

**UNIVERSIDADE DE SÃO PAULO**

Instituto de Ciências Matemáticas e de Computação

**Lambert-F Univariate Distributions for Asymmetrical Data**

**Yuri Antonio Iriarte Salinas**

Tese de Doutorado do Programa Interinstitucional de Pós-Graduação em Estatística (PIPGEs)



SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: \_\_\_\_\_

**Yuri Antonio Iriarte Salinas**

## Lambert-F Univariate Distributions for Asymmetrical Data

Dissertation submitted to the Institute of Mathematics and Computer Science – ICMC-USP and to the Department of Statistics – DEs-UFSCar – in accordance with the requirements of the Statistics Interagency Graduate Program, for the degree of Doctor in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Mário de Castro Andrade Filho

**USP – São Carlos**  
**February 2022**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi  
e Seção Técnica de Informática, ICMC/USP,  
com os dados inseridos pelo(a) autor(a)

I681 Iriarte Salinas, Yuri Antonio  
Lambert-F Univariate Distributions for  
Asymmetrical Data / Yuri Antonio Iriarte Salinas;  
orientador Mário de Castro Andrade Filho. -- São  
Carlos, 2021.  
86 p.

Tese (Doutorado - Programa Interinstitucional de  
Pós-graduação em Estatística) -- Instituto de Ciências  
Matemáticas e de Computação, Universidade de São  
Paulo, 2021.

1. Distribuições de probabilidade. 2. Função W de  
Lambert . I. de Castro Andrade Filho, Mário ,  
orient. II. Título.

**Yuri Antonio Iriarte Salinas**

**Distribuições Univariadas de Lambert-F para Dados  
Assimétricos**

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

Área de Concentração: Estatística

Orientador: Prof. Dr. Mário de Castro Andrade Filho

**USP – São Carlos**  
**Fevereiro de 2022**





# UNIVERSIDADE FEDERAL DE SÃO CARLOS

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

---

## Folha de Aprovação

---

Defesa de Tese de Doutorado do candidato Yuri Antonio Iriarte Salinas, realizada em 16/12/2021.

### Comissão Julgadora:

Prof. Dr. Mário de Castro Andrade Filho (USP)

Prof. Dr. Josemar Rodrigues (USP)

Prof. Dr. Filidor Edilfonso Vilca Labra (UNICAMP)

Prof. Dr. Fredy Walther Castellares Cáceres (UFMG)

Profa. Dra. Maria do Carmo Soares de Lima (UFPE)

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

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa Interinstitucional de Pós-Graduação em Estatística.





# ACKNOWLEDGEMENTS

---

---

I want to use this space to sincerely thank the Institute of Mathematical and Computing Sciences of the University of Sao Paulo (ICMC - USP), and the Department of Statistics of the Federal University of Sao Carlos (DEs - UFSCar), for the opportunity to expand my knowledge and for providing the ideal environment for the development of this dissertation.

Heartfelt thanks to my supervisor, Mário de Castro Andrade Filho, for sharing his knowledge and for his innumerable suggestions given in developing this dissertation. Likewise, I want to thank Professor Héctor W. Gómez for the valuable suggestions provided in the formulation and development of this dissertation.

I also want to thank in a very special way the professors Josemar Rodrigues, Fredy Castellares, Maria do Carmo and Filidor Vilca who provided suggestions that helped to significantly improve this dissertation.

This work was funded by CONICYT PAI/INDUSTRIA 79090016, Chile.



# RESUMO

IRIATE, Y. A. **Distribuições Univariadas de Lambert-F para Dados Assimétricos**. 2022. 103 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

Nesta tese, propomos novas distribuições contínuas univariadas para modelar dados assimétricos. Inicialmente, partindo de uma transformação paramétrica não linear de uma variável aleatória uniforme, propomos uma nova distribuição assimétrica de um parâmetro que estende a distribuição uniforme, a chamada distribuição Lambert-uniforme. A transformação é expressa analiticamente em termos do ramo principal da função Lambert  $W$  de tal forma que a transformação inversa é expressa em termos de uma função exponencial. Conseqüentemente, a função densidade da distribuição Lambert-uniforme tem uma forma fechada simples e exhibe um comportamento monótono crescente ou decrescente. Posteriormente, com base na distribuição Lambert-uniforme, propomos um novo gerador de distribuição que permite adicionar um parâmetro de forma a uma distribuição de referência arbitrária. O parâmetro adicionado permite uma variedade de formas para a função de densidade da distribuição resultante, levando a uma expansão dos intervalos de assimetria e curtose da distribuição de referência. Observamos que o parâmetro induzido pelo gerador atua como parâmetro de assimetria quando a distribuição de referência é simétrica. Por outro lado, quando a distribuição de referência tem suporte positivo, observamos que a função taxa de risco da distribuição resultante corresponde a uma modificação nos tempos iniciais da função taxa de risco da distribuição de referência. Isso é exemplificado através do estudo de quatro casos especiais obtidos considerando as distribuições bimodal generalizada, slash, exponencial e Rayleigh como distribuições de referência. Discutimos a estimação de parâmetros pelo método de máxima verossimilhança e avaliamos o comportamento dos estimadores por meio de experimentos de simulação. Finalmente, consideramos alguns exemplos de aplicação que ilustram a utilidade das distribuições propostas em diferentes ambientes reais.

**Palavras-chave:** Distribuição de referência, Estimador de máxima verossimilhança, Função densidade de probabilidade, Função de taxa de risco, Função Lambert  $W$ , Gerador de distribuição, Parâmetro de forma.



# ABSTRACT

IRIATE, Y. A. **Lambert-F Univariate Distributions for Asymmetrical Data**. 2022. 103 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

In this dissertation, we propose new univariate continuous distributions for modeling asymmetrical data. Initially, starting from a non-linear parametric transformation of a uniform random variable, we propose a new asymmetric one-parameter distribution that extends the uniform distribution, the so-called Lambert-uniform distribution. The transformation is expressed analytically in terms of the principal branch of the Lambert  $W$  function in such a way that the inverse transformation is expressed in terms of an exponential function. Consequently, the density function of the Lambert-uniform distribution has a simple closed form and exhibits increasing or decreasing monotonic behavior. Subsequently, based on the Lambert-uniform distribution, we propose a new distribution generator that allows adding one shape parameter to an arbitrary baseline distribution. The added parameter allows a variety of shapes for the density function of the resulting distribution, leading to an expansion of the skewness and kurtosis ranges of the baseline distribution. We observe that the parameter induced by the generator acts as a skewness parameter when the baseline distribution is symmetric. On the other hand, when the baseline distribution has positive support, we observe that the hazard rate function of the resulting distribution corresponds to a modification in the early times of the hazard rate function of the baseline distribution. This is exemplified through the study of four special cases obtained by considering the generalized-bimodal, slash, exponential and Rayleigh distributions as baseline distributions. We discuss the parameter estimation via the maximum likelihood method and evaluate the behavior of the estimators through simulation experiments. Finally, we consider some application examples that illustrate the usefulness of the proposed distributions in different real settings.

**Keywords:** Baseline distribution, Probability density function, Distribution generator, Hazard rate function, Lambert  $W$  function, Maximum likelihood estimator, Shape parameter.



# LIST OF FIGURES

---

---

Figure 1	– Probability density function of the LU distribution. . . . .	23
Figure 2	– Probability density function of the LU distribution. . . . .	29
Figure 3	– Skewness and kurtosis coefficients of the LU distribution . . . . .	31
Figure 4	– Summary measures for 1000 ML estimates of the parameter of the LU distribution obtained from simulated samples . . . . .	35
Figure 5	– The AE, SD, SE and RMSE for each of the 1000 estimates of the coefficients $\beta$ s obtained in Scenario A, under the different sample size. . . . .	36
Figure 6	– The AE, SD, SE and RMSE for each of the 1000 estimates of the coefficients $\beta$ s obtained in Scenario B, under the different sample size. . . . .	37
Figure 7	– The CPs for the estimates of the coefficients $\beta$ s in: Scenario A ( <b>left</b> ); and Scenario B ( <b>right</b> ). . . . .	37
Figure 8	– QQ-plots for LU, K, B, MOEU, P and SU distributions fitted to peak horizontal acceleration data . . . . .	40
Figure 9	– Histogram of the peak horizontal acceleration data fitted with the LU, K, B, MOEU, P and SU distributions . . . . .	40
Figure 10	– Coefficient estimates and its 95% confidence intervals for variables AS, CA, SI, IN, CE and SO in different LU quantile regression models considering $q = 0.1, 0.2, \dots, 0.9$ and response variable FI. . . . .	43
Figure 11	– Density and hazard rate functions of the LE and LR distributions. . . . .	51
Figure 12	– Skewness and kurtosis coefficients for the LE and LR distributions. . . . .	55
Figure 13	– TTT-plot for survival times of individuals diagnosed with apparently benign monoclonal gammopathy. . . . .	61
Figure 14	– Histogram of the violent crime rate data with the fitted density curves and the survival curves estimated by Kaplan-Meier and with the LE distribution for individuals diagnosed with monoclonal gammopathy. . . . .	62
Figure 15	– Density function of the LGB distribution. . . . .	67
Figure 16	– Density function of the LSL distribution. . . . .	67
Figure 17	– Critical and inflection points of the density function of the LGB distribution. . . . .	69
Figure 18	– Unimodality and bimodality regions for LGB distribution. . . . .	69
Figure 19	– Critical and inflection points of the density function of the LSL distribution. . . . .	70
Figure 20	– Skewness and kurtosis coefficients of the LGB and LSL distributions. . . . .	71
Figure 21	– Skewness coefficients for the LGB and LSL distributions for $\gamma > 2$ and $\kappa > 5$ . . . . .	74
Figure 22	– Pdf curves for some aliases of the LGB distribution. . . . .	74

Figure 23 – Pdf curves for some aliases of the LSL distribution. . . . . 75

Figure 24 – Histograms of the inflation rate and real per-capita income data with the fitted density curves. . . . . 80



# LIST OF TABLES

---



---

Table 1 – Fit measurements for peak horizontal acceleration data. . . . .	39
Table 2 – The $\ell$ , AIC, CAIC and BIC values for the ASHN, K and LU quantile regression models fitted to the risk managements practice data and the $p$ -values of the AD and CvM tests for the randomize residuals. . . . .	42
Table 3 – Coefficient estimates for the LU quantile regression model fitted to the risk managements practice data and significance tests of individual regression coefficients. . . . .	43
Table 4 – Comparison of the LE, LR, W and G distributions in terms of the probability density function. . . . .	53
Table 5 – Comparison of the LE, LR, W and G distributions in terms of the hazard rate function. . . . .	54
Table 6 – Scenarios considered in the simulation studies for some Lambert- $F$ distributions.	57
Table 7 – Summary measures obtained in the simulation study for the LE distribution. . . . .	58
Table 8 – Summary measures obtained in the simulation study for the LR distribution. . . . .	59
Table 9 – Maximum likelihood estimates for distributions fitted to different real data sets	62
Table 10 – Kullback-Leibler divergence measures for some LGB family members. . . . .	75
Table 11 – Kullback-Leibler divergence measures for some LSL family members. . . . .	76
Table 12 – Scenarios considered in the simulation studies for some Lambert- $F$ distributions.	76
Table 13 – Summary measures obtained in the simulation study for the LGB distribution.	78
Table 14 – Summary measures obtained in the simulation study for the LSL distribution.	79
Table 15 – Excess mass tests and some descriptive statistics for the inflation rate, per-capita income and violent crime rate data. . . . .	80
Table 16 – Maximum likelihood estimates for distributions fitted to the real data sets . . . . .	81
Table 17 – Modified Cramer-von Mises and Anderson-Darling statistics and AIC and BIC values for distributions fitted to different real data sets . . . . .	81
Table 18 – Probability density function and cumulative distribution function for the GB, SL, E and R distributions. . . . .	95
Table 19 – Partial derivatives for the E baseline distribution. . . . .	99
Table 20 – Partial derivatives for the R baseline distribution. . . . .	99
Table 21 – Partial derivatives for the GB baseline distribution. . . . .	100
Table 22 – Partial derivatives for the SL baseline distribution. . . . .	101



# LIST OF ABBREVIATIONS AND ACRONYMS

---

---

ASHN	arc-secant-hyperbolic-normal
B	beta
cdf	cumulative distribution function
GB	generalized bimodal
GE	generalized exponential
GR	generalized Rayleigh
GSC	gamma sinh-Cauchy
GSL	generalized slash
hrf	hazard rate function
K	Kumaraswamy
ML	maximum likelihood
MN	mixture-normal
MOEU	Marshall-Olkin extended uniform
MSSL	modified skew-slash
OLLSN	odd log-logistic skew-normal
P	power
pdf	probability density function
qf	quantile function
rv	random variable
SL	slash
SSL	skew-slash
SU	skew-uniform



# CONTENTS

---

---

1	INTRODUCTION . . . . .	21
2	LAMBERT-UNIFORM DISTRIBUTION . . . . .	25
2.1	Motivation . . . . .	25
2.2	Lambert-Uniform Random Variable . . . . .	26
2.3	Characterization of Skewness and Kurtosis . . . . .	29
2.4	Maximum Likelihood Estimation . . . . .	30
2.5	Quantile Regression Model . . . . .	32
2.5.1	<i>The LU Model</i> . . . . .	32
2.5.2	<i>ML Estimation</i> . . . . .	33
2.6	Simulation Studies . . . . .	34
2.6.1	<i>First Simulation Study</i> . . . . .	34
2.6.2	<i>Second Simulation Study</i> . . . . .	34
2.7	Data Analysis . . . . .	36
2.7.1	<i>Peak Horizontal Acceleration Data</i> . . . . .	38
2.7.2	<i>Risk Managements Practice Data</i> . . . . .	41
2.8	Final Comments . . . . .	42
3	LAMBERT-F DISTRIBUTION GENERATOR . . . . .	45
3.1	Motivation . . . . .	45
3.2	The Generator . . . . .	46
3.3	Shapes . . . . .	47
3.4	Moments . . . . .	48
3.5	Quantile Function . . . . .	49
3.6	Stochastic Order . . . . .	50
3.7	Two Special Cases of Positive Supports . . . . .	50
3.8	Maximum Likelihood Estimation . . . . .	54
3.9	Simulation Studies . . . . .	57
3.10	Data Analysis . . . . .	60
3.11	Final Comments . . . . .	61
4	TWO UNBOUNDED-SUPPORT LAMBERT-F DISTRIBUTIONS . . . . .	65
4.1	Motivation . . . . .	65
4.2	Lambert-Generalized Bimodal and Lambert-Slash Distributions . . . . .	66

4.3	Alias Distribution	72
4.4	Considerations on the Maximum Likelihood Estimate and Simulations Studies	76
4.5	Data Analysis	77
4.6	Final Comments	82
5	CONCLUDING REMARKS	83
6	FUTURE STUDIES	85
6.1	Lambert-F Quantile Regression Models	85
6.2	Lambert Transformed U distributions	86
6.3	Alternative Distribution Generators	87
BIBLIOGRAPHY		89
APPENDIX A	BASELINE FUNCTIONS FOR THE SPECIAL CASES GIVEN IN SECTIONS 3.7 AND 4.2	95
APPENDIX B	MAXIMUM LIKELIHOOD ESTIMATION FOR THE LAMBERT-F DISTRIBUTION	97
APPENDIX C	FUNCTIONS FOR THE COMPUTATION OF THE SCORE FUNCTIONS AND THE OBSERVED INFORMATION MATRIX OF THE SPECIAL CASES REPORTED IN SECTIONS 3.7 AND 4.2	99
APPENDIX D	SECOND PARTIAL DERIVATIVES OF THE LOG-LIKELIHOOD FUNCTION GIVEN IN EQUATION (2.11)	103

---

# INTRODUCTION

---

Univariate probability distributions play an important role in data analysis in various fields of knowledge, such as economics, biology, medicine, epidemiology, engineering, among others. For an analyst, having a wide range of distributions at his disposal turns out to be an extremely favorable scenario, since through statistical tests it is possible to decide which is the probability distribution that best represents the theoretical behavior of the phenomenon under analysis.

An important variety of probability distributions can be found in the statistical literature. These distributions are usually classified under different standard criteria, according to the discrete or continuous nature of the corresponding random variable (rv), for example. In the continuous case, a distribution is also usually classified according to the type of support it has (bounded or unbounded) or according to the shapes exhibited by the probability density function (pdf) (Monotonic, unimodal, etc.), or by the hazard rate function (hrf) in the case of a lifetime distribution. Although the variety of distributions available in the literature is considerable, the constant increase in the information to be analyzed in the modern globalized world requires the use of new distributions to make an appropriate description of many of the phenomena or experiments studied.

In this context, several studies proposing methods for the construction of univariate continuous distributions have been developed. In the case of lifetime distributions, a good summary of these methods can be found in [Lai \(2013\)](#). Looking through the history of mathematical statistics, it is possible to highlight two very popular approaches to constructing univariate continuous distributions. The first consists of applying a transformation on a rv to obtain a new distribution with a transformed support. The well-known log-normal distribution is a good example of this approach considering that it can be derived from the exponential functional transformation of a normal rv. The second approach is based on the principle of adding parameters to a baseline distribution to obtain a new distribution with a larger parameter dimension, but which inherits the support of the baseline distribution. In this case, the new distribution is usually generated by

parametrically transforming the pdf, the cumulative distribution function (cdf), or the quantile function (qf) of a particular baseline distribution. The new parameters allow a variety of new shapes for the resulting pdf, which leads to a widening of the skewness and kurtosis ranges of the baseline distribution. Consequently, in the analysis of certain random phenomena, direct inferences about the data through new parameters open the possibility of an improvement in the analyzes. Furthermore, the possibility of linking these new parameters with a set of covariates through a regression type framework is an important justification for the development of studies in this line.

This last approach is very popular in the literature because it has allowed the generation of several asymmetric unbounded-support distributions that are derived from classical symmetric distributions (See [Azzalini \(1985\)](#), [Eugene, Lee and Famoye \(2002\)](#), [Ferreira and Steel \(2006\)](#) and [Cordeiro and de Castro \(2011\)](#), among others), which has favored the adequate description of the theoretical behavior of random phenomena that present an intrinsic characteristic of skewness evidenced in the sample data. A general approach that allows constructing univariate continuous distributions can be found in [Alzaatreh, Lee and Famoye \(2013\)](#).

A function that plays an important role in this dissertation is the Lambert  $W$  function, which is defined as the inverse function of  $f(z) = ze^z$ ,  $z \in \mathbb{C}$ . By restricting  $z$  to be a real number, this function is defined for  $z \geq -1/e$ , where  $e$  is the Euler's number. In this case, it is possible to distinguish three cases: If  $z < -1/e$ , then no solution exists in the reals; If  $z \in (-1/e, 0)$ , then there are two solutions given by the principal branch,  $W_0(z)$ , and the non-principal branch,  $W_{-1}(z)$ ; If  $z \geq 0$ , then the solution is unique,  $W_0(z) = W_{-1}(z)$ .

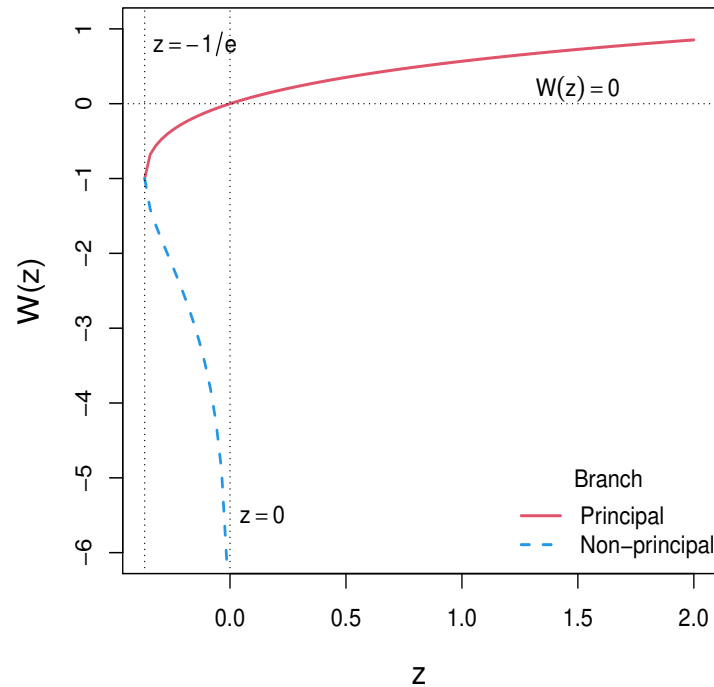
Figure 1 illustrates the behavior of the branches of the Lambert  $W$  function, where it can be seen that the principal branch exhibits an increasing monotonic behavior. As we will see in later chapters, the results of this dissertation are largely based on the use of the principal branch of the Lambert  $W$  function.

The Lambert  $W$  function has applications in many areas, in materials science, chemical engineering, thermodynamics, and statistics, to name a few. In statistics, its use is wide. For example, it has been used in hypothesis tests ([STEHLÍK, 2003](#)), in obtaining pseudo-random numbers ([JODRÁ, 2010](#)) and to define new probability distributions ([GOERG, 2011](#)).

Regarding the latter, based on a systems theory and an input/output point of view, [Goerg \(2011\)](#) proposes a new class of generalized distributions to model asymmetry, the Lambert  $W \times F$  distributions. Specifically, [Goerg \(2011\)](#) focuses on a system represented by the transformation  $Y = g(X, \delta) = X \exp(\delta X)$ , where  $X$  is a symmetric input rv,  $Y$  is an asymmetric output rv, and  $\delta \in \mathbb{R}$  is a shape parameter that controls the skewness level of the output  $Y$ . Here, the inverse transformation  $X = g^{-1}(Y, \alpha)$  is expressed in closed form in terms of the principal and non-principal branches of the Lambert  $W$  function, which leads to the pdf of the output  $Y$  also having closed form in terms of the branches of this function.



Figure 1 – Lambert W function.



Goerg's transformation is attractive in the sense that it allows asymmetric unbounded-support distributions to be derived by transforming symmetric unbounded-range rv's and not the pdf, cdf, or qf of symmetric baseline distributions. On the other hand, a disadvantage of this transformation is that the support of the distribution of  $X$  depends on the parameters. More details of the Lambert  $W \times F$  distributions can be found in [Goerg \(2015\)](#).

In this dissertation, we propose new families of univariate continuous distributions for modeling asymmetrical data. Initially, we derive a new asymmetric one-parameter distribution that generalizes to the uniform distribution, the Lambert-uniform (LU) distribution. It arises as a nonlinear parametric transformation of a uniform rv bounded to the interval  $(0, 1)$ . The transformation is expressed in terms of the principal branch of the Lambert  $W$  function such that the inverse transformation is expressed in terms of an exponential function. Consequently, the pdf of the LU distribution has a simple closed form and its support does not depend on parameters. Similar to the idea developed in [Goerg \(2011\)](#), the skewness characteristic of the LU rv is a consequence of the parametric condition of the transformation applied on a symmetric rv. However, our transformation is applied on a rv with a range bounded to the interval  $(0, 1)$ , resulting in an asymmetric rv with the same range.

Second, we propose a distribution generator that allows adding one shape parameter to an arbitrary baseline distribution, without restriction on the type of support it has, bounded or unbounded. Thus, from a particular baseline distribution, it is possible to generate a more

flexible distribution in terms of the skewness of the distribution. We use the proposed generator to derive new asymmetric distributions of unbounded supports and positive supports. The new distributions can be understood as extensions with one extra parameter of the well-known generalized bimodal, slash, exponential and Rayleigh distributions. We discuss the parameter estimation via the maximum likelihood (ML) method and evaluate the behavior of the estimators through simulation experiments. Finally, we present some application examples in order to illustrate the usefulness of the proposed distributions in different real-world settings.

The dissertation is structured with five interconnected chapters that are organized as follows: In Chapter 2, the LU distribution is proposed. The main structural properties are studied. The parameter estimation via the ML method is discussed. The behavior of the estimators is evaluated through simulation experiments. An application example is presented. In Chapter 3, the Lambert- $F$  distribution generator is proposed. The main structural properties are studied. Four special cases are derived. Parameter estimation via the ML method is discussed. The behavior of the estimators is evaluated through simulation experiments. Two application examples are presented. In Chapter 4, we propose two asymmetric unbounded-support distributions that are derived from the Lambert- $F$  generator. These distributions exhibit skewness while capturing bimodality and high levels of kurtosis, respectively. In Chapter 5, we provide some final comments associated with the main results of Chapters 2 and 3. Finally, in Chapter 6, we present some ideas for future work associated with the distributions proposed in this dissertation.

---

## LAMBERT-UNIFORM DISTRIBUTION

---

The beta (B) and Kumaraswamy (K) distributions ([JOHNSON; KOTZ; BALAKRISHNAN, 1994](#); [KUMARASWAMY, 1980](#)) are two of the most widely used distributions for modeling asymmetrical bounded data. When the histogram of a certain data set exhibits increasing or decreasing behavior, the one-parameter special case called the power (P) distribution becomes a viable alternative.

In this chapter, we propose a new one-parameter distribution for modeling asymmetric bounded data, the Lambert-uniforme (LU) distribution. Using the principal branch of the Lambert  $W$  function, the LU distribution arises from a nonlinear transformation of a uniform rv bounded to the interval  $(0, 1)$ , hence the name of the distribution.

We study the main structural properties, such as the pdf, the cdf and the moment generating function that present simple closed forms. In addition, we derive the qf that can be written in closed form in terms of the principal branch of the Lambert  $W$  function. From this result, we observe that the pdf can be parameterized in terms of the  $q$ th quantile, which allows us to formulate a model of quantile regression in a simple way. The parameter estimation is carried out using the maximum likelihood method and the behavior of the estimators is evaluated through simulation experiments. Finally, two application examples are considered in order to illustrate the usefulness of the proposal.

### 2.1 Motivation

Consider the following adaptation for bounded rv's in the system approach (Input/Output) proposed by [Goerg \(2011\)](#).

Let  $X$  be a continuous input rv bounded to the interval  $(0, b)$ ,  $b > 0$ , and with standard deviation  $s_X$ . If  $Y$  is the output rv with standard deviation  $s_Y = \alpha s_X$ ,  $\alpha > 0$ , the location-scale Lambert- $W \times F$  rv is given by

$$\frac{Y-b}{\sigma_Y} = \frac{X-b}{\sigma_X} e^{\delta \frac{X-b}{\sigma_X}} = \frac{\alpha(X-b)}{\sigma_Y} e^{\delta \alpha \frac{X-b}{\sigma_Y}},$$

where  $\delta \in \mathbb{R}$  is the skewness parameter.

Now, it can be verified that if  $\delta = \log(\alpha)\sigma_Y/(\alpha b)$ , then  $Y$  has the same rank as  $X$ . Thus, the output and input variables can be written as

$$Y = b - (b - X)\alpha^{\frac{X}{b}}, \quad \alpha > 0, \quad (2.1)$$

and

$$X = \begin{cases} \frac{b}{\log(\alpha)} W_0\left(\frac{\log(\alpha)(Y-b)}{b\alpha}\right) + b, & \text{if } \alpha \in (0, 1) \cup (1, \infty), \\ Y, & \text{if } \alpha = 1, \end{cases} \quad (2.2)$$

respectively, where  $W(\cdot)$  is the principal branch of the Lambert  $W$  function.

In Equation 2.1 (Input/Output system), the output  $Y$  is asymmetric and its skewness is controlled only by  $\alpha$  when input  $X$  is symmetric. Here,  $X$  can be obtained explicitly using the principal branch of the Lambert  $W$  function (See the inverse system given in Equation 2.2), which leads to the pdf of  $Y$  also having a closed form in terms of this function.

On the other hand, Equation 2.2 can also define a new asymmetric rv. For this, it is enough to consider Equation 2.2 as a new system, where the output is now  $X$  and the input is  $Y$ . Here, the output  $X$  is asymmetric and its skewness is only controlled by  $\alpha$  when the input  $Y$  is symmetric. One advantage that we see when working with Equation 2.2 is that some of the main properties of the output  $X$ , for example its pdf and cdf, have closed forms that do not depend on special functions.

In what follows, we will focus on studying the distribution of  $X$  considering the special case  $b = 1$  in which  $Y$  has a uniform distribution. In this way, we obtain a new asymmetric one-parameter distribution that can be considered as an alternative to commonly used bounded support distributions.

## 2.2 Lambert-Uniform Random Variable

In this section, we define the Lambert-uniform rv and derive some of its structural properties.

**Definition 1.** A rv  $X$  follows the Lambert-uniform distribution, denoted as  $X \sim \text{LU}(\alpha)$ , if it can be represented as

$$X = \begin{cases} \frac{1}{\log(\alpha)} W_0\left(\frac{\log(\alpha)(U-1)}{\alpha}\right) + 1, & \text{if } \alpha \in (0, 1) \cup (1, e), \\ U, & \text{if } \alpha = 1, \end{cases} \quad (2.3)$$

where  $W_0(\cdot)$  is the principal branch of the Lambert  $W$  function,  $U$  is a rv with standard uniform (U) distribution and  $e$  is the Euler's number.

In Definition 1, it can be seen that the parameter  $\alpha$  must necessarily be greater than 0 since the logarithm function is involved in the transformation. On the other hand,  $\alpha$  must be less than  $e$  to ensure that the cdf of  $X$  is a nondecreasing function. Considering  $\alpha$  as known, and remembering that  $W_0(\cdot)$  is a monotonic function, it is easy to see that  $X$  is a one-to-one transformation of  $U$ . Thus, we concentrate on obtaining the limit values of  $X$  to determine its range.

We see that

$$\lim_{U \rightarrow 1^-} X = 1 \quad \text{and} \quad \lim_{U \rightarrow 0^+} X = \frac{1}{\log(\alpha)} W_0\left(-\frac{\log(\alpha)}{\alpha}\right) + 1.$$

The first limit is direct. As for the second, considering the change of variable  $\beta = -\log(\alpha)$ , it is observed that  $\lim_{U \rightarrow 0^+} X = -\frac{1}{\beta} W_0(e^\beta) + 1 = -1 + 1 = 0$ . Therefore, the range of  $X$  is the interval  $(0, 1)$ .

**Proposition 1.** Let  $X \sim \text{LU}(\alpha)$ . Then, the cdf of  $X$  is given by

$$F_X(x; \alpha) = 1 - (1-x)\alpha^x, \quad x \in (0, 1) \text{ and } \alpha \in (0, e). \quad (2.4)$$

**Proof.** From Equation (2.3), it is observed that

$$F_X(x; \alpha) = P(X \leq x) = \begin{cases} P\left(W_0\left[\frac{\log(\alpha)(U-1)}{\alpha}\right] \leq \log(\alpha)(x-1)\right), & \text{if } \alpha \in (1, e). \\ P\left(W_0\left[\frac{\log(\alpha)(U-1)}{\alpha}\right] \geq \log(\alpha)(x-1)\right), & \text{if } \alpha \in (0, 1). \end{cases}$$

Thus, by definition of the Lambert  $W$  function, it follows that

$$\begin{aligned} P(X \leq x) &= P\left(\frac{\log(\alpha)(U-1)}{\alpha} \leq \log(\alpha)(x-1) \exp\{-\log(\alpha)(x-1)\}\right) \\ &= P(U \leq 1 - [1-x]\alpha^x), \end{aligned}$$

and the result is obtained considering that  $P(U \leq u) = u$ , once  $U$  follows a standard U distribution. Note that the analytic expression obtained for the cdf of  $X$  is also valid for  $\alpha = 1$ , since  $F_X(x; 1) = x$ .  $\square$

The pdf of  $X$  can be obtained in a straightforward way from Equation (2.4).

