^r\right) &= \int_0^{+\infty} x^r\frac{e^{\mu}\beta x^{\beta-1}\left[\left(1+e^{\mu}x^{\beta}\right)-\lambda\left(e^{\mu}x^{\beta}-1\right)\right]}{\left(1+e^{\mu}x^{\beta}\right)^3}\\
&= (1+\lambda)\int_0^{+\infty}\frac{e^{\mu}\beta x^{r+\beta-1}}{(1+e^{\mu}x^{\beta})^3}dx+(1-\lambda)\int_0^{+\infty}\frac{e^{2\mu}\beta x^{r+2\beta-1}}{(1+e^{\mu}x^{\beta})^3}dx\\
&= (1+\lambda)e^{-\frac{\mu r}{\beta}}\int_0^{+\infty}\frac{y^{\frac{r}{\beta}}}{(1+y)^3}dy+(1-\lambda)e^{-\frac{\mu r}{\beta}}\int_0^{+\infty}\frac{y^{\frac{r}{\beta}+1}}{(1+y)^3}dy.
\end{aligned}$$

Then, the r^{th} order moments of a transmuted log-logistic is

$$E\left(X^r\right)=\frac{\pi r}{\beta^2}e^{-\frac{\mu r}{\beta}}(\beta-r\lambda)\csc\left[\frac{\pi r}{\beta}\right].$$

Therefore, the expected value $E(X)$ and variance $Var(X)$ of a transmuted log-logistic random variable X are, respectively, given by

$$E(X) = e^{\frac{-\mu}{\beta}} \frac{\pi}{\beta^2}(\beta - \lambda) \csc\left[\frac{\pi}{\beta}\right] \tag{2.10}$$

and

$$Var(X) = e^{\frac{-2\mu}{\beta}} \frac{\pi}{\beta^2}\left[2(\beta - 2\lambda)\csc\left(\frac{2\pi}{\beta}\right) - \frac{\pi}{\beta^2}(\beta-\lambda)^2 \csc^2\left(\frac{\pi}{\beta}\right)\right]. \tag{2.11}$$

The q^{th} quantile x_q of the transmuted log-logistic distribution can be obtained from (2.5) as

$$x_q = \left[\frac{\sqrt{(1+\lambda)^2 - 4q\lambda} + 2q - (1+\lambda)}{2e^{\mu}(1-q)}\right]^{1/\beta}. \tag{2.12}$$

In particular, the median is given by

$$x_{0.5} = e^{-\mu/\beta}\left[\sqrt{1+\lambda^2} - \lambda\right]^{1/\beta}. \tag{2.13}$$

2.2.2 Random number generation

According to Ross (2009), we have the following Proposition 2.2.1.

Proposition 2.2.1 *Suppose that U is a random variable with a uniform distribution pattern, $U \sim Uniform(0,1)$. Therefore, for a distribution function F, the random variable X defined by $X = F^{-1}(U)$ has distribution function $F(x)$.*

Proposition 2.2.1 shows that we can generate a continuous random variable X from its cumulative distribution function. For that, we generate an uniform value $(0,1)$ and obtain $X = F^{-1}(U)$. Using this method, we generate random numbers of a transmuted log-logistic distribution when parameters μ, β and λ are known as follows,

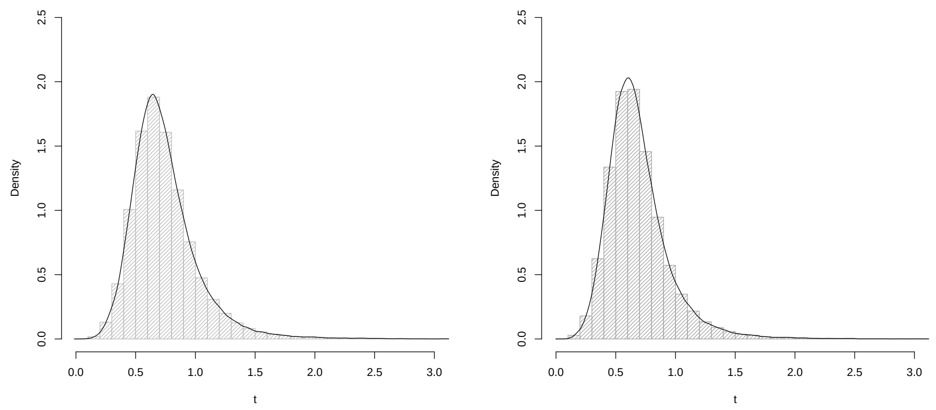
$$\frac{e^{\mu}x^{\beta}}{(1+e^{\mu}x^{\beta})^2}\left(1 + e^{\mu}x^{\beta} + \lambda\right) = u, \tag{2.14}$$

where $u \sim U(0,1)$. Isolating x in the equation above, we obtain

$$x = \left[\frac{\sqrt{(1+\lambda)^2 - 4u\lambda} + 2u - (1+\lambda)}{2e^{\mu}(1-u)}\right]^{1/\beta}, \tag{2.15}$$

similar to 2.12.

The histograms of Figures 2.2-a and 2.2-b were generated using Equation (2.15), from which we observe a full agreement to the theoretical curve with generated values of the transmuted log-logistic distribution.



(a) (b)

Figure 2.2: *Comparing the theoretical curve with generated values of transmuted log-logistic distribution for* $\mu = 2$, $\beta = 5$ *and: (a)* $\lambda = -0.2$*; (b)* $\lambda = 0.2$.

2.2.3 Survival analysis

The reliability function $R(t)$, which is the probability of an item not failing prior to some time t, is defined by $R(t) = 1 - F(t)$. The reliability function of a transmuted log-logistic distribution is given by

$$R(t) = \frac{1 + e^{\mu} t^{\beta}(1-\lambda)}{(1 + e^{\mu} t^{\beta})^2}. \tag{2.16}$$

The other characteristic of interest of a random variable is the hazard rate function defined by

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{R(t)}, \tag{2.17}$$

which is an important quantity characterizing a life phenomenon. It can be loosely interpreted as the conditional probability of failure, given the fact that it has survived to time t; see Lawless (2011). The hazard rate function for a transmuted log-logistic random variable is given by

$$h(t) = \frac{\beta e^{\mu} t^{\beta-1}\left[1 + e^{\mu} t^{\beta} - \lambda(e^{\mu} t^{\beta} - 1)\right]}{(1 + e^{\mu} t^{\beta})(1 + e^{\mu} t^{\beta} - \lambda e^{\mu} t^{\beta})}. \tag{2.18}$$

The hazard function of the transmuted log-logistic distribution has the following properties:

(i) If $\lambda = 0$, we have the log-logistic hazard as a particular case given by

$$h(t) = \frac{\beta e^{\mu} t^{\beta-1}}{(1 + e^{\mu} t^{\beta})}.$$

It can be easily shown that $h(t)$ is increasing for $\beta \leq 1$ and, for $\beta > 1$, the hazard function initially increases to the maximum $T_{\max} = e^{-\mu/\beta}(\beta - 1)^{1/\beta}$ and tends to zero for $t \to \infty$, Figure 2.3-a.

(ii) If $\lambda = -1$, we have

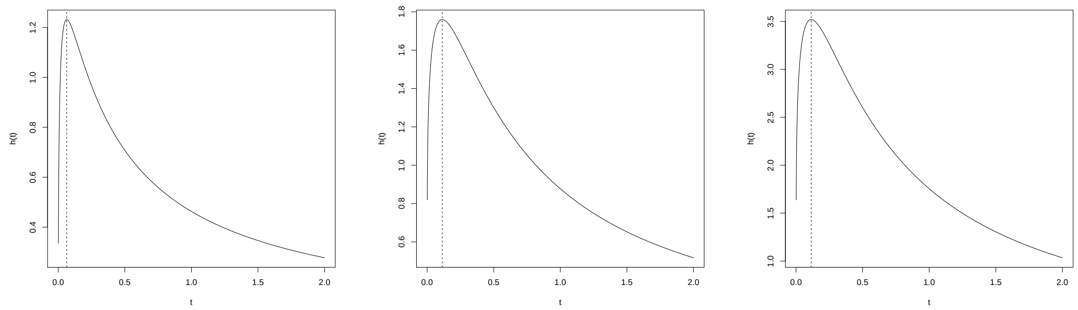
$$h(t) = \frac{2\beta e^{2\mu} t^{2\beta-1}}{(1 + e^{\mu} t^{\beta})(1 + 2e^{\mu} t^{\beta})},$$

which is increasing for $\beta \leq \frac{1}{2}$ and unimodal for $\beta > \frac{1}{2}$ with the maximum in $T_{\max} = \left[\frac{3(\beta-1)+\sqrt{9\beta^2-2\beta+1}}{4e^{\mu}}\right]^{1/\beta}$, Figure 2.3-b.

(iii) If $\lambda = 1$, we have

$$h(t) = \frac{2\beta e^{\mu} t^{\beta-1}}{1 + e^{\mu} t^{\beta}},$$

which is increasing for $\beta \leq 1$ and unimodal for $\beta > 1$ with the maximum in $T_{\max} = \exp\left[\left(\ln(\beta - 1) - \mu\right)/\beta\right]$, Figure 2.3-c.

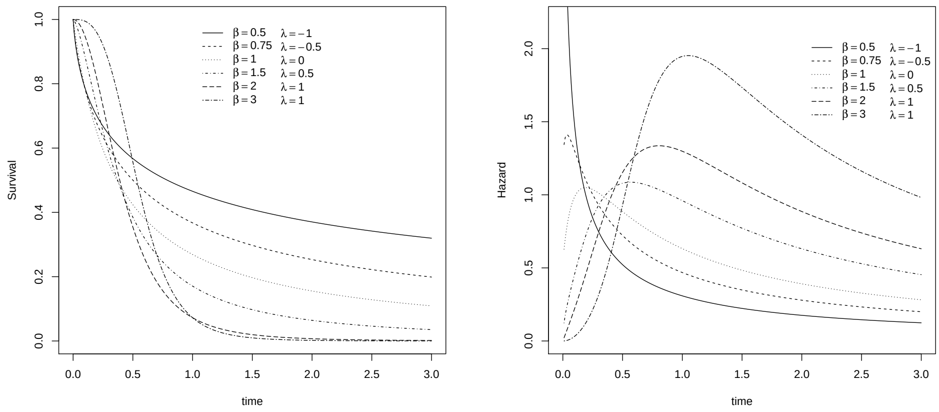


(a) (b) (c)

Figure 2.3: *Some examples of hazard curves with the maximum points for* $\lambda = 0, -1$ *and* 1, *respectively.*

Some properties and characterization of the risk of an event occurring at a given point, conditioned by the fact that the event has not already occurred can be seen in Chechile (2003). Figure 2.4-a and b illustrate, respectively, the reliability and hazard behavior of a transmuted log-logistic distribution as the value of the parameter λ ranges from -1 to 1.

In the survival analysis literature, the main relationship between the hazard and the reliability function is given by the cumulative hazard rate function $H(t)$, that is $H(t) =$



(a) (b)

Figure 2.4: *Survival and hazard curves.*

$-\ln R(t)$. In our case,

$$H(t) = 2\ln(1 + e^{\mu}t^{\beta}) - \ln(1 + e^{\mu}t^{\beta} - \lambda e^{\mu}t^{\beta}), \tag{2.19}$$

where $H(0) = 0$, $H(t)$ is nondecreasing for all $t \geq 0$ and $\lim_{t\to\infty} H(t) = \infty$.

2.2.4 Order statistics

According to Aryal and Tsokos (2011), suppose that we have a system containing two components with each of them having independent and identical "base" distribution, for example, a log-logistic. If the components are connected in series, then the overall system will have the transmuted baseline distribution with $\lambda = 1$ whereas if the components are parallel, then the overall system will have a transmuted baseline.

It has been observed that a transmuted log-logistic distribution with $\lambda = 1$ is the distribution of $\min(X_1, X_2)$ and a transmuted log-logistic distribution with $\lambda = -1$ is the distribution of the $\max(X_1, X_2)$ where X_1 and X_2 are independent and identically distributed 2-parameter log-logistic random variables.

Now, suppose that $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ denotes the order statistics of a random sample $X_1, X_2, \ldots, X_n$ from a continuous population with cumulative density function $F_X(x)$ and probability density function $f_X(x)$. Therefore, the pdf of $X_{(j)}$ is given by

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x)[F_X(x)]^{j-1}[1 - F_X(x)]^{n-j}, \tag{2.20}$$

for $j = 1, 2, \ldots, n$.

Then, the pdf of the j^{th} order transmuted log-logistic random variable $X_{(j)}$ is given by

$$
\begin{aligned}
f_{X_{(j)}}(x) \;&=\; \frac{n!}{(j-1)!(n-j)!}\frac{e^{\mu}\beta x^{\beta-1}}{(1+e^{\mu}x^{\beta})^{2n+1}} \\
&\left\{\left[(1+e^{\mu}x^{\beta})-\lambda(e^{\mu}x^{\beta}-1)\right][e^{\mu}x^{\beta}(1+e^{\mu}x^{\beta}+\lambda)]^{j-1}\times\right. \\
&\left.[(1+e^{\mu}x^{\beta})^{2}+e^{\mu}x^{\beta}(1+e^{\mu}x^{\beta}+\lambda)]^{n-j}\right\}.
\end{aligned} \tag{2.21}
$$

Using Equation (2.21), the pdf of the n^{th} order transmuted log-logistic statistics $X_{(n)}$ is given by

$$
f_{X_{(n)}}(x)=\frac{ne^{\mu}\beta x^{\beta-1}}{(1+e^{\mu}x^{\beta})^{2n+1}}\left\{\left[(1+e^{\mu}x^{\beta})-\lambda(e^{\mu}x^{\beta}-1)\right]\left[e^{\mu}x^{\beta}(1+e^{\mu}x^{\beta}+\lambda)\right]^{n-1}\right\}, \tag{2.22}
$$

while the 1^{st} order is given by

$$
\begin{aligned}
f_{X_{(1)}}(x) \;&=\; \frac{ne^{\mu}\beta x^{\beta-1}}{(1+e^{\mu}x^{\beta})^{2n+1}}\left\{\left[(1+e^{\mu}x^{\beta})-\lambda(e^{\mu}x^{\beta}-1)\right]\times\right. \\
&\left.\left[(1+e^{\mu}x^{\beta})^{2}+e^{\mu}x^{\beta}(1+e^{\mu}x^{\beta}+\lambda)\right]^{n-1}\right\}.
\end{aligned} \tag{2.23}
$$

Note that $\lambda=0$ yields the order statistics of the two parameter log-logistic distribution.

2.3 Parameter estimation of the regression model

In order to present the parameter estimation, let us first define the transmuted log-logistic regression model. For that, let Y be a random variable denoting the lifetimes with pdf (2.6) and $\mu=\gamma(\boldsymbol{x})$ as a parameter depending on a covariate vector $\boldsymbol{x}=(1,x_1,\ldots,x_p)'$ such as

$$
\gamma(\boldsymbol{x})=\gamma_0+\gamma_1x_1+\ldots+\gamma_px_p.
$$

Then, the pdf (2.6) may be written as

$$
f\left(y\mid\gamma(\boldsymbol{x})\right) \;=\; \frac{e^{\gamma(\boldsymbol{x})}\beta y^{\beta-1}\left[\left(1+e^{\gamma(\boldsymbol{x})}y^{\beta}\right)-\lambda\left(e^{\gamma(\boldsymbol{x})}y^{\beta}-1\right)\right]}{\left(1+e^{\gamma(\boldsymbol{x})}y^{\beta}\right)^{3}}, \tag{2.24}
$$

where $\lambda>0$, $\beta>0$ and $\gamma(\boldsymbol{x})$ is a regression defined above.

The maximum likelihood estimation of the parameters that are inherent within the transmuted log-logistic regression probability distribution function is given by the following: Let $y_1,y_2,\ldots,y_n$ be a sample of size n from a transmuted log-logistic distribution and $\boldsymbol{x}=(1,x_1,x_2,\ldots,x_p)'$ be a vector of covariates. Moreover, consider the following relationship between the vector of covariates and the parameters $\gamma(\boldsymbol{x})=\gamma_0+\gamma_1x_1+\ldots+\gamma_px_p$.

Hence, the log-likelihood function is given by

$$\begin{aligned} l &= n\ln\beta+\sum_{i=1}^{n}\sum_{j=0}^{p}\gamma_j x_{ij}+\sum_{i=1}^{n}\ln\left[1+e^{\sum_{j=0}^{p}\gamma_j x_{ij}}y_i^{\beta}-\lambda\left(e^{\sum_{j=0}^{p}\gamma_j x_{ij}}y_i^{\beta}-1\right)\right] \\ &\quad +(\beta-1)\sum_{i=1}^{n}\ln y_i-3\sum_{i=1}^{n}\ln\left(1+e^{\sum_{j=0}^{p}\gamma_j x_{ij}}y_i^{\beta}\right). \end{aligned} \tag{2.25}$$

Therefore, the normal equations are given by

$$\begin{aligned} \frac{\partial l}{\partial \gamma_j} &= \sum_{i=1}^{n}\left[\frac{(1-\lambda)y_i^{\beta}x_{ij}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]}{1+\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]y_i^{\beta}-\lambda\left[\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]y_i^{\beta}-1\right]}\right] \\ &\quad +\sum_{i=1}^{n}x_{ij}-3\sum_{i=1}^{n}\left[\frac{x_{ij}y_i^{\beta}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]}{1+y_i^{\beta}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]}\right], \qquad j=0,1,\ldots,p \end{aligned}$$

$$\begin{aligned} \frac{\partial l}{\partial \beta} &= \frac{n}{\beta}+\sum_{i=1}^{n}\left[\frac{(1-\lambda)y_i^{\beta}\ln(y_i)\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]}{1+y_i^{\beta}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]-\lambda\left[y_i^{\beta}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]-1\right]}\right] \\ &\quad +\sum_{i=1}^{n}\ln(y_i)-3+\sum_{i=1}^{n}\left[\frac{y_i^{\beta}\ln(y_i)\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]}{1+y_i^{\beta}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]}\right], \\ \frac{\partial l}{\partial \lambda} &= \sum_{i=1}^{n}\frac{1-y_i^{\beta}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]}{1+y_i^{\beta}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]-\lambda\left[y_i^{\beta}\exp\left[\sum_{j=0}^{p}\gamma_j x_{ij}\right]-1\right]}. \end{aligned}$$

The maximum likelihood estimator (MLE) $\hat{\boldsymbol{\theta}}=(\hat{\gamma}_0,\ldots,\hat{\gamma}_p,\hat{\beta},\hat{\lambda})'$ can be obtained by solving the above nonlinear system of equations. It is usually more convenient to use nonlinear optimization algorithms such as quasi-Newton or Newton-Raphson algorithms to numerically maximize the log-likelihood function given in (2.25).

Following Aryal and Tsokos (2011), in order to compute the standard error and asymptotic confidence interval, we use the usual large sample approximation in which the MLEs of θ can be treated as being approximately a $(p+3)$-variate normal distribution. Hence as $n\to\infty$, the asymptotic distribution of the MLE $(\hat{\gamma}_0,\ldots,\hat{\gamma}_p,\hat{\beta},\hat{\lambda})$ is given by

$$\begin{pmatrix}\hat{\gamma}_0\\ \hat{\gamma}_1\\ \vdots\\ \hat{\gamma}_p\\ \hat{\beta}\\ \hat{\lambda}\end{pmatrix}\sim N\left[\begin{pmatrix}\hat{\gamma}_0\\ \hat{\gamma}_1\\ \vdots\\ \hat{\gamma}_p\\ \hat{\beta}\\ \hat{\lambda}\end{pmatrix},\begin{pmatrix}\hat{V}_{11} & \hat{V}_{12} & \cdots & \hat{V}_{1(p+3)}\\ \hat{V}_{21} & \hat{V}_{22} & \cdots & \hat{V}_{2(p+3)}\\ \vdots & \vdots & \ddots & \vdots\\ \hat{V}_{(p+3)1} & \hat{V}_{(p+3)2} & \cdots & \hat{V}_{(p+3)(p+3)}\end{pmatrix}\right], \tag{2.26}$$

where, $\hat{V}_{ij} = V_{ij} \mid_{\theta=\hat{\theta}}$ is determined by the inverse of the Hessian matrix.

Therefore, the approximate $100(1-\alpha)\%$ two sided confidence intervals for γ_j, β and λ are, respectively, given by

$$\hat{\gamma}_j \pm z_{\alpha/2}\sqrt{\hat{V}_{(j+1),(j+1)}}, \qquad \hat{\beta} \pm z_{\alpha/2}\sqrt{\hat{V}_{(p+2),(p+2)}} \qquad \text{and} \qquad \hat{\lambda} \pm z_{\alpha/2}\sqrt{\hat{V}_{(p+3),(p+3)}},$$

where $j = 0, \ldots, p$, z_α is the upper α^{th} percentile of the standard normal distribution. The needed Hessian matrix is shown in Appendix B.

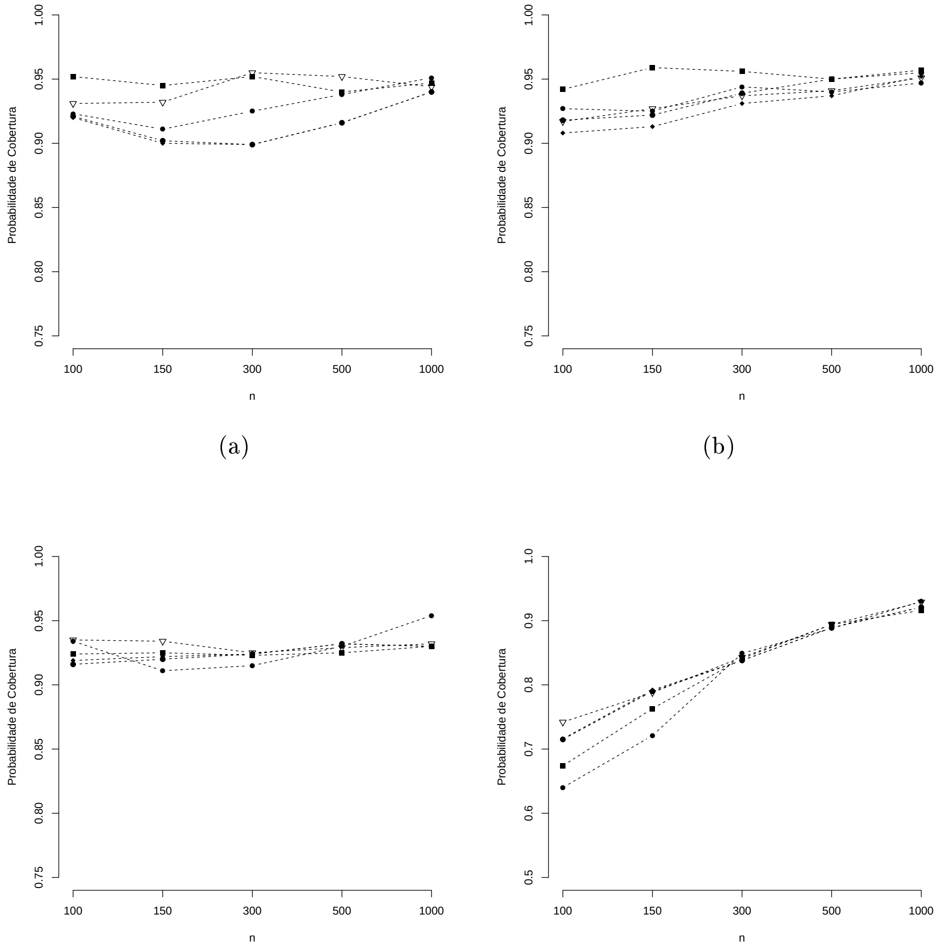
Different models can be compared by penalizing over-fitting by using the Akaike information criterion (Akaike, 1973), which intends to minimize the Kullback-Leibler divergence between the true distribution and the estimate from a candidate model. It is given by AIC $= -2l(\hat{\boldsymbol{\theta}}) + 2size(\boldsymbol{\theta})$, where $l(\hat{\boldsymbol{\theta}})$ denotes the log likelihood function evaluated at the maximum and $size(\boldsymbol{\theta})$ is the number of model parameters. The model with the lowest value of this criterion (among all considered models) is regarded as the preferred model for describing the given dataset.

2.4 Simulation study

This section presents the results of a Monte Carlo experiment on the finite sample behavior of the MLEs. For that, we generated according to a transmuted log-logistic regression distribution in the presence of two covariates. All results were obtained from 1000 Monte Carlo replications and the results are summarized in two tables, presented in Appendix B. Table B.1 shows the relative difference of generated and estimated parameters and their respective standard errors over the 1000 MLEs, which decreased as the sample size increased. Table B.2 shows the coverage probability of 95% two sided confidence intervals for the model parameters, which are close to the nominal coverage for large sample sizes, though they usually differ less than 5% from the nominal coverage probability for the smallest sample size considered. In order to summarize the results of the simulation study, Figure 2.5 presents the coverage probability for $\gamma_1, \gamma_2, \beta$ and λ, respectively.

2.5 Application: Tabapua cattle breed data

Economic results related to beef cattle are directly related to their genetic prepotency, aiming at increasing the production of kilograms of meat per hectare, at a certain time and at less cost. Therefore, the sexual precocity is an important factor in the choice of cattle breed. According to the literature, heifers from temperate regions have sexual precocity, having their first calving when they are little more than two years old, half the time



(a) (b)

(c) (d)

Figure 2.5: *Coverage probability for $\gamma_1, \gamma_2, \beta$ and λ, respectively, by considering the nominal value of 95%.*

observed in heifers from tropical regions, which have their first calving when they are approximately 4 years old on average (Pereira, 2000).

In this context, a long-term study conducted on the Tabapua cattle breed was held at EMBRAPA, a Brazilian agricultural research institute, in order to infer the time up to the first calving in polled Tabapua cows. The Tabapua breed is strong, rustic and chunky, with excellent mothering ability and potential to gain weight. It originated from a crossover made in the municipality of Tabapua, in the state of São Paulo, Brazil, with crossbred zebu, which showed characteristics of spontaneous polled cattle coming from European temperate regions with high genetic potential for meat production and milk (Paro *et al.*, 2013). The data consist of the time up to first calving, in days, of 17026 animals observed from 1983 to 2007.

This breed is distinguished by meekness, sexual precocity, high milk production, fertility, good meat quality, and adaptability to various regions and climate. However, it is mainly bred for meat purposes because it also has very favorable characteristics for this. It is also widely used in crosses with other breeds, generating hardy animals and good productivity (Pereira, 2000).

According to the characteristics mentioned and especially its sexual precocity, the polled Tabapua breed provides an important economic result related to beef cattle since it aims at increasing production of kilograms of meat per hectare, at a certain time and at less cost. The resulting heifers from tropical regions have their first calving when they are approximately 4 years old, twice the time of heifers from temperate regions (Paro *et al.*, 2013; Pereira, 2000).

In this section, we illustrated the usefulness of the transmuted log-logistic distribution modeling the polled Tabapua breed data. Moreover, we provided a comparison model with the log-logistic, Weibull, and exponential distributions.

Firstly, a brief descriptive analysis was made. The minimum observed time was 721 days, or approximately 24 months; the maximum observed was 6184 days or 206 months. The median time for the first calving is 1140 days (38 months) and even the first and third quantiles are 1068 and 1365 days, respectively. Figure 2.6-a shows the TTT plot (for example see Barlow and Campo (1975)), in order to verify the possible shape for the hazard function. If the TTT plot is concave, convex and then concave again, it indicates unimodal hazard, which is our case.

The transmuted log-logistic, log-logistic, Weibull and exponential distributions were fitted to the data. Table 2.1 provides the MLEs, their corresponding standard errors and 95% confidence intervals. For these models, the computed $-2\log$ likelihood and AIC are shown in Table 2.2. Both criteria provide evidence in favor of the transmuted log-logistic distribution. This result is corroborated by the fitted distribution under the data histogram in Figure 2.6-b and by the fitted survival functions under the Kaplan-Meier estimator, Figure 2.6-c.

Table 2.1: *MLEs for the parameters of the transmuted log-logistic, log-logistic, Weibull and exponential distributions.*

Model	Parameter	Estimate	Standard Error	95% Confidence Interval Lower	Upper
Transmuted	α	-17.9079	0.1369	-18.1763	-17.6395
Log-Logistic	β	3.0503	0.0220	3.0073	3.0933
	λ	-0.8139	0.0120	-0.8374	-0.7904
Log-Logistic	α	-16.7482	0.1003	-16.9448	-16.5516
	β	2.5832	0.0160	2.5519	2.6145
Weibull	μ	615.7300	3.1161	609.6300	621.8400
	β	1.6078	0.0082	1.5916	1.6239
Exponential	μ	547.2500	4.1940	539.0300	555.4700

After selecting the most appropriate model for describing the time up to first calving (the transmuted log-logistic distribution), our interest turned to estimating the most probable time the first calving would occur, defined as the $T_{\max}$ according to (2.18). The $\hat{T}_{\max}$ is equal to 1267 days (42.2 months). Its 95% confidence interval is given by $\mathrm{IC}[T_{\max}, 95\%] = (1182.05; 1370.83)$ days. Figure 2.6-d shows the hazard estimate curve, with the $\hat{T}_{\max}$ and the $T_{\max}$ 95% confidence interval. Furthermore, the median time up to first calving is equal to 1173.42 days. Therefore, half of the cows experiment the event of interest up to approximately 39 months and the mean time is 1259.63 days (or approximately 42 weeks), with a standard deviation of 398.33 days.

Table 2.2: *Computed* $-2\log$ *likelihood and AIC values for the log-logistic, transmuted log-logistic, Weibull and Exponential distributions.*

Model	Criteria	
	$-2\log$ likelihood	AIC
Transmuted	236201	236207
Log-Logistic	237996	238000
Weibull	242588	242592
Exponential	248747	248749

From the practical point of view, the use of a misspecificated distribution to represent the random behavior of the time up to first calving would, of course, lead to a misleading conclusion concerning the parameters of interest. For the sake of argument, by considering the log-logistic distribution instead of the transmuted log-logistic one, the $\hat{T}_{\max}$ is equal to 1501.62 days (approximately 50.0 months); almost 8 months longer than the $\hat{T}_{\max}$ obtained by considering the transmuted log-logistic one. This is another drawback that corroborates in favor of the proposed distribution.

2.5.1 Including covariates in the model

In this section, we used a random sample of 500 animals observed and two covariates: the time the calf was born, before or after 2000 (period) and the age the first oestrus occurred (prp), up to or after one year of age.

By using the nonparametric Log-Rank test to compare the survival distributions of two samples, we observed p-values < 0.0001 and 0.0248, respectively, for covariates prp and period (see the survival curves in Figure 2.10). These p-values showed us the significance of these covariates to describe the time of the first calf. Another nonparametric test was carried out (Peto-Peto test; see for example Lawless (2011)), showing almost the same

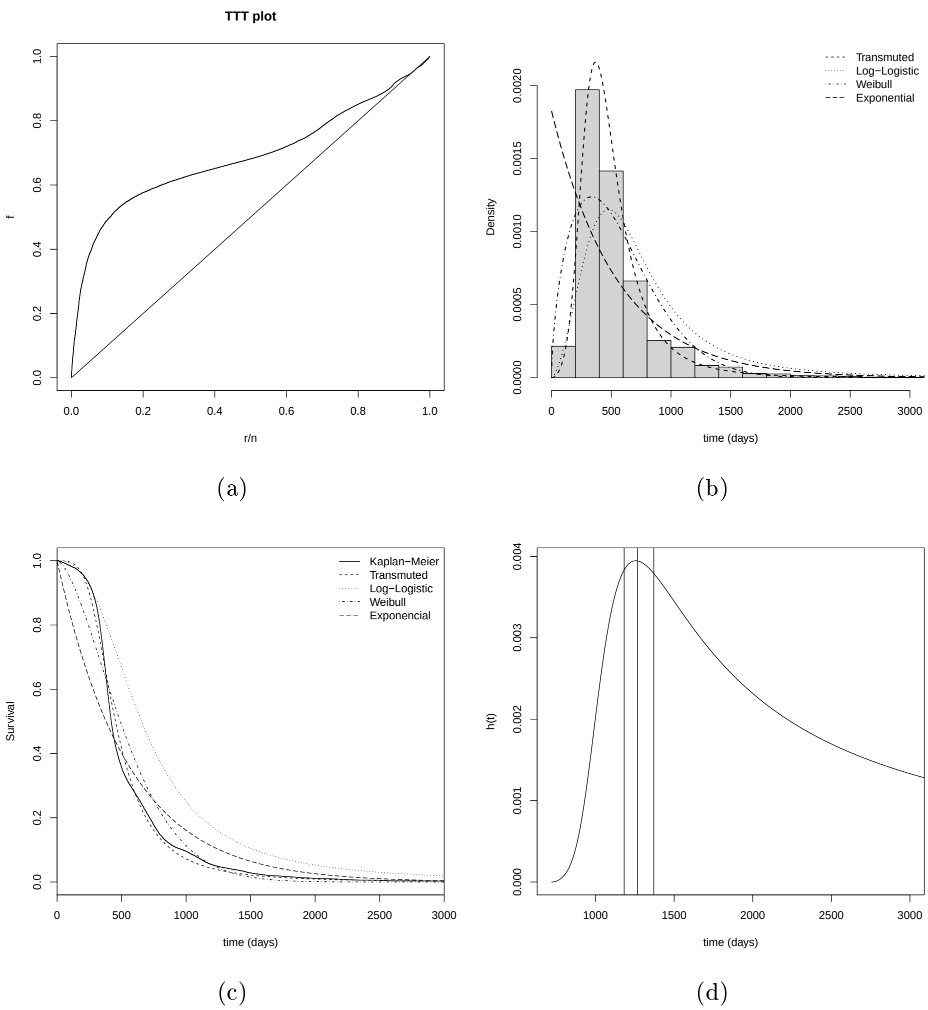


Figure 2.6: *(a) TTT Plot, (b) histogram, (c) reliability curves and (d) hazard estimate curve, with the $\hat{T}_{\max}$ and the $T_{\max}$ 95% confidence interval.*

results (p-values < 0.0001 and 0.0006, respectively).

