University of São Paulo
"Luiz de Queiroz" College of Agriculture

Extensions of the normal distribution using the odd log-logistic family: theory and applications

Altemir da Silva Braga

Thesis presented to obtain the degree of Doctor in Science. Area: Statistics and Agricultural Experimentation

Piracicaba
2017

Altemir da Silva Braga
Degree in Mathematics

Extensions of the normal distribution using the odd log-logistic family: theory and applications

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Prof. Dr. **EDWIN MOISES MARCOS ORTEGA**

Thesis presented to obtain the degree of Doctor in Science. Area: Statistics and Agricultural Experimentation

Piracicaba
2017

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Braga, Altemir da Silva
Extensions of the normal distribution using the odd log-logistic family: theory and applications / Altemir da Silva Braga. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2017.
166 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. Família de distribuição 2. Log-logística 3. Modelo de Regressão 4. Estatística experimental. I. Título

"Joy lies in the fight, in the attempt, in the suffering involved, not in the victory itself."

Mahatma Gandhi

DEDICATION

To my dear parents,
Sebastião da Silva Braga *and* **Maria Lucy**,
the " support", of my life.

My wife, **Carmem Silvia Bandeira Teixeira,** and my children **Tays Bandeira Braga, Eric Henrrir Bandeira Braga and Caio Bandeira Braga**, *for always being by my side.*

My brothers, **Geane da Silva Braga, Leane da Silva Braga, Lucina da Silva Braga, Vanusa da Silva Braga e Adalzemir da Silva Braga**, *for the friendship, love and support.*

To them, I dedicate this work.

ACKNOWLEDGEMENTS

The God for always is blessing my family.

To my wife, Carmen Silvia Bandeira Teixeira and my children Tays Bandeira Braga, Eric Enrrir Bandeira Braga and Caio Bandeira Braga, for all help and support that has given me.

To my parents, Sebastião da Silva Braga and Maria Lucy, for love, teaching and for showing me that the work is a virtue that can make our dreams come true.

To my advisor, Prof. Dr. Edwin Moises Marcos Ortega, for the continuous support of my Doctorate, motivation and knowledge to develop new methods that contribute to the development of the Statistics.

To Prof. Dr. Gauss Moutinho Cordeiro by scientific contributions and support that are always helping to develop new theories in the area of Statistics.

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, the Fundação de Amparo a Pesquisas do Estado do Acre - FAPAC and Universidade Federal do Acre - UFAC for the financial support that contributed for that I could dedicate exclusively to the activities of the doctorate.

To Prof. Dr. Sônia Maria De Stefano Piedade, Prof. Dr. César Gonçalves de Lima and Prof. Dr. Victor Hugo Lachos Dávila for the suggestions and contributions for the development of this work.

To my friends José Nilton, Rodrigo Pescim, Thiago Gentil and Djair Durand for the friendship and discussions about survival analysis that enriched my knowledge.

To my friends from Piracicaba and from the Department of Exact Sciences at ESALQ/USP, Alessandra Santos, Tiago Santana, Ana Júlia, Natália, Natalie Verónika, Maria Cristina, Ricardo Klein, Rafael Moral, Reginaldo, Thiago Oliveira, Erasnilson Vieira, Rick Mangueira, Douglas Toledo, Gislaine, Fábio Prataviera, Julio Cesar and Raquel.

To my friends Edcarlos Miranda de Souza and José Roberto Guimarães de Souza for friendship and support.

To employees of the Department of Exact Sciences at ESALQ/USP, the secretaries Solange de Assis Paes Sabadin, Mayara Segatto and Luciane Brajão, and the computer technicians Jorge Alexandre Wiendl and Eduardo Bonilha.

To all those that contributed directly or indirectly to this work thanks so much.

SUMMARY

Extensões do normal distribuição utilizando a família odd log-logística: teoria e aplicações

A distribuição normal é uma das mais importantes na área de estatística. Porém, não é adequada para ajustar dados que apresentam características de assimetria ou de bimodalidade, uma vez que tal distribuição possui apenas os dois primeiros momentos, diferentes de zero, ou seja, a média e o desvio-padrão. Por isso, muitos estudos são realizados com a finalidade de criar novas famílias de distribuições que possam modelar ou a assimetria ou a curtose ou a bimodalidade dos dados. Neste sentido, é importante que estas novas distribuições tenham boas propriedades matemáticas e, também, a distribuição normal como um submodelo. Porém, ainda, são poucas as classes de distribuições que incluem a distribuição normal como um modelo encaixado. Dentre essas propostas destacam-se: a skew-normal, a beta-normal, a Kumarassuamy-normal e a gama-normal. Em 2013 foi proposta a nova família X de distribuições Odd log-logística-G com o objetivo de criar novas distribuições de probabildade. Assim, utilizando as distribuições normal e a skew-normal como função base foram propostas três novas distribuições e um quarto estudo com dados longitudinais. A primeira, foi a distribuição Odd log-logística normal: teoria e aplicações em dados de ensaios experimentais; a segunda foi a distribuição Odd log-logística t Student: teoria e aplicações; a terceira foi a distribuição Odd log-logística skew-bimodal com aplicações em dados de ensaios experimentais e o quarto estudo foi o modelo de regressão com efeito aleatório para a distribuição distribuição Odd log-logística skew-bimodal: uma aplicação em dados longitudinais. Estas distribuições apresentam boas propriedades tais como: assimetria, curtose e bimodalidade. Algumas delas foram demonstradas como: simetria, função quantílica, algumas expansões, os momentos incompletos ordinários, desvios médios e a função geradora de momentos. A flexibilidade das novas distrições foram comparada com os modelos: skew-normal, beta-normal, Kumarassuamy-normal e gama-normal. A estimativas dos parâmetros dos modelos foram obtidas pelo método da máxima verossimilhança. Nas aplicações foram utilizados modelos de regressão para dados provenientes de delineamentos inteiramente casualizados (DIC) ou delineamentos casualizados em blocos (DBC). Além disso, para os novos modelos, foram realizados estudos de simulação para verificar as propriedades assintóticas das estimativas de parâmetros. Para verificar a presença de valores extremos e a qualidade dos ajustes foram propostos os resíduos quantílicos e a análise de sensibilidade. Portanto, os novos modelos estão fundamentados em propriedades matemáticas, estudos de simulação computacional e com aplicações para dados de delineamentos experimentais. Podem ser utilizados em ensaios inteiramente casualizados ou em blocos casualizados, principalmente, com dados que apresentem evidências de assimetria, curtose e bimodalidade.

Palavras-chave: Famílias de distribuições; Extensões da distribuição normal; Verossimilhança; Função quantílica; Função geradora; Modelo de regressão; Delineamentos experimentais; Estudos de simulação; Resíduos quantílicos

ABSTRACT

Extensions of the normal distribution using the odd family logistics log: theory and applications

In this study we propose three new distributions and a study with longitudinal data. The first was the Odd log-logistic normal distribution: theory and applications in analysis of experiments, the second was Odd log-logistic t Student: theory and applications, the third was the Odd log-logistic skew normal: the new distribution skew-bimodal with applications in analysis of experiments and the fourth regression model with random effect of the Odd log-logistic skew normal distribution: an application in longitudinal data. Some have been demonstrated such as symmetry, quantile function, some expansions, ordinary incomplete moments, mean deviation and the moment generating function. The estimation of the model parameters were approached by the method of maximum likelihood. In applications were used regression models to data from a completely randomized design (CRD) or designs completely randomized in blocks (DBC). Thus, the models can be used in practical situations for as a completely randomized designs or completely randomized blocks designs, mainly, with evidence of asymmetry, kurtosis and bimodality.

Keywords: Families distributions; Extensions of the normal distribution; Quantile function; Generating function; Regression model; Likelihood; Experimental designs; Simulation studies; Quartiles residuals

1 INTRODUCTION

Experimental statistics aims to study experimental data, contributing to scientific research from the planning to interpretation of the results (BANZATTO; KRONKA, 2006). This makes the statistical analysis a key part of engineering, chemical and physical sciences, because these techniques are used in almost all research stages. In this context, many studies have been carried out of experimental statistics to improve the analysis and interpretations of results (MONTGOMERY, 1997).

To carry out an experiment, control conditions should be planned to allow for significant comparisons of treatments and check if the effect of the independent variable is responding to the research problem. Therefore, it is extremely important that the experimenter to know the characteristics, the possibilities and the limits of experimental designs, mainly the experimental tests they want to use in their research (JOHNSTON; PENNYPACKER, 1993).

The randomized designs were used when the variability between the experimental units is very small, that is, practically non-existent. This design is used where the experimental conditions can be well controlled (FISHER, 1928). The major disadvantage of this design is the existing variability can inflate the experimental error other than, the variation source due to the treatment effects, which can compromise the inferences of the research (SERLE; CASELLA; McCULLOCH, 2009).

The normal distribution is the most used in the field of experimentation (SERLE; CASELLA; McCULLOCH, 2009). However, this distribution does not fit data with asymmetries well, since only the first two moments are different from zero, namely the mean and standard deviation (MOOD; BOES, 1974). In addition, some assumptions need to be satisfied, such as homoscedasticity, normality of residuals, and independence of errors, to check the quality of the fit.

However, many phenomena do not meet these assumptions. In these cases, asymmetric distributions are a good option, since they can better fit the data by modeling asymmetry, kurtosis and possibly extreme observations (CAMPOS, 2011). Many studies in the literature have proposed distributions that are more flexible to asymmetry or the presence of outliers (AZZALINI; CAPITANIO, 1999), (DICICCIO; MONTE, 2004) and

(GENTON, 2004).

Many studies have been conducted to create new distributions that accommodate asymmetry, kurtosis and bimodality. In this sense, it is important for these new classes of distributions to have good mathematical properties and also to be able to use the properties of the normal distribution, meaning that the new classes have the normal distribution as an embedded model (ALZAATREH; LEE; FAMOYE, 2013).

However, few classes of distributions include the normal distribution as a sub-model. The literature presents some extensions, such as the skew-normal, beta-normal, Kumaraswamy-normal and gamma-normal (ALZAATREH; FAMOYE; LEE, 2014).

These asymmetric distributions have been used in various areas of knowledge, such as the skew-normal in the field of agricultural insurance, the beta-normal and Kumaraswamy-normal in survival analysis, and more recently the gamma-normal in testing ultimate tensile strength of carbon fibers.

The first extension of the normal model proposed was the family of skew-normal skew distributions (AZZALINI, 1985). Several authors studied this model independently (ROBERTS, 1966), (O'HAGAN; LEONARD, 1976) and (AIGNER; LOVELL; SCHMIDT, 1977).

However, Azzalini formally presented the skew-normal distribution and demonstrated some of its properties. This new distribution class is an extension of the normal distribution through an additional parameter that makes it asymmetric both to the right and to the left.

Eugene et al. (2002) proposed a new method to generalize distributions, which became known as the widespread beta (beta - G). The authors used this method and introduced the beta-normal distribution. One advantage of the new distribution in relation to the normal distribution is its flexibility, since it can assume both unimodal and bimodal shapes. (FAMOYE; LEE; EUGENE, 2003) studied properties of the new distribution and (ALZAATREH; LEE; FAMOYE, 2013) and (GUPTA; NADARAJAH, 2004), (GUPTA; NADARAJAH, 2004) developed some mathematical formulas for the moments.

Cordeiro and Castro (2011) proposed another extension, called the generalized Kumaraswamy distribution (Kw-G) class. This family is based on generalization of the

Kumaraswamy distribution proposed by (KUMARASWAMY, 1980). Using the new class of models, the authors considered the normal distribution as a basis function and created the Kumaraswamy-normal. An advantage of this distribution in relation to the beta-normal model is that its density does not depend on a numerical function, while the beta-normal model has this dependence.

Other studies have considered this distribution. For example, (CORDEIRO; CASTRO, 2011) proposed the Kw-gamma, Kw-Gumbel, Kw-inverse Gaussian, Kw-Weibull and Kw-normal distributions, while (SANTANA, 2010) proposed the log-logistic Kumaraswamy and Kumaraswamy logistic distributions.

Recently, Alzaatreh, Famoye and Lee (2013) presented the gamma-normal distribution. The authors proposed a new method to generate continuous distributions called the T-X family of distributions. They used this method to generate the new gamma-normal distribution with the normal distribution as a submodel (ALZAATREH; LEE; FAMOYE, 2013).

Despite the existence of these distributions, very few studies have been conducted on experimental statistics. Analysis involving experimental tests usually present problems at the time of residual analysis, mainly with extreme values. This often occurs when the assumptions required to perform analysis of variance are violated.

One of the ways suggested in the literature to solve problems of lack of normality is data transformation. However, many times a problem and normality is solved, only to create a problem of heteroscedasticity, thus violating one of the main assumptions of ANAVA. In addition, the response variable is used in a new scale, one in which the researcher often has no interest (MONTGOMERY, 1997) and (BANZATTO; KRONKA, 2006).

A major problem in experimental trials can be caused by small sample sizes, such as only four, five or six repetitions. Therefore, when a problem of lack of fit occurs, it becomes difficult to identify the actual cause. The adjustment may require a more flexible distribution, namely platykurtic, leptokurtic, asymmetrical to the left or right or bimodal. We believe that a distribution with heavier tails can solve the problem.

Therefore, this study was carried out to adjust models that are more flexible

than the normal distribution for data from experimental designs, especially completely randomized designs or completely randomized block designs. Three new distributions are proposed: Odd log-logistic normal, Odd log-logistic t Student and Odd log-logistic skew normal. The symmetrical and asymmetrical models are adjusted and their results are compared with the normal submodel.

These new models can be used for settings that require heavy-tailed distributions or symmetric bimodality. Some structural properties are demonstrated, including symmetry, quantile function and expansions for the density function and the distribution. For the Odd log- logistic t Student and Odd log-logistic skew-normal, we also present analysis of residuals.

For the Odd log-logistic skew-normal model, a fourth study is presented with random effects regression models for longitudinal data. The techniques of polynomials and orthogonal contrasts are used to compare the effects of treatments. The adaptive Gauss-Hermite method is used for numerical integration. The adjustments, a simulation study and analysis of residuals are presented.

One problem worth noting is the difficulty of practical interpretation of the settings of these tests for researchers. The normal distribution and multiple comparison tests are reported in academic papers, developed very intuitively, while the interpretation of estimated values of the model parameters requires knowledge, especially in the field of linear models.

This thesis is organized into chapters. The first chapter refers to The odd log-logistic normal distribution: theory and applications in analysis of experiments. The second covers The odd log-logistic Student t distribution: theory and applications. The third discusses The new distribution skew-bimodal with applications in analysis of experiments. The fourth chapter refers to the Regression model with random effect of the odd log-logistic skew-normal distribution: an application in longitudinal data.

Density graphs are presented for each distribution to demonstrate the flexibility of the new models. Adjustments to actual research data are also emphasized, primarily, experimental trials. To make the studies more comprehensive, mathematical demonstrations, simulation studies, applications and analysis of residuals were performed.

To estimate the parameters of the regression models studied, the maximum likelihood function logarithm is used and optimized through numerical methods (L-BFGS-B and Nelder-Mead) implemented in the R software. Hypothesis tests are also presented and confidence intervals are constructed from the asymptotic distribution of the maximum likelihood estimators to compare the new models in relation to sub-models and compare treatments.

Measurements of influence diagnostics are used to detect influential observations in the proposed models. An analysis of local influence, as proposed by (COOK, 1986), is presented and the results are evaluated according to (LABRA *et al.*, 2012) based on (POON; POON, 1999). This analysis is performed for theory and applications in analysis of experiments with heavy-tailed distribution. The odd log-logistic skew-normal distribution: theory and applications in the analysis of experiments and Regression model with random effect of the odd log-logistic skew normal distribution: application of orthogonal contrasts in repeated measurements.

In addition, to check for deviation from the assumptions made for the model and to detect atypical points, a residual analysis using quantile residuals is conducted in accordance with (DUNN; SMYTH, 1996). A study is also carried out by Monte Carlo simulation to identify the empirical distribution of the quantile residual for the regression model, to use the confidence bands constructed from the generation of simulated envelopes (ATKINSON, 1985). This analysis is important to check the quality of fit of the proposed regression model.

This study aims to propose new extensions of the normal model with applications to data obtained mainly from experiments with different designs. The first two extensions, odd log- logistic normal and odd log-logistic t Student, are symmetrical and bimodal, while the odd log-logistic skew-normal has an extra property: asymmetry. In addition, regression models are studied with random effects by means of orthogonal polynomials and contrasts to the third distribution are proposed. After some preliminary studies, the new distributions showed good results and can be used when sample data have kurtosis, symmetric bimodality or asymmetric bimodality.

References

AIGNER, D.; LOVELL, C.; SCHMIDT, P. (1977). Formulation and estimation of stochastic frontier production function model, **Journal of Econometrics**, 6.1: 21-37

ALZAATREH, A.; LEE, C.; FAMOYE, F. (2013). A new method for generating families of continuous distributions. **Metron 71**: 1, 63-79.

ALZAATREH, A.; FAMOYE, F.; LEE, C. (2014). The gamma-normal distribuition: Properties and applications. **Computational Statistics and Data Analysis**. New York, v.69, n.4, p.67-80.

AZZALINI, A.; CAPITANIO, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. **Journal of the Royal Statistical Society**: Series B (Statistical Methodology), 65.2: 367-389.

ATKINSON, A. C. (1985). **Plots, transformations and regression: an introduction to graphical methods of diagnostics regression analysis**. 2nd ed. Oxford: Clarendon Press, 282p.

AZZALINI, A. (1985). A class of distributions which includes the normal ones. **Scandinavian Journal of Statistics**, **12**, 171-178.

KRONKA, S. N.; BANZATTO, A. D. (2006). **Experimentação agrícola**. Jaboticabal: FUNEP, 249 p.

CAMPOS, A. M. **Uma abordagem bayesiana para alguns modelos de crescimento na presença de assimetria e heterocedasticidade**. (2011). Dissertação (Mestrado) - Instituto de Ciências Matemáticas e de Computação - USP, São Carlos. 63 p.

CORDEIRO, G .M.; de CASTRO, M. A. (2011). New family of generalized distribution. **Journal of Statistical Computation and Simulation**, London, v. 81, p. 883-893.

COOK, R.D. (1986). Assessment of local influence (with discussion). **Journal of the Royal Statistical Society B 48**, 133-169.

DICICCIO T. J.; MONTI, A. C.(2004). Inferential aspects of the skew exponential power distribution. **Journal of the American Statistical Association**, 99.466: 439-450.

DUNN. P. K.; SMYTH G. K. (1996). Randomized quantile residuals. **Journal Computational and Graphical Statistics**, New York, v.5, p. 236-244.

EUGENE, N.; LEE, C.; FAMOYE, F. (2002). Beta-normal distribution and its applications. **Communications in Statistics - Theory and Methods**, **31**, 497-512.

FAMOYE, F.; LEE, C.; EUGENE, N. (2003). Beta-normal distribution: Bimodality properties and application. **Journal of Modern Applied Statisticas Methods**, Wayne, v.2, p.314-326.

FISHER, R. A. (1928). The general sampling distribution of the multiple correlation coefficient. **Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character**, **121** (788), 654-673.

GENTON, M. G. (2004). **Skew-elliptical distributions and their applications: a journey beyond normality**. CRC Press.

JOHNSTON J. M.; PENNYPACKER, H. S. (1993). **Strategies and tactics of behavioral research**. 2. ed. Hillsdale: New Jersey.

KUMARASWAMY, P. (1980). Generalized probability density function for double bounded random processes, **Jornal of Hydrollogy**, v.462, p.79-88.

LABRA, F. V; GARAY, A. M.; LACHOS, V. H.; ORTEGA, E. M. M. (2012). Estimation and diagnostics for heteroscedastic nonlinear regression models based on scale mixtures of skew-normal distributions. **Journal of Statistical Planning and Inference**, **142**, 2149-2165.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. (1974). **Introduction to the Theory of Statistics**.

MONTEGOMERY, D. C. (1997). **Design and Analysis of Experiments**. John Wiley and Sons: New York.

NADARAJAH, S.; GUPTA, A. K. (2004). **The beta Fréchet distribution. Far East Journal of Theretica Statistics**. New Delhi, v.15, p. 15-24.

O'HAGAN, A.; LEONARD, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. **Biometrika**, 63: 201-202.

POON, W.; POON, Y. S. (1999). Conformal normal curvature and assessment of local influence. **Journal of the Royal Statistical Society**, Series B **61**, 51-61.

ROBERTS, C. (1966). A correlation model useful in the study of twins, **Journal of the American Statistical Society 61**: 1184-1190.

SANTANA, T. V. F. (2010). **As distribuições Kumaraswamy log-logística e Kumaraswamy logística**. 103. Dissertação (Mestrado em Estatística e ExperimentaçãoAgronômica - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba.

SERLE, S. R.; CASELLA, G.; McCULLOCH, C. E. (2009). Variance componentes (Vol. 391). John Wiley & Sons.

2 THE ODD LOG-LOGISTIC NORMAL DISTRIBUTION: THEORY AND APPLICATION IN ANALYSIS OF EXPERIMENTS

Abstract

Providing a new distribution is always precious for statisticians. A new three-parameter model called the odd log-logistic normal (OLLN) distribution is defined and studied. The distribution is symmetric, maybe platykurtic or leptokurtic and unimodal or bimodal. Various structural properties are derived including explicit expressions for the ordinary and incomplete moments, generating function and mean deviations. We use maximum likelihood to estimate the parameters of the new model. It was proposed regression models to completely randomized designs and completely randomized block designs. We show that the proposed distribution is a very competitive model to other classical models by means of three real data sets with one example in analysis of experiments. Simulation studies were conducted in order to verify the quality of the maximum likelihood estimates. We can conclude that this model can give more realistic fits than other special regression models.

Keywords: Generating function; Log-logistic distribution; Maximum likelihood estimation; Mean deviation; Normal distribution; Regression model

2.1 Introduction

In Statistics, the normal distribution is the widely used model in applications to real data. When the number of observations is large, it can serve as a good approximation to other models. The probability density function (pdf) and cumulative distribution function (cdf) of the normal (for $x \in \mathbb{R}$) model are given by

$$g(x;\mu,\sigma) = \frac{1}{\sqrt{2\,\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} = \frac{1}{\sigma}\,\phi\left(\frac{x-\mu}{\sigma}\right), \tag{2.1}$$

and

$$G(x;\mu,\sigma) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{2}\left[1+\operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right], \tag{2.2}$$

where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively, and $\operatorname{erf}(\cdot)$ is the error function.

Alzaatreh *et al.* (2013) proposed the T-X family of distributions. Let $r(t)$ be the pdf of a random variable $T \in [a,b]$ for $-\infty < a < b < \infty$ and let $W[G(X)]$ be a function

of the cdf of a random variable X such that $W[G(X)]$ satisfies the following conditions:

i) $W[G(X)] \in [a, b]$;

ii) $W[G(X)]$ is differentiable and monotonically non-decreasing;

iii) $W[G(X)] \to a$ as $x \to -\infty$ and $W[G(X)] \to b$ as $x \to \infty$.

The T-X family is defined by

$$F(x) = \int_a^{W[G(x)]} r(t)dt, \tag{2.3}$$

where $W[G(X)]$ satisfies the conditions i) ii) and iii). The pdf corresponding to (2.3) is given by

$$f(x) = \left\{ \frac{d}{dx} W[G(x)] \right\} r\left(W[G(x)]\right).$$

In this chapter, we propose a new extended normal distribution with heavier tails called the odd log-logistic normal ("OLLN" for short) model. The goal is to work with smaller samples from experimental designs, e.g., completely randomized design (CRD) or completely randomized block design (CRBD). We use $W[G(X)] = \Phi(x; \mu, \sigma)/\bar{\Phi}(x; \mu, \sigma)$ and $r(t) = \alpha t^{\alpha-1}/(1 + t^{\alpha})^2$ $(t > 0)$ in equation (2.3). Thus, the cdf of the OLLN model with an additional shape parameter $\alpha > 0$ is defined by

$$F(x; \mu, \sigma, \alpha) = \int_0^{\frac{\Phi(x;\mu,\sigma)}{\bar{\Phi}(x;\mu,\sigma)}} \frac{\alpha\, t^{\alpha-1}}{(1 + t^{\alpha})^2} dt = \frac{\Phi^{\alpha}\left(\frac{x-\mu}{\sigma}\right)}{\Phi^{\alpha}\left(\frac{x-\mu}{\sigma}\right) + \left[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\alpha}}, \tag{2.4}$$

where $\bar{\Phi}(x; \mu, \sigma) = 1 - \Phi(x; \mu, \sigma)$.

The normal distribution $\Phi(x; \mu, \sigma)$ is clearly a special case of (2.4) when $\alpha = 1$. We note that there is no complicated function in (2.4) in contrast with the beta generalized family, which includes two extra parameters and also involves the beta incomplete function. The OLLN density is given by

$$f(x;\mu,\sigma,\alpha)=\frac{\alpha\phi\left(\frac{x-\mu}{\sigma}\right)\Phi^{\alpha-1}\left(\frac{x-\mu}{\sigma}\right)\left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\alpha-1}}{\sigma\left\{\Phi^{\alpha}\left(\frac{x-\mu}{\sigma}\right)+\left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\alpha}\right\}^{2}}. \tag{2.5}$$

Note that $\alpha > 0$ is a shape parameter. Henceforth, a random variable with density function (2.5) is denoted by $X \sim \text{OLLN}(\mu,\sigma^2,\alpha)$. For $\mu = 0$ and $\sigma = 1$, we obtain the standard OLLN distribution. We can write by omitting the parameters of the cdf

$$\alpha=\frac{\log\left[\frac{F(x;\mu,\sigma,\alpha)}{\bar{F}(x;\mu,\sigma,\alpha)}\right]}{\log\left[\frac{\Phi(x;\mu,\sigma)}{\bar{\Phi}(x;\mu,\sigma)}\right]},$$

where $\bar{\Phi}(x;\mu,\sigma) = 1-\Phi(x;\mu,\sigma)$ and $\bar{F}(x;\mu,\sigma,\alpha) = 1-F(x;\mu,\sigma,\alpha)$. So, the parameter α represents the quotient of the log odds ratio for the generated and baseline distributions.

In next section, we study some structural properties of the OLLN distribution, which generalizes the normal distribution.

2.2 Properties and useful expansions for the OLLN model

In this section, we discuss some of the shapes of the OLLN distribution. It is not possible to study the shapes analytically in terms of the model parameters by taking derivatives. Then, we verify the distribution's bimodality by combining some values of μ, σ and α as shown in Figure 2.1.

Lemma 1. The OLLN (μ,σ^2,α) density is symmetric about μ.

Proof. Let $X \sim \text{OLLN}(\mu,\sigma^2,\alpha)$ and $\Phi(\mu-x)$ = 1 - $\Phi(\mu+x)$, where $\sigma^2 > 0$ and $\alpha > 0$. Then,

$$\begin{aligned} f(\mu-x) &= \frac{\alpha}{\sigma\sqrt{2\pi}}\frac{\Phi^{\alpha-1}(x-\mu)\left[1-\Phi(x-\mu)\right]^{\alpha-1}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}}{\left\{\Phi(x-\mu)^{\alpha}+\left[1-\Phi(x-\mu)\right]^{\alpha}\right\}^{2}} \\ &= \frac{\alpha}{\sigma\sqrt{2\pi}}\frac{\Phi^{\alpha-1}(x+\mu)\left[1-\Phi(x+\mu)\right]^{\alpha-1}\exp\left\{-\frac{x^2}{2\sigma^2}\right\}}{\left\{\Phi(x+\mu)^{\alpha}+\left[1-\Phi(x+\mu)\right]^{\alpha}\right\}^{2}} = f(\mu+x). \end{aligned}$$

So, we prove that the OLLN distribution is symmetric about μ. Them, its skewness is zero. The parameters σ and α characterize the kurtosis and bimodality of this distribution. Plots of the OLLN density function for some parameter values are displayed in Figure 2.1. Based on several plots, we conclude that this density is bimodal when $\alpha \leq 0.5$.

The quantile function (qf) is in widespread use in Statistics. Let $F(x;\mu,\sigma,\alpha) = u$ and $\Phi^{-1}\left(\frac{x-\mu}{\sigma}\right)$ be the inverse function $\Phi\left(\frac{x-\mu}{\sigma}\right)$. Thus, the qf of X is given by

$$Q(u) \;=\; \mu + \sigma\, Q_N\left(\frac{u^{1/\alpha}}{[1-u]^{1/\alpha} + u^{1/\alpha}}\right), \tag{2.6}$$

where $Q_N(u) = \Phi^{-1}(u;\mu,\sigma)$ is the normal qf. The equation (2.6) has tractable properties especially for simulations.

In Figure 2.1, we display some plots of the OLLN density for selected values of μ, σ and α. The six cases provide all forms of the OLLN density. The first three cases and the sixth emphasize the bimodality, whereas the other two reveal weighter tails than the normal ones. Specifically, in the graphic (a), the OLLN density is bimodal for $\mu = 0$ and $\sigma = 1$ and the values of α between 0.1 and 0.5. In the graphic (b), we consider the same values for α and μ but σ varies in $(1.5; 2; 3; 4; 5)$, and the density has more kurtosis. In the graphic (c), we take the same values of σ and α but μ varies in $(-2; -1; 0; 1; 2)$ and, clearly, the location of the density does not depend on the values of σ and α. In the graphic (d), the OLLN density is more leptokurtic, whereas it is platykurtic in the graphic (e). Finally, in the graphic (f), we note bimodality and both leptokurtic and platykurtic forms.

2.2.1 Useful expansions

First, we define the *exponentiated-normal* ("Exp-N"), say $W \sim \text{Exp}^c\left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]$ with power parameter $c > 0$, if W has cdf and pdf given by

$$H_c(x) = \Phi^c\left(\frac{x-\mu}{\sigma}\right) \qquad \text{and} \qquad h_c(x) = c\,\sigma^{-1}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi^{c-1}\left(\frac{x-\mu}{\sigma}\right),$$

respectively. In a general context, the properties of the *exponentiated-G* (Exp-G) distributions have been studied by several authors for some baseline G models, see (MUDHOL-

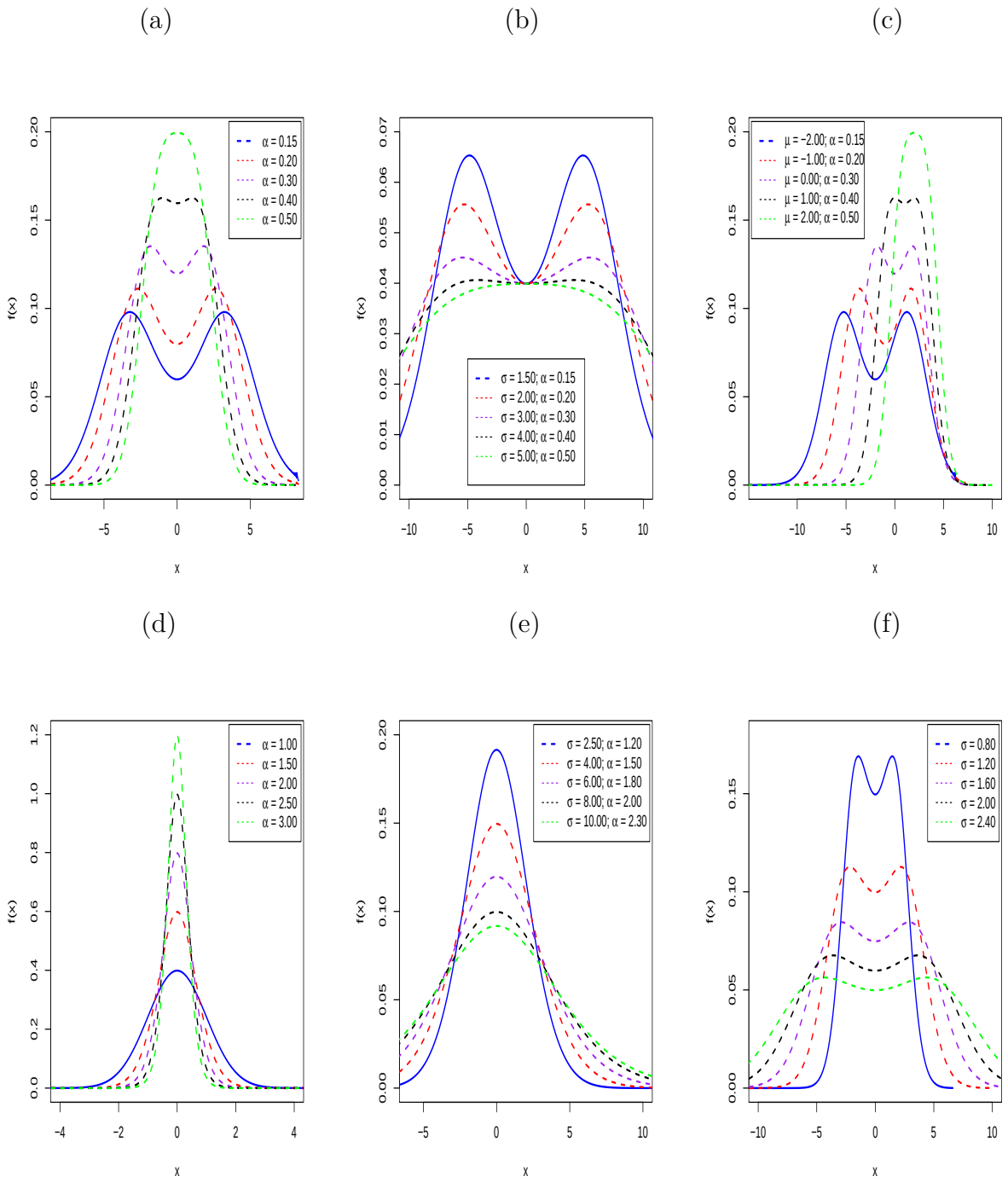


Figura 2.1 - Plots of the OLLN pdf for some parameter values. (a) For different values of $0 < \alpha < 1$ with $\mu = 0$ and $\sigma = 1$. (b) For different values of σ and α with $\mu = 0$. (c) For different values of μ and α with $\sigma = 1$. (d) For different values of $\alpha > 1$ with $\mu = 0$ and $\sigma = 1$. (e) For different values of σ and α with $\mu = 0$. (f) For different values of σ with $\mu = 0$ and $\alpha = 0.30$.

KAR; SRIVASTAVA, 1993) and (MUDHOLKAR; SRIVASTAVA; FRIEMER, 1995) for exponentiated Weibull, (NADARAJAH, 2005) for exponentiated Gumbel, (SHIRKE; KAKADE, 2006) for exponentiated log-normal and (NADARAJAH; GUPTA, 2007) for expo-

nentiated gamma distributions. See, also, (NADARAJAH; KOTZ, 2006), among others.

First, we obtain an expansion for $F(x;\mu,\sigma,\alpha)$ using a power series for $\Phi\left(\frac{x-\mu}{\sigma}\right)^{\alpha}$ ($\alpha > 0$ real)

$$\Phi\left(\frac{x-\mu}{\sigma}\right)^{\alpha} = \sum_{k=0}^{\infty} a_k\, \Phi\left(\frac{x-\mu}{\sigma}\right)^{k}, \tag{2.7}$$

where

$$a_k = a_k(\alpha) = \sum_{j=k}^{\infty} (-1)^{k+j} \binom{\alpha}{j} \binom{j}{k}.$$

For any real $\alpha > 0$, we consider the generalized binomial expansion

$$\left[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{\alpha} = \sum_{k=0}^{\infty} (-1)^{k} \binom{\alpha}{k} \Phi\left(\frac{x-\mu}{\sigma}\right)^{k}. \tag{2.8}$$

Inserting (2.7) and (2.8) in equation (2.4), we obtain

$$F(x;\mu,\sigma,\alpha) = \frac{\sum_{k=0}^{\infty} a_k\, \Phi\left(\frac{x-\mu}{\sigma}\right)^{k}}{\sum_{k=0}^{\infty} b_k\, \Phi\left(\frac{x-\mu}{\sigma}\right)^{k}},$$

where $b_k = a_k + (-1)^k \binom{\alpha}{k}$ for $k \geq 0$.

The ratio of the two power series can be expressed as

$$F(x;\mu,\sigma,\alpha) = \sum_{k=0}^{\infty} c_k\, \Phi\left(\frac{x-\mu}{\sigma}\right)^{k}, \tag{2.9}$$

where $c_0 = a_0/b_0$ and the coefficients c_k's (for $k \geq 1$) are determined from the recurrence equation

$$c_k = b_0^{-1}\left(a_k - \sum_{r=1}^{k} b_r\, c_{k-r}\right).$$

The pdf of X is obtaining by differentiating (2.9) as

$$f(x;\mu,\sigma,\alpha) = \sum_{k=0}^{\infty} c_{k+1}\, h_{k+1}(x), \tag{2.10}$$

where

$$h_{k+1}(x) = \frac{(k+1)}{\sigma}\,\Phi\left(\frac{x-\mu}{\sigma}\right)^k \phi\left(\frac{x-\mu}{\sigma}\right)$$

is the Exp-N density function with power parameter $k+1$.

Equation (2.10) reveals that the OLLN density function is a mixture of Exp-N densities. Thus, some of its structural properties such as the ordinary and incomplete moments and generating function can be obtained from well-established properties of the Exp-N distribution. This equation is the main result of this section.

2.3 Moments and mean deviations

The mathematical results in this section can be evaluated numerically in most symbolic softwares by taking in the sums of a large positive integer value in place of ∞.

2.3.1 Ordinary and central moments

Henceforth, let the random error variable defined as $Z = (X-\mu)/\sigma$ where X has density function given by (2.5), then a random variable Z having the OLLN$(0,1,\alpha)$ distribution. The moments of X having the OLLN(μ,σ,α) distribution are easily determined from the moments of Z by $E(X^n) = E[(\mu+\sigma Z)^n] = \sum_{r=0}^{n}\binom{n}{r}\mu^{n-r}\,\sigma^r\,E(Z^r)$. So, we can work with the standardized random variable Z. We derive two theorems for the nth moment of Z, say $\mu'_n = E(Z^n)$ from (2.10) and the monotone convergence theorem with $\mu = 0$ and $\sigma = 1$. The function (2.11) provides the rth ordinary moment and can be implemented in Mathematica software.

Theorem 1. The nth ordinary moment of Z is given by

$$\mu'_n = \sum_{k=0}^{\infty}(k+1)\,c_{k+1}\int_{-\infty}^{\infty} z^n\,\Phi(z)^k\phi(z)\,dz = \sum_{k=0}^{\infty}(k+1)\,c_{k+1}\,\tau_{n,k}, \tag{2.11}$$

where $\tau_{n,k}$ is expressed in terms of the Lauricella function of type A.

Proof. Let's consider the expression given by $\mu'_n = E(Z^n)$ and the mixture of Exp-N

densities (2.10). So, the equation μ'_n reduces to,

$$\mu'_n = \sum_{k=0}^{\infty}(k+1)\, c_{k+1} \int_{-\infty}^{\infty} z^n\, \Phi(z)^k \phi(z)\, dz. \tag{2.12}$$

This first representation for μ'_n is based of the standard normal distribution defined by $\tau_{n,r} = \int_{-\infty}^{\infty} z^n\, \Phi(z)^r \phi(z)\, dz$ for $n, r \geq 0$.

We define the Lauricella function of type A (EXTON, 1978),

$$F_A^{(n)}(a, b_i; c_i; x_i) \;=\; \sum_{m_1=0}^{\infty} \cdots \sum_{m_n=0}^{\infty} \frac{(a)_{m_1+\cdots+m_n}(b_1)_{m_1}\cdots(b_n)_{m_n}}{(c_1)_{m_1}\cdots(c_n)_{m_n}} \frac{x_1^{m_1}\cdots x_n^{m_n}}{m_1!\cdots m_n!},$$

where $b_i = b_1, \ldots, b_n$, $c_i = c_1, \ldots, c_n$, $x_i = x_1, \ldots, c_n$, $(a)_k = a(a+1)\ldots(a+k-1)$ denotes the Pochhammer symbol, i.e. the kth rising factorial power of a with the convention $(a)_0 = 1$. According with (NADARAJAH, 2008) the function $\tau_{n,r}$ can be expressed in terms of the Lauricella function of type A as,

$$\begin{aligned} \tau_{n,r} \;=\; & 2^{n/2}\pi^{-(r+1/2)} \sum_{\substack{p=0 \\ (n+r-p)\,\text{even}}}^{r} \left(\frac{\pi}{2}\right)^p \binom{r}{p} \Gamma\left(\frac{n+r-p+1}{2}\right) \times \\ & F_A^{(r-p)}\left(\frac{n+r-p+1}{2}; \frac{1}{2}, \ldots, \frac{1}{2}; \frac{3}{2}, \ldots, \frac{3}{2}; -1, \ldots, -1\right). \end{aligned} \tag{2.13}$$

Rewriting (2.13) in equation (2.12) we can conclude (2.11).