**Corollary 1.** Let  $X \sim \text{LU}(\alpha)$ . Then, the pdf of  $X$  is given by

$$f_X(x; \alpha) = [1 - \log(\alpha)(1-x)]\alpha^x, \quad x \in (0, 1), \quad \alpha \in (0, e). \quad (2.5)$$

Consistent with Definition 1, it is observed that Equations (2.4) and (2.5) reduce to the cdf and the pdf of the standard U distribution, respectively, when  $\alpha = 1$ . Therefore, the LU distribution can be understood as an extension with one extra parameter of the standard U distribution.

The analytical description of the shapes for the LU pdf is simple and leads to establish that

1.  $\lim_{x \rightarrow 0^+} f_X(x; \alpha) = 1 - \log(\alpha)$  and  $\lim_{x \rightarrow 1^-} f_X(x; \alpha) = \alpha$ .
2.  $f_X(x; \alpha)$  is a constant function if  $\alpha = 1$ , a decreasing monotonic function if  $\alpha \in (0, 1)$  and an increasing monotonic function if  $\alpha \in (1, e)$ .

Property 1 shows that the pdf of the LU distribution converges to finite values (greater than 0) as  $x$  tends to the extreme values, 0 and 1, of the support. From Property 2, it follows that the LU distribution is appropriate to fit bounded data whose relative frequency shows increasing or decreasing behavior. Figure 2 shows some pdf curves of the LU distribution for different values of  $\alpha$ . Note that the behavior of the pdf curves is consistent with what is established above. In addition, note that the curvature of the pdf varies depending on its behavior at the ends of the support. This is a behavior similar to that exhibited by other one-parameter distributions known in the literature, such that the P (JOHNSON; KOTZ; BALAKRISHNAN, 1994), skew-uniform (SU) (SHAW; BUCKLEY, 2007) and Marshall-Olkin extended uniform (MOEU) (JOSE; KRISHNA, 2011) distributions.

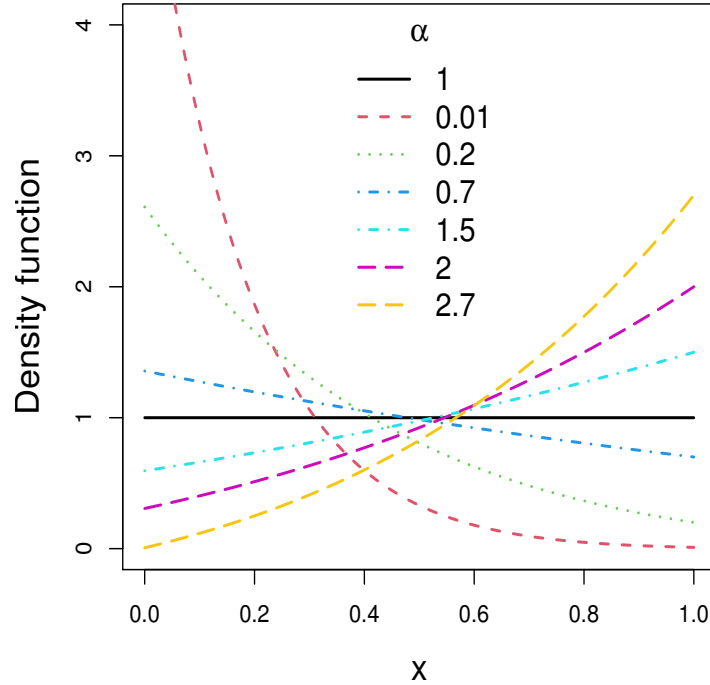
Note that, due to the behavior of the pdf at the ends of the support, the LU distribution may more adequately fit the extreme sample quantiles than a distribution whose pdf tends to  $\infty$  and 0 at the ends of the support. In Section 2.7, we see that the LU distribution may perform better in fitting data than the P, MOEU and SU distributions, even better than the two-parameter beta (B) and Kumaraswamy (K) distributions whose pdf's (in the monotonic case) tend to  $\infty$  and 0 at the ends of the support.

Considering steps very similar to those of the proof of Proposition 1, the qf of the LU distribution can be easily derived by inverting the cdf given in Proposition 1. The resulting analytical expression for this function, for  $u \in (0, 1)$ , is given by

$$Q_X(u; \alpha) = \begin{cases} \frac{1}{\log(\alpha)} W_0 \left( \frac{\log(\alpha)(u-1)}{\alpha} \right) + 1, & \text{if } \alpha \in (0, 1) \cup (1, e), \\ u, & \text{if } \alpha = 1. \end{cases} \quad (2.6)$$

Since the Lambert  $W$  function is implemented in different statistical software, Equation (2.6) can be easily computed.

As a final consideration of this section, we highlight that the linear transformation  $a + bX$ , where  $X \sim \text{LU}(\alpha)$ ,  $a \in \mathbb{R}$  and  $b > 0$ , follows a LU distribution on the continuous range  $(a, a + b)$ . Therefore, the LU distribution can be easily used to fit bounded data to any real range.

Figure 2 – Plot of the LU pdf for different values of  $\alpha$ .

## 2.3 Characterization of Skewness and Kurtosis

In the following, a description of the skewness and kurtosis characteristics of the LU distribution is made by analyzing Fisher's skewness and kurtosis coefficients. For this, the moment generating function is first calculated.

**Proposition 2.** Let  $X \sim \text{LU}(\alpha)$ . Then, in the case  $\alpha = 1$ , the moment generating function of  $X$  is given by  $M_X(t) = (e^t - 1)/t$ . In the case  $\alpha \in (0, 1) \cup (1, e)$ , the moment generating function is given by  $M_X(t) = \{\log^2(\alpha) - [1 - \log(\alpha)]t + \alpha t e^t\} [t + \log(\alpha)]^{-2}$ .

**Proof.** In the case  $\alpha = 1$ , the distribution LU reduces to standard U distribution, thus  $M_x(t) = \mathbb{E}(e^{tX}) = (e^t - 1)/t$ . In the case  $\alpha \in (0, 1) \cup (1, e)$ , we observe that

$$\mathbb{E}(e^{tX}) = \frac{[1 - \log(\alpha)][\alpha \exp(t) - 1]}{t + \log(\alpha)} + \log(\alpha) \int_0^1 x \exp\{x[t + \log(\alpha)]\} dx,$$

and the result is obtained considering the usual method of integration by parts and an appropriate algebra.  $\square$

**Corollary 2.** Let  $X \sim \text{LU}(\alpha)$ . Then, in the case  $\alpha = 1$ , the first four raw moments of  $X$  are  $\mathbb{E}(X) = 1/2$ ,  $\mathbb{E}(X^2) = 1/3$ ,  $\mathbb{E}(X^3) = 1/4$  and  $\mathbb{E}(X^4) = 1/5$ . In the case  $\alpha \in (0, 1) \cup (1, e)$ , for

$\delta = \log(\alpha)$ , the first four raw moments are given by

$$\begin{aligned}\mathbb{E}(X) &= \frac{\alpha - 1 - \delta}{\delta^2}, \quad \mathbb{E}(X^3) = \frac{3[\alpha\delta^2 - 2\delta(2\alpha + 1) - 6(1 - \alpha)]}{\delta^4}, \\ \mathbb{E}(X^2) &= \frac{2[(\alpha + 1)\delta + 2(1 - \alpha)]}{\delta^3}, \quad \mathbb{E}(X^4) = \frac{4[\alpha\delta^2(\delta - 6) + 6\delta(3\alpha + 1) + 24(1 - \alpha)]}{\delta^5}.\end{aligned}$$

**Corollary 3.** Let  $X \sim \text{LU}(\alpha)$ . Then, in the case  $\alpha = 1$ , the skewness ( $\gamma_1(\alpha)$ ) and kurtosis ( $\gamma_2(\alpha)$ ) coefficients assume the values 0 and 9/5, respectively. In the case  $\alpha \in (0, 1) \cup (1, e)$ , the coefficients are given by

$$\gamma_1(\alpha) = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2 - \mu_1)^{3/2}} \quad \text{and} \quad \gamma_2(\alpha) = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2 - \mu_1)^2}.$$

where  $\mu_r = \mathbb{E}(X^r)$ , with  $r = 1, 2, 3, 4$ , are as in Corollary 2.

The skewness and kurtosis ranges for the LU distribution are  $(3e - 6e^2 + 2e^3 - 4)[(2 + 2e - e^2)^{-3/2}] < \gamma_1(\alpha) < 2$  and  $9/5 < \gamma_2(\alpha) < 9$ .

Figure 3 presents plots of the coefficients given in Corollary 3. The figure shows that the LU distribution is symmetric only in the case  $\alpha = 1$ , has positive skewness when  $\alpha \in (0, 1)$  and has negative skewness when  $\alpha \in (1, e)$ . Furthermore, it is observed that the LU distribution can model kurtosis levels higher than the kurtosis level of the  $U$  distribution.

## 2.4 Maximum Likelihood Estimation

For a random sample  $X_1, \dots, X_n$ , such that  $X_i \sim \text{LU}(\alpha)$ , with  $i = 1, \dots, n$ , the log-likelihood function is given by

$$\ell(\alpha) = \log \prod_{i=1}^n f_X(x_i; \alpha) = \log(\alpha) \sum_{i=1}^n x_i + \sum_{i=1}^n \log[1 - \log(\alpha)(1 - x_i)], \quad (2.7)$$

Thus, the score function is given by

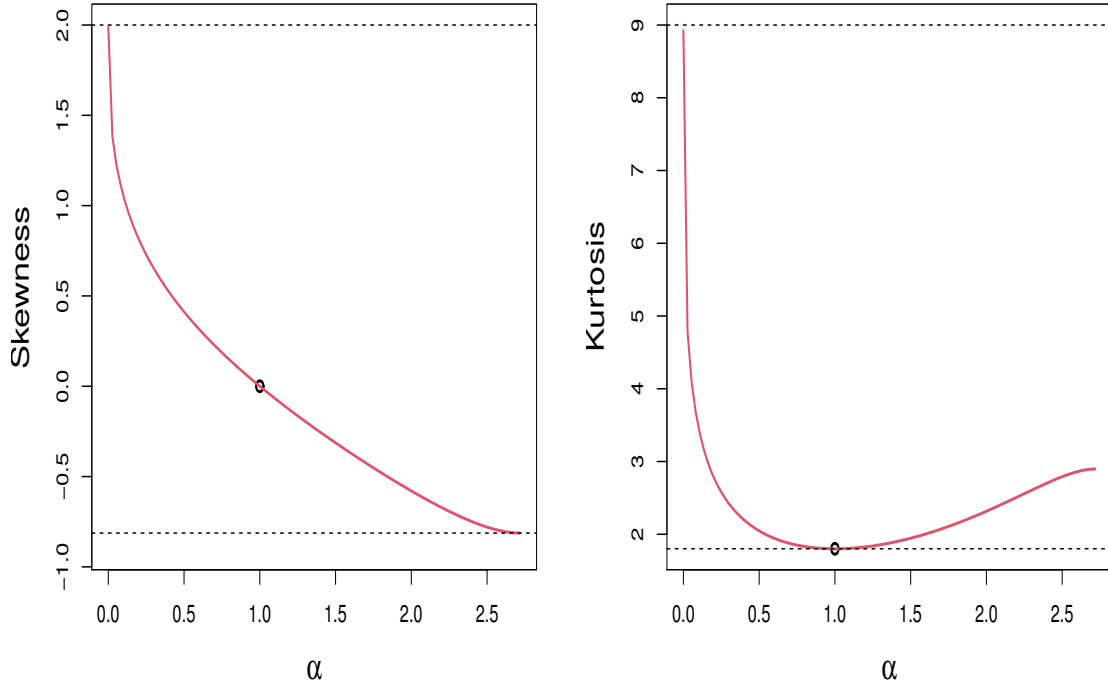
$$\mathcal{U}(\alpha) = \frac{\partial \ell(\alpha)}{\partial \alpha} = \frac{1}{\alpha} \sum_{i=1}^n x_i - \frac{1}{\alpha} \sum_{i=1}^n \frac{1 - x_i}{1 - \log(\alpha)(1 - x_i)}. \quad (2.8)$$

From Equation (2.8), it is observed that the maximum likelihood (ML) estimator for  $\alpha$  cannot be explicitly expressed. Therefore, the ML estimate of  $\alpha$  must be obtained by solving the equation  $\mathcal{U}(\alpha) = 0$  by numerical procedures. The `uniroot.all` function available in the `rootSolve` package of the R programming language (R Core Team, 2019) is a good option to tackle this task.

Since the ML estimator of  $\alpha$  does not have a closed form, a good alternative to obtain the ML estimate is to solve the optimization problem  $\max_{\alpha} \ell(\alpha)$ , subject to  $\alpha \in (0, e)$ . To solve



Figure 3 – Plots of the skewness and kurtosis coefficients of the LU distribution (red color) and the U distribution (circle)



this problem, we use the optim function in the R programming language under the L-BFGS-B algorithm (BYRD *et al.*, 1995). This algorithm requires the specification of a value in the range of  $\alpha$  to initialize the iterative process. Through simulation experiments, we observe that the initial value  $\alpha_0 = 1$  is a good initial value.

The second partial derivative of the  $\ell(\alpha)$  function, with respect to  $\alpha$ , is given by

$$\frac{\partial^2 \ell(\alpha)}{\partial \alpha^2} = -\frac{1}{\alpha^2} \sum_{i=1}^n x_i + \frac{1}{\alpha^2} \sum_{i=1}^n \frac{1-x_i}{1-\log(\alpha)(1-x_i)} - \frac{1}{\alpha^2} \sum_{i=1}^n \left( \frac{1-x_i}{1-\log(\alpha)(1-x_i)} \right)^2.$$

Thus, under regularity conditions, we observe that the Fisher information is given by

$$\mathcal{I}(\alpha) = -\mathbb{E} \left( \frac{\partial^2 \ell(\alpha)}{\partial \alpha^2} \right) = \frac{n}{\alpha} \int_0^1 \frac{u^2}{\alpha^u [1 - \log(\alpha)u]} du. \quad (2.9)$$

The integral in Equation (2.9) can be calculated by numerical integration, for example, the integrate function of the R programming language can be used. Then, under regularity conditions, the asymptotic distribution of  $(\hat{\alpha} - \alpha)$  is  $N(0, \mathcal{I}^{-1}(\alpha))$ . Thus, the asymptotic standard error of  $\hat{\alpha}$  is given by  $1/\sqrt{\mathcal{I}(\hat{\alpha})}$  and the asymptotic  $100(1-\gamma)\%$  confidence interval for  $\hat{\alpha}$  is given by  $\hat{\alpha} \pm z_{\gamma/2}/\sqrt{\mathcal{I}(\hat{\alpha})}$ , where  $\gamma/2$  is the  $\gamma/2$  upper quantile of the standard normal distribution.

## 2.5 Quantile Regression Model

In statistical modeling, the regression technique is used to quantify the relationship between the dependent variable (response) and one or more independent variables (covariates). In the case in which the interest lies in quantifying the effect on the conditional mean response, given the covariates, the classical least squares regression model and the generalized linear models are especially valued. These models have been shown to be very useful when analyzing data in various areas of knowledge. However, there are scenarios where it is equally or even more important to quantify the effect on some other measure such as the conditional median or some extreme conditional quantile of the response, see, e.g., [Girma and Görg \(2003\)](#) and [Chunying \(2011\)](#). In this scenario, a quantile regression model is appropriate because it allows quantifying the effect of the covariates on any quantile of the response.

In this section, we propose a quantile regression model formulated from a reparameterized version of the LU distribution proposed in Section 2.2. In this model, it is only necessary to estimate the regression coefficients, since it is formulated from a distribution with a single shape parameter linked to the linear predictor through an appropriate link function. We highlight that the performance of the proposed model is appropriate in scenarios where the histogram of the observed values of the response variable exhibits a decreasing or increasing behavior.

### 2.5.1 The LU Model

The parameter  $\alpha$  of the LU distribution can be explicitly expressed as a function of the  $q$ th quantile, which allows reparameterizing the LU pdf in terms of the  $q$ th quantile and, consequently, formulate a quantile regression model in a simple way.

Denoting by  $\eta$  the  $q$ th quantile of the LU distribution, from Equation (2.6), we obtain that  $\alpha = [(1 - q)/(1 - \eta)]^{1/\eta}$ . Thus, the LU distribution can be easily reparameterized in terms of the  $q$ th quantile, obtaining (for  $q \in (0, 1)$  is known) the pdf given by

$$f_X(x; \eta) = \left( \frac{1 - q}{1 - \eta} \right)^{\frac{x}{\eta}} \left[ 1 - \frac{1}{\eta} \log \left( \frac{1 - q}{1 - \eta} \right) (1 - x) \right], \quad x, \eta \in (0, 1). \quad (2.10)$$

Let  $X_1, \dots, X_n$  be  $n$  random variables and denote by  $x_1, \dots, x_n$  the observed values. Assume that each  $X_i$  has pdf  $f_{X_i}(x; \eta_i)$  given in Equation (2.10). The LU quantile regression model is defined by establishing that the  $q$ th quantile  $\eta_i$  of  $X_i$  satisfies the functional relationship  $g(\eta_i) = w_i^t \beta$ ,  $i = 1, \dots, n$ , where  $w_i = (1, w_{i1}, \dots, w_{i(k-1)})^t$  is the vector of covariates associated to the response  $x_i$ ,  $\beta = (\beta_0, \beta_1, \dots, \beta_{(k-1)})^t$  is a  $k$ -dimensional vector of unknown regression coefficients and  $g(\cdot)$  is a strictly increasing and twice differentiable function that maps  $(0, 1)$  into  $\mathbb{R}$  (link function). For instance, the most useful well-known link functions are the logit, log-log and probit functions.

### 2.5.2 ML Estimation

From Equation (2.10), the log-likelihood function is given by

$$\ell(\beta) = \log(1-q) \sum_{i=1}^n \frac{x_i}{\eta_i} - \sum_{i=1}^n \frac{(1-\eta_i)x_i}{\eta_i} + \sum_{i=1}^n \log \left[ 1 - \frac{1}{\eta_i} \log \left( \frac{1-q}{1-\eta_i} \right) (1-x_i) \right], \quad (2.11)$$

and the score functions are given by

$$\begin{aligned} \frac{\partial \ell(\beta)}{\partial \beta_r} = & -\log(1-q) \sum_{i=1}^n \frac{x_i \eta_{i,r}}{\eta_i^2} + \sum_{i=1}^n \frac{x_i \log(1-\eta_i) \eta_{i,r}}{\eta_i^2} + \sum_{i=1}^n \frac{x_i \eta_{i,r}}{\eta_i (1-\eta_i)} \\ & + \sum_{i=1}^n \frac{\log \left( \frac{1-q}{1-\eta_i} \right) (1-x_i) \eta_{i,r}}{\eta_i^2 \left[ 1 - \frac{1}{\eta_i} \log \left( \frac{1-q}{1-\eta_i} \right) (1-x_i) \right]} - \sum_{i=1}^n \frac{(1-x_i) \eta_{i,r}}{(1-\eta_i) \left[ 1 - \frac{1}{\eta_i} \log \left( \frac{1-q}{1-\eta_i} \right) (1-x_i) \right]}, \end{aligned} \quad (2.12)$$

where  $\eta_{i,r} = \partial \eta_i / \partial \beta_r$ ,  $\eta_i = g^{-1}(w_i^t \beta)$ , with  $r = 0, 1, \dots, k-1$ . Note that  $\eta_{i,r}$  depends on the link function. For example, if the logit link is considered, that is,  $g(u) = \log(u/(1-u))$ , for  $u \in (0, 1)$ , then  $\eta_{i,r} = \eta_i(1-\eta_i)w_{i,r}$ , where  $\eta_i = \exp(w_i^t \beta) / [1 + \exp(w_i^t \beta)]$ ,  $w_{i0} = 1$ , with  $i = 1, 2, \dots, n$ ,  $r = 0, 1, \dots, k-1$ .

We observe from Equation (2.11) that the ML estimators for the coefficients  $\beta$ s cannot be expressed in closed form. Thus, the ML estimates must be obtained by solving the system of score equations using numerical procedures. In the R programming language, the `multirrot` function of the `rootSolve` package is a good alternative to solve this system of equations.

In this case, since the ML estimators do not have a closed form, a good alternative to obtain ML estimates is to solve the following optimization problem  $\max_{\beta} \ell(\beta)$ , subject to  $\beta_r \in \mathbb{R}$ ,  $r = 0, 1, \dots, k-1$ , where  $\ell(\beta)$  is given in Equation (2.11). We solved this problem using the function `optim` of the R programming language and, specifically, the BFGS algorithm was applied.

Under regularity conditions, the asymptotic distribution of  $(\widehat{\beta}_{\text{ML}} - \beta)$  is  $N_k(0, K(\beta)^{-1})$ , where  $K(\beta)$  is the expected information matrix. Since the function  $\ell(\beta)$  is not simple, it is not easy to obtain the analytical expression of this matrix. However, we obtain an approximation from the observed information matrix, whose elements are computed as minus the second partial derivatives of the log-likelihood function with respect to all the parameters (evaluated at the ML estimates). Thus, the observed information matrix is given by

$$I(\beta) = \begin{pmatrix} \varepsilon_{\beta_0 \beta_0} & \varepsilon_{\beta_0 \beta_1} & \cdots & \varepsilon_{\beta_0 \beta_{k-1}} \\ \varepsilon_{\beta_0 \beta_1} & \varepsilon_{\beta_1 \beta_1} & \cdots & \varepsilon_{\beta_1 \beta_{k-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{\beta_0 \beta_{k-1}} & \varepsilon_{\beta_1 \beta_{k-1}} & \cdots & \varepsilon_{\beta_{k-1} \beta_{k-1}} \end{pmatrix},$$

$$\varepsilon_{\beta_r \beta_p} = - \left. \frac{\partial^2 \ell(\beta)}{\partial \beta_r \partial \beta_p} \right|_{\beta = \widehat{\beta}_{\text{ML}}}, \quad r, p = 0, 1, \dots, k-1,$$

where the second derivatives are presented in Appendix D.

## 2.6 Simulation Studies

In this section, we initially carry out a simulation study to evaluate the behavior of the ML estimators of the shape parameter of the LU distribution. Subsequently, we carried out a second simulation study to evaluate the behavior of the ML estimators for the coefficients of the LU quantile regression model.

### 2.6.1 First Simulation Study

In this study, 1000 random samples from the LU distribution were simulated considering the sample sizes  $n = 10, 20, \dots, 1000$ , respectively, in the scenarios A ( $\alpha = 0.5$ ) and B ( $\alpha = 1.5$ ). These scenarios were chosen arbitrarily taking into account only the condition that in one the skewness of the LU distribution is positive and in the other negative. The samples were generated using the qf given in Equation (2.6). The LambertW package (GOERG, 2011) available in the R programming language was used to compute the principal branch of the Lambert  $W$  function. The estimates were obtained by maximizing the log-likelihood function under the considerations of Section 2.4.

Figure 4 shows the average estimate (AE), the empirical standard deviation (SD) and the roots of the mean squared error (RMSE) for each of the 1000 estimates obtained for each scenario and sample size considered. In addition, the average of asymptotic standard errors (SE) and the coverage probability (CP) of the 95% asymptotic confidence intervals are also reported. In the figures, it is observed that the AE's tend to be close to the true values of  $\alpha$  as the sample size increases. The SD's, RMSE's and SE's are close and decrease to zero as the sample size increases, as expected in standard asymptotic theory. It is also observed that the CP's converge to the nominal values as the sample size increases.

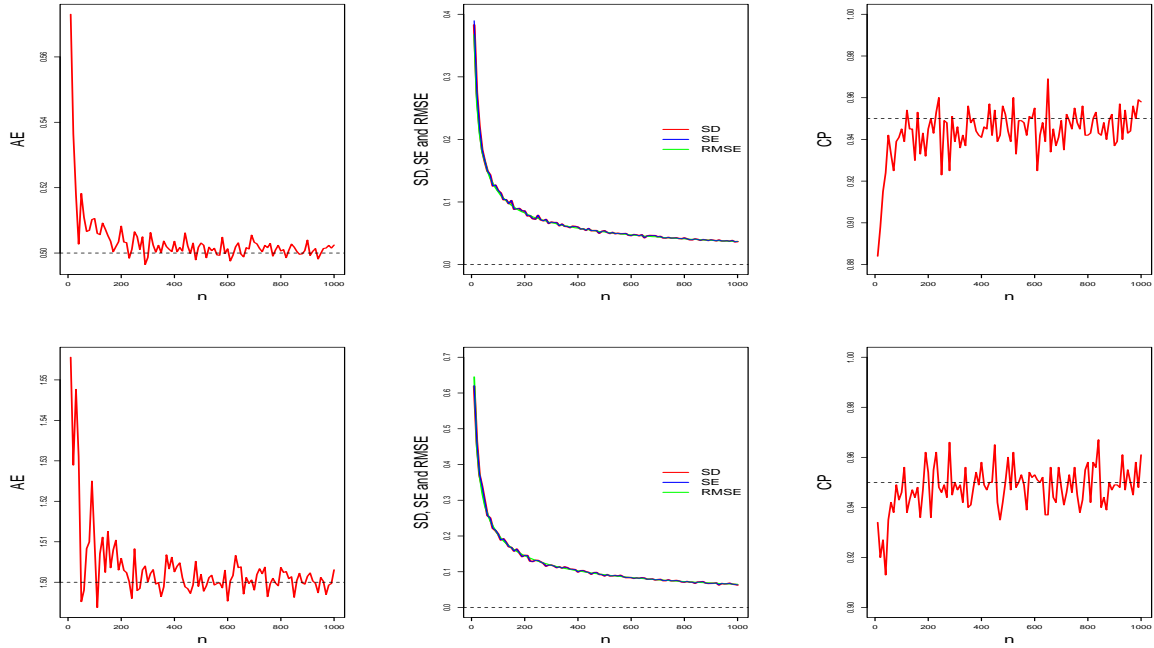
Regarding the comparison of the results obtained in both scenarios, we observe that: The mean estimate in the smaller sample sizes is closer to the true value in scenario A; Consequently, the SD, SE and RMSE values are higher for small sample sizes in scenario B; Finally, the PC values in the smaller sample sizes are closer to the nominal value in scenario B. This is partly explained because in this scenario the SE values are bigger.

### 2.6.2 Second Simulation Study

In this study, we simulated 1000 random samples from the LU distribution (with pdf given in Equation (2.10)), considering the sample sizes  $n = 10, 20, \dots, 1000$ , respectively. Here, the shape parameter is linked via the logit function with three covariates. The samples were generated as follows:

1. *Definition of covariates:* Generate  $w_1 = (w_{11}, \dots, w_{1n})^t$ ,  $w_2 = (w_{21}, \dots, w_{2n})^t$  and  $w_3 = (w_{31}, \dots, w_{3n})^t$ , where  $(w_{1j}, w_{2j})$  follows a bivariate normal distribution with parameters

Figure 4 – The AE, SD, SE, RMSE and CP for each of the 1000 estimates of  $\alpha$  obtained in the scenarios A (top) and B (bottom), under the different sample sizes.



$\mu_1 = \mu_2 = 0$ ,  $\sigma_1 = \sigma_2 = 1$  and  $\rho = 0.7$ , with  $j = 1, \dots, n$  and  $w_3$  is a binary variable with probability of success depending on the variable  $w_1$  through the logistic function, that is,  $w_{3j} \sim \text{Bernoulli}(p_j)$ , where  $p_j = 1/[1 + \exp(-w_{1j})]$   $j = 1, \dots, n$ .

2. *Definition of scenarios:* We considered two scenarios, A and B, where in both we picked  $\beta_0 = -2$ ,  $\beta_1 = 0.1$ ,  $\beta_2 = 0.5$  and  $\beta_3 = -2.5$ . Regarding the choices for  $q$ , we chose the values 0.25 for Scenario A and 0.75 for Scenario B.
3. *Simulate the response variable:* Generate  $(u_1, \dots, u_n)^t$ ,  $u_j \sim \text{uniform}(0, 1)$ ,  $j = 1, \dots, n$ , and calculate

$$x_j = \frac{\eta_j}{\log\left(\frac{1-q}{1-\eta_j}\right)} W_0 \left[ \log\left(\frac{1-q}{1-\eta_j}\right) \frac{u_j - 1}{\eta_j} \left(\frac{1-\eta_j}{1-q}\right)^{1/\eta_j} \right] + 1, \quad j = 1, \dots, n,$$

where  $\eta_j = \exp(w^t \beta) / [1 + \exp(w^t \beta)]$ , such that  $w^t \beta = \beta_0 + \beta_1 w_{1j} + \beta_2 w_{2j} + \beta_3 w_{3j}$ , with  $j = 1, \dots, n$ .

In both scenarios, under the different sample sizes, it is possible to verify that the histogram of the simulated values has a decreasing behavior.

For each simulated sample, we calculated the ML estimates for the coefficients  $\beta$ s under the considerations of Section 2.5.2. Figures 5 and 6 show the AEs, SDs, SEs and RMSEs for the estimates obtained in each scenario and sample size considered. In Figure 7, the CPs of the 95% asymptotic confidence intervals are reported. Similar to the results obtained in the first

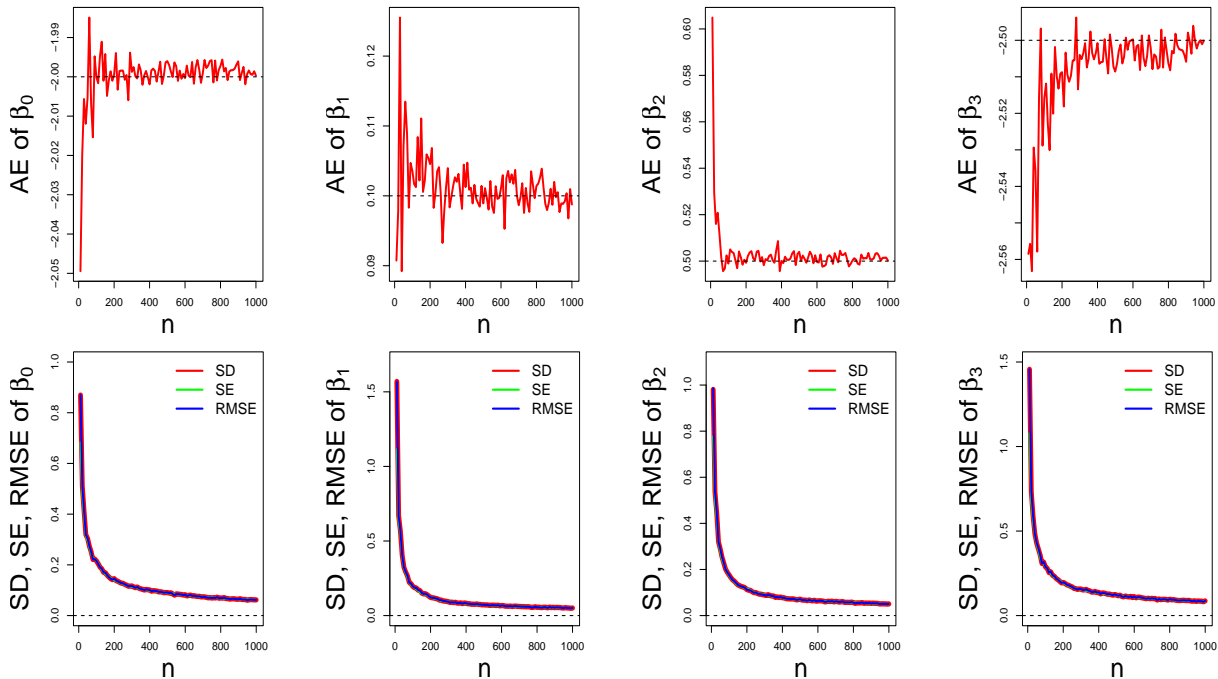


Figure 5 – The AE, SD, SE and RMSE for each of the 1000 estimates of the coefficients  $\beta$ s obtained in Scenario A, under the different sample size.

simulation study, in the figures, it is observed that the AEs tend to be close to the true values of the coefficients  $\beta$ s as the sample size increases. The SDs, RMSEs and SEs are close and decrease to zero as the sample size increases, as expected in standard asymptotic theory. It is observed that the CPs converge to the nominal values as the sample size increases.