In order to verify the behavior of the hazard function, Figures 2.7 show the TTT plots according to two covariates: the time the calf was born (after or before 2000) (period) and the age the first oestrus occurred (prp), up to or after one year. Interested readers can refer to Barlow and Campo (1975) for more information on TTT plotting. Overall, if the TTT plot is concave, convex and then concave again, it indicates an unimodal hazard, which is our case.

Then, the transmuted log-logistic and the simple log-logistic distributions were fitted to the data. Table 2.3 provides the MLEs, their corresponding standard errors and 95% confidence intervals, as well as the computed $-2\log$ likelihood $(-2l)$ and the AIC. Both criteria provide evidence in favor of the transmuted log-logistic regression distribution. Observe that, as the transmutation map presented skewness and the log-logistic distribution

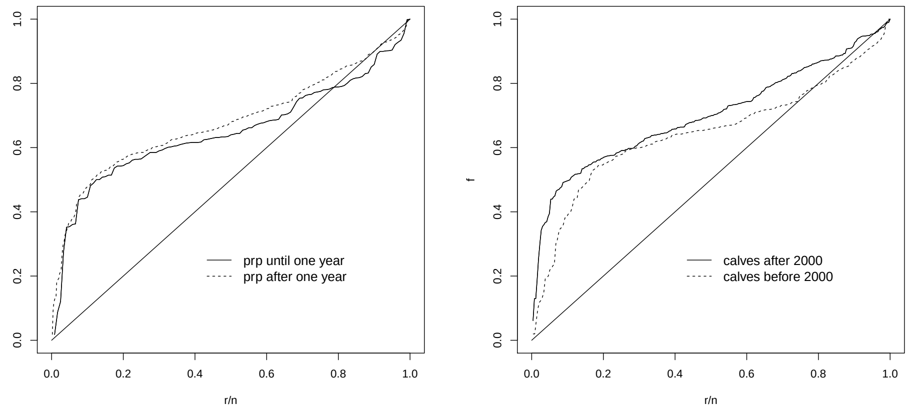


Figure 2.7: *TTT Plot by considering the covariates: prp (left panel) and period (right panel).*

is an asymmetric model, the mode point is higher, which is totally interpreted in our application. In general, the crossbreeds are at similar periods, increasing the chance of a first calf around a certain period of time and, thus, increasing the mode point.

Table 2.3: *MLEs considering the log-logistic and transmuted log-logistic regression model.*

Model	Parameter	Estimate	Standard Error	IC 95% Lower	IC 95% Upper	$-2l$	AIC
	γ_0	−17.092	0.801	−18.667	−17.860	6925.3	6935.3
Transmuted	γ_1	−0.364	0.137	−0.632	−0.095		
	γ_2	3.281	0.794	1.722	4.840		
	β	2.954	0.128	2.703	3.205		
	λ	−0.764	0.084	−0.930	−0.598		
	γ_0	−19.319	0.742	−20.777	−17.860	6936.8	6944.8
Log-Logistic	γ_1	−0.419	0.154	−0.722	−0.117		
	γ_2	3.213	0.809	1.623	4.804		
	β	3.208	0.122	2.968	3.448		

Furthermore, according to Burnham and Anderson (2002), we verified that if the MLE values are influenced by a specific experimental observation, they can be regarded as an outlier. We considered the one-leaved-out approach on the basis of a cross validation scheme, in the sense that the parameters of the transmuted log-logistic regression were re-estimated 500 times. Each time, one specific observed time was withdrawn. We also estimated the standard errors, confidence intervals and the $-2\log$ likelihood and AIC criteria.

Moreover, the differences, Δ_i, given by $\Delta_i = AIC_i - AIC_{\min}$, were calculated. After, the Akaike weights given by $\omega_i = \exp(-\Delta_i/2)$ were obtained and plotted as shown on the left panel of Figure 2.8. The presence of outliers is clear. We decide to remove all the experimental points with $\omega_i > 0.20$, corresponding to observations 15, 143, 231, 242, 289 and 360. The Akaike weights in the sample, without the outliers, are shown on the right panel of Figure 2.8.

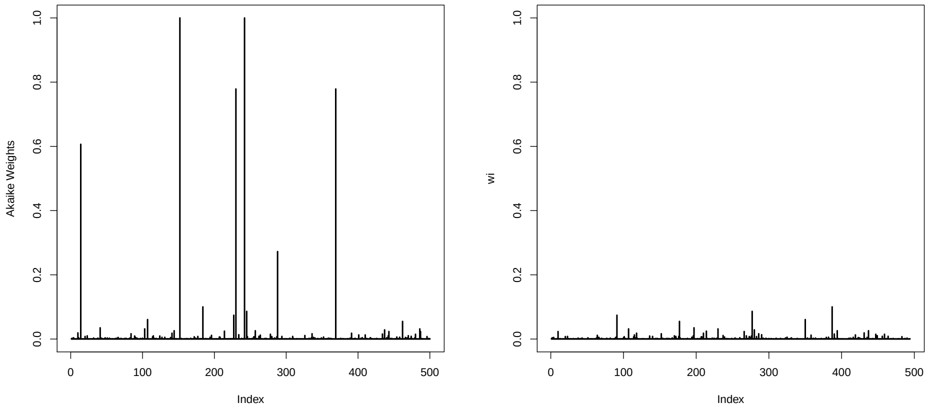


Figure 2.8: *Cross Validation results: Akaike weights for the complete sample (left panel) and without influential points (right panel).*

Table 2.4 shows the MLEs for the parameters of the transmuted log-logistic regression model in the sample without the outliers. There seem to be no important changes in the MLEs, althought the $-2l$ and AIC values are much smaller than those obtained when the complete sample is considered.

Table 2.4: *MLEs, standard error and the 95% confidence interval considering the transmuted log-logistic regression model in the sample without outliers.*

Model	Parameter	Estimate	Standard Error	IC 95% Lower	IC 95% Upper	$-2l$	AIC
	γ_0	-20.611	1.065	-20.704	-18.518	6366.9	6376.9
Transmuted	γ_1	-0.430	0.145	-0.552	-0.145		
without	γ_2	3.646	0.913	1.851	5.440		
outliers	β	3.549	0.168	3.218	3.880		
	λ	-0.710	0.118	-0.943	-0.478		

Moreover, as a goodness-of-fit procedure, we performed a residual analysis for both models (transmuted log-logistic and simple log-logistic regression ones) using the Cox-Snell residuals (Cox and Snell, 1968). The Cox-Snell residuals are defined as $\hat{e}_i = \hat{\Lambda}(t_i \mid x_i)$, where $\hat{\Lambda}(.)$ is the cumulative hazard function of the adjusted model. If we consider the log-logistic and transmuted log-logistic models, the Cox-Snell residuals are, respectively, given by

$$\hat{e}_{iLL} = \ln\left[1 + \exp(\gamma(\boldsymbol{x}))t_i^{\beta}\right] \tag{2.27}$$

and

$$\hat{e}_{iTLL} = 2\ln\left[1 + \exp(\gamma(\boldsymbol{x}))t_i^{\beta}\right] - \ln\left[1 + \exp(\gamma(\boldsymbol{x}))t_i^{\beta} - \lambda\exp(\gamma(\boldsymbol{x}))t_i^{\beta}\right], \tag{2.28}$$

where $\gamma(\boldsymbol{x}) = \gamma_0 + \gamma_1 x_1 + \ldots + \gamma_p x_p$.

Figure 2.9 shows the estimated residuals versus the estimated empirical survival for the residuals. The criteria provide clear evidence in favor of the transmuted log-logistic regression model.

Figure 2.10 shows the estimated hazard curve and the estimated most probable time for the first calf ($T_{\max}$) according to both covariates. The $\hat{T}_{\max}$ is equal to 29.52 months with a 95% confidence interval equal to $(26.56, 32.49)$ if the prp occurs before the first year. For calves born until the year 2000, the $\hat{T}_{\max}$ is equal to 42.97 months with a 95% confidence interval equal to $(41.43, 44.50)$. For calves born after the year 2000 or prp occurring after the first year, the $\hat{T}_{\max}$ is equal to 40.77 months with a 95% confidence interval equal to $(39.35, 42.18)$.

Considering the period prior to the year 2000, we observed an older age at the first calving, occurring close to 4 years. This time decreased to less than 3.5 years showing that the current reproductive techniques of livestock are more efficient than in the previous period. As the reproductive life begins earlier, cows have an increase number of conceptions and consequently increased their reproductive life by reducing the number of cull cows.

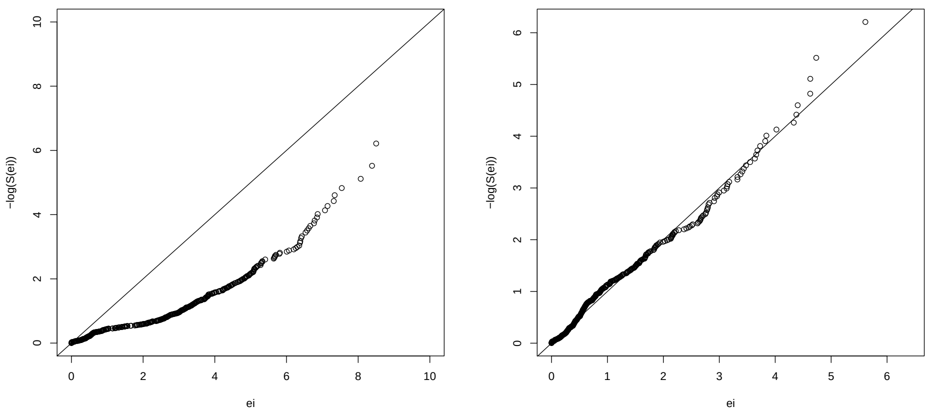


Figure 2.9: *Estimated residuals versus the estimated empirical survival for the residuals: log-logistic (left panel), transmuted log-logistic distributions (right panel).*

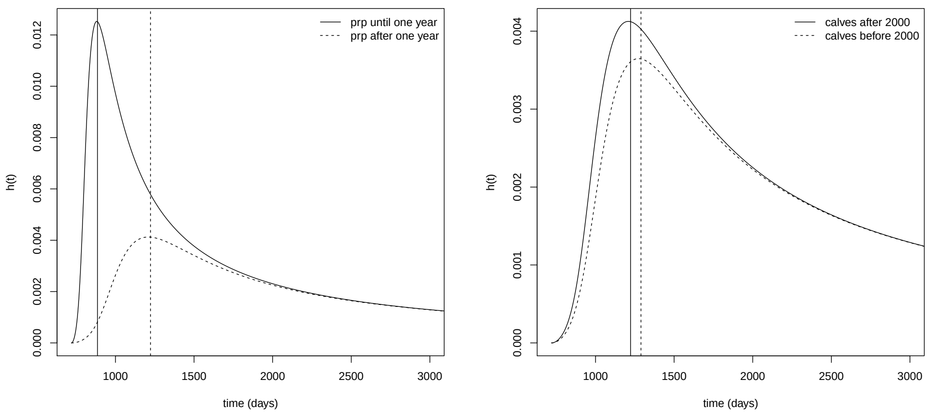


Figure 2.10: *Hazard estimate curve, with the* $\hat{T}_{\max}$.

Chapter 3

Bayesian and profile analysis for the transmuted log-logistic model

In this Chapter, we present the transmuted log-logistic model in a Bayesian context. For that, we used the hierarchical structure using the half-Cauchy prior proposed by Gelman (2006). Furthermore, we presented the transmuted log-logistic model in the presence of the censored lifetime and the profile method was used in the estimation process.

3.1 Hierarchical transmuted log-logistic model

According to Chen and Ibrahim (2006), one of most common ways of combining several sources of information is though hierarchical modeling. Thus, the authors show us the relationship between the power prior and hierarchical models using regression models as an example.

Moreover, Gelman (2006) points out that several studies using multilevel models are central to modern Bayesian statistics for both conceptual and practical reasons. The authors suggested the use of half-t family as a prior distribution for variance parameters such as the half-Cauchy distribution, which is a special conditionally-conjugate folded-non central-t family case of prior distributions for parameters representing the discrepancy. Even though several studies use the half-Cauchy prior for scale parameters (see for example Polson and Scott (2012)), Gelman (2006) used this prior not for scale, but for variance parameters and showed serious problems with the inverse-Gamma prior, the most commonly used prior for variance component, Daniels (1999).

In this study, we proposed to use the hierarchical models in two levels; to do so, suppose the hierarchical model is given as $[X|\mu,\beta,\lambda] \sim f(x|\mu,\beta,\lambda)$, $\mu|\sigma^2 \sim \pi_\mu(\mu|\sigma^2)$, $\beta|\theta \sim \pi_\beta(\beta|\theta)$, $\lambda \sim \pi_\lambda(\lambda)$, $\sigma^2 \sim \psi_\sigma(\sigma^2)$ and $\theta \sim \psi_\theta(\theta)$. The posterior distribution can be constructed as follows.

Proposition 3.1.1 *Let us suppose that, in the first stage, we considered a class* Γ *of priors* *that led to the following*

$$\begin{aligned}\Gamma &= \big\{\pi(\mu,\beta,\lambda|\sigma^2,\theta) : \pi(\mu,\beta,\lambda|\sigma^2,\theta) = \pi_\mu(\mu|\sigma^2)\pi_\beta(\beta|\theta)\pi_\lambda(\lambda)\big| \\ &\quad \pi_\mu \text{ being } \mathrm{N}(\tau,\sigma^2), (\tau,\sigma^2)\in\mathbb{R}\times\mathbb{R}^+; \quad \pi_\beta \text{ being } \mathrm{HC}(\theta), \theta\in\mathbb{R}^+; \\ &\quad \pi_\lambda \text{ being } \mathrm{U}(a,b), (a,b)\in\mathbb{R}\times\mathbb{R}, a<b\big\}.\end{aligned}$$

Moreover, in the second stage (sometimes called a hyperprior), that would consist of putting a prior distribution $\psi_k(\cdot)$ *on the hyperparameters* σ^2 *and* θ *where*

$$\begin{aligned}\Psi &= \big\{\psi(\sigma^2,\theta) : \psi(\sigma^2,\theta) = \psi_{\sigma^2}(\sigma^2)\psi_\theta(\theta); \quad \psi_{\sigma^2} \text{ is Gamma}(\alpha,\zeta), \\ &\quad (\alpha,\zeta)\in\mathbb{R}^+\times\mathbb{R}^+; \quad \psi_\beta \text{ is Gamma}(\eta,\vartheta), (\eta,\vartheta)\in\mathbb{R}^+\times\mathbb{R}^+; \\ &\quad \alpha,\zeta,\eta,\vartheta \text{ are known and do not depend on any other hyperparameter}\big\}.\end{aligned}$$

Thus, the hierarchical log-logistic posterior distribution is written as

$$\begin{aligned}\pi(\mu,\beta,\lambda|\boldsymbol{x}) &\propto \frac{e^\mu\beta\theta}{\sigma}\left[x^{\beta+\alpha+\eta-3}\frac{\left[(1+e^\mu x^\beta)-\lambda(e^\mu x^\beta-1)\right]}{(1+e^\mu x^\beta)^3}\right] \\ &\quad \times\exp\left\{-\left(\zeta+\vartheta+\frac{x}{2\sigma^2}\right)\right\}. \end{aligned} \quad (3.1)$$

Proof: The demonstration is direct.

Note that the β parameter is supposed to be a half-Cauchy distribution whose probability density function is given by

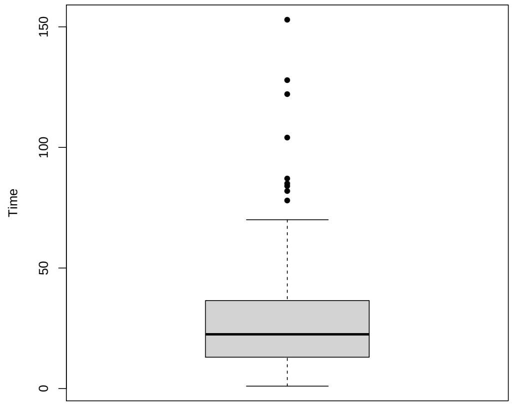
$$f(x) = \frac{2\theta}{\pi\left(x^2+\theta^2\right)}, \qquad x>0, \theta>0, \quad (3.2)$$

where θ is a scale parameter which has a broad peak at zero and, in limit, $\theta\to\infty$ this becomes a uniform prior density. However, large finite values for θ represent prior distributions we call "weakly informative". For example, Gelman (2006) shows us that, for $\theta = 25$, the half-Cauchy is nearly flat, although not completely.

3.2 Tabapua cattle breed data

In this section, we showed the usefulness of the hierarchical transmuted log-logistic model on modeling of the polled Tabapua breed data using Bayesian methods.

The data consist of the time up to first calving, in days, of 17026 animals observed from 1983 to 2007, which was presented in Chapter 2. Figure 3.1-b shows the distribution of the time when the first calving occurred. The median time for the first calving was 1140 days (38 months) and the first and third quantiles are 1068 and 1365 days, respectively.



(a)

Figure 3.1: *Boxplot of times.*

After an initial analysis, the hierarchical transmuted log-logistic was fitted to the data, as we specified in Section 3.1, The posterior samples were generated by the Metropolis-Hastings technique. Three chains of the dimension $100,000$ were considered for each parameter, discarding the first $10,000$ iterations in order to eliminate the effect of the initial values and to avoid correlation problems. A lag size 10 was used, resulting in a final sample size of $10,000$. Tables 3.1 and 3.2 show the posterior summaries for the parameters and the 95% credible intervals considering the mentioned priors.

Table 3.1: *Posterior model summary of the hierarchical transmuted log-logistic model parameters.*

Parameter	Mean	Standard Deviation	Percentiles		
			25%	50%	75%
α	-17.865	0.138	-17.958	-17.871	-17.775
β	3.043	0.022	3.029	3.044	3.058
λ	-0.815	0.012	-0.823	-0.815	-0.807
σ^2	900.100	822.500	333.800	640.400	1208.100
θ	198.800	195.100	60.171	139.800	273.400

The chain convergence was verified by the Gelman and Rubin's convergence diagnostic criterion, see for example Gelman and Rubin (1992), which demonstrated that these criteria are satisfied (Table 3.3). The convergence can also be seen in Figures 3.2-a to j.

Furthermore, the marginal posterior densities for μ, β and λ, respectively, can be analyzed in Figures 3.3-a to e.

After estimating and analyzing the model convergence, Figure 3.4-a and b show, respectively, the hazard estimate curve, with the $\hat{T}_{\max}$ and the $T_{\max}$ 95% credible interval;

Table 3.2: *95% Credible Interval of parameters estimated.*

Parameter	Equal-Tail Interval		HPD Interval	
α	−18.123	−17.576	−18.118	−17.571
β	2.997	3.085	2.997	3.084
λ	−0.838	−0.791	−0.838	−0.790
σ^2	96.474	3044.600	37.837	2516.300
θ	5.395	733.800	0.008	594.800

Table 3.3: *Gelman and Rubin's criterion to verify the parameters convergence of the hierarchical transmuted log-logistic distribution.*

Parameter	Estimate	Upper Bound
μ	1.0085	1.0060
β	1.0082	1.0057
λ	1.0020	1.0017
σ^2	1.0016	1.0019
θ	1.0004	1.0009

the estimated vs empirical survival curves and the histograms allow us to see how well it fits a set of observations.

Considering the hierarchical transmuted log-logistic fitting, the $\hat{T}_{\max}$ is equal to 1246.77 days (41.56 months) and its 95% credible interval is given by
$\text{IC}[T_{\max}, 95\%] = (1160.04; 1352.86)$ days (see Figure 3.4-a). Furthermore, the median time up to first calving is equal to 1152.48 days (or approximately 38.42 months), and the mean time is 1240.13 days (or approximately 41.34 months), with a standard deviation equal to 13.34 months.

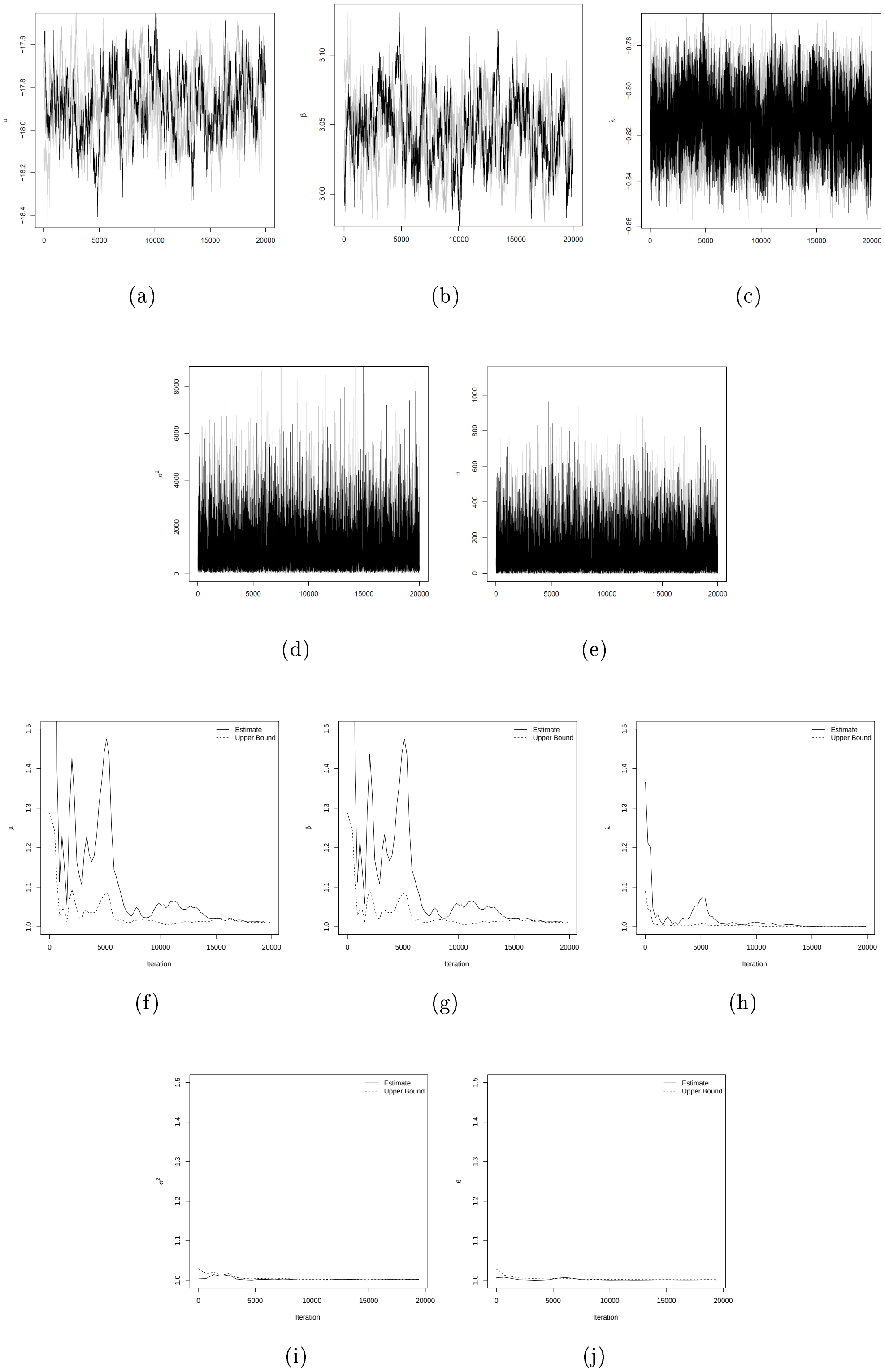
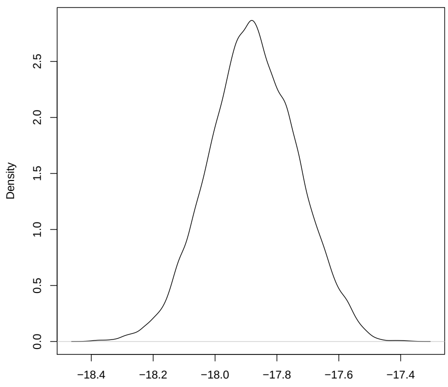
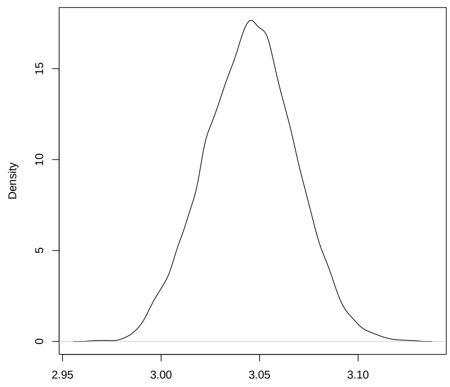


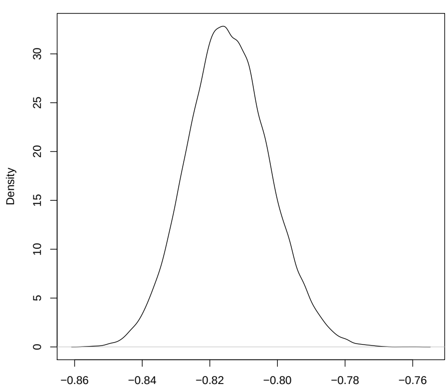
Figure 3.2: *Traceplots and convergence plots, respectively, for: (a,f): μ; (b,g): β; (c,h): λ; (d,i): σ^2 and; (e,j): θ*



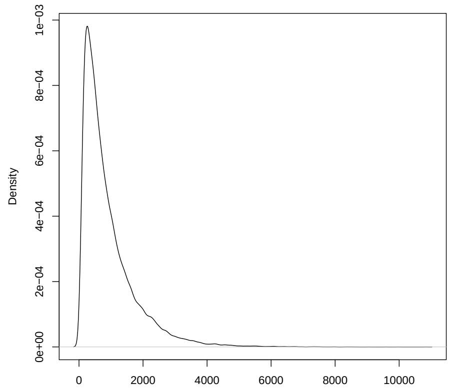
(a)



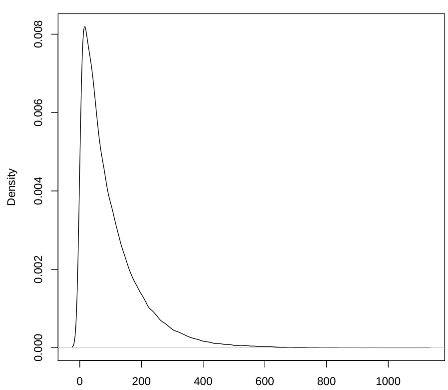
(b)



(c)

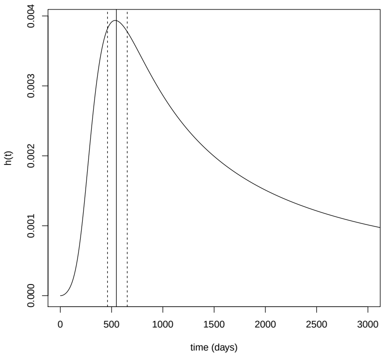


(d)

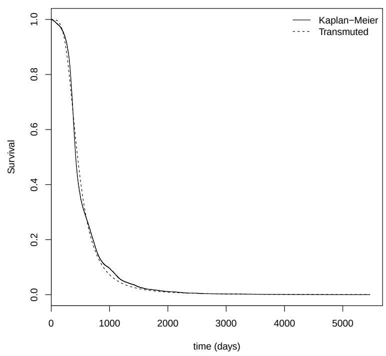


(e)

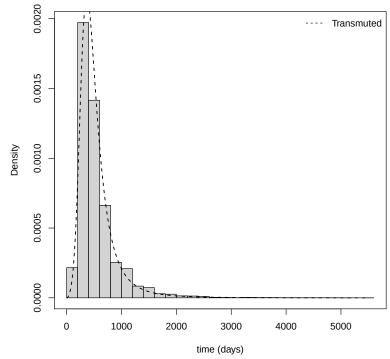
Figure 3.3: *Marginal posterior densities for: (a) μ, (b) β, (c) λ, (d) σ^2 and (e) θ*



(a)



(b)



(c)

Figure 3.4: *(a) hazard estimate curve, with the* $\hat{T}_{\max} - 700$ *and the* $T_{\max} - 700$ *95% credible interval, (b) survival curves and (c) histogram.*

3.2.1 Influence analysis

In this section, we will make an analysis of global influence for the data set given, using the transmuted log-logistical model in a Bayesian context.

The first tool to assess the sensitivity analysis is measuring the global influence. We started with the case-deletion, in which we study the effect of withdrawing the ith element sampled. The first measure of global influence analysis is known as the generalized Cook's distance, which is defined as the standard norm of $\boldsymbol{\zeta}_i = (\alpha_i, \beta_i, \lambda_i, \sigma_i^2, \theta_i)$ and $\hat{\boldsymbol{\zeta}} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda}, \hat{\sigma}^2, \hat{\theta})$ and is given by

$$CD_i(\boldsymbol{\zeta}) = \left[\ \boldsymbol{\zeta}_i - \hat{\boldsymbol{\zeta}}\ \right]^T \left[-\ddot{L}(\boldsymbol{\zeta})\right] \left[\ \boldsymbol{\zeta}_i - \hat{\boldsymbol{\zeta}}\ \right] \tag{3.3}$$

where $\ddot{L}(\boldsymbol{\zeta})$ can be approximated by the estimated covariance and variance matrix. Another way to measure the global influence is by the difference in likelihoods given by

$$LD_i(\boldsymbol{\zeta}) = 2\left\{l(\hat{\boldsymbol{\zeta}}) - l(\boldsymbol{\zeta}_i)\right\}. \tag{3.4}$$

Figures 3.5 show the likelihood distances where some possible influence points can be observed.

In order to reveal the impact of these, the relative changes were measured as

$$\mathrm{RC}_{\zeta_j} = \left|\frac{\hat{\zeta}_j - \hat{\zeta}_{j(I)}}{\hat{\zeta}_j}\right| \times 100\%, \qquad j = 1, \ldots, p+1$$

where $\hat{\zeta}_{j(I)}$ denotes the MLE of ζ_j after set I of observations was removed. As suggested by Lee *et al.* (2006), we use the total and maximum relative changes and the likelihood displacement.

To reveal the impact of the detected influential observations, we used three measures as defined by Lee *et al.* (2006),

$$\mathrm{TRC} = \sum_{i=1}^{n_p}\left|\frac{\hat{\boldsymbol{\zeta}}_i - \hat{\boldsymbol{\zeta}}_i^0}{\hat{\boldsymbol{\zeta}}_i}\right|, \quad \mathrm{MRC} = \max_i \left|\frac{\hat{\boldsymbol{\zeta}}_i - \hat{\boldsymbol{\zeta}}_i^0}{\hat{\boldsymbol{\zeta}}_i}\right| \quad \text{and} \quad LD_{(l)}(\boldsymbol{\zeta}) = 2\{l(\hat{\boldsymbol{\zeta}} - l(\hat{\boldsymbol{\zeta}}^0)\}$$

where TRC is the total relative changes, MRC the maximum relative changes and LD the likelihood displacement, with n_p (the number of parameters) and $\hat{\boldsymbol{\zeta}}^0$ denoting MLE of $\boldsymbol{\zeta}$ after set I of observations was removed.