Theorem 2. Let's a random variable Z having the OLLN$(0, 1, \alpha)$ distribution. Thus, derive the (n, k)th probability weighted moment (PWM) (for n and k positive integers) of the standard normal distribution. The nth ordinary moment of Z can also be expressed as

$$\mu'_n = \sum_{k,i=0}^{\infty} \frac{(k+1)\, c_{k+1}}{(k+i+1)}. \tag{2.14}$$

Proof. Consider the expression given by $\mu'_n = E[Z^n\, \Phi(z)^k]$ and the mixture of Exp-N densities (2.10). Thus, the equation μ'_n can be written by,

$$\mu'_n = E[Z^n\, \Phi(z)^k] = \sum_{k=0}^{\infty}(k+1)\, c_{k+1} \int_{-\infty}^{\infty} z^n\, \Phi(z)^k \phi(z)\, dz. \tag{2.15}$$

Setting $u = \Phi(z)$ and $Q_N(u) = \Phi^{-1}(u)$ in(2.15), we can write for the moments of Z in terms of the standard normal qf, say $Q_N(u)$, as

$$\mu'_n = \sum_{k=0}^{\infty}(k+1)\, c_{k+1} \int_0^1 Q_N(u)^n\, u^k du. \tag{2.16}$$

Following (STEINBRECHER, 2002), the function $Q_N(u)$ can be expanded as

$$Q_N(u) = \sum_{k=0}^{\infty} p_k \left[\sqrt{2\pi}\,(u - 1/2)\right]^{2k+1}, \tag{2.17}$$

where the p_k's (for $k \geq 0$) are given by

$$p_{k+1} = \frac{1}{2(2k+3)} \sum_{r=0}^{k} \frac{(2r+1)\,(2k-2r+1)\, p_r\, p_{k-r}}{(r+1)(2r+1)}.$$

We have $p_0 = 1$, $p_1 = 1/6$, $p_2 = 7/120$, $p_3 = 127/7560, \ldots$

By expanding the binomial term and changing the sums in (2.17), we obtain

$$Q_N(u) = \sum_{i=0}^{\infty} s_i\, u^i, \tag{2.18}$$

where the coefficients are

$$s_i = \sum_{m=\delta_i}^{\infty} \frac{(-1)^{2m+1-i}\, \pi^{m+1/2}\, p_m}{2^{m-i+1/2}} \binom{2m+1}{i},$$

where $\delta_{2i} = \delta_{2i+1} = i$ for $i \geq 0$.

Combining (2.16) and (2.18), we can rewrite μ'_n as

$$\mu'_n = \sum_{k=0}^{\infty}(k+1)\, c_{k+1} \int_0^1 u^k \left(\sum_{i=0}^{\infty} s_i\, u^i\right)^n du. \tag{2.19}$$

We use an equation by (GRADSHTEYN; RYZHIK, 2007, Section 0.314) for a power series raised to a positive integer n

$$Q_N(u)^n = \left(\sum_{i=0}^{\infty} s_i\, u^i\right)^n = \sum_{i=0}^{\infty} q_{n,i}\, u^i, \tag{2.20}$$

where the coefficients $q_{n,i}$ (for $i \geq 1$) are obtained from the recurrence equation

$$q_{n,i} = (i\, s_0)^{-1} \sum_{m=1}^{i} [m(n+1) - i]\, s_m\, q_{n,i-m}, \tag{2.21}$$

and $q_{n,0} = s_0^n$. The coefficient $q_{n,i}$ can be determined from $q_{n,0}, \ldots, q_{n,i-1}$ and hence from the quantities $s_0, \ldots, s_i$. Equations (2.20) and (2.21) are used throughout this chapter. The coefficient $q_{n,i}$ can be given explicitly in terms of the coefficients s_i, although it is not necessary for programming numerically our expansions in any algebraic or numerical software.

Rewriting (2.19) using (2.20) and (2.21) and then integrating.

$$\mu'_n = \sum_{k,i=0}^{\infty} \frac{(k+1)\, c_{k+1}}{(k+i+1)}. \tag{2.22}$$

The theorems 1 and 2 are the main results of this section.

The central moments (μ_r) and cumulants (κ_r) of Z can be determined as

$$\mu_r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} \mu_1'^k\, \mu'_{r-k} \qquad \text{and} \qquad \kappa_r = \mu'_r - \sum_{k=1}^{r-1} \binom{r-1}{k-1} \kappa_k\, \mu'_{r-k},$$

respectively, where $\kappa_1 = \mu'_1$. The kurtosis of X is given by $\gamma_2 = \kappa_4/\kappa_2^2$ in terms of the second and fourth cumulants. The cumulants of Z are obtained from those of Z by $\kappa_{Z,1} = \mu + \sigma\kappa_1$ and $\kappa_{Z,r} = \sigma^r\, \kappa_r$ for $r \geq 2$.

The effects of the additional shape parameter α on the kurtosis of X can be based on quantile measures given by (2.6). The shortcomings of the classical kurtosis measure are well-known. One of the earliest kurtosis measures is the Moors' kurtosis given by

$$K = \frac{Q\,(7/8) - Q\,(5/8) + Q\,(3/8) - Q\,(1/8)}{Q\,(6/8) - Q\,(2/8)}. \tag{2.23}$$

The statistic K is less sensitive to outliers and it exists even for distributions without moments. Plots of the kurtosis of X for selected parameter values are displayed in Figures 2.2 and 2.3.

(a) (b)

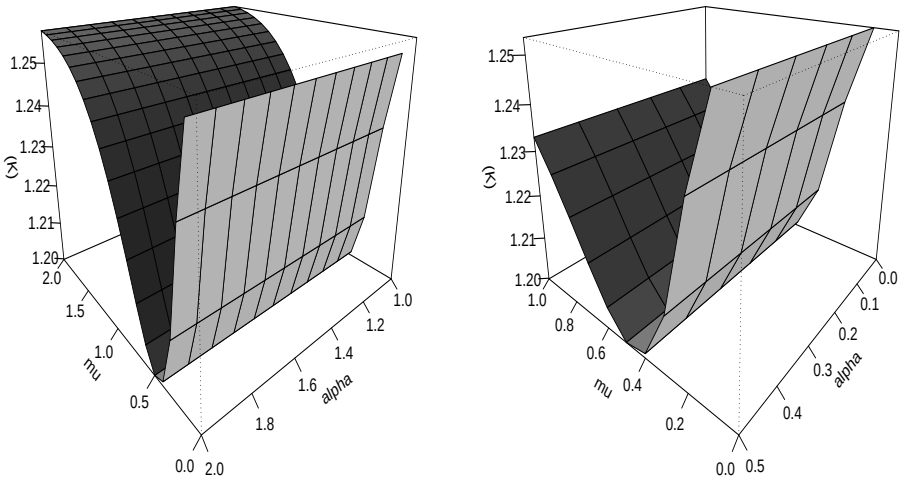


Figura 2.2 - Moors' kurtosis for the OLLN distribution. Plots (a) as functions of $\mu \in [0, 2]$ with $\alpha \in [1, 2]$ and (b) as functions of $\mu \in [0, 1]$ with $\alpha \in [0, 0.5]$.

(a) (b)

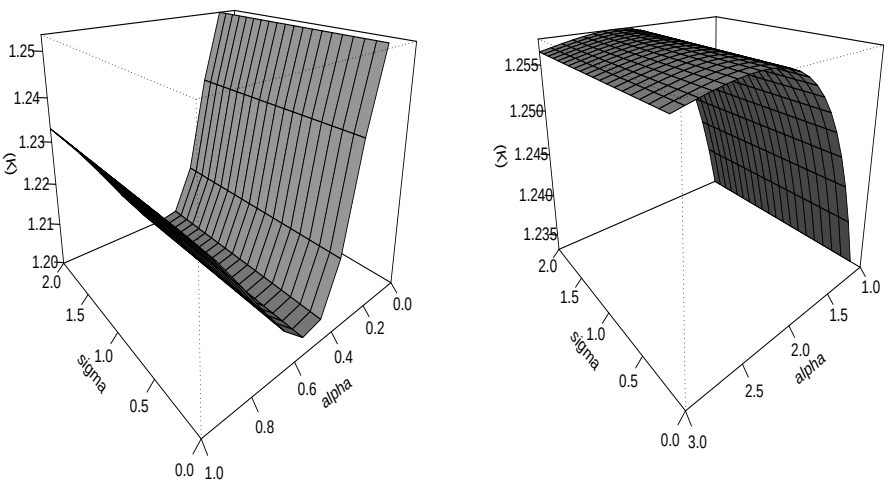


Figura 2.3 - Moors' kurtosis for the OLLN distribution. Plots (a) as functions of $\sigma \in [0, 2]$ with $\alpha \in [0, 1]$ and (b) as functions of $\sigma \in [0, 2]$ with $\alpha \in [1, 3]$.

2.3.2 Incomplete moments

We derive two theorems for the nth incomplete moment of Z given by $m_{Z,n}(y) = \int_0^y z^n f(z)dz$. This theorems are based on the mixture form (2.10) and the monotone

convergence theorem, with $\mu = 0$ and $\sigma = 1$.

Theorem 3. The nth incomplete moment of Z is given by

$$m_{Z,n}(y) = \sum_{j=0}^{\infty} t_j \, A(n+j, y), \tag{2.24}$$

where $A(n+j, y) = \int_{-\infty}^{y} z^{n+j} \, \mathrm{e}^{-\mathrm{z}^2/2} \, \mathrm{dz}$ and $t_j = (\sqrt{2\,\pi})^{-1} \sum_{k=0}^{\infty} (k+1) \, c_{k+1} \, f_{k,j}$.

Proof. Let's consider the expression given by $m_{Z,n}(y) = E(Z^n | Z < y)$ and the mixture of Exp-N densities (2.10). The expression of $m_{Z,n}(y)$ is given by,

$$m_{Z,n}(y) = \sum_{k=0}^{\infty} (k+1) \, c_{k+1} \int_{-\infty}^{y} z^n \, \Phi(z)^k \, \phi(z) \, dz. \tag{2.25}$$

We can write $\Phi(x)$ as a power series $\Phi(z) = \sum_{j=0}^{\infty} e_j \, z^j$, where $e_0 = (1 + \sqrt{2/\pi})^{-1}/2$, $e_{2j+1} = (-1)^j / [\sqrt{2\pi} \, 2^j (2j+1) j!]$ for $j = 0, 1, 2 \ldots$ and $e_{2j} = 0$ for $j = 1, 2, \ldots$. Based on (2.20), we have

$$\Phi(z)^k = \left(\sum_{j=0}^{\infty} e_j \, z^j \right)^k = \sum_{j=0}^{\infty} f_{k,j} \, z^j, \tag{2.26}$$

where the coefficients $f_{k,j}$ (for $j \geq 1$) are determined from (2.21) as $f_{k,j} = (j \, e_0)^{-1} \sum_{m=1}^{j} [m \, (k+1) - i] \;\; e_m \, f_{k,j-m}$ and $f_{k,0} = e_0^k$.

Finally, using (2.26) and changing variable in equation (2.25), the Theorem 3 follows.

$$m_{Z,n}(y) = \sum_{j=0}^{\infty} t_j \, A(n+j, y), \tag{2.27}$$

where $A(n+j, y) = \int_{-\infty}^{y} z^{n+j} \, \mathrm{e}^{-\mathrm{z}^2/2} \, \mathrm{dz}$ and $t_j = (\sqrt{2\,\pi})^{-1} \sum_{k=0}^{\infty} (k+1) \, c_{k+1} \, f_{k,j}$.

We can determine the integral $A(k, y) = \int_{-\infty}^{y} z^k \, \mathrm{e}^{-\mathrm{z}^2/2} \, \mathrm{dz}$ depending if $y < 0$ and $y > 0$. We define

$$G(k) = \int_0^{\infty} z^k \; \mathrm{e}^{-\frac{\mathrm{z}^2}{2}} \mathrm{dz} = 2^{(\mathrm{k}-1)/2} \, \Gamma \left(\frac{\mathrm{k}+1}{2} \right).$$

We consider the confluent hypergeometric function ${}_1F_1(a;b;z) = \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k} \frac{z^k}{k!}$, the hypergeometric function ${}_2F_1(a,b;c;z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!}$ and the Kummer's function given by $U(a,b;z) = z^{-a}\, {}_2F_0(a, 1+a-b; -z^{-1})$. All these functions can be easily computed using Mathematica. See $http://mathworld.wolfram.com/ConfluentHypergeometricFunctionoftheFirstKind.html$.

We can easily prove that $A(k,y) = (-1)^k\, G(k) + (-1)^{k+1}\, H(k,y)$ for $y < 0$ and $A(k,y) = (-1)^k G(k) + H(k,y)$ for $y > 0$, where $H(k,y) = \int_0^y z^k\, \mathrm{e}^{-\mathrm{z}^2/2}\, \mathrm{dz}$ (WHITTAKER; WATSON, 1990) is given by

$$\begin{aligned} H(k,y) &= \frac{2^{k/4+1/4}\, y^{k/2+1/2}\, \mathrm{e}^{-\mathrm{y}^2/4}}{(k/2+1/2)(k+3)}\, N_{k/4+1/4,k/4+3/4}(y^2/2) \\ &\quad + \frac{2^{k/4+1/4}\, y^{k/2-3/2}\, \mathrm{e}^{-\mathrm{y}^2/4}}{k/2+1/2}\, N_{k/4+5/4,k/4+3/4}(y^2/2) \end{aligned}$$

and $N_{k,m}(z)$ is the Whittaker function (WHITTAKER; WATSON, 1990, pp. 339-351) given by

$$N_{k,m}(z) = M_{k,m}(z) = \mathrm{e}^{-z/2}\, z^{m+1/2}\, {}_1F_1\left(\frac{1}{2} + m - k, 1 + 2m; z\right)$$

or

$$N_{k,m}(z) = W_{k,m}(z) = \mathrm{e}^{-z/2}\, z^{m+1/2}\, U\left(\frac{1}{2} + m - k, 1 + 2m; z\right).$$

These functions are implemented in Mathematica as WhittakerM$[k,m,x]$ and WhittakerW$[k,m,x]$, respectively.

Theorem 4. The nth incomplete moment of Z can also be expressed as

$$m_{Z,n}(y) = \sum_{k,i=0}^{\infty} (k+1)\, c_{k+1}\, q_{n,i}\, \frac{\Phi(y)^{i+k+1}}{(i+k+1)}. \tag{2.28}$$

Proof. A second representation for $m_{Z,n}(y)$ is based on the normal qf. Thus, we can rewritten equation (2.25) as

$$m_{Z,n}(y) = \sum_{k=0}^{\infty} (k+1)\, c_{k+1} \int_0^{\Phi(y)} Q_N(u)^n\, u^k du. \tag{2.29}$$

Inserting (2.20) in (2.29) we obtain

$$m_{Z,n}(y) = \sum_{k=0}^{\infty}(k+1)\, c_{k+1} \sum_{i=0}^{\infty} q_{n,i} \int_0^{\Phi(y)} u^{i+k} du. \tag{2.30}$$

Resolving the integral and after some algebra we obtain (2.28).

The nth incomplete moment of Z follows after a binomial expansion

$$m_{X,n}(y) = \sum_{k=0}^{n} \mu'_{n-k}\, \sigma^k \binom{n}{k} m_{Z,k}\left(\frac{y-\mu}{\sigma}\right).$$

The theorems (3) and (4) are the main results of this subsection.

2.3.3 Moment generating function

The mgf $M(-t) = E(\mathrm{e}^{-\mathrm{t}\,\mathrm{Z}})$ of Z can be expressed from (2.10) as

$$M(-t) = \frac{1}{\sqrt{2\pi}} \sum_{k=0}^{\infty} c_{k+1} \int_{-\infty}^{\infty} \exp\left(-tz - \frac{z^2}{2}\right) \Phi(z)^k\, dz.$$

We use (2.26) and write

$$M(-t) = \frac{1}{\sqrt{2\pi}} \sum_{k,j=0}^{\infty} f_{k,j}\, c_{k+1} \int_{-\infty}^{\infty} z^j \exp\left(-tz - \frac{z^2}{2}\right) dz.$$

Based on a result by (PRUDNIKOV; BRYCHKOV; MARICHEV, 1986, equation 2.3.15.8), we obtain

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^j \exp\left(-tz - \frac{z^2}{2}\right) dz = (-1)^j \frac{\partial^j}{\partial t^j}\left(\mathrm{e}^{\mathrm{t}^2/2}\right).$$

Thus, the mgf of Z becomes

$$M(-t) = \sum_{j=0}^{\infty} v_j \frac{\partial^j}{\partial t^j}\left(\mathrm{e}^{\mathrm{t}^2/2}\right), \tag{2.31}$$

where $v_j = (-1)^j \sum_{k=0}^{\infty} f_{k,j}\, c_{k+1}$

A second representation for $M(t)$ can be based on the power of the qf $Q_N(u)$

given by equation (2.20). By expanding the exponential function gives

$$M(t) = \int_0^1 \exp\left[t\, Q_N(u)\right]\, du = \sum_{j=0}^{\infty} \int_0^1 \frac{t^j\, \left(\sum_{i=0}^{\infty} s_i\, u^i\right)^j}{j!}\, du$$

and then using (2.20), we obtain

$$M(t) = \sum_{i,j=0}^{\infty} \frac{q_{j,i}}{j!(i+1)}\, t^j, \tag{2.32}$$

where the quantities $q_{j,i}$ are given by (2.21).

Equations (2.31) and (2.32) are the main results of this section. Equation (2.32) provides an alternative way to obtain the nth ordinary moment of Z as $\mu'_n = \sum_{i=0}^{\infty} q_{n,i}/(i+1)$ for $n \geq 1$.

2.3.4 Mean deviations

We can derive the mean deviations of Z about the mean μ'_1 and about the median M in terms of its first incomplete moment. They can be expressed as

$$\delta_1 = 2\Big[\mu'_1\, F(\mu'_1) - M_{Z,1}(\mu'_1)\Big] \qquad \text{and} \qquad \delta_2 = \mu'_1 - 2m_{Z,1}(M), \tag{2.33}$$

where $\mu'_1 = E(Z)$ and $m_{Z,1}(q) = \int_{-\infty}^{q} x\, f(x)\, dx$. The quantity $m_{Z,1}(q)$ is determined from (2.27) or (2.28) with $n = 1$ and the measures δ_1 and δ_2 are evaluated by setting $q = \mu'_1$ and $q = M$, respectively.

The Bonferroni and Lorenz curves of Z are defined by $B(\pi) = m_{Z,1}(q)/(\pi\mu'_1)$ and $L(\pi) = m_{Z,1}(q)/\mu'_1$, respectively, where $q = Q_Z(\pi)$ is straightforward obtained from the qf (2.6) for a given probability π.

2.4 Maximum likelihood estimation

Let $x_1, \ldots, x_n$ be a random sample of size n from the OLLN(α, μ, σ) distribution. In this section, we determine the maximum likelihood estimates (MLEs) of the model parameters from complete samples only. The log-likelihood function for the vector of

parameters $\boldsymbol{\theta} = (\alpha, \mu, \sigma)^T$ is given by

$$\begin{aligned} l(\boldsymbol{\theta}) &= n[\log(\alpha) - \log(\sigma)] + \sum_{i=1}^{n} \log[\phi(z_i)] + \alpha \sum_{i=1}^{n} \log\{\Phi(z_i)[1 - \Phi(z_i)]\} \\ &\quad - \sum_{i=1}^{n} \log\{\Phi(z_i)[1 - \Phi(z_i)]\} - 2\sum_{i=1}^{n} \log\{\Phi^{\alpha}(z_i) + [1 - \Phi(z_i)]^{\alpha}\}, \end{aligned} \tag{2.34}$$

where $z_i = (x_i - \mu)/\sigma$.

The components of the score vector $U(\boldsymbol{\theta})$ are given by

$$\begin{aligned} U_{\sigma}(\boldsymbol{\theta}) &= \frac{(\alpha - 1)}{\sigma} \sum_{i=1}^{n} \frac{z_i \phi(z_i)\,[1 - 2\Phi(z_i)]}{\Phi(z_i)\,[1 - \Phi(z_i)]} - \frac{2\alpha}{\sigma} \sum_{i=1}^{n} \frac{\phi(z_i)\left\{\Phi^{\alpha-1}(z_i) - [1 - \Phi(z_i)]^{\alpha-1}\right\}}{\Phi^{\alpha}(z_i) + [1 - \Phi(z_i)]^{\alpha}} \\ &\quad - \frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^{n} z_i^2 \\ ,U_{\alpha}(\boldsymbol{\theta}) &= \frac{n}{\alpha} - 2\sum_{i=1}^{n} \frac{\Phi^{\alpha}(z_i) \log\,[\Phi(z_i)] + [1 - \Phi(z_i)]^{\alpha} \log\,[1 - \Phi(z_i)]}{\Phi^{\alpha}(z_i) + [1 - \Phi(z_i)]^{\alpha}} \\ &\quad + \sum_{i=1}^{n} \log\left\{\Phi(z_i)\,[1 - \Phi(z_i)]\right\}, \\ U_{\mu}(\boldsymbol{\theta}) &= \frac{1}{\sigma} \sum_{i=1}^{n} z_i - \frac{2\alpha}{\sigma} \sum_{i=1}^{n} \frac{\phi(z_i)\{\Phi^{\alpha-1}(z_i) - [1 - \Phi(z_i)]^{\alpha-1}\}}{\Phi^{\alpha}(z_i) + [1 - \Phi(z_i)]^{\alpha}}, \\ &\quad + \frac{(\alpha - 1)}{\sigma} \sum_{i=1}^{n} \frac{\phi(z_i)[1 - 2\Phi(z_i)]}{\Phi(z_i)[1 - \Phi(z_i)]}. \end{aligned}$$

If we set these equations to zero and solve them simultaneously, we can compute the MLEs of the parameters in $\boldsymbol{\theta}$. The 3×3 total observed information matrix is given by

$$J(\boldsymbol{\theta}) = \begin{pmatrix} J_{\alpha\alpha} & J_{\alpha\mu} & J_{\alpha\sigma} \\ J_{\mu\alpha} & J_{\mu\mu} & J_{\mu\sigma} \\ J_{\sigma\alpha} & J_{\alpha\mu} & J_{\sigma\sigma} \end{pmatrix},$$

whose elements can be evaluated numerically. We choose as initial values for μ and σ their MLEs $\hat{\mu}$ and $\hat{\sigma}$ under the special normal model. The initial value for α is obtained by trial and error. First, we choose values for α in the interval $[0, 1]$ and then in the sequel values greater than one. This is a kind of profile log-likelihood maximization to obtain the MLE of α. Further, we check if the global maximum is obtained using the NLMixed command of the SAS software. In the future we intend to write a script in the R software

to make the OLLN distribution user-friendly.

Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ is $N_3(0, K(\boldsymbol{\theta})^{-1})$, where $K(\boldsymbol{\theta}) = E[J(\boldsymbol{\theta})]$ is the expected information matrix. The approximate multivariate normal $N_3(0, J(\widehat{\boldsymbol{\theta}})^{-1})$ distribution, where $J(\widehat{\boldsymbol{\theta}})^{-1}$ is the observed information matrix evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$, can be used to construct approximate confidence regions for the model parameters.

The likelihood ratio (LR) statistic test can be used to compare the OLLN distribution with some of its special models. We can compute the maximum values of the unrestricted and restricted log-likelihoods to obtain LR statistics for testing some of its sub-models. In any case, hypothesis tests of the type $H_0 : \boldsymbol{\psi} = \boldsymbol{\psi}_0$ versus $H_1 : \boldsymbol{\psi} \neq \boldsymbol{\psi}_0$, where $\boldsymbol{\psi}$ is a vector formed with some components of $\boldsymbol{\theta}$ of interest and $\boldsymbol{\psi}_0$ is a specified vector, can be performed via LR statistics. For example, the test of $H_0 : \alpha = 1$ versus $H_1 : H_0$ *is not true* is equivalent to compare the OLLN and normal distributions and the LR statistic becomes

$$w = 2\{\ell(\widehat{\alpha}, \widehat{\mu}, \widehat{\sigma}) - \ell(1, \widetilde{\mu}, \widetilde{\sigma})\},$$

where $\widehat{\alpha}$, $\widehat{\mu}$ and $\widehat{\sigma}$ are the MLEs under H_1 and $\widetilde{\mu}$ and $\widetilde{\sigma}$ are the estimates under H_0.

Simulation study

We examine the performance of the OLLN distribution by simulations for different sample sizes. We conduct a Monte Carlo simulation study to assess the finite sample behavior of the MLEs of α, μ and σ. The results are obtained from 1,000 Monte Carlo simulations carried out using the matrix programming Ox language see, (DOORNIK, 2007). In each replication, a random sample of size n is drawn from the OLLN (α, μ, σ) model and the parameters are estimated by maximum likelihood. A simulation study was performed for $n = 20, 30, 80, 150$ and 300, $\alpha = 1.2$ and 0.5, $\mu = 0$ and $\sigma = 1$ and 2.5.

From the results of the simulations in Table 2.1, we note that the root mean squared errors (RMSEs) of the MLEs of μ, σ and α (given in parentheses) decay toward zero as the sample size increases, as usually expected under standard regularity conditions. As the sample size n increases, the mean estimates of the parameters tend to be closer

Tabela 2.1 - Mean estimates and RMSEs of the MLEs of the parameters of the OLLN model.

	n	20	30	150
	α	1.1109 (0.3999)	1.1364 (0.3737)	1.2493 (0.2642)
$\alpha = 1.2$	μ	-0.0019 (0.2175)	0.0026 (0.1507)	-0.0017 (0.0291)
	σ	2.2074 (1.3535)	2.3076 (1.2708)	2.5875 (0.9045)
	n	80	150	300
	α	0.6965 (0.1164)	0.6341 (0.0485)	0.6054 (0.0258)
$\alpha = 0.5$	μ	0.0364 (0.0327)	0.0196 (0.0174)	0.0133 (0.0087)
	σ	1.2560 (0.2172)	1.1833 (0.0937)	1.1478 (0.0518)

to the true parameter values. This fact supports that the asymptotic normal distribution provides an adequate approximation to the finite sample distribution of the estimates. The usual normal approximation can be improved by making bias adjustments to the MLEs. Approximations to the biases of the MLEs in simple models may be obtained analytically. Bias correction typically does a very good job in reducing the biases of the estimates. However, it may either increase the RMSEs. Whether bias correction is useful, in practice, it depends basically on the shape of the bias function and on the variance of the MLE. In order to improve the accuracy of the MLEs using analytical bias reduction one needs to obtain several cumulants of log likelihood derivatives, which are notoriously cumbersome for the proposed model. Recently, several computational methods have been developed to improve the performance of the estimators. Future works can be conducted to explore these methods such as the bootstrap, Jackknife and Bayesian, among others, in order to reduce the biases of the estimators. For example, (HASHIMOTO; ORTEGA; CANCHO; CORDEIRO, 2010) adopt the frequentist method, Jackknife estimator, bootstrap parametric and bayesian analysis to the log-exponentiated Weibull regression model for interval-censored data.

2.5 Completely randomized block design models

Normal linear regression models are usually applied in science and engineering to model symmetrical data for which linear functions of unknown parameters are used to explain or describe the phenomena under study. However, it is well known that several phenomena are not always in agreement with the normal model due to lack of symmetry in the distribution or the presence of heavy and lightly tailed distributions. Thus, in this

section, we use the OLLN model in experimental design.

The completely randomized block design (CRBD) linear model is used for local control, where the treatments are applied randomly within blocks, and frequently this type of assay represents a restriction from randomization. Consequently, the fitted statistical model must take into consideration this type of control, i.e., the model should have the form

$$y_{ij} = m + \tau_i + \beta_j + \epsilon_{ij}, \tag{2.35}$$

where y_{ij} represents the observed value of the group that received treatment i in block j, m is a common constant effect, τ_i is the effect of treatment i applied to the treated group, β_j is the effect of block j found in that group and ϵ_{ij} is the effect of the uncontrolled factors in the experimental group assuming having the OLLN distribution. Here, $i = 1, \ldots, I$ and $j = 1, \ldots, J$, where I denotes the number of treatments and J the number of blocks.

Let $y_{11}, \ldots, y_{IJ}$ be a sample of size $n = IJ$ from the OLLN distribution. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (m, \boldsymbol{\tau}^T, \boldsymbol{\beta}^T, \sigma, \nu, \alpha)^T$, where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_I)^T$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J)^T$ is given by

$$\begin{aligned} l(\boldsymbol{\theta}) &= IJ[\log(\alpha) - \log(\sigma)] + \sum_{i=1}^{I}\sum_{j=1}^{J} \log\left[\phi\left(z_{ij}\right)\right] + \alpha \sum_{i=1}^{I}\sum_{j=1}^{J} \log\left\{\Phi\left(z_{ij}\right)\left[1 - \Phi\left(z_{ij}\right)\right]\right\} \\ &\quad -2\sum_{i=1}^{I}\sum_{j=1}^{J} \log\left\{\Phi^{\alpha}\left(z_{ij}\right) + \left[1 - \Phi\left(z_{ij}\right)\right]^{\alpha}\right\}, \\ &\quad -\sum_{i=1}^{I}\sum_{j=1}^{J} \log\left\{\Phi\left(z_{ij}\right)\left[1 - \Phi\left(z_{ij}\right)\right]\right\}, \end{aligned} \tag{2.36}$$

where $z_{ij} = \left(y_{ij} - m - \tau_i - \beta_j\right)/\sigma$.

The completely randomized design (CRD) model is the first special case of the model (2.35) that follows when the local control effect $\beta_j = 0$ and the second model is the effect of the overall average that occurs when $\beta_j = 0$ and $\tau_i = 0$. By setting $\alpha = 1$, we obtain the usual standard CRBD model.

$$\frac{\partial\phi(z_{ij})}{\partial m} = \frac{\partial\phi(z_{ij})}{\partial\tau_i} = \frac{\partial\phi(z_{ij})}{\partial\beta_j} = \frac{1}{\sigma} z_{ij}\phi(z_{ij}), \qquad \frac{\partial\phi(z_{ij})}{\partial\sigma} = \frac{1}{\sigma} z_{ij}^2\phi(z_{ij}),$$

$$\frac{\Phi(z_{ij})}{\partial\sigma} = \frac{1}{\sigma} z_{ij}\phi(z_{ij}), \qquad \frac{\partial\Phi(z_{ij})}{\partial m} = \frac{\partial\Phi(z_{ij})}{\partial\tau_i} = \frac{\partial\Phi(z_{ij})}{\partial\beta_j} = \frac{1}{\sigma}\phi(z_{ij}).$$

The components of the score vector $U(\boldsymbol{\theta})$ are given by

$$\begin{aligned} U_{\beta_j}(\boldsymbol{\theta}) &= \frac{\alpha}{\sigma}\sum_{i=1}^{I}\frac{\phi(z_{ij})\left[1-2\Phi(z_{ij})\right]}{\Phi(z_{ij})\left[1-\Phi(z_{ij})\right]} - \frac{2\alpha}{\sigma}\sum_{i=1}^{I}\frac{\phi(z_{ij})\left\{\Phi^{\alpha-1}(z_{ij}) - \left[1-\Phi(z_{ij})\right]^{\alpha-1}\right\}}{\Phi^{\alpha}(z_{ij}) + \left[1-\Phi(z_{ij})\right]^{\alpha}} \\ &\quad + \frac{1}{\sigma}\sum_{i=1}^{I} z_{ij} - \frac{1}{\sigma}\sum_{i=1}^{I}\frac{\phi(z_{ij})\left[1-2\Phi(z_{ij})\right]}{\Phi(z_{ij})\left[1-\Phi(z_{ij})\right]}, \end{aligned}$$

$$\begin{aligned} U_{\sigma}(\boldsymbol{\theta}) &= -\frac{n}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J} z_i^2 - \frac{2\alpha}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{z_{ij}\phi(z_{ij})\left\{\Phi^{\alpha-1}(z_{ij}) - \left[1-\Phi(z_{ij})\right]^{\alpha-1}\right\}}{\Phi^{\alpha}(z_{ij}) + \left[1-\Phi(z_{ij})\right]^{\alpha}} \\ &\quad + \frac{(\alpha-1)}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{z_{ij}\phi(z_{ij})\left[1-2\Phi(z_{ij})\right]}{\Phi(z_{ij})\left[1-\Phi(z_{ij})\right]}, \end{aligned}$$

$$\begin{aligned} U_{m}(\boldsymbol{\theta}) &= \frac{1}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J} z_{ij}, -\frac{2\alpha}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\phi(z_{ij})\left\{\Phi^{\alpha-1}(z_{ij}) - \left[1-\Phi(z_{ij})\right]^{\alpha-1}\right\}}{\Phi^{\alpha}(z_{ij}) + \left[1-\Phi(z_{ij})\right]^{\alpha}} \\ &\quad + \frac{(\alpha-1)}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\phi(z_{ij})\left[1-2\Phi(z_{ij})\right]}{\Phi(z_{ij})\left[1-\Phi(z_{ij})\right]} \\ U_{\tau_i}(\boldsymbol{\theta}) &= \frac{1}{\sigma}\sum_{j=1}^{J} z_{ij} - \frac{2\alpha}{\sigma}\sum_{j=1}^{J}\frac{\phi(z_{ij})\left\{\Phi^{\alpha-1}(z_{ij}) - \left[1-\Phi(z_{ij})\right]^{\alpha-1}\right\}}{\Phi^{\alpha}(z_{ij}) + \left[1-\Phi(z_{ij})\right]^{\alpha}} \\ &\quad \frac{(\alpha-1)}{\sigma})\sum_{j=1}^{J}\frac{\phi(z_{ij})\left[1-2\Phi(z_{ij})\right]}{\Phi(z_{ij})\left[1-\Phi(z_{ij})\right]} \\ U_{\alpha}(\boldsymbol{\theta}) &= \frac{n}{\alpha} - 2\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\Phi^{\alpha}(z_{ij})\log\left[\Phi(z_{ij})\right] + \left[1-\Phi(z_{ij})\right]^{\alpha}\log\left[1-\Phi(z_{ij})\right]}{\Phi^{\alpha}(z_{ij}) + \left[1-\Phi(z_{ij})\right]^{\alpha}} \\ &\quad + \sum_{i=1}^{I}\sum_{j=1}^{J}\log\left\{\Phi(z_{ij})\left[1-\Phi(z_{ij})\right]\right\}. \end{aligned}$$

If we set these equations to zero and solve them simultaneously, we can compute the MLEs of the parameters in $\boldsymbol{\theta}$. The $(I+J+3)\times(I+J+3)$ total observed information

matrix is given by

$$J(\boldsymbol{\theta}) = \begin{pmatrix} J_{mm} & J_{m\tau_1} & \cdots & J_{m\tau_I} & J_{m\beta_1} & \cdots & J_{m\beta_J} & J_{m\sigma} & J_{m\alpha} \\ . & J_{\tau_1\tau_1} & \cdots & J_{\tau_1\tau_I} & J_{\tau_1\beta_1} & \cdots & J_{\tau_1\beta_J} & J_{\tau_1\sigma} & J_{\tau_1\alpha} \\ . & . & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ . & . & . & J_{\tau_I\tau_I} & \vdots & \vdots & \vdots & \vdots & \vdots \\ . & . & . & . & J_{\beta_1\beta_1} & \vdots & \vdots & \vdots & \vdots \\ . & . & . & . & . & \ddots & \vdots & \vdots & \vdots \\ . & . & . & . & . & . & J_{\beta_J\beta_J} & \vdots & \vdots \\ . & . & . & . & . & . & . & J_{\sigma\sigma} & \vdots \\ . & . & . & . & . & . & . & . & J_{\alpha\alpha} \end{pmatrix}.$$

The elements of $J(\boldsymbol{\theta})$ are evaluated numerically. We use the Optim script in R to maximize the log-likelihood function in (2.36). In addition to the goodness of fit statistics, we obtain by means of the "L-BFGS-B" or "Nelder-Mead" options the parameter estimates and their standard errors, which are the square roots of the diagonal entries of the estimated covariance matrix. We estimate the vector of parameters $\boldsymbol{\theta}$ under two restrictions $\tau_1 = 0$ and $\beta_1 = 0$. The initial values are chosen as $\hat{\boldsymbol{\theta}} = (\hat{m}, \hat{\boldsymbol{\tau}}^T, \hat{\boldsymbol{\beta}}^T, \hat{\sigma}, \hat{\alpha})^T$, where $\hat{\boldsymbol{\tau}} = (\hat{\tau}_1, \ldots, \hat{\tau}_I)^T$ and $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \ldots, \hat{\beta}_J)^T$ are obtained from the ANOVA procedure in R. In the Optim script there is the "Convergence" function, which indicates: **0** if the convergence is achieved; 1 if the maximum number of iterations "maxit" is reached; 10 if occurs degeneration of the Nelder-Mead method; 51 if occurs any warning of the "L-BFGS-B" method; and 52 if the "L-BFGS-B" method gives any errors. Then, after maximizing the log-likelihood, we check if the output was "Convergence" "**0**" thus indicating that a maximum was obtained.

2.5.1 Applications

In this section, we show the potentiality of the proposed model by means of three real data sets. In each case, the parameters are estimated by maximum likelihood (Section 2.4) using the NLMixed subroutine in SAS. First, we describe the data sets and obtain the MLEs (and the corresponding standard errors in parentheses) of the parameters and the values of the Akaike Information Criterion (AIC), Consistent Akaike Information

Criterion (CAIC) and Bayesian Information Criterion (BIC). The lower the values of these criteria, the better the fitted model. Second, we perform LR tests (Section 2.4) for the additional shape parameters. Third, we provide histograms of the data sets to show a visual comparison of the fitted density functions.

We compare the performance of the new model with those of the skew-normal, beta normal, Kumaraswamy normal and gamma normal models described below.

- **The skew normal (SN) distribution**

$$f(x) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left[\lambda\left(\frac{x-\mu}{\sigma}\right)\right], \tag{2.37}$$

where $\lambda \in \mathbb{R}$ is the parameter of asymmetry. The density function (2.37) holds for $x \in \mathbb{R}$ and it is symmetric if $\lambda = 0$ (AZZALINI, 1985).

- **The beta normal (BN) distribution**

$$f(x) = \frac{1}{\sigma B(a,b)}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\frac{x-\mu}{\sigma}\right)^{a-1}\left\{1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{b-1},$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter and $a > 0$ and $b > 0$ are shape parameters. Note that $B(a,b) = \Gamma(a)\,\Gamma(b)/\Gamma(a+b)$ is the beta function. For $\mu = 0$ and $\sigma = 1$, we obtain the standard BN distribution. The properties of the BN distribution have been studied by some authors in recent years, for example, see (EUGENE; LEE; FAMOYE, 2002), (CORDEIRO; CASTRO, 2011) and (CORDEIRO *et al.*, 2012).

- **The Kumaraswamy normal (KN) distribution**

$$f(x) = \frac{ab}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\left\{\Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{a-1}\left\{1-\Phi\left(\frac{x-\mu}{\sigma}\right)^{a}\right\}^{b-1},$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter and $a > 0$ and $b > 0$ are shape parameters. For $\mu = 0$ and $\sigma = 1$, we obtain the standard KN

distribution.

- **The gamma normal (GN) distribution**

  Recently, Alzaatreh *et al.* (2014) proposed a new three-parameter distribution called the GN distribution, with location parameter $\mu \in \mathbb{R}$, dispersion parameter $\sigma > 0$ and shape parameter $a > 0$, given by

$$f(x) = \frac{1}{\sigma\Gamma(a)}\,\phi\left(\frac{x-\mu}{\sigma}\right)\left\{-\log\left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]\right\}^{a-1}, \tag{2.38}$$

  where $\Gamma(a) = \int_0^\infty t^{a-1}\,\mathrm{e}^{-\mathrm{t}}\mathrm{dt}$ is the gamma function. The density function (2.38) does not involve any complicated function and the normal distribution arises as the basic exemplar for $a = 1$. It is a positive point of the current generalization.

2.5.2 Application 1: agronomic data

In the first application, the following response variables are used: temperature and radiation.

- Temperature and radiation - The temperature (^{o}C) and overall daily radiation ($cm^{-2}d^{-1}$) variables correspond to daily data (for the period from January 1 to December 31, 2011) obtained from the weather station of the Department of Biosystem Engineering of the Luiz de Queiroz College of Agriculture (ESALQ) of the University of São Paulo (USP), located in the city of Piracicaba, at latitude 22°42'30"S, longitude 47°38'30"W and altitude of 546 meters.

Tabela 2.2 - Descriptive statistics for dataset of temperature and radiation.

Data	n	Mean	Median	Mode	S.d.	Variance	Skewness	Kurtosis
Temperature	365	22.32	22.90	19.25	2.95	8.71	-0.50	-0.73
Radiation	365	482.5	492.0	718.0	141.3	19966.1	-0.26	-0.45

Table 2.2 gives a descriptive summary of each sample. The temperature and radiation variables have negative kurtosis. This fact can justify distributions with heavier tails required to be used to model these data. Note that the radiation variable has large variance. We compute the MLEs and the AIC, BIC and CAIC statistics for some models.

The MLEs of μ and σ for the normal distribution are taking as starting values for the numerical iterative procedure.

Tabela 2.3 - MLEs and information criteria.

Temperature	μ	σ	α		AIC	CAIC	BIC
OLLN	21.9071	0.8915	0.1861		1790.4	1790.5	1802.1
	(0.1221)	(0.1326)	(0.1861)				
	μ	σ	λ				
SN	26.3603	4.9948	-9.7087		1764.7	1764.8	1776.4
	(0.1332)	(0.2139)	(2.4430)				
	μ	σ	a	b			
BN	22.0267	0.6336	0.0197	0.0269	2180.9	2181.0	2196.5
	(0.0708)	(0.2321)	(0.0012)	(0.0018)			
KN	21.9668	24.7572	13.2443	4915.20	1800.9	1801.0	1816.5
	(6.9980)	(2.8238)	(4.2949)	(234.7)			
GN	27.4447	1.5393	0.0236		2384.5	2384.6	2396.2
	(0.0729)	(0.0001)	(0.0012)				
Normal	22.3271	2.9463			1828.7	1828.8	1836.5
	(0.1542)	(0.1090)					
Radiation	μ	σ	α		AIC	CAIC	BIC
OLLN	478.02	90.6549	0.5715		4650.5	4650.6	4662.2
	(7.4381)	(15.7894)	(0.1339)				
	μ	σ	λ				
SN	481.99	141.01	0.0049		4655.1	4655.2	4666.8
	(170.72)	(5.2572)	(1.5160)				
	μ	σ	a	b			
BN	665.21	66.3680	0.1260	0.9506	4648.8	4648.9	4664.4
	(43.3878)	(17.3035)	(0.0812)	(0.2060)			
KN	566.56	63.3518	0.1443	0.5135	4647.6	4647.7	4663.2
	(151.85)	(11.2784)	(0.0784)	(0.5084)			
GN	689.11	73.2364	0.1372		4647.1	4647.2	4658.8
	(15.1112)	(3.0157)	(0.0080)				
Normal	482.55	141.11			4653.1	4653.2	4660.9
	(7.3859)	(5.2226)					

Tabela 2.4 - LR tests.