Note that the increasing and decreasing patterns exhibited in the top panels of Figures 5 and 6 suggest an underestimation and overestimation, respectively, of the individual effect of the covariates on the 0.25th and 0.75th quantiles of the response when the sample size is small.

When comparing the upper right panels of the figures, a different behavior pattern is observed for the AEs of  $\beta_3$ . Therefore, the effect of the covariate  $w_3$  on the 0.25th quantile of the response can be underestimated when the sample size is small, while the effect on the 0.75th quantile can be overestimated.

## 2.7 Data Analysis

In this section, two application examples are presented in order to illustrate the usefulness of the LU distribution and the LU quantile regression model.

In the first, we compare the performance of the LU, P, MOEU, SU, B and K distributions in fitting a real data set. In the second example, based on a real data frame, and in order to quantify the effect of the covariates on the 0.25th, 0.5th and 0.75th quantiles of the response, we compare the performance of the LU quantile regression model with the performance of other models such as arc-secant-hyperbolic-normal (ASHN) (KORKMAZ; CHESNEAU; KORKMAZ,

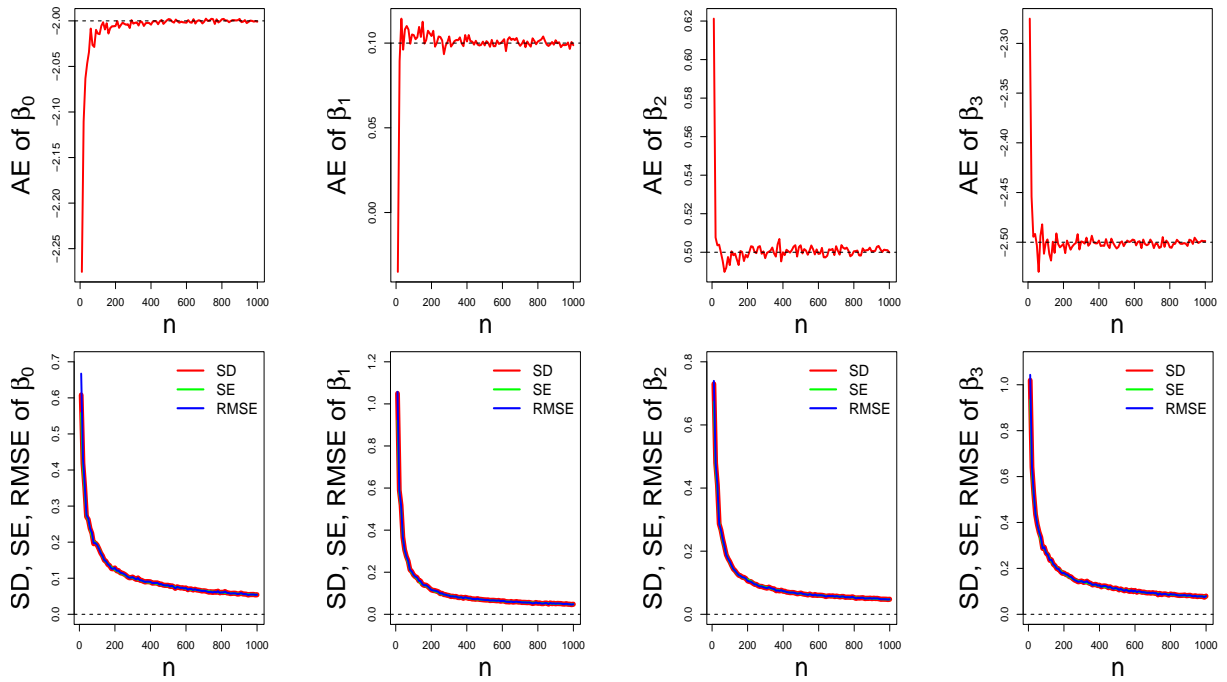


Figure 6 – The AE, SD, SE and RMSE for each of the 1000 estimates of the coefficients  $\beta_s$  obtained in Scenario B, under the different sample size.

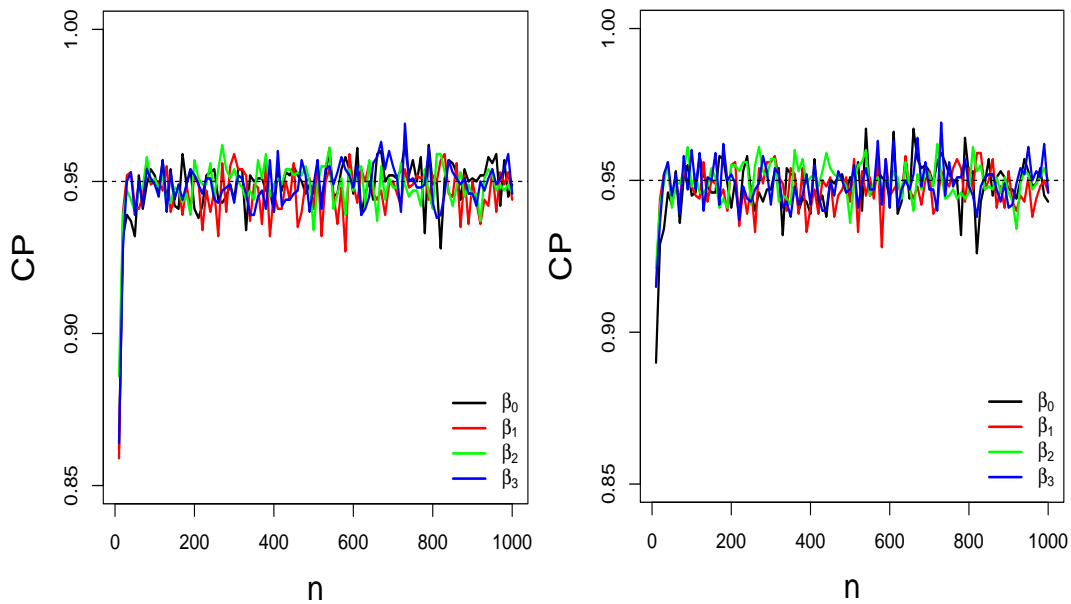


Figure 7 – The CPs for the estimates of the coefficients  $\beta_s$  in: Scenario A (left); and Scenario B (right).

2021) and K (MITNIK; BAEK, 2013) quantile regression models. In all three models, the logit function is considered to relate the  $q$ th quantile of the response to the linear predictor. The pdfs of the ASHN and K distributions are given, respectively, by

$$f(x; \alpha, \eta) = \frac{2\alpha\Phi^{-1}(q_\alpha)}{A(\eta)x\sqrt{1-x^2}} \phi\left(\frac{\Phi^{-1}(q_\alpha)A(x)}{A(\eta)}\right) \left[2 - 2\Phi\left(\frac{\Phi^{-1}(q_\alpha)A(x)}{A(\eta)}\right)\right]^{\alpha-1},$$

where  $x \in (0, 1)$ ,  $q_\alpha = 1 - (q^{1/\alpha})/2$ ,  $A(z) = \log[(1 + \sqrt{1-z^2})/z]$  is the hyperbolic arcsecant function,  $\alpha > 0$  is a shape parameter,  $\eta \in (0, 1)$  is a quantile parameter and  $q \in (0, 1)$  is known, and

$$f(x; \beta, \eta) = \frac{\beta \log(1-q)}{\log(1-\eta^\beta)} x^{\beta-1} (1-x^\beta)^{\log(1-q)/\log(1-\eta^\beta)-1}, \quad x \in (0, 1),$$

where  $\beta > 0$  is a shape parameter,  $\eta \in (0, 1)$  is a quantile parameter and  $q \in (0, 1)$  is known.

The regression framework for bounded responses based on the K and ASHN distributions is very similar to the regression framework based on the LU distribution presented in Section 2.5.1. However, the main difference with LU regression is that it depends on a shape parameter that gives great flexibility to the modeling.

In both examples, the parameters are estimated by maximizing the corresponding likelihood functions with the optim function in the R language. We compared the performance of the distributions by contrasting the maximum value of the log-likelihood function ( $\ell$ ) and contrasting the values associated with the Akaike Information Criterion (AIC) (AKAIKE, 1974) and the Bayesian Information Criterion (BIC) (SCHWARZ, 1978). In general, the best model can be chosen as the one that shows the highest  $\ell$  value and the lowest AIC and BIC values. In addition, we consider the usual Anderson–Darling (AD) and Cramer–von Mises (CvM) goodness-of-fit tests to assess the quality of the fits in the first example. In the second example, we use these tests to assess the overall quality of fit of the regression models, by testing the hypothesis that the randomize residuals (DUNN; SMYTH, 1996) follow the standard normal distribution. These residuals follow a standard normal distribution when the overall quality of fit is appropriate. We use the ad.test and cvm.test functions available in the goftest (FARAWAY *et al.*, 2019) package in the R programming language to calculate the statistics and  $p$ -values of these tests.

### 2.7.1 Peak Horizontal Acceleration Data

We consider a data set consisting of 182 observations on the peak horizontal acceleration (g) in earthquakes recorded by observation stations in California, USA. These data were originally analyzed by Joyner and Boore (1981) and can be found with the name attenu in the data sets package of the R language. For these data, we note that the minimum value is equal to 0.003, the maximum value is equal to 0.810, and the Fisher skewness value is equal to 1.641. Thus, added to the fact that the histogram of the data exhibits a decreasing behavior (See Figure 9), we expect that the LU distribution presents a good performance in the fit of these data.



Table 1 – The parameter estimates (with standard errors in parentheses), the  $\ell$ , AIC, CAIC and BIC values and the  $p$ -values of the AD and CvM goodness-of-fit tests for the SU, P, MOEU, B, K and LU distributions fitted to the peak horizontal acceleration data.

Parameter	LU	K	B	MOEU	P	SU
$\alpha$	0.005 (0.002)	0.890 (0.062)	0.877 (0.080)	0.112 (0.013)	0.412 (0.030)	1.000 (0.170)
$\beta$	-	4.423 (0.571)	4.699 (0.533)	-	-	-
$\ell$	158.5	157.0	156.6	149.3	98.1	91.7
AIC	-315.1	-310.1	-309.3	-296.7	-194.3	-181.5
BIC	-311.9	-303.7	-302.9	-293.5	-191.1	-178.2
AD	0.978	0.882	0.486	0.069	<0.001	<0.001
CvM	0.965	0.884	0.576	0.140	<0.001	<0.001

Table 1 reports the ML estimates, the  $\ell$ , AIC and BIC values and the  $p$ -values of the AD and CvM goodness of fit tests for each distribution fitted to the peak horizontal acceleration data. In this table, based on  $p$ -values, considering a significance level of 5%, we observe that the SU and P distributions are not appropriate to fit the peak horizontal acceleration data. Note that the MOEU, SU, P and LU distributions are uni-parametric, however the performance shown by the LU distribution is clearly better. We also observe in Table 1 that the LU distribution is the one with the lowest AIC and BIC values and the one with the highest  $\ell$  value, evidencing that this distribution must be selected over the others for the fit of the peak horizontal acceleration data.

Calculating  $\lim_{x \rightarrow 0^+} f_X(x)$  for each fitted distribution, we observe the limit value 6.298 for the LU distribution, 8.923 for the MOEU distribution, 2.000 for the SU distribution and  $\infty$  for the B, K and P distributions. Now, calculating  $\lim_{x \rightarrow 1^-} f_X(x)$ , we observe the limit values 0.005 for the LU distribution, 0.112 for the MOEU distribution, 0.412 for the P distribution and 0 for the SU, B and K distributions. The fact that the LU pdf tends to finite values at both ends of the support is an important feature when explaining the good performance of the LU distribution in fitting this data set.

In Figure 9, the histogram of the data fitted with the LU, P, MOEU, SU, B and K distributions is presented. Here, it can be seen that the curvature characteristics of the LU, B and K pdf's are similar, exhibiting a similar performance in the fit of the most central sample quantiles. However, the LU pdf more appropriately fits the more extreme quantiles by tending to the values 6.298 and 0.005 at the ends of the support.

Figure 8 presents the qqplots for the fitted distributions. This figure shows that the LU distribution fits the peak horizontal acceleration data appropriately.

Figure 8 – QQ-plots: (a) LU distribution; (b) K distribution; (c) B distribution; (d) MOEU distribution; (e) P distribution; and (f) SU distribution.

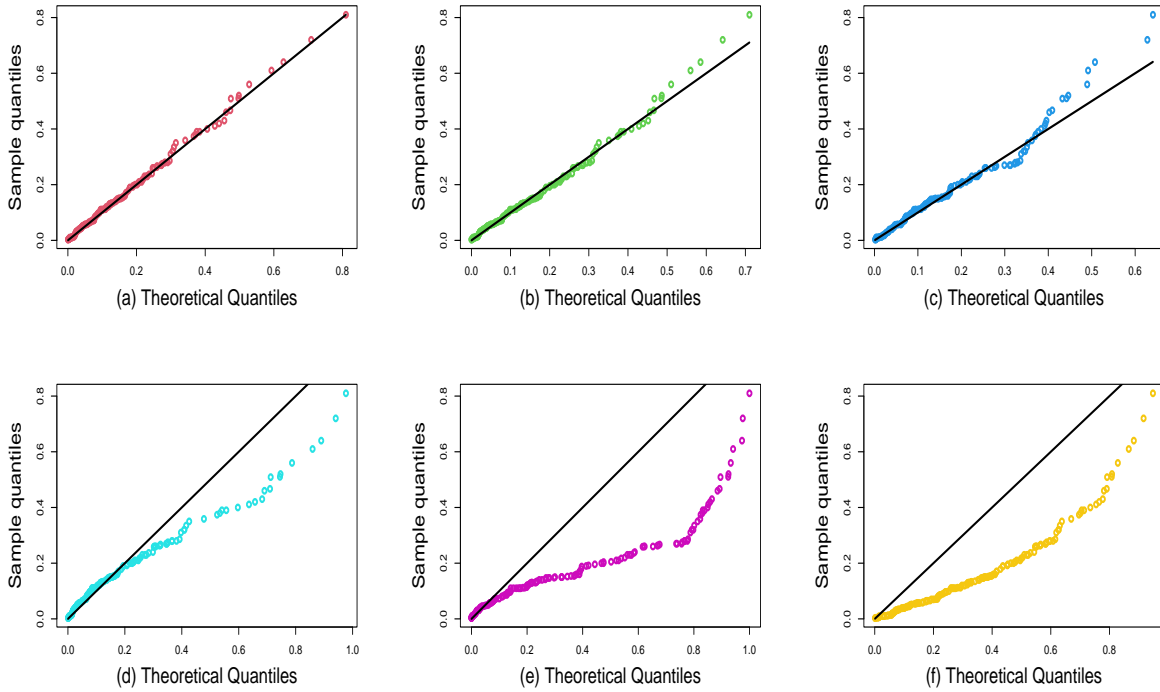
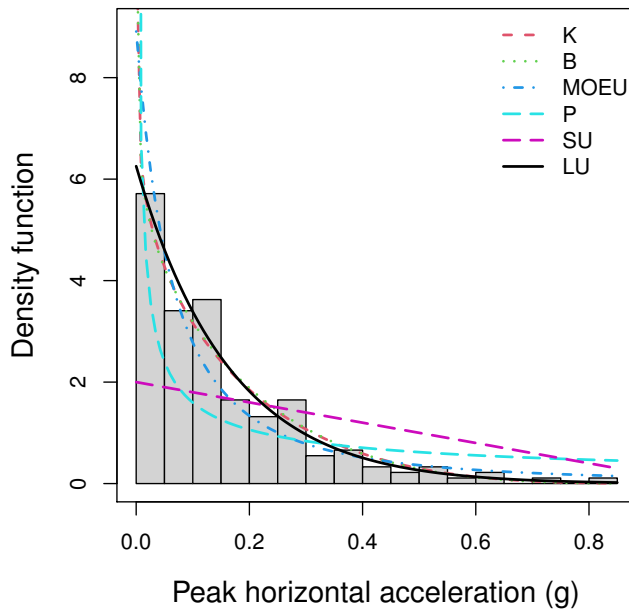


Figure 9 – Histogram of the peak horizontal acceleration data fitted with the LU, K, B, MOEU, P and SU distributions.



### 2.7.2 Risk Managements Practice Data

In this example, we consider the data frame presented by [Schmit and Roth \(1990\)](#) consisting of observations of seven variables (73 observations for each variable) consulted by means of a questionnaire sent to 374 risk managers of large US-based organizations. The variables consulted are described below: FI represents the measure of the firm's risk management cost effectiveness, defined as total property and casualty premiums and uninsured losses as a percentage of total assets; AS represents the per occurrence retention amount as a percentage of total assets; CA indicates that the firm owns a captive insurance company; SI represents the logarithm of total assets; IN represents a measure of the firm's industry risk; CE represents a measure of the importance of the local managers in choosing the amount of risk to be retained; and SO represents a measure of the degree of importance in using analytical tools.

[Gómez-Déniz, Sordo and Calderín-Ojeda \(2014\)](#) considered a Log-Lindley regression model to quantify the effect of the variables AS, CA, SI, IN, CE and SO on the mean of the variable FI. In our case, we consider the LU quantile regression model to quantify the effect of such covariates on the 0.25th, 0.5th and 0.75th quantiles of the variable FI, providing a very informative scenario (which complements the one proposed by [Gómez-Déniz, Sordo and Calderín-Ojeda \(2014\)](#)) to explain the behavior of the FI response in terms of the covariates already described. We observe that the histogram of the variable FI shows a decreasing behavior. Thus, we hope that the LU model can appropriately quantify the effect of the covariates on the 0.25th, 0.5th and 0.75th quantiles of the response variable.

As already mentioned, we compare the results with those obtained with the K and ASHN quantile regression models. The regression structure assumed for  $\eta_i$  is given by  $\text{logit}(\eta_i) = \beta_1(q) + \beta_2(q) AS_i + \beta_3(q) CA_i + \beta_4(q) SI_i + \beta_5(q) IN_i + \beta_6(q) CE_i + \beta_7(q) SO_i$ ,  $i = 1, \dots, 73$ , where  $\eta_i$  denotes the  $q$ th quantile of the LU, K and ASHN distributions.

Table 2 reports the  $\ell$ , AIC and BIC values for the ASHN, K and LU models fitted to the risk managements practice data. The  $p$ -values of the AD and CvM tests of the hypothesis that the randomize residuals follow a standard normal distribution are also reported in Table 2. In this table, we see that the  $\ell$ , AIC and BIC values change as the value of  $q$  changes. This shows that the rate of change in the conditional quantile of the response FI, expressed by the regression coefficients, depends on the value of  $q$ . On the other hand, based on the  $p$ -values and considering a significance level of 5%, we observe that the assumption of normality of the randomize residuals of the LU and K models is not rejected. Thus, under this significance level, the global fit of these models is appropriate. Furthermore, we observe that the LU model is the one with the highest  $\ell$  value and the one with the lowest AIC and BIC values, suggesting that this model should be selected to quantify the effect of the covariates on the 0.25th, 0.5th and 0.75th quantiles of the response.

Table 3 reports the estimates for the coefficients of the LU models fitted to the risk

Table 2 – The  $\ell$ , AIC, CAIC and BIC values for the ASHN, K and LU quantile regression models fitted to the risk managements practice data and the  $p$ -values of the AD and CvM tests for the randomize residuals.

$q$	Model	$\ell$	Criterion		$p$ -Value for the	
			AIC	BIC	AD Test	CvM Test
0.25	ASHN	80.3	-144.7	-126.4	<0.001	<0.001
	K	97.9	-179.8	-161.5	0.166	0.198
	LU	107.8	-201.6	-185.6	0.134	0.231
0.5	ASHN	80.1	-144.2	-125.9	<0.001	<0.001
	K	98.8	-181.6	-163.3	0.150	0.169
	LU	108.1	-202.2	-186.2	0.154	0.252
0.75	ASHN	81.8	-147.7	-129.4	<0.001	<0.001
	K	99.9	-183.9	-165.6	0.151	0.147
	LU	108.6	-203.2	-187.2	0.158	0.230

managements practice data. In addition, the  $z$  statistics and the  $p$ -values of the significance tests of the individual regression coefficients are reported. Here, we observe that the covariates SI and IN (the logarithm of total assets and the measure of the firm's industry risk) are statistically significant at usual nominal levels. Additionally, it is important to point out that there is a negative relationship between the 0.25th, 0.5th and 0.75th quantiles of the response (the firm's risk management cost effectiveness) and the covariate SI, while there is a positive relationship between the 0.25th, 0.5th and 0.75th quantiles of the response and the covariate IN. On the other hand, the covariates AS, CA, CE and SO are not significant. Figure 10 shows the estimates of the coefficients with the 95% confidence intervals of the LU regression model assuming different values for  $q$ . Here, we observe that the estimates of the coefficients of the covariates SI and IN decrease and increase, respectively, distancing themselves more and more from the value 0 as  $q$  increases. This illustrates a greater effect of these covariates on the high quantiles of the firm's risk management cost effectiveness.

## 2.8 Final Comments

In this section, we propose a new asymmetric one-parameter distribution for the modeling of bounded data whose histograms show increasing or decreasing behavior. The new distribution, called the Lambert-uniform distribution (LU), arises from a nonlinear transformation of a uniform rv bounded to the interval  $(0, 1)$ , where the transformation is expressed in term of the principal branch of the Lambert  $W$  function. An important aspect to highlight about the LU distribution is that its pdf tends to finite values at the ends of the support, which allows the extreme sample quantiles to be adequately modeled.

We derive the main structural properties of the LU distribution, including the moment

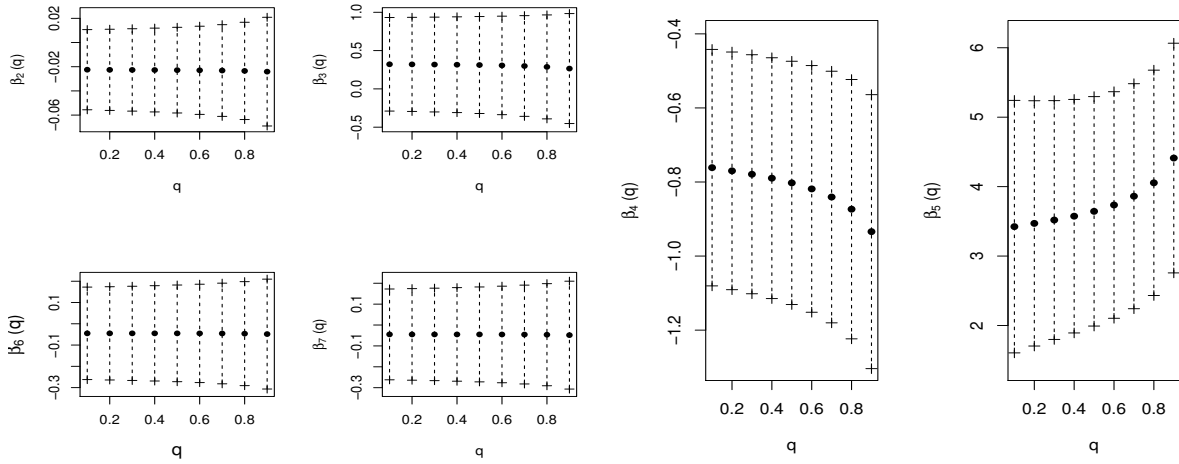


Figure 10 – Coefficient estimates and its 95% confidence intervals for variables AS, CA, SI, IN, CE and SO in different LU quantile regression models considering  $q = 0.1, 0.2, \dots, 0.9$  and response variable FI.

Table 3 – Coefficient estimates for the LU quantile regression model fitted to the risk managements practice data and significance tests of individual regression coefficients.

$q$	Parameter	Estimate	SE	$z$	$p$ -Value
0.25	Intercept	1.619	1.543	1.049	0.293
	AS	-0.022	0.017	-1.311	0.189
	CA	0.318	0.314	1.013	0.310
	SI	-0.774	0.164	-4.717	<0.001
	IN	3.494	0.888	3.932	<0.001
	CE	-0.044	0.112	-0.399	0.689
	SO	-0.009	0.028	-0.322	0.747
0.50	Intercept	2.730	1.601	1.705	0.088
	AS	-0.022	0.018	-1.263	0.206
	CA	0.311	0.322	0.965	0.334
	SI	-0.802	0.167	-4.788	<0.001
	IN	3.644	0.842	4.323	<0.001
	CE	-0.044	0.115	-0.388	0.697
	SO	-0.009	0.029	-0.333	0.738
0.75	Intercept	3.850	1.715	2.245	0.024
	AS	-0.023	0.019	-1.173	0.240
	CA	0.293	0.340	0.863	0.387
	SI	-0.855	0.175	-4.869	<0.001
	IN	3.946	0.826	4.777	<0.001
	CE	-0.045	0.122	-0.374	0.707
	SO	-0.010	0.031	-0.347	0.728

generating function that is used to describe the features of skewness and kurtosis. The quantile function of the LU distribution can be expressed in terms of Lambert  $W$  function, which allows the pdf to be reparameterized in terms of its  $q$ th quantile, resulting in a pdf with a simple analytic structure, easy to compute. Thus, we propose the LU quantile regression model that relates the  $q$ th quantile of the response to a linear predictor through a link function. The parameter estimation for the cases with and without covariates is performed with the maximum likelihood method. The estimators of the parameters do not have a closed form, so the use of some computational routine is required to obtain the estimates. We use the `optim` function in the R programming language to obtain the estimates. We evaluate the behavior of the estimators through two simulation studies, concluding that the maximum likelihood method provides acceptable estimates. Finally, we present two application examples in order to illustrate the usefulness of the proposal. In the first, considering data whose histogram shows a decreasing behavior, we illustrate that the LU distribution can present a better fit than the P, MOEU, SU, B and K distributions. In the second example, based on a real data frame, we illustrate that a quantile regression model formulated from the LU distribution performs well when modeling the 0.25th, 0.5th and 0.75th quantiles of the response (given a set of covariates), being a viable alternative to the other models such as the ASHN and K quantile regression models.

---

## LAMBERT-F DISTRIBUTION GENERATOR

---

In this chapter, we propose a new distribution generator formulated from the LU distribution. For any continuous baseline distribution  $F$ , the corresponding Lambert- $F$  version is generated by using the new generator. The result is a new class of distributions with one extra parameter that generalizes the baseline distribution and whose quantile function can be expressed in closed form in terms of the Lambert  $W$  function. The hazard rate function of a Lambert- $F$  distribution corresponds to a modification of the baseline hazard rate function, greatly increasing or decreasing the baseline hazard rate for earlier times. We study the main structural properties of the new class of distributions. Special attention is given to two particular cases that can be understood as two-parameter extensions of the well-known exponential and Rayleigh distributions. We discuss parameter estimation for the proposed models considering the maximum likelihood method. Finally, two applications are developed to illustrate the usefulness of the proposed distributions in the analysis of data from different real settings.

### 3.1 Motivation

One of the popular methods for constructing continuous probability distributions is to define a new cdf by considering an appropriate transformation of the cdf of a rv of interest.

More specifically, if  $X$  is a rv with cdf  $F(x; \theta)$ , parameterized by the vector  $\theta$ , then

$$F_X(x; \theta, \delta) = r(F(x; \theta), \delta), \quad (3.1)$$

where  $\delta$  is an extra parameter vector and  $r(\cdot, \cdot)$  is a nondecreasing function such that  $r(u, \delta) \rightarrow 0$  as  $u \downarrow 0$ , and  $r(u, \delta) \rightarrow 1$  as  $u \uparrow 1$ , for  $u \in (0, 1)$ , is a cdf.

In the literature, Equation (3.1) is usually referred to as a distribution generator. Once the analytic structure of the function  $r(\cdot, \cdot)$  is established, each choice of  $F(\cdot; \cdot)$  gives rise to a more flexible cdf  $F_X(\cdot; \cdot, \cdot)$  by possessing a larger parameter dimension. For example, if  $r(u, \delta = \alpha) = u^\alpha$ ,  $\alpha > 0$ , then Equation (3.1) corresponds to the cdf of the exponentiated distributions

class; see Gupta and Kundu (1999), Nadarajah and Kotz (2006), Al-hussaini (2010), Castillo *et al.* (2018) and Gómez-Déniz *et al.* (2019), among others. If  $r(u, \delta = (\alpha, \beta)) = I_u(\alpha, \beta)$ , where  $\alpha, \beta > 0$  and  $I_u(\cdot, \cdot)$  denotes the incomplete beta function ratio, then Equation (3.1) corresponds to the cdf of the beta-generated distributions class; see Eugene, Lee and Famoye (2002), Jones (2004), Jones (2007) and Nadarajah and Kotz (2004), among others. Similarly, other alternative expressions for  $r(u)$  can be found in Marshall and Olkin (1997), Shaw and Buckley (2007), Zografos and Balakrishnan (2009), Cordeiro and de Castro (2011), Lai (2013), to name just a few.

In what follows, we study a special case of Equation (3.1) obtained by considering a new analytic expression for the function  $r(u, \delta)$ , the expression  $r(u, \delta = \alpha) = 1 - (1 - u)\alpha^u$ , for  $\alpha \in (0, e)$ . The proposed generator can be understood as the cdf of a rv represented by the qf of an arbitrary baseline distribution, where the input of the qf corresponds to a rv with LU distribution.

## 3.2 The Generator

In this section, we propose the Lambert-F distribution generator and derive the corresponding density, survival, hazard rate and quantile functions.

**Proposition 3.** Let  $Z$  be a continuous baseline rv with cdf  $F(z; \eta)$  and qf  $F^{-1}(z; \eta)$ , where  $\eta$  is a parameter vector. Then, the cdf of  $X = F^{-1}(U; \eta)$ , where  $U \sim \text{LU}(\alpha)$ , is given by

$$F_X(x; \eta, \alpha) = 1 - [1 - F(x; \eta)]\alpha^{F(x; \eta)}, \quad \alpha \in (0, e). \quad (3.2)$$

**Proof.** If  $X = F^{-1}(U; \eta)$ , where  $U \sim \text{LU}(\alpha)$  and  $F^{-1}(\cdot; \eta)$  is the qf of  $Z$ , then the cdf of  $X$  is given by  $F_X(x; \eta, \alpha) = P(X \leq x) = F_U(F(x; \eta); \alpha)$ , where  $F(\cdot; \eta)$  is the cdf of  $Z$  and  $F_U(\cdot; \alpha)$  is the cdf of  $U$ .  $\square$

We refer to Equation (3.2) as the Lambert-F distribution generator since for each choice of  $F(\cdot; \eta)$  a new cdf is generated. We denote  $X \sim \text{LF}(\eta, \alpha)$  to indicate that a rv  $X$  has a cdf as in Equation (3.2).

Note that the cdf  $F_X(x; \eta, \alpha)$  inherits the support of the baseline cdf  $F(x; \eta)$  and that  $F_X(x; \eta, \alpha) = F(x; \eta)$  when  $\alpha = 1$ . Since the baseline cdf is arbitrary (without any conditions on the type of support), the Lambert-F generator can be used to obtain new distributions with unbounded or semi-bounded (positive) support.