In order to analyze these influential points, we withdrew the identified points in Figure 3.5. Moreover, a percentage of points (0.01% to 5%) that stood out was also withdrawn. The results can be seen in Table 3.4.

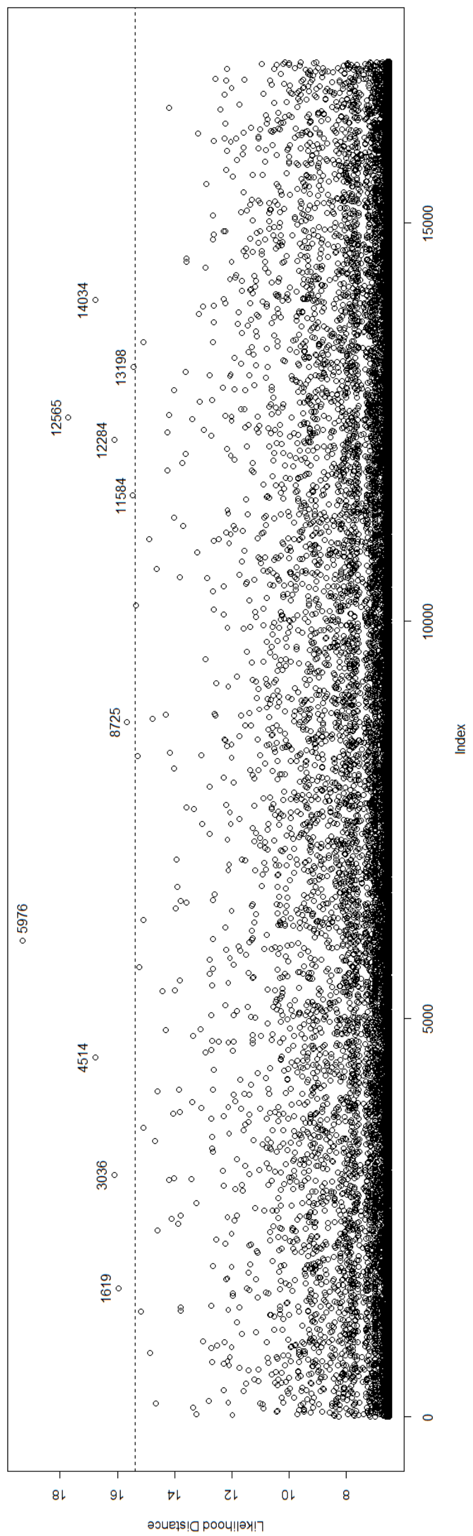


Figure 3.5: *Likelihood distance.*

Table 3.4: *RC (in %) and the corresponding TRC, MRC and* $LD_{(I)}$.

Removed Case	Parameter	RC	TRC	MRC	$LD_{(I)}$
	μ	8.271	86.438	68.336	142
Identify	β	5.142			
Points	λ	68.336			
	σ_2	2.877			
	θ	1.811			
	μ	0.844	3.598	1.358	217
	β	0.818			
0.1%	λ	0.123			
	σ_2	0.456			
	θ	1.358			
	μ	2.098	8.634	3.219	949
	β	2.073			
0.5%	λ	0.098			
	σ_2	1.144			
	θ	3.219			
	μ	3.257	9.782	3.257	1821
	β	3.250			
1%	λ	0.368			
	σ_2	1.700			
	θ	1.207			
	μ	5.843	17.064	5.855	3519
	β	5.855			
2%	λ	0.393			
	σ_2	2.911			
	θ	2.062			
	μ	8.096	22.086	8.162	5178
	β	8.162			
3%	λ	1.460			
	σ_2	3.111			
	θ	1.258			
	μ	10.458	28.558	10.547	6775
	β	10.547			
4%	λ	0.785			
	σ_2	6.566			
	θ	0.201			
	μ	12.463	34.232	12.627	8383

	β	12.627
5%	λ	2.160
	σ_2	6.277
	θ	0.704

Notice that when we withdrew the 10 most influential points, λ parameter was the most affected, $RC \simeq 68\%$. Then, the hierarchical transmuted log-logistic was re-fitted to the data. The posterior samples were generated by the Metropolis-Hastings technique with one chain of the dimension 100000, discarding the first 10000 iterations in order to eliminate the effect of the initial values. To avoid the correlation problems, a lag size 10 was used resulting in a final sample size of 10000. Tables 3.5 and 3.6 show the posterior summaries for the parameters and the 95% credible intervals considering the mentioned priors for all removed cases, respectively.

Table 3.5: *Posterior model summary of the hierarchical transmuted log-logistic model parameters.*

Removed Case	Parameter	Mean	Standard Deviation	Percentiles		
				25%	50%	75%
	α	−19.343	1.269	−20.522	−20.327	−17.975
Identify	β	3.200	0.130	3.062	3.285	3.319
Points	λ	−0.258	0.519	−0.813	0.153	0.232
	σ^2	926.00	840.30	348.70	664.10	1224.00
	θ	202.40	204.70	59.08	139.20	282.00
	α	−18.016	0.129	−18.098	−18.015	−17.933
	β	3.068	0.021	3.055	3.068	3.082
0.1%	λ	−0.814	0.012	−0.822	−0.814	−0.806
	σ^2	904.20	826.40	334.50	651.90	1207.20
	θ	196.10	194.80	56.84	137.20	270.50
	α	−18.240	0.135	−18.336	−18.238	−18.146
	β	3.107	0.022	3.091	3.106	3.122
0.5%	λ	−0.816	0.012	−0.824	−0.816	−0.808
	σ^2	910.40	842.80	339.00	643.80	1203.30
	θ	205.20	203.70	61.83	142.40	284.10
	α	−18.447	0.131	−18.536	−18.441	−18.352
	β	3.142	0.021	3.127	3.142	3.157
1%	λ	−0.818	0.012	−0.826	−0.818	−0.810
	σ^2	915.40	853.70	337.50	647.80	1221.70
	θ	196.40	195.20	56.19	136.90	273.10
	α	−18.909	0.144	−19.007	−18.907	−18.811
	β	3.222	0.023	3.206	3.221	3.238

2%	λ	−0.818	0.013	−0.827	−0.818	−0.810
	σ^2	926.30	826.60	353.60	667.40	1235.00
	θ	202.90	199.40	60.53	143.40	279.70
	α	−19.312	0.142	−19.410	−19.304	−19.222
	β	3.292	0.023	3.277	3.291	3.308
3%	λ	−0.827	0.012	−0.835	−0.827	−0.819
	σ^2	928.10	849.10	352.20	668.20	1215.20
	θ	196.30	193.30	56.37	139.00	275.50
	α	−19.734	0.154	−19.834	−19.737	−19.641
	β	3.364	0.025	3.350	3.365	3.381
4%	λ	−0.821	0.013	−0.830	−0.821	−0.813
	σ^2	959.20	858.50	365.20	698.70	1267.90
	θ	198.40	195.60	58.00	137.50	277.00
	α	−20.092	0.175	−20.221	−20.093	−19.967
	β	3.428	0.028	3.407	3.428	3.448
5%	λ	−0.832	0.013	−0.842	−0.833	−0.824
	σ^2	956.60	854.00	372.90	704.10	1259.10
	θ	200.20	198.30	59.91	141.00	273.20

It could be clearly seen that parameter λ was estimated in the first moment -0.838. It changed abruptly only when the identified points were removed, remaining close to -0.838 in all other cases, see in Tables 3.1 and 3.5.

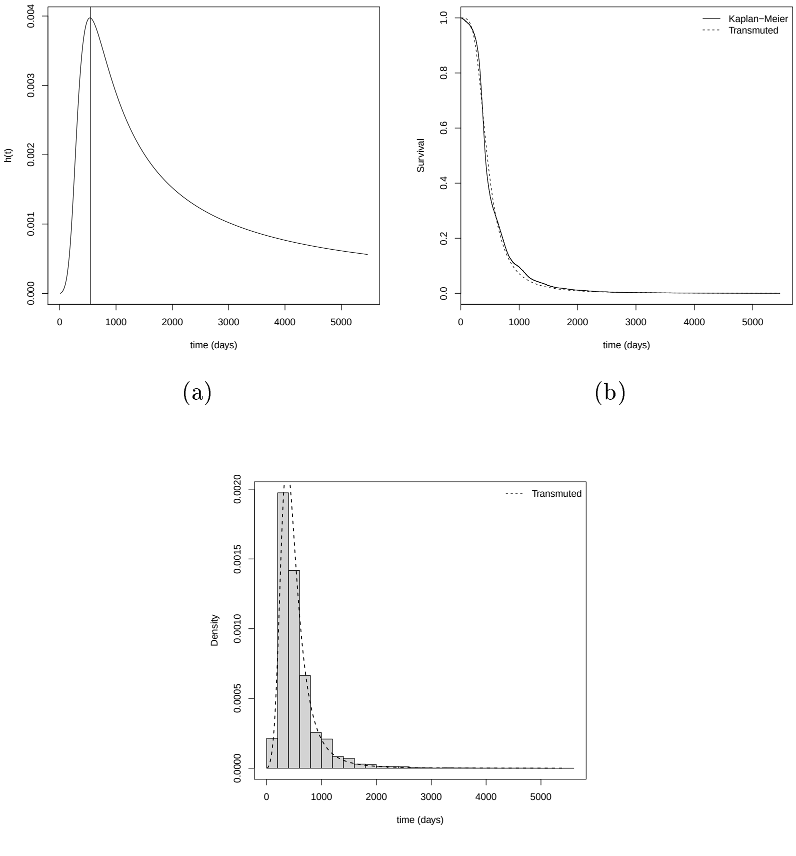
Furthermore, when we withdrew 0.1% from the sample, i.e., just 17 observations, we did not lose much information and we improved the fitted model. Figures 3.6-a, b and c show the fitted model, when using this number of removed cases.

Finally, considering the hierarchical transmuted log-logistic fitting, the $\hat{T}_{\max}$ changes to 1267.71 compared to 1266.77 days (42.22 months) and its 95% credible interval is given by

$\mathrm{CI}[T_{\max}, 95\%] = (39.56; 45.47)$ months. Furthermore, the median time up to the first calving is equal to 1172.41 days (or approximately 39.08 months), and the mean time is 1258.54 days (or approximately 41.95 months), with standard deviation equal to 13.11 months.

Table 3.6: *95% Credible Interval of parameters estimated.*

Removed Case	Parameter	Equal-Tail Interval		HPD Interval	
	α	−20.709	−17.754	−20.713	−17.760
Identify	β	3.026	3.351	3.027	3.352
Points	λ	−0.833	0.293	−0.837	0.284
	σ^2	104.300	3096.600	43.226	2603.000
	θ	5.211	746.000	0.143	601.600
	α	−18.2762	−17.7678	−18.2845	−17.7776
	β	3.0289	3.11	3.0283	3.1093
0.1%	λ	−0.8365	−0.7896	−0.8372	−0.7906
	σ^2	94.0797	3151.6	40.6435	2574.8
	θ	5.2045	728.6	0.1098	586.3
	α	−18.4967	−17.9843	−18.4981	−17.9863
	β	3.0654	3.1476	3.0664	3.1483
0.5%	λ	−0.8391	−0.7905	−0.84	−0.7918
	σ^2	98.81	3143.7	28.4125	2605.9
	θ	5.1449	745.1	0.0156	613.7
	α	−18.7121	−18.205	−18.7066	−18.2029
	β	3.103	3.1855	3.1015	3.1829
1%	λ	−0.8407	−0.793	−0.8421	−0.7948
	σ^2	102.2	3172.8	34.7775	2578.2
	θ	4.6184	727.6	0.0235	591.8
	α	−19.1832	−18.6319	−19.184	−18.6338
	β	3.1769	3.266	3.1779	3.2663
2%	λ	−0.8415	−0.7927	−0.8428	−0.7944
	σ^2	106.2	3111.3	39.5096	2606.1
	θ	5.4845	729.9	0.0306	600.2
	α	−19.5917	−19.0343	−19.5817	−19.0259
	β	3.247	3.3364	3.2482	3.3373
3%	λ	−0.8504	−0.8013	−0.851	−0.8023
	σ^2	107.5	3224.1	30.9968	2597.3
	θ	4.8582	710.8	0.0153	578.9
	α	−20.0252	−19.4068	−20.0321	−19.4161
	β	3.312	3.4113	3.3143	3.4124
4%	λ	−0.8458	−0.7955	−0.8459	−0.7958
	σ^2	111	3307.1	50.2358	2671.5
	θ	5.2325	734.3	0.0631	604.8
	α	−20.4078	−19.7676	−20.418	−19.7839
	β	3.3762	3.4789	3.3778	3.4802
5%	λ	−0.8572	−0.8055	−0.8581	−0.8066
	σ^2	119	3273.2	56.2007	2652.8
	θ	5.0109	743.8	0.0992	605.8



(a) (b)

(c)

Figure 3.6: *(a) Hazard estimate curve, with the* $\hat{T}_{\max} - 700$*, (b) survival curves and (c) histogram.*

3.3 Transmuted log-logistic in the presence of censored lifetime

Let us consider the transmuted log-logistic model presented in Chapter 2. Furthermore, consider the n lifetimes observed $t_1, t_2, \ldots, t_n$ and the observed censor indicators $c_1, c_2, \ldots, c_n$, where $c_i = 1$ if t_i is exactly observed, or $c_i = 0$ otherwise. For the right censoring, we are currently considering in this study, we have $c_i = 1$ if $t_i \leq t_{(r)}$, and $c_i = 0$ otherwise. Then, the likelihood function can be written as

$$L_n(\boldsymbol{\theta}|t_1, t_2, \ldots, t_n) = k \prod_{i=1}^{n} f(t_i|\boldsymbol{\theta})^{c_i} \left(1 - F(t_i|\boldsymbol{\theta})\right)^{(1-c_i)},$$

where K is a constant factor not depending on t_i, c_i or any unknown parameters, see for example Lawless (2011). By considering the transmuted log-logistic model, the likelihood function is written as

$$\begin{aligned} L_n(\mu, \beta, \lambda|t_1, \ldots, t_n, c_1, \ldots, c_n) &= (\beta e^{\mu})^{\sum_{i=1}^{n} c_i} \\ \times \prod_{i=1}^{n} &\left\{ t_i^{c_i(\beta-1)} \left[\frac{(1 + e^{\mu} t_i^{\beta}) - \lambda(e^{\mu} t_i^{\beta} - 1)}{(1 + e^{\mu} t_i^{\beta})^3} \right]^{c_i} \cdot \left[\frac{1 + e^{\mu} t_i^{\beta}(1 - \lambda)}{(1 + e^{\mu} t_i^{\beta})^2} \right]^{1-c_i} \right\}. \end{aligned} \tag{3.5}$$

Hence, the log-likelihood function $l_n(\cdot) = \ln L_n(\cdot)$ becomes

$$\begin{aligned} l_n(\mu, \beta, \lambda \quad | \quad t_1, \ldots, t_n, c_1, \ldots, c_n) &= \sum_{i=1}^{n} c_i \ln \beta + \mu \sum_{i=1}^{n} c_i + (\beta - 1) \sum_{i=1}^{n} c_i \ln t_i \\ &+ \sum_{i=1}^{n} c_i \ln \left[(1 + e^{\mu} t_i^{\beta}) - \lambda(e^{\mu} t_i^{\beta} - 1) \right] - 3 \sum_{i=1}^{n} c_i \ln(1 + e^{\mu} t_i^{\beta}) \\ &+ \sum_{i=1}^{n} (1 - c_i) \ln \left[1 + (1 - \lambda) e^{\mu} t_i^{\beta} \right] - 2 \sum_{i=1}^{n} (1 - c_i) \ln(1 + e^{\mu} t_i^{\beta}). \end{aligned}$$

3.3.1 Profile likelihood

Estimating the parameters of a model using the likelihood method can be quite difficult in some special cases. In these cases, computational alternatives, such as the profile likelihood function, are required. They are used to calculate of classical estimates that allow us to estimate parameters by eliminating the numerical problems of the classical likelihood function; see for example Barndorfe-Nielsen and Cox (1994).

Following Venzon and Moolgavkar (1988), suppose there are two sets of parameters, ψ and ϕ, where ψ represents the vector of factor loadings and ϕ represents all the other model parameters (in some cases, ϕ is a nuisance parameter or a vector of nuisance parameters). The profile method estimating parameters (ψ, ϕ) in two stages is given below.

Let us suppose that $X_1, X_2, \dots, X_n$ are iid random variables with density $f(\boldsymbol{x}; \psi, \phi)$. As our objective is to estimate ψ and ϕ, the log-likelihood function is given by

$$l_P(\psi, \phi | x_1, \dots, x_n) = \sum_{i=1}^{n} \log f(\boldsymbol{x}; \psi, \phi). \tag{3.6}$$

Then, to estimate ψ and ϕ the following can be used $(\hat{\psi}_P, \hat{\phi}_P) = \arg\max_{\psi,\phi} l_P(\psi, \phi | x_1, \dots, x_n)$ what can be quite difficult and can lead to expression which are hard to maximize. Instead let us consider a different method where suppose for now, ψ is known, then we rewrite the likelihood as $l_P(\psi, \phi | x_1, \dots, x_n) = l_\psi(\phi | x_1, \dots, x_n)$ (to show that ψ is fixed and ϕ varies). To estimate ϕ, we maximize $l_\psi(\phi | x_1, \dots, x_n)$ with respect to ϕ, ie,

$$\hat{\phi}_\psi = \arg\max_{\phi} l_\psi(\phi | x_1, \dots, x_n). \tag{3.7}$$

In reality ψ is unknown, hence for each ψ we have a new curve $l_\psi(\phi | x_1, \dots, x_n)$ over ϕ. Now, to estimate ψ, we evaluate the maximum $l_\psi(\phi | x_1, \dots, x_n)$, over ϕ, and choose the ψ, which is the maximum over all these curves, i.e.,

$$\hat{\psi}_P = \arg\max_{\psi} l_\psi(\hat{\phi}_\psi | x_1, \dots, x_n) = \arg\max_{\psi} l_P(\psi, \hat{\phi}_\psi | x_1, \dots, x_n), \tag{3.8}$$

and, logically, $\hat{\psi}$ and $\phi_{\hat{\psi}_P}$ are the maximum likelihood estimators $(\hat{\psi}, \hat{\phi}) = \arg\max_{\psi,\phi}$ $l_P(\psi, \phi | x_1, \dots, x_n)$.

For the transmuted log-logistic model, consider the n lifetimes observed $(t_1, t_2, \dots, t_n)$, with observed right censorship, i.e., $\delta_i = 1$ if the lifetime is observed or 0 otherwise, $i = 1 \dots, n$. Notice that in the presence of a high number of censored lifetimes, it is quite difficult to estimate the model parameter using the pure likelihood. In addition, there is the presence of the nuisance parameter in the model, given by parameter λ, and a vector of parameters (μ, β). Thus, in our case,

$$(\hat{\mu}_P, \hat{\beta}_P) = \arg\max_{\mu,\beta} l_{\mu,\beta}(\hat{\lambda}_{\mu,\beta} | x_1, \dots, x_n) = \arg\max_{\mu,\beta} l_P(\mu, \beta, \hat{\lambda}_{\mu,\beta} | x_1, \dots, x_n), \tag{3.9}$$

where

$$l_P(\mu, \beta, \hat{\lambda}_{\mu,\beta} | \boldsymbol{t}) = \begin{cases} (\beta - 1)\sum_{i=1}^{n} \ln t_i - 3\sum_{i=1}^{n} \ln(1 + e^{\mu} t_i^{\beta}) + k(\ln\beta + \mu) \\ +\sum_{i=1}^{n} \ln\left[(1 + e^{\mu} t_i^{\beta}) - \hat{\lambda}_{\mu,\beta}(e^{\mu} t_i^{\beta} - 1)\right], \text{if } i : \delta_i = 1, \\ \\ -2\sum_{i=1}^{n} \ln(1 + e^{\mu} t_i^{\beta}) + \sum_{i=1}^{n} \ln\left[1 + (1 - \hat{\lambda}_{\mu,\beta}) e^{\mu} t_i^{\beta}\right], \\ \text{if } i : \delta_i = 0, \end{cases}$$

where k is the number of censored lifetimes.

3.3.2 Adjusted profile likelihood

We shall now consider a different adjustment to the profile likelihood function. Suppose that the parameters that index the model are orthogonal, i.e., the elements of the score vector, $\partial l/\partial\mu$, $\partial l/\partial\beta$ and $\partial l/\partial\lambda$, are uncorrelated. As proposed by Cox and Reid (1987), the adjustment can be applied to the profile likelihood function in this setting and it is an approximation to a conditional probability density function of the observations given, the maximum likelihood estimator of μ and β and can be written as

$$L_{AP}(\mu,\beta,\lambda|\boldsymbol{x}) = L_P(\mu,\beta,\hat{\lambda}_{\mu,\beta}|\boldsymbol{x})\left|J_{\lambda\lambda}(\mu,\beta,\hat{\lambda}_{\mu,\beta})\right|^{-1/2} \tag{3.10}$$

where $J_{\lambda\lambda}(\mu,\beta,\hat{\lambda}_{\mu,\beta})$ is the (λ,λ) element of the observed Fischer information $J(\mu,\beta,\lambda)$, see for example Barndorff-Nielsen and McCullagh (1993). In order to obtain the $J_{\lambda\lambda}(\mu,\beta,\lambda)$ element, the Hessian matrix is presented in Appendix A. Then, the expected value is given by

$$\begin{aligned} J_{\lambda\lambda}(\mu,\beta,\lambda) &= \int_0^{+\infty}\left(\frac{1-e^{\mu}x^{\beta}}{1+e^{\mu}x^{\beta}-\lambda e^{\mu}x^{\beta}+\lambda}\right)^2 f(x|\mu,\beta,\lambda)dx \\ &\simeq \frac{\pi e^{-\mu/\beta}}{3\beta^2}(\beta+1)\left[2(\beta+1)^2+1\right]\csc\left(\frac{\pi(\beta+1)}{\beta}\right). \end{aligned} \tag{3.11}$$

By considering n lifetimes observed $(t_1,t_2,\ldots,t_n)$, with observed right censored, i.e., $\delta_i = 1$ if the lifetime is observed or 0 otherwise, $i = 1\ldots,n$, the adjusted profile log-likelihood function can be written as

$$l_{AP}(\mu,\beta,\lambda|\boldsymbol{t}) = \begin{cases} (\beta-1)\sum_{i=1}^n \ln t_i - 3\sum_{i=1}^n \ln(1+e^{\mu}t_i^{\beta})+ \\ +k(\ln\beta+\mu)+\sum_{i=1}^n \ln\left[(1+e^{\mu}t_i^{\beta})-\hat{\lambda}_{\mu,\beta}(e^{\mu}t_i^{\beta}-1)\right]+\zeta, \\ \text{if } i:\delta_i=1, \\ \\ -2\sum_{i=1}^n \ln(1+e^{\mu}t_i^{\beta})+\sum_{i=1}^n \ln\left[1+(1-\hat{\lambda}_{\mu,\beta})e^{\mu}t_i^{\beta}\right]+\zeta, \\ \text{if } i:\delta_i=0, \end{cases} \tag{3.12}$$

where

$$\zeta = \frac{\beta}{e^{-\mu/2\beta}}\left[-\frac{3}{\pi\left[2(\beta+1)^2+1\right]}\sin\left(\frac{\pi(\beta+1)}{\beta}\right)\right]^{1/2}.$$

3.3.3 Modified profile likelihood

The modified profile likelihood was proposed by Barndorff-Nielsen (1993) as an improvement to the profile likelihood and it is defined as

$$L_{MP} = L_P(\theta, \boldsymbol{x}) \left| J_{\phi\phi}(\theta, \hat{\phi}(\theta)) \right|^{-1/2} \left| \frac{\partial \hat{\phi}}{\partial \hat{\phi}(\theta)} \right|. \tag{3.13}$$

The main difficulty in computing the modified profile likelihood function lies in obtaining $|\partial\hat{\phi}/\partial\hat{\phi}(\theta)|$ and several approximations were proposed in order to simplify its evaluation. Severini (1998) proposed an approximation based on empirical covariates and L_{MP} can be rewritten as

$$L_{MP} = L_P(\theta, \boldsymbol{x}) \frac{\left| J_{\phi\phi}(\theta, \hat{\phi}(\theta)) \right|^{1/2}}{\left| I(\theta, \hat{\phi}(\theta); \hat{\theta}, \hat{\phi}) \right|}, \tag{3.14}$$

and replacing $I(\theta, \hat{\phi}(\theta); \hat{\theta}, \hat{\phi})$ by

$$\breve{I}(\theta, \hat{\phi}(\theta); \hat{\theta}, \hat{\phi}) = \sum_{j=1}^{n} \ell_\theta^{(j)}(\theta, \hat{\phi}(\theta)) \ell_\theta^{(j)}(\hat{\theta}, \hat{\phi})^T \tag{3.15}$$

where $\ell_\theta^{(j)}(\cdot)$ is the score function for the jth observation. This approximation is particularly useful when the computation of expected values of products of log-likelihood derivatives is cumbersome.

Thus, for the transmuted log-logistic model,

$$\breve{I}(\mu, \beta, \hat{\lambda}(\mu, \beta); \hat{\mu}, \hat{\beta}, \hat{\lambda}) = \sum_{j=1}^{n} \ell_\lambda^{(j)}(\mu, \beta, \hat{\lambda}(\mu, \beta)) \ell_\lambda^{(j)}(\hat{\mu}, \hat{\beta}, \hat{\lambda})^T$$

and

$$\ell_\lambda^{(j)}(\mu, \beta, \lambda) = \frac{(1 - e^{\mu} x_i^{\beta})}{\left(1 + e^{\mu} x_i^{\beta} - \lambda(e^{\mu} x_i^{\beta} - 1)\right)}.$$

3.3.4 Residual analysis

In order to study the presence of atypical observations, or departures from the error assumptions, we consider two kinds of residuals: martingale-type and deviance. (see for example McCullagh and Nelder (1989), Barlow and Prentice (1988) and Therneau *et al.* (1990)).

The first one, martingale-type residual, was introduced by Therneau *et al.* (1990) and was firstly used in a counting process, which is skewed and has a maximum value at +1 and a minimum value at $-\infty$. By considering the transmuted log-logistic model, the martingale-type residual can be written as

$$r_{M_i} = \begin{cases} 1 + \log\left[1 + e^{\mu} t_i^{\beta}(1-\lambda)\right] - 2\log(1 + e^{\mu} t_i^{\beta}), & \text{if } i : \delta_i = 1, \\ \log\left[1 + e^{\mu} t_i^{\beta}(1-\lambda)\right] - 2\log(1 + e^{\mu} t_i^{\beta}), & \text{if } i : \delta_i = 0, \end{cases} \tag{3.16}$$

where $i = 1, \ldots, n$.

In addition, the deviance residual can be used, widely used in GLMs (generalized linear models). This was proposed by the same authors (Therneau *et al.* (1990)) and it is a transformation of the martingale residual to attenuate the skewness. In our case, the deviance residuals are given by

$$r_{D_i} = \begin{cases} sign(r_{M_i})\left[-2\left[1 + \log(1 + e^{\mu} t_i^{\beta}(1-\lambda)) - 2\log(1 + e^{\mu} t_i^{\beta}) \right.\right. \\ \left.\left. + \log\Big(-\log(1 + e^{\mu} t_i^{\beta}(1-\lambda))\Big) + 2\log(1 + e^{\mu} t_i^{\beta})\right]^{1/2}\right], \\ \text{if } i : \delta_i = 1, \\ \\ sign(r_{M_i})\left[-2\left[1 + \log(1 + e^{\mu} t_i^{\beta}(1-\lambda)) - 2\log(1 + e^{\mu} t_i^{\beta})\right]\right], \\ \text{if } i : \delta_i = 0, \end{cases} \tag{3.17}$$

$i = 1, \ldots, n$.

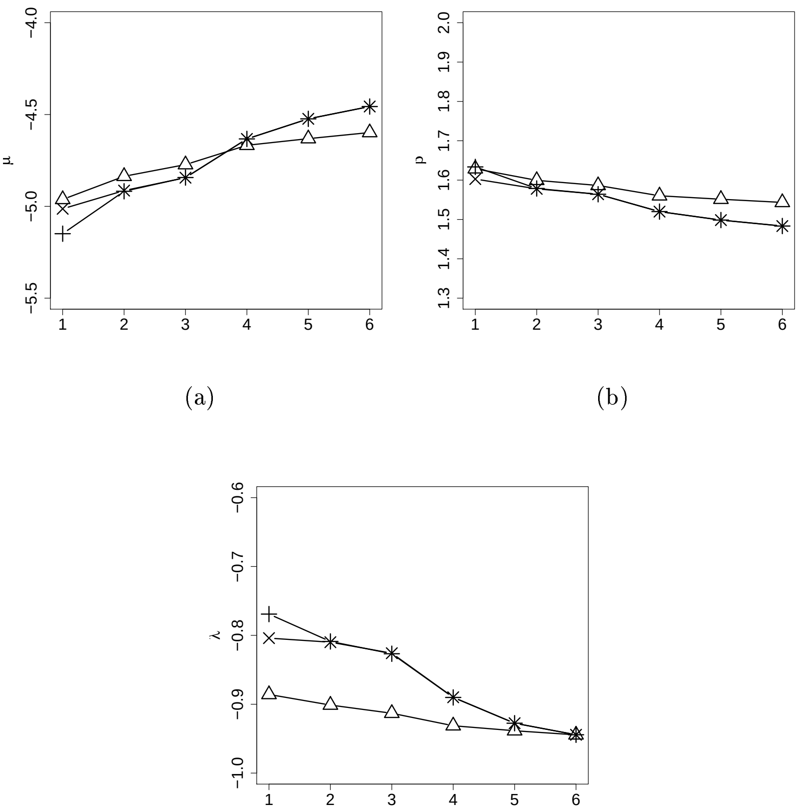
3.3.5 Simulation

In this section, we describe a simulation study designed to assess the frequentist properties of the model in the presence of right censored observations in order to validate the results that will be shown in Section 3.3.6. We must observe that transmuted models have never been built considering the presence of any censor type.

Consider the sample size 148 of patients treated with the drug Linezolid. A bootstrap study was made considering 6 different sizes of re-samples: 50, 80, 100, 200, 300 and 500 (index 1 to 6 respectively). For each sample size, we obtained 1000 re-sampling. The transmuted log-logistic model was fitted considering three estimation methods, presented before: Profile (P), Adjusted Profile (AP) and Modified Profile (MP).

The results of the bootstrap simulations were summarized and are presented in Figures 3.7, 3.8 and 3.9. In the first figure, we present the mean value of estimates of parameters μ, β and λ. Notice that for μ and β parameters, the three profile methods behave similarly. However, in Figure 3.7 the adjusted profile method is better than the other methods for samples less than 200.

As for the mean square error (MSE, in Figures 3.8) it is clear that the adjusted method presents smaller values even though the MSEs are close. Furthermore, with respect to the coverage probability, we expected it to be around 95% in all cases (we can observe that



(a) (b)

(c)

Figure 3.7: *Mean for the parameters estimated: (a) μ, (b) β and (c) λ, by using $\times$ P, $\triangle$ AP and $+$ MP methods.*

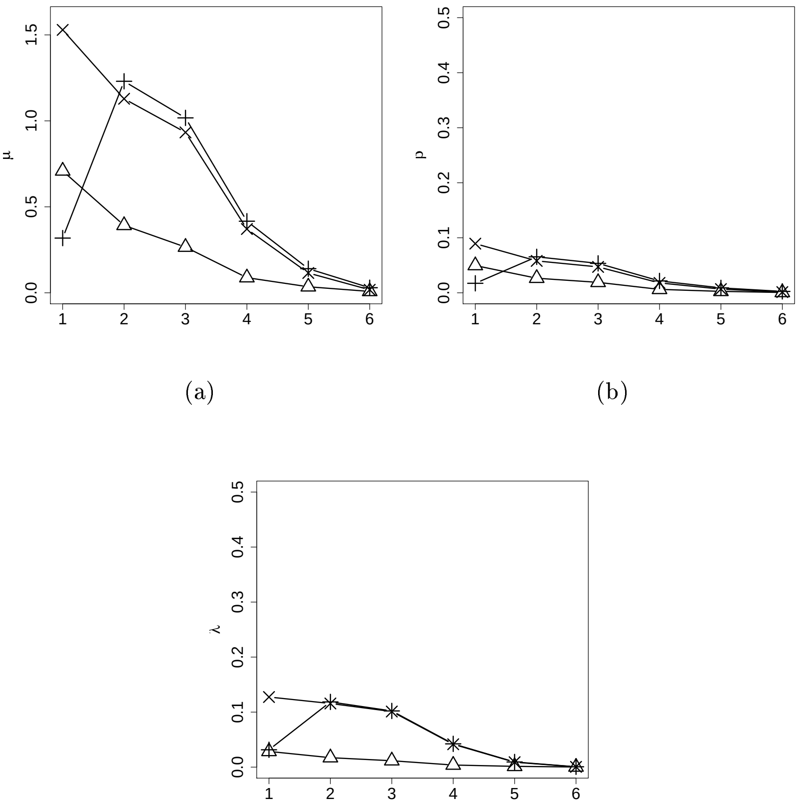
the points are close to the dashed line in Figures 3.9 a, b and c).