Temperature	Hypotheses	LR statistic	p-value
Normal vs OLLN	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	40.3	<0.0001
Radiation	Hypotheses	LR statistic	p-value
Normal vs OLLN	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	4.60	0.0319

The results are reported in Table 2.3. The three information criteria agree on the model's ranking in every case. Clearly, based on the values of these statistics, the best two fitted models for the temperature data are the OLLN and SN distributions, while the BN and GN distributions are far worse. For the radiation data, based on the AIC and

BIC values, the best four models are the OLLN, BN, KN and GN distributions. Overall, the proposed model is a very competitive model to the other extended normal models.

A formal test for the need of the extra parameters in the OLLN model can be performed using LR statistics (Section 2.4). Applying these statistics to both data sets yield the results in Table 2.4. For the temperature and radiation data, we reject the normal model (under both LR tests) in favor of the new distribution. The rejection is extremely highly significant for the temperature data. This gives further evidence of the potential need for the extra parameter in the proposed model when modelling real data.

In order to assess if the model is appropriate, the histograms of these data and the plots of the fitted OLLN and normal density functions are displayed in Figure 2.4 and 2.5. The estimated OLLN density for the temperature data is bimodal. We conclude that the OLLN distribution is very suitable for both data sets.

(a) (b)

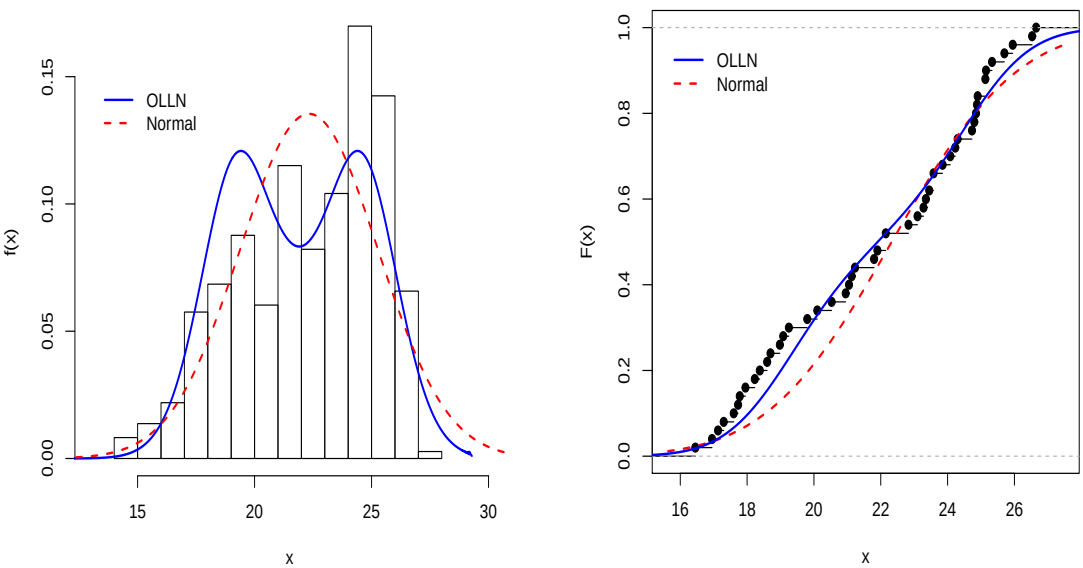


Figura 2.4 - (a) Fitted densities of the OLLN and normal models for temperature data. (b) Fitted cumulative functions and the empirical cdf for temperature data.

(c) (d)

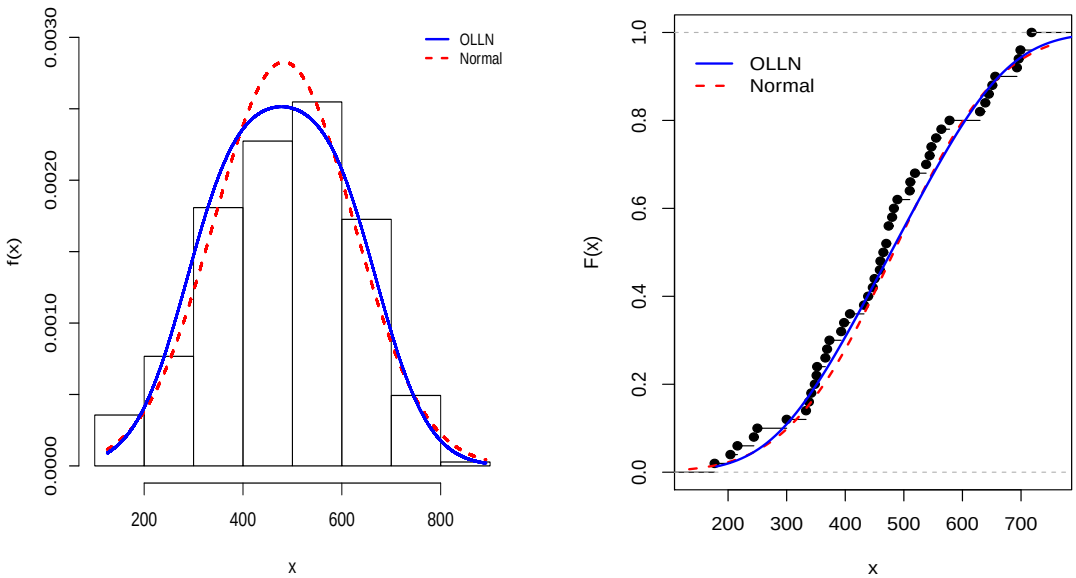


Figura 2.5 - (c) Fitted densities of the OLLN and normal models for radiation data. (d) Fitted cumulative functions and the empirical cdf for radiation data.

2.5.3 Application 2: effect of doses

The data set here comes from an experiment carried out to assess the effects of doses of an anthelmintic compound (ml) to control a parasite (fixed effects) using a CRD with 5 treatments. Treatment 1 and treatment 2 - controls, and treatment 3, treatment 4 and treatment 5 using a new drug, at concentrations of 5%, 10% and 15% and six repetitions. The data are available at Professor Braga's Euclides Malheiros website: http://jaguar.fcav.unesp.br/euclides/. Choose the year 2013 and the option Estatística Experimental-PG em Ciências Animal and download the file A DIC ex2.txt.

Table 2.5 summarizes the main descriptive statistics for each of the five treatments. It can be seen that the mean values are higher than the medians for all the treatments, indicating a positive asymmetric distribution such as the OLLN model with heavier tails of the data. This can be seen better in Figure 2.6, mainly in the boxes referring to treatments 1, 2 and 3. Another fact to be evaluated, both in Table 2.5, and Figure 2.6, is the standard deviation values, which indicate violation of the assumption of the homogeneous variance required for the analysis of variance. For example, the standard

deviation of treatment 1 is approximately 9 times that one of treatment 5.

Tabela 2.5 - Descriptive statistics for each treatment.

Statistics	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5
Average	2477.00	2075.00	527.17	156.33	91.33
Median	2481.00	1972.50	523.50	141.00	76.50
Standard deviation	483.48	274.88	200.31	38.75	52.83

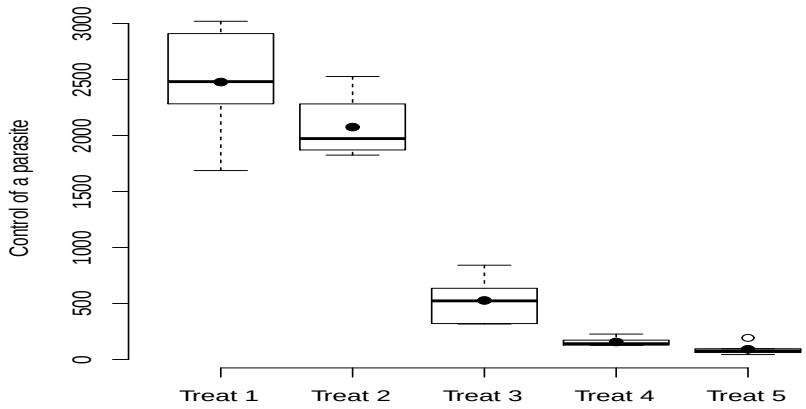


Figura 2.6 - Boxplot for each treatment where • is the average.

Table 2.6 provides the MLEs of the parameters for the OLLN, SN, BN, KN, GN and normal models. Further, we give the standard errors of the estimates and the values of the AIC, CAIC and BIC statistics. The lower these values are, the better the fit is. Table 2.7 gives the values of the LR statistic and descriptive levels, thus indicating the existence of statistical differences between the OLLN and normal distributions. By fitting model (2.35), considering $\beta_j = 0$, we place a constraint on the solution. In other words, we consider the effect of the treatment $\tau_1 = 0$. Hence, the estimates of the parameters is interpreted in relation to the treatment τ_1. For the six distributions (normal, OLLN, SN, BN, KN and GN), the estimates of the parameters are coherent with the plots in Figure 2.6, since as the doses increase (5%, 10% and 15%), the differences in modulus of the treatment values τ_2, τ_3, τ_4, τ_5 also increase as indicated in Table 2.6. This table also shows that the estimated values $\hat{m} = 2477.132$, $\hat{m} = 2466.522$, $\hat{m} = 2384.541$ and $\hat{m} = 2461.895$ for the OLLN , SN and KN distributions, respectively, decrease in relation to the normal model. This fact reveals that the OLLN model with heavier tails or the SN,

BN, KN and GN distributions have positive asymmetry, whereas the SN model stands out for presenting the smallest estimate.

Tabela 2.6 - MLEs and information criteria

	Normal			OLLN			SN	
$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE
m	2477.13	99.08	m	2466.52	110.22	m	2384.54	401.53
τ_2	-402.18	140.14	τ_2	-405.01	146.88	τ_2	-392.26	148.77
τ_3	-1950.01	140.13	τ_3	-1949.32	139.53	τ_3	-1940.79	153.78
τ_4	-2320.57	140.13	τ_4	-2310.00	131.50	τ_4	-2305.92	159.88
τ_5	-2385.70	140.13	τ_5	-2375.43	131.81	τ_5	-2374.20	159.65
σ	242.76	31.33	σ	1022.35	343.89	σ	256.90	115.76
α	-	-	α	5.23	7.05	λ	0.43	1.91
AIC	CAIC	BIC	AIC	CAIC	BIC	AIC	CAIC	BIC
426.67	433.52	435.07	424.72	433.72	434.53	428.72	437.72	438.52
	BN			KN			GN	
$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE
m	2716.72	1339.85	m	2461.89	3304.03	m	2625.18	395.79
τ_2	-371.01	227.25	τ_2	-394.74	331.92	τ_2	-388.86	156.51
τ_3	-1948.81	273.24	τ_3	-1946.10	418.98	τ_3	-1950.01	169.35
τ_4	-2335.52	326.13	τ_4	-2318.27	508.67	τ_4	-2322.01	187.94
τ_5	-2400.13	324.30	τ_5	-2385.70	505.43	τ_5	-2385.70	187.28
σ	355.77	359.77	σ	378.04	1001.68	σ	209.04	87.05
a	1.28	3.20	a	2.03	17.55	a	0.55	0.77
b	3.18	10.13	b	2.31	5.73			
AIC	CAIC	BIC	AIC	CAIC	BIC	AIC	CAIC	BIC
430.17	441.75	441.38	430.18	441.76	441.39	428.41	437.41	438.22

Tabela 2.7 - LR tests.

Models	Hypotheses	LR statistic	p-value
Normal vs OLLN	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	3.950	0.0468
Normal vs SN	$H_0 : \lambda = 0$ vs $H_1 : H_0$ is false	0.048	>0.050

In relation to the estimated standard errors, the normal distribution presents virtually equal values (140.141), a fact that is not observed in Figure 2.6, since by means of the boxplot is observed that the variance of the treatments are different. Thus, the assumption of homogeneity, which is required for ANOVA, is violated thus indicating that the normal distribution does not adequately model the variance within treatments. On the other hand, the OLLN model with heavy tails provides estimates (146.886, 139.539, 131.505, 131.817) more coherent according to the results shown in the boxplots. In this respect, the OLLN and KN distributions stand out, presenting lower standard errors than

the normal, SN, BN and GN distributions. In Table 2.7, the descriptive level (0.0468) reveals the existence of a statically significant difference between the OLLN and normal distributions, which is not noted between anyone of the SN, BN, KN and GN distributions and the normal model.

Therefore, after analysis of the estimates of the parameters, standard errors and comparative criteria, we can adopt the OLLN distribution to explain the data. Further, based on this distribution to explain the data, we can verify with respect to the treatments used as control (1 and 2), that the treatment 2 presents a lower value than treatment 1, of 405.010. The same result is observed for the new drug tested at the concentrations of 5%, 10% and 15%, for which the differences are 1949, 2310 and 2375, respectively.

2.5.4 Application 3: weight gain

This data come from a randomized block experiment with five treatments (substitution of a feed ingredient of 0%, 5%, 10%, 15% and 20%) and four blocks. The data set is available at Professor Euclides Braga's Malheiros website: $http://jaguar.fcav.unesp.br/euclides/$.Choose year 2013 and the option Estatística Experimental-PG em Ciência Animal and downlad the file A DBC ex2.txt.

Table 2.8 summarizes the main descriptive statistics (mean, median and standard deviation) for each treatment (1, 2, 3, 4 and 5). It can be checked that the mean values are higher than the medians for all treatments thus indicating that the weight gain of each animal can be adjusted by a positive asymmetric distribution or by one distribution with heavy tails, for example, the OLLN model. Thus, the behavior data of the normal distribution is not adequate to fit the data. This can be noted better in Figure 2.7, which shows an increase in the statistical values of the treatments 2, 3, 4 and 5 in relation to the first treatment. Another fact to be evaluated, both in Table 2.8 and Figure 2.7, is the standard deviation values, which indicate homogeneity of the variance within the treatments.

Table 2.9 provides the MLEs of the parameters of the normal, OLLN, SN, BN, KN and GN models, the corresponding standard errors and the AIC, CAIC and BIC statistics. Table 2.10 reports the LR statistics and descriptive level values. To fit model

Tabela 2.8 - Descriptive statistics for each treatment.

Statistics	Treat 1	Treat 2	Treat 3	Treat 4	Treat 5
Average	56.57	62.67	65.97	63.12	58.12
Median	55.10	60.85	63.45	60.75	56.00
Standard deviation	8.87	9.19	10.06	8.24	8.56

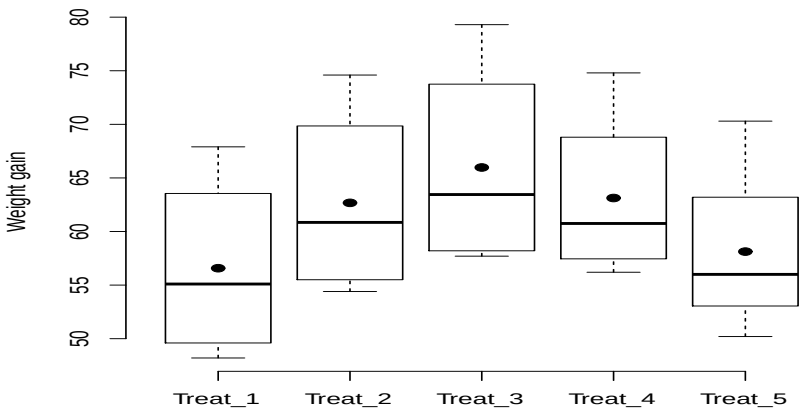


Figura 2.7 - Boxplot for each treatment where • is the average.

(2.35), we impose two constraints on the solution, i.e., we consider the effect of treatment $\tau_1 = 0$ and of block $\beta_1 = 0$. Consequently, the estimates of the treatments' parameters must be interpreted in relation to the treatment 1. The estimates of the parameters of the six distributions (normal, OLLN, SN, BN, KN and GN) are consistent with the results displayed in Figure 2.7, since as the substitution of the feed ingredient increases (5%, 10%, 15% and 20%), the estimates of the treatments (τ_2, τ_3, τ_4, τ_5) increase in relation to the treatment τ_1. This fact is observed for the symmetric normal and OLLN distributions and for all the asymmetric distributions (SN, BN, KN and GN) in Table 2.9.

It also is noted from Table 2.9 that the estimated values $\hat{m} = 49.799$, $\hat{m} = 49.140$ and $\hat{m} = 49.756$ for the normal, OLLN and KN distributions and the values $\hat{m} = 52.628$ and $\hat{m} = 51.329$ for the BN and GN distributions are practically equal, respectively. The same fact is noted for the effects of the treatments τ_2, τ_3, τ_4 and τ_5, mainly for the OLLN, KN and BN distributions. However, the results of the estimated standard errors of the OLLN distribution are smallest than those of the normal, SN, BN, KN and GN distributions. Further, the descriptive level of 0.023 in Table 2.10 indicates that there is a

Tabela 2.9 - MLEs and information criteria

	Normal			OLLN			SN	
$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE
m	49.799	1.593	m	49.140	1.131	m	49.756	7.037
τ_2	6.100	1.782	τ_2	6.038	1.207	τ_2	6.101	1.782
τ_3	9.400	1.782	τ_3	9.407	1.280	τ_3	9.401	1.782
τ_4	6.550	1.782	τ_4	6.511	1.217	τ_4	6.551	1.782
τ_5	1.550	1.782	τ_5	3.192	1.590	τ_5	1.551	1.782
β_2	1.660	1.593	β_2	1.972	1.175	β_2	1.660	1.593
β_3	6.580	1.593	β_3	8.551	1.275	β_3	6.580	1.593
β_4	18.860	1.593	β_4	19.214	1.147	β_4	18.860	1.593
σ	2.520	0.398	σ	19.825	28.729	σ	2.521	0.415
α	-	-	α	11.019	16.430	λ	0.022	3.410
AIC	CAIC	BIC	AIC	CAIC	BIC	AIC	CAIC	BIC
111.731	144.731	120.693	108.546	153.117	118.503	113.731	158.303	123.689
	BN			KN			GN	
$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE
m	52.628	2.527	m	59.979	5.278	m	51.329	1.747
τ_2	5.944	1.744	τ_2	6.425	1.393	τ_2	5.878	1.855
τ_3	9.057	1.740	τ_3	9.880	1.398	τ_3	10.269	1.940
τ_4	6.242	1.760	τ_4	6.808	1.409	τ_4	6.173	1.861
τ_5	2.349	2.232	τ_5	3.835	1.496	τ_5	3.819	1.521
β_2	2.355	1.697	β_2	1.444	1.262	β_2	3.478	1.732
β_3	8.029	1.569	β_3	7.880	1.365	β_3	8.806	1.306
β_4	19.251	1.698	β_4	18.643	1.309	β_4	20.318	1.701
σ	2.081	1.292	σ	0.934	0.885	σ	1.685	0.003
a	0.341	0.412	a	0.043	0.076	a	0.220	0.075
b	2.298	2.610	b	14.593	25.655			
AIC	CAIC	BIC	AIC	CAIC	BIC	AIC	CAIC	BIC
113.6533	174.32	124.6064	110.7614	171.4281	121.7145	114.271	158.842	124.228

Tabela 2.10 - LR tests.

Models	Hypotheses	LR statistic	p-value
Normal vs OLLN	$H_0 : \alpha = 1$ vs $H_1 : H_0$ is false	5.185	0.023
Normal vs SN	$H_0 : \lambda = 0$ vs $H_1 : H_0$ is false	0.000	> 0.050

statistical difference between the OLLN and normal distributions, whereas this difference is not significant for the SN, BN, KN, GN and normal distributions. Therefore, after analyzing the estimates of the parameters, standard errors and comparison criteria, we can use the OLLN distribution to explain the data. Based on this distribution, among the treatments studied, treatment 3 presents the highest value (9.407), followed by treatment 4, treatment 2 and treatment 5 with values of 6.511, 6.038 and 3.192, respectively.

2.6 Concluding remarks

We introduce and study a three-parameter model suitable to data with real support named the *odd log-logistic normal* (OLLN) distribution. The new model is versatile and analytically tractable, and it allows to test as a sub-model the goodness of fit of the normal distribution. The OLLN density function can take various forms depending on its shape parameters. We provide a mathematical treatment of the new distribution including explicit expressions for the ordinary and incomplete moments, generating function and mean deviations. The estimation of the model parameters is approached by the method of maximum likelihood for complete samples. The proposed distribution is capable of improving data fitting substantially over well-known extended normal models. Its flexibility, practical relevance and applicability are well illustrated by means of three applications to real data sets. In fact, it can be much more flexible than the beta normal (BN), skew-normal (SN), gamma normal (GN) and Kumaraswamy normal (KN) models studied recently. Powerful parametric methods and likelihood ratio tests can be effectively applied to the analysis of experiments to which previous attempts are unsatisfactory and unreliable due to lack of wider parametric models.

References

ALZAATREH, A.; LEE, C.; FAMOYE, F. (2013). A new method for generating families of continuous distributions. **Metron 71**: 1, 63-79.

ALZAATREH, A.; FAMOYE, F.; LEE, C. (2014). The gamma-normal distribuition: Properties and applications. **Computational Statistics and Data Analysis**. New York, v.69, n.4, p.67-80.

AZZALINI, A. (1985). A class of distributions which includes the normal ones. **Scandinavian Journal of Statistics**, **12**, 171-178.

CORDEIRO, G. M.; CINTRA, R. J.; RÊGO, L. C; ORTEGA, E. M. M. (2012). The McDonald normal distribution. **Pakistan Journal of Statistics and Operation Research**, **8**, 301-329.

CORDEIRO, G. M.; de CASTRO, M. A. (2011). New family of generalized distribution. **Journal of Statistical Computation and Simulation**, London, v. **81**, p. 883-893.

DOORNIK, J. (2007).**Ox: Object-oriented matrix programming using Ox**. London : Timberlake Consultants Ltd, 461p.

EUGENE, N.; LEE, C.; FAMOYE, F. (2002). Beta-normal distribution and its applications. **Communications in Statistics - Theory and Methods**, **31**, 497-512.

EXTON, H. (1978). **Handbook of Hypergeometric Integrals: Theory, Applications, Tables, Computer Programs**. Halsted Press, New York.

GRADSHTEYN, I. S.; RYZHIK, I. M. (2007). **Table of integrals, series, and products**. Academic Press. San Diego.

HASHIMOTO, E. M.; ORTEGA, E. M. M.; CANCHO, V. G.; CORDEIRO, G. M. (2010). The log-exponentiated Weibull regression model for interval-censored data. **Computational Statistics and Data Analysis**, **54**, 1017-1035.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-real data. **IEEE Transaction on Reliability**, **42**, 299-302.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; FRIEMER, M. (1995). The exponential Weibull family: A reanalysis of the bus-motor failure data. **Technometrics**, **37**, 436-445.

NADARAJAH, S. (2008). Explicit expressions for moments of order statistics. **Statistics and Probability Letters**, **78**, 196-205.

NADARAJAH, S.; GUPTA, A. K. (2007). A generalized gamma distribution with application to drought data. **Mathematics and Computer in Simulation**, **74**, 1-7.

NADARAJAH, S.; KOTZ, S. (2006). The exponentiated type distribution. **Acta Applicandae Mathematicae**, **92**, 97-111.

NADARAJAH, S. (2005). The exponentiated Gumbel distribution with climate application. **Environmetrics**, **17**, 13-23.

PRUDNIKOV, A. P.; BRYCHKOV, Y. A.; MARICHEV, O.I. (1986). **Integrals and series: Elementary functions**. Gordon and Breach Science Publishers, New York.

SHIRKE, D. T.; KAKADE, C. S. (2006). On exponentiated log-normal distribution. **International Journal of Agricultural and Statistical Science**, **2**, 319-326.

STEINBRECHER, G. (2002). **Taylor expansion for inverse error function around origin**. Working paper at University of Craiova.

WHITTAKER, E. T.; WATSON, G.N. (1990). **A Course in Modern Analysis**. Cambridge, University Press, Cambridge.

3 THE ODD LOG-LOGISTIC STUDENT DISTRIBUTION: THEORY AND APPLICATIONS

Abstract

The normal distribution is the most used in analysis of experiments. However, it is not suitable to apply in situations where the data has evidence of bimodality or heavier tails than the normal distribution. So, we propose a new four-parameter model called the odd log-logistic t-Student (OLLS) distribution as an alternative to the normal and t-Student distributions. The new distribution is symmetric, platykurtic, leptokurtic and may be unimodal or bimodal. Various structural properties are derived including explicit expressions for the moments and mean deviations. The estimation of the model parameters is performed by maximum likelihood. The new model can be used as an alternative to the model for completely randomized block design thus providing an analysis of data more realistic than other special regression models. Besides, we use sensitivity analysis to detect influential or outlying observations, and residual analysis with generated envelopes is performed to select appropriate models. We illustrate the importance of the proposed model by means of two real datasets in analysis of experiments carried out in different regions of Brazil.

Keywords: Log-logistic distribution; Maximum likelihood estimation; Mean deviation; Regression model; Student's t distribution

3.1 Introduction

Analysis of experiments are usually applied in science and engineering to model symmetrical data. However, it is well-known that several phenomena are not always in agreement with the normal model due to lack of symmetry or bimodality in the distribution or the presence of heavy and lightly tailed distributions.

The Student's t distribution is the second most popular continuous distribution in Statistics, second only to the normal distribution. Most statistical analysis carried out in experimental investigations are in some way based on the normal distribution. Consequently, besides assuring the assumption of normality for the statistical inferences, it is necessary to verify other conditions, such as symmetry and kurtosis of the data distribution. Extreme values, or outliers, are often present, requiring modeling a probability distribution with heavier tails, such as the Student's t distribution.

However, neither the normal nor the Student's t distribution is suitable to model data with evidence of asymmetry, although for small samples the Student's t distribution

is a more robust model that allows diminishing the influence of outliers on the statistical inferences. Some extensions and modifications of the t-Student distribution are found in the literature, for example, (JONES; FADDY, 2003) proposed a skew extension of the t-Student distribution and (ARELLANO-VALLE; GENTON, 2005) introduced the skew t-Student distribution.

In this paper, we propose the *odd log-logistic t-Student* ("OLLS" for short) distribution, which generalizes the Student's t distribution, and study some of its structural properties. The new distribution shows flexibility in accommodating data that have extreme observations and bimodality. We hope that it will be used in a variety of problems in modeling data. Besides, we propose a linear model based on the OLLS distribution as an alternative to work with smaller samples from experimental designs, e.g., completely randomized design (CRD) or completely randomized block design (CRBD).

After modeling, it is important to check assumptions in the model and to conduct a robustness study in order to detect influential or extreme observations that can cause distortions in the results of the analysis. A solution for the earlier problem can be found in the local influence approach where one again investigates how the results of the analysis are changed under small perturbations in the model or data.

Cook (1986) proposed a general framework to detect influence of observations which indicates how sensitive is the analysis when small perturbations are provoked on the data or in the model. Several authors have applied these methods to nonlinear regression models different to the normal case; see for instance, (LIN; XIE; WEI, 2009), (LABRA *et al.*, 2012) and (GARAY *et al.*, 2014).

We propose a similar methodology to detect influential subjects in linear model to the CRBD. The assessment of the fitted model is an important part of data analysis, particularly in regression models, and residual analysis is a helpful tool to validate the fitted model. In this chapter, we use the quantile residuals in linear model to the CRBD. Finally, in the section on applications of the new model, two datasets are analyzed that present asymmetry and bimodality, so that the normal distribution is not suitable to model these data.

The rest of the chaper is organized as follows. We define and derive properties and

useful expansions for the probability density function (pdf) and cumulative distribution function (cdf) in Subsection 3.2.2. The estimation of the model parameters is performed by maximum likelihood in Subsection 3.3. We also discuss some simulation studies in Subsection 3.3.1. We consider the statistical models of the CRBD in Section 3.4. Several diagnostic measures are presented in Subsection 3.4.1 by considering normal curvatures of local influence under various perturbation schemes. In Subsection 3.4.4, a kind of quantile residual is proposed to assess departures from the underlying OLLS distribution in linear model to CRBD and to detect outliers. Further, three applications are provided in Section 3.5 in analysis of experiments. Concluding remarks are addressed in Section 3.8.

3.2 The model

3.2.1 Definition

In recent decades, many studies were developed with the goal of creating new families of distributions. For example, (AZZALINI, 1985) pioneered the skew-normal distribution and obtained some properties; (EUGENE; LEE; FAMOYE, 2002), the beta normal distribution and applications; (CORDEIRO; CASTRO, 2011), a family of generalized distributions and, recently, (ALZAATREH; FAMOYE; LEE, 2014) obtained the gamma normal distribution with applications to engineering data. According with (ALZAATREH; LEE; FAMOYE, 2013) pointed out that many of these methods aim to generalize continuous distributions.

Thus, various other methods can be found in the literature, for example, the class of generalized beta-generated distributions by (ALEXANDER *et al.*, 2012) and, recently, (ALIZADEH *et al.*, 2015) defined a new family of distributions: the Kumaraswamy odd log-logistic, properties and applications.

Alzaatreh *et al.* (2013) defined the S-X family of distributions. Let $r(s)$ be the pdf of a random variable $S \in [a, b]$ for $-\infty < a < b < \infty$ and let $W[G(X)]$ be a function of the cdf of a random variable X such that $W[G(X)]$ satisfies the following conditions:

- $W[G(X)] \in$ [a,b];
- $W[G(X)]$ is differentiable and monotonically non-decreasing, and

- $W[G(X)] \to a$ as $x \to -\infty$ and $W[G(X)] \to b$ as $x \to \infty$.

Let X be a random variable with pdf $g(x)$ and cdf $G(x)$. If S is a random variable with pdf $r(s)$ defined in the range $[a, b]$, the cdf of the S-X family of distributions is defined by

$$F(x) = \int_a^{W[G(X)]} r(s)ds. \tag{3.1}$$

Thus, the pdf of the S-X family is given by

$$f(x) = r\left\{W[G(X)]\right\} \frac{d}{dx}\left\{W[G(X)]\right\}, \quad -\infty < x < \infty. \tag{3.2}$$

The t-Student distribution is perhaps the most widely applied statistical distribution for date modeling with extremes values. The pdf and cdf of the t-Student (for $t \in \mathbb{R}$) model are given by

$$g(t;\mu,\sigma,\nu) = \frac{\nu^{\frac{\nu}{2}}}{\sigma B\left(\frac{1}{2},\frac{\nu}{2}\right)}\left[\nu + \left(\frac{t-\mu}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} = \frac{1}{\sigma}\phi_{t_\nu}\left(\frac{t-u}{\sigma}\right) \tag{3.3}$$

and

$$G(t;\mu,\sigma,\nu) = \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right) = \begin{cases} \frac{1}{2} - \frac{1}{2} I_{\left[\frac{\left(\frac{t-\mu}{\sigma}\right)^2}{\nu+\left(\frac{t-\mu}{\sigma}\right)^2}\right]}\left(\frac{1}{2},\frac{\nu}{2}\right) & \text{for} \quad -\infty < t < 0 \\ \frac{1}{2} + \frac{1}{2} I_{\left[\frac{\left(\frac{t-\mu}{\sigma}\right)^2}{\nu+\left(\frac{t-\mu}{\sigma}\right)^2}\right]}\left(\frac{1}{2},\frac{\nu}{2}\right) & \text{for} \quad 0 \le t < \infty, \end{cases} \tag{3.4}$$

where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, ν is the number of degrees of freedom, $I_t(a,b)$ is the regularized beta function defined by $I_t(a,b) = B_t(a,b)/B(a,b)$ with $B_t(a,b)$ the incomplete beta function and $B(a,b)$ the (complete) beta function and $\phi_{t_\nu}(\cdot)$ and $\Phi_{T_\nu}(\cdot)$ are the pdf and cdf of the standard Student t distribution, respectively, are presented the demonstration of the $G(t;\mu,\sigma,\nu)$ in Appendix A.

We propose a new extension of the Student's t distribution with heavier tails by taking $W[G(X)]$ and $r(s)$ in equation (3.1) to be $W[G(X)] = G(t;\mu,\sigma,\nu)/\bar{G}(t;\mu,\sigma,\nu)$ and $r(s) = \alpha s^{\alpha-1}/(1+s^\alpha)^2$, $s > 0$, so-called the OLLS model. Then, its cdf with an

additional shape parameter $\alpha > 0$ is given by

$$F(t;\mu,\sigma,\nu,\alpha) = \int_0^{\frac{G(t;\mu,\sigma,\nu)}{\bar{G}(t;\mu,\sigma,\nu)}} \frac{\alpha\, s^{\alpha-1}}{(1+s^{\alpha})^2} ds = \frac{G(t;\mu,\sigma,\nu)^{\alpha}}{G(t;\mu,\sigma,\nu)^{\alpha} + \bar{G}(t;\mu,\sigma,\nu)^{\alpha}}, \tag{3.5}$$

where $\bar{G}(t;\mu,\sigma,\nu) = 1 - G(t;\mu,\sigma,\nu)$.

The pdf corresponding to (3.5) recudes to

$$f(t;\mu,\sigma,\nu,\alpha) = \frac{\alpha\, g(t;\mu,\sigma,\nu)\left\{G(t;\mu,\sigma,\nu)[1-G(t;\mu,\sigma,\nu)]\right\}^{\alpha-1}}{\left\{G(t;\mu,\sigma,\nu)^{\alpha} + [1-G(t;\mu,\sigma,\nu)]^{\alpha}\right\}^2}. \tag{3.6}$$

The shape parameter α can be in the form

$$\alpha = \frac{\log\left[\frac{F(t;\mu,\sigma,\nu,\alpha)}{\bar{F}(t;\mu,\sigma,\nu,\alpha)}\right]}{\log\left[\frac{G(t;\mu,\sigma,\nu)}{\bar{G}(t;\mu,\sigma,\nu)}\right]}.$$

where $\bar{\Phi}(x;\mu,\sigma) = 1 - \Phi(x;\mu,\sigma)$ and $\bar{F}(x;\mu,\sigma,\nu,\alpha) = 1 - F(x;\mu,\sigma,\nu,\alpha)$. So, the parameter α represents the quotient of the log odds ratios for the generated and baseline distributions.

Substituting (3.3) and (3.4) in equations (3.5) and (3.6), the cdf and pdf of the OLLS distribution can be expressed as

$$F(t;\mu,\sigma,\nu,\alpha) = \frac{\Phi^{\alpha}_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)}{\Phi^{\alpha}_{T_\nu}\left(\frac{t-\mu}{\sigma}\right) + \left[1-\Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)\right]^{\alpha}} \tag{3.7}$$

and

$$f(t;\mu,\sigma,\nu,\alpha) = \frac{\alpha\phi_{t_{(\nu)}}\left(\frac{t-\mu}{\sigma}\right)\Phi^{\alpha-1}_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)\left[1-\Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)\right]^{\alpha-1}}{\sigma\left\{\Phi^{\alpha}_{T_\nu}\left(\frac{t-\mu}{\sigma}\right) + \left[1-\Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)\right]^{\alpha}\right\}^2}, \tag{3.8}$$

respectively. Note that $\alpha > 0$ is a shape parameter. Henceforth, a random variable with density function (3.8) is denoted by $T \sim \text{OLLS}(\mu,\sigma,\nu,\alpha)$.

The OLLS distribution contains as special cases several well-known distributions. For example, it becomes the Student's t distribution when $\alpha = 1$. If $\alpha = 1$, in addition to $\nu = \infty$, it reduces to the normal distribution. The odd log logistic Cauchy (OLLC) distribution is also a special case when $\nu = 1$. For $\alpha = 1$ and $\nu = 1$, the OLLS distribution

gives the Cauchy distribution. For $\mu = 0$ and $\sigma = 1$, we obtain the standard OLLS distribution. Further, the OLLS distribution with $\alpha = 1$, $\mu = 0$, $\sigma = 1$ and $\nu = \infty$, reduces to the standard normal distribution. The OLLS distribution allows for greater flexibility of its tails and can be widely applied in many areas of engineering, biology and agronomy.

3.2.2 Properties and useful expansions for the OLLS model

It is not possible to study the behavior of the parameters of the OLLS$(\mu, \sigma, \nu, \alpha)$ distribution by taking derivatives. We verify the distribution bimodality by combining some values of σ, α and ν as shown in the plots of Figure 3.1. Further, the quantile function (qf) is derived, because it is important to obtain some of the OLLS properties and also for performing a simulation study.

Equation (3.7) has tractable properties specially for simulations, since the qf of T has a simple form. Let $F(t; \mu, \sigma, \nu, \alpha) = u$ and $\Phi_{T_\nu}^{-1}\left(\frac{t-\mu}{\sigma}\right)$ be the inverse function of $\Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)$. Then,

$$\begin{aligned} u &= \frac{\Phi_{T_\nu}^{\alpha}\left(\frac{t-\mu}{\sigma}\right)}{\Phi_{T_\nu}^{\alpha}\left(\frac{t-\mu}{\sigma}\right) + \left[1 - \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)\right]^{\alpha}} \\ Q_{\text{OLLS}} &= \mu + \sigma \Phi_{T_\nu}^{-1}\left(\frac{u^{1/\alpha}}{[1-u]^{1/\alpha} + u^{1/\alpha}}\right), \end{aligned} \tag{3.9}$$

where Q_{OLLS} is the qf of T determined by inverting (3.7).

We provide a mixture representation for the OLLS pdf. First, we define the exponentiated - Φ_{T_ν} ("Exp - Φ_{T_ν}") distribution with power parameter $c > 0$ for an arbitrary parent t-Student model, say $W \sim \text{Exp}^c(\Phi_{T_\nu})$, if W has cdf and pdf given by

$$H_c(t) = \Phi_{T_\nu}^{c}\left(\frac{t-\mu}{\sigma}\right) \qquad \text{and} \qquad h_c(t) = c\,\sigma^{-1}\phi_{t_\nu}\left(\frac{t-\mu}{\sigma}\right)\Phi_{T_\nu}^{c-1}\left(\frac{t-\mu}{\sigma}\right),$$

respectively. The properties of Exp-G distributions have been studied by many authors in recent years, see (MUDHOLKAR; SRIVASTAVA, 1993) and (MUDHOLKAR; SRIVASTAVA; FRIEMER, 1995) for exponentiated Weibull, (SHIRKE; KAKADE, 2006) for exponentiated log-normal and (NADARAJAH; GUPTA, 2007) for exponentiated gamma

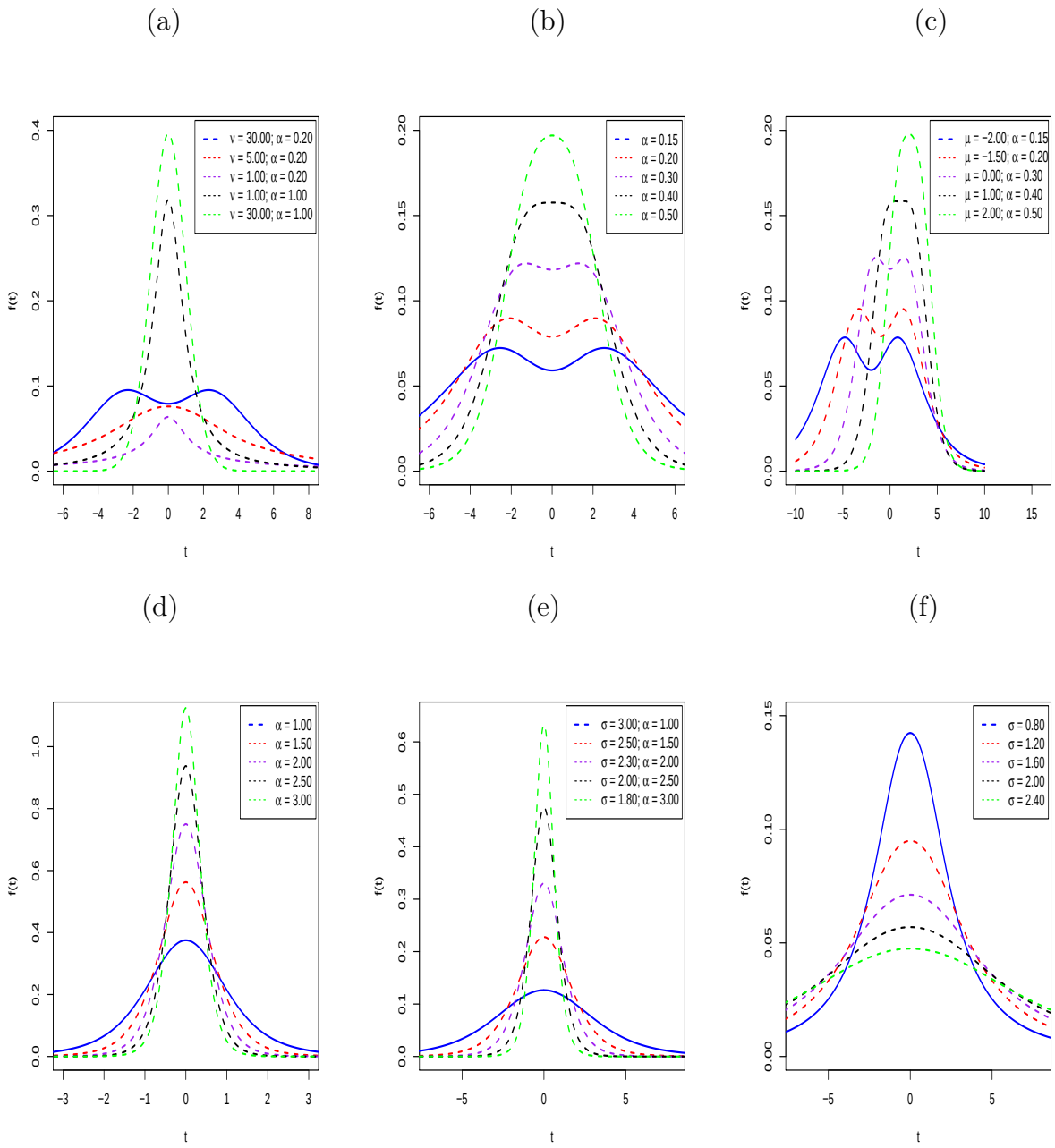


Figura 3.1 - Plots of the OLLS$(\mu, \sigma, \nu, \alpha)$ pdf for some parameter values. (a) For different values of ν and α with $\mu = 0$ and $\sigma = 1$. (b) For different values of α with $\mu = 0$, $\sigma = 1$ and $\nu = 20$. (c) For different values of μ and α with $\sigma = 1$ and $\nu = 30$. (d) For different values of α with $\mu = 0$, $\sigma = 1$ and $\nu = 4$. (e) For different values of α and σ with $\mu = 0$ and $\nu = 5$. (f) For different values of σ with $\alpha = 0.3$, $\mu = 0$ and $\nu = 5$.

distributions. See, also, (NADARAJAH; KOTZ, 2006), among others. We believe that the Exp - Φ_{T_ν} model was not explored yet.