**Corollary 4.** Let  $X \sim \text{LF}(\eta, \alpha)$ . Then, the probability density function (pdf), the survival function (sf) and the hazard rate function (hrf) of  $X$  are given, respectively, by

$$f_X(x; \eta, \alpha) = f(x; \eta)\alpha^{F(x; \eta)} [1 - \log(\alpha)S(x; \eta)], \quad (3.3)$$

$$S_X(x; \eta, \alpha) = S(x; \eta)\alpha^{F(x; \eta)} \quad \text{and} \quad (3.4)$$

$$h_X(x; \eta, \alpha) = h(x; \eta)[1 - \log(\alpha)S(x; \eta)], \quad (3.5)$$



where  $f(x; \eta)$ ,  $F(x; \eta)$ ,  $S(x; \eta) = 1 - F(x; \eta)$  and  $h(x; \eta) = f(x; \eta)/S(x; \eta)$  are the pdf, the cdf, the sf and the hrf of the baseline distribution, respectively.

We derive a linear representation for the pdf of the Lambert- $F$  distributions using the binomial theorem and the exponential expansion series.

**Proposition 4.** Let  $X \sim \text{LF}(\eta, \alpha)$ . Then, the pdf of  $X$  can be expressed as

$$f_X(x; \eta, \alpha) = \sum_{k=0}^{\infty} c_k(\alpha) f_E(x; \eta, k), \quad (3.6)$$

where

$$c_k(\alpha) = (-1)^{k+1} \sum_{i=k-1}^{\infty} \frac{\alpha[-\log(\alpha)]^i}{i!} \binom{i+1}{k},$$

and  $f_E(x; \eta, k) = k f(x; \eta) F^{k-1}(x; \eta)$  is the cdf of the exponentiated distributions with baseline parameter vector  $\eta$  and shape parameter  $k$  (NADARAJAH; KOTZ, 2006).

**Proof.** We observe that Equation (3.4) can be written as

$$\begin{aligned} S_X(x; \eta, \alpha) &= \alpha[1 - F(x; \eta)] \exp\{-\log(\alpha)[1 - F(x; \eta)]\} \\ &= \alpha \sum_{i=0}^{\infty} \frac{[-\log(\alpha)]^i}{i!} [1 - F(x; \eta)]^{i+1} \\ &= \sum_{k=1}^{\infty} \sum_{i=k-1}^{\infty} \frac{[-\log(\alpha)]^i}{i!} (-1)^k \binom{i+1}{k} F^k(x; \eta). \end{aligned}$$

Finally, the result is obtained by calculating  $-\partial S_X(x; \eta, \alpha)/\partial x$ .  $\square$

Equation (3.6) reveals that the pdf of the Lambert- $F$  distribution can be expressed as an infinite sum of exponentiated pdf's.

From Corollary 4, we note that (due to the performance of extra parameter  $\alpha$ ) the pdf and the hrf given in Equations (3.3) and (3.5) exhibit a wider variety of shapes than those exhibited by corresponding baseline functions. These shapes are discussed in detail in the next section.

### 3.3 Shapes

The shapes of the pdf and the hrf presented in Corollary 4 can be described analytically. First of all, we see that (i)  $f_X(0; \eta, \alpha) = [1 - \log(\alpha)]f(0; \eta)$ ; (ii)  $h_X(0; \eta, \alpha) = [1 - \log(\alpha)]h(0; \eta)$  and  $h_X(\infty; \eta, \alpha) = h(\infty; \eta)$ .

Secondly, the critical points of the pdf of  $X$  are the roots of the equation

$$a_1 f'(x; \eta) + a_2 f(x; \eta)(1 + a_1) = 0, \quad (3.7)$$

where  $a_1 = 1 - \log(\alpha)S(x; \eta)$  and  $a_2 = \log(\alpha)f(x; \eta)$ . There may exist more than one root of Equation (3.7). If  $x = x_0$  is a root of Equation (3.7), then it corresponds to a local maximum, a

local minimum or a point of inflexion depending on whether  $\varphi(x_0) < 0$ ,  $\varphi(x_0) > 0$  or  $\varphi(x_0) = 0$ , where  $\varphi(x) = \alpha^{F(x;\eta)} \{(a_1 + 1)a_2[3f'(x;\eta) + a_2f(x;\eta)] + a_2^2f(x;\eta) + a_1f''(x;\eta)\}$ . The critical points of the hrf of  $X$  are the roots of the equation

$$\left[ \frac{f'(x;\eta)}{S(x;\eta)} + h^2(x;\eta) \right] a_1 + a_2h(x;\eta) = 0. \quad (3.8)$$

There may exist more than one root of Equation (3.8). If  $x = x_0$  is a root of Equation (3.8), then it corresponds to a local maximum, a local minimum or a point of inflexion depending on whether  $\tau(x_0) < 0$ ,  $\tau(x_0) > 0$  or  $\tau(x_0) = 0$ , where

$$\tau(x) = \left[ \frac{f''(x;\eta)}{S(x;\eta)} + \frac{3f'(x;\eta)h(x;\eta)}{S(x;\eta)} + 2h^3(x;\eta) \right] a_1 + \left[ \frac{3f'(x;\eta)}{S(x;\eta)} + 2h^2(x;\eta) \right] a_2.$$

In the case where the baseline distribution is symmetric, it can be verified that the parameter  $\alpha$  acts as a skewness parameter allowing asymmetric shapes for the resulting pdf. Thus, the Lambert- $F$  generator can be used to obtain asymmetric distributions from symmetric baseline distributions, see Section 4.

Giving greater attention to the hrf of the Lambert- $F$  distributions, we point out the following:

1. The hrf of the Lambert- $F$  distribution approximates the hrf of the baseline distribution  $F$  when  $x$  is large enough, that is,  $h_X(x;\eta, \alpha)/h(x;\eta) \rightarrow 1$  as  $x \rightarrow \infty$ .
2. The hrf of the Lambert- $F$  distribution is greater than the hrf of the baseline distribution  $F$  if and only if  $\alpha \in (1, e)$ .
3. If  $\alpha \in (1, e)$  and the hrf of the baseline distribution  $F$  is a nondecreasing monotonic function, then the hrf of the Lambert- $F$  distribution is a nondecreasing monotonic function.

In view of the above, the application of the proposed Lambert transformation to a baseline distribution has a simple justification in terms of the hrf of the resulting model. The hrf induced by the Lambert- $F$  generator is more distant from the baseline hrf for lower values of  $X$  (earlier times).

### 3.4 Moments

**Corollary 5.** Let  $X \sim \text{LF}(\eta, \alpha)$ . Then, from Equation (3.6), the  $r$ -th raw moment for the Lambert- $F$  distribution is given by

$$\mathbb{E}(X^r) = \sum_{k=0}^{\infty} c_k(\alpha) m_r(\eta, k), \quad r = 1, 2, 3, \dots, \quad (3.9)$$

where  $m_r(\eta, k)$  is the  $r$ -th raw moment of the exponentiated distributions with baseline parameter vector  $\eta$  and shape parameter  $k$ .

Therefore, the  $r$ -th raw moment for a Lambert- $F$  distribution can be easily derived from Equation (3.9) when the exponentiated version of the baseline distribution  $F$  is known. For example, considering the exponential baseline distribution with hrf equal to  $1/\lambda > 0$ , and using the expression for the  $r$ -th raw moments of the exponentiated-exponential distribution given in Nadarajah (2011), the  $r$ -th raw moment of the corresponding Lambert-exponential (LE) distribution is given by

$$\mathbb{E}(X^r) = \sum_{k=0}^{\infty} (-1)^r \lambda^r k c_k(\alpha) \frac{\partial^r B(k, p+1-k)}{\partial p^r} \Big|_{p=k}.$$

Then, the mean and variance of the LE distribution are given by

$$\begin{aligned} \mathbb{E}(X) &= \lambda \sum_{k=0}^{\infty} c_k(\alpha) [\psi(k+1) + \gamma] \quad \text{and} \\ \text{Var}(X) &= \lambda^2 \sum_{k=0}^{\infty} \frac{c_k(\alpha) [\pi^2 - 6\psi'(k+1)]}{6}. \end{aligned}$$

where  $\psi(\cdot)$ ,  $\psi'(\cdot)$  and  $\gamma$  denote the gamma and digamma functions and Euler's constant, respectively.

The LE distribution will be discussed in more detail in Section 3.7. The following result is useful for cases where the exponentiated version of the baseline distribution  $F$  is not known.

**Proposition 5.** Let  $X \sim \text{LF}(\eta, \alpha)$ . Then, for  $r = 1, 2, \dots$ , the  $r$ -th raw moment of  $X$  can be written as

$$\mu_r = \mathbb{E}(X^r) = \int_0^1 \alpha^u [F^{-1}(u; \eta)]^r [1 - \log(\alpha)(1-u)] du, \quad r = 1, 2, \dots, \quad (3.10)$$

where  $F^{-1}(\cdot; \eta)$  denotes the qf of the baseline distribution.

**Proof.** From the Lambert- $F$  pdf given in Corollary 4, we have that  $\mu_r = \mathbb{E}(X^r) = \int_0^{\infty} x^r f(x; \eta) \alpha^{F(x; \eta)} [1 - \log(\alpha)\{1 - F(x; \eta)\}] dx$ , and applying the change of variable  $u = F(x; \eta)$ , the result is obtained.

## 3.5 Quantile Function

**Proposition 6.** Let  $X \sim \text{LF}(\eta, \alpha)$ . Then, the qf of  $X$  is given by

$$Q_X(u) = \begin{cases} F^{-1} \left[ \frac{1}{\log(\alpha)} W_0 \left( \frac{\log(\alpha)(u-1)}{\alpha} \right) + 1; \eta \right], & \text{if } \alpha \neq 1, \\ F^{-1}(u; \eta), & \text{if } \alpha = 1, \end{cases} \quad (3.11)$$

where  $F^{-1}(\cdot; \eta)$  is the qf of the baseline distribution and  $W_0(\cdot)$  is the principal branch of the Lambert  $W$  function.

**Proof.** From Equation (3.2) it is observed that the baseline cdf can be written as  $F(x; \eta) = [\log(\alpha)]^{-1} W_0[\alpha^{-1}(u-1)\log(\alpha)] + 1$ . Then, solving this equation with respect to  $x$ , we obtain the expression in Equation (3.11) for  $\alpha \neq 1$ . In the special case  $\alpha = 1$ , the result is obtained directly by calculating the inverse function of the baseline cdf.

For  $\alpha \in (0, 1)$ , we have that  $\log(\alpha)(u-1)/\alpha > 0$ . Then, the Lambert  $W$  function has a unique solution (principal branch). For  $\alpha \in (1, e)$ , we have that  $\log(\alpha)(u-1)/\alpha \in (-1/e, 0)$ . Then, the Lambert  $W$  function has two solutions but only the solution of the principal branch guarantees that the input values of the baseline qf correspond to values in the interval  $(0, 1)$ .

### 3.6 Stochastic Order

The ordering of continuous random variables is an important tool for judging comparative behavior. It is well known that a random variable  $X$  is smaller than a random variable  $Y$  in the stochastic order ( $X \leq_{st} Y$ ) if  $F_X(x) \geq F_Y(x)$  for all  $x$ , in the hazard rate order ( $X \leq_{hr} Y$ ) if  $h_X(x) \geq h_Y(x)$  for all  $x$ , and in the likelihood ratio order ( $X \leq_{lr} Y$ ) if  $f_X(x)/f_Y(x)$  decreases in  $x$ . Also, the implications

$$X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{st} Y \quad (3.12)$$

are well known (SHAKED; SHANTHIKUMAR, 2007).

**Proposition 7.** Let  $X_1 \sim \text{LF}(\eta, \alpha_1)$  and  $X_2 \sim \text{LF}(\eta, \alpha_2)$ . If  $\alpha_1 > \alpha_2$ , then  $X_1 \leq_{lr} X_2$  and therefore  $X_1 \leq_{hr} X_2$  and  $X_1 \leq_{st} X_2$ .

**Proof.** First, notice that  $f_{X_2}(x; \eta, \alpha_2)/f_{X_1}(x; \eta, \alpha_1) = (\alpha_2/\alpha_1)^{F(x;\eta)} \{1 - \log(\alpha_2)[1 - F(x; \eta)]\} / \{1 - \log(\alpha_1)[1 - F(x; \eta)]\}$  is a nondecreasing function if and only if  $\mu'(x) \geq 0$ , where  $\mu(x) = \{1 - \log(\alpha_2)[1 - F(x; \eta)]\} / \{1 - \log(\alpha_1)[1 - F(x; \eta)]\}$ . Some calculations show that  $\mu'(x) = \log(\alpha_2/\alpha_1) f(x; \eta) \{1 - \log(\alpha_1)[1 - F(x; \eta)]\}^{-2}$ .

Now, note that  $\alpha_1 > \alpha_2$  implies  $\mu'(x) < 0$ , that is,  $f_{X_2}(x; \eta)/f_{X_1}(x; \eta)$  is decreasing in  $x$ , which means  $X_1 \leq_{lr} X_2$ . The remaining affirmations are derived from the implications in Equation (3.12).

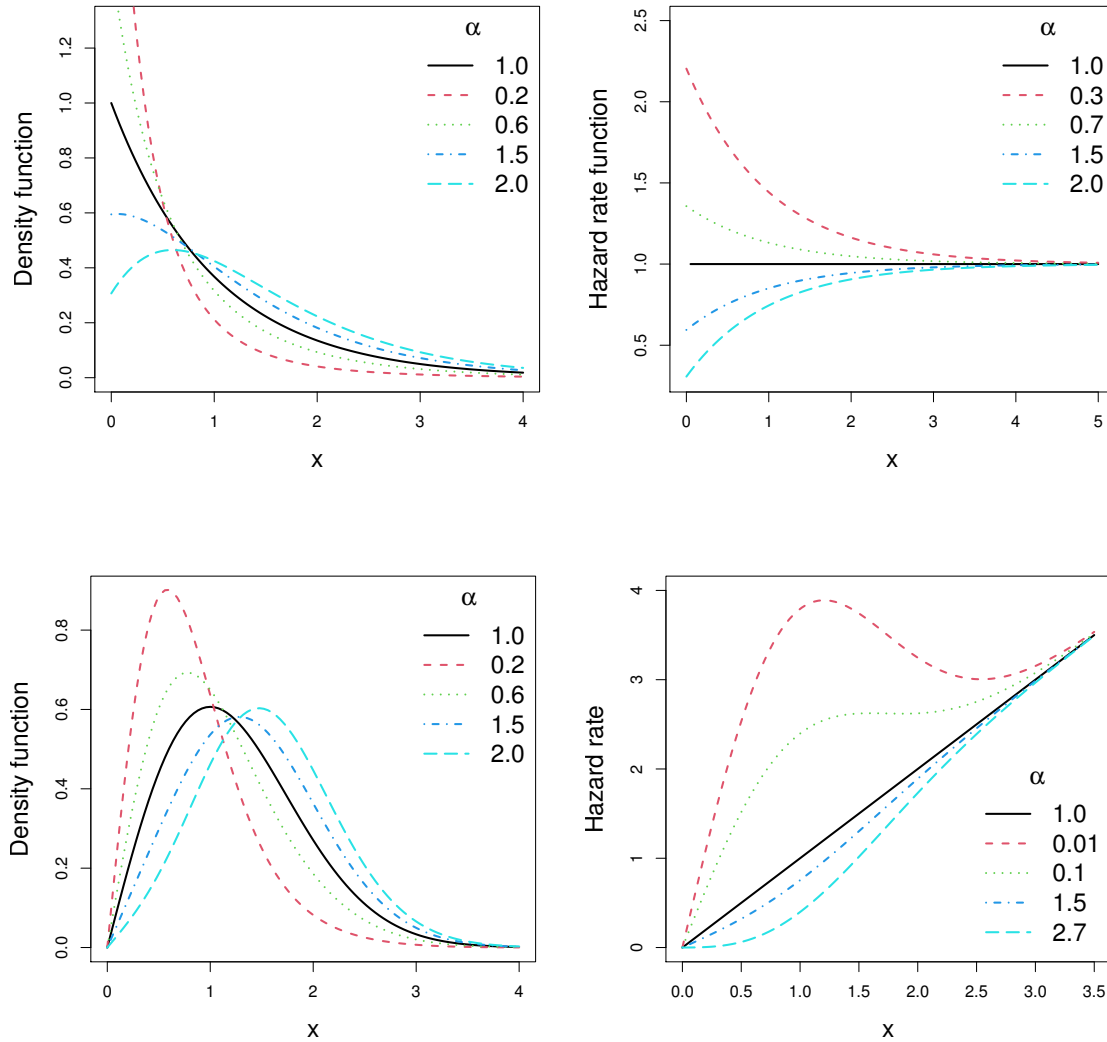
A direct consequence of Proposition 7 is that the hrf of the Lambert- $F$  distribution is less than the hrf of the baseline  $F$  distribution when  $\alpha \in (1, e)$ , which is consistent with what was observed in Section 3.3.

### 3.7 Two Special Cases of Positive Supports

Positive support distributions are commonly considered in the analysis of a great variety of random phenomena. For example, when analyzing the stature of individuals in a population or the lifetime of a certain type of machine. The exponential (E), Rayleigh (R), Gamma (G) and Weibull (W) (JOHNSON; KOTZ; BALAKRISHNAN, 1994; WEIBULL, 1939) distributions are some of the distributions usually considered in data modeling associated with these phenomena.

In this section, we propose two new positive support distributions generated from the Lambert- $F$  generator by considering the classic E and R distributions (See Appendix A) as

Figure 11 – Plots of the density and hazard rate functions of the LE and LR distributions for  $\sigma = 1$  and different values of  $\alpha$ .



baseline distributions. The resulting distributions (having two parameters, scale and shape) can be considered as natural alternatives to the W and G distributions.

**Definition 2.** The random variable  $X$  follows the Lambert-exponential distribution, with scale parameter  $\sigma > 0$  and shape parameter  $\alpha \in (0, e)$ , denoted as  $X \sim \text{LE}(\sigma, \alpha)$ , if its pdf and hrf for  $x > 0$  are given, respectively, by

$$f_X(x) = \frac{1}{\sigma} e^{-\frac{x}{\sigma}} \alpha^{1-e^{-\frac{x}{\sigma}}} [1 - \log(\alpha) e^{-\frac{x}{\sigma}}] \quad \text{and} \quad h_X(x) = \frac{1}{\sigma} [1 - \log(\alpha) e^{-\frac{x}{\sigma}}].$$

Figure 11 (top panels) displays some plots of the pdf and the hrf of the the LE distribution for  $\sigma = 1$  and different values of shape parameter. In the figure, we can see that the LE pdf can have monotonic or unimodal shapes depending on the parameter  $\alpha$ .

**Definition 3.** The random variable  $X$  follows the Lambert-Rayleigh distribution, with scale parameter  $\sigma > 0$  and shape parameter  $\alpha \in (0, e)$ , denoted as  $X \sim \text{LR}(\sigma, \alpha)$ , if its pdf and hrf for

$x > 0$  are given, respectively, by

$$f_X(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \alpha^{1-e^{-\frac{x^2}{2\sigma^2}}} [1 - \log(\alpha) e^{-\frac{x^2}{2\sigma^2}}] \quad \text{and} \quad h_X(x) = \frac{x}{\sigma^2} [1 - \log(\alpha) e^{-\frac{x^2}{2\sigma^2}}].$$

Figure 11 (bottom panels) displays some plots of the pdf and the hrf of the LR distribution for  $\sigma = 1$  and different values of shape parameter. In the figure, we can see that the LR pdf exhibits unimodal shapes with variations of skewness depending on the parameter  $\alpha$ .

The new two-parameter distributions described above are members of the well-known and popular shape and scale distributions family. The scale  $\sigma$  in both models is inherited from the respective baseline distribution, while the shape parameter  $\alpha$  arises from the application of the Lambert transformation to the baseline distribution. From the results in Section 3.3, some details are provided below regarding the shapes exhibited by these distributions.

**Corollary 6.** Let  $X \sim \text{LE}(\sigma, \alpha)$ . Then:

1.  $f_X(0) = [1 - \log(\alpha)]/\sigma$ ,  $h_X(0) = [1 - \log(\alpha)]/\sigma$  and  $h_X(\infty) = 1/\sigma$ ;
2. Considering  $a = \exp[(3 - \sqrt{5})/2]$ , the pdf of  $X$  is a monotonically decreasing function when  $\alpha \in (0, a]$  or  $\alpha = 1$ , and it is a unimodal function for  $\alpha \in (a, 1)$  or  $\alpha \in (1, e)$ . The mode of  $X$  is given by  $\text{Mod}(X) = \sigma \{\log[2 \log(\alpha)] - \log(3 - \sqrt{5})\}$ ;
3. The hrf of  $X$  is a monotonically decreasing function when  $\alpha \in (0, 1)$ , a monotonically increasing function when  $\alpha \in (1, e)$  and a constant function for  $\alpha = 1$ .

**Corollary 7.** Let  $X \sim \text{LR}(\sigma, \alpha)$ . Then:

1.  $f_X(0) = 0$ ,  $h_X(0) = 0$ , and  $h_X(\infty) = \infty$ .
2. The pdf of  $X$  is a unimodal function for  $\alpha \in (0, e)$  (Mode without explicit analytic expression).
3. The hrf of  $X$  is a monotonically increasing function when  $\alpha \in [0.1, e)$  and presents the increasing-decreasing-increasing shape for  $\alpha \in (0, 0.1)$ . The local maximum and the local minimum are given by  $x = \sigma \{1 - 2W_j[e^{1/2}/\log(\alpha^2)]\}^{1/2}$ ,  $j \in \{0, -1\}$ , where  $W_0(\cdot)$  and  $W_{-1}(\cdot)$  denote the principal and non-principal branches of the Lambert  $W$  function, respectively.

From these results, we observe that the shapes of the pdf and the hrf for the LE distribution are similar to those presented by other two-parameter distributions such as the W and G distributions. It is important to note that the LE, W and G distributions are two-parameter extensions of the exponential distribution, so that the comparison between these models is quite natural. A similar observation can be made for the LR and W distributions because both are two-parameter extensions of the classical R distribution.

Table 4 – Comparison of the Lambert-exponential (LE), Lambert-Rayleigh (LR), Weibull (W) and gamma (G) distributions in terms of the pdf.

Distribution	Shape parameter interval	Shape of the pdf, $f_X(x)$	$\lim_{x \rightarrow 0^+} f_X(x)$
LE	$(0, \exp^{\frac{3-\sqrt{5}}{2}})$ or $\{1\}$	Decreasing	$\frac{1}{\sigma}[1 - \log(\alpha)]$
	$(\exp^{\frac{3-\sqrt{5}}{2}}, 1)$ or $(1, e)$	Unimodal	$\frac{1}{\sigma}[1 - \log(\alpha)]$
LR	$(0, e)$	Unimodal	0
W	$(0, 1]$	Decreasing	$\infty$ or $\frac{1}{\sigma}$
	$(1, \infty)$	Unimodal	0
G	$(0, 1]$	Decreasing	$\infty$ or $\frac{1}{\sigma}$
	$(1, \infty)$	Unimodal	0

In Tables 4 and 5, we present a comparison of the shapes of the pdf and the hrf of the LE and LR distributions with those of the W and G distributions. In Table 4, it is seen that the pdf of the LE distribution presents similar shapes to those of the W and G distributions. However, unlike the W and G models, the pdf of the LE distribution tends to  $\sigma^{-1}[1 - \log(\alpha)]$  as  $x \downarrow 0$  when it is unimodal, that is, it tends to a positive finite value. This leads us to establish that the LE distribution can properly fit data sets whose frequency distributions are unimodal but there are observations lumped around 0. On the other hand, the pdf of the LR distribution presents only the unimodal shape but (as will be seen later) with variations of skewness and kurtosis. In Table 5, it is seen that the shapes presented by the hrf of the LE distribution are similar to those of the W and G distributions. However, from the results presented in Table 5, and similar to the behavior of the LE pdf, an important difference is observed in the behavior of the LE hrf for lower values of  $x$  (times close to 0). On the other hand, the LR distribution is the only distribution (among the distributions considered) that has a hrf that can present the increasing-decreasing-increasing shape.

**Corollary 8.** If  $X_1 \sim \text{LE}(\sigma, \alpha)$  and  $X_2 \sim \text{LR}(\sigma, \alpha)$ , from Proposition 5, we obtain that (for  $r = 1, 2, \dots$ ) the  $r$ -th raw moment of  $X_j$  is given by  $E(X_j^r) = a_i(\sigma)m_{r,\alpha}(u_i)$ ,  $j = 1, 2$ , where  $m_{r,\alpha}(u_j) = \int_0^1 \alpha^u u_j^r [1 - \log(\alpha)(1 - u)] du$ ,  $a_1(\alpha) = \sigma^r$ ,  $a_2(\sigma) = (2\sigma^2)^{r/2}$ ,  $u_1 = -\log(1 - u)$  and  $u_2 = \sqrt{-\log(1 - u)}$ . Thus, the skewness ( $\beta_{1(X_j)}$ ) and kurtosis ( $\beta_{2(X_j)}$ ) coefficients for  $X_j$  are given by

$$\beta_{1(X_j)} = \frac{m_{3,\alpha}(u_j) - 3m_{1,\alpha}(u_j)m_{2,\alpha}(u_j) + 2m_{1,\alpha}^3(u_j)}{\left[m_{2,\alpha}(u_j) - m_{1,\alpha}^2(u_j)\right]^{3/2}} \text{ and}$$

$$\beta_{2(X_j)} = \frac{m_{4,\alpha}(u_j) - 4m_{1,\alpha}(u_j)m_{3,\alpha}(u_j) + 6m_{1,\alpha}^2(u_j)m_{2,\alpha}(u_j) - 3m_{1,\alpha}^4(u_j)}{\left[m_{2,\alpha}(u_j) - m_{1,\alpha}^2(u_j)\right]^2}.$$

Table 5 – Comparison of the Lambert-exponential (LE), Lambert-Rayleigh (LR), Weibull (W) and gamma (G) distributions in terms of the hrf.

Distribution	Shape parameter interval	Shape of the hrf, $h_X(x)$	$\lim_{x \rightarrow 0^+} h_X(x)$	$\lim_{x \rightarrow \infty} h_X(x)$
LE	(0, 1)	Decreasing	$\frac{1}{\sigma}[1 - \log(\alpha)]$	$\frac{1}{\sigma}$
	(1, $e$ )	Increasing	$\frac{1}{\sigma}[1 - \log(\alpha)]$	$\frac{1}{\sigma}$
LR	(0, 1)	Increasing-decreasing -increasing	0	$\infty$
	[1, $e$ )	Increasing	0	$\infty$
W	(0, 1)	Decreasing	$\infty$	0
	(1, $\infty$ )	Increasing	0	$\infty$
G	(0, 1)	Decreasing	$\infty$	$\frac{1}{\sigma}$
	(1, $\infty$ )	Increasing	0	$\frac{1}{\sigma}$

If  $\alpha = 1$ , the hrf of the LR distribution reduces to  $h(x) = x/\sigma^2$  (the Rayleigh hrf) and the hrf of the LE, W, and G distributions to  $h(x) = 1/\sigma$  (the exponential hrf).

Figure 12 shows plots of the skewness and kurtosis coefficients for both the LE and LR distributions as well as for the baseline distributions; E and R, respectively. Note that the skewness and kurtosis values of the baseline distributions are extended to a range of values in the respective Lambert versions, showing greater flexibility of these latter distributions. The asymmetry and kurtosis ranges for the LE distribution are (1.456, 4.461) and (6.416, 48.814), respectively, and for the LR distribution (0.342, 1.274) and (3.027, 6.005), respectively. These ranges were calculated by minimizing and maximizing the skewness and kurtosis coefficients. We use the integrate function in the R language to compute the  $m_{r,\alpha}(u_i)$  functions. The optimize function was used to minimize and maximize the coefficients.

A more detailed discussion on the structural properties of the LE and LR distributions can be found in [Iriarte, de Castro and Gómez \(2020\)](#).

### 3.8 Maximum Likelihood Estimation

In this section, we deal with the problem of parameter estimation in the Lambert-F distribution under the maximum likelihood (ML) method.

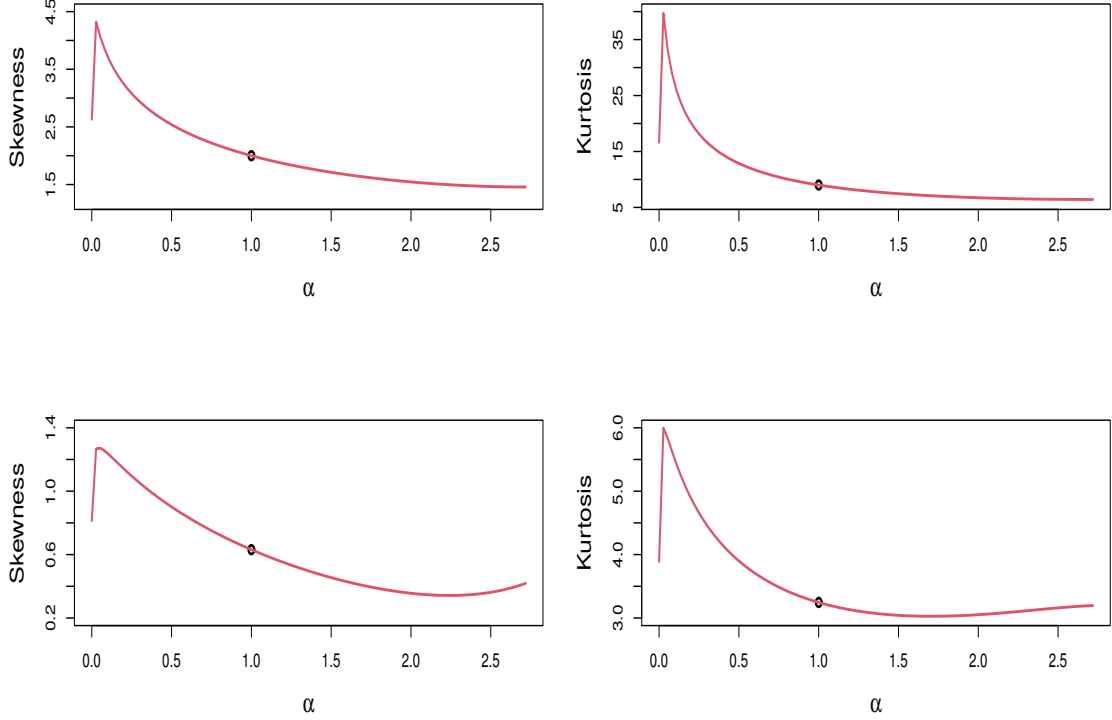
If  $X \sim \text{LF}(\theta)$ , with  $\theta = (\eta, \alpha)^t$  and  $\eta = (\eta_1, \dots, \eta_{k-1})^t$ , then the log-likelihood function is given by

$$\ell(\theta; x) = \log[f(x; \eta)] + F(x; \eta) \log(\alpha) + \log[1 - \log(\alpha)S(x; \eta)], \quad (3.13)$$

where  $f(x; \eta)$ ,  $F(x; \eta)$  and  $S(x; \eta)$  are the pdf, the cdf and the sf, respectively, of the baseline distribution.



Figure 12 – Plots of the skewness and kurtosis coefficients for the Lambert-exponential (red line) and exponential (circle) distributions (top panels) and for the Lambert-Rayleigh (red line) and Rayleigh (circle) distributions (bottom panels).