3.3.6 Application to real dataset

Linezolid is a synthetic antibiotic developed by a team at the Pharmacia and Upjohn Company, used to treat of serious infections, Brickner (1996).

Discovered in the 1990s and first approved for use in 2000, linezolid was the first commercially available 1, 3-oxazolidinone antibiotic. The main indication of this substance is the treatment of severe infections caused by Gram-positive bacteria that are resistant to other antibiotics. In both the popular press and the scientific literature, linezolid has been called a ”reserve antibiotic“ - one that should be used sparingly so that it will remain effective as a drug of last resort against potentially intractable infections, see Wilson *et al.* (2006).

In Brazil, linezolid has been used since 2007 and researchers are concerned about their indiscriminate use. It is known that its use, for short periods of time, is safe but the effects



(a) (b) (c)

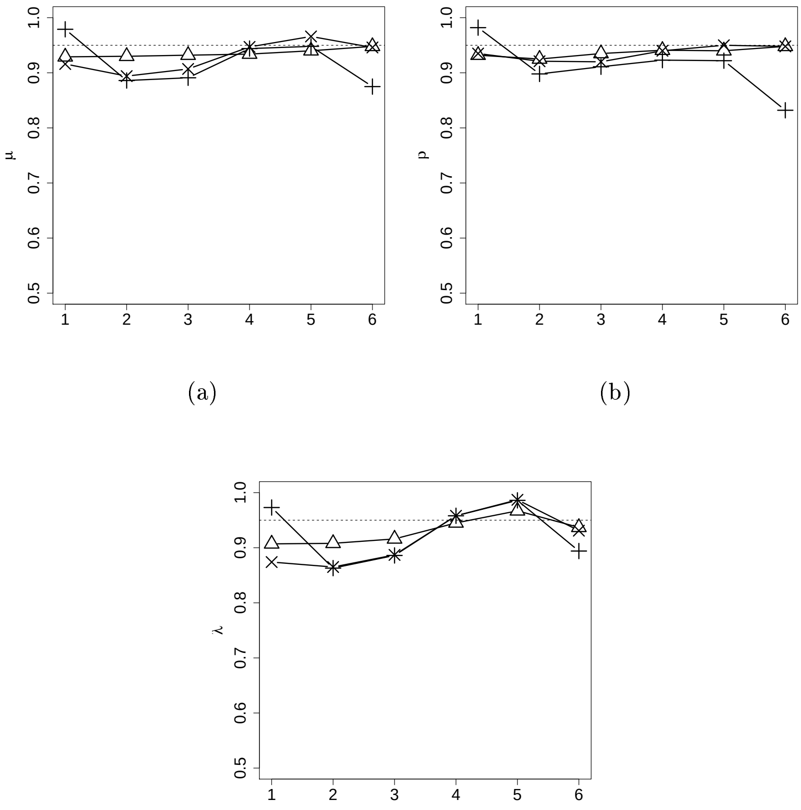
Figure 3.8: *Mean Square Error (MSE) of the parameters estimated: (a) μ, (b) β and (c) λ, by using $\times$ P, $\triangle$ AP and $+$ MP methods.*

of its prolonged use are not known. Given this situation, we conducted a study related to the time of using this drug in patients at the ICU (Intensive Care Unit) at the University Hospital in Maringá city, from 2008 and 2012.

Total of 148 patients were treated with the drug, which is only used in extreme cases of bacterial resistance against other drugs, such as Vancomycin (drug used before linezolid in those cases). An initial analysis was made, in which we noticed that the mean time was 28.4 days and the median time was 22.5 days. Moreover, Figure 3.10-a show the distribution of the time and, it can be seen that the first and third observed quantiles of time of treatment was 13.0 and 36.3 days, respectively.

In order to know the behavior of the failure rate, the TTT plot was plotted in Figure 3.10-b. It is possible to see that the hazard is unimodal, since the TTT plot is initially concave and then convex.

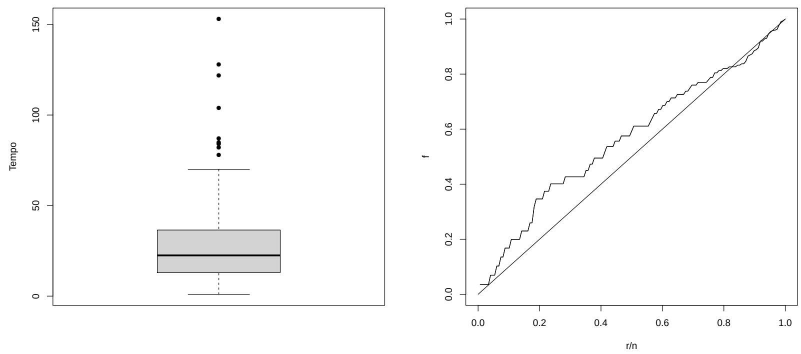
After that, the transmuted log-logistic model was fitted, using the pure likelihood and the profile method. Both results are shown in Table 3.7 below. It can be clearly seen



(a) (b)

(c)

Figure 3.9: *Coverage probability for the parameters estimated: (a) μ, (b) β and (c) λ, by using $\times$ P, $\triangle$ AP and $+$ MP methods.*



(a) (b)

Figure 3.10: *(a) Boxplot of the observed times and (b) TTT plot.*

that the estimated parameters are very close in both likelihood methods. However, the profile confidence intervals are shorter when using the pure likelihood. Furthermore, if we consider the Wald confidence intervals for the profile method, the intervals are even smaller, as shown in Table 3.8.

Table 3.7: *Estmates for the parameters of transmuted log-logistic model estimated by using profile methods, considering right censored lifetimes.*

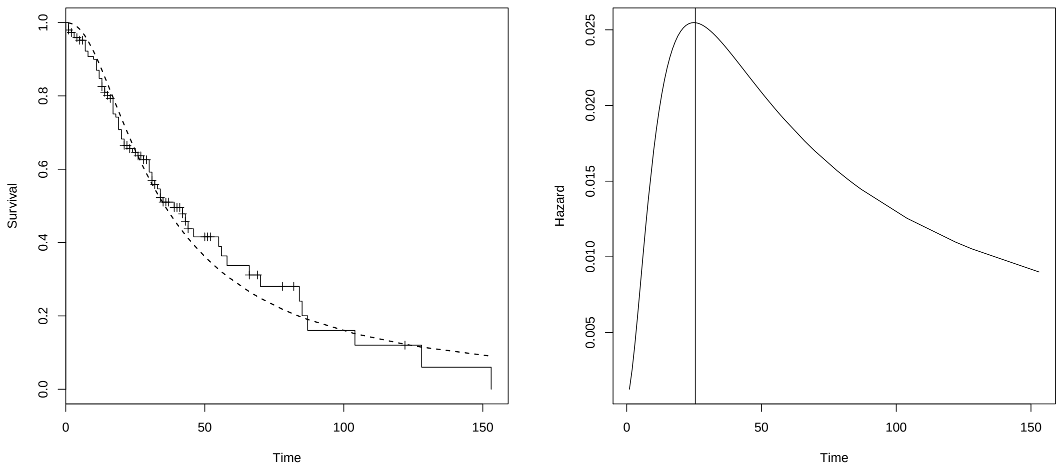
Method	Parameter	Estimate	Error	Confidence Intervals 95%	
				Lower	Upper
Profile	α	-4.4446	0.8072	-5.0528	-3.9723
	β	1.4804	0.2089	1.3526	1.6301
	λ	-0.9483	0.1685	-1.0000	-0.7928
Adjusted	α	-4.5909	0.6548	-5.0649	-4.1828
	β	1.5415	0.1702	1.4319	1.6614
	λ	-0.9473	0.1382	-1.0000	-0.8296
Modified	α	-4.3460	0.8980	-5.0183	-3.8702
	β	1.4480	0.2344	1.3142	1.6135
	λ	-0.9577	0.1796	-1.0000	-0.7888

Table 3.8: *Wald confidence limits by considering 95% of confidence.*

Method	Lower	Upper
Profile	-4.9890	-3.9001
	1.3394	1.6213
	-1.0000	-0.8346
Adjusted	-5.0649	-4.1828
	1.4319	1.6614
	-1.0000	-0.8296
Modified	-4.9517	-3.7403
	1.2899	1.6060
	-1.0000	-0.8366

Figure 3.11-a and b, shows the estimated survival curves versus empirical (by Kaplan-Meier). It can be clearly observed how close the curves are. Moreover the hazard curve can also be seen. Notice that the most probable time for patient discharge is 25.44 days, with 95% confidence interval given by $(15.47; 48.69)$ days.

Furthermore, Figure 3.12 shows the graphics of the Profile, Adjusted and Modified relative log-likelihoods (divided by the maximum absolute value of the log-likelihood). It can be observed that the Profile and Adjusted Profile showed a lower decrease (after the maximum value) than the Modified Profile log-likelihoods.



(a) (b)

Figure 3.11: *(a) Survival curves: estimated versus empirical; (b) Hazard estimate curve, with the* $\hat{T}_{\max}$.

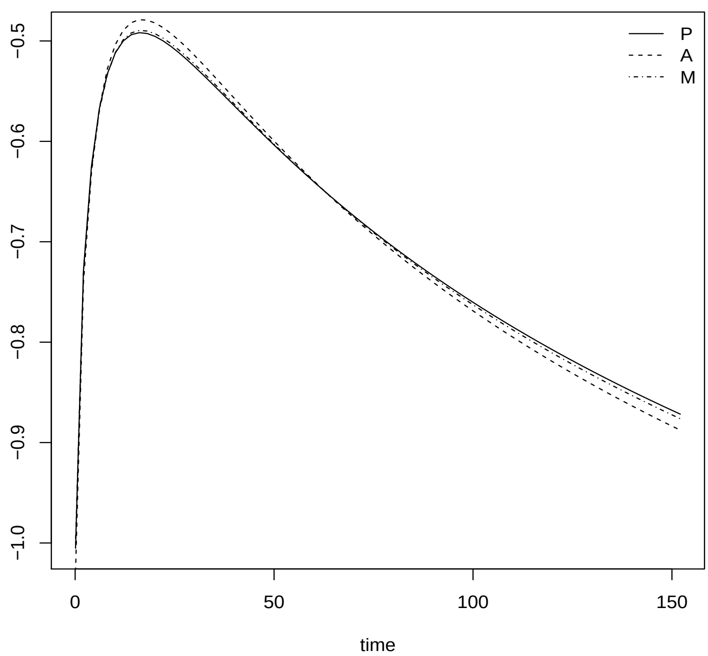


Figure 3.12: *(P) Profile, (A) Adjusted Profile and (M) Modified Profile relative log-likelihood for the Linezolid data.*

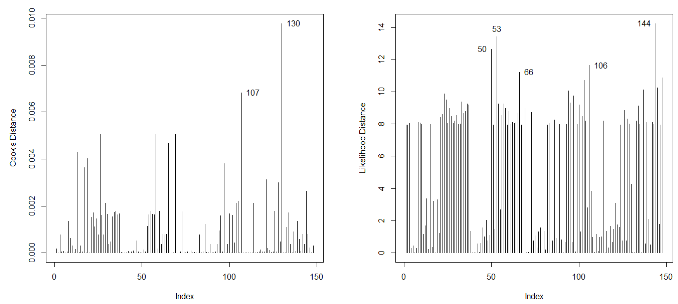
Global and local influence

In this section, we will make an analysis of the global and local influence for the data set given, using the transmuted log-logistical model, as presented in Section 3.2.1.

In order to analyze the local influence, here we consider the response variable perturbation, i.e., we will consider that each t_i is perturbed as $t_{im} = t_i + m_i S_t$, where S_t is a scale factor that may be the estimated standard deviation of T and $m_i \in \mathbb{R}$. Then, the perturbed log-likelihood function becomes

$$\begin{aligned} l_n(\mu, \beta, \lambda \mid \boldsymbol{t}, \boldsymbol{c}, \boldsymbol{m}) = & \sum_{i=1}^{n} c_i \ln \beta + \mu \sum_{i=1}^{n} c_i + (\beta - 1) \sum_{i=1}^{n} c_i \ln t_{im} \\ & + \sum_{i=1}^{n} c_i \ln \left[(1 + e^{\mu} t_{im}^{\beta}) - \lambda (e^{\mu} t_{im}^{\beta} - 1) \right] - 3 \sum_{i=1}^{n} c_i \ln(1 + e^{\mu} t_{im}^{\beta}) \\ & + \sum_{i=1}^{n} (1 - c_i) \ln \left[1 + (1 - \lambda) e^{\mu} t_{im}^{\beta} \right] - 2 \sum_{i=1}^{n} (1 - c_i) \ln(1 + e^{\mu} t_{im}^{\beta}). \end{aligned}$$

Figure 3.13-a and b, show, respectively, the Cook's generalized and likelihood distances and it can be observed that the pertubation causes some disproportionate effects.



(a) (b)

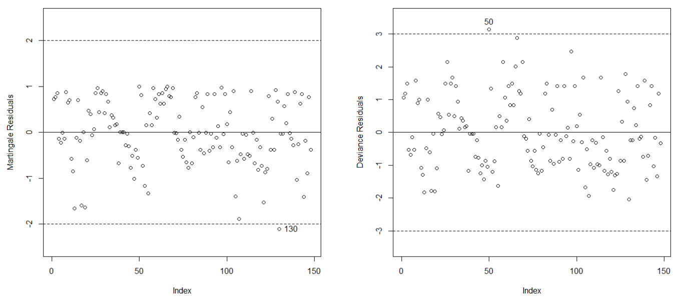
Figure 3.13: *(a) Cooks's distance and (b) Likelihood distance after response pertubation.*

After analyzing Figure 3.13-a and b, we can see the distinction of some observations in relation to others. Furthermore, we made a residual analysis using the Martingale-type and deviance presented in Section 3.3.4. Figures 3.14 show the results of these analyses.

To reveal the impact of the detected influential observations, we use three measures defined by Lee *et al.* (2006): TRC, MRC and LD. In our case, we find TRC $= 22.78$, MRC$= 0.0873$ and $\text{LD}_{(I)} = -32.3$, for $I = \{50, 53, 130, 144\}$. Hence, the results are more sensitive for the influential observation group.

After the influence and residual analysis, the possible influential observations were identified: case 50, 53, 130 and 144. Then, we removed each of these points and fitted the transmuted log-logistic model again for each case. In order to reveal the impact of them, the relative changes were measured RC_{θ_j} and all these results are shown in Table 3.9.

As a complementary analysis, the possible influential observations were removed. Aided by the analysis of local influence and residual analysis, observations $50, 53, 130$ and 144 were



(a) (b)

Figure 3.14: *(a) Martingale residuals; (b) Deviance residuals.*

Table 3.9: *RC (in %) and the corresponding TRC, MRC and* $LD_{(I)}$.

Case removed	Parameter	RC	TRC	MRC	$LD_{(I)}$
{50}	μ	0.682	7.336	5.452	4.300
	β	1.202			
	λ	5.452			
{53}	μ	1.442	3.831	4.144	14.300
	β	1.513			
	λ	0.875			
{130}	μ	4.144	10.704	4.144	13.400
	β	3.702			
	λ	2.858			
{144}	μ	2.187	5.707	2.187	12.600
	β	2.128			
	λ	1.392			
{50; 53; 130; 144}	μ	8.734	22.777	8.734	33.300
	β	8.127			
	λ	5.916			

then removed. The results can be found in Table 3.10.

The graphics of the Profile, Adjusted and Modified relative log-likelihoods (divided by the maximum absolute value of the log-likelihood) were plotted again and they are presented in Figure 3.15. It can be observed that the Profile and Adjusted Profile had a lower decrease (after the maximum value) than the Modified Profile log-likelihoods, as seen in Figure 3.12 showing some advantage, which are not significative, for this method.

Table 3.10: *Estimates for the parameters of transmuted log-logistic model estimated by using profile methods after removing influential observations.*

Method	Parameter	Estimate	Error	Confidence Intervals 95%	
				Lower	Upper
Profile	α	−4.8327	0.9350	−5.4634	−4.2021
	β	1.6006	0.2374	1.4405	1.7607
	λ	−0.8922	0.2139	−1.0000	−0.7479
Adjusted	α	−4.8334	0.7183	−5.3179	−4.3489
	β	1.6248	0.1872	1.4985	1.7511
	λ	−0.9192	0.1542	−1.0000	−0.8152
Modified	α	−4.8259	1.0426	−5.5292	−4.1227
	β	1.5922	0.2624	1.4153	1.7692
	λ	−0.8848	0.2419	−1.0000	−0.7216

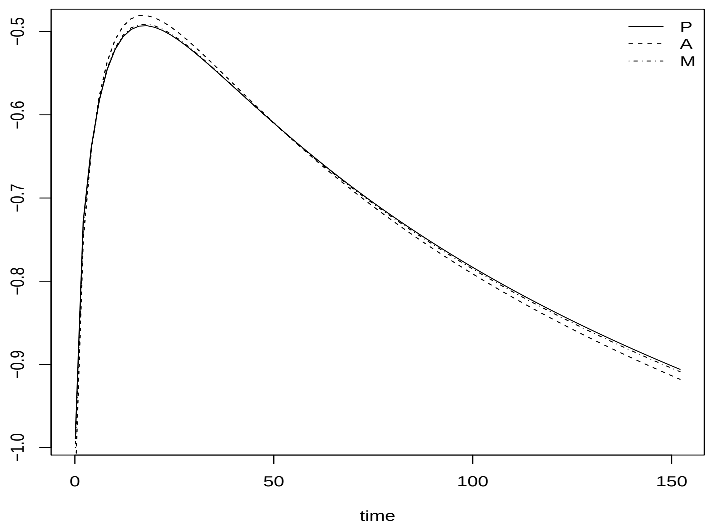


Figure 3.15: *(P) Profile, (A) Adjusted Profile and (M) Modified Profile relative log-likelihood for the Linezolid data.*

Chapter 4

Cubic ranking transmuted model

This Chapter introduces a new order of transmuted distribution, the cubic rank transmutation map distribution. This new order increases the flexibility of transmuted distributions providing the accommodation of more complex data; for instance, data in the presence of bimodal hazard rates. In order to exemplify the applicability of the cubic rank transmutation map, we used two of the best known distributions maps in the reliability field: the Weibull and the log-logistic. Several mathematical properties of these new distributions, namely the transmuted Weibull distribution of order 2 and transmuted log-logistic distribution of order 2, are derived (either cubic rank transmuted Weibull or cubic rank transmuted log-logistic). Inference is based on maximum likelihood and applications were made using real datasets. Analysis of diagnostic and a bootstrap simulation study are part of this work.

4.1 The cubic ranking transmutation map

In this section we propose a cubic ranking transmutation map, following the transmutation described in Section 2.1, but now considering order 2.

Theorem 4.1.1 *Let X_1, X_2 and X_3 be independent and identically random variables with distribution $G(x)$. Then, the ranking cubic transmutation map is given by*

$$F(x) = \lambda_1 G(x) + (\lambda_2 - \lambda_1)G(x) + (1 - \lambda_2)G^3(x), \tag{4.1}$$

with $\lambda_1 \in [0,1]$ and $\lambda_2 \in [-1,1]$.

Proof: *Let X_1, X_2 and X_3 be independent and identically random variables distributed with distribution $G(x)$. Now, consider the following order*

$$X_{1:3} = \min(X_1, X_2, X_3), \qquad X_{2:3} \text{ and } \qquad X_{3:3} = \max(X_1, X_2, X_3),$$

and let

$$\begin{cases} Y \stackrel{d}{=} X_{1:3}, & \textit{with prob } \pi_1, \\ Y \stackrel{d}{=} X_{2:3}, & \textit{with prob } \pi_2, \\ Y \stackrel{d}{=} X_{3:3}, & \textit{with prob } \pi_3, \end{cases}$$

where $\sum_{i=1}^{3} \pi_i = 1 \Rightarrow \pi_3 = 1 - \pi_1 - \pi_2$. *Evidently,* $F_Y(x)$ *is given by*

$$\begin{aligned} F_Y(x) &= \pi_1 Pr(\min(X_1, X_2, X_3) \le x) + \pi_2 Pr(X_{2:3} \le x) + \pi_3 Pr(\max(X_1, X_2, X_3) \le x) \\ &= 3\pi_1 G(x) + 3(\pi_2 - \pi_1) G^2(x) + (1 - \pi_2) G^3(x), \end{aligned}$$

and if $3\pi_1 = \lambda_1$ *and* $3\pi_2 = \lambda_2$*, the cubic ranking transmuted distribution (or transmuted of order 2) becomes the one given in (4.1).* □

Definition 4.1.1 *The density function* $f(x)$ *of the cubic ranking transmutation distribution is given by*

$$f(x) \;=\; g(x)\left[\lambda_1 + 2(\lambda_2 - \lambda_1)G(x) + 3(1 - \lambda_2)G^2(x)\right]. \tag{4.2}$$

□

In what follows, we focus on two particular cases based on the well-known Weibull and log-logistic lifetime distributions.

4.2 Cubic rank transmuted weibull distribution

The Weibull distribution, originally proposed by Weibull (1951), is being used effectively in many applied problems; see for example, Johnson *et al.* (1996), Nichols and Padgett (2006), Pham and Lai (2007), and Aryal and Tsokos (2011). Let X be a positive random variable with Weibull distribution, then, the cumulative function is given by

$$G(x) = 1 - \exp\left[-\left(\frac{x}{\beta}\right)^{\mu}\right], \qquad x > 0, \tag{4.3}$$

where $\mu > 0$ and $\beta > 0$ are shape and scale parameters, respectively. It is important to note that for $\mu < 1$ the failure rate decreases over time (which happens, for instance, in data where there is significant "infant mortality"), $\mu = 1$ indicates that the failure rate is constant over time (i.e., it is an exponential distribution as a special case), and $\mu > 1$ possesses the failure rate increasing over time, (Lawless, 2011).

Proposed by Aryal and Tsokos (2011), the transmuted Weibull distribution was formulated as a generalization of the Weibull distribution. The transmuted quadratic rank Weibull cumulative distribution function has the form

$$F(x) \;=\; \left[1 - \exp\left(-\left(\frac{x}{\beta}\right)^{\mu}\right)\right]\left[1 + \lambda \exp\left(-\left(\frac{x}{\beta}\right)^{\mu}\right)\right],$$

where $\lambda \in [-1, 1]$. The cubic rank transmuted Weibull distribution introduced here has the following form.

Proposition 4.2.1 *Let X have a Weibull distribution with parameters μ and β. Then, the density function of the cubic rank transmuted Weibull distribution is given by*

$$f(x) = \frac{\mu}{\beta^{\mu}} x^{\mu-1} e^{-(x/\beta)^{\mu}} \Big[(3-\lambda_1-\lambda_2) + 2(\lambda_1+2\lambda_2-3)e^{-(x/\beta)^{\mu}} + 3(1-\lambda_2)e^{-2(x/\beta)^{\mu}}\Big], \tag{4.4}$$

for $x > 0$.

Proof: *With X having a Weibull distribution with parameters μ and β, we obtain from equation (4.1) that*

$$\begin{aligned} F(x) &= \lambda_1 G(x) + (\lambda_2 - \lambda_1)G^2(x) + (1-\lambda_2)G^3(x) \\ &= 1 + (\lambda_1+\lambda_2-3)e^{-(x/\beta)^{\mu}} + (3-\lambda_1-2\lambda_2)e^{-2(x/\beta)^{\mu}} + (\lambda_2-1)e^{-3(x/\beta)^{\mu}}, \end{aligned}$$

for $x > 0$. Upon differentiating the above expression with respect to x, we obtain the density function presented in (4.1). □

Note that the cubic rank transmuted Weibull is an extended distribution to analyze more complex data, generalizing the Weibull distribution ($\lambda_1 = \lambda_2 = 1$) and the transmuted Weibull distribution ($\lambda_2 = 1$). Some of the possible shapes that the cubic rank transmuted Weibull distribution can take on are illustrated in the upper panels of Figure 4.1 for some selected values of the parameters. The λ's are responsible for introducing extra skewness into the Weibull distribution. This is in full agreement with Shaw and Buckley (2009) who pointed out that the introduction of extra skewness into a distribution is a direct effect of transmutation maps of order 1, or transmuted quadratic rank ones.

Now, let T be a random variable representing the lifetime of an unit. Then, the survival and hazard functions of T are given, respectively, by

$$R(t) = (3-\lambda_1-\lambda_2)\zeta(t) + (\lambda_1+2\lambda_2-3)(\zeta(t))^2 + (1-\lambda_2)(\zeta(t))^3, \quad t > 0, \tag{4.5}$$

and

$$h(t) = \frac{\mu}{\beta^{\mu}} t^{\mu-1} \left[\frac{(3-\lambda_1-\lambda_2) + 2(\lambda_1+2\lambda_2-3)\zeta(t) + 3(1-\lambda_2)(\zeta(t))^2}{(3-\lambda_1-\lambda_2) + (\lambda_1+2\lambda_2-3)\zeta(t) + (1-\lambda_2)(\zeta(t))^2}\right], \quad t > 0, \tag{4.6}$$

where $\zeta(t) = e^{-(t/\beta)^{\mu}}$. Several plots of the above survival and hazard functions are presented in the lower panels of Figure 4.1.

Moments are essential for inferential proposes. Moreover, they are useful to study some important features and characteristics of a distribution such as tendency, dispersion, skewness and kurtosis. We derive the r^{th} moment for the cubic rank transmuted Weibull dis-

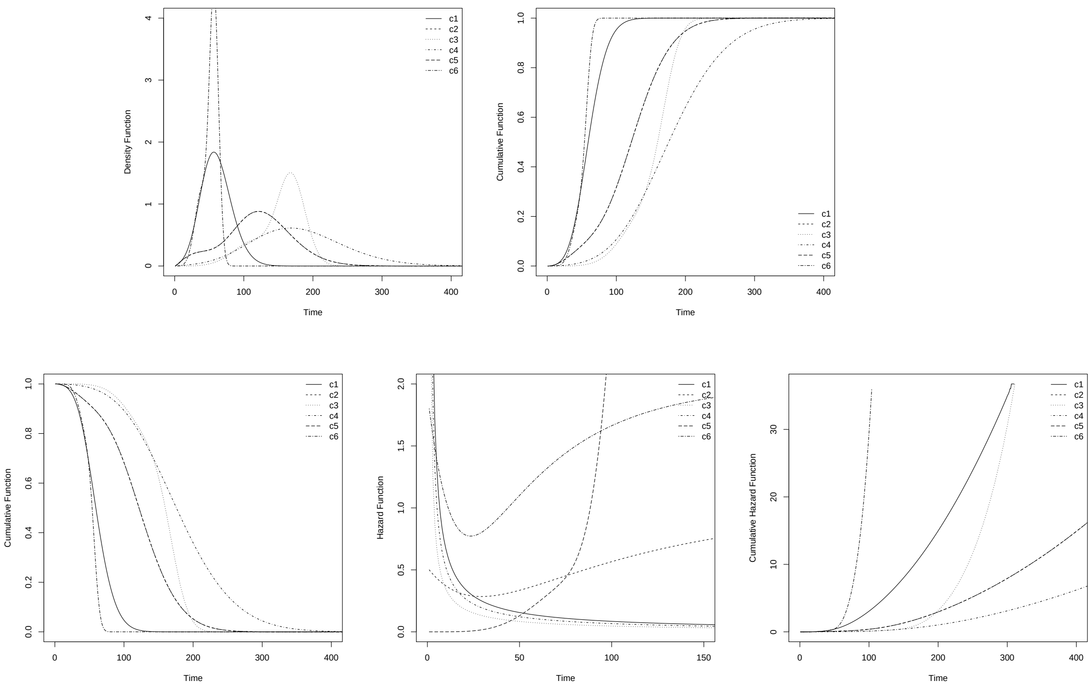


Figure 4.1: *Upper panels: Density and cumulative distribution functions; Lower panels: Survival, hazard and cumulative hazard functions, all of which are plotted for different values of model parameters.*

tribution as follows.

The r^{th} moment is given by

$$\begin{aligned} E\left(X^r\right) &= \int_0^{+\infty} x^r \frac{\mu}{\beta^\mu} x^{\mu-1} e^{-(x/\beta)^\mu} \Big[(3-\lambda_1-\lambda_2) + 2(\lambda_1+2\lambda_2-3)e^{-(x/\beta)^\mu} \\ &\quad +3(1-\lambda_2)e^{-2(x/\beta)^\mu}\Big]\, dx. \end{aligned}$$

With the use of gamma function, we can obtain an expression for the above integral as

$$E\left(X^r\right) = \frac{\mu}{\beta^{1-r}} \Gamma\left(\frac{r-1}{\mu}+2\right)\left[(3-\lambda_1-\lambda_2) + \frac{(\lambda_1+2\lambda_2-3)}{2^{\frac{r-1}{\mu}+1}} + \frac{(1-\lambda_2)}{3^{\frac{r-1}{\mu}+1}}\right],$$

where $\Gamma(\cdot)$ denotes a complete gamma function. In particular, the mean and variance are given, respectively, by

$$E\left(X\right) = \frac{\mu}{6}(11-3\lambda_1-2\lambda_2) \tag{4.7}$$

and

$$\begin{aligned} V\left(X\right) &= \Gamma\left(\frac{1}{\mu}+2\right)\beta\mu\left[(3-\lambda_1-\lambda_2) + \frac{(\lambda_1+2\lambda_2-3)}{2^{\frac{1}{\mu}+1}} + \frac{(1-\lambda_2)}{3^{\frac{1}{\mu}+1}}\right] \\ &\quad -\frac{\mu^2}{36}(11-3\lambda_1-2\lambda_2)^2. \end{aligned}$$

4.3 Cubic rank transmuted log-logistic distribution

Let X be a nonnegative random variable representing the lifetime of an individual in a population log-logistically distributed as presented in Section 2.1. Then, the cubic rank transmuted log-logistic distribution can be defined as follows.

Proposition 4.3.1 *Let X be random variables having a log-logistic distribution with parameters β and μ. Then, the density function of the cubic rank transmuted log-logistic distribution is given by*

$$f(x) = \frac{\mu x^{\mu-1}e^{\beta}}{(1+e^{\beta}x^{\mu})^4}\left[\lambda_1(1-e^{2\beta}x^{2\mu})+\lambda_2 e^{\beta}x^{\mu}(2-e^{\beta}x^{\mu})+3e^{2\beta}x^{2\mu}\right], \quad x>0. \quad (4.8)$$

Proof: *Let X be a log-logistic random variable with parameters β and μ. Then, from equation (4.1), we obtain*

$$\begin{aligned} F(x) &= \lambda_1 G(x)+(\lambda_2-\lambda_1)G^2(x)+(1-\lambda_2)G^3(x) \\ &= \frac{x^{\beta}e^{\mu}}{(1+e^{\mu}x^{\beta})}\left[\lambda_1+(\lambda_2-\lambda_1)\frac{x^{\beta}e^{\mu}}{(1+e^{\mu}x^{\beta})}+(1-\lambda_2)\frac{x^{2\beta}e^{2\mu}}{(1+e^{\mu}x^{\beta})^2}\right], \quad x>0. \end{aligned}$$

Upon differentiating the above expression with respect to x, we obtain the density function as presented in (4.8). □

Note that the cubic rank transmuted log-logistic model is also an extended distribution that is quite useful for analysing more complex data, as presented for the Weibull model in Section 4.2. The log-logistic distribution is clearly a special case when $\lambda_1=\lambda_2=1$. Some of the possible shapes that the cubic rank transmuted log-logistic distribution can take on are illustrated in the upper panels of Figure 4.2 for some selected values of the parameters. The λ's are responsible for introducing extra skewness into the log-logistic distribution.

Furthermore, the reliability and hazard functions of this distribution are given by

$$R(t) = 1-\kappa(t)\left[\lambda_1+(\lambda_2-\lambda_1)\kappa(t)+(1-\lambda_2)\kappa^2(t)\right], \quad t>0, \quad (4.9)$$

and

$$h(t) = \frac{\mu t^{\mu-1}e^{\beta}}{(1+e^{\beta}t^{\mu})^3}\left[\frac{\lambda_1(1-e^{2\beta}t^{2\mu})+\lambda_2 e^{\beta}t^{\mu}(2-e^{\beta}t^{\mu})+3e^{2\beta}t^{2\mu}}{1+t^{\mu}e^{\beta}-t^{\mu}e^{\beta}\left[\lambda_1+(\lambda_2-\lambda_1)\kappa(t)+(1-\lambda_2)\kappa^2(t)\right]}\right], \quad (4.10)$$

$t>0$, respectively, where $\kappa(t)=\frac{e^{\beta}t^{\mu}}{(1+e^{\beta}t^{\mu})}$. Several plots of survival and hazard functions are presented in the lower panels of Figure 4.2.