First, we obtain an expansion for $F(t; \mu, \sigma, \nu, \alpha)$. We use a power series for

$\Phi_{T_\nu}^{\alpha}\left(\frac{t-\mu}{\sigma}\right)$ ($\alpha > 0$ real) given by

$$\Phi_{T_\nu}^{\alpha}\left(\frac{t-\mu}{\sigma}\right) = \sum_{k=0}^{\infty} a_k\, \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)^k, \tag{3.10}$$

where

$$a_k = a_k(\alpha) = \sum_{j=k}^{\infty} (-1)^{k+j} \binom{\alpha}{j} \binom{j}{k}.$$

For any real $\alpha > 0$, we consider the generalized binomial expansion

$$\left[1 - \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)\right]^{\alpha} = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)^k. \tag{3.11}$$

Inserting (3.10) and (3.11) in equation (3.7), we obtain

$$F(t;\mu,\sigma,\nu,\alpha) = \frac{\sum_{k=0}^{\infty} a_k\, \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)^k}{\sum_{k=0}^{\infty} b_k\, \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)^k}, \tag{3.12}$$

where $b_k = a_k + (-1)^k \binom{\alpha}{k}$ for $k \geq 0$. The ratio of the two power series in equation (3.12) can be expressed as

$$F(t;\mu,\sigma,\nu,\alpha) = \sum_{k=0}^{\infty} c_k\, \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)^k, \tag{3.13}$$

where the coefficients c_k's (for $k \geq 0$) are determined from the recurrence equation

$$c_k = b_0^{-1}\left(a_k - \sum_{r=1}^{k} b_r\, c_{k-r}\right).$$

Thus, the pdf of T follows by differentiating (3.13) as

$$f(t;\mu,\sigma,\nu,\alpha) = \sum_{k=0}^{\infty} c_{k+1}\, h_{k+1}(t), \tag{3.14}$$

where $h_{k+1}(t) = (k+1)\,\sigma^{-1}\phi_{t_\nu}\left(\frac{t-\mu}{\sigma}\right)\, \Phi_{T_\nu}\left(\frac{t-\mu}{\sigma}\right)^k$ is the Exp-Φ_{T_ν} density function with power parameter $k+1$.

Some mathematical properties of the OLLS(μ,σ,ν,α) model are presented in Appendix B.

3.3 Maximum likelihood estimation

Let $t_1, \ldots, t_n$ be a random sample of size n from the OLLS$(\mu, \sigma, \nu, \alpha)$ distribution. Here, we determine the maximum likelihood estimates (MLEs) of the model parameters from complete samples only. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\mu, \sigma, \nu, \alpha)^T$ is given by,

$$\begin{aligned} l(\boldsymbol{\theta}) &= n[\log(\alpha) - \log(\sigma)] + \sum_{i=1}^{n} \log[\phi_{t_\nu}(z_i)] + \alpha \sum_{i=1}^{n} \log\{\Phi_{T_\nu}(z_i)[1 - \Phi_{T_\nu}(z_i)]\} \\ &\quad -2\sum_{i=1}^{n} \log\{\Phi_{T_\nu}^{\alpha}(z_i) + [1 - \Phi_{T_\nu}(z_i)]^{\alpha}\}, \\ &\quad -\sum_{i=1}^{n} \log\{\Phi_{T_\nu}(z_i)[1 - \Phi_{T_\nu}(z_i)]\} \end{aligned} \tag{3.15}$$

where $z_i = (t_i - \mu)/\sigma$.

In general, it is reasonable to expect that the shape parameter $\nu > 0$ is fixed in the first step of the iterative process to maximize (3.15). Then, we evaluate the MLEs $\widetilde{\mu}(\nu)$, $\widetilde{\sigma}(\nu)$ and $\widetilde{\alpha}(\nu)$ and the maximized profile log-likelihood function $\hat{l}(\nu)$ for fixed ν. We use in this step the Optim script in R. In the second step, the profile log-likelihood $\hat{l}(\nu)$ is maximized for a grid of values of ν and then $\widehat{\nu}$ is obtained. The MLEs of μ, σ and α are, respectively, given by $\widehat{\mu} = \widetilde{\mu}(\widehat{\nu})$, $\widehat{\sigma} = \widetilde{\sigma}(\widehat{\nu})$, and $\widehat{\alpha} = \widetilde{\alpha}(\widehat{\nu})$.

The procedure discussed in this work is developed by assuming ν fixed. The tests and estimates for $\boldsymbol{\theta} = (\mu, \sigma, \alpha)^T$ can be based on the asymptotic normal approximation for $\widehat{\boldsymbol{\theta}}$.

We obtain the following expressions

$$\frac{\partial \phi_{t_\nu}(z_i)}{\partial \mu} = \frac{(\nu+1)\, z_i}{\sigma\,(\nu + z_i^2)} \phi_{t_\nu}(z_i), \qquad \frac{\partial \phi_{t_\nu}(z_i)}{\partial \sigma} = \frac{(\nu+1)\, z_i^2}{\sigma\,(\nu + z_i^2)} \phi_{t_\nu}(z_i), \qquad \frac{\partial \Phi_{T_\nu}(z_i)}{\partial \mu} = \frac{1}{\sigma} \phi_{t_\nu}(z_i)$$
$$\text{and} \qquad \frac{\partial \Phi_{T_\nu}(z_i)}{\partial \sigma} = \frac{1}{\sigma} z_i \phi_{t_\nu}(z_i),$$

and the components of the score vector $U(\boldsymbol{\theta})$ are given by

$$\begin{aligned} U_\mu(\boldsymbol{\theta}) &= \frac{(\alpha-1)}{\sigma} \sum_{i=1}^{n} \frac{\phi_{t_\nu}(z_i)[1 - 2\Phi_{T_\nu}(z_i)]}{\Phi_{T_\nu}(z_i)[1 - \Phi_{T_\nu}(z_i)]} - \frac{2\alpha}{\sigma} \sum_{i=1}^{n} \frac{\phi_{t_\nu}(z_i)\{\Phi_{T_\nu}^{\alpha-1}(z_i) - [1 - \Phi_{T_\nu}(z_i)]^{\alpha-1}\}}{\Phi_{T_\nu}^{\alpha}(z_i) + [1 - \Phi_{T_\nu}(z_i)]^{\alpha}} \\ &\quad + \sum_{i=1}^{n} \frac{(\nu+1)\, z_i}{\sigma\,(\nu + z_i^2)}, \end{aligned}$$

$$
\begin{aligned}
U_\sigma(\boldsymbol{\theta}) &= \frac{\alpha}{\sigma}\sum_{i=1}^{n}\frac{z_i\phi_{t_\nu}(z_i)\left[1-2\Phi_{T_\nu}(z_i)\right]}{\Phi_{T_\nu}(z_i)\left[1-\Phi_{T_\nu}(z_i)\right]}-\frac{2\alpha}{\sigma}\sum_{i=1}^{n}\frac{\phi_{t_\nu}(z_i)\left\{\Phi_{T_\nu}^{\alpha-1}(z_i)-\left[1-\Phi_{T_\nu}(z_i)\right]^{\alpha-1}\right\}}{\Phi_{T_\nu}^{\alpha}(z_i)+\left[1-\Phi_{T_\nu}(z_i)\right]^{\alpha}} \\
&\quad -\frac{n}{\sigma}+\frac{(\nu+1)}{\sigma}\sum_{i=1}^{n}\frac{z_i^2}{(\nu+z_i^2)}-\frac{1}{\sigma}\sum_{i=1}^{n}\frac{z_i\phi_{t_\nu}(z_i)\left[1-2\Phi_{T_\nu}(z_i)\right]}{\Phi_{T_\nu}(z_i)\left[1-\Phi_{T_\nu}(z_i)\right]},
\end{aligned}
$$

$$
\begin{aligned}
U_\alpha(\boldsymbol{\theta}) &= \frac{n}{\alpha}-2\sum_{i=1}^{n}\frac{\Phi_{T_\nu}^{\alpha}(z_i)\log\left[\Phi_{T_\nu}(z_i)\right]+\left[1-\Phi_{T_\nu}(z_i)\right]^{\alpha}\log\left[1-\Phi_{T_\nu}(z_i)\right]}{\Phi_{T_\nu}^{\alpha}(z_i)+\left[1-\Phi_{T_\nu}(z_i)\right]^{\alpha}} \\
&\quad +\sum_{i=1}^{n}\log\left\{\Phi_{T_\nu}(z_i)\left[1-\Phi_{T_\nu}(z_i)\right]\right\}.
\end{aligned}
$$

By setting these equations to zero and solve them simultaneously, we can obtain the MLEs of the parameters in $\boldsymbol{\theta}$ for fixed ν. The 3×3 total observed information matrix is given by

$$
\mathbf{L}(\boldsymbol{\theta})=\begin{pmatrix} L_{\alpha\alpha} & L_{\alpha\mu} & L_{\alpha\sigma} \\ . & L_{\mu\mu} & L_{\mu\sigma} \\ . & . & L_{\sigma\sigma} \end{pmatrix},
$$

whose elements can be evaluated numerically.

Under standard regularity conditions, the asymptotic distribution of $(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ is given by $N_3(0,K(\boldsymbol{\theta})^{-1})$, where $K(\boldsymbol{\theta})=E[\mathbf{L}(\boldsymbol{\theta})]$ is the expected information matrix. The approximate multivariate normal $N_3(0,\mathbf{L}(\widehat{\boldsymbol{\theta}})^{-1})$ distribution, where $\mathbf{L}(\widehat{\boldsymbol{\theta}})^{-1}$ is the observed information matrix evaluated at $\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}$, allow us to construct approximate confidence intervals for the model parameters.

The likelihood ratio (LR) statistic can be used to compare the OLLS distribution with some of its special models. We can compute the maximum values of the unrestricted and restricted log-likelihoods to obtain LR statistics for testing some of its sub-models. In any case, hypothesis tests of the type $H_0:\ \boldsymbol{\psi}=\boldsymbol{\psi}_0$ versus $H:\ \boldsymbol{\psi}\neq\boldsymbol{\psi}_0$, where $\boldsymbol{\psi}$ is a vector formed with some components of $\boldsymbol{\theta}$ and $\boldsymbol{\psi}_0$ is a specified vector, can be performed using LR statistics. For example, the test of $H_0:\ \alpha=1,\ \nu=\infty$ versus $H:\ H_0$ is not true is equivalent to compare the OLLS and normal distributions and the LR statistic becomes

$$w = 2\{\ell(\widehat{\mu}, \widehat{\sigma}, \widehat{\alpha}) - \ell(\widetilde{\mu}, \widetilde{\sigma}, 1)\},$$

where $\widehat{\mu}$, $\widehat{\sigma}$ and $\widehat{\alpha}$ are the MLEs under H and $\widetilde{\mu}$ and $\widetilde{\sigma}$ are the estimates under H_0.

3.3.1 Simulation study

We conduct a Monte Carlo simulation study to assess the finite sample behavior of the MLEs of μ, σ and α. The results are obtained from 1,000 Monte Carlo simulations carried out using the matrix programming language Ox see (DOORNIK, 2006). In each replication, a random sample of size n is drawn from the OLLS(μ, σ, α) model and the parameters are estimated by maximum likelihood. A simulation study was performed for $n = 20, 30$ and 300, $\alpha = 0.5$ and 0.8, $\mu = 0$ and $\sigma = 1$. For these simulations, the parameter ν is fixed at $\nu = 5$ and $\nu = 6$.

Tabela 3.1 - Mean estimates and RMSEs (given in parentheses) of the MLEs of the parameters of the OLLS model.

Parameters	n	20	30	300
	α	0.4840 (0.0709)	0.5354 (0.0850)	0.5350 (0.0125)
$\alpha = 0.5;\ \nu = 5$	μ	0.0009 (0.2306)	0.0082 (0.1643)	0.0058 (0.0177)
	σ	0.8695 (0.4107)	1.0202 (0.4633)	1.0754 (0.0694)
	α	0.4168 (0.0336)	0.4701 (0.0354)	0.5523 (0.0244)
$\alpha = 0.5;\ \nu = 6$	μ	0.0187 (0.2099)	0.0115 (0.1514)	0.0031 (0.0170)
	σ	0.7420 (0.2156)	0.8805 (0.2002)	1.1076 (0.1150)
	α	0.5651 (0.1664)	0.6615 (0.1608)	0.9041 (0.1331)
$\alpha = 0.8;\ \nu = 5$	μ	0.0063 (0.0941)	0.0113 (0.0696)	0.0047 (0.0067)
	σ	0.6057 (0.3632)	0.7613 (0.3147)	1.1336 (0.2380)
	α	0.4580 (0.1490)	0.5325 (0.1155)	0.8393 (0.0988)
$\alpha = 0.8;\ \nu = 6$	μ	0.0001 (0.0960)	0.0058 (0.0678)	0.0016 (0.0065)
	σ	0.4918 (0.3266)	0.6012 (0.2501)	1.0464 (0.1680)

From the results of the simulations in Table 3.1, we note that the root mean squared errors (RMSEs) of the MLEs of μ, σ and α (given in parentheses) decay toward zero as the sample size increases, as usually expected under standard regularity conditions. As the sample size n increases, the mean estimates of the parameters tend to be closer to the true parameter values. This fact supports that the asymptotic multivariate normal distribution provides an adequate approximation to the finite sample distribution of the

estimators. The usual normal approximation can be improved by making bias corrections to the MLEs. Approximations to the biases of the MLEs in simple models may be obtained analytically. Bias correction typically does a very good job in reducing the biases of the estimates. However, it may either increase the RMSEs. Whether bias correction is useful in practice depends basically on the shape of the bias function and on the variance of the MLE. In order to improve the accuracy of the MLEs using analytical bias reduction one needs to obtain several cumulants of log-likelihood derivatives, which are notoriously cumbersome for the proposed model.

3.4 Statistical models of the completely randomized block design

The normal linear regression models are usually applied in science and engineering to model symmetrical data for which linear functions of unknown parameters are used to explain or describe the phenomena under study. However, it is well known that several phenomena are not always in agreement with the normal model due to lack of symmetry in the distribution or the presence of heavy and lightly tailed distributions. Thus, in this section, we use the OLLS model in experimental design.

The CRBD linear model is used for local control. In this model, the treatments are applied randomly within blocks, and frequently this type of assay represents a restriction from randomization. Consequently, the fitted statistical model must take into consideration this type of control, i.e., the model should have the form

$$Y_{ij} = m + \tau_i + \beta_j + \epsilon_{ij}, \tag{3.16}$$

where Y_{ij} represents the observed value of the group that received treatment i in block j, m is the overall mean effect, τ_i is the effect of treatment i applied to the treated group, β_j is the effect of block j that is found in that group and $\epsilon_{ij} \sim \text{OLLS}(0, \sigma, \nu, \alpha)$ is the effect of the uncontrolled factors in the experimental group, for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, where I denotes the number of treatments and J the number of blocks.

Let $y_{11}, \ldots, y_{IJ}$ be a sample of size n from the OLLS distribution. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (m, \boldsymbol{\tau}^T, \boldsymbol{\beta}^T, \sigma, \nu, \alpha)^T$, where $\boldsymbol{\tau} =$

$(\tau_1, \ldots, \tau_I)^T$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_J)^T$, is given by

$$\begin{aligned} l(\boldsymbol{\theta}) &= IJ\log(\alpha) + \sum_{i=1}^{I}\sum_{j=1}^{J}\log\left[\phi_{t_\nu}\left(z_{ij}\right)\right] + \alpha\sum_{i=1}^{I}\sum_{j=1}^{J}\log\left\{\Phi_{T_\nu}\left(z_{ij}\right)\left[1-\Phi_{T_\nu}\left(z_{ij}\right)\right]\right\} \\ &\quad -2\sum_{i=1}^{I}\sum_{j=1}^{J}\log\left\{\Phi_{T_\nu}^{\alpha}\left(z_{ij}\right)+\left[1-\Phi_{T_\nu}\left(z_{ij}\right)\right]^{\alpha}\right\} - IJ\log(\sigma), \\ &\quad -\sum_{i=1}^{I}\sum_{j=1}^{J}\log\left\{\Phi_{T_\nu}\left(z_{ij}\right)\left[1-\Phi_{T_\nu}\left(z_{ij}\right)\right]\right\}, \end{aligned} \tag{3.17}$$

where $z_{ij} = \left(y_{ij} - m - \tau_i - \beta_j\right)/\sigma$.

The MLEs $\widehat{\boldsymbol{\theta}}$ of the model parameters can be obtained by maximizing the log-likelihood (3.17). In this case, we estimate $\boldsymbol{\theta}$ with ν fixed as described in sub-section 3.3. We use the Optim function in R to maximize the log-likelihood function. In addition, we obtain using the "L-BFGS-B" or "Nelder-Mead" methods the parameter estimates and their standard errors. We can provide upon request the data and the R script. The components of the score vector $U(\boldsymbol{\theta})$ for model (3.17) are given in Appendix C.

The asymptotic distribution of $\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}$ is multivariate normal $N_{I+J+3}(0, K(\boldsymbol{\theta})^{-1})$, where $K(\boldsymbol{\theta})$ is the information matrix. The asymptotic covariance matrix $K(\boldsymbol{\theta})^{-1}$ of $\widehat{\boldsymbol{\theta}}$ can be approximated by the inverse of the $(I+J+3)\times(I+J+3)$ observed information matrix $\mathbf{L}(\boldsymbol{\theta})$, i.e., we can use $\mathbf{L}(\boldsymbol{\theta})^{-1}$ to obtain an approximation for the large-sample covariance matrix of the MLEs. We provide confidence intervals for any parameters using the asymptotic normality of these estimates. Then, the inference for the parameter vector $\boldsymbol{\theta}$ can be based on the multiuvariate normal approximation $N_{I+J+3}(0, \mathbf{L}(\boldsymbol{\theta})^{-1})$ for $\widehat{\boldsymbol{\theta}}$ under the usual asymptotic theory.

The CRD model is the first special case of model (3.16) that follows when the local control effect $\beta_j = 0$ and the second model is the effect of the overall average that occurs when $\beta_j = 0$ and $\tau_i = 0$ for $i = 1, \ldots, I$ and $j = 1, \ldots, J$. By setting $\alpha = 1$ and $\nu > 30$, we obtain the usual standard CRBD model.

3.4.1 Sensitivity analysis

As a first tool for sensitivity analysis, the local influence method can be described for the CRDB linear model. The best known perturbation schemes are based on the local

influence in which the effects are studied by completely removing some cases from the analysis. This reasoning will form the basis for our local influence methodology and it shall be possible to determine which subjects might be influential for the analysis. This approach was suggested by (COOK, 1986), where instead of removing observations, weights are given to them. Local influence can be carried out for model (3.16). If likelihood displacement $LD(\boldsymbol{\omega}) = 2\{l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\}$ is used, where $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ denotes the MLE under the perturbed model, the normal curvature for $\boldsymbol{\theta}$ at the direction $\mathbf{d}, \| \mathbf{d} \| = 1$, is given by $C_{\mathbf{d}}(\boldsymbol{\theta}) = 2|\mathbf{d}^T \boldsymbol{\Delta}^T \left[\ddot{\mathbf{L}}(\boldsymbol{\theta})\right]^{-1} \boldsymbol{\Delta}\mathbf{d}|$. Here, $\boldsymbol{\Delta}$ is a $(I + J + 3) \times (I + J)$ matrix that depends on the perturbation scheme, whose elements are given by $\Delta_{vu} = \partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\theta_v \partial\omega_i$, $v = 1, 2, \ldots, I + J + 3$ and $u = 1, 2, \ldots, I + J$, evaluated at $\hat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}_0$, where $\boldsymbol{\omega}_0$ is the no perturbation vector.

We can also evaluate normal curvatures $C_{\mathbf{d}}(\alpha)$, $C_{\mathbf{d}}(m)$, $C_{\mathbf{d}}(\sigma)$, $C_{\mathbf{d}}(\boldsymbol{\tau})$ and $C_{\mathbf{d}}(\boldsymbol{\beta})$ to perform various index plots, for instance, the index plot of $\mathbf{d}_{max}$, the eigenvector corresponding to $C_{\mathbf{d}_{max}}$, the largest eigenvalue of the matrix $\mathbf{B} = -\boldsymbol{\Delta}^T \left[\ddot{\mathbf{L}}(\boldsymbol{\theta})\right]^{-1} \boldsymbol{\Delta}$ and the index plots of $C_{\mathbf{d}_u}(\alpha)$, $C_{\mathbf{d}_u}(m)$, $C_{\mathbf{d}_u}(\sigma)$, $C_{\mathbf{d}_u}(\tau)$ and $C_{\mathbf{d}_u}(\boldsymbol{\beta})$, called the total local influence see, for example, (LESAFFRE; VERBEKE, 1998), where $\mathbf{d}_u$ denotes an $n \times 1$ vector of zeros with one at the uth position. Thus, the curvature at direction $\mathbf{d}_i$ takes the form $C_i = 2|\boldsymbol{\Delta}_u^T \left[\ddot{\mathbf{L}}(\boldsymbol{\theta})\right]^{-1} \boldsymbol{\Delta}_u|$, where $\boldsymbol{\Delta}_u^T$ denotes the uth row of $\boldsymbol{\Delta}$. It is usual to point out those cases such that $C_u \geq 2\bar{C}$, where $\bar{C} = \frac{1}{n}\sum_{u=1}^{n} C_u$.

Next, we calculate, for two perturbation schemes, the matrix

$$\boldsymbol{\Delta} = \left(\Delta_{vu}\right)_{\left[(I+J+3)\times(I+J)\right]} = \left(\frac{\partial^2 l(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial\theta_v \partial\omega_u}\right)_{\left[(I+J+3)\times(I+J)\right]},$$

where $v = 1, 2, \ldots, I + J + 3$ and $u = 1, \ldots, I + J$, considering the model defined in (3.16) and its log-likelihood function given by (3.17).

3.4.2 Case-weight perturbation

Consider the vector of weights $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^T$. In this case, the perturbed log-likelihood function reduces to

$$\begin{aligned} l(\boldsymbol{\theta}|\boldsymbol{\omega}) &= [\log(\alpha) - \log(\sigma)] \sum_{i=1}^{I}\sum_{j=1}^{J} \omega_{ij} + \sum_{i=1}^{I}\sum_{j=1}^{J} \omega_{ij} \log\left[\phi_{t_\nu}\left(z_{ij}\right)\right] \\ &+(\alpha-1)\sum_{i=1}^{I}\sum_{j=1}^{J} \omega_{ij} \log\left\{\Phi_{T_\nu}\left(z_{ij}\right)\left[1-\Phi_{T_\nu}\left(z_{ij}\right)\right]\right\} \\ &-2\sum_{i=1}^{I}\sum_{j=1}^{J} \omega_{ij} \log\left\{\Phi_{T_\nu}^{\alpha}\left(z_{ij}\right)+\left[1-\Phi_{T_\nu}\left(z_{ij}\right)\right]^{\alpha}\right\}, \end{aligned}$$

where $z_{ij} = (y_{ij} - m - \tau_i - \beta_j)/\sigma$, $0 \leq \omega_i \leq 1$, $\boldsymbol{\omega}_0 = (1, \ldots, 1)^T$ and $i = 1, \ldots, I$ and $j = 1, \ldots, J$. The elements of $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha, \boldsymbol{\Delta}_m, \boldsymbol{\Delta}_\sigma, \boldsymbol{\Delta}_{\boldsymbol{\tau}}, \boldsymbol{\Delta}_{\boldsymbol{\beta}})^T$ can be evaluated numerically.

3.4.3 Response perturbation

Consider that each y_{ij} is perturbed as $y_{ijw} = y_{ij} + \omega_{ij} S_y$, where S_y is a scale factor that may be the estimated standard deviation of Y and $\omega_{ij} \in \mathbb{R}$.

Here, the perturbed log-likelihood function reduces to

$$\begin{aligned} l(\boldsymbol{\theta}|\boldsymbol{\omega}) &= IJ\log(\alpha) + \sum_{i=1}^{I}\sum_{j=1}^{J} \log\left[\phi_{t_\nu}\left(z_{ij}^*\right)\right] + \alpha\sum_{i=1}^{I}\sum_{j=1}^{J} \log\left\{\Phi_{T_\nu}\left(z_{ij}^*\right)\left[1-\Phi_{T_\nu}\left(z_{ij}^*\right)\right]\right\} \\ &-2\sum_{i=1}^{I}\sum_{j=1}^{J} \log\left\{\Phi_{T_\nu}^{\alpha}\left(z_{ij}^*\right)+\left[1-\Phi_{T_\nu}\left(z_{ij}^*\right)\right]^{\alpha}\right\} - IJ\log(\sigma), \\ &-\sum_{i=1}^{I}\sum_{j=1}^{J} \log\left\{\Phi_{T_\nu}\left(z_{ij}^*\right)\left[1-\Phi_{T_\nu}\left(z_{ij}^*\right)\right]\right\}, \end{aligned}$$

where $z_{ij} = \left(y_{ij}^* - m - \tau_i - \beta_j\right)/\sigma$ and $y_{ij}^* = y_{ij} + \omega_{ij} S_y$ for $i = 1, \ldots, I$ and $j = 1, \ldots, J$. The elements of $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_\alpha, \boldsymbol{\Delta}_m, \boldsymbol{\Delta}_\sigma, \boldsymbol{\Delta}_{\boldsymbol{\tau}}, \boldsymbol{\Delta}_{\boldsymbol{\beta}})^T$ can be evaluated numerically.

3.4.4 Residual analysis

When attempting to adjust a model to a dataset, the validation of the fit must be analyzed by a specific statistic, with the purpose of measuring the goodness-of-fit. Once the model is chosen and fitted, the analysis of the residuals is an efficient way to

check the model adequacy. The residuals also serve for other purposes, such as to detect the presence of aberrant points (outliers), identify the relevance of an additional factor omitted from the model and verify if there are indications of serious deviance from the distribution considered for the random error. Further, since the residuals are used to identify discrepancies between the fitted model and the dataset, it is convenient to define residuals that take into account the contribution of each observation to the goodness-of-fit measure used.

In summary, the residuals allow measuring the model fit for each observation and enable studying whether the differences between the observed and fitted values are due to chance or to a systematic behavior that can be modeled. The quantile residual, method proposed by (DUNN; SMYTH, 1996) can be defined as a measure of the discrepancy between y_{ij} and $\hat{\mu}_{ij}$, and it is given by

$$\widehat{rq}_{ij} = \Phi^{-1}\left\{\frac{\Phi_{T_\nu}^{\hat{\alpha}}\left(\frac{y_{ij}-\hat{m}-\hat{\tau}_i-\hat{\beta}_j}{\hat{\sigma}}\right)}{\Phi_{T_\nu}^{\hat{\alpha}}\left(\frac{y_{ij}-\hat{m}-\hat{\tau}_i-\hat{\beta}_j}{\hat{\sigma}}\right)+\left[1-\Phi_{T_\nu}\left(\frac{y_{ij}-\hat{m}-\hat{\tau}_i-\hat{\beta}_j}{\hat{\sigma}}\right)\right]^{\hat{\alpha}}}\right\}, \tag{3.18}$$

where $\Phi(\cdot)^{-1}$ is the inverse cumulative standard normal distribution.

Atkinson (1985) suggested the construction of envelopes to enable better interpretation of the probability normal plot of the residuals. These envelopes are simulated confidence bands that contain the residuals, such that if the model is well-fitted, the majority of points will be within these bands and randomly distributed. The construction of the confidence bands follows the steps:

- Fit the proposed model and calculate the residuals $\widehat{rq}_{ij}$'s;
- Simulate k samples of the response variable using the fitted model;
- Fit the model to each sample and calculate the residuals $\widehat{rq}_{ij}$, $i = 1, \ldots, I$ and $j = 1, \ldots, J$;
- Arrange each sample of IJ residuals in rising order to obtain $\widehat{rq}_{(ij)k}$ for $k = 1, \ldots, K$;

- For each ij, obtain the mean, minimum and maximum $\widehat{rq}_{(ij)k}$, namely

$$\widehat{rq}_{(ij)M} = \sum_{k=1}^{K} \frac{\widehat{rq}_{(ij)k}}{K}\,, \quad \widehat{rq}_{(ij)B} = \min\{\widehat{rq}_{(ij)k} : 1 \leq k \leq K\}$$

$$\text{and} \qquad \widehat{rq}_{(ij)H} = \max\{\widehat{rq}_{(ij)k} : 1 \leq k \leq K\}\,;$$

- Include the means, minimum and maximum together with the values of $\widehat{rq}_{ij}$ against the expected percentiles of the standard normal distribution.

The minimum and maximum values of $\widehat{rq}_{ij}$ form the envelope. If the model under study is correct, the observed values should be inside the bands and distributed randomly.

3.5 Applications

In this section, we provide three applications to real data to prove empirically the flexibility of the OLLS distribution.

- The first application involves data on production of soybeans in the municipality of Rio Verde, state of Mato Grosso, Brazil. These data show bimodality, so that the OLLS distribution is a good candidate to model them.

- The second dataset refers to the production of soybeans in the municipality of Iracemápolis, São Paulo state. The study was carried out by the company Produquimica, in Iracemápolis, São Paulo state, Brazil, during the 2014-2015 growing season. This application shows the flexibility of the OLLS distribution in relation to its submodels: OLLN, Student t and normal.

- The data from the third application refer to weight gain of livestock. The purpose of this application is to fit a linear CRBD model using the OLLS distribution. In this application, diagnostic analysis and examination of the residuals are presented.

In each case, the parameters are estimated by maximum likelihood using the optim function in the R software. First, we describe the datasets and provide the MLEs (and the corresponding standard errors in parentheses) of the parameters and the values of

the Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC) statistics. The lower the values of these criteria, the better the fit. Next, we perform likelihood ratio (LR) tests for the additional shape parameters. Further, we provide histograms of the datasets to show a visual comparison of the fitted density functions.

We motivate the paper by comparing the performances of the skew-normal, beta normal, Kumaraswamy normal and gamma normal models described below:

- **The skew normal (SN) distribution**

  The pdf is given by (for $x \in \mathbb{R}$)

$$f(x) = \frac{2}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left[\lambda\left(\frac{x-\mu}{\sigma}\right)\right], \tag{3.19}$$

  where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter and $\lambda \in \mathbb{R}$ is the parameter of asymmetry. The density function (3.19) holds for $x \in \mathbb{R}$ and it is symmetric if $\lambda = 0$ (Azzalini, 1985).

- **The beta normal (BN) distribution**

  The pdf is given by (for $x \in \mathbb{R}$)

$$f(x) = \frac{1}{\sigma B(a,b)}\phi\left(\frac{x-\mu}{\sigma}\right)\left\{\Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{a-1}\left\{1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{b-1}, \tag{3.20}$$

  where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter and $a > 0$ and $b > 0$ are shape parameters. Note that $B(a,b) = \Gamma(a)\,\Gamma(b)/\Gamma(a+b)$ is the beta function and $\Gamma(a) = \int_0^\infty t^{a-1}\,\mathrm{e}^{-\mathrm{t}}\mathrm{dt}$ is the gamma function. For $\mu = 0$ and $\sigma = 1$, we obtain the standard BN distribution. The properties of the BN distribution have been studied by some authors in recent years, for example, see Eugene *et al.* (2002), Cordeiro and de Castro (2011) and Cordeiro *et al.* (2012).

- **The Kumaraswamy normal (KN) distribution**

  The pdf is given by (for $x \in \mathbb{R}$)

$$f(x) = \frac{ab}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\left\{\Phi\left(\frac{x-\mu}{\sigma}\right)\right\}^{a-1}\left\{1-\Phi\left(\frac{x-\mu}{\sigma}\right)^{a}\right\}^{b-1}, \tag{3.21}$$

  where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter and $a > 0$ and

$b > 0$ are shape parameters. For $\mu = 0$ and $\sigma = 1$, we obtain the standard KN distribution.

- **The gamma normal (GN) distribution**

  Recently, Lima *et al.* (2015) proposed a new three-parameter distribution called the GN distribution with location parameter $\mu \in \mathbb{R}$, dispersion parameter $\sigma > 0$ and shape parameter $a > 0$, whose density (for $x \in \mathbb{R}$) is given by

$$f(x) = \frac{1}{\sigma\Gamma(a)}\,\phi\left(\frac{x-\mu}{\sigma}\right)\left\{-\log\left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]\right\}^{a-1}. \tag{3.22}$$

  The pdf (3.22) does not involve any complicated function and the normal distribution arises as the basic exemplar for $a = 1$. It is a positive point of the current generalization.

3.6 Application 1: soybean production

The first dataset refers to the production of soybeans in the municipality of Lucas do Rio Verde in the period from 1990 to 2012. This town is one of the 15 leading soybean producers in the Mato Grosso state. The data are the crop yields in kilograms of soybean per hectare (Kg/ha) obtained from the Brazilian Institute of Geography and Statistics, see, for details, in **http://www.sidra.ibge.gov.br/bda**.

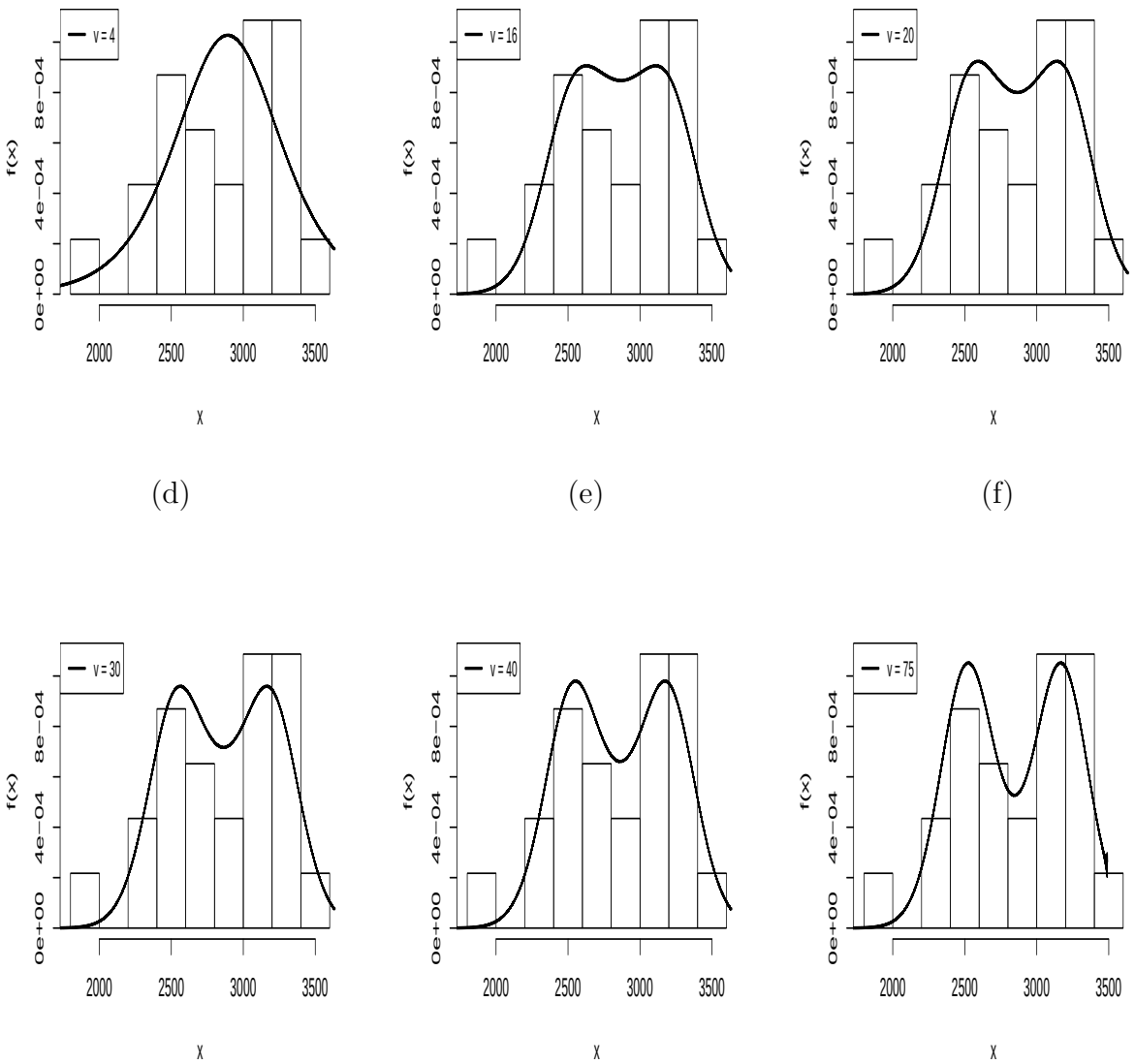
In order to compare the distributions, we consider some goodness-of-fit measures including AIC, CAIC and BIC statistics. The MLEs of the parameters μ, σ and α, for fixed ν, are obtained using the optim function in the R software. These estimates and the corresponding standard errors (SEs) are given in Table 3.2.

The means of the MLEs in Table 3.2 indicate that the values of the estimates of the comparison criteria, e.g., AIC (352.280, 348.767 and 347.771) decrease as the values of the parameter ν (4, 30 and 75) increase. Therefore, it is possible to model the bimodality of the data, because the parameter ν makes the kurtosis more flexible as shown in Figure 3.2.

Tabela 3.2 - MLEs and information criteria for soybean production data

$\nu = 4$			$\nu = 30$			$\nu = 75$		
$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE
μ	2892.402	96.155	μ	2862.814	82.667	μ	2844.953	0.501
σ	51301.353	5844.709	σ	119.889	72.128	σ	79.512	0.512
α	132.082	26.920	α	0.215	0.160	α	0.105	0.018
AIC	CAIC	BIC	AIC	CAIC	BIC	AIC	CAIC	BIC
352.280	355.809	355.687	348.767	352.297	352.174	347.771	351.300	351.177

(a) (b) (c)



(d) (e) (f)

Figura 3.2 - Estimated densities of the OLLS$(\mu, \sigma, \nu, \alpha)$ model for soybean production data for different values of $\nu = 4$, 16, 20, 30, 40 and 75.

3.7 Application 2: completely randomized design model - soybean data

The experiment was conducted at Block 3, of the Geraldo Schultz Research Center, located in the municipality of Iracemápolis, São Paulo state, with average altitude of 570 m (longitude 47° 30' 10.81"W and latitude 22° 38' 49.14"S). The climate in the region is classified as Cwa according to the Köppen classification (tropical highlands, with rain mainly in the summer and dry winters). The soil was classified as Dystrophic Red Latosol according to the Brazilian Soil Classification System (Rhodic Hapludox according to the Soil Taxonomy). The objective of the experiment was to assess the level of boron (B) in soybeans measured in mg/kg (Glycine max L. Merryl) under the effects of different treatments composed of the elements boron (B) and sulfur (S). The higher the concentration in the soil, the greater the uptake of these nutrients by the plants should be. The experimental design was completely randomized with 4 repetitions and 7 treatments. Each plot was composed of 6 rows with length of 7 m. The useful portion of the plot was composed of two rows with length of 5 m. The study was carried out by the company Produquimica, in Iracemápolis, São Paulo state, Brazil, during the 2014-2015 growing season.

- **Response variable** ($\mathbf{Y}_{ij}$): Concentration of B in grains of the soybeans (mg/kg); $i = 1, \ldots, 7$ and $j = 1, \ldots, 4$,
- **Treatments**:
  - **Treat 1** - Controle;
  - **Treat 2** - Sulfurgran (elemental S);
  - **Treat 3** - Sulfugran + Borosol (elemental S + boric acid);
  - **Treat 4** - Sufurgran + ActiveBor (elemental S + sodium octoborate);
  - **Treat 5** - Sulfurgran + Ulexite (elemental S + ulexite);
  - **Treat 6** - Sulfurgran + Produbor (elemental S + partially acidified ulexite);
  - **Treat 7** - Sulfurgran B-MAX (elemental S + ulexite in the same pellet).

The MLEs of the parameters (the SEs are in parentheses) and the values of the AIC, CAIC and BIC statistics for four fitted models are given in Table 3.3. Based on

these results, the three statistics for the OLLS model have the smallest values in relation to the OLLN, Student t and normal sub-models.

Tabela 3.3 - MLEs, SEs (in parentheses) and information criteria for soybean production data.

Soybean	μ	σ	ν	α	AIC	CAIC	BIC
OLLS	22.872	0.317	5	0.310	119.49	122.22	123.48
	(0.2447)	(0.1608)		(0.1098)			
OLLN	23.174	290.583	30	160.526	130.39	133.12	134.39
	(0.3551)	(4616.6968)		(2550.9318)			
Student t	23.073	1.491	5	1	125.49	127.23	128.16
	(0.3131)	(0.2844)					
Normal	23.646	2.631	30	1	137.63	139.37	140.29
	(0.497)	(0.3515)					

Table 3.4 lists the MLEs and SEs of the estimates of the parameters for the normal and OLLS regression models fitted to the current data using the optim function in the R software.