Thus, the score functions are given by

$$\frac{\partial \ell(\theta; x)}{\partial \eta_j} = \frac{f_{\eta_j}(x; \eta)}{f(x; \eta)} + \log(\alpha) F_{\eta_j}(x; \eta) - \frac{\log(\alpha) S_{\eta_j}(x; \eta)}{1 - \log(\alpha) S(x; \eta)} \quad \text{and} \quad (3.14)$$

$$\frac{\partial \ell(\theta; x)}{\partial \alpha} = \frac{1}{\alpha} F(x; \eta) - \frac{S(x; \eta)}{\alpha [1 - \log(\alpha) S(x; \eta)]}, \quad (3.15)$$

where  $f_{\eta_j}(x; \eta) = \partial f(x; \eta) / \partial \eta_j$ ,  $F_{\eta_j}(x; \eta) = \partial F(x; \eta) / \partial \eta_j$  and  $S_{\eta_j}(x; \eta) = \partial S(x; \eta) / \partial \eta_j$ , with  $j = 1, \dots, k - 1$ .

For a random sample  $X_1, \dots, X_n$  from  $X \sim \text{LF}(\theta)$ , we observe that the ML estimator  $\hat{\theta}$  for  $\theta$  cannot be expressed in closed form. The solution of the score equations gives rise to a system of nonlinear equations (See Appendix B) that must be solved with the help of some computational routine in search of ML estimates.

In this case, as the ML estimators do not have a closed form, a good alternative to obtain ML estimates is to solve the optimization problem

$$\max_{\theta} \sum_{i=1}^n \ell(\theta; x_i), \quad \text{subject to } \eta_j \in \Delta_j, \quad j = 1, \dots, k - 1, \quad \text{and } \alpha \in (0, e), \quad (3.16)$$

where  $\ell(\cdot; \cdot)$  is given in Equation (3.13).

Under regularity conditions, the asymptotic distribution of  $(\hat{\theta}_{\text{ML}} - \theta)$  is  $N_k(0, K(\theta)^{-1})$ , where  $K(\theta)$  is the expected information matrix. As the function  $\sum_{i=1}^n \ell(\theta; x_i)$  is not simple, it is

not easy to obtain the analytical expression of this matrix. However, we obtain an approximation from the observed information matrix, whose elements are computed as minus the second partial derivatives of the log-likelihood function with respect to all the parameters (evaluated at the ML estimates). Thus, for a random sample  $X_1, \dots, X_n$  from  $X \sim \text{LF}(\theta)$ , the observed information matrix is given by

$$I(\theta) = \begin{pmatrix} \varepsilon_{\eta_1 \eta_1} & \varepsilon_{\eta_1 \eta_2} & \cdots & \varepsilon_{\eta_1 \eta_{k-1}} & \varepsilon_{\eta_1 \alpha} \\ \varepsilon_{\eta_2 \eta_1} & \varepsilon_{\eta_2 \eta_2} & \cdots & \varepsilon_{\eta_2 \eta_{k-1}} & \varepsilon_{\eta_2 \alpha} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \varepsilon_{\eta_{k-1} \eta_1} & \varepsilon_{\eta_{k-1} \eta_2} & \cdots & \varepsilon_{\eta_{k-1} \eta_{k-1}} & \varepsilon_{\eta_{k-1} \alpha} \\ \varepsilon_{\alpha \eta_1} & \varepsilon_{\alpha \eta_2} & \cdots & \varepsilon_{\alpha \eta_{k-1}} & \varepsilon_{\alpha \alpha} \end{pmatrix}, \quad (3.17)$$

$$\varepsilon_{\theta_r \theta_p} = - \sum_{i=1}^n \frac{\partial^2 \ell(\theta; x_i)}{\partial \theta_r \partial \theta_p} \Big|_{\theta = \hat{\theta}_{\text{ML}}}, \quad r = p = 1, \dots, k,$$

where  $\theta_s = \eta_s$ , with  $s = 1, \dots, k-1$ ,  $\theta_k = \alpha$  and the second partial derivatives are given by

$$\begin{aligned} \frac{\partial^2 \ell(\theta; x)}{\partial \eta_m \partial \eta_j} &= \log(\alpha) F_{\eta_m \eta_j}(x; \eta) + \frac{f_{\eta_m \eta_j}(x; \eta)}{f(x; \eta)} - \frac{f_{\eta_m}(x; \eta) f_{\eta_j}(x; \eta)}{f^2(x; \eta)} \\ &\quad - \frac{\log(\alpha) S_{\eta_m \eta_j}(x; \eta)}{1 - \log(\alpha) S(x; \eta)} - \frac{\log^2(\alpha) S_{\eta_m}(x; \eta) S_{\eta_j}(x; \eta)}{[1 - \log(\alpha) S(x; \eta)]^2}, \\ \frac{\partial^2 \ell(\theta; x)}{\partial \alpha \partial \eta_j} &= \frac{F_{\eta_j}(x; \eta)}{\alpha} - \frac{S_{\eta_j}(x; \eta)}{\alpha [1 - \log(\alpha) S(x; \eta)]} - \frac{\log(\alpha) S(x; \eta) S_{\eta_j}(x; \eta)}{\alpha [1 - \log(\alpha) S(x; \eta)]}, \\ \frac{\partial^2 \ell(\theta; x)}{\partial \eta_j \partial \alpha} &= \frac{\partial^2 \ell(\theta; x)}{\partial \alpha \partial \eta_j}, \\ \frac{\partial^2 \ell(\theta; x)}{\partial \alpha^2} &= \frac{1}{\alpha^2 [1 - \log(\alpha) S(x; \eta)]} - \frac{F(x; \eta)}{\alpha^2} - \frac{S(x; \eta)}{\alpha^2 [1 - \log(\alpha) S(x; \eta)]}, \end{aligned}$$

where  $f_{\eta_m \eta_j}(x; \eta) = \partial^2 f(x; \eta) / \partial \eta_m \partial \eta_j$ ,  $F_{\eta_m \eta_j}(x; \eta) = \partial^2 F(x; \eta) / \partial \eta_m \partial \eta_j$  and  $S_{\eta_m \eta_j}(x; \eta) = \partial^2 S(x; \eta) / \partial \eta_m \partial \eta_j$ , with  $j = 1, \dots, k-1$ .

Partial derivatives  $f_{\eta_j}(x; \eta)$ ,  $f_{\eta_m \eta_j}(x; \eta)$ ,  $F_{\eta_j}(x; \eta)$  and  $F_{\eta_m \eta_j}(x; \eta)$  related to the special cases presented in Section 3.7 are presented in Appendix C. Although the ML estimates for the parameters of the LE and LR distributions can be obtained by numerically solving the corresponding systems of score equations, we prefer to obtain the ML estimates by solving the optimization problem formulated in Equation (3.16). For this, we use the optim function of the R programming language under the consideration of the L-BFGS-B algorithm. We verify through simulation experiments that  $(s_x, 1)$ , where  $s_x$  represents the standard deviation of the observations, is a good starting point for the iterative process.

In the presence of right-censoring, we adopt the following scheme. Assuming that for each individual the failure time is independent of the censoring time (say,  $Y_i$  and  $C_i$  for  $i = 1, \dots, n$  respectively). The observed times are given by  $X_i = \min(Y_i, C_i)$  and the failure indicator is denoted as  $\delta_i = I(Y_i \leq C_i)$ . Given a sample of observed times and failure indicators

Table 6 – Scenarios considered in the simulation studies for the LE and LR distributions.

Scenario	Parameters
A	: $\sigma = 1$ and $\alpha = 1.5$
B	: $\sigma = 1$ and $\alpha = 2$
C	: $\sigma = 1$ and $\alpha = 2.5$
D	: $\sigma = 2$ and $\alpha = 1.5$
E	: $\sigma = 3$ and $\alpha = 2$
F	: $\sigma = 4$ and $\alpha = 2.5$
G	: $\sigma = 0.5$ and $\alpha = 1.5$
H	: $\sigma = 1.5$ and $\alpha = 2$
I	: $\sigma = 2.5$ and $\alpha = 2.5$

$(t_1, \delta_1), (t_2, \delta_2), \dots, (t_n, \delta_n)$  and under the additional assumption of non-informative censoring, the log-likelihood function for a Lambert- $F$  type distribution is given by

$$\begin{aligned} \ell(\theta, x) = & \sum_{i=1}^n \delta_i \log f(x_i; \eta) + \log(\alpha) \sum_{i=1}^n \delta_i F(x_i; \eta) + \sum_{i=1}^n \delta_i \log[1 - \log(\alpha) S(x_i; \eta)] \\ & + \sum_{i=1}^n (1 - \delta_i) \log S(x_i; \eta) + \log(\alpha) \sum_{i=1}^n (1 - \delta_i) F(x_i; \eta), \end{aligned} \quad (3.18)$$

where  $f(\cdot; \cdot)$ ,  $F(\cdot; \cdot)$  and  $S(\cdot; \cdot)$  are the pdf, the cdf and the sf of the baseline distribution. For  $\delta_i = 1, i = 1, \dots, n$ , the Equation (3.18) reduces to a log-likelihood function considered in the optimization problem presented in Equation (3.16). Inference based on (3.18) can be performed in a similar manner as was done in the uncensored case, as described above.

### 3.9 Simulation Studies

In this section, we carry out two simulation studies in order to evaluate the behavior of the ML estimators for the parameters of the LE and LR distributions. For both distributions, from the qf given in Equation (3.11), we generate 1000 random samples considering different sample sizes ( $n$ ) and different choices for the parameters. The LambertW package (GOERG, 2011) in the R language was used to compute the principal branch of the Lambert  $W$  function. Table 12 presents the different scenarios considered for each distribution.

For each simulated sample, we obtain the ML estimates by solving (3.16) under the considerations mentioned in Section 3.8. Tables 7 and 8 report the average estimate (AE), the empirical standard deviation (SD), the asymptotic standard error (SE) and the root mean square error (RMSE) obtained in each scenario and sample size considered. In the tables, it is observed that the AE's tend to approach the real values of the parameters as the size of the sample increases. Furthermore, as expected in the standard asymptotic theory, it is observed that the SD's, SE's and RMSE's are close and decrease towards 0 as the sample size increases.

Table 7 – Averages (AE), standard deviations (SD), root of the simulated mean square errors (RMSE), and averages of asymptotic standard errors (SE) for the estimates of  $\sigma$  and  $\alpha$  of the LE distribution.

Scenario	$\hat{\sigma}$				$\hat{\alpha}$			
	AE	SD	SE	RMSE	AE	SD	SE	RMSE
<i>n</i> = 50								
A	0.981	0.179	0.196	0.143	1.620	0.388	0.449	0.325
B	1.024	0.159	0.173	0.160	1.969	0.372	0.432	0.373
C	1.041	0.129	0.151	0.136	2.373	0.304	0.432	0.329
D	1.976	0.340	0.399	0.340	1.601	0.393	0.451	0.406
E	2.938	0.546	0.590	0.549	1.627	0.405	0.450	0.424
F	3.962	0.748	0.817	0.748	1.572	0.387	0.451	0.393
G	0.494	0.088	0.100	0.088	1.600	0.391	0.451	0.404
H	1.526	0.240	0.254	0.241	1.980	0.375	0.430	0.375
I	2.594	0.345	0.378	0.358	2.368	0.308	0.429	0.335
<i>n</i> = 100								
A	0.994	0.137	0.145	0.137	1.540	0.310	0.325	0.313
B	1.010	0.109	0.115	0.109	2.002	0.285	0.297	0.285
C	1.022	0.091	0.097	0.093	2.426	0.220	0.271	0.232
D	1.991	0.281	0.288	0.281	1.555	0.319	0.323	0.323
E	2.986	0.407	0.432	0.407	1.547	0.310	0.324	0.313
F	3.992	0.558	0.579	0.558	1.551	0.310	0.324	0.315
G	0.498	0.067	0.072	0.067	1.555	0.302	0.323	0.307
H	1.524	0.172	0.174	0.174	1.993	0.285	0.299	0.285
I	2.552	0.222	0.243	0.228	2.431	0.222	0.270	0.232
<i>n</i> = 200								
A	0.998	0.101	0.101	0.101	1.528	0.231	0.231	0.232
B	1.001	0.083	0.080	0.083	1.998	0.223	0.210	0.223
C	1.011	0.062	0.065	0.063	2.457	0.165	0.177	0.171
D	2.004	0.208	0.205	0.208	1.524	0.239	0.231	0.240
E	3.016	0.303	0.310	0.303	1.508	0.221	0.231	0.221
F	4.001	0.417	0.413	0.417	1.513	0.238	0.231	0.238
G	0.500	0.051	0.051	0.051	1.519	0.232	0.232	0.232
H	1.500	0.114	0.118	0.114	2.005	0.202	0.209	0.203
I	2.519	0.155	0.162	0.155	2.468	0.160	0.176	0.163

Table 8 – Averages (AE), standard deviations (SD), root of the simulated mean square errors (RMSE), and averages of asymptotic standard errors (SE) for the estimates of  $\sigma$  and  $\alpha$  of the LR distribution.

Scenario	$\hat{\sigma}$				$\hat{\alpha}$			
	AE	SD	SE	RMSE	AE	SD	SE	RMSE
$n = 50$								
A	0.991	0.124	0.101	0.125	1.626	0.407	0.450	0.426
B	1.002	0.078	0.082	0.078	2.008	0.384	0.421	0.384
C	1.015	0.064	0.071	0.065	2.380	0.291	0.408	0.315
D	1.970	0.176	0.196	0.178	1.636	0.401	0.451	0.423
E	2.969	0.267	0.301	0.269	1.604	0.393	0.455	0.407
F	3.956	0.342	0.397	0.345	1.624	0.401	0.454	0.419
G	0.496	0.049	0.050	0.049	1.613	0.391	0.454	0.407
H	1.504	0.121	0.125	0.121	2.029	0.390	0.428	0.390
I	2.537	0.152	0.178	0.157	2.382	0.276	0.406	0.300
$n = 100$								
A	0.997	0.069	0.072	0.069	1.541	0.299	0.325	0.302
B	1.003	0.055	0.056	0.055	2.010	0.291	0.295	0.291
C	1.006	0.042	0.047	0.043	2.445	0.204	0.260	0.211
D	1.994	0.137	0.144	0.137	1.554	0.309	0.309	0.314
E	2.976	0.193	0.211	0.194	1.565	0.295	0.324	0.302
F	3.989	0.263	0.286	0.263	1.556	0.295	0.325	0.300
G	0.496	0.034	0.035	0.034	1.568	0.317	0.323	0.325
H	1.498	0.084	0.084	0.084	2.014	0.290	0.295	0.290
I	2.519	0.107	0.118	0.109	2.433	0.200	0.261	0.211
$n = 200$								
A	1.001	0.051	0.051	0.051	1.518	0.232	0.232	0.233
B	1.001	0.039	0.039	0.039	2.012	0.214	0.209	0.214
C	1.002	0.030	0.032	0.030	2.467	0.148	0.174	0.152
D	1.999	0.102	0.102	0.102	1.521	0.231	0.232	0.232
E	2.992	0.147	0.152	0.147	1.523	0.226	0.232	0.228
F	3.996	0.204	0.204	0.204	1.518	0.231	0.232	0.232
G	0.499	0.024	0.024	0.024	1.520	0.221	0.231	0.221
H	1.499	0.060	0.059	0.060	2.007	0.213	0.210	0.213
I	2.511	0.079	0.080	0.080	2.472	0.154	0.172	0.156

### 3.10 Data Analysis

In this section, we fit two data sets in order to illustrate the usefulness of the LE and LR distributions in real-world settings. Here, we provide evidence that the LE and LR distributions may perform better than other commonly used distributions. In each case, the parameters are estimated via the ML method using the `optim` function in the R language. We consider the information criteria, AIC and BIC, to evaluate the comparative performance of the fitted distributions.

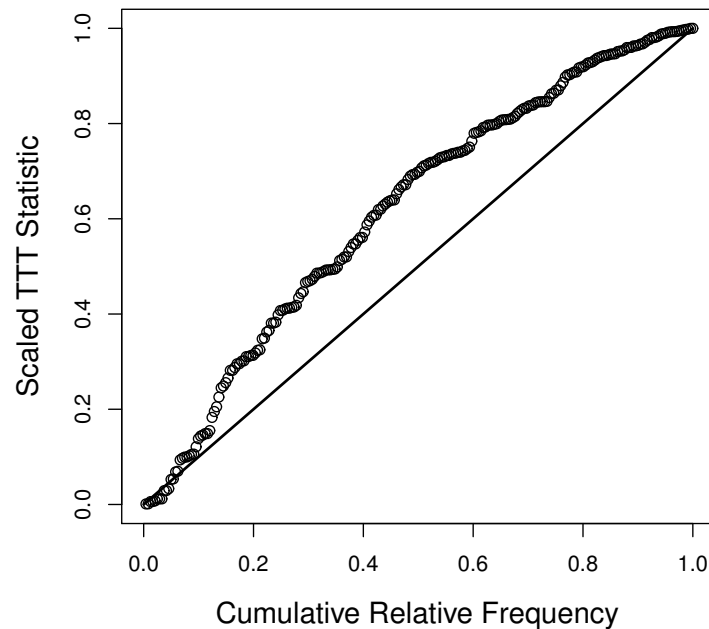
The data considered are briefly described below:

**Violent Crime Rate Data** This data set consists of 1,173 observations on the violent crime rate (incidents per 100,000 residents) recorded in 51 US states. This data set can be found in the Guns database of the AER package (KLEIBER; ZEILEIS, 2008) in the R language. For these data, we test hypothesis  $H_0$ : the data have exactly one mode versus the alternative hypothesis  $H_1$ : the data have at least two modes. For this, we consider the excess mass test (AMEIJEIRAS-ALONSO; CRUJEIRAS; RODRÍGUEZ-CASAL, 2019) using the `modetest` function (AMEIJEIRAS-ALONSO; CRUJEIRAS; RODRÍGUEZ-CASAL, 2018) in R language. Here, the observed statistic and the corresponding  $p$ -value are 0.032 and 0.154, respectively. Thus, we observe that under a level of significance equal to 0.05, the hypothesis that the data have exactly one mode is not rejected. The foregoing, added to the fact that the data exhibit a positive skewness, and that the observations assume strictly positive values, we fit these data with the LR distribution and compare its performance with that of other asymmetric unimodal positive support distributions such as the W, G, R and generalized Rayleigh (GR) (SURLES; PADGETT, 2001) distributions.

**Monoclonal Gammopathy Data** This data set comprises survival times (days) from diagnosis to the last follow-up of 241 subjects diagnosed with apparently benign monoclonal gammopathy at Mayo Clinic (US). Of the 241 subjects, 16 survived until the end of the follow-up and three had monoclonal gammopathy of undetermined significance (MGUS) detected on the day of death. This data set was previously analyzed in Kyle (1993) and is currently available under the name `mgus` in the survival package (THERNEAU, 2015) of the R language. For these data, we elaborated the TTT-plot (Figure 13) to identify which distribution can adequately model survival times. In Figure 13, we observe that it is advisable to consider distributions with increasing  $hrf$ 's for modeling these survival times. Consequently, we fit this data set with the LE distribution and compare its performance with that of the baseline E distribution and with that of other distributions with increasing  $hrf$ 's such as the G, W and generalized exponential (GE) (GUPTA; KUNDU, 1999) distributions.

Table 16 reports ML estimates and AIC and BIC values for each distribution fitted to violent crime rate data and monoclonal gammopathy data. Here, we observe that the LR and LE distributions present the lowest AIC and BIC values, which suggests that these distributions

Figure 13 – TTT-plot for survival times of individuals diagnosed with apparently benign monoclonal gammopathy.



should be chosen over the others to model the corresponding data sets.

Figure 14 (Left panel) present the histogram of the inflation rate data, where it can be seen that the LR density values are close to the empirical frequency values. In the right panel of the same figure, the Kaplan-Meier survival curve and the survival curve fitted with the LE distribution are presented for the subjects diagnosed with monoclonal gammopathy, where it is observed that the survival curves are close.

### 3.11 Final Comments

In this chapter, we proposed a new distribution generator called the Lambert- $F$  generator. The new generator allows us to add one shape parameter to an arbitrary baseline distribution, enabling a variety of shapes for the pdf and the hrf of the resulting distribution.

The Lambert- $F$  generator arises from the qf of an arbitrary baseline distribution, where the input of the qf is a Lambert-uniform rv. This construction leads to Lambert- $F$  distributions having attractive properties. Some of these are the following: First, the pdf, hrf and sf of the Lambert- $F$  distributions correspond to modifications in a multiplicative fashion of the corresponding pdf, hrf and sf of the baseline distributions; Second, the Lambert- $F$  qf can be expressed in closed form in terms of the principal branch of the Lambert  $W$  function and in terms of the qf of the baseline distribution. In this way, pseudo-random numbers can be easily generated using the corresponding qf; Third, the hrf of a Lambert- $F$  distribution can be understood as a

Figure 14 – Left panel: Histogram of the violent crime rate data with the fitted pdf curves. Right panel: Kaplan-Meier survival curve (black line) and the fitted LE survival curve (red line) for the survival times of monoclonal gammopathy data.

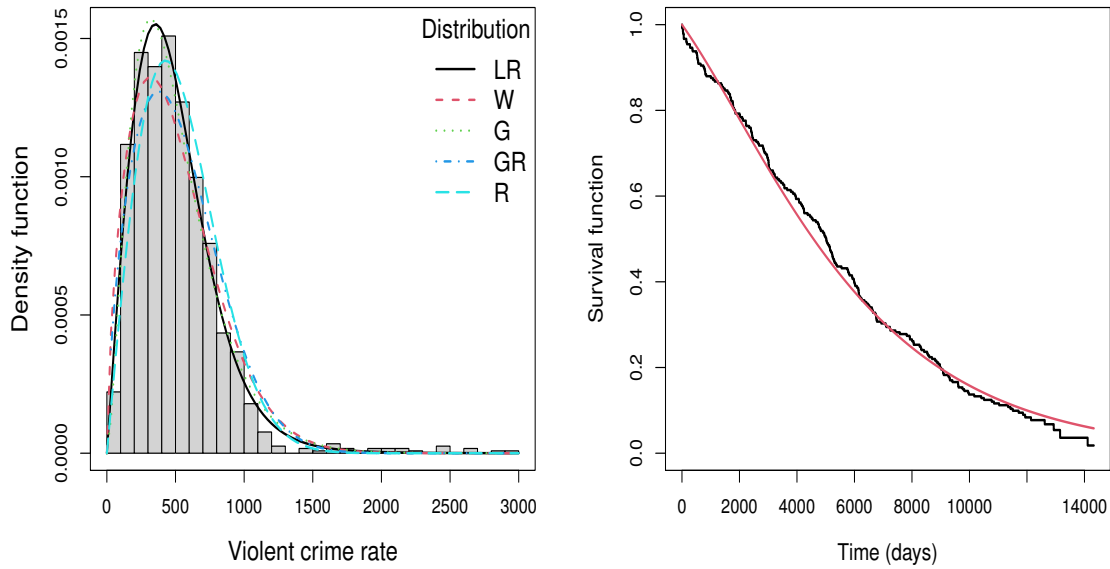


Table 9 – The ML estimates and their standard errors (in parentheses) for each distribution fitted to the different data sets.

Violent crime rate data					
Parameter	LR	W	G	GB	R
$\sigma$	828.803 (46.143)	0.799 (0.024)	183.606 (7.865)	565.315 (10.646)	427.042 (0.6234)
$\alpha$	0.017 (0.008)	1.579 (0.118)	2.739 (0.106)	1.640 (0.033)	-
AIC	16414.4	16521.9	16432.0	16595.5	16626.8
BIC	16424.5	16532.1	16442.1	16605.7	16631.9
Monoclonal gammopathy data					
Parameter	LE	W	G	GE	E
$\sigma$	4173.763 (309.337)	6009.851 (345.316)	4909.396 (534.902)	5270.609 (441.299)	5810.491 (389.431)
$\alpha$	1.830 (0.181)	1.186 (0.067)	1.174 (0.099)	1.159 (0.100)	-
AIC	4338.5	4345.8	4350.8	4351.5	4352.3
BIC	4345.5	4352.7	4357.8	4358.4	4355.8



modification in early times of the hrf of the baseline distribution; As for place, when the baseline distribution is symmetric, the shape parameter  $\alpha$  acts as a skewness parameter, enabling skewed shapes for the pdf; Fifth, the Lambert- $F$  rv's can be ordered according to likelihood ratio order, which implies ordering in relation to the usual stochastic order and the hazard rate order.

We derive two special cases of the Lambert- $F$  generator, the LE and LR distributions. These distributions are obtained considering the exponential and Rayleigh distributions as the baseline distribution. We discussed the maximum likelihood estimators for the parameters of the Lambert- $F$  distributions. We provided guidance on numerical procedures that might be used. Additionally, we carried out a simulation study to assess the behavior of the estimators for the parameters of the LE and LR distribution. We observe that the maximum likelihood method provides acceptable estimates for the parameters of these distributions. Finally, we consider two applications to real data sets, thus providing evidence that the LE and LR distributions can present a better fit than other two-parameter distributions such as Weibull, gamma, generalized exponential, and generalized Rayleigh.



---

## TWO UNBOUNDED-SUPPORT LAMBERT-F DISTRIBUTIONS

---

In this section, we propose two unbounded support distributions that are capable of modeling skewness in the presence of bimodality (the first) and skewness in the presence of high levels of kurtosis (the second). Both distributions are obtained from the Lambert- $F$  generator considering symmetric baseline distributions, the generalized bimodal (GB) and slash (SL) baseline distributions, respectively. The shapes of the pdf and the behavior of Fisher's skewness and kurtosis coefficients are studied. An analysis on the occurrence of alias distributions for both distributions is considered. Parameter estimation is performed via the maximum likelihood method. Simulation studies are developed to evaluate the behavior of the estimators. Finally, two applications are presented in order to illustrate the utility of the proposed distributions in data modeling in real settings.

### 4.1 Motivation

Unbounded support distributions play an important role in data analysis. When the data exhibit unimodality and symmetry, distributions such as the normal, t-student, and Cauchy are highly valued. On the other hand, when the data present another property such as skewness, bimodality or excessive levels of kurtosis, for example, these distributions do not perform properly. Consequently, the use of alternative distributions capable of capturing the properties exhibited by the data is required.

The GB distribution is a bimodal symmetric distribution originally proposed by [Rao \(1987\)](#) and later studied as a special case of the bimodal distribution proposed by [Sarma, Rao and Rao \(1990\)](#). The GB distribution can also be derived as a special case of the class of bimodal distributions proposed by [Hassan and El-Bassiouni \(2016\)](#), which is defined by the cdf  $F(x) = \Phi(x) - \delta(x)\phi(x)$ , where  $\delta(x)$  is a linear function of  $x$ . If  $\delta(x) = x/(1 + \gamma)$ , the cdf of

this class reduces to the cdf of the GB distribution.

On the other hand, the SL distribution is a symmetric bell-shaped distribution that has heavier tails than the normal distribution. Therefore, its use is highly valued when modeling data that exhibit symmetry, unimodality, and high levels of kurtosis. Properties of this distribution can be consulted in [Johnson, Kotz and Balakrishnan \(1994\)](#) and [Gómez, Quintana and Torres \(2007\)](#). In [Appendix A](#) the analytical expressions of the pdf and the cdf for the GB and SL distributions are presented.

The symmetry feature of the GB and SL distributions can be considered a desired feature in some scenarios, but a limitation in others. The distributions proposed in this chapter have the same shape characteristics as the GB and SL distributions, but they can also exhibit asymmetric shapes while capturing bimodality and high levels of kurtosis.

## 4.2 Lambert-Generalized Bimodal and Lambert-Slash Distributions

In this section, we define two unbounded support Lambert- $F$  distributions and discuss the shapes of the corresponding pdf's.

**Definition 4.** The rv  $X$  follows the Lambert-generalized bimodal distribution with location parameter  $\mu \in \mathbb{R}$ , scale parameter  $\sigma > 0$ , and shapes parameters  $\gamma \in (0, 2)$  and  $\alpha \in (0, e)$ , denoted as  $X \sim \text{LGB}(\mu, \sigma, \gamma, \alpha)$ , if its pdf is given by

$$f_X(x; \mu, \sigma, \gamma, \alpha) = \frac{\gamma + z^2}{\sigma(\gamma + 1)} \phi(z) \alpha^{\Phi(z) - \frac{z}{\gamma+1} \phi(z)} \times \left\{ 1 - \log(\alpha) \left[ 1 - \Phi(z) + \frac{z}{\gamma+1} \phi(z) \right] \right\},$$

where  $x \in \mathbb{R}$ ,  $z = (x - \mu)/\sigma$  and  $\phi(\cdot)$  and  $\Phi(\cdot)$  are the pdf and the cdf of the standard normal distribution.

Figure 15 shows some pdf curves for the LGB distribution considering different values of  $\mu$ ,  $\sigma$ ,  $\gamma$  and  $\alpha$ . Here, it can be seen that the LGB pdf is bimodal symmetric when  $\alpha = 1$ . Note that the parameter  $\alpha$  has an effect on the shape of the LGB pdf allowing unimodal or bimodal asymmetric shapes.

**Definition 5.** The rv  $X$  follows the Lambert-slash distribution with location parameter  $\mu \in \mathbb{R}$ , scale parameter  $\sigma > 0$ , and shape parameters  $\kappa > 0$  and  $\alpha \in (0, e)$ , denoted as  $X \sim \text{LSL}(\mu, \sigma, \gamma, \alpha)$ , if its pdf is given by

$$f_X(x; \mu, \sigma, \kappa, \alpha) = \frac{1}{\sigma} f(z; \kappa) \alpha^{\Phi(z) - \frac{z}{\kappa} f(z; \kappa)} \times \left\{ 1 - \log(\alpha) \left[ 1 - \Phi(z) + \frac{z}{\kappa} f(z; \kappa) \right] \right\},$$

Figure 15 – Pdf curves for the LGB distribution for  $\mu = 5, \sigma = 2, \text{ and } \alpha = 1$  in the top left panel;  $\mu = -5, \sigma = 4, \text{ and } \alpha = 0.5$  in the top right panel;  $\mu = 5, \sigma = 4, \text{ and } \alpha = 2$  in the bottom left panel; and  $\mu = 10, \sigma = 5, \text{ and } \gamma = 1.5$  in the bottom right panel.