The r^{th} moment of the cubic rank transmuted log-logistic distribution is given by

$$E\left(X^r\right) = \int_0^{+\infty}\frac{\mu x^{r+\mu-1}e^{\beta}}{(1+e^{\beta}x^{\mu})^4}\left[\lambda_1(1-e^{2\beta}x^{2\mu})+\lambda_2 e^{\beta}x^{\mu}(2-e^{\beta}x^{\mu})+3e^{2\beta}x^{2\mu}\right]dx.$$

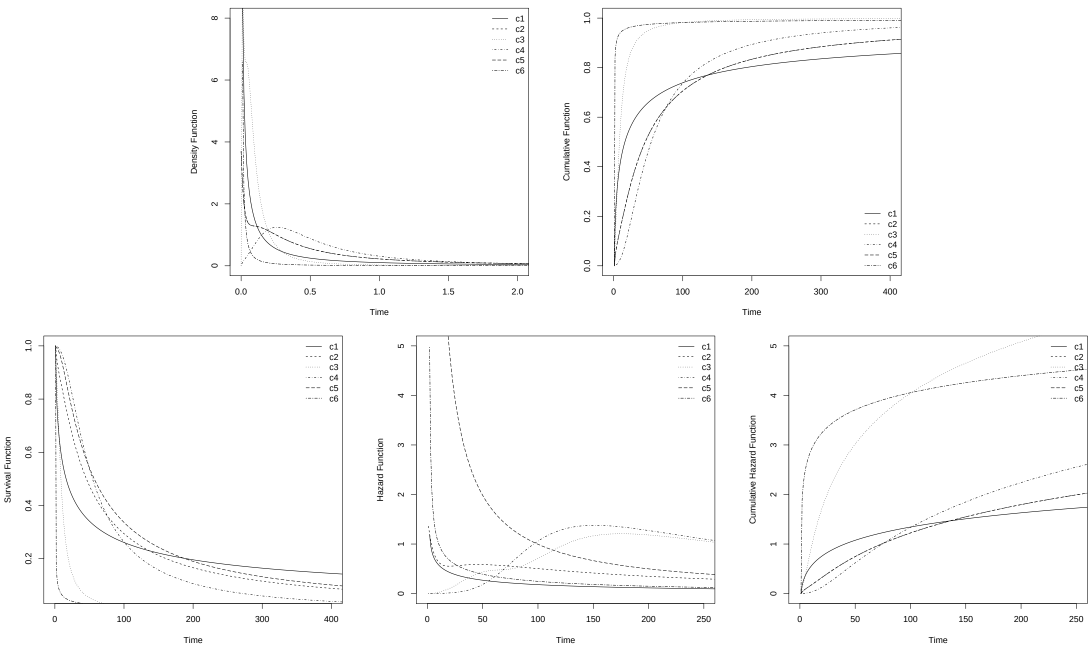


Figure 4.2: *Upper panels: Density and cumulative distribution functions; Lower panels: Survival, cumulative hazard and hazard functions, all of which are plotted for different values of model parameters.*

Upon using beta function, we obtain an expression for the r^{th} moment as

$$E\left(X^{r}\right) = \frac{e^{-\beta r/\mu}}{6} B\left(1-\frac{r}{\mu}, \frac{r}{\mu}+1\right)\left\{\frac{r^{2}}{\mu^{2}}(3-\lambda_2+2\lambda_1)+3\frac{r}{\mu}(3+\lambda_2)+2(2\lambda_1+\lambda_2+3)\right\},$$

where $B(\cdot,\cdot)$ denotes the complete beta function. In particular, the mean and variance are given by

$$E\left(X\right)=\frac{e^{-\beta/\mu}}{6\mu^{2}} B\left(\frac{\mu-1}{\mu}, \frac{\mu+1}{\mu}\right)\left\{2\lambda_1(1+2\mu^{2})+\lambda_2(\mu^{2}+3\mu-1)+7\mu^{2}+9\mu\right\} \quad (4.11)$$

and

$$\begin{aligned} V\left(X\right) = \frac{e^{-2\beta/\mu}}{6\mu^{2}} & \left[B\left(\frac{\mu-2}{\mu}, \frac{\mu+2}{\mu}\right)\left\{4\lambda_1(2+\mu^{2})+2\lambda_2(\mu^{2}+3\mu-2)\right.\right. \\ & \left.+6(\mu^{2}+3\mu+2)\right\} - B^{2}\left(\frac{\mu-1}{\mu}, \frac{\mu+1}{\mu}\right)\left\{2\lambda_1(1+2\mu^{2})+\lambda_2(\mu^{2}+3\mu-1)\right. \\ & \left.\left.+7\mu^{2}+9\mu\right\}^{2}\right]. \end{aligned} \quad (4.12)$$

4.4 Parameter estimation

Maximum likelihood approach can be used for the estimation of model parameters. Let $X_1, \ldots, X_n$ be a random sample of size n so that the likelihood function is given by

$$L = L(\boldsymbol{\theta}|X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} f(X_i|\boldsymbol{\theta}),$$

where $\boldsymbol{\theta}$ is the parameter vector. For the cubic rank transmuted Weibull distribution, the log-likelihood function becomes

$$\begin{aligned} l_W &= n\log(\mu) - n\mu\log(\beta) + (\mu-1)\sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\mu} \\ &\quad + \sum_{i=1}^{n}\log\left[(3-\lambda_1-\lambda_2) + 2(\lambda_1+2\lambda_2-3)e^{-(x_i/\beta)^{\mu}} + 3(1-\lambda_2)e^{-2(x_i/\beta)^{\mu}}\right]. \end{aligned} \tag{4.13}$$

Therefore, the maximum likelihood estimates (MLEs) of μ, β, λ_1 and λ_2, which maximize (4.13), must satisfy the following normal equations:

$$\begin{aligned} \frac{\partial l_W}{\partial \mu} &= \frac{n}{\mu} - \sum_{i=1}^{n}\log x_i - \sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\mu}\log\left(\frac{x_i}{\beta}\right) \\ &\quad + 2\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\mu}\log\left(\frac{x_i}{\beta}\right)e^{-(x_i/\beta)^{\mu}}\left[\frac{\lambda_1+2\lambda_2-3}{\vartheta_i} - 3(1-\lambda_2)e^{-(x_i/\beta)^{\mu}}\right] = 0, \\ \frac{\partial l_W}{\partial \beta} &= \frac{2\mu}{\beta}\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\mu}e^{-(x_i/\beta)^{\mu}}\left[\frac{\lambda_1+2\lambda_2-3}{\vartheta_i} + 3(1-\lambda_2)e^{-(x_i/\beta)^{\mu}}\right] \\ &\quad - \frac{n\mu}{\beta} + \frac{\mu}{\beta}\sum_{i=1}^{n}\left(\frac{x_i}{\beta}\right)^{\mu} = 0, \\ \frac{\partial l_W}{\partial \lambda_1} &= \sum_{i=1}^{n}\frac{2e^{-(x_i/\beta)^{\mu}}-1}{\vartheta_i} = 0, \\ \frac{\partial l_W}{\partial \lambda_2} &= \sum_{i=1}^{n}\frac{4e^{-(x_i/\beta)^{\mu}}-1}{\vartheta_i} - 3e^{-2(x_i/\beta)^{\mu}} = 0, \end{aligned}$$

where $\vartheta_i = (3-\lambda_1-\lambda_2) + 2(\lambda_1 - 2\lambda_2 - 3)e^{-(x_i/\beta)^{\mu}}$. The MLE $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\beta}, \hat{\lambda_1}, \hat{\lambda_2})'$ is obtained by solving the above nonlinear system of equations. It is usually more convenient to use nonlinear optimization algorithms such as quasi-Newton or Newton-Raphson to numerically maximize the log-likelihood function in (4.13).

Similarly, for the cubic rank transmuted log-logistic distribution, the log-likelihood function is given by

$$\begin{aligned} l_{LL} &= n\beta + n\log\mu - 4\sum_{i=1}^{n}\log(1+e^{\beta}x_i^{\mu}) + (\mu-1)\sum_{i=1}^{n}\log x_i + \\ &+ \sum_{i=1}^{n}\log(\lambda_1 - \lambda_1 e^{2\beta}x_i^{2\mu} + 2\lambda_2 e^{\beta}x_i^{\mu} - \lambda_2 e^{2\beta}x_i^{2\mu} + 3e^{2\beta}x_i^{2\mu}). \end{aligned} \tag{4.14}$$

Therefore, the MLE of μ, β, λ_1 and λ_2 which maximize (4.14) must satisfy the following normal equations:

$$\begin{aligned} \frac{\partial l_{LL}}{\partial\mu} &= \frac{n}{\mu} - 4\sum_{i=1}^{n}\frac{\eta_i\log x_i}{1+\eta_i} + \sum_{i=1}^{n}\log x_i \\ &+ \sum_{i=1}^{n}2\eta_i\log x_i\left[\frac{3\eta_i + \lambda_2(1-\eta_i) - \lambda_1\eta_i}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1}\right] = 0, \\ \frac{\partial l_{LL}}{\partial\beta} &= n - 4\sum_{i=1}^{n}\frac{\eta_i}{1+\eta_i} + 2\sum_{i=1}^{n}\eta_i\left[\frac{\eta_i(3-\lambda_1-\lambda_2)+\lambda_2}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1}\right] = 0, \\ \frac{\partial l_{LL}}{\partial\lambda_1} &= \sum_{i=1}^{n}\frac{1-\eta_i^2}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1} = 0, \\ \frac{\partial l_{LL}}{\partial\lambda_2} &= \sum_{i=1}^{n}\frac{\eta_i(2-\eta_i)}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1} = 0, \end{aligned}$$

where $\eta_i = e^{\beta}x_i^{\mu}$.

The MLE $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\beta}, \hat{\lambda}_1, \hat{\lambda}_2)'$ is obtained by solving the above nonlinear system of equations.

In order to compute the asymptotic confidence intervals, we use the usual large-sample approximation for the distribution of the maximum likelihood estimator of $\boldsymbol{\theta}$ which is four-variate normal (Migon *et al.*, 2014). Thus, as $n \to \infty$, the asymptotic distribution of the MLE, for both cubic rank transmuted Weibull and log-logistic distributions is given by,

$$\begin{pmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix} \sim N\left[\begin{pmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{\lambda}_1 \\ \hat{\lambda}_2 \end{pmatrix}, \begin{pmatrix} \hat{V}_{11}, \hat{V}_{12}, \hat{V}_{13}, \hat{V}_{14} \\ \hat{V}_{21}, \hat{V}_{22}, \hat{V}_{23}, \hat{V}_{24} \\ \hat{V}_{31}, \hat{V}_{32}, \hat{V}_{33}, \hat{V}_{34} \\ \hat{V}_{41}, \hat{V}_{42}, \hat{V}_{43}, \hat{V}_{44} \end{pmatrix}\right], \tag{4.15}$$

where $\hat{V}_{ij} = V_{ij}\mid_{\theta=\hat{\theta}}$. The asymptotic variance-covariance matrix V is determined by the inverse of Hessian matrix; see A for the corresponding details for the cubic rank transmuted Weibull and log-logistic distributions. Then, approximate $100(1-\alpha)\%$ two-sided confidence intervals for μ, β, λ_1 and λ_2 are, respectively, given by $\hat{\mu} \pm z_{\alpha/2}\sqrt{\hat{V}_{11}}$, $\hat{\beta} \pm z_{\alpha/2}\sqrt{\hat{V}_{22}}$, $\hat{\lambda}_1 \pm z_{\alpha/2}\sqrt{\hat{V}_{33}}$ and $\hat{\lambda}_2 \pm z_{\alpha/2}\sqrt{\hat{V}_{44}}$, where z_α is the upper $\alpha - th$ percentile of the standard normal distribution.

Fits of different distributions can be compared by penalizing over-fitting by using the Akaike information criterion (Akaike, 1973), which minimizes the Kullback-Leibler divergence between the true distribution and the estimate from a candidate distribution. It is given by AIC $= -2l(\hat{\boldsymbol{\theta}}) + 2size(\boldsymbol{\theta})$, where $l(\hat{\boldsymbol{\theta}})$ denotes the log-likelihood function evaluated at the maximum value and $size(\boldsymbol{\theta})$ is the number of model parameters. The distribution with the lowest value of this criterion (among all considered distributions) is regarded as the preferred distribution for describing a given dataset.

4.5 Numerical studies

In this section, we present the results of a simulation study designed to assess the properties of the proposed estimation procedure. Moreover, we illustrate the usefulness of the proposed models with two real data sets. An analysis of global and local influence is also carried out.

4.5.1 Simulation study

A simulation study was performed by considering samples of size 50, 80, 100, 150, 300 and 500 from the cubic rank transmuted Weibull and cubic rank transmuted log-logistic distributions. A total of 1000 random samples were generated for each set up with the parameters fixed as $\mu = 2.0$, $\beta = 2.5$, $\lambda_1 = 0.2$, and $\lambda_2 = 1.0$ for the cubic rank transmuted Weibull distribution, and $\mu = 3.0$, $\beta = -16.0$, $\lambda_1 = 0.2$, and $\lambda_2 = -0.9$ for the cubic rank transmuted log-logistic distribution.

Tables 4.1 and 4.2 present the means of the estimates and the mean square errors (MSE) as well as the coverage probabilities of 95% two-sided confidence intervals for the model parameters. The MSE decreases with increasing sample size, showing the consistency of the estimators. Moreover, the coverage probabilities become closer to the nominal level as the sample size increases.

4.5.2 Application: Carbon fibers data

In this section, we provide an application of the cubic rank transmuted Weibull distribution. For this purpose, we consider a data set from a study on breaking stress of carbon fibers (in Gba), taken from Nichols and Padgett (2006). The lifetimes are times until the break of the fibers. TTT plot of the lifetimes is presented in the upper left panel of Figure 4.3, indicating a possible increasing hazard function.

Also, for the sake of comparison, we have considered six alternative distributions: the transmuted Weibull distributions of orders 2 and 1, and the prominent Weibull, exponential, gamma and log-normal distributions.

Table 4.1: *Means of estimates of all parameters and mean square errors (MSE), for different sample sizes.*

Distribution	Sample Size	Estimate				MSE			
		μ	β	λ_1	λ_2	μ	β	λ_1	λ_2
	50	2.226	2.338	0.317	0.492	0.895	0.413	0.821	1.787
Trans	80	2.2	2.35	0.308	0.569	0.714	0.416	0.667	1.376
Weibull	100	2.179	2.346	0.289	0.597	0.63	0.397	0.563	1.258
Order2	150	2.176	2.363	0.29	0.652	0.53	0.374	0.457	1.028
	300	2.15	2.364	0.245	0.746	0.402	0.336	0.296	0.771
	500	2.142	2.376	0.228	0.807	0.314	0.281	0.23	0.568
	50	2.893	−16.138	0.293	−0.822	0.304	4.559	0.120	0.257
Trans	80	2.889	−16.115	0.299	−0.823	0.253	3.507	0.107	0.244
Log-Logistic	100	2.890	−16.122	0.313	−0.838	0.166	1.929	0.080	0.187
Order2	150	2.880	−16.061	0.317	−0.845	0.110	1.083	0.059	0.138
	300	2.876	−16.036	0.318	−0.848	0.096	0.915	0.053	0.124
	500	2.876	−16.030	0.321	−0.853	0.040	0.309	0.025	0.059

Table 4.2: *Coverage probabilities of the confidence intervals for different sample sizes.*

Distribution	Sample Size	Coverage Probability			
		μ	β	λ_1	λ_2
	50	0.577	0.671	0.551	0.732
Transmuted	80	0.662	0.72	0.642	0.781
Cubic Rank	100	0.702	0.734	0.690	0.787
Weibull	150	0.783	0.808	0.777	0.815
	300	0.894	0.897	0.903	0.898
	500	0.959	0.959	0.956	0.947
	50	0.830	0.824	0.864	0.521
Transmuted	80	0.852	0.855	0.887	0.624
Cubic Rank	100	0.879	0.882	0.906	0.749
Log-Logistic	150	0.893	0.898	0.902	0.835
	300	0.893	0.894	0.897	0.857
	500	0.901	0.904	0.940	0.950

In Table 4.3, we have presented, for all these distributions, seven different comparison measures used as selection criteria: $-2\times$ log-likelihood (-2log), Akaike's information criterion (AIC), corrected Akaike's information criterion (AICC), Schwarz's Bayesian information criterion (BIC), Kolmogorov-Smirnov statistic (KS), Anderson-Darling statistic (A) and Cramér-von-Mises statistic (W). The calculated values of theses statistics (the smaller the better) all reveal that the cubic rank transmuted Weibull distribution is the most appropriate model according to four different criteria and the Weibull distribution is the most appropriate one according to AIC, AICC and BIC (it must be noted that, in these criteria, the number of estimated parameters has a huge impact).

Figure 4.4 shows P-P plots that represent the empirical cumulative versus estimated cumulative functions of the rank transmuted Weibull distributions of orders 2 and 1 and the Weibull distribution, respectively.

In order to examine the global fit of the distribution, we carried out a residual anal-

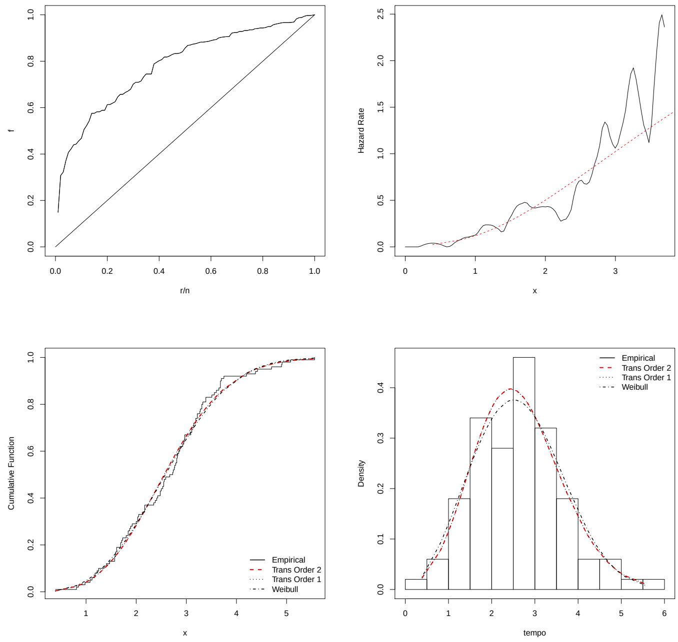


Figure 4.3: *Upper panels: TTT plot and hazard curves empirical and estimated by cubic rank transmuted Weibull distribution; Lower panels: Cuimulative and density functions estimated by transmuted Weibull and Weibull distributions.*

Table 4.3: *Selection criteria estimated for six different lifetime distributions.*

Distribution	$-2\log$	AIC	AICC	BIC	KS	A*	W*
Cubic rank transmuted	282.694	290.694	291.115	301.115	0.647	0.371	0.064
Transmuted quadratic rank	282.702	288.702	288.952	296.518	0.688	0.372	0.074
Weibull	283.059	287.059	287.182	292.269	0.651	0.390	0.068
Exponential	392.742	394.742	394.783	397.347	3.225	17.076	3.392
Gamma	286.467	290.467	290.591	295.678	1.027	0.695	0.147
Log-Normal	296.840	300.840	300.963	306.050	1.313	1.380	0.267

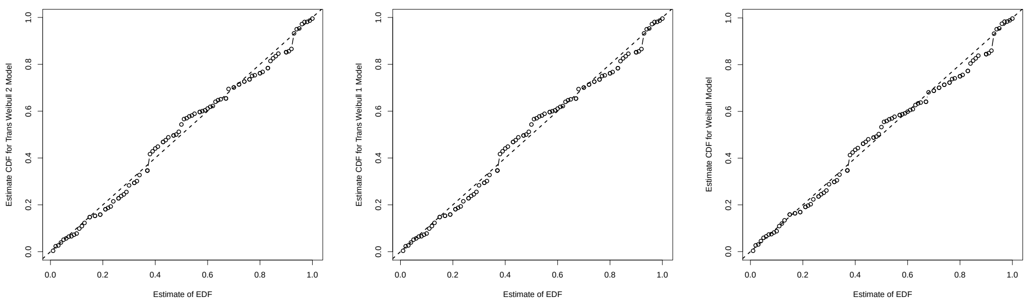


Figure 4.4: *P-P plot estimated for transmuted cubic and quadratic rank Weibull and Weibull distributions.*

yses by using the Martingale-type residual and the deviance residual; see, for example, McCullagh and Nelder (1989), Barlow and Prentice (1988) and Therneau *et al.* (1990) for pertinent details. The first one, the martingale-type residual, was introduced by Therneau *et al.* (1990) and is based on a counting process. For the cubic rank transmuted Weibull distribution, the martingale-type residual can be expressed as

$$r_{M_i} = (3-\lambda_1-\lambda_2)\zeta(t_i) + (\lambda_1+2\lambda_2-3)(\zeta(t_i))^2 + (1-\lambda_2)(\zeta(t_i))^3, t_i > 0, \qquad (4.16)$$

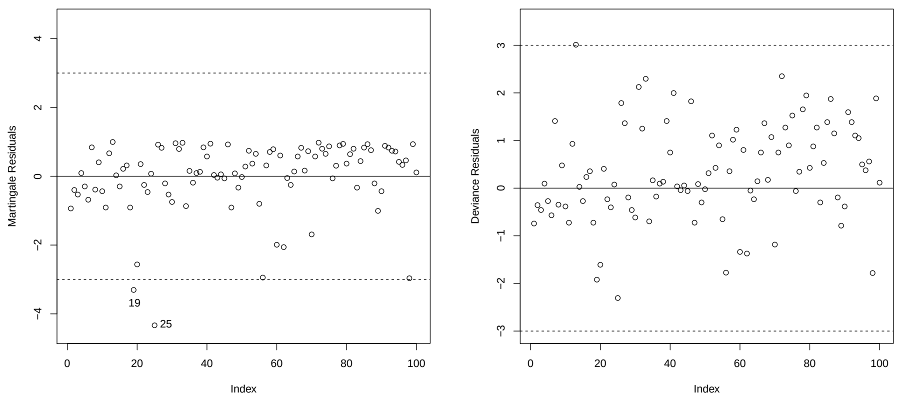
where $\zeta(t_i) = e^{-(t_i/\beta)^\mu}$, $i = 1, \ldots, n$.

In addition, it is possible to use the deviance residual that has been widely used in generalized linear models. This was also proposed by the same authors (Therneau *et al.* (1990)) and it is a transformation of the martingale residual to attenuate the skewness. In our case, the deviance residuals are given by

$$r_{D_i} = \text{sign}(\hat{r}_{M_i})\left[-2\left(\hat{r}_{M_i} + \log\left(1-\hat{r}_{M_i}\right)\right)\right] \qquad (4.17)$$

for $i = 1, \ldots, n$.

Figure 4.5 shows the Martingale and deviance residuals.



(a) (b)

Figure 4.5: *(a) Martingale residuals; (b) Deviance residuals.*

Global and local influence

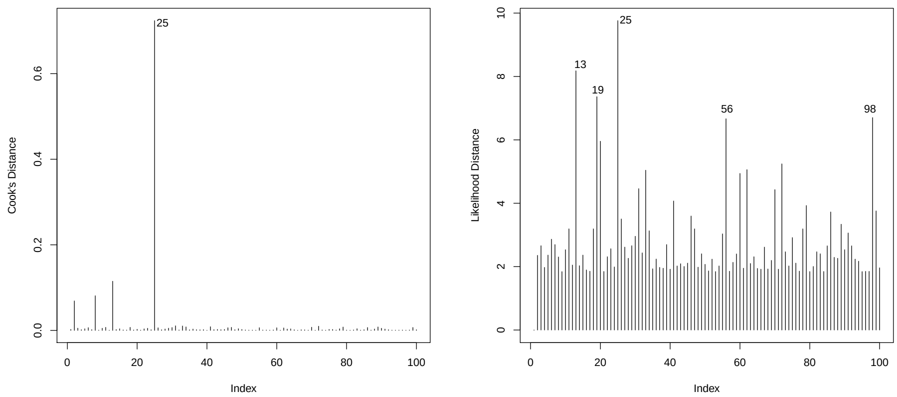
In this section, we perform global and local influence analyses by considering the cubic rank transmuted Weibull distribution as presented in Section 3.2.1.

For analysing the local influence, we consider the response variable perturbation, i.e., we consider that each t_i is peturbed as $t_{im} = t_i + m_i V_t$, where V_t is a scale factor that may be the estimated standard deviation of T and $m_i \in \mathbb{R}$. Then, the perturbed log-likelihood function can be expressed as

$$\ell\left(\psi|\mathbf{t},\mathbf{m}\right) = n\log(\mu) - n\mu\log(\beta) + (\mu-1)\sum_{i=1}^{n}\log t_{im} - \sum_{i=1}^{n}\left(\frac{t_{im}}{\beta}\right)^{\mu} \quad (4.18)$$

$$+\sum_{i=1}^{n}\log\left[(3-\lambda_1-\lambda_2)+2(\lambda_1+2\lambda_2-3)e^{-(t_{im}/\beta)^{\mu}}+3(1-\lambda_2)e^{-2(t_{im}/\beta)^{\mu}}\right].$$

We can see possible presence of influential points in panels a and b of Figure 4.6, which show the generalized Cook's and likelihood distances.



(a) (b)

Figure 4.6: *(a) Cook's distance; (b) Likelihood distance.*

Note that observation 25 appears to be a discrepant point under both the residual analysis of adjusted distribution and Cook's and likelihood distances for perturbed and non-perturbed data. To assess the impact of the detected influential observations, the RC_{θ_j} was calculated and the impact was measured by using the total and maximum relative changes and the likelihood displacement given by TRC $= \sum_{i=1}^{n_p}\left|\mathrm{RC}_{\theta_j}\right| = 4.18$, MRC $= \max_j\left|\mathrm{RC}_{\theta_j}\right| = 3.65$ and $LD_{(I)}(\boldsymbol{\theta}) = 2\{l(\hat{\boldsymbol{\theta}} - l(\hat{\boldsymbol{\theta}}_I)\} = 9.79$, with $n_p = 4$ (the number of parameters). In fact, we consider this point as an outlier and the cubic rank transmuted Weibull distribution was re-fitted. The MLEs of the parameters and their respective standard deviations (in brackets) are found to be $\hat{\mu} = 2.702$ (1.119), $\hat{\beta} = 2.731$ (1.161), $\hat{\lambda}_1 = 0.721$ (1.289), and $\hat{\lambda}_2 = 0.968$ (1.617). The re-fitted distribution can be seen in Figure 4.7. Also, the estimated mean of breaking stress to the carbon fibres turns out to be 3.110 and the corresponding standard error to be 1.025.

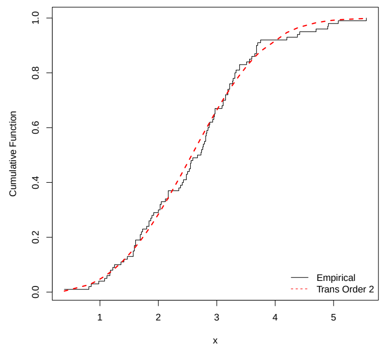


Figure 4.7: *Cumulative curves: empirical vs estimated for cubic rank transmuted Weibull distribution.*

4.5.3 Application: Cattle sexual precocity data

In order to illustrate the usefulness of the cubic rank transmuted log-logistic distribution for modelling, in this section we consider a data set concerned with a study on the economic results concerning beef cattle which are directly related to their genetic prepotency, presented in Section 2.5.

The transmuted quadratic and cubic rank log-logistic distributions and the log-logistic distribution were fitted to these data. Table 4.4 provides the MLEs, their corresponding standard errors and 95% confidence intervals for the model parameters. For these distributions, the computed $-2\log$ -likelihood and AIC values are presented in Table 4.5.

Table 4.4: *MLEs of the parameters of the transmuted quadratic and cubic ranks log-logistic and log-logistic distributions.*

Distribution	Parameter	Estimate	Standard Error	95% Confidence Interval	
				Lower	Upper
	μ	2.8758	0.0208	2.8351	2.9165
Transmuted	β	-16.0305	0.1263	-16.2781	-15.7829
Cubic Rank	λ_1	0.3201	0.0131	0.2942	0.3457
	λ_2	-0.8505	0.0304	-0.9101	-0.7908
Transmuted	μ	3.0503	0.0220	3.0073	3.0933
Quadratic Rank	β	-17.9079	0.1369	-18.1763	-17.6395
	λ	-0.8139	0.0120	-0.8374	-0.7904
Log-Logistic	μ	2.5832	0.0160	2.5519	2.6145
	β	-16.7482	0.1003	-16.9448	-16.5516

Both criteria provide evidence in favor of the cubic rank transmuted log-logistic distribution. This result is corroborated further by the fitted distribution over the histogram of the data presented in the upper right panel of Figure 4.8, and by the fitted survival functions under the Kaplan-Meier estimator on the lower left panel of Figure 4.8. In order to examine the global fit of the distribution, the P-P plot is presented in Figure 4.9.

Table 4.5: *Computed $-2\log$ likelihood and AIC values for the log-logistic and transmuted log-logistic of orders 1 and 2 distributions.*

Distribution	$-2\log$	AIC	AICC	BIC	A^*	W^*
Transmuted cubic	235160	235168	235168	235199	94.15	19.30
Transmuted quadratic	236201	236207	236207	236230	142.53	26.63
Log-Logistic	236745	236749	236749	236764	3113.52	662.43

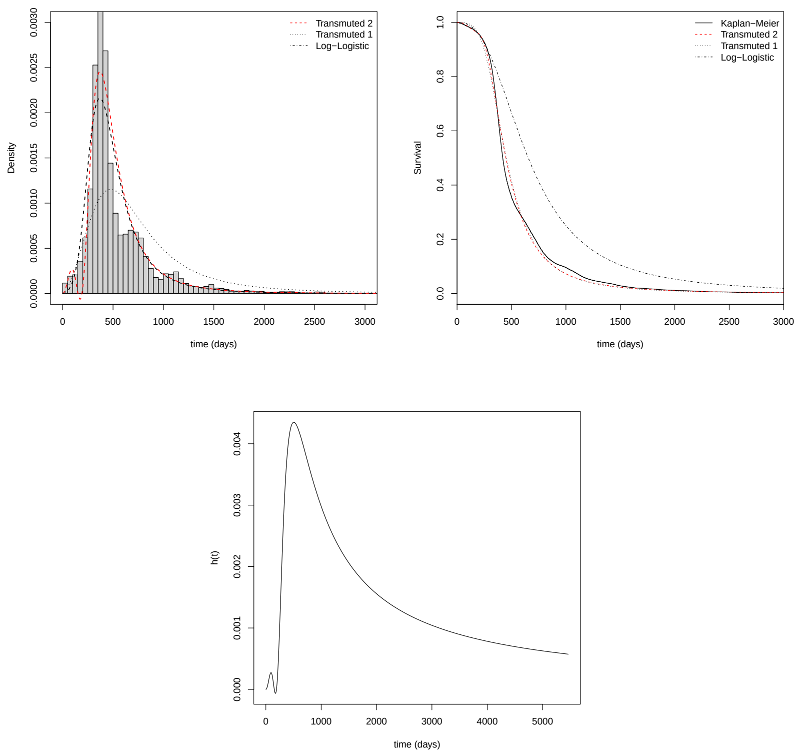


Figure 4.8: *Upper Panels: histogram and survival curves; Lower Panel: hazard estimate curve.*

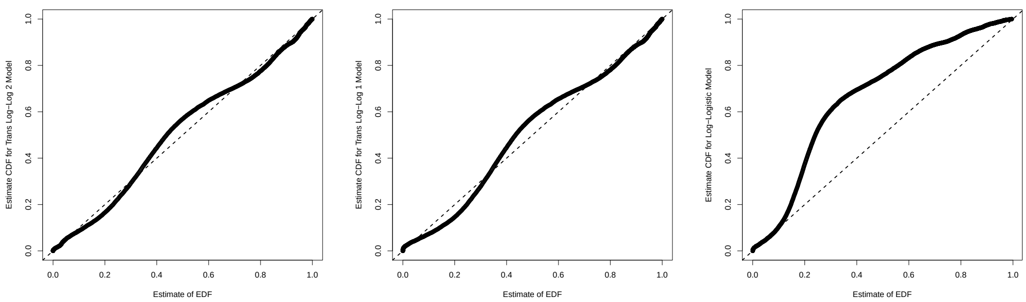


Figure 4.9: *P-P plot estimated for transmuted cubic and quadratic rank log-logistic and log-logistic distributions, respectively.*

Upon using the most appropriate distribution for describing the time up to first calving, viz., the cubic rank transmuted log-logistic distribution, we turn our attention to estimating the mean time for the first calf to occur and it is determined to be 1252.54 days (or approximately 41.75 months), with a standard error of 308.32 days. Also, the hazard curve possesses two modes: the first one occuring at 814.5 days and the second one occuring at 1205 days.