In the fitted model (2.35), a restriction on the solution is imposed, i.e., the effects of $\tau_1 = 0$ and $\beta_j = 0$. Thus, the estimates of the parameters of the treatments (τ_2, τ_3, τ_4, τ_5, τ_6 and τ_7) represent mean differences in relation to the treatment τ_1. This means that the interpretations should be carried out in relation to the treatment under the restriction τ_1 as shown in Table 3.4.

Tabela 3.4 - MLEs and their SEs, AIC and BIC for the normal and OLLS distributions.

	Normal				OLLS		
$\boldsymbol{\theta}$	MLE	SE	p-value	$\boldsymbol{\theta}$	EMVs	SE	p-value
m	23.400	0.986	<0.001	m	23.367	0.377	< 0.001
τ_2	0.625	1.395	0.659	τ_2	-1.797	0.473	< 0.001
τ_3	-1.250	1.395	0.380	τ_3	-1.245	0.429	< 0.001
τ_4	4.125	1.395	0.007	τ_4	2.636	0.510	< 0.001
τ_5	-1.625	1.395	0.257	τ_5	-1.482	0.508	< 0.001
τ_6	-0.150	1.395	0.915	τ_6	-0.002	0.480	> 0.050
τ_7	0.000	1.395	1.000	τ_7	-0.092	0.530	> 0.050
σ	0.679	0.133	<0.001	σ	1.161	0.079	< 0.001
α				α	0.279	0.099	< 0.001
	AIC	CAIC	BIC		AIC	CAIC	BIC
	133.643	146.584	144.3		103.774	120.274	115.764

Then, by using the OLLS distribution to explain the concentration of boron in

soybeans, we note that the Treatment 7 (Sulfurgran B-MAX) presents the smallest average difference (-0.0925), which is not statistically significant, while Treatment 4 (Sufurgran + ActiveBor) presents the largest average difference (2.6362), which is statistically significant. In practical terms, it can be said that the higher the concentration of boron, the better the quality of soybeans. Table 3.5 shows all the comparisons between the treatments. These comparisons are carried out by means of confidence intervals at the 5% significance level. The asterisks (*)indicate the existence of a statistically significant difference between average differences estimated for the treatments, at the 5% level, while (ns) indicates no significant difference.

Tabela 3.5 - Results of the comparison of the 7 treatments for soybean data.

Hypotheses	Estimates	Lwr	Upr	Hypotheses	Estimates	Lwr	Upr
$H_0 : \tau_2 - \tau_1 = 0$	-1.797^{*}	-2.419	-1.176	$H_0 : \tau_4 - \tau_3 = 0$	4.429^{*}	3.838	5.019
$H_0 : \tau_3 - \tau_1 = 0$	-1.245^{*}	-1.809	-0.682	$H_0 : \tau_5 - \tau_3 = 0$	$0.085^{\mathbf{ns}}$	-0.604	0.776
$H_0 : \tau_4 - \tau_1 = 0$	2.636^{*}	1.966	3.306	$H_0 : \tau_6 - \tau_3 = 0$	1.550^{*}	0.826	2.274
$H_0 : \tau_5 - \tau_1 = 0$	-1.482^{*}	-2.148	-0.815	$H_0 : \tau_7 - \tau_3 = 0$	1.557^{*}	0.545	2.568
$H_0 : \tau_6 - \tau_1 = 0$	$-0.002^{\mathbf{ns}}$	-0.633	0.628	$H_0 : \tau_5 - \tau_4 = 0$	-4.283^{*}	-5.125	-3.441
$H_0 : \tau_7 - \tau_1 = 0$	$-0.092^{\mathbf{ns}}$	-0.788	0.603	$H_0 : \tau_6 - \tau_4 = 0$	-2.618^{*}	-3.336	-1.900
$H_0 : \tau_3 - \tau_2 = 0$	$0.393^{\mathbf{ns}}$	-0.609	1.396	$H_0 : \tau_7 - \tau_4 = 0$	-2.522^{*}	-3.266	-1.779
$H_0 : \tau_4 - \tau_2 = 0$	$0.330^{\mathbf{ns}}$	-0.584	1.245	$H_0 : \tau_6 - \tau_5 = 0$	$-0.923^{\mathbf{ns}}$	-1.941	0.095
$H_0 : \tau_5 - \tau_2 = 0$	4.084^{*}	3.023	5.146	$H_0 : \tau_7 - \tau_5 = 0$	$0.515^{\mathbf{ns}}$	-0.446	1.476
$H_0 : \tau_6 - \tau_2 = 0$	1.434^{*}	0.499	2.368	$H_0 : \tau_7 - \tau_6 = 0$	0.942^{*}	0.315	1.568
$H_0 : \tau_7 - \tau_2 = 0$	1.579^{*}	0.640	2.518				

3.7.1 Application 3: weight gain in animals

This dataset comes from a CRBD with five treatments (substitution of a feed increment of 0%, 5%, 10%, 15% and 20%) and four blocks. The data are available at Professor Euclides Malheiros Braga's website: http://jaguar.fcav.unesp.br/euclides/. Choose the year 2013 and the option *Estatística Experimental-PG in Ciências Animal, Universidade Federal do Semi-Árido (UFERSA), Mossoró, Rio Grande do Norte (Brazil)* and download the A DBC ex2.txt file. The original sample size is $n = 20$ observations.

Table 3.6 summarizes the main descriptive statistics (mean, median and standard deviation) for each treatment (1, 2, 3, 4 and 5). It can be noted that the mean values are higher than the medians for all treatments and then we have a positive asymmetric distribution, and then the normal distribution is not adequate to fit these data. This

can be observed better in Figure 3.3, which shows an increase in the statistical values of the treatments 2, 3, 4 and 5 in relation to the first treatment. Another fact to be evaluated, both in Table 3.6 and Figure 3.3, is the standard deviation values, which indicate homogeneity of the variance within the treatments.

The weight gain of the animals in all five treatments is studied by applying the following CRBD linear model (see Section 3.4)

$$y_{ij} = m + \tau_i + \beta_j + \epsilon_{ij} \qquad \begin{cases} i = 1, \ldots, 5 \\ j = 1, \ldots, 4; \end{cases} \tag{3.23}$$

- y_{ij}: denotes the animal's weight gain in the group that received treatment i, observed in block j, with $i = 1, \ldots, 5$ and $j = 1, \ldots, 4$;
- m: indicates the effect of the average population weight gain;
- τ_i: denotes the effect of treatment i applied in the group, with $i = 1, \ldots, 5$, i.e., substitution of the feed ingredient in the proportion of 0%, 5%, 10%, 15% or 20%;
- β_j : is the effect of block j in which the group is found, with $j = 1, \ldots, 4$, i.e., the local control;
- ϵ_{ij}: is the random error, i.e., the effect of the factor not controlled in the experiment;
- In this case, we assume that $\epsilon_{ij} \sim \text{OLLS}(0, \sigma, \nu, \alpha)$.

The interest is to estimate the CRBD linear model for each of the five treatments. In other words, after the exploratory analysis of the data, a priori for each of the five treatments, the interest is on the expected value of model (3.23), which is obtained for each of the five treatments. Therefore, the expected values of the CRBD linear model for each treatment have the following forms: **Trat i (c)** : $E(y_{1j}) = m + \tau_i + \beta_j,\ i = 1, \ldots, 4;,\ j = 1, \ldots, 4;$ with concentrations $c = 0\%$, 5%, 10% ,15% and 20%.

Table 3.7 provides the MLEs of the parameters for the fitted normal, OLLS, SN, BN, KN and GN models, the corresponding SEs and the AIC, CAIC and BIC statistics. For the OLLS model, we used $\nu = 4$. Table 3.8 reports the LR statistics and descriptive levels. To fit model (2.35), we impose two constraints on the solution, i.e., we consider

Tabela 3.6 - Descriptive statistics for each treatment to the weight gain data.

Statistics	Trat 1	Trat 2	Trat 3	Trat 4	Trat 5
Average	56.57	62.67	65.97	63.12	58.12
Median	55.10	60.85	63.45	60.75	56.00
Standard deviation	8.87	9.19	10.06	8.24	8.56
Coefficient of variation	15.69	14.67	15.26	13.06	14.74

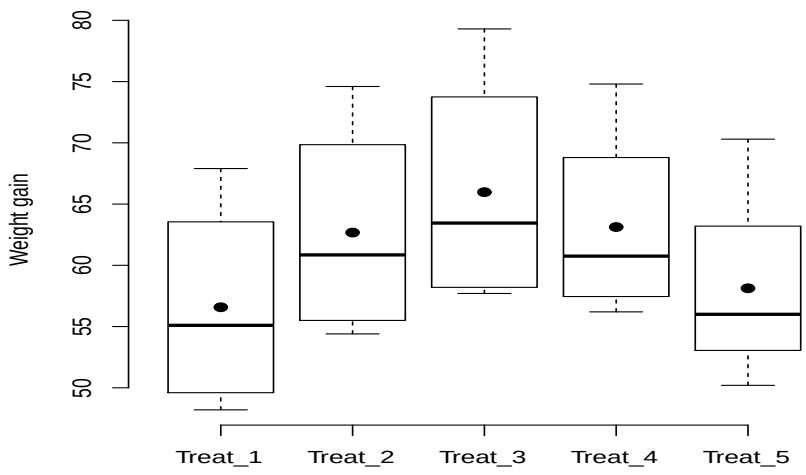


Figura 3.3 - Boxplot for each treatment in which the • is the average.

the effect of treatment $\tau_1 = 0$ and of block $\beta_1 = 0$. Consequently, the estimates of the treatment parameters must be interpreted in relation to the treatment 1. The estimates of the parameters of the six distributions (normal, OLLS, SN, BN, KN and GN) are consistent with the results shown in Figure 3.3, since as the substitution of the feed ingredient increases (5%, 10%, 15% and 20%), the estimates of the treatments (τ_2, τ_3, τ_4, τ_5) increase in relation to the treatment τ_1 for the OLLS model with heavier tails and for the SN, BN, KN and GN asymmetric distributions, as can be seen in Table 3.7. It can also be noted from Table 3.7 that the estimated values $\hat{m} = 49.799$, $\hat{m} = 47.976$ and $\hat{m} = 49.756$ for the normal, OLLS and SN distributions and the values $\hat{m} = 52.628$, $\hat{m} = 59.979$ and $\hat{m} = 51.329$ for the BN, KN and GN distributions are practically equal, respectively. The same fact can be seen for the effects of the treatments τ_2, τ_3, τ_4 and τ_5, mainly for the normal, SN, BN and KN distributions. However, the results of the estimates of the SEs for the OLLS distribution are lower than those for the normal, SN, BN, KN and GN distributions.

Tabela 3.7 - MLEs for the parameters of the CRBD linear model fitted to the weight gain data.

Normal			OLLS			SN		
$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE
m	49.799	1.593	m	47.976	0.656	m	49.756	7.037
τ_2	6.100	1.782	τ_2	7.200	0.810	τ_2	6.101	1.782
τ_3	9.400	1.782	τ_3	9.742	0.864	τ_3	9.401	1.782
τ_4	6.550	1.782	τ_4	7.726	0.612	τ_4	6.551	1.782
τ_5	1.550	1.782	τ_5	4.406	1.074	τ_5	1.551	1.782
β_2	1.660	1.593	β_2	3.293	1.104	β_2	1.660	1.593
β_3	6.580	1.593	β_3	9.881	1.252	β_3	6.580	1.593
β_4	18.860	1.593	β_4	19.370	1.210	β_4	18.860	1.593
σ	2.520	0.398	σ	0.214	0.125	σ	2.521	0.415
			α	0.285	0.126	λ	0.022	3.410
AIC	CAIC	BIC	AIC	CAIC	BIC	AIC	CAIC	BIC
111.731	144.731	120.693	101.169	145.741	111.126	113.731	158.303	123.689
BN			KN			GN		
$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE
m	52.628	2.527	m	59.979	5.278	m	51.329	1.747
τ_2	5.944	1.744	τ_2	6.425	1.393	τ_2	5.878	1.855
τ_3	9.057	1.740	τ_3	9.880	1.398	τ_3	10.269	1.940
τ_4	6.242	1.760	τ_4	6.808	1.409	τ_4	6.173	1.861
τ_5	2.349	2.232	τ_5	3.835	1.496	τ_5	3.819	1.521
β_2	2.355	1.697	β_2	1.444	1.262	β_2	3.478	1.732
β_3	8.029	1.569	β_3	7.880	1.365	β_3	8.806	1.306
β_4	19.251	1.698	β_4	18.643	1.309	β_4	20.318	1.701
σ	2.081	1.292	σ	0.934	0.885	σ	1.685	0.003
a	0.341	0.412	a	0.043	0.076	a	0.220	0.075
b	2.298	2.610	b	14.593	25.655			
AIC	CAIC	BIC	AIC	CAIC	BIC	AIC	CAIC	BIC
113.214	157.785	123.171	113.761	158.333	123.719	114.271	158.842	124.228

Tabela 3.8 - LR tests to the weight gain data.

Models	Hypotheses	LR statistic	p-value
Normal vs OLLS	$H_0 : \alpha = 1 \ e \ \nu = \infty$ vs $H_1 : H_0$ is false	12.562	< 0.001

Besides this, the 5% descriptive level in Table 3.8, reveals that there is a statistical difference between the OLLS and normal distributions. Therefore, after analyzing the estimates of the parameters, SEs and comparison criteria, we can use the OLLS distribution to explain the data.

Therefore, based on this distribution, among the treatments studied, treatment 3 presents the highest value (9.742), followed by treatment 4, treatment 2 and treatment 5 with values of 7.726, 7.200 and 4.406, respectively.

In order to perform other hypothesis tests that the researcher might be interested, we can apply the technique of orthogonal contrasts for possible differences between means of treatments.

Sensibility and residual analysis

To detect possible influential observations in the OLLS regression model fitted to the CRBD data, we investigate the local influence using computational routines in the Ox 6.01 matrix language. Figure 3.4 displays plots of the local influence by considering perturbations of cases ($C_{\mathbf{d}_{max}} = 0.0042$) and the response variable perturbation ($C_{\mathbf{d}_{max}} = 1.2052$). These plots indicate that the points ♯11 and ♯18 are possible influential observations.

(a) (b)

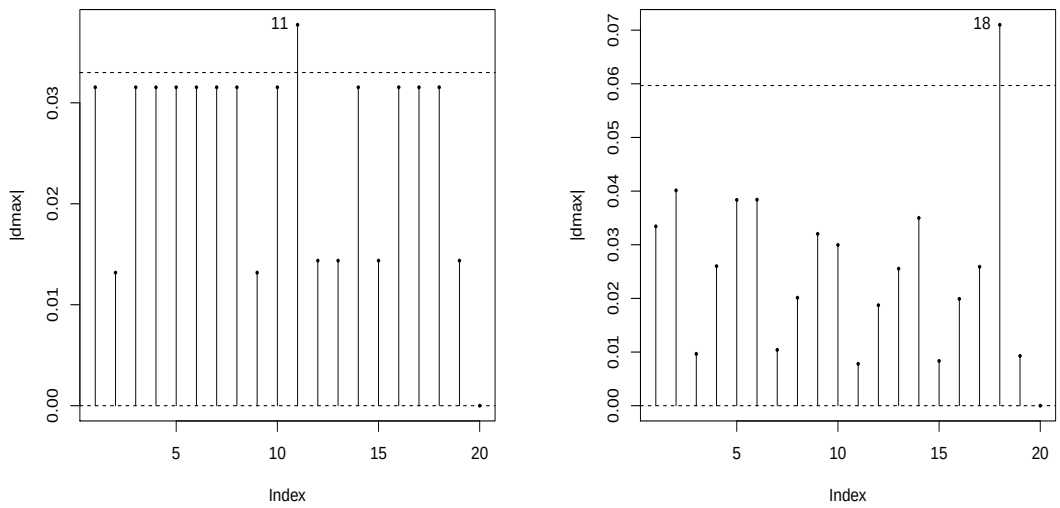


Figura 3.4 - Index plot of $\mathbf{d}_{max}$ from the CRBD linear model fitted to the weight gain data. (a) Case-weight perturbation. (b) Response variable perturbation.

Figure 3.5 displays the plots of the residuals to detect possible influential observations in the OLLS regression model. Figure 3.5,(a) indicates the estimated quantile residuals ($\hat{rq}_{ij}$) versus the adjusted values and Figure 3.5(b) gives the simulated envelopes. It can be seen that the quantile residuals are well distributed around zero and do not have problems of outliers. In the plots of Figure 3.5(b), all the estimated points are within the

simulated envelope, thus indicating that the OLLS distribution provides a good fit to the CRBD dataset. Therefore, based on the sensitivity and residual analysis, we conclude that the cases ♯11 and ♯18 cannot be considered influential observations.

(a) (b)

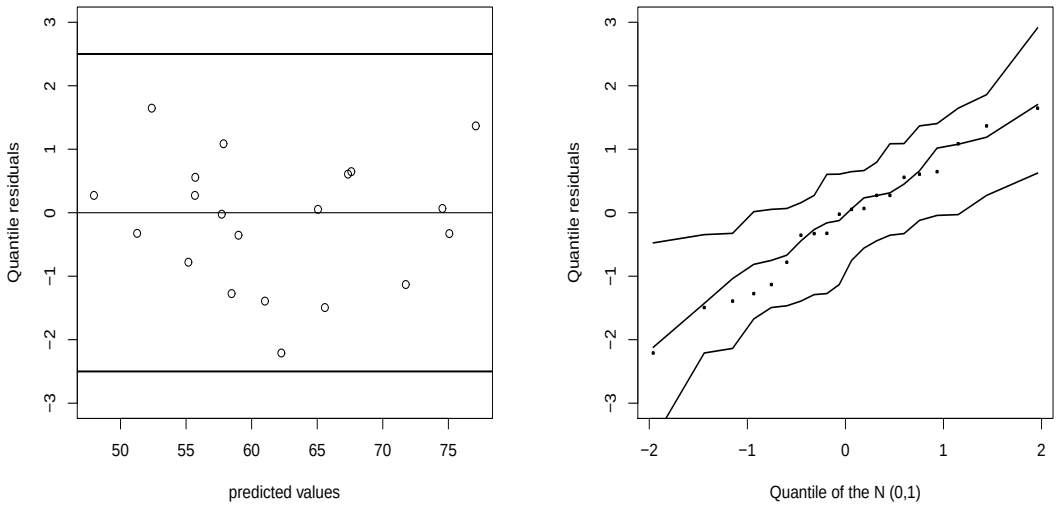


Figura 3.5 - (a) Quantile residuals for the CRBD linear model fitted to the weight gain data. (b) Normal probability plot for the deviance residuals with envelopes.

3.8 Concluding remarks

In this study, we propose a new four-parameters distribution called the odd log-logistic t-Student (OLLS). This distribution is successfully proposed in the study of data from experimental design. The proposed distribution is more versatile than the Student's t distribution, since it can be adjusted to bimodal data. Another characteristic is that the new distribution is analytically tractable and also allows testing for some sub-models: odd log-logistic normal, odd log-logistic Cauchy, Student-t and normal. Further, we obtain some mathematical properties of the distribution such as quantile function and useful expansions. For estimating the parameters we use the maximum likelihood estimation method. The flexibility of the new model is examined by means of two applications to real datasets, one of production of soybeans and other from experimental designs. The new distribution presents good results when compared to the skew-normal (SN), beta normal (BN), Kumaraswamy normal (KN) and gamma normal (GN) models. To verify

the quality of the adjustment we carry out sensitivity and residual analysis. The model can be used in practical situations for as a completely randomized designs or completely randomized blocks designs.

References

ALEXANDER, C.; Cordeiro, G. M.; Ortega, E. M. M.; Sarabia, J.M.(2012). Generalized beta-generated distributions. **Computational Statistics and Data Analysis**, **56**, 1880-189.

ALZAATREH, A.; FAMOYE, F.; LEE, C. (2014). The gamma-normal distribution: Properties and applications. **Computational Statistics and Data Analysis**, **69**, 67-80.

ALZAATREH, A.; LEE, C.; FAMOYE, F. (2013). A new method for generating families of continuous distributions. **Metron**, **71(1)**, 63-79.

ALIZADEH, M.; EMADI, M.; DOOSTPARAST, M., CORDEIRO, G. M., ORTEGA, E.M.M., & PESCIM, R. R.(2015). A new family of distributions: the kumaraswamy odd log-logistic, properties and applications. **Hacettepa Journal of Mathematics and Statistics (to appear)**.

ARELLANO-VALLE, R. B.; GENTON, M. G. (2005). On fundamental skew distributions. **Journal of Multivariate Analysis, 96**, 93-116.

ATKINSON, A.C. (1985). **Plots, transformations and regression: an introduction to graphical methods of diagnostics regression analysis**. 2nd ed.Oxford: Clarendon Press.

AZZALINI, A. (1985). A class of distributions which includes the normal ones. **Scandinavian Journal of Statistics**, **12**, 171-178.

CORDEIRO, G. M.; CINTRA, R. J.; RÊGO, L. C.; ORTEGA, E. M. M. (2012). The McDonald normal distribution. **Pakistan Journal of Statistics and Operation Research**, **8**, 301-329.

CORDEIRO, G. M.; DE CASTRO, M. (2011). A New family of generalized distribution. **Journal of Statistical Computation and Simulation**, 81, 883-893.

COOK, R. D. (1986). Assessment of local influence (with discussion). **Journal of the Royal Statistical Society B**, **48**, 133-169.

DOORNIK, J. A. (2006). **An Object-Oriented Matrix Language - Ox 4**. 5th ed. Timberlake Consultants Press, London.

DUNN. P. K.; SMYTH G. K. (1996). Randomized quantile residuals. **Journal Computational and Graphical Statistics**, **5**, 236-244.

EUGENE, N.; LEE, C.; FAMOYE, F. (2002). Beta-normal distribution and its applications. **Communications in Statistics - Theory and Methods**, 31, 497-512.

GARAY, A. M.; LACHOS, V. H.; LABRA, F. V.; ORTEGA, E. M. M. (2014). Statistical diagnostics for nonlinear regression models based on scale mixtures of skew-normal distributions. **Journal of Statistical Computation and Simulation**, **84**, 1761-1778.

GRADSHTEYN, I. S.; RYZHIK, I. M. (2000). **Table of Integrals, Series, and Products**, sixth edition. Academic Press, San Diego.

IBGE. Instituto Brasileiro de Geografia e Estatística. Accessed June 25, 2014 Available in: http: // www.sidra.ibge.gov.br/bda/tabela/listabl.asp?c=99 &z=t &o=11. Available in: http: // www.sidra. ibge. gov. br/bda/tabela/listabl. asp?c = 99 & z=t &o=11.

JONES, M .C.; FADDY, M. J. (2003). A skew extension of the t-distribution with applications. **Journal of the Royal Statistical Society. Series B (Statistical Methodology)**, **65**, 159-174.

LABRA, F. V.; GARAY, A. M.; LACHOS, V.H.; ORTEGA, E. M. M. (2012). Estimation and diagnostics for heteroscedastic nonlinear regression models based on scale mixtures of skew-normal distributions. **Journal of Statistical Planning and Inference** , **142**, 2149-2165.

LESAFFRE, E.; VERBEKE, G. (1998). Local influence in linear mixed models. **Biometrics**, **54**, 570-582.

LIN, J. G.; XIE, F. C.; WEI, B. C. (2009). Statistical diagnostics for skew-t-normal nonlinear models. **Communications in Statistics - Simulation and Computation**, **38**, 2096-2110.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-real data. **IEEE Transaction on Reliability**, 42, 299-302.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; FRIEMER, M. (1995). The exponential Weibull family: A reanalysis of the bus-motor failure data. **Technometrics**, **37**, 436-445.

NADARAJAH, S.; GUPTA, A. K. (2007). A generalized gamma distribution with application to drought data. **Mathematics and Computer in Simulation**, **74**, 1-7.

NADARAJAH, S.; KOTZ, S. (2006). The exponentiated type distribution. **Acta Applicandae Mathematicae**, **92**, 97-111.

SHAW, W. T. (2006). Sampling Student′s T distribution - use of the inverse cumulative distribution function. **Journal of Computational Finance**, **9**, 37-73.

SHIRKE, D. T.; KAKADE, C. S. (2006). On exponentiated log-normal distribution. **International Journal of Agricultural and Statistical Science**, **2**, 319-326.

4 THE NEW DISTRIBUTION SKEW-BIMODAL WITH APPLICATION IN ANALYSIS OF EXPERIMENTS

Abstract

The modeling and analysis of experiments is an important aspect of statistical work in a wide variety of scientific and technological fields. We introduce and study the odd log-logistic skew-normal model, which can be interpreted as a generalization of the skew-normal distribution. The new distribution contains, as special sub-models, several important models discussed in the literature, such as the normal, skew-normal and odd log-logistic normal distributions. The new distribution can be used effectively in the analysis of experiments data since it accommodates unimodal, bimodal, symmetric, bimodal and right-skewed and bimodal and left-skewed density function depending on the parameter values. We derive some structural properties of the proposed model. The method of maximum likelihood is used to estimate its parameters. Besides, residual analysis with generated envelopes is performed to select appropriate models. We illustrate the importance of the new model by means of real data sets in analysis of experiments carried out in different regions of Brazil. We can conclude that the proposed model can give more realistic fits than other special models.

Keywords: Log-logistic distribution; Maximum likelihood estimation; Mean deviation; Skew distribution; Skew bimodality

4.1 Introduction

The normal distribution is often used in many different areas to model data with symmetric distributions. However, as is widely known, many phenomena cannot always be modeled by the normal distribution, whether by the lack of symmetry or the presence of atypical values. In past decades even when the phenomenon of interest did not present responses for which the assumption of normality was reasonable, the reaction was to try to find some transformation so that the data had at least some semblance of symmetric behavior. The best known of these transformations was the one proposed by (BOX; COX, 1964).

In recent years, many other distributions have been proposed as alternatives for this type of problem. The elliptical distributions (FANG; KOTZ; NG, 1990) are perhaps the best known among the proposals to preserve the symmetric structure of the Gaussian distribution and allow for heavier or lighter tails than the normal distribution. This optimal property has recently been considered in many studies; see, for example,

(CYSNEIROS; PAULA; GALEA, 2007).

Although this new class of models presents good alternatives to the normal distribution, some of them are not suitable when the distribution of the data or errors of the model are asymmetric (HILL; DIXON, 1982). In this context, the skew-elliptical distributions have been used successfully, making this new family a visible alternative to model data sets that have asymmetric behavior. The skew-normal (SN) distribution, introduced by (AZZALINI, 1985), is perhaps the pioneer of this new modeling strategy, in which the normal distribution is a special case.

Many authors have stressed the importance of using more flexible models in this type of modeling. Among the proposed distributions in this respect, the following stand out: skew-Cauchy (ARNOLD; BEAVER, 1993), skew-slash (WANG; GENTON, 2006), skew-slash-t (PUNATHUMPARAMBATH, 2012), the skew-t (SAHU; DEY; BRANCO, 2003), (GUPTA, 2003) and elliptic-skew (AZZALINI; CAPITANIO, 1999), (BRANCO; DEY, 2001), (WANG; BOYER; GENTON, 2004). However, among these classes, a few include the normal distribution as a sub-model.

When the observed data have asymmetry and bi-modality, the skew-normal distribution is not appropriate. Therefore, we propose a more general distribution as an alternative, which can model asymmetry and bi-modality and include as special cases the normal, skew-normal and odd log-logistic normal distributions. The new distribution is called the *odd log-logistic skew-normal* (OLLSN) model.

Regression models can be investigated in different forms in analysis of experiments. In this chapter, we also propose a regression model based on the OLLSN distribution. The inferential part is carried out using the asymptotic distribution of the maximum likelihood estimators (MLEs). The performance of these estimators is verified by means of a simulation study.

The assessment of the fitted model is an important part of data analysis, particularly in regression models, and residual analysis is a helpful tool to validate the fitted model. For example, examination of the residuals can be used to detect the presence of outlying observations, the absence of components in the systematic part of the model and departures from the error and variance assumptions. We define appropriate residuals to

detect influential observations in the OLLSN regression model.

This chapter is organized as follows. In Section 4.2, we introduce the OLLSN model and obtain its quantile function (qf). In Section 4.3, we provide structural properties of the OLLSN distribution. In Section 4.4, we derive explicit expressions for the ordinary and incomplete moments and generating function of the OLLSN distribution. Some inferential tools are discussed in Section 4.5 and the performance of the MLEs. In Section 4.6, we present the completely randomized design model based on the OLLSN distribution. In Section 4.7, a kind of quantile residual is proposed to assess departures from the underlying OLLSN distribution in linear models to completely randomized design and to detect outliers. In Section 4.8, we prove empirically the potentiality of the new model by means of two real data. Finally, Section 4.9 ends with some conclusions.

4.2 The OLLSN model

Statistical distributions are very useful in describing and predicting real world phenomena. Numerous extended distributions have been extensively used over the last decades for modeling data in several areas. Recent developments focus on defining new families that extend well-known distributions and at the same time provide great flexibility in modeling data in practice. So, several classes to generate new distributions by adding one or more parameters have been proposed such as the (MARSHALL; OLKIN, 1997), beta-G by (EUGENE; LEE; FAMOYE, 2002), Kumaraswamy-G (Kw-G) by (CORDEIRO; CASTRO, 2011), McDonald normal distribution by (CORDEIRO *et al.*, 2012), generalized beta-generated distributions by (ALEXANDRE *et al.*, 2012), Weibull-G by (BOURGUIGNON; RODRIGO; GAUSS, 2014), exponentiated half-logistic by (CORDEIRO; MORAD; ORTEGA, 2014a), logistic-X by (TAHIR *et al.*, 2016) and Lomax generator by (CORDEIRO *et al.*, 2014b).

Let $g(x)$ and $G(x)$ be the probability density function (pdf) and cumulative distribution function (cdf) of the skew-normal (for $x \in \mathbb{R}$) model given by

$$g(z;\lambda) = \frac{2}{\sigma}\phi\left(z\right)\Phi\left(\lambda\, z\right) \tag{4.1}$$

and

$$G(z;\lambda) = \Phi\left(z\right) - 2\,T\left(z;\lambda\right) = \Phi_{SN}\left(z;\lambda\right), \tag{4.2}$$

where $z = (x-\mu)/\sigma$, $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a dispersion parameter, and $T(z;\lambda)$ is the Owen's function given by (for for $z, \lambda \in \mathbb{R}$)

$$T(z;\lambda) = (2\pi)^{-1} \int_0^{\lambda} \frac{\exp\left\{-\frac{1}{2}z^2(1+x^2)\right\}}{1+x^2} dx.$$

The parameter λ regulates the skewness and it varies in $\mathbb{R}$. For $\lambda = 0$, we have the $N(\mu, \sigma^2)$ density. The function $\Phi_{SN}(\cdot)$ denotes the standard skew-normal cdf.

Based on the transformer odd log-logistic generator (GLEATON; LYNCH, 2006), we propose a new continuous distribution called the OLLSN model by integrating the log-logistic density function. It has cdf given by

$$F(z;\lambda,\alpha) = \int_0^{\frac{\Phi_{SN}(z;\lambda)}{\bar{\Phi}_{SN}(z;\lambda)}} \frac{\alpha\, t^{\alpha-1}}{(1+t^{\alpha})^2} dt = \frac{\Phi_{SN}^{\alpha}\left(z;\lambda\right)}{\Phi_{SN}^{\alpha}\left(z;\lambda\right) + \left[1-\Phi_{SN}\left(z;\lambda\right)\right]^{\alpha}}, \tag{4.3}$$

where $\alpha > 0$ is an extra shape parameter.

The pdf corresponding to (5.3) is given by

$$f(z;\lambda,\alpha) = \frac{2\alpha\phi\left(z\right)\Phi\left(\lambda\, z\right)\Phi_{SN}^{\alpha-1}\left(z;\lambda\right)\left[1-\Phi_{SN}\left(z;\lambda\right)\right]^{\alpha-1}}{\sigma\left\{\Phi_{SN}^{\alpha}\left(z;\lambda\right) + \left[1-\Phi_{SN}\left(z;\lambda\right)\right]^{\alpha}\right\}^2}. \tag{4.4}$$

So, we prove that the OLLSN distribution is symmetric about 0 for $\lambda = 0$, then the parameters σ and α characterize the kurtosis and skew-bi-modality of this distribution.

Hereafter, a random variable Z with density function (5.4) is denoted by $Z \sim$ OLLSN$(0, 1, \lambda, \alpha)$. Clearly, the random variable $X = \mu + \sigma Z$ follows the OLLSN$(\mu, \sigma^2, \lambda, \alpha)$ distribution.

The density function (5.4) allows greater flexibility of its tails and can be widely applied in many areas of engineering and biology. The normal distribution is a special case of (5.4) when $\lambda = 0$ and $\alpha = 1$. If $\lambda \neq 0$ and $\alpha = 1$, we obtain the SN distribution

and, for $\lambda = 0$ and $\alpha \neq 1$, it reduces to the OLLN distribution (BRAGA *et al.*, 2016), respectively.

We can write

$$\alpha = \frac{\log\left\{\Phi_{SN}^{\alpha}(z;\lambda)/[1-\Phi_{SN}(z;\lambda)]^{\alpha}\right\}}{\log\left\{\Phi_{SN}(z;\lambda)/[1-\Phi_{SN}(z;\lambda)]\right\}},$$

and then the parameter α represents the quotient of the log odds ratio for the generated and baseline distributions.

Equation (5.3) has tractable properties specially for simulations, since the qf of Z has a simple form. Let $F(z;\lambda,\alpha) = u$ and $\Phi_{SN}^{-1}(z;\lambda)$ be the inverse function of $\Phi_{SN}(z;\lambda)$. We have

$$Q(u) = Q_{SN}\left[h(u,\alpha);\lambda\right], \tag{4.5}$$

where $u \sim U(0,1)$ and $Q_{SN}[h(u,\alpha);\lambda] = \Phi_{SN}^{-1}\left[h(u,\alpha);\lambda\right]$ is the qf of the SN distribution at

$$h(u,\alpha) = u^{\frac{1}{\alpha}}\left[u^{\frac{1}{\alpha}} + (1-u)^{\frac{1}{\alpha}}\right]^{-1}.$$

It is not possible to study the behavior of the parameters of the OLLSN distribution by taking derivatives. We can verify skew-bi-modality of the new distribution in the plots of Figures 4.1 and 4.2 by combining some values of λ, α and μ. Figure 4.1 reveals different types of bi-modality, whereas Figure 4.2a and 4.2b reveal the different types of asymmetrical bi-modality.

Let $f(z;0,\alpha)$ be the density of the OLLN$(0,1,\alpha)$ distribution. By using $\Phi(-z) = 1-\Phi(z)$, we have

$$\begin{aligned} f(-z;0,\alpha) &= \frac{2\alpha}{\sqrt{2\pi}}\frac{\Phi^{\alpha-1}(-z)\left[1-\Phi(-z)\right]^{\alpha-1}\exp(-z^2/2)}{\left\{\Phi(-z)^{\alpha}+\left[1-\Phi(-z)\right]^{\alpha}\right\}^2} \\ &= \frac{2\alpha}{\sqrt{2\pi}}\frac{\Phi^{\alpha-1}(z)\left[1-\Phi(z)\right]^{\alpha-1}\exp(-z^2/2)}{\left\{\Phi(z)^{\alpha}+\left[1-\Phi(z)\right]^{\alpha}\right\}^2} \\ &= f(z;0,\alpha). \end{aligned}$$

So, we prove that the OLLSN distribution is symmetric about 0 for $\lambda = 0$, then

(a) (b) (c)

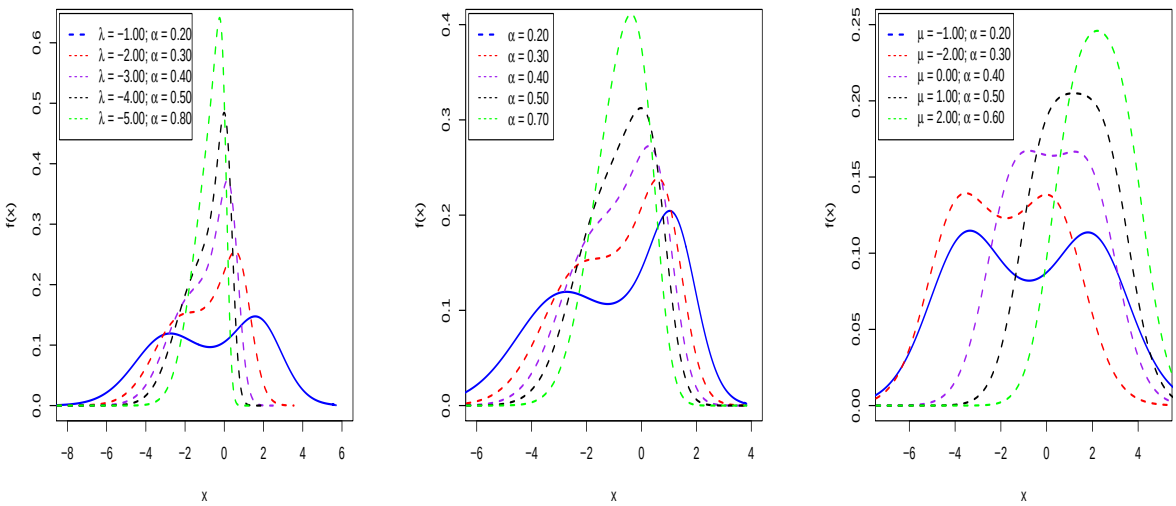


Figura 4.1 - Plots of the OLLSN density function for some parameter values. (a) For different values of the λ and α with $\mu = 0$ and $\sigma = 1$. (b) For different values of the α with $\lambda = -1.8$, $\mu = 0$ and $\sigma = 1$. (c) For different values of the μ and α with $\lambda = 0.3$ and $\sigma = 1$.

(a) (b) (c)

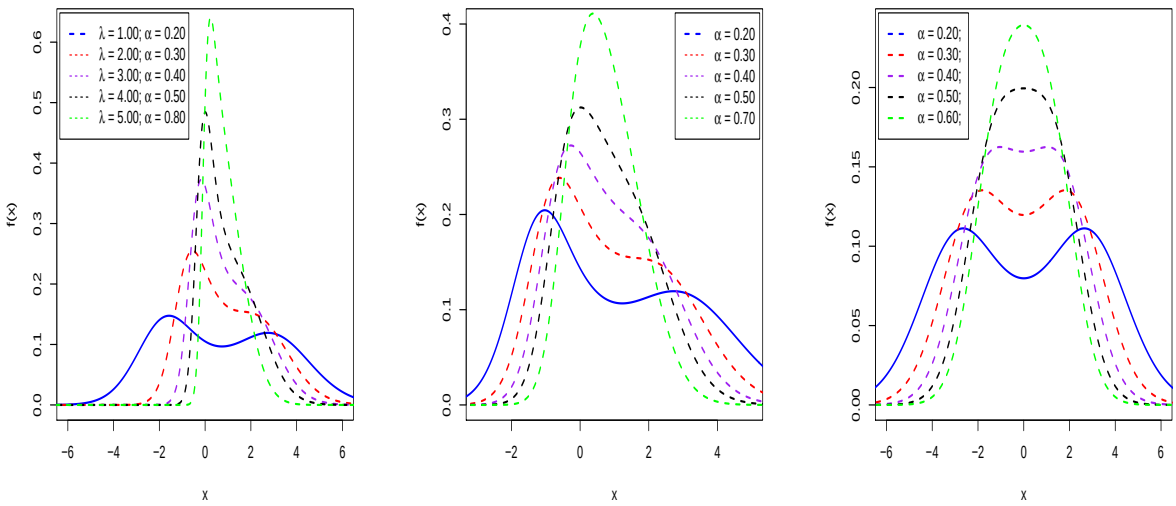


Figura 4.2 - Plots of the OLLSN density function for some parameter values. (a) For different values of the α and λ with $\mu = 0$ and $\sigma = 1$. (b) For different values of the α and $\lambda = 1.8$, $\mu = 0$ and $\sigma = 1$. (c) For different values of the α with $\lambda = 0$, $\mu = 0$ and $\sigma = 1$.

the parameters σ and α characterize the kurtosis and skew-bi-modality of this distribution.

4.3 Useful expansions

First, we define the exponentiated-SN ("Exp-SN" for short) distribution with power parameter $c > 0$ by raising the baseline skew-normal cdf $G(z;\lambda)$ to a power parameter c, say $W_c \sim \text{Exp}^c(SN)$. Then, the cdf and pdf of W_c are given by

$$H_c(z) = \Phi_{SN}\left(z;\lambda\right)^c \qquad \text{and} \qquad h_c(z) = 2\,c\,\sigma^{-2}\,\phi\left(z\right)\,\Phi\left(\lambda z\right)\,\Phi_{SN}\left(z;\lambda\right)^{c-1},$$

respectively. In a general context, the properties of the exponentiated-G (Exp-G) distributions have been studied by several authors for some baseline G models, see (NADARAJAH; KOTZ, 2006).

First, we obtain an expansion for $F(z;\lambda,\alpha)$ using a power series for $\Phi_{SN}^{\alpha}\left(z\right)$ (for $\alpha > 0$)

$$\Phi_{SN}^{\alpha}\left(z;\lambda\right) = \sum_{k=0}^{\infty} a_k\,\Phi_{SN}\left(z;\lambda\right)^k, \tag{4.6}$$

where

$$a_k = a_k(\alpha) = \sum_{j=k}^{\infty}(-1)^{k+j}\binom{\alpha}{j}\binom{j}{k}.$$

For any real $\alpha > 0$, we consider the generalized binomial expansion

$$\left[1-\Phi_{SN}\left(z;\lambda\right)\right]^{\alpha} = \sum_{k=0}^{\infty}(-1)^k\binom{\alpha}{k}\Phi_{SN}\left(z;\lambda\right)^k. \tag{4.7}$$

Inserting (4.6) and (4.7) in equation (5.3), we obtain

$$F(z;\lambda,\alpha) = \frac{\sum_{k=0}^{\infty} a_k\,\Phi_{SN}\left(z;\lambda\right)^k}{\sum_{k=0}^{\infty} b_k\,\Phi_{SN}\left(z;\lambda\right)^k},$$

where $b_k = a_k + (-1)^k\binom{\alpha}{k}$ for $k \geq 0$.