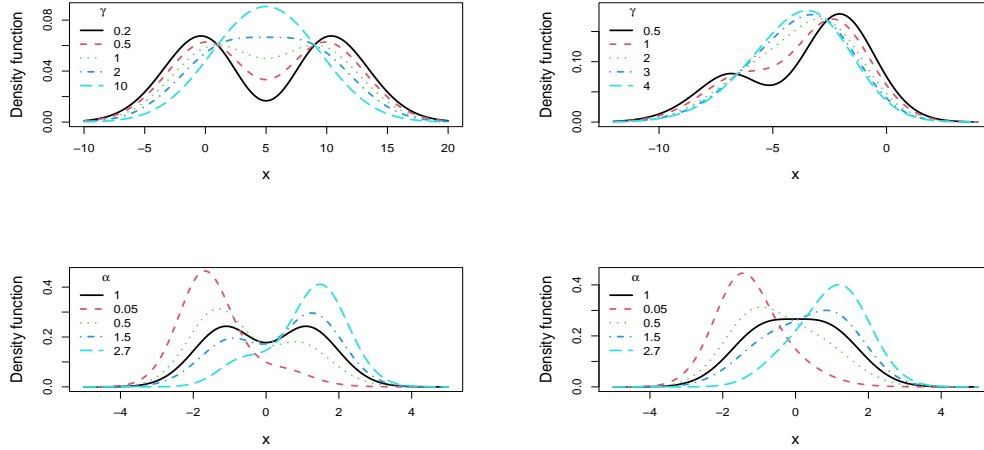
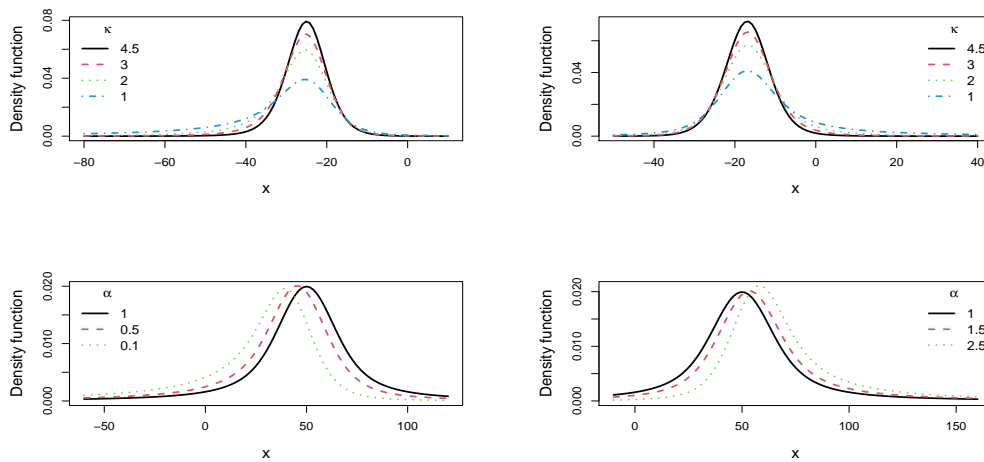


Figure 16 – Pdf curves for the LSL distribution for  $\mu = -20, \sigma = 5 \text{ and } \alpha = 0.1$  in the top left panel,  $\mu = 20, \sigma = 5 \text{ and } \alpha = 2$  in the top right panel and  $\mu = 50, \sigma = 10 \text{ and } q = 1$  in the bottom panels.



where  $x \in \mathbb{R}, z = (x - \mu)/\sigma$  and  $\Phi(\cdot)$  is the standard normal cdf and  $f(z; \kappa) = \kappa \int_0^1 u^\kappa \phi(uz) du$  is the SL baseline cdf.

Figure 16 presents some curves of the LSL pdf for different values of the parameters. Here, it can be seen that the LSL pdf has symmetric or asymmetric shape depending on the value assumed by  $\alpha$  and that the weight in the tails varies depending on the value of  $\kappa$ .

**Corollary 9.** If  $(x^*, f_X(x^*; \gamma, \alpha))$  and  $(x^{**}, f_X(x^{**}; \gamma, \alpha))$  are respectively a critical point and an inflection point of the LGB pdf, then  $x^*$  is a root of the equation

$$\sigma \log(\alpha)(\gamma + z^2)(1 + z_3)z_1 + z(2 - \gamma - z^2)z_3 = 0, \tag{4.1}$$

and  $x^{**}$  a root of the equation

$$\begin{aligned} \sigma^2 \log^2(\alpha)(\gamma + z^2)(2 + z_3)z_1^2 + 3\sigma \log(\alpha)z(2 - \gamma - z^2)(1 + z_3)z_1 \\ + [2 + z^4 - (5 - \gamma)z^2 - \gamma]z_4 = 0, \end{aligned} \quad (4.2)$$

where  $z_1 = \{(\gamma + z^2)/[\sigma(1 + \gamma)]\}\phi(z)$ ,  $z_2 = \Phi(z) - [z/(1 + \gamma)]\phi(z)$ ,  $z_3 = 1 - \log(\alpha)(1 - z_2)$  and  $z_4 = 1 - \log(\alpha)(1 + z_2)$ .

In the case  $\alpha = 1$ , Equations (4.1) and (4.2) lead to establish that the LGB pdf is bimodal with modes given by  $\mu \pm \sigma\sqrt{2 - \gamma}$ , antimode given by  $\mu$ , and abscissa of inflection points given by  $\mu \pm \sigma\sqrt{w_j}$ , with  $j = 1, 2$ , where  $w_1 = [(5 - \gamma) + \sqrt{(5 - \gamma)^2 - 4(2 - \gamma)}]/2$  and  $w_2 = [(5 - \gamma) - \sqrt{(5 - \gamma)^2 - 4(2 - \gamma)}]/2$ .

In the case  $\alpha \neq 1$ , it is not possible to obtain closed expressions for the modes, antimode and abscissa of inflection points. Therefore, these values must be obtained by solving Equations (4.1) and (4.2) by numerical procedures.

Figure 17 presents some profiles of Equations (4.1) and (4.2) and the corresponding LGB pdf curves for different values of the parameters. In Figure 18, the regions of unimodality and bimodality established in the plane defined by the ranges of  $\gamma$  and  $\alpha$  are presented. This figure was drawn by solving Equation (4.1) using the `uniroot.all` function (SOETAERT, 2009) in the R language. From these figures, we observe that the LGB pdf is unimodal or bimodal depending on the values of  $\gamma$  and  $\alpha$ . In the unimodal case, we observe that the LGB pdf can have two or four inflection points.

**Corollary 10.** If  $(x^*, f_X(x^*; \gamma, \alpha))$  and  $(x^{**}, f_X(x^{**}; \gamma, \alpha))$  are respectively a critical point and an inflection point of the LGB pdf, then  $x^*$  is a root of the equation

$$\log(\alpha)(z_1 + 1) - \frac{\sigma\kappa}{\kappa + 2} \frac{zf(z; \kappa + 2)z_1}{f^2(z; \kappa)} = 0, \quad (4.3)$$

and  $x^{**}$  a root of the equation

$$\begin{aligned} \log(\alpha)f(z; \kappa) \left[ \log(\alpha)(z_1 + 1)f^2(z; \kappa) - \frac{\kappa}{\kappa + 2}zf(z; \kappa + 2)z_1 \right] \\ + \log(\alpha)f(z; \kappa) \left[ \frac{\kappa}{\kappa + 2}z(z_1 + 1)f(z; \kappa + 2) + \log(\alpha)f^2(z; \kappa) \right] \\ - \frac{\kappa}{\kappa + 2}z_1 \left[ f(z; \kappa + 2) + \frac{\kappa}{\kappa + 4}z^2z_1f(z; \kappa + 4) \right] \\ - \frac{\kappa \log(\alpha)}{\kappa + 2}z(z_1 + 2)f(z; \kappa)f(z; \kappa + 2) = 0, \end{aligned} \quad (4.4)$$

where  $z_1 = 1 - \log(\alpha)[1 - \Phi(z) + (z/\kappa)f(z; \kappa)]$ .

Figure 19 presents some profiles of Equations (4.3) and (4.4) and the corresponding LSL pdf curves for different values of the parameters. From this figure, we observe that the LSL pdf is unimodal with two inflection points.

Figure 17 – Critical and inflection points and pdf curve for the Lambert generalized bimodal (LGB) distribution: LGB(5,2,1.5,0.005) (solid black curves), LGB(5,2,1.5,0.5) (dashed red curves), LGB(5,2,0.5,0.5) (solid green curves), and LGB(5,2,0.5,1) (dashed blue curves).

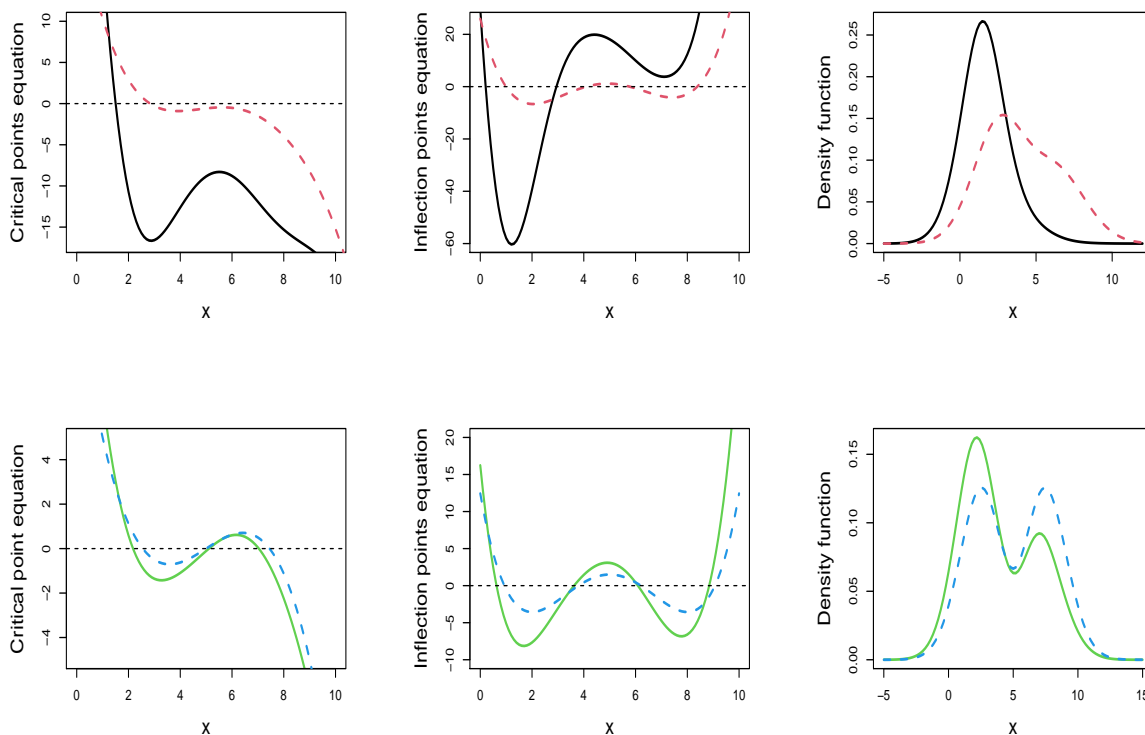


Figure 18 – Regions of unimodality (white region) and bimodality (gray region) for the Lambert generalized bimodal (LGB) distribution.

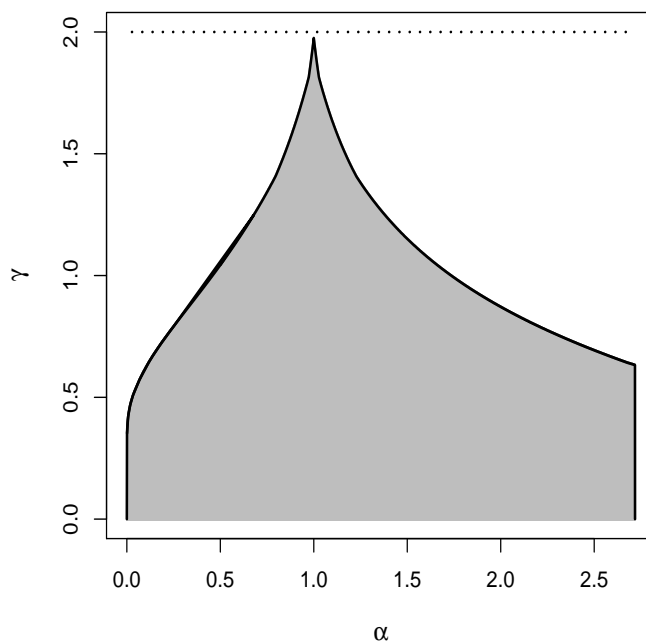
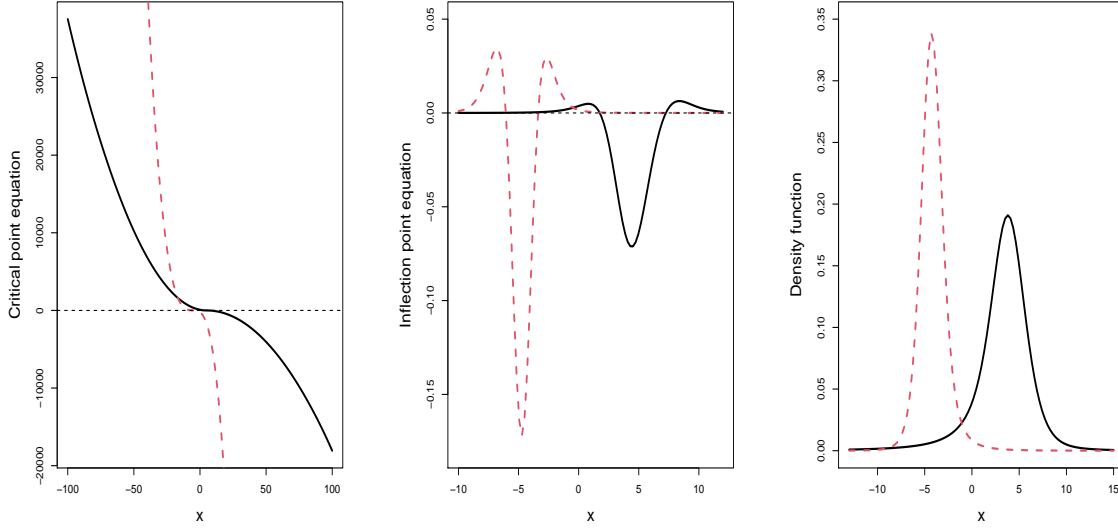


Figure 19 – Critical and inflection points and pdf curve for the Lambert-slash (LSL) distribution: LSL(5,1.5,2,0.5) (solid black curves) and LSL(-5,1,3,1.5) (dashed red curves).



On the other hand, in order to provide more details on the shape of the pdf of the LGB and LSL distributions, we analyze the behavior of the Fisher's skewness and kurtosis coefficients. For this, from the result given in Proposition 5, we first derive the raw moments for these distributions.

**Corollary 11.** Let  $X_1 \sim \text{LGB}(\mu, \sigma, \gamma, \alpha)$  and  $X_2 \sim \text{LSL}(\mu, \sigma, \kappa, \alpha)$ . Then, for  $r = 1, 2, \dots$ , the  $r$ -th raw moment of  $X_j$  ( $j = 1, 2$ ) is given by

$$\mathbb{E}(X_j^r) = \sum_{k=0}^r \binom{r}{k} \mu^{r-k} \sigma^k a_k(\theta_j, \alpha),$$

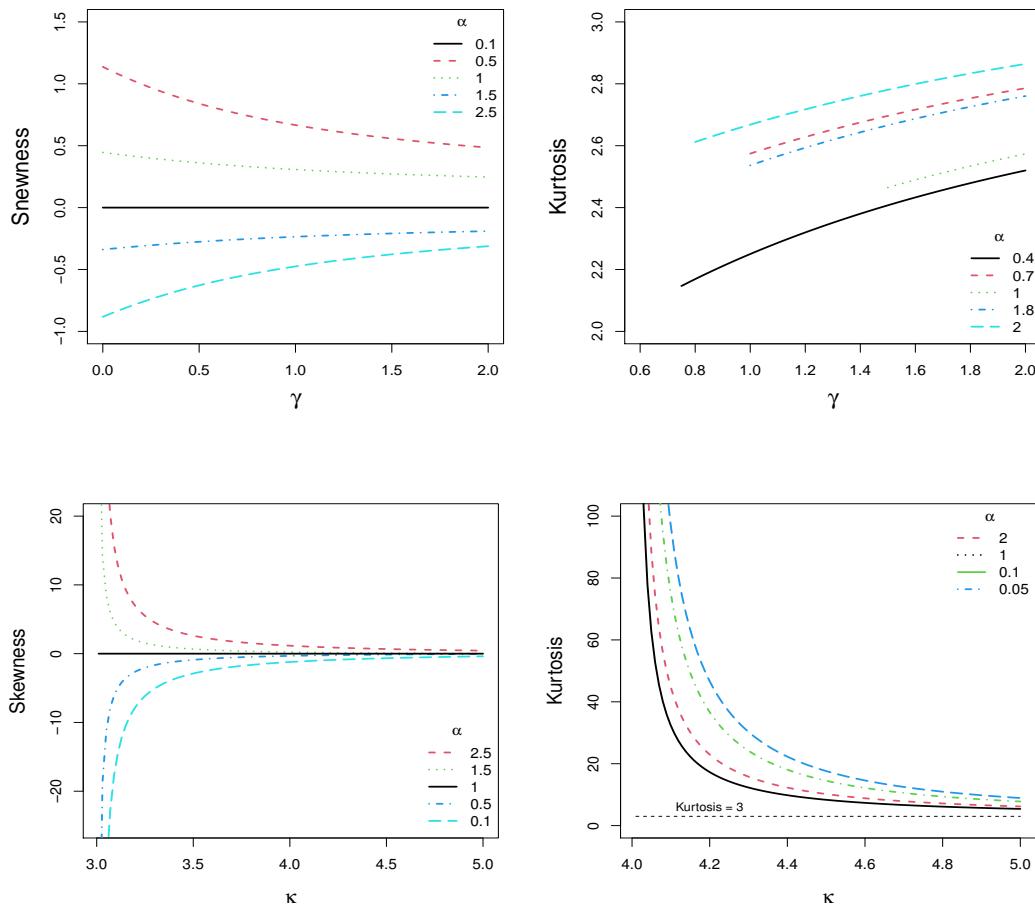
where  $a_r(\theta_j, \alpha) = \int_0^1 [Q_j(u; \theta_j)]^r \alpha^u [1 - \log(\alpha)(1 - u)] du$ , such that  $Q_1(u; \theta_1)$ , with  $\theta_1 = \gamma$ , is the qf of the GB distribution and  $Q_2(u; \theta_2)$ , with  $\theta_2 = \kappa$ , is the qf of the SL distribution.

**Corollary 12.** Let  $X_1 \sim \text{LGB}(\mu, \sigma, \gamma, \alpha)$  and  $X_2 \sim \text{LSL}(\mu, \sigma, \kappa, \alpha)$ . Then, the skewness ( $\beta_1(\theta_j, \alpha)$ ) and kurtosis ( $\beta_2(\theta_j, \alpha)$ ) coefficients of  $X_j$  ( $j = 1, 2$ ) are given by

$$\begin{aligned} \beta_1(\theta_j, \alpha) &= \frac{\mathbb{E}\{[X_j - \mathbb{E}(X_j)]^3\}}{[\text{Var}(X_j)]^{3/2}} \\ &= \frac{a_3(\theta_j, \alpha) - 3a_1(\theta_j, \alpha)a_2(\theta_j, \alpha) + 2a_1^3(\theta_j, \alpha)}{[a_2(\theta_j, \alpha) - a_1^2(\theta_j, \alpha)]^{3/2}} \text{ and} \\ \beta_2(\theta_j, \alpha) &= \frac{\mathbb{E}\{[X_j - \mathbb{E}(X_j)]^4\}}{[\text{Var}(X_j)]^2} \\ &= \frac{a_4(\theta_j, \alpha) - 4a_1(\theta_j, \alpha)a_3(\theta_j, \alpha) + 6a_1^2(\theta_j, \alpha)a_2(\theta_j, \alpha) - 3a_1^4(\theta_j, \alpha)}{[a_2(\theta_j, \alpha) - a_1^2(\theta_j, \alpha)]^2}. \end{aligned}$$



Figure 20 – Plots of the skewness and kurtosis coefficients of the LGB (top panels) and LSL (bottom panels) distributions.



Note that the  $r$ -th raw moment of the LGB and LSL distributions must be computed using numerical integration. The function `integrate` in the R language is a good choice for this task.

Figure 20 (top left panel) presents some curves for the skewness coefficient of the LGB distribution considering different values of  $\gamma$  and  $\alpha$ . The top right panel of the same figure presents some curves for the kurtosis coefficient of the LGB distribution when it is unimodal. Figure 20 (bottom panels) presents some curves for the skewness and kurtosis coefficients of the LSL distribution. In the figures, we observe the following: 1) The LGB and LSL pdf's are symmetric when  $\alpha = 1$ , regardless of the value assumed by the parameters  $\gamma$  and  $\kappa$ ; 2) The LGB and LSL pdf's can be skewed (positively or negatively) depending on the value assumed by  $\alpha$ . If  $\alpha \in (0, 1)$  or  $\alpha \in (1, e)$ , then the pdf's are skewed and the skewness is also controlled by the parameters  $\gamma$  and  $\kappa$ . However, the effects of  $\gamma$  and  $\kappa$  on the skewness of the distributions are especially important if these parameters assume lower values within their ranges. The lower the values assumed for  $\gamma$  and  $\kappa$ , the more skewed the LGB and LS distributions will be; 3) If the LGB pdf is unimodal, then it is asymmetric; 4) For both distributions, the excess kurtosis

$\beta_2(\theta_j, \alpha) - 3$ ,  $j = 1, 2$ , is less than 0, that is, the LGB (when is unimodal) and LSL distributions are platykurtic distributions; 5) Finally, we observe that the LSL distribution is able to model kurtosis levels greater than the kurtosis levels of the SL baseline distribution (case  $\alpha = 1$ ).

A more detailed discussion on the structural properties of the LGB distribution can be found in [Iriarte, de Castro and Gómez \(2021b\)](#).

### 4.3 Alias Distribution

From a distance perspective, the concept “alias distribution” is used to indicate that members belonging to a particular family of distributions exhibit virtually identical behavior. This means that for a given base distribution parameterization, it can be found sets of parameter values for which two distributions are practically identical from a numerical standpoint, that is, no matter what the sample size is, it is possible to find pairs of distributions that are so close in terms of having nearly the same probability measure that a typical fitting routine would not be able to distinguish between the two sets of parameters, not even for massively large sample sizes.

The analysis considered in this section is inspired by the cautionary note provided by [Hutson and Vexler \(2018\)](#), a note referring to the occurrence of aliases in the beta-normal and beta-logistic families. [Hutson and Vexler \(2018\)](#) state that the occurrence of alias distributions is highly relevant in the context of testing subclasses of distributions within the family of beta distributions, for example, normality tests as a submodel of the beta-distribution. normal. Furthermore, they illustrate that the large-sample variance approximations of the maximum likelihood estimators are very different between a given distribution and its respective alias, even when the actual density functions are nearly identical.

In this section, we discuss the occurrence of alias distributions for the LGB and LSL distributions. The justifications for carrying out an analysis of this type are the following: First, taking into account that these families have four parameters, with two shape parameters, the occurrence of aliases in the LGB and LSL families is not unreasonable. Second, we observe that the skewness coefficients of these distributions behave as a non-monotonic function of the shape parameter  $\alpha$  when the shape parameter inherited from the baseline distribution ( $\gamma$  and  $\kappa$ , respectively) assumes certain values. Consequently, there may be two or more members of the LGB and LSL families that have the same skewness value.

The latter is a serious problem for any unimodal Lambert- $F$  distribution that is generated from a symmetric baseline distribution, since here the level of skewness is determined by the value of  $\alpha$ . Thus, taking into account that  $\alpha$  has a slight effect on kurtosis, it is enough to make an appropriate choice of locations and scales in a location-scale Lambert- $F$  family to identify two members that have a virtually identical behavior.

In what follows, like [Hutson and Vexler \(2018\)](#), we use the Kullback-Leibler divergence

to identify aliases for certain members of the LGB and LSL families. The Kullback-Leibler divergence measures the degree of divergence between the distributions of two rv's. For two LGB or LSL rv's consider the following proposition.

**Proposition 8.** Let  $X_{1j} \sim \text{LGB}(\mu_j, \sigma_j, \delta_{1j}, \alpha_j)$  and  $X_{2j} \sim \text{LSL}(\mu_j, \sigma_j, \delta_{2j}, \alpha_j)$ , with  $j = 1, 2$ , such that  $\delta_{11} = \gamma_1$ ,  $\delta_{12} = \gamma_2$ ,  $\delta_{21} = \kappa_1$  and  $\delta_{22} = \kappa_2$ . Then, the Kullback-Leibler divergence for  $X_{1j}$  and  $X_{2j}$  is given by

$$\mathcal{K}(X_{i1}, X_{i2}) = \log(\sigma_{i2}/\sigma_{i1}) + \int_{-\infty}^{\infty} w_i^* f_{X_i}(w; 0, 1, \delta_{i1}, \alpha_1) dw, \quad i = 1, 2, \quad (4.5)$$

where  $w_i^* = \log[f_{X_i}(w; 0, 1, \delta_{i1}, \alpha_1)/f_{X_i}(h(w); 0, 1, \delta_{i2}, \alpha_2)]$  and  $h(w) = (w\sigma_1 + \mu_1 - \mu_2)/\sigma_2$ , such that  $f_{X_1}(\cdot; \cdot, \cdot, \cdot, \cdot)$  is as in Definition 4 and  $f_{X_2}(\cdot; \cdot, \cdot, \cdot, \cdot)$  is as in Definition 5.

**Proof.** From the definition of the Kullback-Leibler divergence, we obtain  $\mathcal{K}(X_{i1}, X_{i2}) = \mathbb{E}[\log\{f_{X_i}(x; \mu_1, \sigma_1, \delta_{i1}, \alpha_1)/f_{X_i}(x; \mu_2, \sigma_2, \delta_{i2}, \alpha_2)\}]$ , where the expectation is taken with respect to  $X_{i1}$ . Then, the result is obtained by considering the change of variable  $w = z_1 = (x - \mu_1)/\sigma_1$ , once  $z_2 = (x - \mu_2)/\sigma_2$  can be written as  $z_2 = (z_1\sigma_1 + \mu_1 - \mu_2)/\sigma_2$ .  $\square$

We observed that the range of the parameters  $\gamma$  of the LGB distribution is the interval  $(0, 2)$ . However, the range of  $\alpha$  can be extended to  $(0, \infty)$  so that the normal distribution can be derived as the limiting case  $\gamma \rightarrow \infty$  (when  $\alpha = 1$ ) of the LGB distribution. Despite this interesting property, this is not recommended since the skewness coefficient of the LGB distribution is not a monotonic function of  $\alpha$ .

In the left panels of Figure 20, it can be verified that the skewness coefficient of the LGB and LSL distributions exhibit monotonous behavior when  $\gamma$  and  $\kappa$  assume values in the intervals  $(0, 2)$  and  $(0, 5)$ , respectively, which is not true when  $\gamma > 2$  and  $\kappa > 5$ . Figure 21 presents some skewness curves for the LGB and LSL distributions when  $\gamma > 2$  and  $\kappa > 5$ . Here, it can be clearly seen that the coefficients exhibit non-monotonous behavior.

Based on the minimization of Equation (4.5), we analyze the existence of alias distributions for the LGB and LSL families in different scenarios. In each scenario,  $X_{1j} \sim \text{LGB}(\theta_{1j})$  and  $X_{2j} \sim \text{LSL}(\theta_{2j})$ , where  $\theta_{ij} = (\mu_{ij}, \sigma_{ij}, \delta_{ij}, \alpha_{ij})^t$ ,  $i, j = 1, 2$ ,  $\delta_{11} = \gamma_{11}$ ,  $\delta_{12} = \gamma_{12}$ ,  $\delta_{21} = \kappa_{21}$ ,  $\delta_{22} = \kappa_{22}$ ,  $\theta_{i1}$  is known and  $\theta_{i2} \neq \theta_{i1}$  minimizes Equation (4.5).

Table 10 reports the values of  $\theta_{11}$  and  $\theta_{12}$  considered in each scenario together with the corresponding Kullback-Leibler divergence values. Figure 22 reports the pdf curves of  $X_{11} \sim \text{LGB}(\theta_{11})$  and  $X_{12} \sim \text{LGB}(\theta_{12})$ , where it is observed that in scenarios E and F the distribution of  $X_{12}$  is an alias of the distribution of  $X_{11}$ . On the other hand, it can be seen that in scenarios A to D ( $\gamma < 2$ ) the pdf of  $X_{12}$  has a smaller amount of inflection points than the pdf of  $X_{11}$ . Thus, the distribution of  $X_{12}$  is not an alias of the distribution of  $X_{11}$ .

Table 11 reports the values of  $\theta_{21}$  and  $\theta_{22}$  considered in each scenario together with the corresponding Kullback-Leibler divergence values. Figure 23 reports the pdf curves of

Figure 21 – Some skewness curves for the LGB distribution with  $\gamma > 2$  (left) and for the LSL distribution with  $\kappa > 5$  (right).

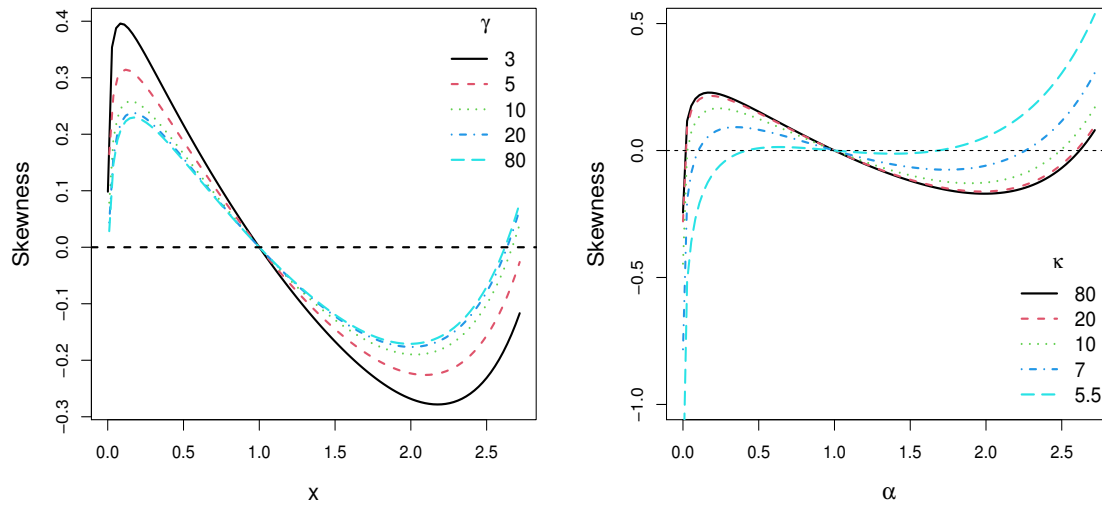


Figure 22 – Pdf curves of  $X_{11} \sim \text{LGB}(\theta_{11})$  (black solid line) and  $X_{12} \sim \text{LGB}(\theta_{12})$  (red dashed line) in scenarios A to F.