Chapter 5

e-Transmuted family of distribution

Although transmutation maps are a convenient way to construct new distributions, the restricted parametric space of the extra parameter λ may be a problem in some situations. As an alternative to this class of model, in this Chapter we present the regression e-Transmuted family of model (or Exponential Transmuted), which has the property that the extra parameter λ can take any real value.

5.1 Formulation of the model

Let Y be a continuous and positive random variable that represents the lifetime of an individual in some population of interest. Consider the parametric family of cumulative distribution functions $\mathcal{G} = \{G(y|\boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ where $\boldsymbol{\theta}$ contains a set of k unknown parameters with $\Theta \subseteq \mathbb{R}^k$. The first objective of this work is to construct a larger family of distributions that contains the baseline family $\mathcal{G}$. More specifically, we would like to find a family of distributions $F(y|\boldsymbol{\theta}, \omega)$ where $\omega \in \mathbb{R}$ that satisfies

$$F(y|\boldsymbol{\theta}, \omega = 0) = G(y|\theta)\,, \quad \text{for all } y\,,\, \theta \in \Theta\,. \tag{5.1}$$

In this sense, the new family $\mathcal{F}_{\mathcal{G}} = \{F(\cdot|\boldsymbol{\theta}, \omega) : \boldsymbol{\theta} \in \Theta\,,\, \omega \in \mathbb{R}\}$ not only generalizes $\mathcal{G}$ but is also "centred" around $\mathcal{G}$ in the scale defined by the parameter ω. The proposed generalized family is defined next.

Definition 5.1.1 *Given a parametric family of distributions $\mathcal{G}$, the cumulative distribution functions in the extended family $\mathcal{F}_{\mathcal{G}}$ are defined as follows:*

$$F(y|\boldsymbol{\theta}, \omega) = \frac{1 - e^{-\omega\, G(y|\boldsymbol{\theta})}}{1 - e^{-\omega}}, \qquad \omega \in \mathbb{R}\,, \boldsymbol{\theta} \in \Theta. \tag{5.2}$$

We will call the generated family $\mathcal{F}_{\mathcal{G}}$, the exponential extension (or e-extension) of the baseline family $\mathcal{G}$ and we will use the notation $Y \sim$ e-extension$(\omega, G(\cdot|\boldsymbol{\theta}))$.

It is trivial to see that $F(y|\boldsymbol{\theta},\omega)$ is in fact, a cumulative distribution function. In order to graphically see that the e-extended family is centred at $\mathcal{G}$, we can take, for example, G to be an exponential distribution with unit mean, that is $G(y|\theta) = 1 - \exp(-y)$. Cumulative distribution functions in $\mathcal{F}_{\mathcal{G}}$ have been plotted in Figure 5.1 where we can clearly see that the functions are centred around the baseline distribution G either above or below, depending on the sign of ω.

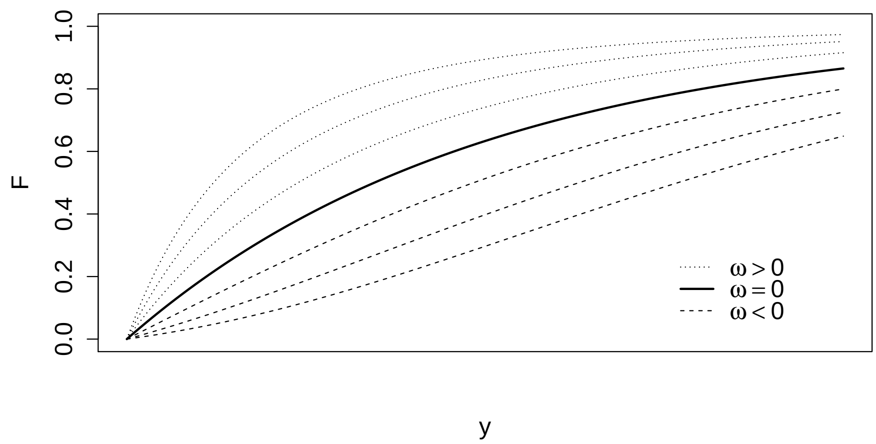


Figure 5.1: *The effect of ω in the cdf curves of the e-extended family with $G(y)$ exponentially distributed.*

In more generality, that is, for any distribution $G(y|\boldsymbol{\theta})$, a simple application of L'Hospital's rule shows that $F(y|\boldsymbol{\theta},\omega=0) = G(y|\boldsymbol{\theta})$ since

$$\lim_{\omega\to 0} F(y|\boldsymbol{\theta},\omega) = \lim_{\omega\to 0} \frac{\partial/\partial\omega\left[\omega e^{-\omega G(y|\boldsymbol{\theta})}\right]}{\partial/\partial\omega\left[1-e^{-\omega}\right]} = G(y|\boldsymbol{\theta}) \qquad \text{for all } y,\boldsymbol{\theta}$$

In this Chapter, we focus on distributions defined on the positive real line but clearly our construction is valid for any support of the random variable Y. Also, we will assume from now on that the functional form of the distributions in $\mathcal{G}$ is known. Now, the probability density functions in an e-extended family can be written as:

$$f(y|\boldsymbol{\theta},\omega) = \frac{\omega g(y|\boldsymbol{\theta})e^{-\omega G(y|\boldsymbol{\theta})}}{1-e^{-\omega}}, \qquad \boldsymbol{\theta}\in\Theta\,,\,\omega\in\mathbb{R}, \tag{5.3}$$

where $g(y|\boldsymbol{\theta})$ is the density corresponding to $G(y|\boldsymbol{\theta})$. Applying L'Hospital rule again, shows that $f(y|\boldsymbol{\theta},\omega=0) = g(y|\boldsymbol{\theta})$.

The following result, provides a simple probabilistic interpretation of our construction (5.2) and (5.3). More explicitly, a random variable Y that follows an e-extension of a given $\mathcal{G}$ can be constructed via a truncated (to lie in the unit interval $(0,1)$) and then transformed (by the quantile function G^{-1}) exponential random variable.

Result 1 *Let Z be an exponential random variable with rate ω which has been truncated to lie in the unit interval* $(0,1)$. *Then $Y = G^{-1}(Z|\boldsymbol{\theta}) \sim$e-Transmuted$(\omega, G)$.*

We clearly have that

$$P(Z \leq z) = \frac{1-\exp(-\omega\, z)}{1-\exp(-\omega)}$$

for any $z \in (0,1)$ so that for any $y > 0$ we have

$$P(Y \leq y) = P(G^{-1}(Z|\boldsymbol{\theta}) \leq y) = P\left(Z \leq G(y|\boldsymbol{\theta})\right) = \frac{1-\exp(-\omega\, G(y|\boldsymbol{\theta}))}{1-\exp(-\omega)}.$$

This result can be seen as a particular case of the construction of extended parametric models presented by Verdinelli *et al.* (1998). Clearly, the result also provides a simple method to generate random samples from an e-extended family. In particular, the lower qth quantile of e-extended model is given by

$$y_q = G^{-1}(z_q(\omega)|\boldsymbol{\theta})$$

where

$$z_q(\omega) = -\frac{\log\left[1-q(1-e^{-\omega})\right]}{\omega}$$

that is, the composition of two quantiles, first for the truncated exponential $(z_q(\omega))$ and then the quantile $G^{-1}(\cdot|\boldsymbol{\theta})$ of the baseline family $\mathcal{G}$. Generation a random samples from an e-extended family is then trivial by using the inverse method.

5.2 Hazard and related functions

Given that we are interested in distributions with positive support we develop here expressions for survival and hazard functions. The survival function, which is the probability of the survival time Y being larger than or equal to some time y, has a simple expression in terms of the baseline survival function $S(y|\boldsymbol{\theta}) = 1 - G(y|\boldsymbol{\theta})$, namely

$$S(y|\boldsymbol{\theta},\omega) = \frac{e^{\omega\, S(y|\boldsymbol{\theta})}-1}{e^{\omega}-1}, \tag{5.4}$$

so that clearly $S(y|\boldsymbol{\theta},\omega=0) = S(y|\boldsymbol{\theta})$. The other function of interest is the hazard rate function, which for an e-extended model can be written as

$$h(y|\boldsymbol{\theta},\omega) = \frac{\omega\, h(y|\boldsymbol{\theta})\, S(y|\boldsymbol{\theta})}{1-e^{-\omega S(y|\boldsymbol{\theta})}}, \tag{5.5}$$

where $h(y|\boldsymbol{\theta}) = g(y|\boldsymbol{\theta})/S(y|\boldsymbol{\theta})$ is the baseline hazard function. Figure 5.2 shows the behaviour of the hazard rate functions for different values of ω when the baseline distribution is exponential with mean one.

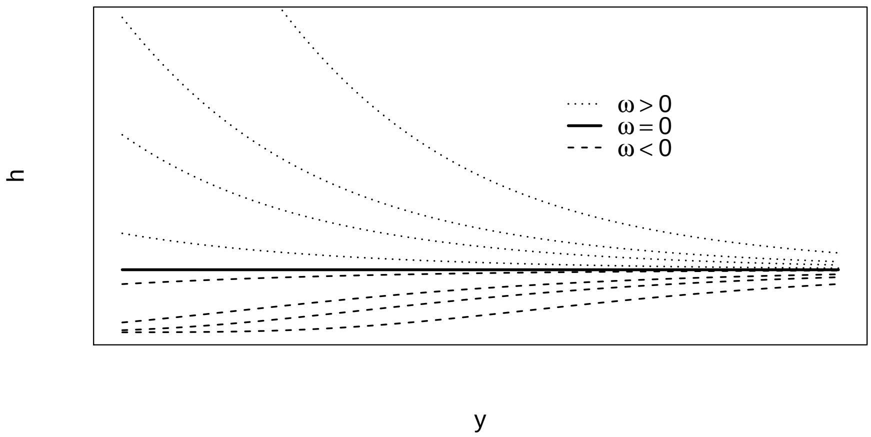


Figure 5.2: *The effect of ω in the hazard curves with $G(y)$ exponential.*

We can see from the plot that the hazard rate curves are all asymptote to the constant baseline when y is large. This result can be generalized as follows.

Proposition 5.2.1 *For all $\boldsymbol{\theta}$ and ω, we have that $h(y|\omega, \boldsymbol{\theta}) \sim h(y|\boldsymbol{\theta})$ when $y \to \infty$.*

Proof: The hazard rate function (5.5) can be written as

$$h(y|\boldsymbol{\theta}, \omega) = h(y|\boldsymbol{\theta}) \times R(y, \boldsymbol{\theta}) \quad \text{where} \quad R(y, \boldsymbol{\theta}) = \frac{\omega S(y|\boldsymbol{\theta})}{1 - e^{-\omega S(y|\boldsymbol{\theta})}}.$$

For any $\boldsymbol{\theta} \in \Theta$ and $\omega \in \mathbb{R}$, we have that $S(y|\boldsymbol{\theta}) \sim 0$ and $1 - \exp(-\omega\, S(y|\boldsymbol{\theta})) \sim 0$ when y is large, so an application of L'Hospital rule shows that $R(y, \boldsymbol{\theta}) \sim 1$, which proves the result. □

5.2.1 Related distributions

The modified negative Gompertz distribution (see Marshall and Olkin (2007b) page 390 and Dahiya and Hossain (1996)) can be obtained from the construction (5.2) in the particular case where $\mathcal{G}$ is the exponential distribution family with rate θ, that is, when $G(y|\theta) = 1 - \exp(-\theta\, y)$. It is interesting to note that we naturally allow the shape parameter ω to be negative but this does not seem to be acknowledged in the work of Marshall and Olkin (2007b) and Dahiya and Hossain (1996). The corresponding density can be written as

$$f(y|\omega, \theta) = \frac{\operatorname{sign}(\omega) e^{-\omega}}{\sigma(\theta)(1 - e^{-\omega})} \exp\left\{-\frac{y - \kappa(\omega, \theta)}{\sigma(\theta)} + \operatorname{sign}(\omega) \exp\left(-\frac{y - \kappa(\omega, \theta)}{\sigma(\theta)}\right)\right\}, \tag{5.6}$$

with $\omega \in \mathbb{R}$, $\kappa(\omega,\theta) = \log|\omega|/\theta$ and $\sigma(\omega) = 1/\theta$. Therefore, if $\omega < 0$, expression (5.6) shows that the e-extended model corresponds to a subset of truncated (to the positive real line) Gumbel distribution.

For a given value of $\boldsymbol{\theta} \in \Theta$, the densities (5.3) form an exponential family with natural parameter ω and sufficient statistic $G(y|\boldsymbol{\theta})$. This will clearly have important implications for making inference on the family $\mathcal{F}$, namely, for maximum likelihood. In fact, our construction can be seen as a particular case of a more general construction of the form

$$F(y|\omega,\boldsymbol{\theta}) = g(y|\boldsymbol{\theta}) \exp\left(\omega\, R(y|\boldsymbol{\theta}) - \psi(\omega,\boldsymbol{\theta})\right)$$

(see, for example, Rayner *et al.* (2009)), where our construction corresponds to the choice $R(y|\boldsymbol{\theta}) = -G(y|\boldsymbol{\theta})$ and as a consequence of that choice, the function $\psi(\omega,\boldsymbol{\theta}) = \log((1 - e^{-\omega})/\omega)$ does not depend on $\boldsymbol{\theta}$.

Another related construction is the so-called rank transmutation mapping introduced by Shaw and Buckley (2009). The authors define a rank transmutation map, in the continuous case, as $T(u) = F(G^{-1}(u))$ where F and G are continuous cumulative distribution functions with the same support and varying in different parametric families. The authors use this, and other related mappings, with copula-base simulation in mind, which we do not pursue here, but we can think of their construction as similar to ours in the following sense. Assume, for illustration purposes, that the functions F and G vary in parametric families so that the mapping T can be indexed by a parameter ω so that we have $T(u|\omega)$. If the identity mapping is included into consideration, that is, $T(u|\omega = 0) = u$ then clearly $T(G(y|\boldsymbol{\theta})|\omega)$ will generate a family that includes $\mathcal{G}$ as a particular case when $\omega = 0$. The simplest example is the quadratic transmutation

$$T(u|\omega) = u + \frac{\omega}{2}\, u(1-u).$$

Then we can construct a larger family via

$$F(y|\boldsymbol{\theta},\omega) := T(G(y|\boldsymbol{\theta})|\omega) = G(y|\boldsymbol{\theta}) + \frac{\omega}{2}\, G(y|\boldsymbol{\theta}) S(y|\boldsymbol{\theta}), \tag{5.7}$$

with the only difference that this instance requires that $\omega \in [-2, 2]$ since any cumulative distribution function is bounded between 0 and 1. There is a simple heuristic relationship between the two approaches which helps to understand the parametrizations used. We can write the densities corresponding to (5.7) as:

$$f(y|\boldsymbol{\theta},\omega) = f(y|\boldsymbol{\theta}) \left[1 + \omega \left(\frac{1}{2} - G(y|\boldsymbol{\theta})\right)\right]$$

For small values of $\omega(1 - G(y|\boldsymbol{\theta}))$ we have that

$$1 + \omega\left(\frac{1}{2} - G(y|\boldsymbol{\theta})\right) \approx \exp(1/2 - G(y|\boldsymbol{\theta}))$$

and a renormalization (in order to integrate to one) will arrive at the same expression as (5.3), hence the use of the same parameter ω in both (5.2) and (5.7).

There also is a relationship with the so-called local mixture models of Anaya-Izquierdo and Marriott (2007). In this case the extension of the densities is as follows

$$f(y|\boldsymbol{\theta}, \omega) = f(y|\boldsymbol{\theta})\left[1 + \omega\, R(y, \boldsymbol{\theta})\right],$$

where the function $R(y, \boldsymbol{\theta})$ requires that the expectation with respect to $f(y|\boldsymbol{\theta})$ is zero. In this case the choice $R(y, \boldsymbol{\theta}) = 1/2 - G(y|\boldsymbol{\theta})$ gives clearly the same distribution as (5.7).

Our construction has some advantages over the alternative construction in the sense of having a wider range of behaviours. Figure 5.3 shows the comparison between the e-extended and the quadratic rank transmuted model, respectively, by assuming that $G(y|\boldsymbol{\theta})$ is exponentially distributed. We can observe that the e-extended model has a wider range of shapes when the parameter $|\omega| \geq 2$ and a similar behaviour to the rank transmuted one when the parameter $|\omega| < 2$.

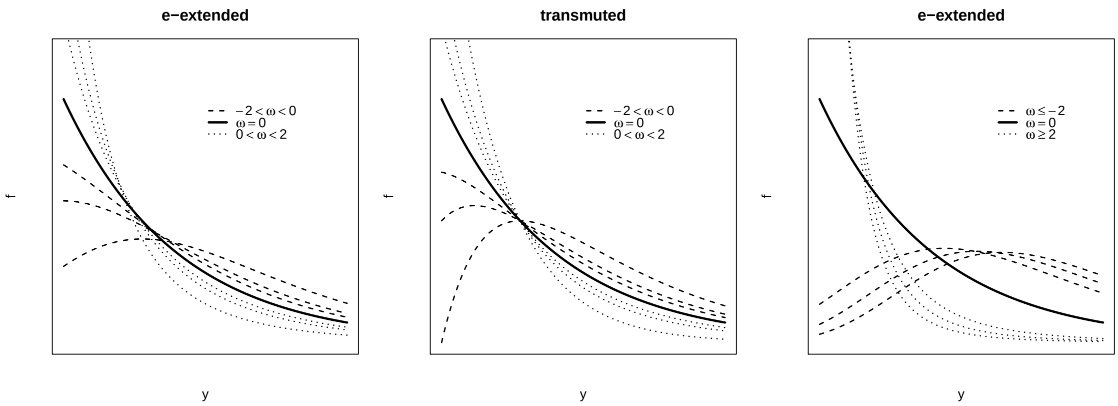


Figure 5.3: *The effect of ω in the density curves of the e-extended and transmuted families, with $G(y|\boldsymbol{\theta})$ exponential.*

5.3 E-extended Weibull

Given its central place in survival analysis, we focus here on e-extensions of the Weibull distribution. If Y is a positive random variable following a Weibull distribution then the probability density and cumulative distribution functions are given by

$$g(y|\phi,\beta) = \frac{\beta}{\phi^{\beta}} y^{\beta-1} e^{-(y/\phi)^{\beta}}, \qquad G(y|\phi,\beta) = 1 - e^{-(y/\phi)^{\beta}}, \tag{5.8}$$

where $\phi > 0$ and $\beta > 0$ are scale and shape parameters, respectively, so in this case we have that $\boldsymbol{\theta} = (\phi, \beta)$. The shape parameter β is linked to the hazard behavior: if $\beta \in [0, 1)$ the hazard is decreasing; if $\beta > 1$ the hazard is increasing and; the exponential distribution is a particular case if $\beta = 1$ and the hazard is constant. By using the expressions above the Weibull e-extended model density can be written as:

$$f(y|\phi,\beta,\omega) = \frac{\omega\beta e^{-\omega}}{(1-e^{-\omega})\phi^{\beta}} y^{\beta-1} \exp\left\{-\left(\frac{y}{\phi}\right)^{\beta} + \omega \exp\left[-\left(\frac{y}{\phi}\right)^{\beta}\right]\right\}, \tag{5.9}$$

with $\omega \in \mathbb{R}$ and $(\phi, \beta) \in \mathbb{R}_{+}^{2}$. Figure 5.4 shows some examples of probability density Weibull e-extended curves for different values of parameters β and ω, with fixed $\phi = 0.7$.

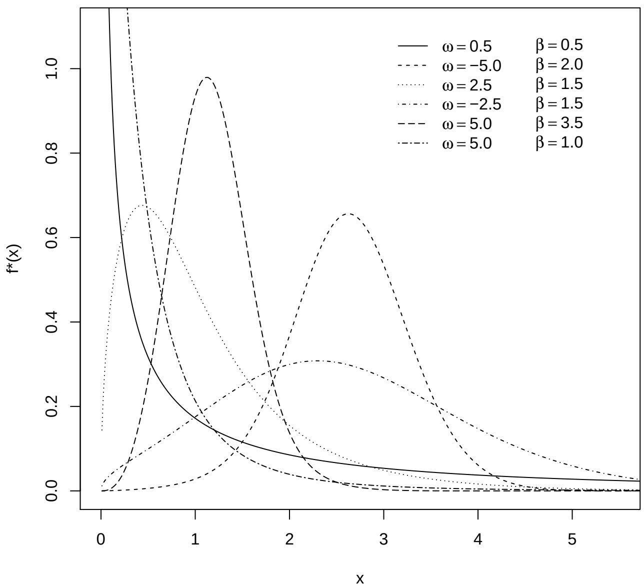


Figure 5.4: *Probability density and cumulative distribution function Weibull e-transmuted model, respectively, for different values of* ω *and* $\phi = 0.7$.

The corresponding hazard rate function is given by

$$h(y|\omega,\phi,\beta) = \frac{\omega\beta y^{\beta-1} e^{-(y/\phi)^{\beta}}}{e^{\phi\beta}\left[e^{-\gamma(\boldsymbol{y})e^{-(y/\phi)^{\beta}}} - 1\right]}. \tag{5.10}$$

There are three different behaviours of the hazard rate function, which are described next:

Case I: if the baseline model is Weibull with parameter $\beta > 1$, i.e., the base distribution hazard is increasing, see for example Figure 5.5-(a).

• For $\omega \leq 0$ the hazard curves continue being increasing but they are dominated by the Weibull baseline model;

• For $\omega > 0$ the hazard curves present two inflexion points, i.e., initially increase until the point t_1 when start to decrease up to t_2 and increase again for times greater than t_2 (the changes of curves behavior are evident for greater values of ω).

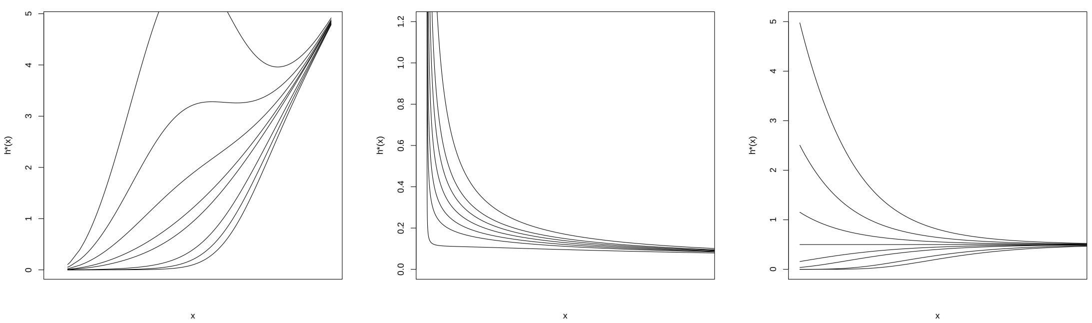
Case II: if the baseline model is Weibull with parameter $\beta < 1$, i.e. the baseline distribution hazard is decreasing, then the behavior of the Weibull e-extended hazard is the same for all ω, see for example Figure 5.5-(b).

Case III: if the baseline model is Weibull with parameter $\beta = 1$, i.e the baseline distribution is exponential and the hazard is constant, see for example Figure 5.5-(c).

• For $\omega > 0$ the hazard of Weibull e-extended model is decreasing;

• For $\omega < 0$ the hazard increasing.

Note that in all cases, as explained by Proposition 5.2.1, the e-extended hazard converges to the baseline hazard for large values of y.



(a) (b) (c)

Figure 5.5: *The effect of ω in the hazard curves with $G(y)$ Weibull(μ, β): (a) $\beta > 1$; (b) $\beta < 1$; and (c) $\beta = 1$.*

5.4 Parameter estimation

5.4.1 Underlying exponential family structure

One of the advantages of this choice is that we can found easily expressions for entries of the expected Fisher information.

This particular choice of R ($R(y|\boldsymbol{\theta}) = -G(y|\boldsymbol{\theta})$), presented in Section 5.2.1 will give orthogonality between the parameters $\boldsymbol{\theta}$ and ω at the underlying base model $\mathcal{G}$. Note that,

$$-\frac{\partial^2 \log f(y|\boldsymbol{\theta},\omega)}{\partial\boldsymbol{\theta}\,\partial\omega} = \frac{\partial}{\partial\boldsymbol{\theta}}G(y|\boldsymbol{\theta}).$$

Assuming the standard regularity conditions on the parametric family $\mathcal{G}$ we can differentiate under the integral sign so that

$$E\left[-\frac{\partial^2 \log f(Y|\boldsymbol{\theta},\omega)}{\partial\boldsymbol{\theta}\,\partial\omega}\right] = E\left[\frac{\partial}{\partial\boldsymbol{\theta}}G(Y|\boldsymbol{\theta})\right] = \frac{\partial}{\partial\boldsymbol{\theta}}\,E\left[G(Y|\boldsymbol{\theta})\right],$$

where the expectation is taken with respect to $f(y|\boldsymbol{\theta},\omega)$. At the baseline model ($\omega = 0$) we have that $f(y|\boldsymbol{\theta},\omega) = f(y|\boldsymbol{\theta})$ so that clearly $E[G(y|\boldsymbol{\theta})] = 1/2$ since, in that case, $G(Y|\boldsymbol{\theta})$ follows a uniform distribution on the unit interval. Therefore it follows that

$$\frac{\partial}{\partial\boldsymbol{\theta}}\,E\left[G(Y|\boldsymbol{\theta})\right] = 0.$$

The above result implies that maximum likelihood can proceed iteratively and also using the so-called profile likelihood. Fixing a value of $\boldsymbol{\theta}$ the MLE of ω is easily found since the resultant family is of exponential form.

Let $y_1, \ldots, y_n$ be a random sample of size n from an e-extended distribution. We will allow for unknown parameters in the baseline distribution G, namely $G(y_i|\boldsymbol{\theta})$.

Given the exponential family structure defined above, if we fix the values of $\boldsymbol{\theta}$ then we can find the corresponding maximum value of ω trivially by solving the non-linear equation

$$k(\omega) = \frac{\sum_{i=1}^{n} G(y_i|\boldsymbol{\theta})}{n} := \bar{G}(\boldsymbol{\theta}). \tag{5.11}$$

Now, suppose $\boldsymbol{X} = (X_1, \cdots, X_p)'$ a vector of covariates and consider the following relationship between the vector of covariates and the parameters $\gamma(\mathbf{x}) = \gamma_1 x_1 + \ldots + \gamma_p x_p$. Then the likelihood function is given by

$$L(\omega,\boldsymbol{\theta}|\boldsymbol{y}) = \prod_{i=1}^{n}\prod_{j=1}^{p} \frac{\gamma_j x_{ij}}{(1-e^{-\gamma_j x_{ij}})} g(y_i|\boldsymbol{\theta}) e^{-\gamma_j x_{ij} G(y_i|\boldsymbol{\theta})}. \tag{5.12}$$

Hence, the log-likelihood function becomes

$$\begin{aligned} l = \log L(\omega,\boldsymbol{\theta}|\boldsymbol{y}) \quad &= \quad \sum_{i=1}^{n}\sum_{j=1}^{p}\log\left(\frac{\gamma_j x_{ij}}{1-e^{-\gamma_j x_{ij}}}\right) + p\sum_{i=1}^{n}\log g(y_i|\boldsymbol{\theta}) \\ &\quad - \sum_{i=1}^{n}\sum_{j=1}^{p}\gamma_j x_{ij} G(y_i|\boldsymbol{\theta}). \end{aligned} \tag{5.13}$$

Therefore, the maximum likelihood estimates of $\boldsymbol{\gamma}$ and the vector of parameters $\boldsymbol{\theta}$, inherent to the baseline distribution, which maximize (5.13) must satisfy the following normal equations

$$\begin{aligned}
\frac{\partial l}{\partial \gamma_j} &= \sum_{i=1}^{n} x_i^* \sum_{j=1}^{p} \frac{(1 - e^{-\gamma_j x_{ij}} - \gamma_j x_{ij} e^{-\gamma_j x_{ij}})}{\gamma_j x_{ij}\,(1 - e^{-\gamma_j x_{ij}})} - \sum_{i=1}^{n} x_i^* G(y_i|\boldsymbol{\theta}) = 0, \\
\frac{\partial l}{\partial \theta_j} &= p \sum_{i=1}^{n} \frac{g'_{\theta_j}(y_i|\boldsymbol{\theta})}{g(y_i|\boldsymbol{\theta})} - \sum_{i=1}^{n} \sum_{j=1}^{p} \gamma_j x_{ij} \sum_{i=1}^{n} G'_{\theta_j}(y_i|\boldsymbol{\theta}) = 0,
\end{aligned}$$

with x_i^* the values of x_i fixed at the point j. Note that, the maximum likelihood estimator $(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\theta}})$ is obtained by solving the above nonlinear system of equations. Furthermore, the observed Fisher Information matrix is given in Appendix A.4.

In order to compute the standard error and asymptotic confidence interval we use the usual large sample approximation in which the maximum likelihood estimators of $\boldsymbol{\gamma}, \boldsymbol{\theta}$ can be treated as being approximately $(p+k)-$variate normal, p is the number of covariates and k the number of parameters of the baseline distribution. For example, as $n \to \infty$ the asymptotic distribution of the MLE $(\hat{\boldsymbol{\gamma_p}}, \hat{\theta_k})$, for $p = k = 1$, is given by,

$$\begin{pmatrix} \hat{\gamma} \\ \hat{\theta} \end{pmatrix} \sim N \left[\begin{pmatrix} \hat{\gamma} \\ \hat{\theta} \end{pmatrix}, \begin{pmatrix} \hat{V}_{11}, \hat{V}_{12} \\ \hat{V}_{21}, \hat{V}_{22} \end{pmatrix} \right], \tag{5.14}$$

with, $\hat{V}_{ij} = V_{ij}\mid_{\theta=\hat{\theta}}$ and it is determined by the inverse of Fisher information matrix.

Thereby, an approximate $100(1-\alpha)\%$ two sided confidence intervals for γ and θ are, respectively, given by

$$\hat{\gamma} \pm z_{\alpha/2}\sqrt{\hat{V}_{11}} \qquad \text{and} \qquad \hat{\theta} \pm z_{\alpha/2}\sqrt{\hat{V}_{22}}\ ,$$

where z_α is the upper $\alpha - th$ percentile of the standard normal distribution.