The ratio of the two power series in last equation can be expressed as

$$F(z;\lambda,\alpha) = \sum_{k=0}^{\infty} c_k\, \Phi_{SN}\,(z;\lambda)^k\,, \tag{4.8}$$

where $c_0 = a_0/b_0$ and the coefficients c_k's (for $k \geq 0$) are determined from the recurrence equation

$$c_k = b_0^{-1}\left(a_k - \sum_{r=1}^{k} b_r\, c_{k-r}\right).$$

The pdf of X follows by differentiating (4.8) as

$$f(z;\lambda,\alpha) = \sum_{k=0}^{\infty} c_{k+1}\, h_{k+1}(z;\lambda), \tag{4.9}$$

where $h_{k+1}(z;\lambda) = 2(k+1)\sigma^{-2}\,\phi(z)\,\Phi(\lambda z)\,\Phi_{SN}^{k}(z;\lambda)$ is the Exp-SN density function with power parameter $k+1$.

Equation (4.9) reveals that the OLLSN density function is a linear combination of the Exp-SN densities. Thus, some of its structural properties such as the ordinary and incomplete moments and generating function can be determined from well-established properties of the Exp-SN distribution. This equation is the main result of this section.

Let Y_{k+1} have the Exp-SN density function $h_{k+1}(x)$ with power parameter $k+1$. In this section, we prove that the mathematical properties of Y_{k+1} can be determined from those of the normal distribution.

Next, we give some power series for the SN cdf. Let Z be a $SN(\lambda)$ random variable. First, for $0 < \lambda < 1$, the cdf of Z can be expressed as (CASTELLARES *et al.*, 2012)

$$\Phi_{SN}(z;\lambda) = \sum_{r=0}^{\infty} p_r\, z^r, \tag{4.10}$$

where the coefficients $p_r = p_r(\lambda)$ are functions of λ given by

$$p_0 = \frac{1}{2} - \frac{1}{\pi}\arctan(\lambda),$$

$$p_{2r} = \frac{(1+\lambda^2)^r}{\lambda^{2r-1}\pi\, 2^r} \sum_{k=r}^{\infty} \frac{(-1)^{k+1}\binom{2k+1}{2r}\lambda^{2k}\Gamma\left((k-r)+1\right)}{(2k+1)k!}$$

and

$$p_{2r+1} = \frac{(1+\lambda^2)^{r+1/2}}{\lambda^{2r}\pi\, 2^{r+1/2}} \sum_{k=r}^{\infty} \frac{(-1)^{k}\binom{2k+1}{2r+1}\lambda^{2k}\Gamma\left((k-r)+1/2\right)}{(2k+1)k!}.$$

Second, for $\lambda \geq 1$, the cdf of Z can be expressed as

$$\Phi_{SN}(z;\lambda) = 2\Phi(z)\Phi(\lambda z) - \Phi_{SN}(\lambda z;\lambda^{-1}), \tag{4.11}$$

where $\Phi_{SN}(z;\lambda^{-1})$ can be expanded as in (4.10).

The standard normal cdf $\Phi(z)$ can be expressed as a power series $\Phi(z) = \sum_{i=0}^{\infty} e_i\, z^i$, $|z| < \infty$, where $e_0 = (1+\sqrt{2/\pi})^{-1}/2$, $e_{2i+1} = \frac{(-1)^i}{\sqrt{2\pi}(2i+1)2^i i!}$ for $i = 0, 1, 2\ldots$ and $e_{2i} = 0$ for $i = 1, 2, \ldots$

Then, for $\lambda \geq 1$, we can write using (4.10) and (4.11)

$$\Phi_{SN}(z;\lambda) = 2\sum_{i,j=0}^{\infty} \lambda^j\, e_i\, e_j\, z^{i+j} - \sum_{r=0}^{\infty} \lambda^r\, q_r\, z^r, \tag{4.12}$$

where $q_r = q_r(\lambda) = p_r(\lambda^{-1})$.

We define the set $I_r = \{(i,j); i+j = r, i,j = 0,1,2,\ldots\}$ for $r \geq 0$. Then, we can rewrite (4.12) as

$$\Phi_{SN}(z;\lambda) = \sum_{r=0}^{\infty} w_r\, z^r, \tag{4.13}$$

where $w_r = 2\lambda^j\, e_i\, e_j - \lambda^r\, q_r$ for any $(i,j) \in I_r$, $r \geq 0$.

Third, if $\lambda < 0$, we can use the relation $\Phi_{SN}(z;-\lambda) = 2\Phi(z) - \Phi_{SN}(z;\lambda)$ and apply the previous expansions to obtain $\Phi_{SN}(z;\lambda) = \sum_{r=0}^{\infty} g_r^{\star}\, z^r$, where $g_r^{\star} = 2e_r - g_r$ for $r \geq 0$.

Equations (4.10) and (4.13) have the same form except for the coefficients. The results below hold for all $\lambda > 0$ and we can only work with (4.10) and define $\Phi_{SN}(z;\lambda) = \sum_{r=0}^{\infty} g_r\, z^r$, where $g_r = p_r$ when $0 < \lambda < 1$ and $g_r = w_r$ when $\lambda > 1$. If $\lambda < 0$, we have only to change g_r by $g_r^{\star}$.

We use throughout a result of (GRADSHTEYN; RYZHIK, 2000, Section 0.314) for a power series raised to a positive integer k

$$\Phi_{SN}(z;\lambda)^k = \left(\sum_{r=0}^{\infty} g_r\, z^r\right)^k = \sum_{r=0}^{\infty} d_{k,r}\, z^r, \tag{4.14}$$

where the coefficients $d_{k,r}$ (for $r = 1, 2, \ldots$) are easily determined from the recurrence equation

$$d_{k,r} = (r g_0)^{-1} \sum_{m=1}^{r} [m(k+1) - r]\, g_m\, d_{k,r-m}, \tag{4.15}$$

and $d_{k,0} = g_0^k$. The coefficient $d_{k,r}$ can be obtained from $d_{k,0}, \ldots, d_{k,r-1}$ and then from the quantities $g_0, \ldots, g_r$.

By using (4.10) and (5.28) and the previous power series for $\Phi(z)$, the density $h_{k+1}(z;\lambda)$ can be expressed as

$$h_{k+1}(z;\lambda) = 2(k+1)\sigma^{-2}\, \phi(z) \sum_{i,r=0}^{\infty} e_i\, \lambda^i\, d_{k,r}\, z^{i+r}. \tag{4.16}$$

Combining (4.9) and (4.16), the OLLSN density can be reduced to

$$f(z;\lambda,\alpha) = \phi(z) \sum_{i,r=0}^{\infty} m_{i,r}\, z^{i+r}, \tag{4.17}$$

where $m_{i,r} = 2\sigma^{-2}\, \lambda^i\, e_i\, \sum_{k=0}^{\infty}(k+1)\, c_{k+1}\, d_{k,r}$.

Equation (4.17) reveals that some mathematical properties of the OLLSN distribution can be determined from those of the standard normal distribution. This equation is the main result of this section.

4.4 Mathematical properties

In this section, we provide some mathematical properties of $Z \sim \text{OLLSN}(0, 1, \lambda, \alpha)$ based on equation (4.17). The properties of $X \sim \text{OLLSN}(\mu, \sigma^2, \lambda, \alpha)$ follow from those of Z by simple linear transformation. The formulae derived throughout the paper can be easily handled in software platforms such as Maple, Mathematica and Matlab because of their ability to deal with analytic expressions of formidable size and complexity.

4.4.1 Moments

The nth ordinary moment of Z is given by

$$E(Z^n) = \sum_{i,r=0}^{\infty} m_{i,r}\, \delta_{n+i+r}, \tag{4.18}$$

where $\delta_{n+i+r} = 0$ if $n+i+r$ is odd and $\delta_{n+i+r} = \sigma^{n+i+r}\,(n+i+r)!!$ if $n+i+r$ is even, where $p!! = p(p-2)\ldots 3$ denotes the double factorial.

The central moments (μ_n) and cumulants (κ_n) of Z are determined from (4.18) as

$$\mu_n = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \mu_1'^k\, \mu_{n-k}' \qquad \text{and} \qquad \kappa_n = \mu_n' - \sum_{k=1}^{n-1} \binom{n-1}{k-1} \kappa_k\, \mu_{n-k}',$$

respectively, where $\kappa_1 = \mu_1'$. The skewness $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ and kurtosis $\gamma_2 = \kappa_4/\kappa_2^2$ are obtained from the third and fourth standardized cumulants.

4.4.2 Incomplete moments

The nth incomplete moment of Z is given by $m_n(y) = \int_0^y x^n\, f(x)dx$. Based on the linear representation (4.17) and the monotone convergence theorem, we obtain

$$m_n(y) = \frac{1}{\sqrt{2\pi}} \sum_{i,r=0}^{\infty} m_{i,r} \int_{-\infty}^{y} A(i+r, y), \tag{4.19}$$

where $A(k,y) = \int_{-\infty}^{y} z^k\, \mathrm{e}^{-\mathrm{z}^2/2}\, \mathrm{dx}$ (for $k \geq 0$) depends if $y < 0$ and $y > 0$.

For determining the integral $A(k,y)$ we use some known results on special functions. First,

$$G(k) = \int_0^{\infty} x^k\, \mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\, \mathrm{dx} = 2^{(\mathrm{k}-1)/2}\, \Gamma\left(\frac{\mathrm{k}+1}{2}\right).$$

Second, we consider the confluent hypergeometric function ${}_1F_1(a;b;z) = \sum_{k=0}^{\infty} \frac{(a)_k}{(b)_k} \frac{z^k}{k!}$, the hypergeometric function ${}_2F_1(a,b;c;z) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{z^k}{k!}$ and the Kummer's function given by $U(a,b;z) = z^{-a}\, {}_2F_0(a, 1+a-b; -z^{-1})$. All these functions can be easily computed using Mathematica. See

http://mathworld.wolfram.com/ConfluentHypergeometricFunctionoftheFirstKind.html.

We can prove that $A(k,y) = (-1)^k\, G(k) + (-1)^{k+1}\, H(k,y)$ for $y < 0$ and $A(k,y) = (-1)^k G(k) + H(k,y)$ for $y > 0$, where $H(k,y) = \int_0^y x^k\, \mathrm{e}^{-\mathrm{x}^2/2}\, \mathrm{dx}$ (WHITTAKER; WATSON, 1990) is given by

$$\begin{aligned} H(k,y) &= \frac{2^{k/4+1/4}\, y^{k/2+1/2}\, \mathrm{e}^{-\mathrm{y}^2/4}}{(k/2+1/2)(k+3)}\, N_{k/4+1/4,k/4+3/4}(y^2/2) \\ &\quad + \frac{2^{k/4+1/4}\, y^{k/2-3/2}\, \mathrm{e}^{-\mathrm{y}^2/4}}{k/2+1/2}\, N_{k/4+5/4,k/4+3/4}(y^2/2) \end{aligned}$$

and $N_{k,m}(x)$ is the Whittaker function (WHITTAKER; WATSON, 1990, pp. 339-351) given by

$$N_{k,m}(x) = M_{k,m}(x) = \mathrm{e}^{-x/2}\, x^{m+1/2}\; {}_1F_1\left(\frac{1}{2} + m - k, 1 + 2m; x\right)$$

or

$$N_{k,m}(x) = W_{k,m}(x) = \mathrm{e}^{-x/2}\, x^{m+1/2}\; U\left(\frac{1}{2} + m - k, 1 + 2m; x\right).$$

These functions are implemented as WhittakerM$[k,m,x]$ and WhittakerW$[k,m,x]$ in Mathematica, respectively. Based on these results, we can obtain $m_n(y)$ from (4.19).

An important application of the first incomplete moment, say $m_1(y)$, is related to the Bonferroni and Lorenz curves of Z defined by $B(\pi) = m_1(q)/(\pi\mu_1')$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $q = Q_Z(\pi)$ is obtained from the qf (5.5) for a given probability π.

A second application of $m_1(y)$ refers to the mean residual life, which represents the expected additional life length for a unit which is alive at age t, and the mean inactivity time, which represents the waiting time elapsed since the failure of an item on condition that this failure had occurred in $(0,t)$, given by $\nu_1(t) = [1 - m_1(t)]\,/[1 - F(t)] - t$ and $\xi(t) = t - m_1(t)/F(t)$, respectively.

A third one is related to the mean deviations of Z about the mean μ_1' and about the median M given by $\delta_1 = 2\Big[\mu_1'\, F(\mu_1') - M_1(\mu_1')\Big]$ and $\delta_2 = \mu_1' - 2m_1(M)$, where $\mu_1' = E(Z)$ and $M = Q(1/2)$ is the median obtained from (5.5).

4.4.3 Moments based on quantiles

We provide a further insight of the effects of the parameters α and λ on the skewness and kurtosis of Z by considering these measures based on quantiles. The shortcomings of the classical kurtosis measure are well-known. There are many heavy-tailed distributions for which this measure is infinite, so it becomes uninformative precisely when it needs to be.

The Bowley's skewness is based on quartiles

$$B = \frac{Q(3/4) - 2Q(1/2) + Q(1/4)}{Q(3/4) - Q(1/4)},$$

whereas the Moors' kurtosis is based on octiles

$$M = \frac{Q(7/8) - Q(5/8) - Q(3/8) + Q(1/8)}{Q(6/8) - Q(2/8)}.$$

These measures are less sensitive to outliers and they exist even for distributions without moments. For the classical Student t distribution with 10 and 5 degrees of freedom, these measures are zero (Bowley) and 1.27705 and 1.32688 (Moors), respectively. For the standard normal distribution, these measures are zero (Bowley) and 1.2331 (Moors).

In Figure 4.3 and 4.4, we plot B and M respectively, for different values of α and λ. These plots indicate how these measures can be sensitive to the two shape parameters α and λ.

4.4.4 Generating function

In this section, we calculate the moment generating function (mgf) of the OLLSN distribution. First, we require the formula, which is a special case of the result by (PRUDNIKOV; BRYCHKOV; MARICHEV, 1986, equation 2.3.15.8),

$$J(n,t) = \frac{1}{\sqrt{2\,\pi}} \int_{-\infty}^{\infty} z^n \exp\left(-tz - \frac{z^2}{2}\right) dz = (-1)^n \frac{\partial^n}{\partial t^n} \left(\mathrm{e}^{\mathrm{t}^2/2}\right). \tag{4.20}$$

(a) (b)

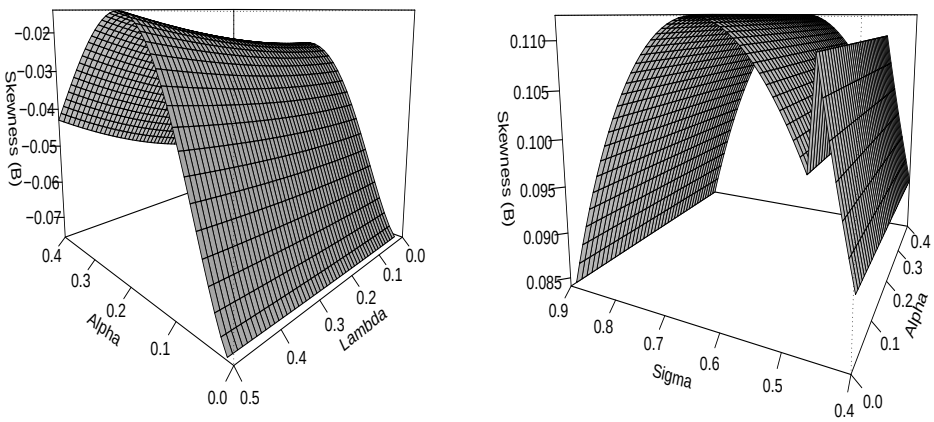


Figura 4.3 - Bowley's skewness for the OLLSN distribution. Plots (a) as functions of $\lambda \in [0, 0.5]$ with $\alpha \in [0, 0.4]$ and (b) as functions of $\lambda \in [0, 0.9]$ with $\alpha \in [0, 0.4]$

(c) (d)

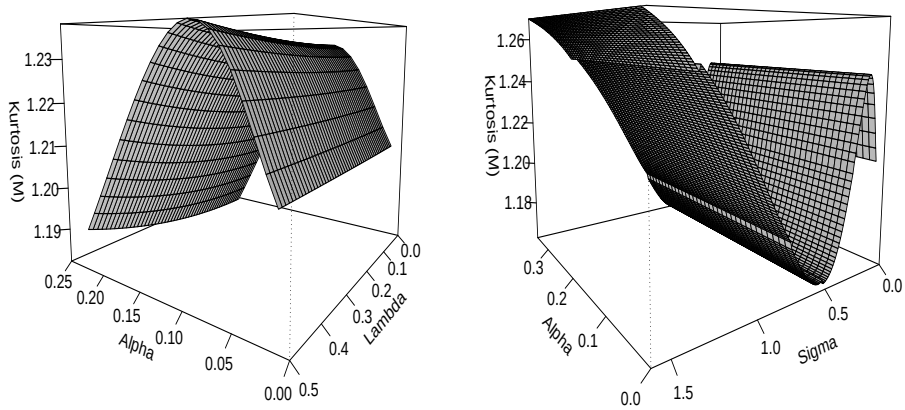


Figura 4.4 - Moors' kurtosis for the OLLSN distribution. Plots (a) as functions of $\lambda \in [0, 0.5]$ with $\alpha \in [0, 0.25]$ and (b) as functions of $\lambda \in [0, 1.6]$ with $\alpha \in [0, 0.35]$

Theorem. Let Z have the OLLSN distribution. The mgf of Z is given by

$$M_Z(-t) = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{tz}}\, \mathrm{f(z)dz} = \sum_{\mathrm{i,r}=0}^{\infty} (-1)^{\mathrm{i+r}}\, \mathrm{m_{i,r}}\, \mathrm{J(i+r,t)},$$

where $J(i+r,t)$ is given by (4.20).

Proof:

We can write using (4.17)

$$M_Z(-t) = (2\pi)^{-\frac{1}{2}} \sum_{i,r=0}^{\infty} m_{i,r} \int_{-\infty}^{\infty} z^{i+r} \exp\left(-\frac{z^2}{2} - tz\right) dz.$$

By interchanging the sums and the integral and using (4.20), the theorem is proved. Clearly, the mgf of $Y = \mu + \sigma Z$ is given by $M_Y(-t) = \mathrm{e}^{\mu \mathrm{t}}\, \mathrm{M_Z}\,(-\sigma \mathrm{t})$.

4.5 Maximum likelihood estimation

Let $x_1, \ldots, x_n$ be a random sample of size n from the OLLSN$(\mu, \sigma, \lambda, \alpha)$ distribution. In this section, we determine the maximum likelihood estimates (MLEs) of the model parameters from complete samples only. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\mu, \sigma, \lambda, \alpha)^T$ is given by

$$\begin{aligned} l(\boldsymbol{\theta}) &= (\alpha-1)\sum_{i=1}^{n} \log[1-\Phi_{SN}(z_i;\lambda)] - 2\sum_{i=1}^{n} \log\left\{\Phi_{SN}^{\alpha}(z_i;\lambda) + [1-\Phi(z_i;\lambda)]^{\alpha}\right\} \\ &+ \sum_{i=1}^{n} \log[\phi(z_i)] + \sum_{i=1}^{n} \log[\Phi(\lambda\, z_i)] + (\alpha-1)\sum_{i=1}^{n} \log[\Phi_{SN}(z_i;\lambda)] \\ &+ n \log\left(\frac{2\,\alpha}{\sigma}\right), \end{aligned} \tag{4.21}$$

where $z_i = (x_i - \mu)/\sigma$.

The components of the score vector $U(\boldsymbol{\theta})$ are given by

$$\begin{aligned} U_{\mu}(\boldsymbol{\theta}) &= -2\alpha \sum_{i=1}^{n} \frac{\phi\left(z_i\right)\left[-1 - \frac{24}{5\sigma}\mathrm{erf}\left(-\frac{\lambda z_i}{\sqrt{2}}\right)\right]\left\{\Phi_{SN}^{\alpha-1}\left(z_i;\lambda\right) + \left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha-1}\right\}}{\sigma\left\{\Phi_{SN}^{\alpha}\left(z_i;\lambda\right) + \left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha}\right\}} \\ &\frac{(\alpha-1)}{\sigma}\sum_{i=1}^{n}\left\{\frac{\phi\left(\lambda z_i\right)\left\{-1 + 2\Phi_{SN}\left(z_i;\lambda\right) - \frac{24}{5\sigma}\mathrm{erf}\left(-\frac{\lambda z_i}{\sqrt{2}}\right)\right\}}{\Phi_{SN}\left(z_i;\lambda\right)\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]}\right\} \\ &+ \sum_{i=1}^{n} z_i + \sum_{i=1}^{n} \frac{\phi\left(\lambda z_i\right)}{\sqrt{2}\Phi\left(\lambda z_i\right)} +, \end{aligned}$$

$$
\begin{aligned}
U_{\sigma}(\boldsymbol{\theta}) \;=\; & \frac{4\alpha}{\sigma}\sum_{i=1}^{I}\frac{\Phi_{SN}\left(z_i;\lambda\right)\left[1-\frac{3\phi(z_i)}{5\sigma^2}\mathrm{erf}\left(-\frac{\lambda z_i}{\sqrt{2}}\right)(\mu-x_i)\right]\phi\left(z_i\right)\left[z_i-\frac{3\phi(z_i)}{5\sigma^2}\mathrm{erf}\left(-\frac{\lambda z_i}{\sqrt{2}}\right)\right]}{\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)+\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha}} \\
& +\sum_{i=1}^{I}\frac{\phi\left(\lambda z_i\right)\Phi\left(\lambda z_i\right)\left[1-z_i\sqrt{2}\right]}{\sqrt{2}\Phi\left(\lambda z_i\right)\Phi\left(\lambda z_i\right)}+\frac{2\alpha}{\sigma}\sum_{i=1}^{I}\frac{\left[\Phi_{SN}^{\alpha-1}\left(z_i\right)+\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha-1}\right]}{\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)+\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha}} \\
& +\frac{n}{\sigma}+\frac{1}{\sigma}\sum_{i=1}^{I}z_i+(\alpha-1)\sum_{i=1}^{I}\frac{2\Phi_{SN}\left(z_i;\lambda\right)\left[1-\frac{3\phi(z_i)}{5\sigma^2}\mathrm{erf}\left(-\frac{\lambda z_i}{\sqrt{2}}\right)(\mu-x_i)\right]}{\Phi_{SN}\left(z_i;\lambda\right)\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]} \\
& +(\alpha-1)\sum_{i=1}^{I}\frac{z_i\phi\left(z_i\right)\sigma^{-1}+2\sigma^{-1}z_i\phi\left(z_i\right)\Phi_{SN}\left(z_i;\lambda\right)}{\Phi_{SN}\left(z_i;\lambda\right)\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]},
\end{aligned}
$$

$$
\begin{aligned}
U_{\lambda}(\boldsymbol{\theta}) \;=\; & -\frac{\alpha}{\pi\left(1+\lambda^2\right)}\sum_{i=1}^{n}\frac{\left\{\Phi_{SN}^{\alpha-1}\left(z_i;\lambda\right)-\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha-1}\right\}}{\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)+\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha}}\exp\left\{-\frac{1}{2}z_i^2\left(1+\lambda^2\right)\right\} \\
& +\sum_{i=1}^{n}\Phi_{SN}^{-1}\left(z_i;\lambda\right)\left[\frac{1}{2\lambda^2\sqrt{\pi}}\mathrm{erf}\left(-\frac{\lambda z_i}{\sqrt{2}}\right)\right]+\sum_{i=1}^{n}\Phi_{SN}^{-1}\left(z_i;\lambda\right)\left[\frac{z_i\phi\left(\lambda z_i\right)}{\lambda}\right] \\
& \frac{(\alpha-1)}{\pi\left(1+\lambda^2\right)}\sum_{i=1}^{n}\frac{\left[1-4\Phi_{SN}\left(z_i;\lambda\right)\right]\exp\left\{-\frac{1}{2}z_i^2\left(1+\lambda^2\right)\right\}}{\Phi_{SN}\left(z_i;\lambda\right)\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]},
\end{aligned}
$$

$$
\begin{aligned}
U_{\alpha}(\boldsymbol{\theta}) \;=\; & -\frac{2\alpha\left\{\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)\left[\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)\right]+\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha}\log\left[\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)\right]\right\}}{\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)+\left[1-\Phi_{SN}\left(z_i;\lambda\right)\right]^{\alpha}} \\
& +\sum_{i=1}^{n}\log\left[1-\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)\right]+\sum_{i=1}^{n}\log\left[\Phi_{SN}^{\alpha}\left(z_i;\lambda\right)\right].
\end{aligned}
$$

The MLE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained from the nonlinear likelihood equations $U_{\mu}(\boldsymbol{\theta})=0$, $U_{\sigma}(\boldsymbol{\theta})=0$, $U_{\lambda}(\boldsymbol{\theta})=0$ and $U_{\alpha}(\boldsymbol{\theta})=0$. These equations cannot be solved analytically and statistical software can be used to solve them numerically. We can use iterative techniques such as a Newton-Raphson type algorithm to obtain the $\widehat{\boldsymbol{\theta}}$. We employed the Optim script in the R software. We choose as initial values for μ and σ their MLEs $\hat{\mu}$ and $\hat{\sigma}$ under the special normal model.

The 4×4 total observed information matrix is given by $J(\boldsymbol{\theta})$, whose elements can be evaluated numerically. The asymptotic distribution of $(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta})$ is $N_4(0,K(\boldsymbol{\theta})^{-1})$ under standard regularity conditions, where $K(\boldsymbol{\theta})=E[L(\boldsymbol{\theta})]$ is the expected information matrix. The approximate multivariate normal $N_4(0,J(\widehat{\boldsymbol{\theta}})^{-1})$ distribution, where $J(\widehat{\boldsymbol{\theta}})^{-1}$ is the observed information matrix evaluated at $\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}$, can be used to construct approximate confidence intervals for the model parameters.

The likelihood ratio (LR) statistic can be used to compare the OLLN distribution with some of its special models. We can compute the maximum values of the unrestricted and restricted log-likelihoods to obtain LR statistics for testing some of its sub-models. In any case, hypothesis tests of the type $H_0 : \boldsymbol{\psi} = \boldsymbol{\psi}_0$ versus $H_1 : \boldsymbol{\psi} \neq \boldsymbol{\psi}_0$, where $\boldsymbol{\psi}$ is a vector formed with some components of $\boldsymbol{\theta}$ of interest and $\boldsymbol{\psi}_0$ is a specified vector, can be performed via LR statistics. For example, the test of $H_0 : \alpha = 1$ versus $H_1 :$ H_0 *is not true* is equivalent to compare the OLLSN and SN distributions and the LR statistic becomes

$$w = 2\{\ell(\widehat{\mu}, \widehat{\sigma}, \widehat{\lambda}, \widehat{\alpha}) - \ell(\widetilde{\mu}, \widetilde{\sigma}, \widetilde{\lambda}, 1)\},$$

where $\widehat{\mu}$, $\widehat{\sigma}$, $\widehat{\lambda}$ and $\widehat{\alpha}$ are the MLEs under H_1 and $\widetilde{\mu}$, $\widetilde{\sigma}$ and $\widetilde{\lambda}$ are the estimates under H_0.

Simulation study

We perform a simulation study in order to evaluate some properties of the MLE $\widehat{\boldsymbol{\theta}}$. The data are simulated from the OLLSN$(\mu, \sigma^2, \lambda, \alpha)$ distribution. We use the qf given in equation (5.5). We consider the following values for the parameters:

- **Scenario 1**: $n = 100, 150$ and 300, $\mu = 0.00$, $\sigma = 1.00$, $\lambda = 1.00$, $\alpha = 0.20$.
- **Scenario 2**: $n = 20, 30$ and 50, $\mu = 0.00$, $\sigma = 1.00$, $\lambda = 1.00$, $\alpha = 0.20$.

For scenarios, the observations are generated by taking $n = 20, 30, 50, 100, 150$ and 300. The results are obtained from $1,000$ Monte Carlo simulations performed using the **R** software with the Optim function. We determine the average estimates (AEs), biases and means squared errors (MSEs). For each generated sample, the parameters are estimated by maximum likelihood. The results are reported in Table 4.1.

The figures given in Table 4.1 indicate that the MSEs and the biases of the estimates of μ, σ, λ and α decay toward zero when the sample size increases, as expected under standard asymptotic theory. As n increases, the AEs of the parameters tend to be closer to the true parameter values. This fact supports that the asymptotic normal distribution provides an adequate approximation to the finite sample distribution of the MLEs. The normal approximation can be oftentimes improved by making bias adjustments to the MLEs. Approximations to the biases of these estimates in simple models

Tabela 4.1 - The AEs and MSEs based on $1,000$ simulations of the OLLSN distribution

Scenario 1								
n=100			n=150			n=300		
$\boldsymbol{\theta}$	AE	MSE	$\boldsymbol{\theta}$	AE	MSE	$\boldsymbol{\theta}$	AE	MSE
μ	-0.0119	0.2033	μ	-0.0085	0.0383	μ	-0.0053	0.0293
σ	1.0680	0.5407	σ	1.0270	0.0408	σ	1.0154	0.0291
λ	1.0165	0.0390	λ	0.9931	0.0146	λ	1.0221	0.0241
α	0.2449	0.7054	α	0.2132	0.6436	α	0.2081	0.6994
Scenario 2								
n=20			n=30			n=50		
$\boldsymbol{\theta}$	AE	MSE	$\boldsymbol{\theta}$	AE	MSE	$\boldsymbol{\theta}$	AE	MSE
μ	-0.0166	0.1254	μ	-0.0194	0.2192	μ	0.0171	0.1313
σ	1.0578	0.2261	σ	1.1682	1.5601	σ	1.0662	0.1897
λ	1.0232	0.0507	λ	1.0089	0.0222	λ	1.0087	0.0315
α	0.2311	0.7279	α	0.2832	0.6763	α	0.2346	0.6854

(a) (b) (c)

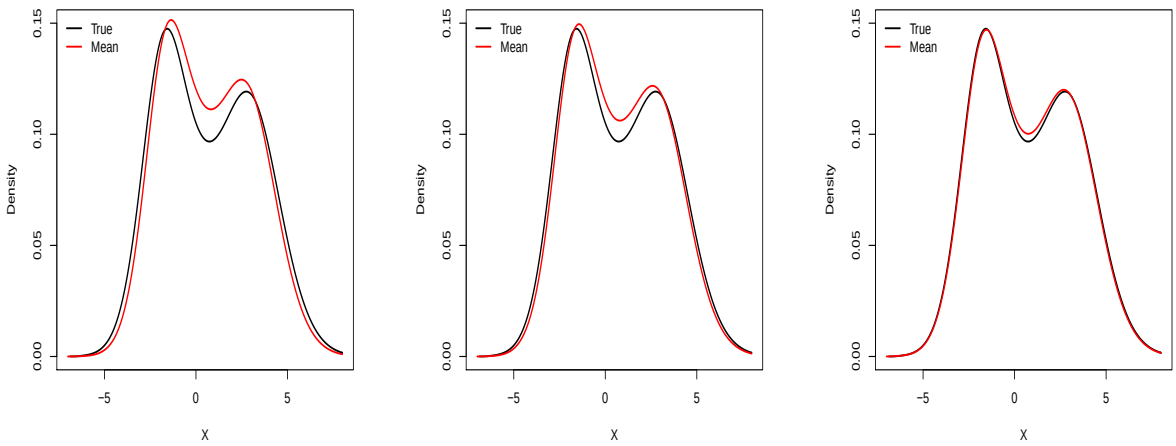


Figura 4.5 - Density functions of the OLLSN distribution at the true parameter values and at the average estimates for $\mu = 0$, $\sigma = 1$, $\lambda = 1$ and $\alpha = 0.20$. (a) n=1020; (b) n=150; (c) n=300.

may be obtained analytically. Bias correction typically does a very good job for correcting the MLEs. However, it may either increase the MSEs. Whether bias correction is useful in practice depends basically on the shape of the bias function and on the variance of the MLE. In order to improve the accuracy of the MLEs using analytical bias reduction one needs to obtain several cumulants of log likelihood derivatives, which are notoriously cumbersome for the proposed model. in Figure 4.5 we display plots of the true densities for selected parameter values and the density functions evaluated at the AEs given in Ta-

ble 4.1 for some sample sizes. These plots are in agreement with the standard asymptotic theory for the MLEs.

4.6 Completely randomized design model

In this section, we use the OLLSN model in experimental design. It is often necessary to plan an experiment so that the treatments are not systematically subject to any bias (defect), thus invalidating the conclusions. The proper design of experiments allows simultaneously investigating various treatments in a single experiment without invalidating the presuppositions required by the mathematical model. One of the model assumptions. However, when the data present asymmetry, kurtosis or bimodality, the normal model will not be suitable. For this reason, we consider that the distribution of the errors is the OLLSN distribution. The completely randomized design (CRD) model is characterized by not imposing any restriction on randomization of the treatments. In this case, all the treatments have the same chance of occupying any experimental unit or plot. This design is used when all the conditions for the performance of the experiment are controlled and the experimental units are considered homogeneous. The mathematical model associated with experiments with one factor is:

$$Y_{ij} = m + \tau_i + \epsilon_{ij}, \tag{4.22}$$

where Y_{ij} represents the observed value of the group that received treatment i, m is the overall mean effect, τ_i is the effect of treatment i applied to the treated group and ϵ_{ij} $\sim$ OLLSN$(0, \sigma, \lambda, \alpha)$ is the effect of the uncontrolled factors in the experimental group, for $i = 1, \ldots, I$ and $j = 1, \ldots, J$, where I denotes the number of treatments and J the number of repetitions.

Let $y_{11}, \ldots, y_{IJ}$ be a sample of size n from the OLLSN distribution. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (m, \boldsymbol{\tau}^T, \sigma, \lambda, \alpha)^T$, where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_I)^T$,

is given by

$$
\begin{aligned}
l(\boldsymbol{\theta}) &= \sum_{i=1}^{I}\sum_{j=1}^{J}\log[\phi(z_{ij})] + (\alpha-1)\sum_{i=1}^{I}\sum_{j=1}^{J}\log\{\Phi_{SN}(z_{ij};\lambda)[1-\Phi_{SN}(z_{ij};\lambda)]\} \\
&+ \sum_{i=1}^{I}\sum_{j=1}^{J}\log[\Phi(\lambda z_{ij})] - 2\sum_{i=1}^{I}\sum_{j=1}^{J}\log\{\Phi_{SN}^{\alpha}(z_{ij};\lambda) + [1-\Phi_{SN}(z_{ij};\lambda)]^{\alpha}\} \\
&+ n\log\left(\frac{2\,\alpha}{\sigma}\right), \qquad (4.23)
\end{aligned}
$$

where $z_{ij} = (y_{ij} - m - \tau_i)/\sigma$.

The MLE $\widehat{\boldsymbol{\theta}}$ of the model parameters can be obtained by maximizing the log-likelihood (4.23) using the Optim function in R. In addition, we obtain using the "L-BFGS-B" or "Nelder-Mead" methods the parameter estimates and their standard errors. We can provide upon request the data and the R script for doing the calculations. The components of the score vector $U(\boldsymbol{\theta})$ are given by

$$
\begin{aligned}
U_m(\boldsymbol{\theta}) &= -2\alpha\sum_{i=1}^{n}\frac{\phi(z_i)\left[-1-\frac{12}{5\sigma}\mathrm{erf}\left(-\frac{\lambda z_i}{\sqrt{2}}\right)\right]\left\{\Phi_{SN}^{\alpha-1}(z_{ij};\lambda) + [1-\Phi_{SN}(z_{ij};\lambda)]^{\alpha-1}\right\}}{\sigma\left\{\Phi_{SN}^{\alpha}(z_{ij};\lambda) + [1-\Phi_{SN}(z_{ij};\lambda)]^{\alpha}\right\}} \\
&+ \frac{(\alpha-1)}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\left\{\frac{\phi(\lambda z_{ij})\left\{-1+2\Phi_{SN}(z_{ij};\lambda) - \frac{24}{5\sigma}\mathrm{erf}\left(-\frac{\lambda z_{ij}}{\sqrt{2}}\right)\right\}}{\Phi_{SN}(z_{ij};\lambda)[1-\Phi_{SN}(z_{ij};\lambda)]}\right\} \\
&+ \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\phi(\lambda z_{ij})}{\sqrt{2}\Phi(z_{ij})} + \sum_{i=1}^{I}\sum_{j=1}^{J} z_{ij}
\end{aligned}
$$

$$
\begin{aligned}
U_{\tau_i}(\boldsymbol{\theta}) &= \frac{(\alpha-1)}{\sigma}\sum_{i=1}^{I}\left\{\frac{\phi(\lambda z_{ij})\left\{-1+2\Phi_{SN}(z_{ij};\lambda) - \frac{12}{5\sigma}\mathrm{erf}\left(-\frac{\lambda z_{ij}}{\sqrt{2}}\right)\right\}}{\Phi_{SN}(z_{ij};\lambda)[1-\Phi_{SN}(z_{ij};\lambda)]}\right\} \\
&- \frac{2}{\sigma}\sum_{i=1}^{I}\frac{\alpha\phi(z_{ij})\left[-1-\frac{12}{5\sigma}\mathrm{erf}\left(-\frac{\lambda z_{ij}}{\sqrt{2}}\right)\right]}{\left\{\Phi_{SN}^{\alpha}(z_{ij};\lambda) + [1-\Phi_{SN}(z_{ij};\lambda)]^{\alpha}\right\}} + \sum_{i=1}^{I} z_{ij} \\
&+ \sum_{i=1}^{I}\frac{\phi(\lambda z_{ij})}{\sqrt{2}\Phi(\lambda z_{ij})},
\end{aligned}
$$

$$
\begin{aligned}
U_{\sigma}(\boldsymbol{\theta}) &= \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\Phi_{SN}\left(z_{ij},\lambda\right)\left[\phi\left(z_{ij}\right)-\frac{3\phi^{2}(z_{ij})}{5\sigma}\mathrm{erf}\left(-\frac{\lambda z_{ij}}{\sqrt{2}}\right)z_{ij}\right]\left[z_{ij}-\frac{3\phi(z_{ij})}{5\sigma^{2}}\mathrm{erf}\left(-\frac{\lambda z_{ij}}{\sqrt{2}}\right)\right]}{(4\alpha)^{-1}\sigma\{\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)+\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]^{\alpha}\}} \\
&\quad+\frac{1}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}z_{ij}+(\alpha-1)\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{z_{ij}\phi\left(z_{ij}\right)\sigma^{-1}+2\sigma^{-1}z_{ij}\phi\left(z_{ij}\right)\Phi_{SN}\left(z_{ij};\lambda\right)}{\Phi_{SN}\left(z_{ij};\lambda\right)\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]} \\
&\quad+(\alpha-1)\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{2\Phi_{SN}\left(z_{ij};\lambda\right)\left[1-\frac{3\phi(z_{ij})}{5\sigma^{2}}\mathrm{erf}\left(-\frac{\lambda z_{ij}}{\sqrt{2}}\right)(\mu-x_{i})\right]}{\Phi_{SN}\left(z_{ij};\lambda\right)\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]} \\
&\quad+\frac{2\alpha}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\left[\Phi_{SN}^{\alpha-1}\left(z_{ij}\right)+\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]^{\alpha-1}\right]}{\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)+\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]^{\alpha}} \\
&\quad+\frac{n}{\sigma}+\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\phi\left(\lambda z_{ij}\right)\Phi\left(\lambda z_{ij}\right)\left[1-z_{ij}\sqrt{2}\right]}{\sqrt{2}\Phi\left(\lambda z_{ij}\right)\Phi\left(\lambda z_{ij}\right)}, \\
U_{\lambda}(\boldsymbol{\theta}) &= \sum_{i=1}^{I}\sum_{j=1}^{J}\Phi_{SN}^{-1}\left(z_{ij};\lambda\right)\left[\frac{1}{2\lambda^{2}\sqrt{\pi}}\mathrm{erf}\left(-\frac{\lambda z_{ij}}{\sqrt{2}}\right)\right]+\sum_{i=1}^{I}\sum_{j=1}^{J}\Phi_{SN}^{-1}\left(z_{ij};\lambda\right)\left[\frac{z_{ij}\phi\left(\lambda z_{ij}\right)}{2\lambda}\right] \\
&\quad-\frac{\alpha}{\pi\left(1+\lambda^{2}\right)}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\left\{\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)-\left[1-\Phi_{SN}\left(z_{ij}\right)\right]^{\alpha-1}\right\}}{\Phi_{SN}^{\alpha}\left(z_{ij}\right)+\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]^{\alpha}}\exp\left\{-\frac{1}{2}z_{ij}^{2}\left(1+\lambda^{2}\right)\right\} \\
&\quad+\frac{(\alpha-1)}{\pi\left(1+\lambda^{2}\right)}\sum_{i=1}^{n}\frac{\left[1-4\Phi_{SN}\left(z_{ij};\lambda\right)\right]\exp\left\{-\frac{1}{2}z_{ij}^{2}\left(1+\lambda^{2}\right)\right\}}{\Phi_{SN}\left(z_{ij};\lambda\right)\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]} \\
&\quad+\sum_{i=1}^{I}\sum_{j=1}^{J}\Phi_{SN}^{-1}\left(z_{ij};\lambda\right)\left[\frac{z_{ij}\phi\left(\lambda z_{ij}\right)}{2\lambda}\right], \\
U_{\alpha}(\boldsymbol{\theta}) &= -2\alpha\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\left\{\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)\left[\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)\right]+\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]^{\alpha}\log\left[\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)\right]\right\}}{\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)+\left[1-\Phi_{SN}\left(z_{ij};\lambda\right)\right]^{\alpha}} \\
&\quad+\sum_{i=1}^{I}\sum_{j=1}^{J}\log\left[\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)\right]+\sum_{i=1}^{I}\sum_{j=1}^{J}\log\left[1-\Phi_{SN}^{\alpha}\left(z_{ij};\lambda\right)\right].
\end{aligned}
$$

The asymptotic distribution of $\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}$ is multivariate normal $N_{I+4}(0,K(\boldsymbol{\theta})^{-1})$, where $K(\boldsymbol{\theta})$ is the information matrix. The asymptotic covariance matrix $K(\boldsymbol{\theta})^{-1}$ of $\widehat{\boldsymbol{\theta}}$ can be approximated by the inverse of the $(I+4)\times(I+4)$ observed information matrix $J(\boldsymbol{\theta})$, i.e., we can use $J(\boldsymbol{\theta})^{-1}$ to obtain an approximation for the large-sample covariance matrix of the MLEs. We provide confidence intervals for any parameters using the asymptotic normality of these estimates. Then, the inference for the parameter vector $\boldsymbol{\theta}$ can be based on the multivariate normal approximation $N_{I+4}(0,J(\boldsymbol{\theta})^{-1})$ for $\widehat{\boldsymbol{\theta}}$ under the usual asymptotic theory.