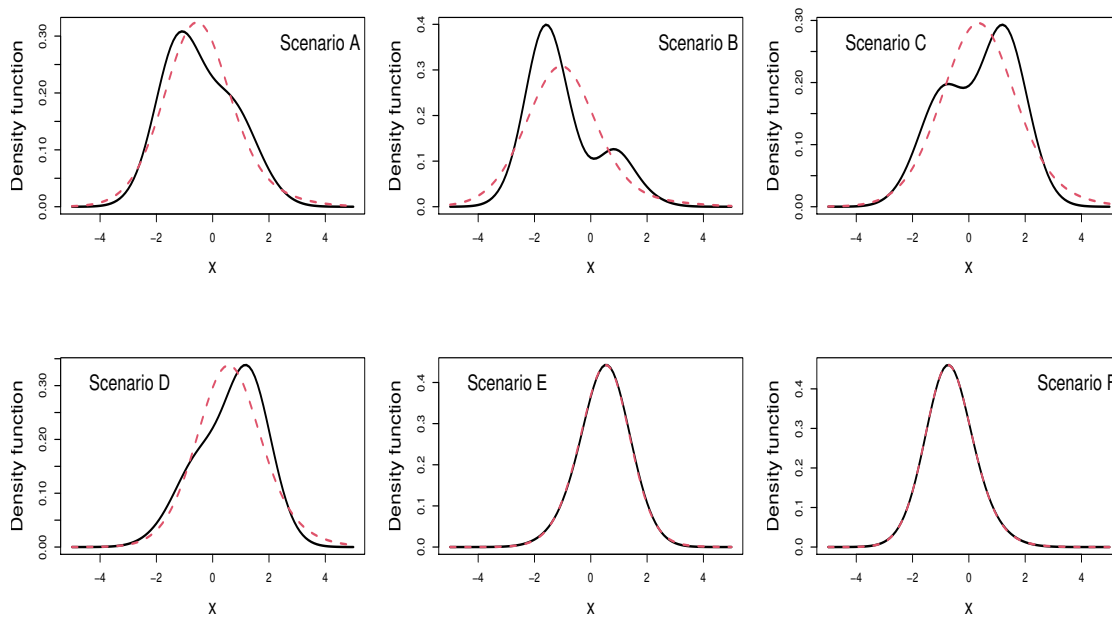
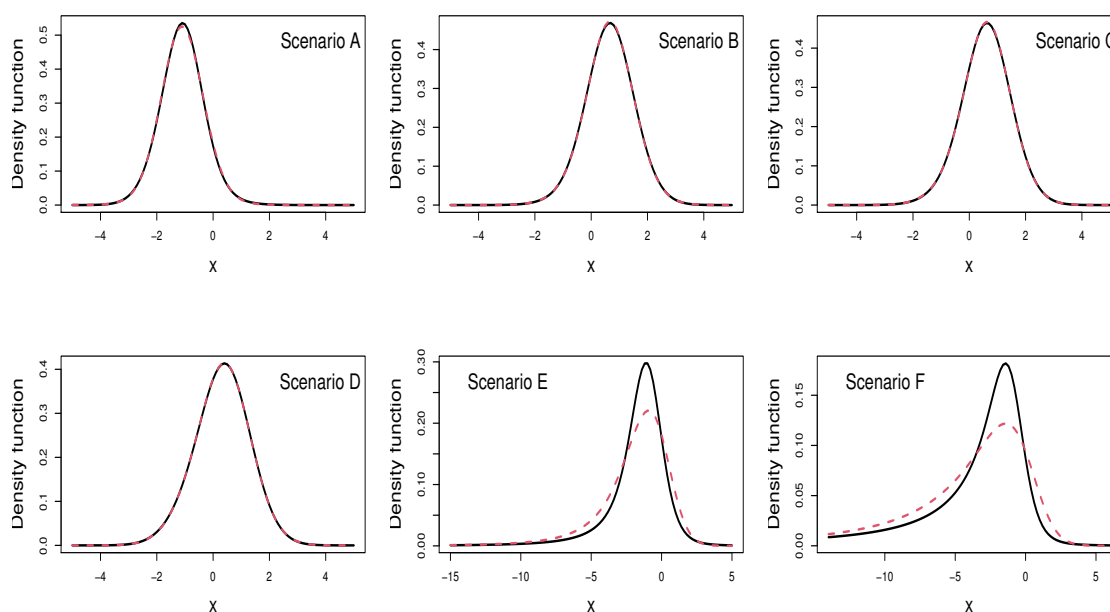


Table 10 – The Kullback-Leibler divergence values for  $X_{11} \sim \text{LGB}(\theta_{11})$  and  $X_{12} \sim \text{LGB}(\theta_{12})$  obtained in each scenario.

Scenario	$\theta_{11}$				$\theta_{12}$				$\mathcal{K}(X_{11}, X_{12})$
	$\mu_{11}$	$\sigma_{11}$	$\gamma_{11}$	$\alpha_{11}$	$\mu_{12}$	$\sigma_{12}$	$\gamma_{12}$	$\alpha_{12}$	
A	0	1	1.5	0.5	2.568	1.719	1.462	0.002	0.029
B	0	1	0.5	0.2	2.600	1.891	0.999	0.001	0.091
C	0	1	1	1.5	3.686	1.896	1.999	0.001	0.054
D	0	1	1.5	2	3.402	1.631	1.999	0.002	0.039
E	0	1	100	2	2.221	1.359	95.000	0.003	$1.077 \times 10^{-3}$
F	0	1	50	0.2	0.488	1.020	5.010	0.043	$2.210 \times 10^{-4}$

Figure 23 – Pdf curves of  $X_{21} \sim \text{LGB}(\theta_{21})$  (black solid line) and  $X_{22} \sim \text{LSL}(\theta_{22})$  (red dashed line) in scenarios A to F.

of  $X_{21} \sim \text{LSL}(\theta_{21})$  and  $X_{22} \sim \text{LSL}(\theta_{22})$ , where it is observed that in scenarios A to D the distribution of  $X_{22}$  is an alias of the distribution of  $X_{21}$ . On the other hand, it can be seen that in scenarios E and F ( $\kappa < 5$ ) the pdf of  $X_{22}$  has heavier tails than the pdf of  $X_{21}$ . Thus, the distribution of  $X_{22}$  is not an alias of the distribution of  $X_{21}$ .

Therefore, based on the previous results, it is observed that considering in the LGB and LSL distributions the intervals  $(0, 2)$  and  $(0, 5)$  as the ranges of  $\gamma$  and  $\kappa$ , respectively, avoids the occurrence of aliases for members of these families.

Table 11 – The Kullback-Leibler divergence values for  $X_{21} \sim \text{LSL}(\theta_{21})$  and  $X_{22} \sim \text{LSL}(\theta_{22})$  obtained in each scenario.

Scenario	$\theta_{21}$				$\theta_{22}$				$\mathcal{K}(X_{21}, X_{22})$
	$\mu_{21}$	$\sigma_{21}$	$\kappa_{21}$	$\alpha_{21}$	$\mu_{22}$	$\sigma_{22}$	$\kappa_{22}$	$\alpha_{22}$	
A	0	1	30	0.158	-1.056	0.657	6.444	0.960	$8.493 \times 10^{-4}$
B	0	1	40	1.555	0.044	0.927	12.894	1.528	$1.236 \times 10^{-4}$
C	0	1	80	1.513	0.063	0.917	14.105	1.469	$7.456 \times 10^{-5}$
D	0	1	50	1.275	0.018	0.959	17.999	1.261	$3.077 \times 10^{-5}$
E	0	1	2	0.300	3.000	1.999	3.5000	0.003	0.049
F	0	1	1	0.151	3.545	2.128	1.889	0.001	0.078

Table 12 – Scenarios considered in the simulation studies for the LGB, LSL, LE and LR distributions.

Distribution	Scenario
LGB	A : $\mu = 5, \sigma = 2, \gamma = 0.5$ and $\alpha = 0.5$
	B : $\mu = -5, \sigma = 4, \gamma = 0.75$ and $\alpha = 1.5$
LS	A : $\mu = -5, \sigma = 1, \kappa = 2$ and $\alpha = 1.5$
	B : $\mu = 5, \sigma = 2, \kappa = 1$ and $\alpha = 0.5$

## 4.4 Considerations on the Maximum Likelihood Estimate and Simulations Studies

From Section 3.8, we have that the ML estimates for the parameters of the LGB and LSL distributions can be obtained by solving the system of equations given in Appendix B, where the baseline functions and the corresponding derivatives are presented in Appendices A and C. However, as we have done so far, we prefer to obtain the estimates by solving the optimization problem given in Equation (3.16). For this, we use the optim function of the R language via the L-BFGS-B algorithm. We verify by simulation experiments that  $(\bar{x}, s_x, 0, 1)$  and  $(\bar{x}, s_x, 1, 1)$ , where  $\bar{x}$  is the mean of the observations, and  $s_x$  the corresponding standard deviation, are good starting points for the iterative process in the LGB and LSL cases, respectively.

Based on the above considerations, we carry out two simulation studies in order to evaluate the behavior of the ML estimators for the parameters of the LGB and LSL distributions. For each of these distributions, from the qf given in Equation (3.11), we generate 1000 random samples considering different sample sizes ( $n$ ) and different choices for the parameters. The LambertW package (GOERG, 2011) in the R language was used to compute the principal branch of the LambertW function. Table 12 presents the different scenarios considered for each distribution. Note that for each distribution we consider one value of  $\alpha$  less than 1 and one greater than 1. In this way, we are considering a scenario where the skewness is positive and another where the skewness is negative.

Tables 13 and 14 report the average estimate (AE), the empirical standard deviation (SD), the asymptotic standard error (SE) and the root mean square error (RMSE) obtained in each scenario and sample size considered. In the tables, it is observed that the AE's tend to approach the real values of the parameters as the size of the sample increases. Furthermore, as expected in the standard asymptotic theory, it is observed that the SD's, SE's and RMSE's are close and decrease towards 0 as the sample size increases.

## 4.5 Data Analysis

In this section, we provide evidence that the LGB and LSL distributions can perform better in fitting real data than other commonly used distributions. Here, the ML estimates for the parameters of each fitted distribution are obtained using the optim function in the R language. We also calculate the modified Cramer-von Mises ( $W^*$ ) and Anderson-Darling ( $A^*$ ) statistics (CHEN; BALAKRISHNAN, 1995) to test the hypothesis  $H_0 : X_1, \dots, X_n$  is a random sample from a continuous distribution  $F(x; \theta)$ , where  $F(\cdot; \cdot)$  is known but  $\theta$  is unknown. In these tests,  $H_0$  is rejected at a significance level equal to 0.05 if  $W^* > 0.126$  and  $A^* > 0.752$ .

The data considered are briefly described below:

**Inflation Rate Data** This data set consists of 188 observations on the inflation rate (in %) registered quarterly between the years 1950 and 1996 in Canada. These data can be found in the Tbrate database of the Ecdat package (CROISSANT; GRAVES, 2019) in the R language.

**Per-Capita Income Data** This data set consists of 342 observations on the real per-capita income (on the logarithmic scale) recorded in member countries of the Organization for Economic Cooperation and Development (OECD). These observations can be found in the Gasoline database of the Ecdat package in the R language.

As in Section 3.10, for both data sets we test the hypothesis  $H_0$ : the data have exactly one mode versus the alternative hypothesis  $H_1$ : the data have at least two modes. The observed statistics and  $p$ -values associated with the excess mass test are presented in Table 15. In addition, some descriptive statistics are also reported in this table.

From the figures of Table 15, we observe that the distribution of the inflation rate data is at least bimodal and that it exhibits a positive skewness level. Consequently, we fit this data set with the LGB distribution and compare its performance with that of other bimodal asymmetric distributions such as the odd log-logistic skew-normal (OLLSN) (BRAGA; CORDEIRO; ORTEGA, 2018), gamma sinh-Cauchy (GSC) (GÓMEZ *et al.*, 2019) and mixture-normal (MN) (MCLACHLAN; PEEL, 2004) distributions. Regarding the per-capita income data, we observe that the data have a high level of kurtosis and that (at a significance level equal to 0.05) the hypothesis that the data have exactly one mode is not rejected. Consequently, we fit the per-capita income data with the LSL distribution and compare its performance with that of the other

Table 13 – Averages (AE), standard deviations (SD), root of the simulated mean square errors (RMSE), and averages of asymptotic standard errors (SE) for the estimates of  $\mu$ ,  $\sigma$ ,  $\gamma$ , and  $\alpha$  of the LGB distribution.

$n$	$\hat{\mu}$					$\hat{\sigma}$					$\hat{\gamma}$					$\hat{\alpha}$				
	AE	SD	SE	RMSE	AE	SD	SE	RMSE	AE	SD	SE	RMSE	AE	SD	SE	RMSE	AE	SD	SE	RMSE
Scenario A																				
100	5.005	0.329	0.329	0.329	1.992	0.126	0.129	0.126	0.126	0.497	0.260	0.247	0.260	0.517	0.168	0.169	0.169	0.169	0.169	0.169
200	5.004	0.234	0.234	0.235	1.993	0.092	0.092	0.091	0.091	0.498	0.177	0.170	0.177	0.508	0.120	0.118	0.119	0.119	0.119	0.119
300	5.003	0.192	0.188	0.192	1.995	0.076	0.074	0.076	0.076	0.498	0.139	0.135	0.139	0.504	0.097	0.096	0.097	0.097	0.097	0.097
500	5.001	0.148	0.145	0.148	1.997	0.058	0.057	0.058	0.058	0.499	0.106	0.104	0.106	0.503	0.077	0.075	0.077	0.077	0.077	0.077
1000	5.001	0.107	0.103	0.107	1.999	0.041	0.041	0.041	0.041	0.500	0.073	0.073	0.073	0.500	0.052	0.053	0.052	0.053	0.053	0.052
Scenario B																				
100	-4.943	0.774	0.798	0.776	3.977	0.284	0.293	0.285	0.285	0.745	0.413	0.388	0.412	1.478	0.315	0.336	0.316	0.316	0.316	0.316
200	-4.984	0.556	0.559	0.556	3.989	0.217	0.205	0.217	0.217	0.748	0.276	0.258	0.276	1.505	0.233	0.237	0.233	0.233	0.233	0.233
300	-4.996	0.464	0.455	0.464	3.996	0.167	0.168	0.167	0.167	0.749	0.227	0.210	0.227	1.504	0.193	0.193	0.193	0.193	0.193	0.193
500	-4.999	0.342	0.347	0.342	3.997	0.126	0.128	0.126	0.126	0.750	0.160	0.158	0.160	1.502	0.143	0.148	0.143	0.143	0.143	0.143
1000	-5.000	0.236	0.245	0.236	3.999	0.094	0.091	0.094	0.094	0.750	0.115	0.111	0.115	1.500	0.101	0.105	0.101	0.105	0.105	0.101



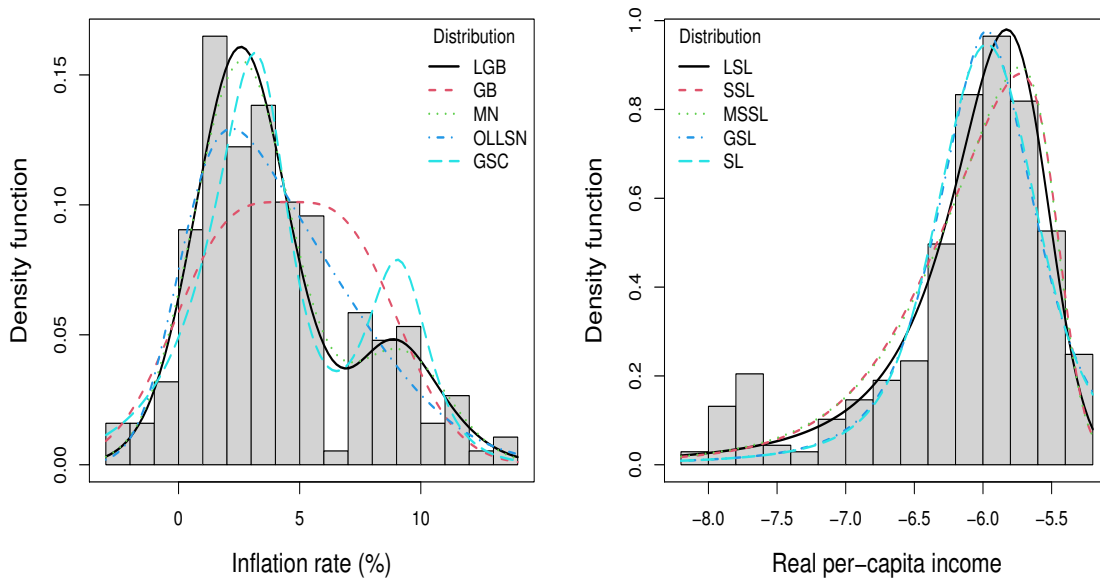
Table 14 – Averages (AE), standard deviations (SD), root of the simulated mean square errors (RMSE), and averages of asymptotic standard errors (SE) for the estimates of  $\mu$ ,  $\sigma$ ,  $\kappa$ , and  $\alpha$  of the LSL distribution.

$n$	$\hat{\mu}$			$\hat{\sigma}$			$\hat{\kappa}$			$\hat{\alpha}$						
	AE	SD	SE	RMSE	AE	SD	SE	RMSE	AE	SD	SE	RMSE				
Scenario A																
100	-4.901	0.713	0.608	0.720	1.061	0.212	0.222	0.221	2.306	0.139	0.879	1.083	1.543	0.701	0.621	0.701
200	-4.987	0.497	0.445	0.496	1.042	0.152	0.145	0.157	2.141	0.487	0.425	0.506	1.522	0.544	0.500	0.545
300	-4.995	0.391	0.365	0.391	1.026	0.115	0.114	0.118	2.096	0.360	0.325	0.372	1.520	0.464	0.425	0.466
500	-4.998	0.291	0.281	0.291	1.013	0.085	0.084	0.086	2.044	0.247	0.235	0.251	1.514	0.347	0.339	0.347
1000	-5.000	0.212	0.200	0.212	1.007	0.060	0.058	0.061	2.026	0.166	0.162	0.168	1.505	0.258	0.245	0.258
Scenario B																
100	5.116	0.992	0.907	0.998	2.046	0.398	0.401	0.380	1.054	0.205	0.187	0.212	0.509	0.236	0.222	0.236
200	5.048	0.616	0.606	0.618	2.010	0.260	0.259	0.260	1.019	0.132	0.123	0.134	0.507	0.161	0.156	0.161
300	5.033	0.515	0.493	0.516	2.023	0.222	0.212	0.223	1.022	0.106	0.101	0.108	0.502	0.131	0.127	0.131
500	5.022	0.380	0.473	0.382	2.001	0.161	0.162	0.161	1.008	0.079	0.076	0.079	0.501	0.099	0.097	0.099
1000	5.009	0.307	0.305	0.307	2.000	0.133	0.132	0.133	1.007	0.062	0.062	0.063	0.500	0.080	0.080	0.080

Table 15 – Observed statistics and the  $p$ -value for the excess mass tests and descriptive statistics for the inflation rate, per-capita income and violent crime rate data.

Data	Excess mass test		Descriptive statistics			
	Statistic	$p$ -value	Minimum	Maximum	Skewness	Kurtosis
Inflation Rate	0.050	0.050	-2.930	14	0.699	3.007
Per-Capita Income	0.033	0.126	-8.073	-5.221	-1.315	4.245

Figure 24 – Left panel: Histogram of the inflation rate data with the fitted pdf curves. Right panel: Histogram of the real per-capita income data with the fitted pdf curves.



heavy-tailed distributions such as the SL, skew-slash (SSL) (WANG; GENTON, 2006), modified skew-slash (MSSL) (REYES; GÓMEZ; VIDAL, 2016) and generalized slash (GSL) (GENC, 2007) distributions.

Table 16 reports the ML estimates and their standard errors (in parentheses) for each fitted distribution. Table 17 reports the values associated with the statistics  $A^*$  and  $W^*$  and with the information criteria. Here, based on the values of  $W^*$  and  $A^*$ , we observe that the quality of fit performed by the LGB and LSL distributions is appropriate in each case. In addition, we observe that the LGB and LSL distributions present the lowest AIC and BIC values, which suggests that these distributions should be chosen over the others to model the corresponding data sets.

Figure 14 presents the histograms for both data sets with the corresponding fitted pdf's. In the figures, it can be seen that the LGB and LSL density values are close to the empirical frequency values.

Table 16 – The ML estimates and their standard errors (in parentheses) for each distribution fitted to the inflation rate data and real per-capita income data.

Inflation rate data					
Parameter	LGB	MN	OLLSN	GSC	GB
$\mu$	6.682 (0.307)	2.600 (0.220)	0.834 (1.048)	6.441 (0.274)	4.454 (0.379)
$\mu_2$	-	9.086 (0.585)	-	-	-
$\sigma$	2.545 (0.131)	2.001 (0.161)	4.129 (1.401)	1.176 (0.104)	2.631 (0.230)
$\sigma_2$	-	2.026 (0.410)	-	-	-
$\gamma$	0.434 (0.175)	-	3.003 (0.880)	0.177 (0.044)	2.000 (1.321)
$\alpha$	0.183 (0.060)	0.777 (0.045)	0.772 (0.294)	0.572 (0.065)	-
Real per-capita income data					
Parameter	LSL	SSL	MSSL	GSL	SL
$\mu$	-5.011 (0.063)	-5.446 (0.044)	-5.457 (0.022)	-5.973 (0.027)	-5.972 (0.027)
$\sigma$	0.398 (0.041)	0.609 (0.106)	0.580 (0.028)	0.268 (0.028)	0.266 (0.025)
$\kappa$	2.900 (0.307)	3.107 (0.940)	3.680 (0.122)	1.751 (0.751)	1.727 (0.223)
$\alpha$	$6.975 \times 10^{-6}$ ( $3.220 \times 10^{-8}$ )	-4.907 (1.471)	-4.594 (0.120)	1.812 (0.418)	-

Table 17 – Statistics  $W^*$  and  $A^*$  and AIC and BIC values for each distribution fitted to the different data sets.

Distribution	Inflation rate data				Real per-capita income data				
	$W^*$	$A^*$	AIC	BIC	Distribution	$W^*$	$A^*$	AIC	BIC
LGB	0.045	0.248	962.5	975.4	LSL	0.044	0.630	543.5	558.9
MN	0.040	0.246	965.5	981.7	SSL	0.160	1.268	548.4	563.8
OLLSN	0.085	0.558	973.7	986.6	MSSL	0.149	1.204	545.5	560.9
GSC	0.158	1.083	982.1	995.0	GSL	0.657	4.662	619.6	634.9
GB	0.610	3.318	996.7	1006.4	SL	0.647	4.580	617.7	629.2

## 4.6 Final Comments

In this chapter, we proposed two new distributions to model skewness under the presence of bimodality and high levels of kurtosis, respectively, the LGB and LSL distributions. These distributions are generated considering as symmetric baseline distributions, the generalized bimodal and slash distributions.

The ability to capture bimodality and high levels of kurtosis by the LGB and LSL distributions, respectively, are capabilities inherited from the baseline distributions, while the ability to capture skewness is due to the performance of  $\alpha$  induced by the Lambert- $F$  generator.

We observe that the skewness coefficient of the LGB and LSL distributions behaves as a non-monotonic function of  $\alpha$  when the shape parameter inherited by the baseline distribution assumes a large value. Consequently, there may be different LGB (or LSL) distributions that have the same skewness value, and it is enough to make an appropriate choice of locations and scale to identify pairs of distributions that exhibit virtually identical behavior. By minimizing the Kullback-Leibler divergence we see that the alias distributions in the LGB and LSL families can be avoided if the ranges of  $\gamma$  and  $\kappa$  are set to be the intervals  $(0, 2)$  and  $(0, 5)$ , respectively.

Parameter estimation of the LGB and LSL distributions is discussed using the ML method. Through simulation experiments, we observe that the ML method provide acceptable estimates of the parameters of these distributions. Finally, by fitting two real data sets, we provide evidence that the LGB and LSL distributions can perform better than commonly used distributions.

---

## CONCLUDING REMARKS

---

In this dissertation, we propose new distributions to model asymmetry; the Lambert-uniform (LU), Lambert-exponential (LE), Lambert-Rayleigh (LR), Lambert-generalized bimodal (LGB), and Lambert-slash (LSL) distributions.

The LU distribution arises from a nonlinear transformation of a uniform rv, where the transformation is expressed by the principal branch of the Lambert  $W$  function. The LU pdf exhibits monotonic increasing/increasing shapes and tends to finite values at the extremes of the support, which in certain scenarios allows fitting the extreme empirical quantiles more appropriately than other distributions with bounded supports whose pdf's tend to 0 and  $\infty$  at the ends of the support. We note that the shape parameter of the LU distribution can be explicitly expressed as a function of the  $q$ th quantile, allowing us to parameterize the LU pdf in terms of the  $q$ th quantile. With this result, we propose a regression model that relates the  $q$ th quantile of the response to a linear predictor through an appropriate link function.

On the other hand, the LE, LR, LGB and LSL distributions arise as special cases of a new distribution generator that we call the Lambert- $F$  generator. This generator is formulated from the qf of an arbitrary baseline distribution, where the argument of the qf is a LU rv. Thus, for each choice of the baseline distribution, a new flexible distribution provided with one extra shape parameter is defined.

When the baseline distribution has positive support, such as in the LE and LR distributions, we observe that the Lambert- $F$  hrf corresponds to a modification in the early times of the baseline hrf. When the baseline distribution is symmetric and has unbounded support, such as the LGB and LSL distributions, the Lambert- $F$  pdf corresponds to a modification in a multiplicative fashion of the baseline pdf, allowing asymmetric shapes for the Lambert- $F$  pdf. In the LGB and LSL cases, we observe that these distributions inherit from the baseline distributions the ability to capture bimodality and high levels of kurtosis, respectively, while the ability to capture skewness is due to the performance of the shape parameter induced by the Lambert- $F$  generator.

For all the distributions proposed in this dissertation, the parameter estimation is performed via the maximum likelihood method. The maximum likelihood estimators for the parameters of the proposed distributions do not have closed forms, so the estimates are obtained by maximizing the corresponding likelihood functions using the `optim` function in the R language. Through simulation experiments, we observe that the estimates provided by the ML method are acceptable. Regarding the usefulness of the proposed distributions, by modeling data from different real-world settings, we observe that the Lambert- $F$  distributions may perform better than commonly used distributions. Consequently, we conclude that these new distributions can be considered as a viable alternative for classical distributions. Finally, based on the structural properties of the proposed distributions, as well as on the application examples, we think that a Lambert- $F$  distribution is able to perform appropriately in scenarios where the baseline distribution is commonly used. However, the performances of these new distributions will be especially valued in scenarios where the data exhibit levels of skewness that cannot be captured by the baseline distributions.

---

## FUTURE STUDIES

---

In this section, we propose some ideas for future work associated with the distributions proposed in this dissertation.

### 6.1 Lambert-F Quantile Regression Models

For a Lambert- $F$  type distribution, we observe that the parameter  $\alpha$  can be analytically expressed as a function of the  $q$ -th quantile, which allows the distribution to be parameterized in terms of its  $q$ -th quantile.

Denoting by  $\gamma$  the  $q$ -th quantile of a Lambert- $F$  type distribution, from Equation (3.11), we obtain that

$$\alpha = \left( \frac{1-q}{1-F(\gamma; \eta)} \right)^{\frac{1}{F(\gamma; \eta)}},$$

where  $F(\cdot; \eta)$  is the cdf of the baseline distribution.

Thus, the LU distribution can be easily parameterized in terms of the  $q$ th quantile, obtaining (for  $q \in (0, 1)$  is known) the pdf given by

$$f_X(x; \eta, \gamma) = f(x; \eta) \left( \frac{1-q}{1-F(\gamma; \eta)} \right)^{\frac{F(x; \eta)}{F(\gamma; \eta)}} \times \left[ 1 - \frac{1}{F(\gamma; \eta)} \log \left( \frac{1-q}{1-F(\gamma; \eta)} \right) S(x; \eta) \right], \quad (6.1)$$

where  $f(\cdot; \eta)$  and  $S(\cdot; \eta)$  are the pdf and the sf of the baseline distribution.

Now, for a quantile regression model based on a Lambert- $F$  type distribution, consider the following.

Let  $X_1, \dots, X_n$  be  $n$  random variables and denote by  $x_1, \dots, x_n$  the observed values. Assume that each  $X_i$  has pdf  $f_{X_i}(x; \eta, \gamma_i)$  given in Equation (6.1). The Lambert- $F$  quantile regression model is defined by establishing that the  $q$ th quantile  $\gamma_i$  of  $X_i$  satisfies the functional relationship

$g(\gamma_i) = w_i^t \beta$ ,  $i = 1, \dots, n$ , where  $w_i = (1, w_{i1}, \dots, w_{i(k-1)})^t$  is the vector of covariates associated to the response  $x_i$ ,  $\beta = (\beta_0, \beta_1, \dots, \beta_{(k-1)})^t$  is a  $k$ -dimensional vector of unknown regression coefficients and  $g(\cdot)$  is a strictly increasing and twice differentiable function that must be specified according to the support of the baseline distribution. For example, if the baseline distribution has a bounded support to the interval  $(0, 1)$ , a feasible choice for  $g(\cdot)$  is the logit function; If the baseline distribution has unbounded support, a feasible choice for  $g(\cdot)$  is the identity function.

A quantile regression model based on the LU distribution can be consulted in [Iriarte, de Castro and Gómez \(2021a\)](#).

## 6.2 Lambert Transformed U distributions

In Definition 1, the distribution of the rv  $U$  does not necessarily have to be uniform. By imposing only the conditions that  $U$  is continuous and has a range bounded to the interval  $(0, 1)$ , new distributions that generalize to the LU distribution can be derived.

**Definition 6.** Let  $U$  be a continuous rv with range bonded to the interval  $(0, 1)$  and cdf  $F_U(u; \beta)$ , where  $\beta$  is a parameter vector. Then, the rv  $X$ , represented as

$$X = \begin{cases} \frac{1}{\log(\alpha)} W_0 \left( \frac{\log(\alpha)(U-1)}{\alpha} \right) + 1, & \text{if } \alpha \in (0, 1) \cup (1, e), \\ U, & \text{if } \alpha = 1, \end{cases}$$

where  $W_0(\cdot)$  is the principal branch of the Lambert  $W$  function, is a Lambert transformed  $U$  rv.

In this case, the cdf and the pdf of  $X$  are given respectively by

$$\begin{aligned} F_X(x; \beta, \alpha) &= F_U(1 - (1-x)\alpha^x; \beta) \text{ and} \\ f_X(x; \beta, \alpha) &= f_U(1 - [1-x]\alpha^x; \beta) [1 - \log(\alpha)(1-x)] \alpha^x, \end{aligned}$$

where  $x \in (0, 1)$ ,  $\alpha \in (0, e)$  and  $F_U(\cdot; \beta)$  and  $f_U(\cdot; \beta)$  are the cdf and the pdf of  $U$ .

In Definition 1, the skewness of the rv  $X$  depends only on the value assumed by  $\alpha$ . On the other hand, in Definition 6, we suspect that the skewness of  $X$  may also be influenced by some shape parameter of the distribution of the rv  $U$ . We leave the following questions open:

If  $U$  has a power distribution with shape parameter  $\gamma$ , what effect will  $\gamma$  have on the skewness of the resulting Lambert transformed power distribution? Will the parameter  $\gamma$  allow different shapes for the resulting pdf other than those displayed by the LU pdf?

If  $U$  has a beta distribution, with shape parameters  $\gamma_1$  and  $\gamma_2$  such that  $\gamma_1 = \gamma_2$ , what effect will  $\gamma_1$  have on the skewness of the resulting Lambert transformed beta distribution? Will the parameter  $\gamma_1$  allow different shapes for the resulting pdf other than those displayed by the LU pdf?



Note that the power distribution, unlike the uniform distribution, is an asymmetric distribution. On the other hand, the beta distribution under the restriction imposed on the parameters is a symmetric distribution.

### 6.3 Alternative Distribution Generators

Similar to the construction of the Lambert- $F$  generator based on the LU rv, it is possible to define a new distribution generator based on a Lambert transformed  $U$  rv.

**Proposition 9.** Let  $Z$  be an arbitrary rv with cdf  $F(z; \eta)$ , where  $\eta$  is a parameter vector. The rv  $Y$  follows the Lambert- $F_U F$  distribution, if its cdf is given by

$$F_Y(y; \eta, \beta, \alpha) = F_U\left(1 - [1 - F(y; \eta)]\alpha^{F(y; \eta)}; \beta\right),$$

where  $F_U(\cdot; \cdot)$  is the cdf of a Lambert transformed  $U$  rv.

*Demonstration.* If  $Y = F^{-1}(U; \eta)$ , where  $F^{-1}(\cdot; \eta)$  is the qf of an arbitrary baseline rv and  $U$  is a Lambert transformed  $U$  rv, then the cdf of  $Y$  is given by  $F_Y(y; \eta, \alpha) = P(Y \leq y) = F_U(F(y; \eta); \beta, \alpha)$ , where  $F(\cdot; \eta)$  is the inverse function of the  $F^{-1}(\cdot; \eta)$  and  $F_U(u; \beta, \alpha) = F_U(1 - (1 - u)\alpha^u; \beta)$  is the cdf of  $U$ .  $\square$

Proposition 9 defines a new distribution generator class. Once the functions  $F(\cdot; \eta)$  and  $F_U(\cdot; \beta)$  are specified, a new family of distributions indexed by the parameter vector  $(\eta, \beta, \alpha)^t$  is defined.