5.5 Weibull e-extended Model

5.5.1 Parameter estimation

The maximum likelihood estimates, MLEs, of the parameters of the Weibull regression e-transmuted probability distribution function are given by the following: Let $Y_1, \ldots, Y_n$ be a sample of size n from a Weibull e-transmuted distribution and $\boldsymbol{x} = (x_1, \ldots, x_p)'$ the vector of covariates with $\gamma(\mathbf{x}) = \gamma_1 x_1 + \ldots + \gamma_p x_p$. Then the likelihood function of the Weibull regression e-transmuted model is given by

$$\begin{aligned} L(\omega,\phi,\beta|\boldsymbol{y}) &= \frac{\beta^{np}}{\phi^{np\beta}}\prod_{i=1}^{n}\prod_{j=1}^{p}\left\{\left(\frac{\gamma_j x_{ij}}{1-e^{-\gamma_j x_{ij}}}\right)e^{-\gamma_j x_{ij}}y_i^{\beta-1}\right.\\ &\quad \left.\times\exp\left\{-\left(\frac{y_i}{\phi}\right)^{\beta}+\gamma_j x_{ij}\exp\left[-\left(\frac{y_i}{\phi}\right)^{\beta}\right]\right\}\right\}. \qquad (5.15)\end{aligned}$$

Hence, the log-likelihood function becomes

$$\begin{aligned} l=\log L(\omega,\phi,\beta|\boldsymbol{y}) &= np\log\beta-np\beta\log\phi+\sum_{i=1}^{n}\sum_{j=1}^{p}\log\left(\frac{\gamma_j x_{ij}}{1-e^{-\gamma_j x_{ij}}}\right)-\sum_{i=1}^{n}\sum_{j=1}^{p}\gamma_j x_{ij}\\ &\quad +(\beta-1)p\sum_{i=1}^{n}\log y_i-\sum_{i=1}^{n}\sum_{j=1}^{p}\left[\left(\frac{y_i}{\phi}\right)^{\beta}-\gamma_j x_{ij}e^{-(y_i/\phi)^{\beta}}\right]. \qquad (5.16)\end{aligned}$$

Therefore, the maximum likelihood estimates of γ_j, ϕ and β which maximize (5.16) must satisfy the following normal equations:

$$\begin{aligned} \frac{\partial l}{\partial\gamma_j} &= \sum_{i=1}^{n}x_i^*\sum_{j=1}^{p}\frac{(1-e^{-\gamma_j x_{ij}}-\gamma_j x_{ij}e^{-\gamma_j x_{ij}})}{\gamma_j x_{ij}\,(1-e^{-\gamma_j x_{ij}})}-\sum_{i=1}^{n}x_i^*+p\sum_{i=1}^{n}x_i^*e^{-(y_i/\phi)^{\beta}}=0,\\ \frac{\partial l}{\partial\phi} &= \frac{np\beta}{\phi}+\frac{p}{\phi}\sum_{i=1}^{n}\varphi_i+\sum_{i=1}^{n}\sum_{j=1}^{p}\frac{\gamma_j x_{ij}}{\phi}\varphi_i e^{-\varphi_i}=0,\\ \frac{\partial l}{\partial\beta} &= \frac{np}{\beta}-np\phi+p\sum_{i=1}^{n}\log y_i-\sum_{i=1}^{n}\sum_{j=1}^{p}\varphi_i\log\left(\frac{y_i}{\phi}\right)-\sum_{i=1}^{n}\sum_{j=1}^{p}\gamma_j x_{ij}\varphi_i\log\left(\frac{y_i}{\phi}\right)e^{-\varphi_i}=0,\end{aligned}$$

with $\varphi_i=\left(\frac{y_i}{\phi}\right)^{\beta}$ and x_i^* the values of the covariate x_i fixed at the point j.

We denote, the maximum likelihood estimator $(\hat{\gamma_j},\hat{\phi},\hat{\beta})$ and the observed Fisher Information is indicated in Appendix A.4.

In order to compute the standard error and asymptotic confidence intervals we use the usual large sample approximation in which the maximum likelihood estimators of γ_j,ϕ,β can be treated as being approximately $(p+2)-$variate normal. For example, hence as $n\to\infty$, fixing $j=1$, the asymptotic distribution of the MLE $(\hat{\gamma_1},\hat{\phi},\hat{\beta})$, are given by,

$$\begin{pmatrix}\hat{\gamma}_1\\ \hat{\phi}\\ \hat{\beta}\end{pmatrix}\sim N\left[\begin{pmatrix}\gamma_1\\ \phi\\ \beta\end{pmatrix}\begin{pmatrix}V_{11} & V_{12} & V_{13}\\ V_{21} & V_{22} & V_{23}\\ V_{31} & V_{32} & V_{33}\end{pmatrix}\right], \qquad (5.17)$$

with $\hat{V}_{ij}=V_{ij}\mid_{\theta=\hat{\theta}}$ and it is determined by the inverse of observed Fisher information matrix.

Therefore, an approximate $100(1-\alpha)\%$ two sided confidence intervals for γ_1, ϕ and θ are, respectively, given by

$$\hat{\gamma_1} \pm z_{\alpha/2}\sqrt{\hat{V}_{11}}, \qquad \hat{\phi} \pm z_{\alpha/2}\sqrt{\hat{V}_{22}} \text{ and } \qquad \hat{\beta} \pm z_{\alpha/2}\sqrt{\hat{V}_{33}},$$

where z_α is the upper $\alpha - th$ percentile of the standard normal distribution.

5.5.2 Numerical experiments for the Weibull e-transmuted model

This section presents the results of a Monte Carlo experiment to investigate the finite sample behavior of the MLEs. First of all, we consider the Weibull e-transmuted model without covariates and all results were obtained from 5000 Monte Carlo replications. The sample sizes n range from 50 to 1000, generated according to a Weibull e-transmuted distribution as presented in Section 5.3. For that, since ϕ is a location parameter the results of the Monte Carlo experiment are unaffected by the choice of ϕ. We fixed the value of parameter $\phi = 2$ and we consider different combinations of ω and β parameters values. We consider a wide range of hazard behaviors, namely $\beta = 0.5, 1.0$ and 1.5 (decreasing, constant and increasing ones, respectively). Also, we considered positive and negative ω's.

In order to check the coverage probability of the confidence intervals a boostrap study was made. For that, we generated 200 samples with sample sizes ranging from 50 to 500, generated according to a Weibull e-transmuted distribution, without covariates, and considering several values of parameters. We fixed the value of parameter $\phi = 2$ and considered different combinations of ω and β. For each of the 200 samples, we resampled $B = 1000$ and constructed the intervals by using the p-boostrap method; Gentle (2009). The coverage probability of these intervals is presented in Figure 5.6, for the parameters ω, μ and β, left, middle and right panels respectively, where we can see the closeness to the nominal value, 95%, in all simulated situations.

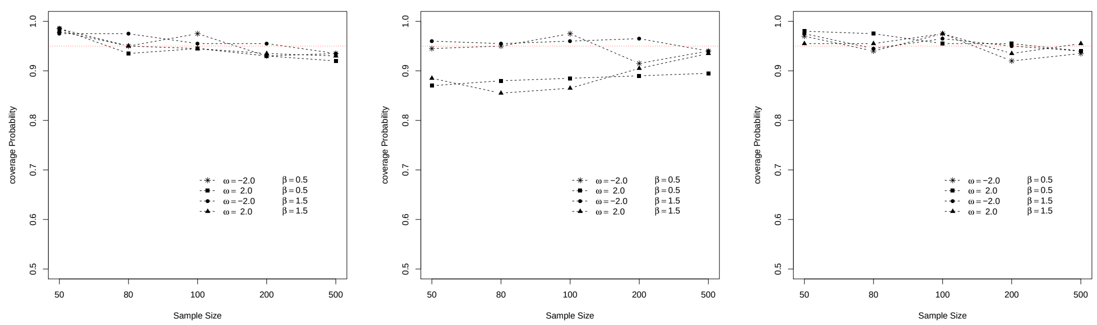


Figure 5.6: *Coverage probability of intervals for ω, μ and β parameters, respectively, generated by using the p-bootstrap method for $\phi = 2$ fixed and different values of ω and β.*

Now, we consider the Weibull regression e-transmuted model in the presence of two continuous covariates and all results were obtained from 5000 Monte Carlo replications.

Table 5.1: *Results of a Monte Carlo experiment for the Weibull e-transmuted model with* $\phi = 2$ *and different combinations of* ω *and* β *values. The estimated values presented are the median values of the* 5000 *generated samples. Within each horizontal block, the generated samples were subsampled from the larger one which was generated by using the inverse CDF method.*

Generated	Sample	Estimated				Standard Error			MSE		
$(\omega;\beta)$	Size	ω	ϕ	β	$-2\log$	ω	ϕ	β	ω	ϕ	β
$(-0.5;0.5)$	50	−0.409	2.047	0.513	338.2	1.900	1.057	0.104	3.556	1.696	0.116
	80	−0.451	2.027	0.506	542.3	1.624	0.923	0.088	3.029	1.461	0.098
	100	−0.444	2.033	0.507	679.5	1.516	0.858	0.082	2.785	1.338	0.091
	200	−0.467	2.016	0.502	1359.2	1.180	0.680	0.066	2.057	1.010	0.072
	500	−0.468	2.014	0.502	3404.0	0.832	0.483	0.049	1.311	0.661	0.052
	1000	−0.490	2.002	0.501	6803.6	0.629	0.368	0.038	0.904	0.472	0.040
$(0.5;0.5)$	50	0.390	1.917	0.501	280.2	1.911	1.131	0.090	3.599	1.770	0.100
	80	0.420	1.948	0.498	447.2	1.659	0.988	0.075	2.985	1.500	0.083
	100	0.468	1.961	0.499	559.4	1.540	0.925	0.070	2.710	1.380	0.076
	200	0.465	1.979	0.498	1121.7	1.185	0.711	0.054	1.970	1.003	0.057
	500	0.491	1.992	0.500	2805.6	0.800	0.479	0.037	1.125	0.609	0.038
	1000	0.477	1.987	0.499	5612.0	0.575	0.344	0.027	0.756	0.414	0.028
$(-2;1)$	50	−1.577	2.103	1.074	334.6	1.994	0.481	0.253	4.574	0.683	0.327
	80	−1.699	2.076	1.048	536.6	1.719	0.421	0.218	3.678	0.586	0.275
	100	−1.778	2.062	1.040	671.6	1.603	0.397	0.202	3.356	0.552	0.254
	200	−1.845	2.038	1.029	1346.2	1.258	0.317	0.162	2.493	0.434	0.196
	500	−1.904	2.024	1.015	3369.9	0.891	0.230	0.119	1.576	0.298	0.138
	1000	−1.953	2.012	1.009	6744.2	0.682	0.179	0.093	1.061	0.220	0.103
$(2;1)$	50	1.554	1.843	0.975	240.3	2.011	0.612	0.143	3.912	0.810	0.163
	80	1.671	1.887	0.978	385.2	1.748	0.533	0.115	3.328	0.693	0.126
	100	1.708	1.905	0.977	482.4	1.628	0.495	0.102	2.977	0.637	0.112
	200	1.854	1.946	0.985	966.8	1.274	0.386	0.073	2.128	0.478	0.077
	500	1.904	1.973	0.992	2422.8	0.865	0.262	0.046	1.266	0.305	0.048
	1000	1.943	1.984	0.996	4849.0	0.627	0.188	0.032	0.827	0.211	0.033
$(-0.5;1.5)$	50	−0.367	2.022	1.546	282.1	1.881	0.349	0.309	3.477	0.430	0.403
	80	−0.393	2.025	1.535	452.9	1.614	0.303	0.266	2.984	0.373	0.343
	100	−0.391	2.020	1.528	566.9	1.496	0.284	0.246	2.785	0.347	0.316
	200	−0.429	2.011	1.515	1136.4	1.185	0.227	0.199	2.095	0.273	0.248
	500	−0.461	2.005	1.504	2845.3	0.840	0.162	0.147	1.341	0.187	0.173
	1000	−0.486	2.001	1.499	5692.9	0.633	0.124	0.115	0.909	0.138	0.129
$(0.5;1.5)$	50	0.355	1.970	1.489	269.1	1.908	0.379	0.269	3.538	0.458	0.335
	80	0.406	1.978	1.489	431.5	1.645	0.329	0.227	3.033	0.395	0.271
	100	0.429	1.983	1.491	540.2	1.523	0.305	0.208	2.678	0.360	0.246
	200	0.452	1.988	1.493	1082.4	1.173	0.236	0.162	1.886	0.272	0.182
	500	0.480	1.995	1.498	2709.7	0.794	0.159	0.112	1.129	0.176	0.120
	1000	0.493	1.999	1.501	5424.6	0.570	0.114	0.081	0.739	0.123	0.085

Similar to the first numerical example, the sample sizes n range from 50 to 1000, generated according to a Weibull regression e-transmuted distribution as presented in Section 5.3. Again, we fixed the value of parameter $\phi = 2$ and we consider different combinations of γ_1, γ_2 (the regressors) and β parameters values. We consider a wide range of hazard behaviors, namely $\beta = 0.5, 1.0$ and 1.5 (decreasing, constant and increasing ones, respectively).

We checked the coverage probability of the confidence intervals for sample sizes ranging from 50 to 2000, generated according to a Weibull e-transmuted distribution, in the presence of two covariates, and considering several values of parameters as presented in Table 5.2. Similar to which was made in the first simulation study, we fixed the value of parameter $\phi = 2$ and considered different combinations of γ_1, γ_2 and β. The coverage probability of these intervals is presented in Figure 5.7, for the parameters γ_1 and γ_2 (upper panels), μ and β (lower panels), respectively, where we can see the closeness to the nominal value, 95%, in all simulated situations.

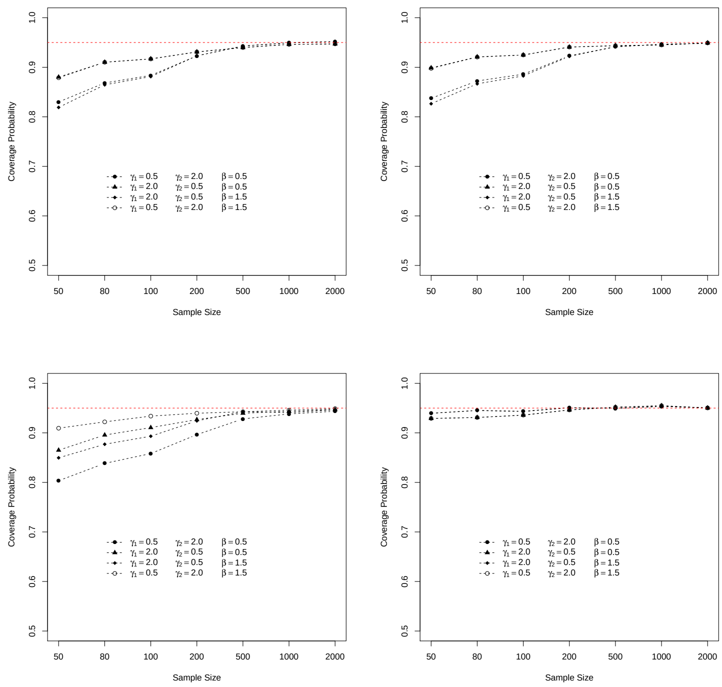


Figure 5.7: *Coverage probability of intervals for γ_1, γ_2, μ and β parameters, respectively, based on Monte Carlo experiment for $\phi = 2$ fixed and different values of γ_1, γ_2 and β.*

5.5.3 Goodness-of-fit to the Weibull regression e-transmuted model

One of the most complex problems in parametric statistical methods is the identification of the distribution that can best fit the data. In general words, the goodness-of-fit of a statistical model describes how well it fits into a set of observations. The goodness-of-fit indices summarize the discrepancy between the observed values and the values expected under a statistical model.

Also, the goodness-of-fit statistics are usually obtained using asymptotic methods, that are used in statistical hypothesis testing. As large sample approximations may behave poorly in small samples, a great deal of research using simulation studies has been devoted to investigate under which conditions the asymptotic p-values of goodness-of-fit statistics are accurate (i.e., how large the sample size must be for models of different sizes), Tollenaar and Mooijaart (2003).

Some general indices and statistics can be used in goodness-of-fit tests like Akaike information criteria, bayesian factor, Anderson-Darling, Cramér-von Mises, among others.

Table 5.2: *Results of a Monte Carlo experiment for the Weibull regression e-transmuted model with $\phi = 2$ and different combinations of γ_1, γ_2 and β values. The estimated values presented are the median values of the 5000 generated samples. Within each horizontal block, the generated samples were subsampled from the larger one which was generated by using the inverse CDF method.*

Generated	Sample	Estimated				Estimated Error				MSE			
$(\gamma_1;\gamma_2;\beta)$	Size	γ_1	γ_2	ϕ	β	γ_1	γ_2	ϕ	β	γ_1	γ_2	ϕ	β
$(0.5, 2, 0.5)$	50	0.524	2.018	1.924	0.509	1.169	0.972	1.579	0.055	1.969	3.346	2.450	0.056
	80	0.513	2.026	1.987	0.505	0.882	0.799	1.347	0.043	1.301	3.164	1.968	0.043
	100	0.505	2.018	1.986	0.505	0.772	0.721	1.224	0.038	1.085	3.061	1.747	0.039
	500	0.505	1.994	1.990	0.501	0.330	0.329	0.574	0.017	0.383	2.558	0.695	0.017
	1000	0.500	1.995	1.993	0.500	0.231	0.233	0.408	0.012	0.257	2.470	0.474	0.012
$(2, 0.5, 0.5)$	50	2.207	0.508	2.047	0.514	1.354	0.450	1.284	0.060	2.089	2.726	1.833	0.062
	80	2.095	0.512	2.059	0.509	1.035	0.349	1.048	0.047	1.490	2.573	1.386	0.048
	100	2.078	0.505	2.035	0.507	0.914	0.309	0.923	0.042	1.270	2.548	1.215	0.043
	500	2.009	0.499	2.000	0.501	0.398	0.136	0.409	0.018	0.470	2.389	0.478	0.019
	1000	2.001	0.502	1.997	0.501	0.280	0.096	0.291	0.013	0.315	2.340	0.324	0.013
$(0.5, 2, 1)$	50	0.525	2.020	1.963	1.018	1.106	0.908	0.764	0.109	1.971	3.350	1.170	0.115
	80	0.514	2.027	1.994	1.011	0.866	0.773	0.667	0.086	1.301	3.165	0.906	0.089
	100	0.505	2.018	1.993	1.010	0.762	0.707	0.608	0.076	1.084	3.061	0.801	0.079
	500	0.505	1.994	1.995	1.002	0.330	0.329	0.288	0.034	0.383	2.558	0.324	0.034
	1000	0.500	1.995	1.997	1.001	0.231	0.233	0.204	0.024	0.257	2.470	0.223	0.024
$(2, 0.5, 1)$	50	2.207	0.508	2.023	1.029	1.344	0.447	0.638	0.120	2.089	2.726	0.836	0.127
	80	2.095	0.512	2.029	1.018	1.033	0.349	0.513	0.094	1.490	2.573	0.630	0.098
	100	2.078	0.505	2.017	1.014	0.914	0.309	0.457	0.084	1.270	2.548	0.553	0.087
	500	2.009	0.499	2.000	1.001	0.398	0.136	0.205	0.037	0.470	2.389	0.225	0.038
	1000	2.001	0.502	1.999	1.002	0.280	0.096	0.145	0.026	0.315	2.340	0.155	0.026
$(0.5, 2, 1.5)$	50	0.527	2.019	1.975	1.527	1.109	0.911	0.528	0.164	1.970	3.348	0.738	0.176
	80	0.514	2.027	1.996	1.516	0.864	0.775	0.449	0.128	1.299	3.165	0.572	0.136
	100	0.505	2.018	1.995	1.515	0.764	0.708	0.409	0.114	1.084	3.061	0.509	0.120
	500	0.505	1.994	1.997	1.502	0.330	0.329	0.192	0.050	0.383	2.558	0.210	0.052
	1000	0.500	1.995	1.998	1.501	0.231	0.233	0.136	0.035	0.257	2.470	0.145	0.036
$(2, 0.5, 1.5)$	50	2.207	0.508	2.015	1.543	1.345	0.448	0.426	0.179	2.089	2.726	0.528	0.196
	80	2.095	0.512	2.020	1.527	1.034	0.349	0.341	0.140	1.490	2.573	0.402	0.150
	100	2.078	0.505	2.011	1.521	0.914	0.309	0.306	0.125	1.270	2.548	0.352	0.133
	500	2.009	0.499	2.000	1.502	0.398	0.136	0.137	0.055	0.470	2.389	0.147	0.057
	1000	2.001	0.502	1.999	1.502	0.280	0.096	0.097	0.039	0.315	2.340	0.102	0.040

Another ones are graphical based, for example the probability-probability and QQ plots; or quantify a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution, for example, the widely used Kolmogorov-Smirnov test. In this section, we will focused in two quadratic distance methods: Anderson-Darling and Cramér-von Mises and consider the following description of the test.

First, let $Y_1, \ldots, Y_n$ be random variables from a family of distribution and $F(\cdot)$ the cumulative distribution function. Then, consider the following hypothesis:

$$\begin{cases} H_0 : F(y|\boldsymbol{\theta}) \in \mathcal{F} = \left\{F(\cdot|\omega, \boldsymbol{\theta}) : \omega \in \mathbb{R}, \boldsymbol{\theta} \in \Theta^k\right\} \\ H_1 : F(y|\boldsymbol{\theta}) \notin \mathcal{F} = \left\{F(\cdot|\omega, \boldsymbol{\theta}) : \omega \in \mathbb{R}, \boldsymbol{\theta} \in \Theta^k\right\} \end{cases}. \tag{5.18}$$

It is equivalent to test the alternative hypothesis as being $H_1 : F(y) \neq F(x|\boldsymbol{\theta}), \forall \boldsymbol{\theta} \in \Theta$.

Let $\psi(\boldsymbol{Y})$ be a test statistic, in this particular case two that belong to the quadratic class of empirical distribution function statistics, Anderson-Darling and Cramér-von Mises statistics (see, for example, Anderson (1962) and Duchesne *et al.* (1997)). This class of statistics is based on the squared difference $(F_n(y) - F_0(y))^2$ and have the following general form:

$$\psi(\boldsymbol{Y}) = n \int_{-\infty}^{+\infty} \left(F_n(y) - F_(y|\hat{\boldsymbol{\theta}})\right)^2 \kappa(y) dF(y)$$

where $\kappa(\cdot)$ weights the squared difference and $F_n(\cdot)$ is the empirical distribution function and $F_0(\cdot)$ is the cumulative distribution function under H_0. The Anderson-Darling and the Cramér-von Mises statistics are defined, respectively, by

$$\psi_{AD}(\boldsymbol{Y}) = n \int_{-\infty}^{+\infty} \left(F_n(y) - F_(y|\hat{\boldsymbol{\theta}})\right)^2 \left[F_(y|\hat{\boldsymbol{\theta}}) \left(1 - F_(y|\hat{\boldsymbol{\theta}})\right)\right] dF(y) \tag{5.19}$$

and

$$\psi_{CvM}(\boldsymbol{Y}) = n \int_{-\infty}^{+\infty} \left(F_n(y) - F_(y|\hat{\boldsymbol{\theta}})\right)^2 dF(y). \tag{5.20}$$

The first interest is to investigate the type I error, which is the incorrect rejection of a true null hypothesis, i.e. $P\,(\text{Reject } H_0 | H_0 \text{ is true}) = \alpha_1$. For that, we need the sampling distribution of $\psi_{AD}(\boldsymbol{Y})$ and $\psi_{CvM}(\boldsymbol{Y})$ under the null hypothesis $F(x|\boldsymbol{\theta})$ for some $\theta \in \Theta$. Note that, the sampling distribution of $\psi_{AD}(\boldsymbol{Y})$ and $\psi_{CvM}(\boldsymbol{Y})$ depending on θ. In this case, we use the $\hat{\theta}_0$ and use the sampling distribution of $\psi_{AD}(\boldsymbol{Y})$ and $\psi_{CvM}(\boldsymbol{Y})$ when $\theta = \hat{\theta}_0$. Then, we simulated from $F(y|\hat{\boldsymbol{\theta}})$, i.e., from the Weibull e-transmuted model and the quantile $100(1-\alpha_1)\%$ of the Anderson-Darling and the Cramér-von misses statistics were calculated and given by $\psi_{AD}^{1-\alpha_1}(\boldsymbol{Y})$ and $\psi_{CvM}^{1-\alpha_1}(\boldsymbol{Y})$, respectively.

Then, the initial Monte Carlo experiment with 1000 samples was used to set the corresponding nominal values when $\alpha_1 = 5\%$ and 10% for different combinations of the parameters. In order to check the rejection rate behaviors, we use the non-parametric bootstrap method to resample $B = 1000$ times from Monte Carlo samples and the rejection rates are showed in the Figure 5.8. Note that, we analysed for different situations: decreasing and increasing hazards ($\beta = 0.5$ and 1.5, respectively) and ω parameter being negative and positive. From Figure 5.8, it is clear that in all situations the rejection rates are close to the nominal values.

Also, extensions to the theory of hypothesis testing include the study of the power of tests. Then, let's define the type II error as the probability of failure to reject a false null hypothesis, i.e., when the null hypothesis is false but erroneously fails to be rejected,

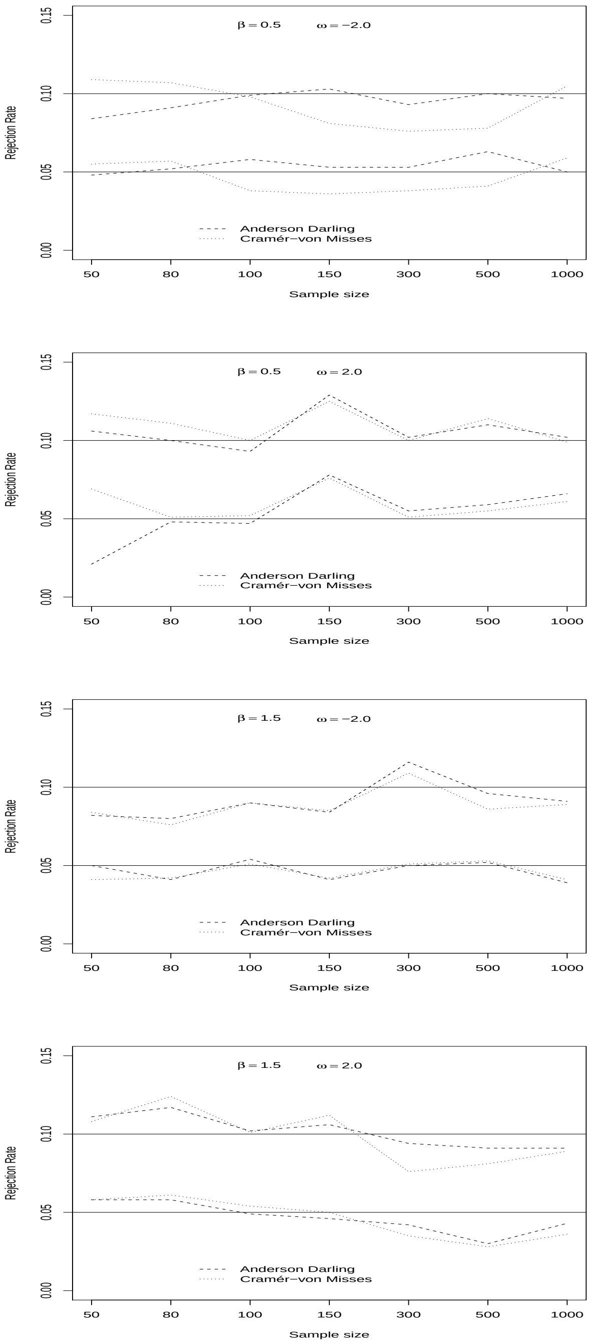


Figure 5.8: *Probability of error type I by using Anderson-Darling and Cramér-von-Mises statistics for different combinations of the parameters β and ω and fixed $\phi = 2$ by considering $\alpha_1 = 5\%$ and* 10% *.*

P(Not reject $H_0|H_0$ is false) $= 1 - \alpha_2$. Then the probability of correctly rejecting the null hypothesis given that it is false is the complement of the false negative rate, α_2.

In order to check the power of the test, initially, let's assume a more general test, presented in the composite hypothesis (5.18). As described before, a Monte Carlo experiment with 1000 samples was used to verify the sampling distribution of the tests statistics, for different combinations of the parameters. After, we used the parametric bootstrap method to resample $B = 1000$ times from Monte Carlo samples and the critical regions were defined. Now, we simulated by using the Monte Carlo method from the log-logistic distribution with location parameter $\phi = \hat{\phi}$, the MLE under the null hypothesis and $\beta = \beta^*$ varying as shown in Figure 5.9. Finally, 1000 $\psi_{AD}(\boldsymbol{Y})$ and $\psi_{CvM}(\boldsymbol{Y})$ were calculated for each β^* and the rejection rate was used to compute the power of the test as presented in Figure 5.9.

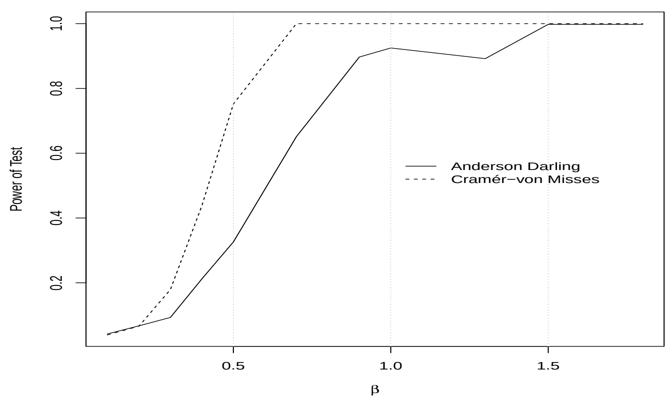


Figure 5.9: *Power of the tests Anderson-Darling and Cramér-von Misses for* $n = 100$ *and different values of* β.

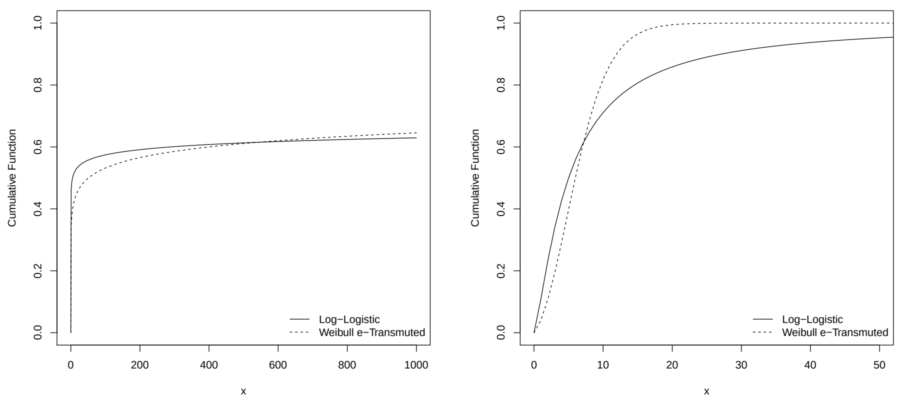


Figure 5.10: *Cumulative distribution curves for the log-logistic and Weibull e-Transmuted models when* $\beta = 0.1$ *and* 1.5*, respectively;* ω *and* ϕ *were set in the same values then estimated in power of the test experiment.*

In Figure 5.10 we show the cumulative distribution curves for the models log-logistic and Weibull e-transmuted. It is important to note that the left panel, when β is close to zero, the curves are close which explain the low Anderson-Darling and Cramér-von Mises power of tests. However, when β is more then 1, in left panel we set $\beta = 1.3$, the curves start having different shapes, which is consistent with the observed test power, close to 1.

Furthermore, a second study of power of the test was carried out. For this study, we consider the following directional simple hypothesis:

$$\begin{cases} H_0 : Y \sim \text{Weibull}(\phi, \beta) \\ H_1 : Y \sim \text{e-Transmuted Weibull}(\omega, \phi, \beta) \end{cases} \Rightarrow \begin{cases} H_0 : \omega = 0 \\ H_1 : \omega \neq 0 \end{cases}$$

In order to test these hypothesis, a Monte Carlo experiment with 1000 samples with size $n = 100$ was used to check the sampling distribution of the statistics Anderson-Darling and Cramér-von Mises, i.e. to obtain the points $\psi_{AD}^{1-\alpha_1}(\boldsymbol{Y})$ and $\psi_{CvM}^{1-\alpha_1}(\boldsymbol{Y})$. Then, for each sample, we used the bootstrap non-parametric method resampling $B = 1000$ times and, by using this estimates, we calculated the quadratic distances presented in equations (5.20) and (5.19). The results of this experiment are showed in Figure 5.11.

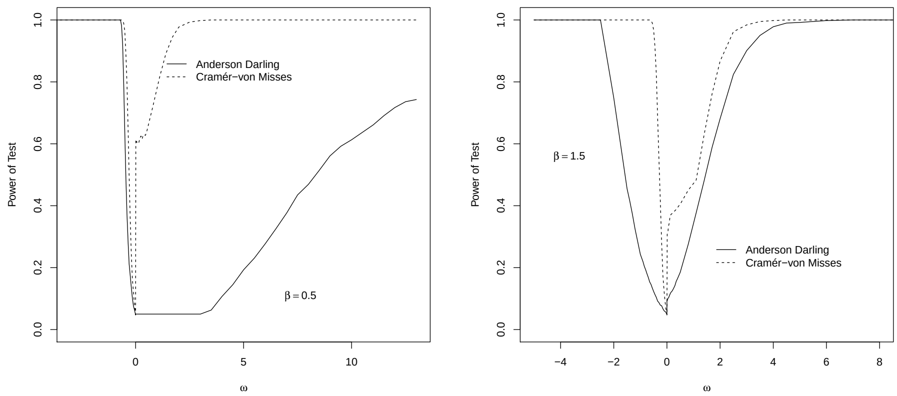


Figure 5.11: *Power of the tests Anderson-Darling and Cramér-von Mises with sample size $n = 100$ and for different combinations of the parameters β and ω and fixed $\phi = 2$ when $Y \sim$Weibull(ϕ, β).*

It is important to note that, when the parameter ω is close to zero, the power of the test is about 5%. The power of the test increases when ω moves away from zero positively or negatively. Also, note that the test is more sensible for negative ω's than for positive ones since the curves are asymmetric.