The $(I+4)\times(I+4)$ total observed information matrix is given by

$$J(\boldsymbol{\theta}) = \begin{pmatrix} \mathrm{J}_{mm} & \mathrm{J}_{m\tau_1} & \dots & \mathrm{J}_{m\tau_I} & \mathrm{J}_{m\sigma} & \mathrm{J}_{m\lambda} & \mathrm{J}_{m\alpha} \\ \cdot & \mathrm{J}_{\tau_1\tau_1} & \cdots & \mathrm{J}_{\tau_1\tau_I} & \mathrm{J}_{\tau_1\sigma} & \mathrm{J}_{\tau_1\lambda} & \mathrm{J}_{\tau_1\alpha} \\ \cdot & \cdot & \ddots & \vdots & \vdots & \vdots & \vdots \\ \cdot & \cdot & \cdot & \mathrm{J}_{\tau_1\tau_I} & \vdots & \vdots & \vdots \\ \cdot & \cdot & \cdot & \cdot & \mathrm{J}_{\sigma\sigma} & \vdots & \vdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \mathrm{J}_{\lambda\lambda} & \vdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \mathrm{J}_{\alpha\alpha} \end{pmatrix},$$

where $i = i^{'} = 1, \dots, I$ and $j = j^{'} = 1, \dots, J$ are evaluated numerically.

An asymptotic confidence interval with significance level γ for each parameter θ_r is given by

$$IC_r = \left(\hat{\theta}_r - z_{\gamma/2}\sqrt{-\widehat{J}^{r,r}}, \hat{\theta}_r + z_{\gamma/2}\sqrt{-\widehat{J}^{r,r}}\right),$$

where$-\widehat{J}^{r,r}$ is the rth diagonal of $-J(\widehat{\boldsymbol{\theta}})^{-1}$ estimated at $\hat{\boldsymbol{\theta}}$, for $r = 1, \dots, I+4$ and $z_{\gamma/2}$ is the quantile $1-\gamma/2$ of the standard normal distribution

4.7 Residual analysis

When attempting to adjust a model to a dataset, the validation of the fit showed be analyzed by a specific statistic, with the purpose of measuring the goodness-of-fit. Once the model is chosen and fitted, the analysis of the residuals is an efficient way to check the model adequacy. The residuals also serve for other purposes, such as to detect the presence of aberrant points (outliers), identify the relevance of an additional factor omitted from the model and verify if there are indications of serious deviance from the distribution considered for the random error. Further, since the residuals are used to identify discrepancies between the fitted model and the dataset, it is convenient to define residuals that take into account the contribution of each observation to the goodness-of-fit measure used.

In summary, the residuals allow measuring the model fit for each observation and enable studying whether the differences between the observed and fitted values are due to chance or to a systematic behavior that can be modeled. The quantile residual proposed

by (DUNN; SMYTH, 1996) can be defined as a measure of the discrepancy between y_{ij} and $\hat{\mu}_{ij}$, and it is given by

$$\widehat{qr}_{ij} = \Phi^{-1}\left\{\frac{\Phi_{SN}^{\hat{\alpha}}\left(\frac{y_{ij}-\hat{m}-\hat{\tau}_i}{\hat{\sigma}};\hat{\lambda}\right)}{\Phi_{SN}^{\hat{\alpha}}\left(\frac{y_{ij}-\hat{m}-\hat{\tau}_i}{\hat{\sigma}};\hat{\lambda}\right)+\left[1-\Phi_{SN}\left(\frac{y_{ij}-\hat{m}-\hat{\tau}_i}{\hat{\sigma}};\hat{\lambda}\right)\right]^{\hat{\alpha}}}\right\}, \tag{4.24}$$

where $\Phi(\cdot)^{-1}$ is the inverse cumulative standard normal distribution.

Atkinson (1985) suggested the construction of envelopes to enable better interpretation of the probability normal plot of the residuals. These envelopes are simulated confidence bands that contain the residuals, such that if the model is well-fitted, the majority of points will be within these bands and randomly distributed. The construction of the confidence bands follows the steps:

- Fit the proposed model and calculate the residuals $\widehat{qr}_{ij}$'s;
- Simulate k samples of the response variable using the fitted model;
- Fit the model to each sample and calculate the residuals $\widehat{qr}_{ij}$, $i = 1, \ldots, I$ and $j = 1, \ldots, J$;
- Arrange each sample of IJ residuals in rising order to obtain $\widehat{qr}_{(ij)k}$ for $k = 1, \ldots, K$;
- For each ij, obtain the mean, minimum and maximum $\widehat{qr}_{(ij)k}$, namely

$$\widehat{qr}_{(ij)M} = \sum_{k=1}^{K}\frac{\widehat{qr}_{(ij)k}}{K}, \quad \widehat{qr}_{(ij)B} = \min\{\widehat{qr}_{(ij)k} : 1 \leq k \leq K\}$$

$$\text{and} \qquad \widehat{qr}_{(ij)H} = \max\{\widehat{qr}_{(ij)k} : 1 \leq k \leq K\};$$

- Include the means, minimum and maximum together with the values of $\widehat{qr}_{ij}$ against the expected percentiles of the standard normal distribution.

The minimum and maximum values of $\widehat{qr}_{ij}$ form the envelope. If the model under study is correct, the observed values should be inside the bands and distributed randomly.

4.8 Aplications

In this section are three applications that show the flexibility of the new OLLSN distribution in relation to the, OLLN, SN and normal models.

4.8.1 Application 1: Temperature and production of soybeans data

- The first dataset refers to the temperature variable (°C) obtained from daily readings for the period from January 1 to December 31, 2011 for the city of Piracicaba. The data were supplied by the Department of Biosystems Engineering of the Luiz de Queiroz Higher School of Agriculture (ESALQ) associated with the University of São Paulo (USP). The temperatures were measured at the Piracicaba Meteorological Station, located at latitude 22°42'30"S, longitude 47°38'30"W and altitude of 546 meters.

- This second dataset refers to the production of soybeans in the municipality of Lucas do Rio Verde in the period from 1990 to 2012. This town is one of the 15 leading soybean producers in the state of Mato Grosso. The data are the crop yields in kilograms of beans per hectare (Kg/ha) obtained from the Brazilian Institute of Geography and Statistics (IBGE).

In order to compare the distributions, we consider some goodness-of-fit measures including the Akaike information criterion (AIC), consistent Akaike information criterion (CAIC) and Bayesian information criterion (BIC). The MLEs of the parameters μ, σ and λ and α are obtained using the Optim script in the R software. These estimates and the corresponding standard errors (SEs) are given in Table 4.2. These results inidicate that the OLLSN distribution has the lowest AIC, CAIC and BIC values among all fitted models, and so it could be chosen as the best model in both applications. Additionally, it is evident that the normal distribution presents the worst fit to the current data and then the proposed models outperform this distribution.

In addition to comparing the models, we consider LR statistics and formal tests. First, the OLLSN model includes some sub-models thus allowing their evaluation relative to each other and to a more general model. The values of the LR statistics for testing

Tabela 4.2 - MLEs, SEs (in parentheses) and information criteria for temperature and production of soybeans data.

Temperature	μ	σ	λ	α	AIC	CAIC	BIC
OLLSN	24.020	1.466	-2.530	0.228	1748.3	1748.5	1763.9
	(0.2778)	(0.2635)	(0.4037)	(0.0578)			
	μ	σ	α		AIC	CAIC	BIC
OLLN	21.9071	0.8915	0.1861		1790.4	1790.5	1802.1
	(0.1221)	(0.1326)	(0.1861)				
	μ	σ	λ				
SN	26.3603	4.9948	-9.7087		1764.7	1764.8	1776.4
	(0.1332)	(0.2139)	(2.4430)				
Normal	22.3271	2.9463			1828.7	1828.8	1836.5
	(0.1542)	(0.1090)					
Soybean	μ	σ	λ	α	AIC	CAIC	BIC
OLLSN	3134.400	160.919	-2.328	0.155	345.5	350.7	350.0
	(0.002)	(0.001)	(0.002)	(0.027)			
	μ	σ	α		AIC	CAIC	BIC
OLLN	2823.607	115.850	0.1522		348.3	351.8	351.7
	(15.331)	(14.811)	(0.0419)				
	μ	σ	λ				
SN	2863.500	437.520	-0.00006		351.0	354.5	354.4
	(15428.161)	(522.636)	(35.993)				
Normal	2863.478	437.529			349.0	351.2	351.2
	(91.231)	(64.509)					

some sub-models of the OLLSN distribution are given in Table 4.3. The figures in this table indicate the superiority of the OLLSN model in relation to the others to the current data in both applications.

Tabela 4.3 - LR tests.

Temperature	Hypotheses	LR statistic	p-value
OLLSN vs normal	$H_0 : \alpha = 1,\ \lambda = 0$ vs $H_1 : H_0$ is false	84.3	< 0.001
OLLSN vs OLLN	$H_0 : \alpha \neq 1,\ \lambda = 0$ vs $H_1 : H_0$ is false	43.9	< 0.001
OLLSN vs SN	$H_0 : \alpha = 1,\ \lambda \neq 0$ vs $H_1 : H_0$ is false	18.3	< 0.001
Soybean	Hypotheses	LR statistic	p-value
OLLSN vs normal	$H_0 : \alpha = 1,\ \lambda = 0$ vs $H_1 : H_0$ is false	7.5	0.024
OLLSN vs OLLN	$H_0 : \alpha \neq 1,\ \lambda = 0$ vs $H_1 : H_0$ is false	4.7	0.006
OLLSN vs SN	$H_0 : \alpha = 1,\ \lambda \neq 0$ vs $H_1 : H_0$ is false	7.4	0.023

In order to assess if the model is appropriate, the estimated pdf and cdf of the fitted distributions are displayed in Figures 4.6 and 4.7. From these plots, we conclude that the OLLSN model yields the best fits and that they could be adequate for these data

in both applications..

(a) (b)

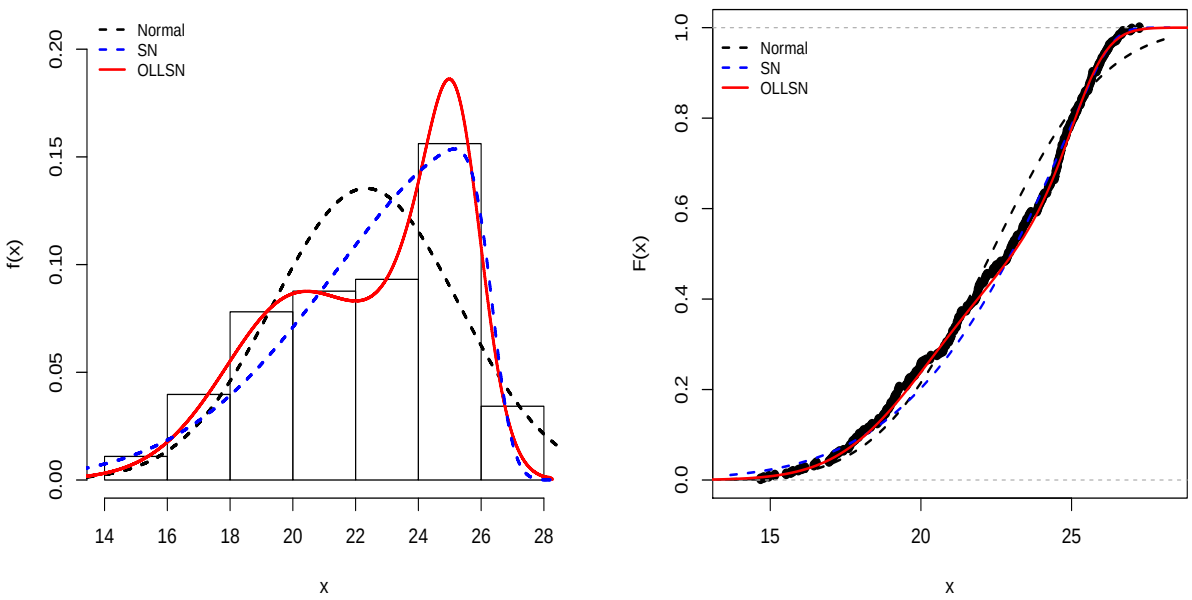


Figura 4.6 - (a) Estimated density of the OLLSN, OLLN, SN and normal models for the temperature data. (b) Estimated cdf and the empirical cdf of the OLLSN, OLLN, SN and normal models for the temperature data.

4.8.2 Application 2: Completely randomized design model - Soybean data

The experiment was conducted at Block 3, of the Geraldo Schultz Research Center, located in the municipality of Iracemápolis, São Paulo state, with average altitude of 570 m (longitude 47° 30' 10.81"W and latitude 22° 38' 49.14"S). The climate in the region is classified as Cwa according to the Köppen classification (tropical highlands, with rain mainly in the summer and dry winters). The soil was classified as Dystrophic Red Latosol according to the Brazilian Soil Classification System (Rhodic Hapludox according to the Soil Taxonomy). The objective of the experiment was to assess the level of nitrogen and protein in soybeans (Glycine max L. Merryl) under the effects of different treatments composed of the elements boron (B) and sulfur (S). The higher the concentration in the soil, the greater the uptake of these nutrients by the plants should be. The experimental design was completely randomized with 4 repetitions and 7 treatments. Each plot was composed of 6 rows with length of 7 m. The useful portion of the plot was composed of

(a) (b)

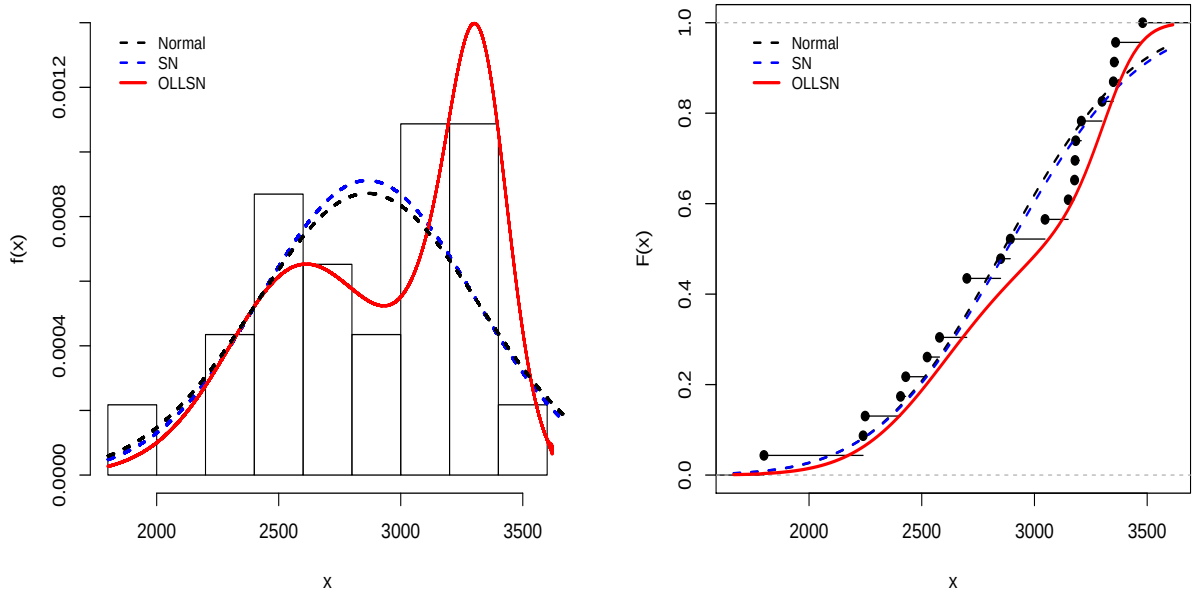


Figura 4.7 - (a) Estimated density of the OLLSN, OLLN, SN and normal models for the production of soybeans data. (b) Estimated cdf and the empirical cdf of the OLLSN, OLLN, SN and normal models for the production of soybeans data.

two rows with length of 5 m. The study was carried out by the company Produquimica, in Iracemápolis, São Paulo state, Brazil, during the 2014-2015 growing season.

- **Response variable** ($\mathbf{Y}_{ij}$): Concentration of nitrogen and protein the soybeans; $i = 1, \ldots, 7$ and $j = 1, \ldots, 4$,

- **Treatments**:

  - **Treat 1** - Controle;
  - **Treat 2** - Sulfurgran (elemental S);
  - **Treat 3** - Sulfurgran + Borosol (elemental S + boric acid);
  - **Treat 4** - Sufurgran + ActiveBor (elemental S + sodium octoborate);
  - **Treat 5** - Sulfurgran + Ulexite (elemental S + ulexite);
  - **Treat 6** - Sulfurgran + Produbor (elemental S + partially acidified ulexite);
  - **Treat 7** - Sulfurgran B-MAX (elemental S + ulexite in the same pellet).

According to the general model given in equation (4.22), the OLLSN regression model considering the dummy variables is given by

$$Y_{ij} = m + \sum_{k=2}^{I} \beta_k D_k + \epsilon_{ij}, \qquad k = 2, \ldots, 7, \tag{4.25}$$

where m is the effect of the treatment of reference for comparison, $D_2, D_3, D_4, D_5, D_6, D_7$ are the variables *dummies* for the treatment levels, the coefficients $\beta_2 = \tau_2 - \tau_1$, $\beta_3 = \tau_3 - \tau_1$,..., $\beta_7 = \tau_7 - \tau_1$ are the effects of treatment differences, τ_i is the effect of treatment i and $\epsilon_{ij} \sim \text{OLLSN}(0, \sigma, \lambda, \alpha)$ is the effect of the uncontrolled factors in the experimental group, for $i = 1, \ldots, 7$ and $j = 1, \ldots, 4$.

Figure 4.8 shows the average differences between the treatments. Then, using model (4.25), the mathematical expectations for each of the differences are given by:

- $E(Y_{ij}) = m$ se $\mathrm{D}_2 = \ldots = \mathrm{D}_7 = 0$ (represents the effect of the τ_1)
- $E(Y_{ij}) = m + \beta_2$ se $\mathrm{D}_2 = 1, \mathrm{D}_3 = \mathrm{D}_4, \ldots, = \mathrm{D}_7 = 0$ $(\beta_2 = \tau_2 - \tau_1)$
- $E(Y_{ij}) = m + \beta_3$ se $\mathrm{D}_3 = 1, \mathrm{D}_2 = \mathrm{D}_4, \ldots, = \mathrm{D}_7 = 0$ $(\beta_3 = \tau_3 - \tau_1)$
- $E(Y_{ij}) = m + \beta_4$ se $\mathrm{D}_4 = 1, \mathrm{D}_2 = \mathrm{D}_3, \ldots, = \mathrm{D}_7 = 0$ $(\beta_4 = \tau_4 - \tau_1)$
- $E(Y_{ij}) = m + \beta_5$ se $\mathrm{D}_5 = 1, \mathrm{D}_2 = \ldots = \mathrm{D}_7 = 0$ $(\beta_5 = \tau_5 - \tau_1)$
- $E(Y_{ij}) = m + \beta_6$ se $\mathrm{D}_6 = 1, \mathrm{D}_2 = \ldots = \mathrm{D}_7 = 0$ $(\beta_6 = \tau_6 - \tau_1)$
- $E(Y_{ij}) = m + \beta_7$ se $\mathrm{D}_7 = 1, \mathrm{D}_2 = \ldots = \mathrm{D}_6 = 0$ $(\beta_7 = \tau_7 - \tau_1)$

Table 4.4 lists the MLEs and SEs of the parameters for the OLLSN, SN and normal regression models fitted to these data using the Optim script in the R software.

In fitting model (4.22), a restriction on the solution was imposed, i.e., only the effect of $\tau_1 = 0$ was considered. Thus, the estimates of the parameters of the treatments (τ_2, τ_3, τ_4, τ_5, τ_6 e τ_7) represent mean differences in relation to treatment τ_1. This means that the interpretations should be carried out in relation to the treatment that was carried out with the restriction τ_1. Furthermore, for the three fitted models (Normal, SN and OLLSN), the estimates of the parameters are coherent with the results shown in Figure

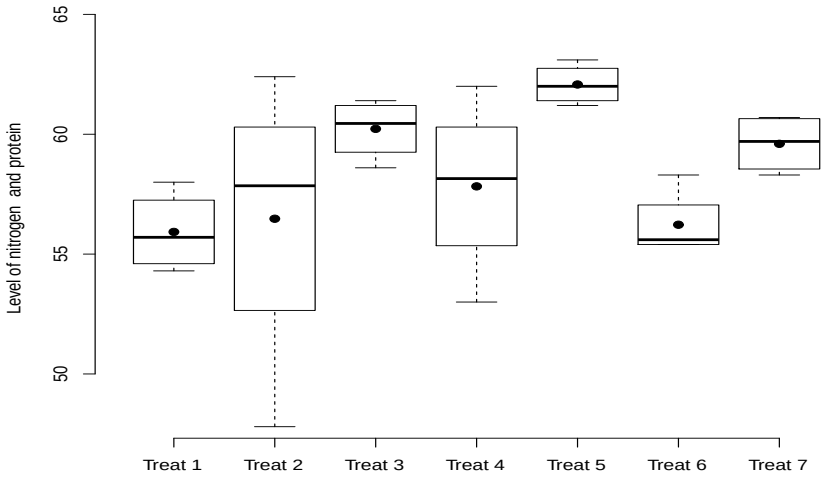


Figura 4.8 - Boxplot for each treatment where • is the average.

4.8, since the average effects of all the estimated differences present positive values, as can be verified in Table 4.4.

The values of the AIC, CAIC and BIC statistics to compare the OLLSN, SN and normal regression models are given in Table 4.4. Note that the OLLSN regression model outperforms the SN and normal models irrespective of the criteria and then the proposed regression model can be used effectively in the analysis of these data. A comparison of the OLLSN regression model with some of its sub-models using LR statistics is performed in Table 4.5. The figures in this table, specially the p-values, indicate that the OLLSN regression model yields better fits to these data than the other sub-models.

With respect to the estimates of the standard errors, the normal distribution presents equal values (1.799), indicating homogeneity within the treatments, something not verified in Figure 4.8. The results of the SN model are very close, while the OLLSN distributions provides letter variance within the treatments. For example, treatment (τ_2) was the one that presented the largest variance and also the highest standard error (1.501), while τ_5 presents the smallest variance and also the lowest standard error (1.293). Therefore, the OLLSN distribution presents smaller estimates of the standard errors than the normal and SN models.

Therefore, by using the OLLSN distribution to explain the concentration of nitro-

Tabela 4.4 - MLEs, SEs and information criteria for soybean data

Normal			SN			OLLSN		
$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE	$\boldsymbol{\theta}$	MLE	SE
m	55.925	1.272	m	55.904	8.283	m	52.318	5.531
β_2	0.550	1.799	β_2	0.549	1.799	β_2	2.062	1.501
β_3	4.300	1.799	β_3	4.299	1.799	β_3	4.356	1.320
β_4	1.899	1.799	β_4	1.899	1.799	β_4	2.306	1.485
β_5	6.150	1.799	β_5	6.149	1.799	β_5	6.166	1.293
β_6	0.300	1.799	β_6	0.300	1.799	β_6	0.221	1.326
β_7	3.675	1.799	β_7	3.674	1.799	β_7	3.714	1.323
σ	2.544	0.340	σ	2.544	0.346	σ	3.090	1.553
λ			λ	0.010	4.031	λ	0.501	0.196
α			α			α	2.516	1.557
AIC	CAIC	BIC	AIC	CAIC	BIC	AIC	CAIC	BIC
147.76	160.70	158.42	149.76	166.26	161.75	139.21	160.01	152.53

Tabela 4.5 - LR tests for soybean data.

Models	Hypotheses	LR statistic	p-value
OLLSN vs normal	$H_0 : \lambda = 0$ and $\alpha = 1$ vs $H_1 : H_0$ é false	12.5532	<0.001
OLLSN vs SN	$H_0 : \lambda \neq 0$ and $\alpha = 1$ vs $H_1 : H_0$ is false	12.8172	<0.001

gen and protein in soybeans, it can be stated that Treatment 6 (Sulfurgran + Produbor) presents the smallest average difference (0.221), which is not statistically significant, while Treatment 5 (Sulfurgran + ulexite) presents the largest average difference (6.166), which is statistically significant. In practical terms, it can be said that the higher the concentration of nitrogen and protein, the better the quality of soybeans is. Table 4.6 shows all the comparisons between the treatments. These comparisons are carried out by means of confidence intervals at 5% significance. The asterisks (*) indicate the existence of a statistically significant difference between average differences estimated for the treatments, at the 5% level, while (ns) indicates no significant difference.

The results of the Table 4.6 can be shown in a summary table. The comparing of the treatments it was made two to two in relation to the references (**a**, **b**, **c**, **d**), according

Tabela 4.6 - Results of the comparison of the 7 treatments for soybean data.

Hypotheses	Estimates	Lwr	Upr	Hypotheses	Estimates	Lwr	Upr
$H_0 : \tau_2 - \tau_1 = 0$	$2.062^{\mathbf{ns}}$	-0.34	4.47	$H_0 : \tau_4 - \tau_3 = 0$	$-2.059^{\mathbf{ns}}$	-4.38	0.26
$H_0 : \tau_3 - \tau_1 = 0$	4.365^{*}	2.24	6.48	$H_0 : \tau_5 - \tau_3 = 0$	$1.801^{\mathbf{ns}}$	-0.20	3.80
$H_0 : \tau_4 - \tau_1 = 0$	$2.307^{\mathbf{ns}}$	-0.07	4.69	$H_0 : \tau_6 - \tau_3 = 0$	-4.135^{*}	-6.19	-2.07
$H_0 : \tau_5 - \tau_1 = 0$	6.167^{*}	4.09	8.24	$H_0 : \tau_7 - \tau_3 = 0$	$-0.651^{\mathbf{ns}}$	-2.70	1.40
$H_0 : \tau_6 - \tau_1 = 0$	$0.232^{\mathbf{ns}}$	-1.89	2.36	$H_0 : \tau_5 - \tau_4 = 0$	3.856^{*}	1.568	6.14
$H_0 : \tau_7 - \tau_1 = 0$	3.715^{*}	1.59	5.83	$H_0 : \tau_6 - \tau_4 = 0$	$-2.081^{\mathbf{ns}}$	-4.414	0.25
$H_0 : \tau_3 - \tau_2 = 0$	$2.305^{\mathbf{ns}}$	-0.04	4.65	$H_0 : \tau_7 - \tau_4 = 0$	$1.401^{\mathbf{ns}}$	-0.928	3.73
$H_0 : \tau_4 - \tau_2 = 0$	$0.247^{\mathbf{ns}}$	-2.32	2.81	$H_0 : \tau_6 - \tau_5 = 0$	-5.937^{*}	-7.95	-3.92
$H_0 : \tau_5 - \tau_2 = 0$	4.110^{*}	1.79	6.42	$H_0 : \tau_7 - \tau_5 = 0$	-2.453^{*}	-4.46	-0.44
$H_0 : \tau_6 - \tau_2 = 0$	$-1.828^{\mathbf{ns}}$	-4.19	0.53	$H_0 : \tau_7 - \tau_6 = 0$	3.489^{*}	1.42	5.55
$H_0 : \tau_7 - \tau_2 = 0$	$1.655^{\mathbf{ns}}$	-0.70	4.01				

shown below:

$$\bar{X}_{Treat\,1} = 55.9\ \mathbf{a}$$
$$\bar{X}_{Treat\,2} = 56.4\ \text{a}\ \mathbf{b}$$
$$\bar{X}_{Treat\,3} = 60.2\ \ \text{b}\ \ \mathbf{c}$$
$$\bar{X}_{Treat\,4} = 57.8\ \text{a}\ \text{b}\ \ \text{c}\ \ \mathbf{d}$$
$$\bar{X}_{Treat\,5} = 62.0\ \ \ \ \text{c}\ \ \ \ \mathbf{e}$$
$$\bar{X}_{Treat\,6} = 56.2\ \text{a}\ \text{b}\ \ \ \ \text{d}\ \ \ \ \mathbf{f}$$
$$\bar{X}_{Treat\,7} = 59.6\ \ \text{b}\ \ \text{c}\ \ \text{d}$$

Therefore, averages followed by the same letter no differ significantly at the level of 5 %.

4.8.3 Residual analysis

To detect possible outlying observations by fitting the OLLSN regression models for soybean data, Figure 4.9(a) provides the index plot of rq_{ij}. By analyzing the quantile residual plot, one observation appears as a possible outlier, indicating that the model is well-fitted.

To detect possible departures from the assumptions of distribution errors for model (4.25) as well as outlying observations, we display in Figure 4.9(b) the normal probability plot for the quantile residual with the generated envelope, as suggested by (ATKINSON, 1985). As we can see, the plot in Figure 4.9(b) indicates that the OLLSN regression model for soybean data does not seem unsuitable to fit the data. Also, no observation appears as a possible outlier.

(a) (b)

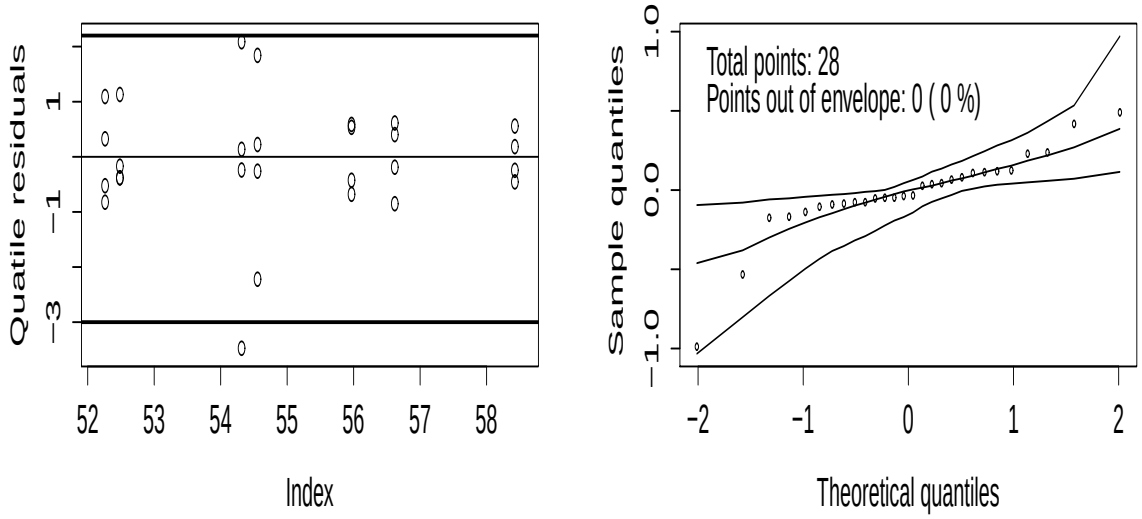


Figura 4.9 - (a) Index plot of the quantile residual for the soybean data. (b) Normal probability plot with envelope for the quantile residual from the fitted OLLSN regression model fitted to soybean data.

4.9 Concluding remarks

We introduce a four-parameter continuous distribution, called the odd log-logistic skew-normal (OLLSN) distribution, which extends the normal, skew-normal (SN) and the odd log-logistic normal (OLLN) distributions. The proposed distribution is more versatile than the SN and OLLN distributions, since it can be adjusted to bimodal data. We provide a mathematical treatment of the new distribution including expansions for the density function, moments, generating and quantile functions. The model parameters are estimated by the method of maximum likelihood. Further, based on the OLLSN distribution, we propose an extended regression model for completely randomized design. This extended regression model is very flexible and can be used in many practical situations. Three applications of the new models to real data are given to prove empirically that they can provide consistently better fits than other special models. Our formulas related to the new distribution and the extended regression model are manageable, and with the use of modern computer resources and their analytic and numerical capabilities, the proposed models may to be useful additions to the arsenal of applied statisticians.

References

ALEXANDER, C.; CORDEIRO, G. M.; ORTEGA, E. M. M.; SARABIA, J. M. (2012). Generalized beta-generated distributions. **Computational Statistics and Data Analysis**, 56, 1880-189.

ATKINSON, A. C. (1985). **Plots, transformations and regression: an introduction to graphical methods of diagnostics regression analysis**. 2nd ed. Oxford: Clarendon Press.

ARNOLD, B. C.; BEAVER, R. J.; GROENEVELD, R. A.; MEEKER, W. Q. (1993). The nontruncated marginal of a truncated bivariate normal distribution. **Psychometrika**, 58, 471-488.

AZZALINI, A. (1985). A class of distributions which includes the normal ones. **Scandinavian Journal of Statistics**, 12, 171-178.

AZZALINI, A.; CAPITANIO, A. (1999). Statistical application of the multivariate skew normal distribution. **Journal of the Royal Statistical Society**, 61, 579-602.

BRAGA, A. S.; CORDEIRO, G. M.; ORTEGA, E. M. M.; da CRUZ, J. N. (2016). The odd log-logistic normal distribution: Theory and applications in analysis os experiments. **Journal of Statistical Theory and Practice**, 10, 311-335.

BOURGUIGNON, M.; RODRIGO B. S.; GAUSS M. C. (2014). The Weibull-G family of probability distributions. **Journal of Data Science**, 12, 53-68.

BOX, G.; COX, D. R. (1964). An analysis of transformations. **Journal of the Royal Society**, 26, 211-252.

BRANCO, M. D.; DEY D. K. (2001). A general class of multivariate skew-elliptical distributions. **Journal of Multivariate Analysis**, 79, 99-113.

CASTELLARES, F.; CORDEIRO, G. M.; SANTOS, M. C.; MONTENEGRO, L. C. (2012). A power series expansion for the skew normal distribution function. **Journal of Applied Probability and Statistics**, 7, 1-17.

CORDEIRO, G. M.; CINTRA, R. J.; RÊGO, L. C.; ORTEGA, E. M. M. (2012). The McDonald Normal Distribution. **Pakistan Journal of Statistics and Operation Research**, 8, 301-329.

CORDEIRO, G. M.; MORAD A.; ORTRGA, E. M. M. (2014a). The exponentiated half-logistic family of distributions: Properties and applications. **Journal of Probability and Statistics**, 1, 1-21.

CORDEIRO, G. M.; ORTEGA, E. M.; POPOVIC, B. V.; PESCIM, R. R. (2014b). The Lomax generator of distributions: Properties, minification process and regression model. **Applied Mathematics and Computation**, 247, 465-486.

CORDEIRO, G. M.; DE CASTRO, M. (2011). A New family of generalized distribution. **Journal of Statistical Computation and Simulation**, 81, 883-893.

CYSNEIROS, F. J. A.; PAULA, G. A.; GALEA, M. (2007). Heteroscedastic symmetrical linear models. **Statistics and Probability Letters**, 77, 1084-1090.

DUNN P. K.; SMYTH G. K. (1996). Randomized quantile residuals. **Journal Computational and Graphical Statistics**, 5, 236-244.

EUGENE, N.; LEE, C.; FAMOYE, F. (2002). Beta-normal distribution and its applications. **Communications in Statistics - Theory and Methods**, **31**, 497-512.

FANG, K.; KOTZ, S.; NG, K. (1990). **Symmetric Multivariate and Related Distribution**. 36. ed. London: Chapman & Hall.

GLEATON, J. U.; LYNCH, J. D. (2006). Properties of generalized log-logistic families of lifetime distributions. **Journal of Probability and Statistical Science**, 4, 51-64.

GRADSHTEYN, I. S.; RYZHIK, I. M. (2000). **Table of Integrals, Series, and Products**, sixth edition. Academic Press, San Diego.

GUPTA, A. (2003). Multivariate skew t-distribution. **A Journal of Theoretical and Applied Statistics**, 37, 359-363.

HILL, M.; DIXON, W. J. (1982). Robustness in real life: A study of clinical laboratory data. **Biometrics**, 38, 377-396.

MARSHALL, A. W.; OLKIN I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. **Biometrika**, 84, 641-652.

NADARAJAH, S.; KOTZ, S. (2006). The exponentiated type distribution. **Acta Applicandae Mathematicae**, 92, 97-111.

PRUDNIKOV, A. P.; BRYCHKOV, Y. A.; MARICHEV, O. I. (1986). **Integrals and series: Elementary functions.** Gordon and Breach Science Publishers, New York.

PUNATHUMPARAMBATH, B. (2012). The multivariate skew-slash t and skew-slash Cauchy distributions. **Model Assisted Statistics and Appplications**, 7, 33-40.

SAHU, S. K.; DEY, D. K.; BRANCO, M. D. (2003). A new class of multivariate skew distributions with applications to bayesian regression models. **The Canadian Journal of Statistic**, 31, 129-150.

SHIRKE, D. T.; KAKADE, C. S. (2006). On exponentiated log-normal distribution. **International Journal of Agricultural and Statistical Science**, 2, 319-326.

TAHIR, M. H.; CORDEIRO, G. M.; ALZAATREH, A.; MANSOOR, M.; ZUBAIR, M. (2016). The Logistic-X family of distributions and its applications. **Communications in Statistics-Theory and Methods**, 45, 7326-7349.

WANG, J.; GENTON, M. G. (2006). The multivariate skew-slash distribution. **Journal os Statistical Planning and Inference**, 136, 209-220.

WANG, J.; BOYER, J.; GENTON, M. G. (2004). A skew-symmetric representation of multivariate distributions. **Statistica Sinica**, 14, 1259-1270.

WHITTAKER, E. T.; WATSON, G.N. (1990). **A Course in Modern Analysis**. Cambridge, University Press, Cambridge.

5 REGRESSION MODEL WITH RANDOM EFFECT OF THE ODD LOG-LOGISTIC SKEW NORMAL DISTRIBUTION: AN APPLICATION IN LONGITUDINAL DATA

Abstract

Regression models with random effects are often used when the data have a longitudinal or grouped structure, because they are more flexible to model the correlation between or within individuals. Normality of the random errors is a common assumption in regression models with random effects, but this may not be realistic and can hide important characteristics of the variation between and within individuals. In this paper, we neglect the normality assumption and consider that the errors follow a new model called the odd log-logistic skew-normal (OLLSN) distribution. This distribution includes the normal and skew-normal as special cases. The estimates of the parameters of the proposed model are obtained using the maximum likelihood method. We illustrate the performance of the OLLSN regression model with random effect by means of a real longitudinal data set (Postharvest storage of lychee). Using the proposed model, we evaluate the physicochemical, physiological, biochemical and sensorial answers associated to the techniques of postharvest lychee conservation.

Keywords: Log-logistic distribution; maximum likelihood estimation; method numeric; mixed model; skew distribution.

5.1 Introduction

The normal regression model with random effect has a structure that is often utilized to analyze data from repeated, grouped and longitudinal measures (among others). They have wide applicability in areas like agriculture, biology, economics and medicine, among others. Studies are often conducted with experimental designs that draw samples about the same experimental unit, such as evaluation of the growth of animals or the time for a determined fruit species to ripen under the effect of a treatment, etc.

These experimental tests allow evaluating the changes that occur over time, such as (DIGGLE, 1988), (CROWDER; HAND, 1990) and (FERREIRA; MORAIS, 2013). Because the measures are performed on the same experimental units, correlation between the measures in time is expected, as well as lack of homogeneity of the variances. In this sense, it is possible to consider responses measured at shorter intervals to be more strongly correlated than those measured at longer intervals. This is a common feature of data measured over time; see (LITTELL; HENRY; AMMERMAN, 1998).

The growing popularity of these regression models with random effect is explained by the flexibility they offer to model the correlation between and within sample units. These models accommodate a variance and covariance structure to model the correlation that exists in longitudinal, grouped, balanced and unbalanced data. However, in such studies, depending on the structure presented by the data, the number of parameters to be estimated can be very large, often being impossible to estimate it. However, for less complex models, the majority of the inferences for the fixed parameters and random effects can be treated with a wide variety of computational programs presented in the literature.

Furthermore, although these models offer more modeling flexibility see (VERBEKE; LESAFFRE, 1997), they must satisfy the assumptions of analysis of variance for their results to be correct. The most common assumptions required are that both the errors and random effects must be normally distributed and homoscedastic. In this case, the method most often adopted to attain a normal distribution is transformation of the response variable, which works well in many cases (BOX; COX, 1964). Although this model can lead to reasonable result, it should be avoided when a more suitable theoretical model can be found, according to (AZZALINI; CAPITANIO, 1999).

An interesting alternative to the technique presented above is to fit the regression model with random effect based on generalizations of the normal distribution, meaning that the aim is to deal with distributions that show good behavior and model asymmetric and bimodal data besides symmetric data.