## BIBLIOGRAPHY

---

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Ieee, v. 19, p. 716–723, 1974. Citation on page 38.
- AL-HUSSAINI, E. K. Inference based on censored samples from exponentiated populations. **Test**, Springer, v. 19, p. 487–513, 2010. Citation on page 46.
- ALZAATREH, A.; LEE, C.; FAMOYE, F. A new method for generating families of continuous distributions. **Metron**, Springer, v. 71, p. 63–79, 2013. Citation on page 22.
- AMEIJEIRAS-ALONSO, J.; CRUJEIRAS, R. M.; RODRÍGUEZ-CASAL, A. Multimode: An r package for mode assessment. **arXiv preprint arXiv:1803.00472**, 2018. Citation on page 60.
- \_\_\_\_\_. Mode testing, critical bandwidth and excess mass. **Test**, Springer, v. 28, p. 900–919, 2019. Citation on page 60.
- AZZALINI, A. A class of distributions which includes the normal ones. **Scandinavian journal of statistics**, JSTOR, p. 171–178, 1985. Citation on page 22.
- BRAGA, A. d. S.; CORDEIRO, G. M.; ORTEGA, E. M. A new skew-bimodal distribution with applications. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 47, p. 2950–2968, 2018. Citation on page 77.
- BYRD, R. H.; LU, P.; NOCEDAL, J.; ZHU, C. A limited memory algorithm for bound constrained optimization. **SIAM Journal on scientific computing**, SIAM, v. 16, p. 1190–1208, 1995. Citation on page 31.
- CASTILLO, N. O.; GALLARDO, D. I.; BOLFARINE, H.; GÓMEZ, H. W. Truncated power-normal distribution with application to non-negative measurements. **Entropy**, Multidisciplinary Digital Publishing Institute, v. 20, p. 433, 2018. Citation on page 46.
- CHEN, G.; BALAKRISHNAN, N. A general purpose approximate goodness-of-fit test. **Journal of Quality Technology**, Taylor & Francis, v. 27, p. 154–161, 1995. Citation on page 77.
- CHUNYING, Z. A quantile regression analysis on the relations between foreign direct investment and technological innovation in china. In: IEEE. **2011 International Conference of Information Technology, Computer Engineering and Management Sciences**. [S.l.], 2011. v. 4, p. 38–41. Citation on page 32.
- CORDEIRO, G. M.; de Castro, M. A new family of generalized distributions. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 81, p. 883–898, 2011. Citations on pages 22 and 46.
- CROISSANT, Y.; GRAVES, S. **Ecdat: Data Sets for Econometrics**. [S.l.], 2019. R package version 0.3-4. Available: <<https://CRAN.R-project.org/package=Ecdat>>. Citation on page 77.
- DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**, Taylor & Francis, v. 5, p. 236–244, 1996. Citation on page 38.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. **Communications in Statistics-Theory and methods**, Taylor & Francis, v. 31, p. 497–512, 2002. Citations on pages 22 and 46.

FARAWAY, J.; MARSAGLIA, G.; MARSAGLIA, J.; BADDELEY, A. **gofstest: Classical Goodness-of-Fit Tests for Univariate Distributions**. [S.l.], 2019. R package version 1.2-2. Available: <<https://CRAN.R-project.org/package=gofstest>>. Citation on page 38.

FERREIRA, J. T. S.; STEEL, M. F. J. A constructive representation of univariate skewed distributions. **Journal of the American Statistical Association**, Taylor & Francis, v. 101, p. 823–829, 2006. Citation on page 22.

GENC, A. I. A generalization of the univariate slash by a scale-mixed exponential power distribution. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 36, p. 937–947, 2007. Citation on page 80.

GIRMA, S.; GÖRG, H. Foreign direct investment, spillovers and absorptive capacity: Evidence from quantile regressions. Discussion Paper 1. GEP Working Paper 2002/14, 2003. Citation on page 32.

GOERG, G. M. Lambert W random variables-a new family of generalized skewed distributions with applications to risk estimation. **The Annals of Applied Statistics**, Institute of Mathematical Statistics, v. 5, p. 2197–2230, 2011. Citations on pages 22, 23, 25, 34, 57, and 76.

\_\_\_\_\_. The lambert way to gaussianize heavy-tailed data with the inverse of tukey's h transformation as a special case. **The Scientific World Journal**, Hindawi, v. 2015, 2015. Citation on page 23.

GÓMEZ-DÉNIZ, E.; IRIARTE, Y. A.; CALDERÍN-OJEDA, E.; GÓMEZ, H. W. Modified power-symmetric distribution. **Symmetry**, Multidisciplinary Digital Publishing Institute, v. 11, p. 1410, 2019. Citation on page 46.

GÓMEZ-DÉNIZ, E.; SORDO, M. A.; CALDERÍN-OJEDA, E. The log–lindley distribution as an alternative to the beta regression model with applications in insurance. **Insurance: Mathematics and Economics**, Elsevier, v. 54, p. 49–57, 2014. Citation on page 41.

GÓMEZ, H. W.; QUINTANA, F. A.; TORRES, F. J. A new family of slash-distributions with elliptical contours. **Statistics & probability letters**, Elsevier, v. 77, p. 717–725, 2007. Citation on page 66.

GÓMEZ, Y. M.; GÓMEZ-DÉNIZ, E.; VENEGAS, O.; GALLARDO, D. I.; GÓMEZ, H. W. An asymmetric bimodal distribution with application to quantile regression. **Symmetry**, Multidisciplinary Digital Publishing Institute, v. 11, p. 899, 2019. Citation on page 77.

GUPTA, R. D.; KUNDU, D. Theory & methods: Generalized exponential distributions. **Australian & New Zealand Journal of Statistics**, Wiley Online Library, v. 41, p. 173–188, 1999. Citations on pages 46 and 60.

HASSAN, M.; EL-BASSIOUNI, M. Bimodal skew-symmetric normal distribution. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 45, p. 1527–1541, 2016. Citation on page 65.

HUTSON, A. D.; VEXLER, A. A cautionary note on beta families of distributions and the aliases within. **The American Statistician**, Taylor & Francis, v. 72, p. 121–129, 2018. Citation on page 72.

IRIARTE, Y. A.; de Castro, M.; GÓMEZ, H. W. The Lambert-F distributions class: An alternative family for positive data analysis. **Mathematics**, Multidisciplinary Digital Publishing Institute, v. 8, p. 1398, 2020. Citation on page 54.

\_\_\_\_\_. An alternative one-parameter distribution for bounded data modeling generated from the Lambert transformation. **Symmetry**, Multidisciplinary Digital Publishing Institute, v. 13, p. 1190, 2021. Citation on page 86.

\_\_\_\_\_. A unimodal/bimodal skew/symmetric distribution generated from Lambert's transformation. **Symmetry**, Multidisciplinary Digital Publishing Institute, v. 13, p. 269, 2021. Citation on page 72.

JODRÁ, P. Computer generation of random variables with lindley or poisson–lindley distribution via the lambert w function. **Mathematics and Computers in Simulation**, Elsevier, v. 81, p. 851–859, 2010. Citation on page 22.

JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Probability Distributions**. [S.l.]: Wiley, New York, 1994. Citations on pages 25, 28, 50, and 66.

JONES, M. On a class of distributions defined by the relationship between their density and distribution functions. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 36, p. 1835–1843, 2007. Citation on page 46.

JONES, M. C. Families of distributions arising from distributions of order statistics. **Test**, Springer, v. 13, p. 1–43, 2004. Citation on page 46.

JOSE, K.; KRISHNA, E. Marshall-olkin extended uniform distribution. In: **ProbStat Forum**. [S.l.: s.n.], 2011. v. 4, p. 78–88. Citation on page 28.

JOYNER, W. B.; BOORE, D. M. Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 imperial valley, california, earthquake. **Bulletin of the Seismological Society of America**, The Seismological Society of America, v. 71, p. 2011–2038, 1981. Citation on page 38.

KLEIBER, C.; ZEILEIS, A. **Applied Econometrics with R**. New York: Springer-Verlag, 2008. ISBN 978-0-387-77316-2. Available: <<https://CRAN.R-project.org/package=AER>>. Citation on page 60.

KORKMAZ, M. Ç.; CHESNEAU, C.; KORKMAZ, Z. S. On the arcsecant hyperbolic normal distribution. properties, quantile regression modeling and applications. **Symmetry**, Multidisciplinary Digital Publishing Institute, v. 13, p. 117, 2021. Citation on page 38.

KUMARASWAMY, P. A generalized probability density function for double-bounded random processes. **Journal of hydrology**, Elsevier, v. 46, p. 79–88, 1980. Citation on page 25.

KYLE, R. A. “Benign” monoclonal gammopathy – after 20 to 35 years of follow-up. **Mayo Clinic Proceedings**, v. 68, p. 26–36, 1993. Citation on page 60.

- LAI, C. Constructions and applications of lifetime distributions. **Applied Stochastic Models in Business and Industry**, Wiley Online Library, v. 29, p. 127–140, 2013. Citations on pages 21 and 46.
- MARSHALL, A. W.; OLKIN, I. A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. **Biometrika**, Oxford University Press, v. 84, p. 641–652, 1997. Citation on page 46.
- MCLACHLAN, G. J.; PEEL, D. **Finite mixture models**. [S.l.]: John Wiley & Sons, 2004. Citation on page 77.
- MITNIK, P. A.; BAEK, S. The kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation. **Statistical Papers**, Springer, v. 54, p. 177–192, 2013. Citation on page 38.
- NADARAJAH, S. **The exponentiated exponential distribution: a survey**. [S.l.]: Springer, 2011. Citation on page 49.
- NADARAJAH, S.; KOTZ, S. The beta gumbel distribution. **Mathematical Problems in engineering**, Hindawi, v. 2004, p. 323–332, 2004. Citation on page 46.
- \_\_\_\_\_. The exponentiated type distributions. **Acta Applicandae Mathematica**, Springer, v. 92, p. 97–111, 2006. Citations on pages 46 and 47.
- R Core Team. **R: A Language and environment for statistical computing**. Vienna, Austria, 2019. Available: <<http://www.R-project.org/>>. Citation on page 30.
- RAO, K. On a bivariate bimodal distribution. In: **Presented at ISPS Annual Conference**. Delhi, India: [s.n.], 1987. Citation on page 65.
- REYES, J.; GÓMEZ, H. W.; VIDAL, I. Modified skew-slash distribution. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 45, p. 1070–1080, 2016. Citation on page 80.
- SARMA, P.; RAO, K. S.; RAO, R. P. On a family of bimodal distributions. **Sankhyā: The Indian Journal of Statistics, Series B**, JSTOR, p. 287–292, 1990. Citation on page 65.
- SCHMIT, J. T.; ROTH, K. Cost effectiveness of risk management practices. **Journal of Risk and Insurance**, JSTOR, p. 455–470, 1990. Citation on page 41.
- SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, Institute of Mathematical Statistics, p. 461–464, 1978. Citation on page 38.
- SHAKED, M.; SHANTHIKUMAR, J. G. **Stochastic Orders**. [S.l.]: Springer, New York, 2007. Citation on page 50.
- SHAW, W. T.; BUCKLEY, I. R. The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. **UCL discovery repository**, p. 1–16, 2007. Citations on pages 28 and 46.
- SOETAERT, K. rootsolve: Nonlinear root finding, equilibrium and steady-state analysis of ordinary differential equations. **R package version**, v. 1, 2009. Citation on page 68.
- STEHLÍK, M. Distributions of exact tests in the exponential family. **Metrika**, Springer, v. 57, p. 145–164, 2003. Citation on page 22.

SURLES, J.; PADGETT, W. Inference for reliability and stress-strength for a scaled Burr Type X distribution. **Lifetime Data Analysis**, Springer, v. 7, p. 187–200, 2001. Citation on page 60.

THERNEAU, T. M. **A Package for Survival Analysis in S**. [S.l.], 2015. Version 2.38. Available: <<https://CRAN.R-project.org/package=survival>>. Citation on page 60.

WANG, J.; GENTON, M. G. The multivariate skew-slash distribution. **Journal of Statistical Planning and Inference**, Elsevier, v. 136, p. 209–220, 2006. Citation on page 80.

WEIBULL, W. A statistical theory of the strength of materials. **Ingeniors Vetenskaps Akademiens, Stockholm 151**, 1939. Citation on page 50.

ZOGRAFOS, K.; BALAKRISHNAN, N. On families of beta-and generalized gamma-generated distributions and associated inference. **Statistical methodology**, Elsevier, v. 6, p. 344–362, 2009. Citation on page 46.





## BASELINE FUNCTIONS FOR THE SPECIAL CASES GIVEN IN SECTIONS 3.7 AND 4.2

Table 18 – The pdf, the cdf and the sf for the generalized-bimodal (GB), slash (SL), exponential (E) and Rayleigh (R) baseline distributions.

Distribution	$\eta_1$	$\eta_2$	$\eta_3$	Pdf $f(x; \eta)$	Cdf $F(x; \eta)$
GB	$\mu$	$\sigma$	$\gamma$	$\frac{\gamma + \left(\frac{x-\mu}{\sigma}\right)^2}{\sigma(1+\gamma)} \phi\left(\frac{x-\mu}{\sigma}\right), x \in \mathbb{R},$	$\Phi\left(\frac{x-\mu}{\sigma}\right) - \frac{x-\mu}{\sigma(1+\gamma)} \phi\left(\frac{x-\mu}{\sigma}\right)$
SL	$\mu$	$\sigma$	$\kappa$	$\frac{\kappa}{\sigma} \int_0^1 u^\kappa \phi\left(\frac{x-\mu}{\sigma} u\right) du, x \in \mathbb{R},$	$\Phi\left(\frac{x-\mu}{\sigma}\right) - \frac{x-\mu}{\sigma\kappa} f(x; \eta)$
E	$\sigma$	-	-	$\frac{1}{\sigma} e^{-\frac{x}{\sigma}}, x > 0,$	$1 - e^{-\frac{x}{\sigma}}$
R	$\sigma$	-	-	$\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, x > 0,$	$1 - e^{-\frac{x^2}{2\sigma^2}}$

For  $\mu \in \mathbb{R}$  (location),  $\sigma > 0$  (scale),  $\gamma \in (0, 2)$  and  $\kappa > 0$  (shape).

$\Phi(\cdot)$  and  $\phi(z)$  denote the cdf and pdf of the standard normal distribution.



## MAXIMUM LIKELIHOOD ESTIMATION FOR THE LAMBERT-F DISTRIBUTION

---



---

Let  $X_1, \dots, X_n$  be a random sample of  $X \sim \text{LF}(\eta, \alpha)$  with  $\eta = (\eta_1, \dots, \eta_{k-1})^t$  and  $(\eta, \alpha)^t$  unknown. Then, the ML estimate of  $(\eta, \alpha)^t$  satisfies the system of equations

$$\begin{aligned} \log(\alpha) \sum_{i=1}^n \frac{S_{\eta_1}(x_i; \eta)}{1 - \log(\alpha)S(x_i; \eta)} - \log(\alpha) \sum_{i=1}^n F_{\eta_1}(x_i; \eta) &= \sum_{i=1}^n \frac{f_{\eta_1}(x_i; \eta)}{f(x_i; \eta)}, \\ &\vdots \\ \log(\alpha) \sum_{i=1}^n \frac{S_{\eta_{k-1}}(x_i; \eta)}{1 - \log(\alpha)S(x_i; \eta)} - \log(\alpha) \sum_{i=1}^n F_{\eta_{k-1}}(x_i; \eta) &= \sum_{i=1}^n \frac{f_{\eta_{k-1}}(x_i; \eta)}{f(x_i; \eta)}, \\ \sum_{i=1}^n \frac{S(x_i; \eta)}{1 - \log(\alpha)S(x_i; \eta)} &= \sum_{i=1}^n F(x_i; \eta), \end{aligned}$$

where  $f_{\eta_j}(x; \eta) = \partial f(x; \eta) / \partial \eta_j$ ,  $F_{\eta_j}(x; \eta) = \partial F(x; \eta) / \partial \eta_j$  and  $S_{\eta_j}(x; \eta) = -F_{\eta_j}(x; \eta)$ , with  $j = 1, \dots, k-1$ , obtained from the score equation system.



## FUNCTIONS FOR THE COMPUTATION OF THE SCORE FUNCTIONS AND THE OBSERVED INFORMATION MATRIX OF THE SPECIAL CASES REPORTED IN SECTIONS **3.7 AND 4.2**

Table 19 – Partial derivatives for the pdf and the cdf of the E baseline distribution.

Derivative	Expression
$f_{\sigma}(x; \sigma)$	$-\frac{1}{\sigma} \left(1 - \frac{x}{\sigma}\right) f(x; \sigma)$
$f_{\sigma\sigma}(x; \sigma)$	$\frac{1}{\sigma^2} \left[ \left(1 - \frac{x}{\sigma}\right) \left(2 - \frac{x}{\sigma}\right) - \frac{x}{\sigma} \right] f(x; \sigma)$
$F_{\sigma}(x; \sigma)$	$-\frac{x}{\sigma} f(x; \sigma)$
$F_{\sigma\sigma}(x; \sigma)$	$\frac{x}{\sigma^2} f(x; \sigma) - \frac{x}{\sigma} f_{\sigma}(x; \sigma)$

$f(\cdot; \cdot)$  being the pdf of the E distribution (See Table 18).

Table 20 – Partial derivatives for the pdf and the cdf of the R baseline distribution.

Derivative	Expression
$f_{\sigma}(x; \sigma)$	$-\frac{1}{\sigma} \left(2 - \frac{x^2}{\sigma^2}\right) f(x; \sigma)$
$f_{\sigma\sigma}(x; \sigma)$	$\frac{1}{\sigma^2} \left[ \left(3 - \frac{x^2}{\sigma^2}\right) \left(2 - \frac{x^2}{\sigma^2}\right) - \frac{2x^2}{\sigma^2} \right] f(x; \sigma)$
$F_{\sigma}(x; \sigma)$	$-\frac{x}{\sigma} f(x; \sigma)$
$F_{\sigma\sigma}(x; \sigma)$	$\frac{x}{\sigma^2} f(x; \sigma) - \frac{x}{\sigma} f_{\sigma}(x; \sigma)$

$f(\cdot; \cdot)$  being the pdf of the E distribution (See Table 18).

Table 21 – Partial derivatives for the pdf and the cdf of the GB baseline distribution.

Derivative	Expression
$f_{\mu}(x; \mu, \sigma, \gamma)$	$-\frac{2z}{\sigma^2(1+\gamma)}\phi(z) + \frac{z}{\sigma}f(x; \mu, \sigma, \gamma)$
$f_{\sigma}(x; \mu, \sigma, \gamma)$	$-\frac{\gamma+3z^2}{\sigma^2(1+\gamma)}\phi(z) + \frac{z^2}{\sigma}f(x; \mu, \sigma, \gamma)$
$f_{\gamma}(x; \mu, \sigma, \gamma)$	$\frac{1}{\sigma(1+\gamma)}\phi(z) - \frac{1}{1+\gamma}f(x; \mu, \sigma, \gamma)$
$F_{\mu}(x; \mu, \sigma, \gamma)$	$-\frac{1}{\sigma}f(x; \mu, \sigma, \gamma)$
$F_{\sigma}(x; \mu, \sigma, \gamma)$	$-\frac{z}{\sigma}f(x; \mu, \sigma, \gamma)$
$F_{\gamma}(x; \mu, \sigma, \gamma)$	$\frac{z}{(1+\gamma)^2}\phi(z)$
$f_{\mu\mu}(x; \mu, \sigma, \gamma)$	$\frac{2(1-z^2)}{\sigma^3(1+\gamma)}\phi(z) - \frac{1}{\sigma^2}f(x; \mu, \sigma, \gamma) + \frac{z}{\sigma}f_{\mu}(x; \mu, \sigma, \gamma)$
$f_{\mu\sigma}(x; \mu, \sigma, \gamma)$	$\frac{z}{\sigma^3}\left(\frac{6}{1+\gamma} + z^2\right)\phi(z) - \frac{2z}{\sigma^2}f(x; \mu, \sigma, \gamma) + \frac{z}{\sigma}f_{\sigma}(x; \mu, \sigma, \gamma)$
$f_{\mu\gamma}(x; \mu, \sigma, \gamma)$	$\frac{2z}{\sigma^2(1+\gamma)^2} + \frac{z}{\sigma}f_{\gamma}(x; \mu, \sigma, \gamma)$
$f_{\sigma\sigma}(x; \mu, \sigma, \gamma)$	$\frac{6z^2}{\sigma^3(1+\gamma)}\phi(z) + \frac{(2-z^2)(\gamma+3z^2)}{\sigma^3(1+\gamma)}\phi(z) - \frac{3z^2}{\sigma^2}f(x; \mu, \sigma, \gamma) + \frac{z^2}{\sigma}f_{\sigma}(x; \mu, \sigma, \gamma)$
$f_{\sigma\gamma}(x; \mu, \sigma, \gamma)$	$-\frac{1-3z^2}{\sigma^2(1+\gamma)^2}\phi(z) + \frac{z^2}{\sigma}f_{\gamma}(x; \mu, \sigma, \gamma)$
$f_{\gamma\gamma}(x; \mu, \sigma, \gamma)$	$-\frac{1}{\sigma(1+\gamma)^2}\phi(z) + \frac{1}{(1+\gamma)^2}f(x; \mu, \sigma, \gamma) - \frac{1}{1+\gamma}f_{\gamma}(x; \mu, \sigma, \gamma)$
$F_{\mu\mu}(x; \mu, \sigma, \gamma)$	$-\frac{1}{\sigma}f_{\mu}(x; \mu, \sigma, \gamma)$
$F_{\mu\sigma}(x; \mu, \sigma, \gamma)$	$\frac{1}{\sigma^2}f(x; \mu, \sigma, \gamma) - \frac{1}{\sigma}f_{\sigma}(x; \mu, \sigma, \gamma)$
$F_{\mu\gamma}(x; \mu, \sigma, \gamma)$	$-\frac{1}{\sigma}f_{\gamma}(x; \mu, \sigma, \gamma)$
$F_{\sigma\sigma}(x; \mu, \sigma, \gamma)$	$\frac{2z}{\sigma^2}f(x; \mu, \sigma, \gamma) - \frac{z}{\sigma}f_{\sigma}(x; \mu, \sigma, \gamma)$
$F_{\sigma\gamma}(x; \mu, \sigma, \gamma)$	$-\frac{z}{\sigma}f_{\gamma}(x; \mu, \sigma, \gamma)$
$F_{\gamma\gamma}(x; \mu, \sigma, \gamma)$	$-\frac{2z}{(1+\gamma)^3}\phi(z)$

$f(\cdot; \cdot, \cdot, \cdot)$  being the pdf of the GB distribution (See Table 18),  $\phi(\cdot)$  the pdf of the standard normal distribution and  $z = (x - \mu)/\sigma$ .

Table 22 – Partial derivatives for the pdf and the cdf of the SL baseline distribution.

Derivative	Expression
$f_{\mu}(x; \mu, \sigma, \kappa)$	$\frac{\kappa z}{\sigma(\kappa+2)} f(x; \mu, \sigma, \kappa+2)$
$f_{\sigma}(x; \mu, \sigma, \kappa)$	$-\frac{1}{\sigma} f(x; \mu, \sigma, \kappa) + \frac{\kappa z^2}{\sigma(\kappa+2)} f(x; \mu, \sigma, \kappa+2)$
$f_{\kappa}(x; \mu, \sigma, \kappa)$	$\frac{1}{\kappa} f(x; \mu, \sigma, \kappa) + \frac{\kappa}{\sigma} \int_0^1 u^{\kappa} \log(u) \phi(z) du$
$F_{\mu}(x; \mu, \sigma, \kappa)$	$-\frac{1}{\sigma} f(x; \mu, \sigma, \kappa)$
$F_{\sigma}(x; \mu, \sigma, \kappa)$	$-\frac{z}{\sigma} f(x; \mu, \sigma, \kappa)$
$F_{\kappa}(x; \mu, \sigma, \kappa)$	$\frac{z}{\kappa^2} \phi(z)$
$f_{\mu\mu}(x; \mu, \sigma, \kappa)$	$-\frac{\kappa}{\sigma^2(\kappa+2)} f(x; \mu, \sigma, \kappa+2) + z f_{\mu}(x; \mu, \sigma, \kappa+2)$
$f_{\mu\sigma}(x; \mu, \sigma, \kappa)$	$-\frac{2\kappa z}{\sigma^2(\kappa+2)} f(x; \mu, \sigma, \kappa+2) + \frac{\kappa z}{\kappa+2} f_{\sigma}(x; \mu, \sigma, \kappa+2)$
$f_{\mu\kappa}(x; \mu, \sigma, \kappa)$	$\frac{2z}{\sigma(\kappa+2)^2} f(x; \mu, \sigma, \kappa+2) + \frac{\kappa z}{\sigma(\kappa+2)} f_{\kappa}(x; \mu, \sigma, \kappa+2)$
$f_{\sigma\sigma}(x; \mu, \sigma, \kappa)$	$\frac{1}{\sigma^2} f(x; \mu, \sigma, \kappa) - \frac{1}{\sigma} f_{\sigma}(x; \mu, \sigma, \kappa) - \frac{3\kappa z^2}{\sigma^2(\kappa+2)} f(x; \mu, \sigma, \kappa+2)$ $+ \frac{\kappa z^2}{(\kappa+2)} f_{\sigma}(x; \mu, \sigma, \kappa+2)$
$f_{\sigma\kappa}(x; \mu, \sigma, \kappa)$	$-\frac{1}{\sigma} f_{\kappa}(x; \mu, \sigma, \kappa) + \frac{2z^2}{\sigma(\kappa+2)^2} f(x; \mu, \sigma, \kappa+2)$ $+ \frac{\kappa z^2}{\sigma(\kappa+2)} f_{\kappa}(x; \mu, \sigma, \kappa+2)$
$f_{\kappa\kappa}(x; \mu, \sigma, \kappa)$	$-\frac{1}{\kappa^2} f(x; \mu, \sigma, \kappa) + \frac{1}{\kappa} f_{\kappa}(x; \mu, \sigma, \kappa) + \frac{1}{\sigma} \int_0^1 u^{\kappa} \log(u) \phi(zu) du$ $+ \frac{\kappa}{\sigma} \int_0^1 u^{\kappa} \log^2(u) \phi(zu) du$
$F_{\mu\mu}(x; \mu, \sigma, \kappa)$	$-\frac{1}{\sigma} f_{\mu}(x; \mu, \sigma, \kappa)$
$F_{\mu\sigma}(x; \mu, \sigma, \kappa)$	$-\frac{1}{\sigma^2} f(x; \mu, \sigma, \kappa) - \frac{1}{\sigma} f_{\sigma}(x; \mu, \sigma, \kappa)$
$F_{\mu\kappa}(x; \mu, \sigma, \kappa)$	$-\frac{1}{\sigma} f(x; \mu, \sigma, \kappa)$
$F_{\sigma\sigma}(x; \mu, \sigma, \kappa)$	$\frac{2z}{\sigma^2} f(x; \mu, \sigma, \kappa) + \frac{z}{\sigma} f_{\sigma}(x; \mu, \sigma, \kappa)$
$F_{\sigma\kappa}(x; \mu, \sigma, \kappa)$	$-\frac{z}{\sigma} f_{\kappa}(x; \mu, \sigma, \kappa)$
$F_{\kappa\kappa}(x; \mu, \sigma, \kappa)$	$-\frac{2z}{\kappa^3} \phi(z)$
$f(\cdot; \cdot, \cdot, \cdot)$ being the pdf of the SL distribution (See Table 18), $\phi(\cdot)$ the pdf of the standard normal distribution and $z = (x - \mu)/\sigma$ .	





## SECOND PARTIAL DERIVATIVES OF THE LOG-LIKELIHOOD FUNCTION GIVEN IN EQUATION (2.11)

The second partial derivative of Equation (2.11) is given by

$$\begin{aligned}
\frac{\partial^2 \ell(\beta)}{\partial \beta_p \partial \beta_r} = & -\log(1-q) \sum_{i=1}^n \frac{x_i \eta_{i,r,p}}{\eta_i^2} + 2 \log(1-q) \sum_{i=1}^n \frac{x_i \eta_{i,r} \eta_{i,p}}{\eta_i^3} - \sum_{i=1}^n \frac{x_i \eta_{i,r} \eta_{i,p}}{\eta_i^2 (1-\eta_i)} \\
& + \sum_{i=1}^n \frac{x_i \log(1-\eta_i) \eta_{i,r,p}}{\eta_i^2} - 2 \sum_{i=1}^n \frac{x_i \log(1-\eta_i) \eta_{i,r} \eta_{i,p}}{\eta_i^3} + \sum_{i=1}^n \frac{x_i (1-\eta_i) \eta_{i,r,p}}{\eta_i^3} \\
& - \sum_{i=1}^n \frac{x_i \eta_{i,r} (1-\eta_i) \eta_{i,p}}{\eta_i^4} + \sum_{i=1}^n \frac{x_i \eta_{i,r} \eta_{i,p}}{\eta_i^3} + \sum_{i=1}^n \frac{(1-x_i) \eta_{i,r} \eta_{i,p}}{\eta_i^2 (1-\eta_i) H(x_i; \eta_i)} \\
& + \sum_{i=1}^n \frac{(1-x_i) \log\left(\frac{1-q}{1-\eta_i}\right) \eta_{i,r,p}}{\eta_i^2 H(x_i; \eta_i)} - 2 \sum_{i=1}^n \frac{(1-x_i) \log\left(\frac{1-q}{1-\eta_i}\right) \eta_{i,r} \eta_{i,p}}{\eta_i^3 H(x_i; \eta_i)} \\
& - \sum_{i=1}^n \frac{(1-x_i)^2 \log^2\left(\frac{1-q}{1-\eta_i}\right) \eta_{i,r} \eta_{i,p}}{\eta_i^4 H^2(x_i; \eta_i)} - \sum_{i=1}^n \frac{(1-x_i) \eta_{i,p}}{(1-\eta_i) H(x_i; \eta_i)} \\
& - \sum_{i=1}^n \frac{(1-x_i) \eta_{i,r} \eta_{i,p}}{(1-\eta_i)^2 H(x_i; \eta_i)} - \sum_{i=1}^n \frac{(1-x_i)^2 \log\left(\frac{1-q}{1-\eta_i}\right) \eta_{i,r} \eta_{i,p}}{\eta_i^2 (1-\eta_i) H^2(x_i; \eta_i)} \\
& + \sum_{i=1}^n \frac{(1-x_i)^2 \eta_{i,r} \eta_{i,p}}{\eta_i (1-\eta_i)^2 H^2(x_i; \eta_i)},
\end{aligned}$$

where  $H(x_i; \eta_i) = 1 - (1/\eta_i) \log[(1-q)/(1-\eta_i)](1-x_i)$ ,  $\eta_{i,r}$  and  $\eta_{i,s}$  are as in Equation (2.12) and  $\eta_{i,r,p} = \partial^2 \eta_i / (\partial \beta_p \partial \beta_r)$ , with  $r, p = 0, 1, \dots, k-1$ . Thus, under the consideration of the link logit, we observe that  $\eta_{i,0,p} = \delta_i w_{ip}$ ,  $\eta_{i,r,0} = \delta_i w_{ir}$  and  $\eta_{i,r,p} = \delta_i w_{ir} w_{ip}$ , where  $\delta_i = \eta_i(1-\eta_i)(1-2\eta_i)$ , with  $i = 1, 2, \dots, n$ , and  $r, p = 0, 1, \dots, k-1$ .