Chapter 6

Conclusions and perspectives

In this research, we propose a new generalization of the log-logistic, the transmuted log-logistic distribution. The proposed distribution is constructed by using a quadratic rank transmutation map and taking the log-logistic distribution with two parameters as the baseline distribution. We also propose the use of a new generalization of the log-logistic distribution, the transmuted log-logistic distribution, in a Bayesian context and using a regression approach.

We have provided closed expressions for several probabilistic measures including the probability density function, function hazard, moments, quantile function, mean, variance, and median.

For each presented model, a simulation study was performed, from which we learned that the maximum likelihood estimate bias and the standard error decrease when the sample size is increased. On the other hand, the coverage probability of 95% two sided confidence intervals for the model parameters becomes closer to the nominal ones as the sample size increases.

Furthermore, we proposed to analyze the behavior of the transmuted log-logistic model in the presence of censored observations. Note that this is the first time that this class of model (transmuted) was applied in the presence of censored data.

In order to fit the parameters of the model, profile methods were used and compared to each other: Pure, Adjusted and Modified profile. All profile methods can be used to fit this class of model in the presence of censors, but the Adjusted Profile was better when compared to others.

The applicability of the model was shown using three real datasets. The first one refers to a long term study involving 17026 cows of Tabapua breed studied by EMBRAPA. This dataset was used in order to show the usefulness of the proposed log-logistic transmuted model in a classical and Bayesian context, in the presence of covariates (regression approach) and by considering the cubic log-logistic model.

The second one refers to 148 patients treated with the drug Linezolid. This drug is used to treat serious infections, therefore, all patients were treated with Linezolid at an ICU (Intensive Care Unit), in the city of Maringá. Furthermore, an analysis of influential points, residuals and a bootstrap study were made in order to improve the fitting quality of the model and validate our results. This dataset was used in order to show the usefulness of the model in the presence of right censor times.

The last one corresponds to an uncensored study on the breaking stress of carbon fibers (in Gba), from Nichols and Padgett (2006) and was used to provide an example of applicability of the cubic Weibull transmuted model.

In all of these applications, we observed that the most probable time of the event of interest to occur could not be fitted by an usual model in survival analysis. This happens because the risk function has a high peak at its point of mode, which does not occur with any usual model.

In addition, to solve the problem of the restricted parametric space of λ parameter, we proposed a new family of distribution called e-transmuted or exponential transmuted family of models. Goodness-of-fit by using two quadratic distance measures were presented to validate the results to the e-transmuted regression Weibull model.

In order to continue this research, we are interested in studying the mixture transmuted models and its behavior in the presence of a latent variable. In parallel, we are interested in generalizing the ranking of transmutation maps proposing, for example, the k^{th} order of a ranking transmutation map.

Appendix A

Hessian matrix

A.1 The hessian matrix of transmuted log-logistic model of order 1

The Hessian matrix is given by

$$\mathbf{A} = \begin{pmatrix} A_{11}, A_{12}, A_{13} \\ A_{21}, A_{22}, A_{23} \\ A_{31}, A_{32}, A_{33} \end{pmatrix}$$

where

$$\begin{pmatrix} \hat{V}_{11}, \hat{V}_{12}, \hat{V}_{13} \\ \hat{V}_{21}, \hat{V}_{22}, \hat{V}_{23} \\ \hat{V}_{31}, \hat{V}_{32}, \hat{V}_{33} \end{pmatrix} = \begin{pmatrix} A_{11}, A_{12}, A_{13} \\ A_{21}, A_{22}, A_{23} \\ A_{31}, A_{32}, A_{33} \end{pmatrix}^{-1}$$

and

$$\begin{aligned} A_{11} = -\frac{\partial^2 l}{\partial \mu^2} \quad &= \quad -\sum_{i=1}^{n} \left(\frac{e^{\mu} x_i^{\beta} - \lambda e^{\mu} x_i^{\beta}}{1 + e^{\mu} x_i^{\beta} - \lambda e^{\mu} x_i^{\beta} + \lambda} \right) \left(1 - \frac{e^{\mu} x_i^{\beta} - \lambda e^{\mu} x_i^{\beta}}{1 + e^{\mu} x_i^{\beta} - \lambda e^{\mu} x_i^{\beta} + \lambda} \right) \\ &\quad +3 \sum_{i=1}^{n} \left(\frac{e^{\mu} x_i^{\beta}}{1 + e^{\mu} x_i^{\beta}} \right) \left(1 - \frac{e^{\mu} x_i^{\beta}}{1 + e^{\mu} x_i^{\beta}} \right) \end{aligned}$$

$$
\begin{aligned}
A_{12} &= A_{21} = -\frac{\partial^2 l}{\partial\mu\partial\beta} \\
&= -\sum_{i=1}^{n}\left(\frac{e^{\mu}x_i^{\beta}\ln x_i - \lambda e^{\mu}x_i^{\beta}\ln x_i}{1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda}\frac{\left(e^{\mu}x_i^{\beta}\ln x_i - \lambda e^{\mu}x_i^{\beta}\ln x_i\right)\left(e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}\right)}{\left(1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda\right)^2}\right) \\
&\quad +3\sum_{i=1}^{n}\frac{e^{\mu}x_i^{\beta}\ln x_i}{1+e^{\mu}x_i^{\beta}}\left(1-\frac{e^{\mu}x_i^{\beta}}{1+e^{\mu}x_i^{\beta}}\right)
\end{aligned}
$$

$$
\begin{aligned}
A_{13} &= A_{31} = -\frac{\partial^2 l}{\partial\mu\partial\lambda} \\
&= \sum_{i=1}^{n}\left(\frac{e^{\mu}x_i^{\beta}}{1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda}\right)\left(1+\frac{(1-\lambda)(1-e^{\mu}x_i^{\beta})}{1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda}\right)
\end{aligned}
$$

$$
\begin{aligned}
A_{22} = -\frac{\partial^2 l}{\partial\beta^2} &= \frac{n}{\beta^2}-\sum_{i=1}^{n}\left[\frac{e^{\mu}x_i^{\beta}\ln x_i^2 - \lambda e^{\mu}x_i^{\beta}\ln x_i^2}{1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda}-\left(\frac{e^{\mu}x_i^{\beta}\ln x_i - \lambda e^{\mu}x_i^{\beta}\ln x_i}{1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda}\right)^2\right] \\
&\quad +3\sum_{i=1}^{n}\frac{e^{\mu}x_i^{\beta}\ln x_i^2}{1+e^{\mu}x_i^{\beta}}\left(1-\frac{e^{\mu}x_i^{\beta}}{1+e^{\mu}x_i^{\beta}}\right)
\end{aligned}
$$

$$
\begin{aligned}
A_{23} &= A_{32} = -\frac{\partial^2 l}{\partial\beta\partial\lambda} \\
&= \sum_{i=1}^{n}\left(\frac{e^{\mu}x_i^{\beta}\ln x_i}{1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda}+\frac{(1-e^{\mu}x_i^{\beta})(e^{\mu}x_i^{\beta}\ln x_i-\lambda e^{\mu}x_i^{\beta}\ln x_i)}{(1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda)^2}\right)
\end{aligned}
$$

$$
A_{33} = -\frac{\partial^2 l}{\partial\lambda^2} = x\sum_{i=1}^{n}\left(\frac{1-e^{\mu}x_i^{\beta}}{1+e^{\mu}x_i^{\beta}-\lambda e^{\mu}x_i^{\beta}+\lambda}\right)^2
$$

A.2 The hessian matrix of transmuted Weibull of order 2

The Hessian matrix is given by

$$
\mathbf{A} = \begin{pmatrix} A_{11}, A_{12}, A_{13}, A_{14} \\ A_{21}, A_{22}, A_{23}, A_{24} \\ A_{31}, A_{32}, A_{33}, A_{34} \\ A_{41}, A_{42}, A_{43}, A_{44} \end{pmatrix}
$$

where

$$\begin{pmatrix} \hat{V}_{11}, \hat{V}_{12}, \hat{V}_{13}, \hat{V}_{14} \\ \hat{V}_{21}, \hat{V}_{22}, \hat{V}_{23}, \hat{V}_{24} \\ \hat{V}_{31}, \hat{V}_{32}, \hat{V}_{33}, \hat{V}_{34} \\ \hat{V}_{41}, \hat{V}_{42}, \hat{V}_{43}, \hat{V}_{44} \end{pmatrix} = \begin{pmatrix} A_{11}, A_{12}, A_{13}, A_{14} \\ A_{21}, A_{22}, A_{23}, A_{24} \\ A_{31}, A_{32}, A_{33}, A_{34} \\ A_{41}, A_{42}, A_{43}, A_{44} \end{pmatrix}^{-1}$$

and

$$\begin{aligned} A_{11} &= -\frac{n}{\mu^2} + 2(\lambda_1 + 2\lambda_2 - 3)\sum_{i=1}^{n}\frac{k_i}{\omega_i}\left[S_i^\mu \log S_i - \log S_i - 2(\lambda_1 + 2\lambda_2 - 3)k_i\right] + \\ &\quad -6(1-\lambda_2)\sum_{i=1}^{n} k_i\left[Z_i^\mu \log S_i + 2k_i\right] \end{aligned}$$

$$\begin{aligned} A_{12} = A_{21} &= -\frac{n}{\beta} + \frac{1}{\beta}\sum_{i=1}^{n}\left[\mu S_i^\mu \log S_i + S_i^\mu\right] + \\ &\quad + \frac{2(\lambda_1 + 2\lambda_2 - 3)}{\beta}\sum_{i=1}^{n}\left[\frac{k_i}{\omega_i}\left(\mu - \mu S_i^\mu + S_i^\mu Z_i\right) + S_i^\mu Z_i\right] + \\ &\quad \frac{6(1-\lambda_2)}{\beta}\sum_{i=1}^{n} S_i^\mu Z_i\left[1 + \mu Z_i \log S_i - 2\mu S_i^\mu Z_i \log S_i\right] \end{aligned}$$

$$A_{13} = A_{31} = 2\sum_{i=1}^{n}\frac{k_i}{\omega_i}\left[(\lambda_1 + 2\lambda_2 - 3)\frac{(2Z_i - 1)}{\omega_i} - 1\right]$$

$$A_{14} = A_{41} = 2\sum_{i=1}^{n}\frac{k_i}{S_i}\left[(\lambda_1 + 2\lambda_2 - 3)\frac{(4Z_i - 1)}{\omega_i} - 2\right] + 6\sum_{i=1}^{n} k_i Z_i$$

$$\begin{aligned} A_{22} &= \frac{n\mu}{\beta^2} + 2\mu\beta^2(\lambda_1 + 2\lambda_2 - 3)\frac{S_i^\mu Z_i}{\omega_i}\left[\mu S_i^\mu - 2(\lambda_1 + 2\lambda_2 - 3)\frac{S_i^\mu Z_i}{\omega_i} - \mu - 1\right] + \\ &\quad +6(1-\lambda_2)\frac{\mu}{\beta}\sum_{i=1}^{n} S_i^\mu Z_i^2\left[2\mu S_i^\mu - \mu - 1\right] - \mu(\mu+1)\sum_{i=1}^{n}\frac{S_i^\mu}{\beta^2} \end{aligned}$$

$$A_{23} = A_{32} = 2\frac{\mu}{\beta}\sum_{i=1}^{n}\frac{S_i^\mu Z_i}{\omega_i}\left[1 - 2(\lambda_1 + 2\lambda_2 - 3)\frac{(2Z_i - 1)}{\omega_i}\right]$$

$$A_{24} = A_{42} = 2\frac{\mu}{\beta}\sum_{i=1}^{n}\frac{S_i^{\mu}Z_i}{\omega_i}\left[1 - 2(\lambda_1 + 2\lambda_2 - 3)\frac{(4Z_i - 1)}{\omega_i}\right] - 6\frac{\mu}{\beta}S_i^{\mu}Z_i^2$$

$$A_{33} = -\sum_{i=1}^{n}\frac{(2Z_i - 1)^2}{\omega_i^2}$$

$$A_{34} = A_{43} = -\sum_{i=1}^{n}\frac{(2Z_i - 1)(1 + 4Z_i)}{\omega_i^2}$$

$$A_{44} = -\sum_{i=1}^{n}\frac{(4Z_i - 1)^2}{\omega_i^2}$$

where $Z_i = e^{-(x_i/\beta)^{\mu}}$, $S_i = \frac{x_i}{\beta}$ and $k_i = \left(\frac{x_i}{\beta}\right)^{\mu}\log\left(\frac{x_i}{\beta}\right)e^{-(x_i/\beta)^{\mu}}$.

A.3 The hessian matrix of transmuted log-logistic of order 2

The Hessian matrix is given by

$$\mathbf{A} = \begin{pmatrix} A_{11}, A_{12}, A_{13}, A_{14} \\ A_{21}, A_{22}, A_{23}, A_{24} \\ A_{31}, A_{32}, A_{33}, A_{34} \\ A_{41}, A_{42}, A_{43}, A_{44} \end{pmatrix}$$

where

$$\begin{pmatrix} \hat{V}_{11}, \hat{V}_{12}, \hat{V}_{13}, \hat{V}_{14} \\ \hat{V}_{21}, \hat{V}_{22}, \hat{V}_{23}, \hat{V}_{24} \\ \hat{V}_{31}, \hat{V}_{32}, \hat{V}_{33}, \hat{V}_{34} \\ \hat{V}_{41}, \hat{V}_{42}, \hat{V}_{43}, \hat{V}_{44} \end{pmatrix} = \begin{pmatrix} A_{11}, A_{12}, A_{13}, A_{14} \\ A_{21}, A_{22}, A_{23}, A_{24} \\ A_{31}, A_{32}, A_{33}, A_{34} \\ A_{41}, A_{42}, A_{43}, A_{44} \end{pmatrix}^{-1}$$

and

$$
\begin{aligned}
A_{11} &= 4\sum_{i=1}^{n}\frac{\eta_i}{(1+\eta_i)}\left[\frac{\eta_i}{(1+\eta_i)-1}\right]+2\sum_{i=1}^{n}\eta_i\left\{\frac{\lambda_2+\eta_i(6-2\lambda_2-2\lambda_1)}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1}+\right.\\
&\left.-\frac{(\lambda_2+\eta_i(3-\lambda_2-\lambda_1))^2}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}\right\}
\end{aligned}
$$

$$
\begin{aligned}
A_{12}=A_{21} &= 4\sum_{i=1}^{n}\log x_i\frac{\eta_i}{(1+\eta_i)}\left[\frac{\eta_i}{(1+\eta_i)-1}\right]+\\
&+2\sum_{i=1}^{n}\eta_i\log x_i\left[\frac{\lambda_2+\eta_i(6-2\lambda_2-2\lambda_1)}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1}\right]+\\
&-\sum_{i=1}^{n}\eta_i\frac{(\lambda_2+\eta_i(3-\lambda_2-\lambda_1))(1+\log x_i)}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}
\end{aligned}
$$

$$
A_{13}=A_{31}=\sum_{i=1}^{n}\left[\frac{\eta_i(\lambda_2+\eta_i(3-\lambda_2-\lambda_1))(\eta_i-1)}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}-\frac{2\eta_i^2}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1}\right]
$$

$$
A_{14}=A_{41}=\sum_{i=1}^{n}\left[\frac{\eta_i^2(\lambda_2+\eta_i(3-\lambda_2-\lambda_1))(\eta_i-2)}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}-\frac{2\eta_i^2(1-\eta_i)}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1}\right]
$$

$$
\begin{aligned}
A_{22} &= 4\sum_{i=1}^{n}(\log x_i)^2\frac{\eta_i}{(1+\eta_i)}\left[\frac{\eta_i}{(1+\eta_i)-1}\right]-\frac{n}{\beta}+\\
&+2\sum_{i=1}^{n}\eta_i\left\{\frac{(\log x_i)^2\left[\lambda_2+\eta_i(6-2\lambda_2-2\lambda_1)\right]}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1}+\right.\\
&\left.-\frac{\log x_i(\lambda_2+\eta_i(3-\lambda_2-\lambda_1))^2}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}\right\}
\end{aligned}
$$

$$
\begin{aligned}
A_{23}=A_{32} &= -\sum_{i=1}^{n}\left\{\frac{2\eta_i\log x_i(\lambda_2-\lambda_1\eta_i-\lambda_2\eta_i+3\eta_i)(1-\eta_i^2)}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}+\right.\\
&\left.+\frac{2\eta_i^2\log x_i}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1}\right\}
\end{aligned}
$$

$$A_{24} = A_{42} = -\sum_{i=1}^{n} \left\{ \frac{2\eta_i \log x_i(\lambda_2 - \lambda_1\eta_i - \lambda_2\eta_i + 3\eta_i)[\eta_i(2-\eta_i)]^2}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2} + -\frac{2\eta_i \log x_i(1-\eta_i)}{\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1} \right\}$$

$$A_{33} = -\sum_{i=1}^{n} \frac{(1-\eta_i^2)^2}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}$$

$$A_{34} = A_{43} = -\sum_{i=1}^{n} \frac{(1-\eta_i^2)(\eta_i(2-\eta_i))^2}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}$$

$$A_{44} = -\sum_{i=1}^{n} \frac{(\eta_i(2-\eta_i))^2}{\left[\eta_i^2(3-\lambda_1-\lambda_2)+2\eta_i+\lambda_1\right]^2}$$

where $\eta_i = e^{\mu} x_i^{\beta}$.

A.4 Hessian matrix of e-transmuted model

In order to present the hessian matrix, we consider the model e-transmuted without covariates as follows:

$$\begin{bmatrix} I_{11} & \cdots & I_{1(p+k)} \\ & \ddots & \vdots \\ & & I_{(p+k)(p+k)} \end{bmatrix}_{(p+k)\times(p+k)} \tag{A.1}$$

with p the number of covariates and k the number of parameters of base distribution and the elements I as follow:

$$I_{11} = \frac{\partial^2 l}{\partial \omega^2} = \frac{n\left[\omega^2 e^{-\omega} - e^{-2\omega} + 2e^{-\omega} - 1\right]}{\omega^2\left(1-e^{-\omega}\right)^2}$$

$$I_{1j} = \frac{\partial^2 l}{\partial \omega \partial \theta_j} = -\sum_{i=1}^{n} G'_{\theta_j}(y_i|\boldsymbol{\theta})$$

$$I_{kj} = \frac{\partial^2 l}{\partial \theta_j \partial \theta_k} = -\sum_{i=1}^{n} \frac{g''_{\theta_j\theta_k}(y_i|\boldsymbol{\theta})g(y_i|\boldsymbol{\theta}) - g'_{\theta_j}(y_i|\boldsymbol{\theta})g'_{\theta_k}(y_i|\boldsymbol{\theta})}{g^2(y_i|\boldsymbol{\theta})} - \omega\sum_{i=1}^{n} G''_{\theta_j\theta_k}(y_i|\boldsymbol{\theta})$$

$$I_{jj} = \frac{\partial^2 l}{\partial \theta_j^2} = -\sum_{i=1}^{n} \frac{\left[g'_{\theta_j}(y_i|\boldsymbol{\theta})\right]^2}{g^2(y_i|\boldsymbol{\theta})} + \sum_{i=1}^{n} \frac{g''_{\theta_j}(y_i|\boldsymbol{\theta})}{g(y_i|\boldsymbol{\theta})} + \omega \sum_{i=1}^{n} G''_{\theta_j}(y_i|\boldsymbol{\theta}).$$

Appendix B

Tables

Table B.1: *MLEs for the original sample generated with different parameter values and sample sizes.*

Parameters Generated	Sample Size	Estimated Relative Difference(%)				Estimated Standard Error			
$(\gamma_1, \gamma_2, \beta, \lambda)$	n	γ_1	γ_2	β	λ	γ_1	γ_2	β	λ
$(-2, -1, 0.5, -0.8)$	50	11.136	20.499	3.644	12.094	0.707	0.634	0.070	0.527
	100	7.479	14.098	2.158	7.631	0.518	0.463	0.053	0.372
	150	6.029	10.908	1.175	5.175	0.431	0.379	0.044	0.291
	300	4.068	6.587	0.747	3.721	0.320	0.278	0.034	0.197
	500	2.430	3.919	0.270	1.781	0.248	0.213	0.027	0.142
	1000	1.244	2.397	0.228	0.987	0.173	0.149	0.019	0.096
$(-2, -0.5, 1.5, -0.5)$	50	6.667	29.552	0.051	15.737	0.727	0.643	0.210	0.598
	100	4.237	21.604	0.860	2.789	0.535	0.466	0.155	0.437
	150	4.479	19.244	0.609	4.435	0.465	0.400	0.134	0.386
	300	3.155	10.868	0.466	1.911	0.347	0.294	0.098	0.291
	500	2.710	8.546	0.202	4.239	0.282	0.237	0.079	0.240
	1000	1.516	5.446	0.133	3.468	0.204	0.169	0.057	0.173
$(-2, 0.5, 1, -0.5)$	50	6.570	27.020	0.015	14.991	0.724	0.623	0.140	0.595
	100	4.188	21.382	0.880	3.046	0.534	0.447	0.103	0.438
	150	4.341	18.216	0.657	3.862	0.464	0.377	0.089	0.386
	300	3.155	10.491	0.466	1.911	0.347	0.274	0.065	0.291
	500	2.710	7.654	0.202	4.239	0.282	0.217	0.053	0.240
	1000	1.516	4.722	0.133	3.468	0.204	0.152	0.038	0.173
$(-3, -1, 0.5, 0.8)$	50	2.204	0.718	3.115	20.596	0.677	0.592	0.069	0.491
	100	1.009	0.442	1.984	14.506	0.471	0.416	0.052	0.364
	150	1.031	0.397	1.570	10.668	0.378	0.333	0.046	0.294
	300	0.829	0.079	0.646	6.547	0.263	0.231	0.033	0.197
	500	0.730	0.094	0.284	4.466	0.201	0.176	0.026	0.147
	1000	0.423	0.430	0.087	2.725	0.141	0.123	0.019	0.098
$(-2, -0.5, 1.5, 0.5)$	50	2.453	1.889	0.544	27.549	0.653	0.618	0.207	0.538
	100	2.413	1.101	0.610	30.890	0.473	0.455	0.156	0.454
	150	1.615	2.003	0.328	29.270	0.386	0.379	0.134	0.403
	300	1.139	2.257	0.479	19.601	0.272	0.273	0.098	0.310

	500	0.300	1.557	0.330	13.072	0.207	0.213	0.080	0.248
	1000	0.056	0.477	0.181	7.248	0.144	0.152	0.058	0.179
$(-3, 0.5, 1, 0.5)$	50	3.048	0.123	0.512	28.512	0.713	0.631	0.138	0.536
	100	1.663	1.233	0.596	31.439	0.514	0.471	0.104	0.455
	150	1.112	1.809	0.349	29.053	0.417	0.400	0.089	0.403
	300	0.645	0.515	0.479	19.601	0.292	0.293	0.066	0.310
	500	0.137	0.336	0.342	12.917	0.220	0.233	0.053	0.248
	1000	0.015	0.247	0.181	7.248	0.152	0.168	0.039	0.179

Table B.2: *Probability of coverage by considering* 95% *of confidence.*

Generated Parameters	Sample Size	Coverage Probability			
$(\gamma_1, \gamma_2, \beta, \lambda)$	n	γ_1	γ_2	β	λ
$(-2, -1, 0.5, -0.8)$	50	0.916	0.914	0.941	0.498
	100	0.923	0.927	0.934	0.640
	150	0.911	0.925	0.911	0.721
	300	0.925	0.944	0.915	0.849
	500	0.938	0.940	0.930	0.888
	1000	0.951	0.947	0.954	0.930
$(-2, -0.5, 1.5, -0.5)$	50	0.914	0.909	0.937	0.634
	100	0.920	0.908	0.919	0.716
	150	0.900	0.913	0.922	0.792
	300	0.899	0.931	0.924	0.838
	500	0.916	0.937	0.932	0.889
	1000	0.940	0.952	0.930	0.921
$(-2, 0.5, 1, -0.5)$	50	0.913	0.919	0.932	0.626
	100	0.921	0.918	0.916	0.715
	150	0.902	0.922	0.920	0.790
	300	0.899	0.939	0.924	0.838
	500	0.916	0.950	0.932	0.889
	1000	0.940	0.955	0.930	0.921
$(-3, -1, 0.5, 0.8)$	50	0.933	0.934	0.933	0.494
	100	0.952	0.942	0.924	0.674
	150	0.945	0.959	0.925	0.763
	300	0.952	0.956	0.923	0.841
	500	0.940	0.950	0.925	0.894
	1000	0.947	0.957	0.930	0.916
$(-2, -0.5, 1.5, 0.5)$	50	0.918	0.915	0.918	0.625
	100	0.931	0.917	0.935	0.742
	150	0.932	0.927	0.934	0.788

	300	0.955	0.937	0.925	0.843
	500	0.952	0.941	0.929	0.894
	1000	0.944	0.951	0.932	0.929
$(-3, 0.5, 1, 0.5)$	50	0.917	0.922	0.919	0.625
	100	0.937	0.919	0.934	0.743
	150	0.938	0.915	0.933	0.786
	300	0.953	0.913	0.925	0.843
	500	0.948	0.933	0.929	0.894
	1000	0.947	0.955	0.932	0.929

References

Akaike, H. (1973). Information theory and the maximum likelihood principle. *International Simposium on Information Theory, eds. V. Petrov and F.* 15, 59

Anaya-Izquierdo, K. and Marriott, P. (2007). Local mixture models of exponential families. *Bernoulli*, pages 623–640. 72

Anderson, T. W. (1962). On the distribution of the two-sample cramer-von mises criterion. *The Annals of Mathematical Statistics*, pages 1148–1159. 82

Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1992). *A first course in order statistics*, volume 54. Siam. 6

Aryal, G. R. and Tsokos, C. P. (2009). On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods & Applications*, **71**(12), e1401–e1407. 1, 6

Aryal, G. R. and Tsokos, C. P. (2011). Transmuted weibull distribution: A generalization of theweibull probability distribution. *European Journal of Pure and Applied Mathematics*, **4**(2), 89–102. 1, 12, 14, 52

Barlow, R. E. and Campo, R. A. (1975). Total time on test processes and applications to failure data analysis. Technical report, DTIC Document. 17, 19

Barlow, W. E. and Prentice, R. L. (1988). Residuals for relative risk regression. *Biometrika*, **75**(1), 65–74. 41, 62

Barndorfe-Nielsen, O. E. and Cox, D. R. (1994). *Inference and asymptotics*. Chapman & Hall. 38

Barndorff-Nielsen, O. E. (1993). On a formula for the distribution of the maximum likelihood estimator. *Biometrika*. 41

Barndorff-Nielsen, O. E. and McCullagh, P. (1993). A note on the relation between modified profile likelihood and the cox-reid adjusted profile likelihood. *Biometrika*, **80**(2), 321–328. 2, 40

Bennett, S. (1983). Log-logistic regression models for survival data. *Applied Statistics*, pages 165–171. 2

Brickner, S. J. (1996). Oxazolidinone antibacterial agents. *Current Pharmaceutical Design 2*, **2**, 175–194. 43

Burnham, K. P. and Anderson, D. R. (2002). *Model selection and multimodel inference: a practical information- theoretic approach, Second Edition*. Springer Series in Statistics. 20

Chechile, R. A. (2003). Mathematical tools for hazard function analysis. *Journal of Mathematical Psychology*, **47**, 478–494. 11

Chen, M. h. and Ibrahim, J. G. (2006). The relationship between the power prior and hierarchical models. *Bayesian Anal.*, **1**(3), 551–574. 24

Cox, D. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*. 40

Cox, D. R. and Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society*, **30**, 248–275. 21

Dahiya, R. C. and Hossain, S. A. (1996). A modification of goel-okumoto model. pages 77–84. 70

Daniels, M. J. (1999). A prior for the variance in hierarchical models. 24

Duchesne, T., Rioux, J., and Luong, A. (1997). Minimum cramér-von mises distance methods for complete and grouped data. *Communications in statistics-theory and methods*, **26**(2), 401–420. 82

Fachini, B. J., Ortega, E. M. M., and Louzada, F. (2008). Influence diagnostics for polyhazard models in the presence of covariates. *Stat Meth Appl*, **17**, 413–433. 2

Ferrari, S. L. P., Silva, M. F., and Cribari-Neto, F. (2007). Adjusted profile likelihoods for the Weibull shape parameter. *Journal of Statistical Computation and Simulation*. 2

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Communications in Statistics. Theory and Methods*, **1**(3), 515–533. 2, 24, 25

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, (7), 457–511. 26

Gentle, J. (2009). *Computational Statistics*. Statistics and Computing. Springer New York. 78

Ghitany, M. E. (2001). A compound rayleigh survival model and its application to randomly censored data. *Statistical Paper, Springer*, **42**, 437–450. 1

Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2005). *Bayesian survival analysis*. Wiley Online Library. 1, 6

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1996). *Continuous univariate distributions*, volume 2. Wiley New York. 52

Lawless, J. F. (2011). *Statistical models and methods for lifetime data*, volume 362. John Wiley & Sons. 2, 10, 18, 38, 52

Lee, S. Y., Lu, B., and Song, X. Y. (2006). Assessing local influence for nonlinear structural equation models with ignorable missing data. *Computational Statistics & Data Analysis*. 31, 48

Marshall, A. and Olkin, I. (2007a). *Life distributions: structure of nonparametric, semi-parametric and parametric families*. Springer, New York. 1

Marshall, A. and Olkin, I. (2007b). *Life Distributions: Structure of Nonparametric, Semi-parametric, and Parametric Families*. Springer Series in Statistics. Springer New York. 70

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. McGraw Hill, London. 41, 62

Migon, H. S., Gamerman, D., and Louzada, F. (2014). *Statistical inference: an integrated approach*. CRC press. 58

Mood, A. M., Graybill, F. A., and Boes, D. C. (1974). *Introduction to the Theory of Statistics*. McGraw Hill, 3 edition. 7

Nichols, M. and Padgett, W. (2006). A Bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, **22**, 141–151. 52, 59, 87

Ortega, E. M., Bolfarine, H., and Paula, G. A. (2003). Influence diagnostics in generalized log-gamma regression models. *Computational Statistics and Data Analysis*, **42**, 165–186. 2

Paro, P. a. Z., Santos, A. L. Q., Maximiniano-Neto, A., Paro, J. L. N., Rodrigues, D. C., Cruz, G. C., Malta, T. S., Ribeiro, F. M., and Andrade, M. A. (2013). Anatomic study of the vascular casts of the testicular arteries in bovines of Tabapua race. *Bioscience Journal*, **19**(1), 123–132. 16, 17

Pereira, J. C. C. (2000). Contribuição genética do Zebu na pecuária bovina do Brasil. *Informe Agropecuário*, **21**, 30–38. 16, 17

Pham, H. and Lai, C.-D. (2007). On Recent Generalizations of the Weibull Distribution. *IEEE Transactions on RELIABILITY*, **56**(3), 454–458. 52

Polson, N. G. and Scott, J. G. (2012). On the half-Cauchy prior for a global scale parameter. *Bayesian Anal.*, **7**(4), 887–902. 2, 24

Rayner, J. C., Thas, O., and Best, D. J. (2009). *Smooth tests of goodness of fit: using R*. John Wiley & Sons. 71

Ross, S. M. (2009). *A First Course in Probability*. Prentice Hall. 9

Severini, T. (1998). An approximation to the modified profile likelihood function. *Biometrika*. 2, 41

Shaw, W. T. and Buckley, I. R. C. (2009). The alchemy of probability distributions: beyond gram-charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint arXiv:0901.0434*. 1, 5, 6, 53, 71

Therneau, T. M., Grambsch, P. M., and Fleming, T. R. (1990). Martingale-based residuals for survival models. *Biometrika*. 41, 42, 62

Tollenaar, N. and Mooijaart, A. (2003). Type i errors and power of the parametric bootstrap goodness-of-fit test: Full and limited information. *British Journal of Mathematical and Statistical Psychology*, **56**(2), 271–288. 80

Venzon, D. J. and Moolgavkar, S. H. (1988). A Method for Computing Profile-Likelihood Based Confidence Intervals. *Applied Statistics*. 38

Verdinelli, I., Wasserman, L., *et al.* (1998). Bayesian goodness-of-fit testing using infinite-dimensional exponential families. *The Annals of Statistics*, **26**(4), 1215–1241. 69

Wilson, A. P. R., Cepeda, J. A., Hayman, S., Whitehouse, T., Singer, M., and Bellingan, G. (2006). In vitro susceptibility of Gram-positive pathogens to linezolid and teicoplanin and effect on outcome in critically ill patients. *Journal of Antimicrobial Chemotherapy*. 43