In this respect, the normal-asymmetric family of distributions is important to accommodate measures of asymmetry and kurtosis. It can be used to model data so that the statistical analysis is more robust. Many authors, such as (AZZALINI; CAPITANIO, 1999), (AZZALINI; KOTZ, 2002), (SAHU; DEY, 2001) and (SAHU *et al.*, 2003), have studied regression models with asymmetric distributions. However, the literature does not contain articles on families of distributions that model bimodality and asymmetry jointly.

In this work, we propose to study regression models with random effect assuming that the errors follow a new distribution called the *odd log-logistic skew-normal* (OLLSN). It allows modeling asymmetric and bimodal data, and also has the normal and

normal asymmetric distributions as particular cases. This distribution can be symmetric, asymmetric, platykurtic, leptokurtic, unimodal or asymmetric bimodal depending on the parameter values. These properties make the distribution more flexible, allowing modeling asymmetry, kurtosis and bimodality.

To motivate this paper, the regression models with random intercept are fitted to data on antioxidants of Bengal lychees. These data came from a completely randomized experiment in which the measures are obtained over time. The estimates of the parameters were obtained using the maximum likelihood method with computational functions implemented in the R software. Besides this, we propose two studies of standardized residuals: marginal quantile residuals and conditional quantile residuals; then simulated envelope for marginal quantile residuals are built. The new proposed model is very useful and gives better results than the normal and skew-normal models. Therefore, the proposed model can be recommended to analyze data with the presence of asymmetry, kurtosis or bimodality.

The chapter is organized as follows. In Section 5.2, we introduce of the OLLSN regression models with random effect. In Section 5.3, we estimate the model parameters by a maximum likelihood. In Section 5.4, a kind of quantile residual is proposed to assess departures from the underlying OLLSN distribution to detect outliers. In Section 5.5, we fit a real longitudinal data set (Postharvest storage of lychee) to prove empirically the potentiality of the new model. Finally, Section 5.6 ends with some conclusions.

5.2 The OLLSN model

In statistics, the normal distribution is the most popular model in applications to real data. When the number of observations is large, it can serve as a good approximation to other models. In the last decade, the skew-normal distribution (AZZALINI, 1985) has been widely used to model asymmetric data. Let $g(z)$ and $G(z)$ be the probability density function (pdf) and cumulative distribution function (cdf) of the skew-normal (for $x \in \mathbb{R}$) model given by

$$g(z;\lambda) = 2\,\phi\,(z)\,\Phi\,(\lambda\,z) \tag{5.1}$$

and

$$G(z;\lambda) = \Phi\left(z\right) - 2\,T\left(z;\lambda\right) = \Phi_{SN}\left(z;\lambda\right), \tag{5.2}$$

where $z = (x-\mu)/\sigma$, $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a dispersion parameter, and $T(z;\lambda)$ is the Owen's function given by (for $z, \lambda \in \mathbb{R}$)

$$T(z;\lambda) = (2\pi)^{-1} \int_0^{\lambda} \frac{\exp\left\{-\frac{1}{2}z^2(1+x^2)\right\}}{1+x^2} dx.$$

The parameter λ regulates the skewness and it varies in $\mathbb{R}$. For $\lambda = 0$, we have the $N(\mu, \sigma^2)$ distribution. The function $\Phi_{SN}(\cdot)$ denotes the standard skew-normal cdf.

Based on the transformer odd log-logistic generator (GLEATON; LYNCH, 2006), we propose a new continuous distribution called the OLLSN model by integrating the log-logistic density function. It has cdf given by

$$F(z;\lambda,\alpha) = \int_0^{\frac{\Phi_{SN}(z;\lambda)}{1-\Phi_{SN}(z;\lambda)}} \frac{\alpha\, t^{\alpha-1}}{(1+t^{\alpha})^2} dt = \frac{\Phi_{SN}^{\alpha}\left(z;\lambda\right)}{\Phi_{SN}^{\alpha}\left(z;\lambda\right) + \left[1-\Phi_{SN}\left(z;\lambda\right)\right]^{\alpha}}, \tag{5.3}$$

where $\alpha > 0$ is an extra shape parameter.

The pdf corresponding to (5.3) is given by

$$f(z;\lambda,\alpha) = \frac{2\alpha\phi\left(z\right)\Phi\left(\lambda\, z\right)\Phi_{SN}^{\alpha-1}\left(z;\lambda\right)\left[1-\Phi_{SN}\left(z;\lambda\right)\right]^{\alpha-1}}{\left\{\Phi_{SN}^{\alpha}\left(z;\lambda\right) + \left[1-\Phi_{SN}\left(z;\lambda\right)\right]^{\alpha}\right\}^2}. \tag{5.4}$$

Hereafter, a random variable Z with density function (5.4) is denoted by $Z \sim \text{OLLSN}(0,1,\lambda,\alpha)$. Clearly, the random variable $X = \mu + \sigma Z$ follows the $\text{OLLSN}(\mu,\sigma^2,\lambda,\alpha)$ distribution.

The density function (5.4) allows greater flexibility of its tails and can be widely applied in many areas of engineering and biology. The normal distribution is a special case of (5.4) when $\lambda = 0$ and $\alpha = 1$. If $\lambda \neq 0$ and $\alpha = 1$, we obtain the SN distribution and, for $\lambda = 0$ and $\alpha \neq 1$, it reduces to the OLLN distribution (Braga *et al.*, 2016), respectively.

We can write

$$\alpha = \frac{\log\left\{\Phi_{SN}^{\alpha}(z;\lambda)/[1-\Phi_{SN}(z;\lambda)]^{\alpha}\right\}}{\log\left\{\Phi_{SN}(z;\lambda)/[1-\Phi_{SN}(z;\lambda)]\right\}},$$

and then the parameter α represents the quotient of the log odds ratio for the generated and baseline distributions.

Equation (5.3) has tractable properties specially for simulations, since the quantile function of Z has a simple form. Let $F(z;\lambda,\alpha) = u$ and $\Phi_{SN}^{-1}(z;\lambda)$ be the inverse function of $\Phi_{SN}(z;\lambda)$. We have

$$Q(u) = Q_{SN}\left[h(u,\alpha);\lambda\right], \tag{5.5}$$

where $u \sim U(0,1)$ and $Q_{SN}[h(u,\alpha);\lambda] = \Phi_{SN}^{-1}\left[h(u,\alpha);\lambda\right]$ is the qf of the SN distribution at

$$h(u,\alpha) = u^{\frac{1}{\alpha}}\left[u^{\frac{1}{\alpha}} + (1-u)^{\frac{1}{\alpha}}\right]^{-1}.$$

It is not possible to study the behavior of the parameters of the OLLSN distribution by taking derivatives. We can verify skew-bi-modality of the new distribution in the plots of Figure 5.1 by combining some values of λ, α and μ. Figure 5.1 (b) reveal the different types of asymmetrical bi-modality.

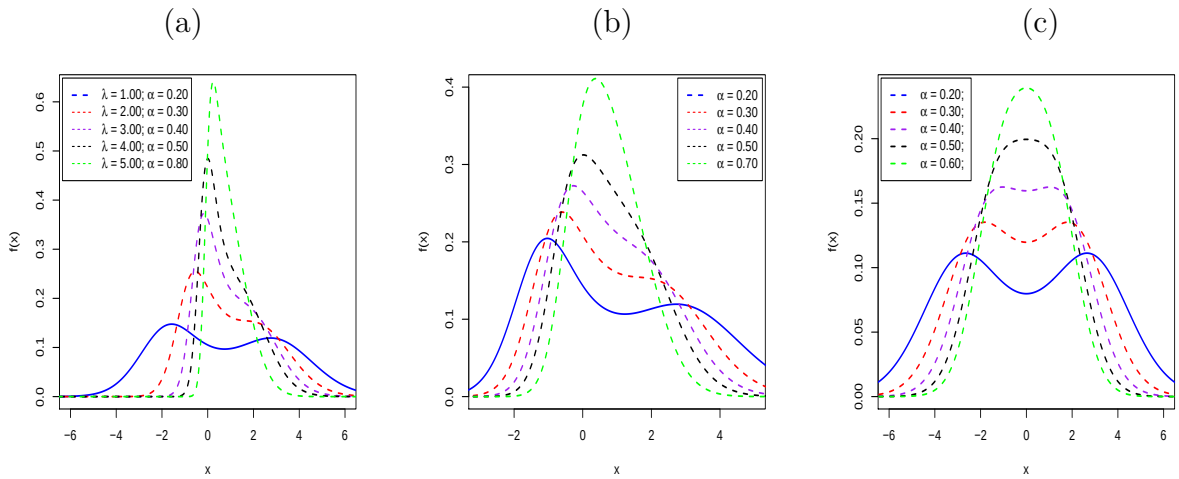


Figura 5.1 - Plots of the OLLSN density function for some parameter values. (a) For different values of α and λ with $\mu = 0$ and $\sigma = 1$. (b) For different values of α and $\lambda = 1.8$, $\mu = 0$ and $\sigma = 1$. (c) For different values of α with $\lambda = 0$, $\mu = 0$ and $\sigma = 1$.

Let $f(z;0,\alpha)$ be the density of the OLLN$(0,1,\alpha)$ distribution. By using $\Phi(-z) = 1-\Phi(z)$, we can prove that $f(-z;0,\alpha) = f(z;0,\alpha)$. So, we prove that the OLLSN distribution is symmetric about zero for $\lambda = 0$, and then the parameters σ and α characterize the kurtosis and skew-bi-modality of this distribution.

5.2.1 The OLLSN regression model with random effect

Let Y_{ijk} be the response variable associated with the i-th individual ($i = 1,\ldots,m$), for $j = 1,\ldots,n_i$ (for example, indicating measures repeated over time) and $k = 1,\ldots,K$ repetitions. Then, individual i can be represented by the vector of response variables $\mathbf{Y}_i = [Y_{i11},\ldots,Y_{i1k};Y_{i21},\ldots,Y_{i2k};\ldots;Y_{in_i1},\ldots,Y_{in_ik}]^T$. So, assuming the OLLSN model given by (5.4) with random effect, we have the following hierarchical structure:

$$Y_{ijk}|\mathrm{w}_i \overset{iid}{\sim} OLLSN(\mu_{ij},\sigma^2,\lambda,\alpha) \text{ and } w_i \overset{iid}{\sim} N(0,\sigma^2_{\beta_0}). \tag{5.6}$$

On the other hand, response variable can be related to other variables, known as covariables. Assuming that all the response from the same individual have a common random effect, denoted by w_i, and considering that the random effects are unobserved random variables, the regression model for correlated data is expressed in the following form:

$$Y_{ijk} = \mathrm{x}^T_{ijk}\boldsymbol{\beta} + w_i + \sigma z_{ijk}, \tag{5.7}$$

where $\mathbf{x}_{ijk}$ is the vector of covariables ($p\times 1$), the $Z_{ijk}'s$ are independent and identically distributed random errors with having the OLLSN distribution, i.e., $z_{ijk} \sim$ OLLSN$(0,\sigma^2,\lambda,\alpha)$, and the elements of the vector $(\boldsymbol{\beta}^\top,\sigma,\sigma_{\beta_0},\lambda,\alpha)^\top$ are unknown parameters. Each individual i has a random effect w_i, which is represented by i.d.d. random variables with normal distribution, i.e., $w_i\sim$ N$(0,\sigma^2_{\beta_0})$.

In this structure, equation (5.7) is called the OLLSN regression model with random effect. Considering the conditional probability function $Y_{ijk}|w_i$ for $f(y_{ijk}|w_i)$ and the pdf $f(w_i;\sigma^2_{\beta_0})$ for w_i, we have that the joint probability function (Y_{ijk},w_i) is given by$f(y_{ijk},w_i;\boldsymbol{\theta}) = f(y_{ijk}|w_i)f(w_i;\sigma^2_{\beta_0})$, thus resulting in some important structures:

i) $Y_{ijk}|w_i \backsim \text{OLLSN}(\mu_{ijk}, \sigma, \lambda, \alpha)$, with conditional density function given by

$$f(y_{ijk}|w_i) = \frac{2\,\alpha\,\phi\left(\frac{y_{ijk}-\mu}{\sigma}\right)\Phi\left[\lambda\left(\frac{y_{ijk}-\mu}{\sigma}\right)\right]\Phi_{SN}^{\alpha-1}\left[\left(\frac{y_{ijk}-\mu}{\sigma}\right);\lambda\right]\Gamma^{\alpha-1}}{\sigma\left\{\Phi_{SN}^{\alpha}\left[\left(\frac{y_{ijk}-\mu}{\sigma}\right);\lambda\right]+\Gamma^{\alpha}\right\}^2}, \tag{5.8}$$

where $\Gamma = \left\{1 - \Phi_{SN}\left[\left(\frac{y_{ijk}-\mu}{\sigma}\right);\lambda\right]\right\}$;

ii) $\mu_{ijk} = \mathrm{x}_{ijk}^T\boldsymbol{\beta} + w_i$ represents the conditional expectation, where w_i denotes the random effects associated with the i-th group of the regression model (5.7);

iii) The random effect distribution is $w_i \sim N(0, \sigma^2_{\beta_0})$ defined by

$$f(w_i, \sigma^2_{\beta_0}) = \frac{1}{\sigma_{\beta_0}\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{w_i}{\sigma_{\beta_0}}\right)^2\right\}, \tag{5.9}$$

where $w_i \in \mathbb{R}$ and the parameter $\sigma_{\beta_0} > 0$ represents the random effect intercept term.

Thus, the model (5.7) considers that the random effect has the normal distribution. The normal random-effects distribution is a special case of (5.6) when $\lambda = 0$ and $\alpha = 1$. If, $\lambda \neq 0$ and $\alpha = 1$, we obtain the skew normal random-effects distribution and for $\lambda = 0$ and $\alpha \neq 1$ reduces to the OLLN normal random-effects distribution.

5.3 Maximum likelihood estimation

We estimate the parameters of the OLLSN regression model with random effect using the maximum likelihood method. In this case, the method consists of maximizing the logarithm of the marginal likelihood function obtained by integrating the likelihood function in relation to the random effects w_i. For each individual i, the vector of the response variable is represented by $\mathbf{Y}_i = [Y_{i11}, \ldots, Y_{i1k}; Y_{i21}, \ldots, Y_{i2k}; \ldots; Y_{in_i1}, \ldots, Y_{in_ik}]^T$. The likelihood function conditional on the random effect (independence within the individuals) for the individual i is given by

$$L_i(y_{ijk}|w_i) = \prod_{j=1}^{n_i}\prod_{k=1}^{K} f(y_{ijk}|w_i), \tag{5.10}$$

where $f(y_{ijk}|w_i)$ is the pdf defined in (5.8). In this way, assuming that the terms w_i and Y_{ijk} are independent random variables, the contribution of the i-th individual to the

marginal likelihood function is

$$\int L_i(y_{ijk}|w_i)\, f(w_i;\sigma^2_{\beta_0})dw_i,$$

where $f(.)$ is the pdf of the random effect defined in equation (5.9) and $L_i(y_{ijk}|w_i)$ is given in equation (5.10).

Hence, under the assumption of independence between the groups, the marginal likelihood function considering the vector of parameters $\boldsymbol{\theta} = (\mu_{ijk}, \sigma^2, \sigma^2_{\beta_0}, \lambda, \alpha)^\top$ is given by

$$L(\boldsymbol{\theta}) = \frac{1}{\sigma_{\beta_0}\sqrt{2\pi}} \prod_{i=1}^{m} \int \prod_{j=1}^{n_i} \prod_{k=1}^{K} f(y_{ijk}|w_i)\, \exp\left\{-\frac{1}{2}\left(\frac{w_i}{\sigma_{\beta_0}}\right)^2\right\} dw_i. \tag{5.11}$$

So, for a dataset consisting of n observations, $(y_{11k}, \mathbf{x}_{11k}), \ldots, (y_{1n_ik}, \mathbf{x}_{1n_ik}), \ldots,$ $(y_{m1k}, \mathbf{x}_{m1k}), \ldots, (y_{mn_ik}, \mathbf{x}_{mn_ik})$, where $n = n_1 + \ldots + n_i$ and $\mathbf{x}_{ijk}$ is the vector of covariables associated with the j-th observation of the i-th individual and k repetitions, the logarithm of the marginal likelihood function in (5.11), supposing the distribution for the random effect, is expressed in the following form

$$\begin{aligned} l(\boldsymbol{\theta}) \;=\; & \sum_{i=1}^{m} \log \left\{ \int \prod_{j=1}^{n_i} \prod_{k=1}^{K} \frac{2\alpha\phi\left(z_{ijk}\right)\Phi\left[\lambda z_{ijk}\right]\Phi_{SN}^{\alpha-1}\left(z_{ijk},\lambda\right)\left[1-\Phi_{SN}\left(z_{ijk},\lambda\right)\right]^{\alpha-1}}{\sigma\left\{\Phi_{SN}^{\alpha}\left(z_{ijk},\lambda\right)+\left[1-\Phi_{SN}\left(z_{ijk},\lambda\right)\right]^{\alpha}\right\}^2} \times \right. \\ & \left. \frac{1}{\sigma_{\beta_0}\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{w_i}{\sigma_{\beta_0}}\right)^2\right\} dw_i \right\}, \end{aligned} \tag{5.12}$$

where $z_{jik} = (y_{jik} - \mathbf{x}^\top_{jik}\boldsymbol{\beta} - w_i)/\sigma$

The problem of maximizing the likelihood function defined in (5.12) is the presence of m integrals of the random effects w_i, which in many situations do not have analytic solutions. Therefore, to estimate the parameters of the models with random effect, it is necessary to use numerical integration methods or estimation methods that can in some way resolve the integrals, such as Bayesian methods or methods that use the EM algorithm (VAIDA; XU, 2000). We consider the Gauss-Hermite numerical integration method, described by (BONAT *et al.*, 2012). This method is an extension of the Gaussian

quadrature method to resolve integrals of the form

$$\int_{-\infty}^{+\infty} g(x)\,dx = \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2} f(x)\,dx. \tag{5.13}$$

In this case, this integral is approximated by a weighted sum of the function evaluated at the Gauss points with corresponding integration weights, i.e.

$$\int_{-\infty}^{+\infty} g(x)\,dx \approx \sum_{p=1}^{q} v_p\,\mathrm{e}^{s_p^2}\,g(s_p),$$

where q is the number of points employed for the approximation. The values of s_p are the roots of the Hermite polynomial $H(x)(p = 1, \ldots, q)$ and the associated weights v_p are given by

$$v_p = \frac{2^{q-1} q! \sqrt{\pi}}{q^2 [H_{q-1}(s_p)]^2}. \tag{5.14}$$

To approximate integrals by the Gauss-Hermite method it is necessary to obtain the integration weights v_p and the Gauss points s_p. The function `gauss.quad()` of the `statmod` routine of the `R` statistical software calculates the weights and Gauss-Hermite points.

In this work, we consider adaptive Gauss-Hermite numerical integration methods. The authors (LESAFFRE; SPIESSENS, 2001) studied models with random intercept using adaptive Gauss-Hermite quadrature, and also found that the adaptive version is more computationally efficient than the non-adaptive version. Therefore, for optimizing $g(x)$, supported by the appropriate form of the function (5.13), we have the solution:

$$\hat{\mu} = \arg\max_x g(x) \;\text{ and }\; \hat{\sigma}^2 = \left\{ \frac{d^2}{dx^2} [\log g(x)] \right\}^{-1} \Bigg|_{x=\hat{\mu}}, \tag{5.15}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ are the mean and variance of the distribution, respectively. In this case, the adaptive version can be presented in the following form:

$$\int_{-\infty}^{\infty} g(x)dx = \int_{-\infty}^{\infty} f(x)v(x)dx, \quad \text{where} \quad v(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2} \left(\frac{x-\hat{\mu}}{\hat{\sigma}} \right)^2 \right\}. \tag{5.16}$$

Resolving the integral (5.16) and using expression (5.13), we conclude that the adaptive

Gauss-Hermite quadrature can be expressed by (see Appendix G for more details)

$$\int_{-\infty}^{\infty} g(x)dx \approx \sqrt{2}\hat{\sigma}\sum_{p=1}^{q} v_p^+ g(s_p^+) \tag{5.17}$$

where $v_p^+ = v_p e^{s_p^2}$ and $s_p^+ = \sqrt{2}\hat{\sigma}s_p + \hat{\mu}$.

Therefore, by using the adaptive Gauss-Hermite quadrature numerical integration method, the log-likelihood function reduces to

$$\begin{aligned} l(\boldsymbol{\theta}) &= \sum_{i=1}^{m}\log\left\{\frac{\hat{\sigma}2\sqrt{2}}{\sigma_{\beta_0}\sqrt{\pi}}\sum_{p=1}^{q} v_p^+ \prod_{j=1}^{n_i}\prod_{k=1}^{K}\frac{\alpha\phi\left(z_{ijk}\right)\Phi\left[\lambda z_{ijk}\right]\Phi_{SN}^{\alpha-1}\left(z_{ij},\lambda\right)\left[1-\Phi_{SN}\left(z_{ijk},\lambda\right)\right]^{\alpha-1}}{\sigma\left\{\Phi_{SN}^{\alpha}\left(z_{ijk},\lambda\right)+\left[1-\Phi_{SN}\left(z_{ijk},\lambda\right)\right]^{\alpha}\right\}^2}\right. \\ &\quad \left.\times\exp\left\{-\frac{1}{2}\left(\frac{s_p^+}{\sigma_{\beta_0}}\right)^2\right\}dw_i\right\}, \end{aligned} \tag{5.18}$$

where $z_{jik} = (y_{jik} - \mathbf{x}_{jik}^\top\boldsymbol{\beta} - s_p^+)/\sigma$, $v_p{}'s$ are the weights and $s_p{}'s$ are the Gauss points that defined in (5.14).

Hence, by using the a quasi-Newton "BFGS" method in R, it is possible to maximize the marginal likelihood functions (5.18) and thus obtain the the maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$.

Under standard regularily conditions, the asymptotic distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is multivariate normal $N_{p+4}(0, K(\boldsymbol{\theta})^{-1})$, where $K(\boldsymbol{\theta})$ is the total information matrix. The asymptotic covariance matrix $K(\boldsymbol{\theta})^{-1}$ of $\widehat{\boldsymbol{\theta}}$ can be approximated by the inverse of the $(p+4)\times(p+4)$ observed information matrix $-\ddot{\mathbf{L}}(\boldsymbol{\theta})$. The elements of this matrix can be determined by simple double differentiation of $l(\boldsymbol{\theta})$ with respect to the model parameters and them evaluated numerically. Then, an asymptotic confidence interval with significance level γ for each parameter θ_r is given by

$$IC_r = \left(\hat{\theta}_r - z_{\gamma/2}\sqrt{-\widehat{L}^{r,r}}, \hat{\theta}_r + z_{\gamma/2}\sqrt{-\widehat{L}^{r,r}}\right),$$

where$-\widehat{L}^{r,r}$ is the rth diagonal element of $-\ddot{L}(\widehat{\boldsymbol{\theta}})^{-1}$ estimated at $\hat{\boldsymbol{\theta}}$, for $r = 1, \ldots, p+4$, and $z_{\gamma/2}$ is the quantile $1-\gamma/2$ of the standard normal distribution.

To maximize the log-likelihood function in (5.18) computational routines were performed in the R statistical software. In addition to the goodness of fit statistics, we obtain by means of the "L-BFGS-B" or "Nelder-Mead" options the parameter estimates and their standard errors, which are the square roots of the diagonal entries of the estimated covariance matrix $-\ddot{\mathbf{L}}(\boldsymbol{\theta})^{-1}$. The initial values are chosen as $\hat{\boldsymbol{\theta}} = (\hat{\mu}_{ijk}, \hat{\sigma}^2)^\top$, after, $\hat{\boldsymbol{\theta}} = (\hat{\mu}_{ijk}, \hat{\sigma}^2, \hat{\lambda})^\top$ and finally,

$\hat{\boldsymbol{\theta}} = (\hat{\mu}_{ijk}, \hat{\sigma}^2, \hat{\sigma}^2_{\hat{\beta}_0}, \hat{\lambda}, \hat{\alpha})^\top$, are obtained from the "*summary*" procedure in software GAMLSS.

We can supply the R script to anyone interested. Many works have shown how to predict the random effect. The authors (VILLEGAS; PAULA; LEIVA, 2011), (FABIO; PAULA; CASTRO, 2011) and (HASHIMOTO, 2013), for example, used the Bayes estimator. In this study, we obtained the predicted values w_i for the OLLSN distribution with the GAMLSS computational package. By this means, using the function $ranef(\cdot)$ with appropriate arguments, predicted values for the fitted model are obtained.

5.4 Residual analysis

Analysis of the residuals is an important step in statistical modeling, since it is an efficient way to detect many types of problems in the fitted model. Among the problems that can be checked are the presence of discrepant observations, deviations from suppositions of the errors and heteroscedasticity of variances. The authors (WEST *et al.*, 2007) studied the marginal and conditional residuals by the standard deviation of the fitted model. Based on this, they proposed marginal quantile residuals $\hat{rq}_{smarg(ijk)}$ and conditional quantile residuals $\hat{rq}_{sc(ijk)}$, then the simulated envelope for marginal quantile residuals are constructed.

The residuals $\hat{rq}_{smarg(ijk)}$ are used to check the linearity between the response variable Y and the time variable in observation ijk. In this way, we plot the estimated values $\hat{rq}_{smarg(ijk)}$ versus the time covariable in repetition j of each individual i and repetition k. When random behavior around zero exists, a linear relationship is expected according to (NOBRE; SINGER, 2007).

The residuals $\hat{rq}_{sc(ijk)}$ are used to check the assumption of homoscedasticity. We then plot the estimated values $\hat{rq}_{oc(ijk)}$ versus the response variable estimated for time in the repeated measure j of each individual i and repetition k. In this case, homoscedasticity is expected when random behavior around zero exists according to (PINHEIRO; BATES, 1995).

The marginal quantile residuals $\hat{rq}_{smarg(ijk)}$ considering the OLLSN regression model with random effect (5.7) are given by

$$\hat{rq}_{smarg(ijk)} = \Phi^{-1}\left\{\frac{\Phi^{\hat{\alpha}}_{SN}\left(\frac{y_{ijk}-\hat{\mu}_{ijk}}{\hat{\sigma}}, \hat{\lambda}\right)}{\Phi^{\hat{\alpha}}_{SN}\left(\frac{y_{ijk}-\hat{\mu}_{ijk}}{\hat{\sigma}}, \hat{\lambda}\right) + \left[1 - \Phi_{SN}\left(\frac{y_{ijk}-\hat{\mu}_{ijk}}{\hat{\sigma}}\right), \hat{\lambda}\right]^{\hat{\alpha}}}\right\}, \tag{5.19}$$

where $\hat{\mu}_{ijk} = \mathrm{x}^T_{ijk}\hat{\boldsymbol{\beta}}$ corresponds to the marginal average estimated for the observation ijk of the

model (5.6) and $\Phi(\cdot)^{-1}$ is the inverse cumulative standard normal distribution. The residuals $\hat{rq}_{smarg(ijk)}$ depend on the vector of estimated parameters $\widehat{\boldsymbol{\theta}}$ and the predicted values of the random effects w_i .

The conditional quantile residuals $\hat{rq}_{sc(ijk)}$ are important because they incorporate the effects of the presence of atypical observations in the fitted model. Their expression is given by:

$$\hat{rq}_{sc(ijk)} = \Phi^{-1}\left\{\frac{\Phi_{SN}^{\hat{\alpha}}\left(\frac{y_{ijk}-\hat{\mu}_{ijk}}{\hat{\sigma}}, \hat{\lambda}\right)}{\Phi_{SN}^{\hat{\alpha}}\left(\frac{y_{ijk}-\hat{\mu}_{ijk}}{\hat{\sigma}}, \hat{\lambda}\right) + \left[1 - \Phi_{SN}\left(\frac{y_{ijk}-\hat{\mu}_{ijk}}{\hat{\sigma}}\right), \hat{\lambda}\right]^{\hat{\alpha}}}\right\}, \tag{5.20}$$

where $\hat{\mu}_{ijk} = \mathrm{x}_{ijk}^T\hat{\boldsymbol{\beta}} + \hat{w}_i$ corresponds to the marginal average estimated for the observation ijk of model (5.6). The residuals $\hat{rq}_{sc(ijk)}$ depend on the vector of estimated parameters $\widehat{\boldsymbol{\theta}}$ and the predicted values of the random effects w_i.

Atkinson (1985) suggested the construction of envelopes to enable better interpretation of the normal plot of probabilities of the residuals. These envelopes are simulated confidence bands that contain the residuals, such that if the model is well fitted, the majority of points will be within these bands and randomly distributed. The construction of the confidence bands follows the steps:

- Fit the proposed model and calculate the residuals $rq_{smarg(ijk)}$'s for each individual i, in the j-th repeated measure in time and repetition k;
- Simulate n samples of the response variable using the fitted model;
- Fit the model to each sample and calculate the residuals $rq_{smarg(ijk)}$, $i = 1, \ldots, m$, and $j = 1, \ldots, n_i$ with repetition k;
- Arrange each sample of IJK residuals in rising order to obtain $y_{smarg(ijk)n}$ for $n = 1, \ldots, N$;
- For each IJK, obtain the mean (M), minimum (B) and maximum (H) $y_{smarg(ijk)n}$, namely $y_{smarg(ijk)M} = \sum_{n=1}^{N} \frac{y_{smarg(ijk)n}}{N}$; $y_{smarg(ijk)B} = \min\{y_{smarg(ijk)n} : 1 \leq n \leq N\}$ and $y_{smarg(ijk)H} = \max\{y_{smarg(ijk)n} : 1 \leq n \leq N\}$;
- Include the means, minimum and maximum together with the values of $rq_{smarg(ijk)}$ against the expected percentiles of the standard normal distribution.

The minimum and maximum values of $rq_{smarg(ijk)}$ form the envelope. If the model under study is correct, the observed values should be inside the bands and distributed randomly.

5.5 Application: Lychees data

The species Litchi chinensis Sonn, a member of the Sapindaceae family, is native to Southern China and is considered a national fruit. The first lychee trees were introduced in Brazil in about 1810, planted in the Botanical Gardens in Rio de Janeiro. The fruit is still relatively unknown to Brazilian consumers and the potential market is substantial due to the quality of the fruit and the typical ripening during the year-end holiday season. Due to the good acceptance of the lychee in many countries, a potential market exists for exports from Brazil, even in producing countries because of the seasonal difference in peak harvesting periods between the northern and southern hemispheres. In particular, the fruit has gained wide acceptance in United States and Europe. The lychee is a drupe with sweet and succulent translucent flesh (aril). Its pericarp is a very attractive red color. However, once harvested when maintained under environmental conditions, its quality deteriorates in only two days. This short post-harvest life greatly limits the market for this fruit.

In Brazil, although a good deal of interest in the fruit exists due to the country's ideal growing conditions and large potential market, little is known about the metabolic and physiological processes for determination of post-harvest conservation techniques. Therefore, the present study was conducted in the Physiology and Biochemistry Laboratory of the Department of Biological Sciences (ESALQ) of the University of São Paulo for the purpose of determining the physical, physiological, biochemical and sensorial responses associated with the use of post-harvest conservation of lychee, aiming to reduce the loss of quality.

Next, we describe the experiments conducted in this study, in which we analyzed the data using the OLLSN regression model with random effects, as discussed in the previous sections. This experiment aimed to evaluate the use of antioxidants in the lychees "Bengal". The experiment adopted the completely randomized design with the factors treatments and time with four replicates for treatments. The fruits were placed in expanded polystyrene trays, wrapped in 14μm thick polyvinyl chloride (PVC) film and were evaluated every 3 days for 15 days of storage at 5°C and 90% RH.

- Let Y_{ijk} represent the ascorbic acid content of the i-th individual ($i = 1, \ldots, 5$), in the j-th ($j = 1, \ldots, 6$) repeated measure in time and in the k-th ($k = 1, \ldots, 4$) group, making a total of 120 sample units;

- Treatments:

- Treat1 - without treatment (distilled water);
- Treat2 - immersion in 4-hexylresorcinol (300 mg L^{-1}) solution;
- Treat3 - immersion in ascorbic acid (300 mg L^{-1}) solution;
- Treat4 - immersion in citric acid (300 mg L^{-1}) solution;
- Treat5 - immersion in de citric acid + ascorbic acid (300 mg L^{-1}) solution;

- T: longitudinal factor (t): (0, 3, 6, 9, 12 and 15 days).

First we present a descriptive analysis of the response variable data. Figure 5.2 contains the frequency histograms for the ascorbic acid concentration for each treatment, indicating evidence that the data have bimodal distribution. In turn, Table (5.1) summarizes the main descriptive statistics for ascorbic acid concentration and for Treat 4.

Tabela 5.1 - Descriptive statistics.

Data	Mean	Median	Mode 1	Mode 2	Variance	Skewness	Kurtosis	Min.	Max.
Y	30.84	31.19	26.54[a]	32.81[a]	24.92	-0.29	-0.65	18.83	39.86
Treat 4	29.01	28.34	25.00[a]	32.66[a]	30.00	0.11	-1.19	19.85	38.94

[a] There are several modes

Table 5.1 gives a descriptive summary of the Y and Treat 4. The Y data has negative skewness and kurtosis, has the lowest variability an the presence of bimodality in the data.

Figure 5.3 shows the graphs of the mean profiles and the treatment x time interaction. Figure 5.3 (a) presents the mean effect of each time interval (days) in the treatments, while Figure 5.3 (b) presents the behavior of each treatment on each day assessed. It can be seen in Figure 5.3 (a) that a linear relationship exists between the ascorbic acid content (Y) and time intervals. In turn, Figure 5.3 (b) indicates that the response in the treatments was different for each time interval, mainly after 9 and 12 days.

The main objectives of this study were:

- to indicate the best treatment for each time interval; and
- to indicate as of what moment the treatments no longer have effects on the fruits.

To attain these objectives, we divided the data analysis into three parts, as described next.

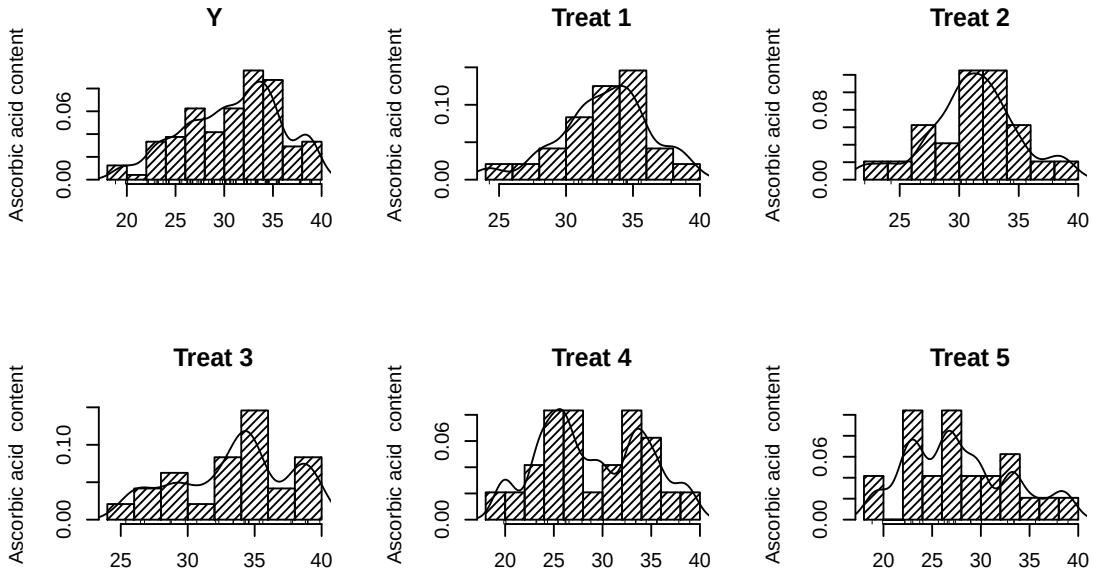


Figura 5.2 - Frequency histogram of the effect of ascorbic acid concentration in function of treatments: Treat 1, Treat 2, Treat 3, Treat 4 and Treat 5.

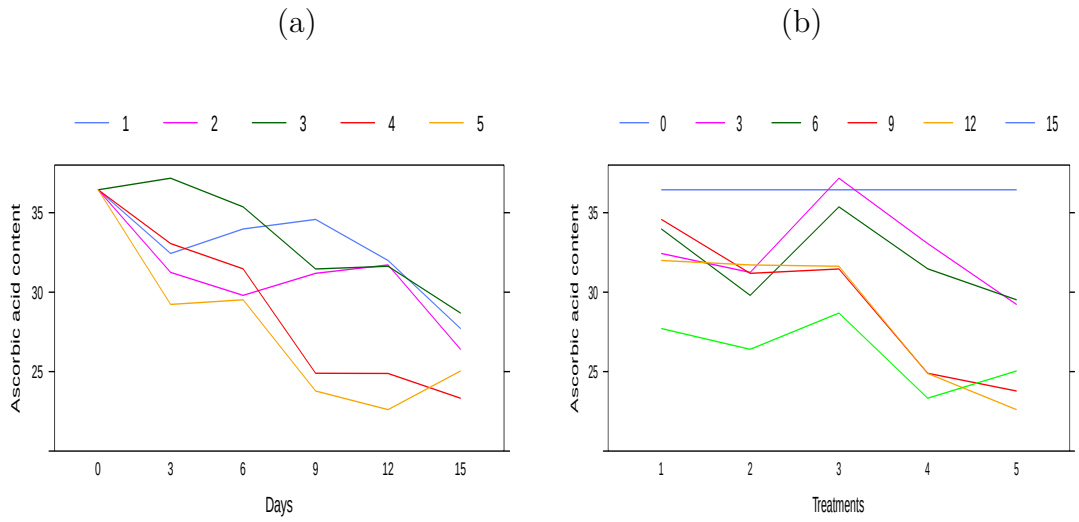


Figura 5.3 - Analysis of profile and cross-effects of the ascorbic acid concentration in function of the time and treatment factors: (a) Represents the average effect of each day in a treatment. (b) Represents the effect of each treatment on a day.

5.5.1 Analysis 1: OLLSN regression model for the interaction effect

To study the effect of the (treatment $\times$ time) interaction, we considered the OLLSN regression model with random effect described in equation (2), where the vector of parameters

is $\boldsymbol{\theta} = (\beta_0, \beta_1, \boldsymbol{\tau}^T, \boldsymbol{\gamma}^T, \sigma^2, \sigma^2_{\beta_0}, \lambda, \alpha)^T$ with $\boldsymbol{\tau} = (\tau_2, \ldots, \tau_5)^T$ and $\boldsymbol{\gamma} = (\gamma_2, \ldots, \gamma_5)^T$. Note that in this case, the vector of covariables $\mathbf{x}_{ijk}$ is composed by the dummy covariables d_i and t, so the OLLSN regression model with random effect (5.7), assumes the following structure:

$$y_{ijk} = \beta_0 + \underbrace{\beta_1 t}_{longitudinal} + \underbrace{\sum_{i=2}^{5} \tau_i d_i}_{treatments} + \underbrace{t \sum_{i=2}^{5} \gamma_i d_i}_{interaction} + w_i + \sigma z_{ijk}, \tag{5.21}$$

where β_0 is the intercept, β_1 is the rate of variation of the ascorbic acid concentration in function of the time covariable t; d_2, d_3, d_4, d_5 are dummy variables for each treatment level ($\boldsymbol{\tau}$), represented by τ_2, τ_3, τ_4, and τ_5 the coefficients $\gamma_2, \gamma_3, \gamma_4$, and γ_5 denote the effects of the (treatment x time) interaction; w_i is the random effect for individual i; and z_{ijk} is a random variable with OLLSN distribution. Each dummy variable d_i represents a treatment in the following form:

$$d_i = \begin{cases} 0, \text{ if treatment for } \neq i \\ 1, \text{ if treatment for } = i. \end{cases}$$

Table 5.2 contains the MLE values for model (5.21) with their standard errors and significance levels, respectively. We highlight the estimated values of $\hat{\gamma}_4 = -0.4938$ (p-value=0.0161) e $\hat{\gamma}_5 = -0.3564$ (p-value=0.0372), which present have significant descriptive values. Therefore, dependence exists between the time and treatment factors.

Tabela 5.2 - MLEs to test the effects of the interaction of the longitudinal factor *versus* treatments (use of antioxidants on Bengal lychee fruits).

Parameter	Estimates	SE	p-value	Parameter	Estimates	SE	p-value
β_0	36.2158	1.3727	0.0000	γ_3	-0.1323	0.1822	0.4692
β_1	-0.4300	0.1383	0.0023	γ_4	-0.4938	0.2021	0.0161
τ_2	-1.4848	1.8374	0.4207	γ_5	-0.3564	0.1690	0.0372
τ_3	1.5836	1.8202	0.3861	σ_{β_0}	$7.33\text{x}10^{-09}$	$8.56\text{x}10^{-05}$	
τ_4	-0.1740	1.9808	0.9301	σ	$9.99\text{x}10^{-01}$	$9.99\text{x}10^{-01}$	
τ_5	-2.5070	1.6328	0.1275	λ	2.0583	0.1590	< 0.0001
γ_2	-0.0349	0.1808	0.8472	α	1.1734	0.0882	0.0724

5.5.2 Analysis 2: the hierarchical OLLSN regression model

To study the behavior of the ascorbic acid level considering the breakdown of a treatment by day, the OLLSN regression model with random effect (5.7) presents the following

hierarchical structure:

$$y_{ijk} = \beta_0 + \underbrace{\sum_{j=2}^{6} c_j u_j}_{\text{longitudinal effects}} + \underbrace{\sum_{j=1}^{6}\sum_{k=2}^{5} \delta_{jk} u_j\, d_i}_{\text{hierarchical effects}} + w_i + \sigma^2 z_{ijk}, \tag{5.22}$$

in which the time covariable t was categorized by the *dummies* variables $u_2, \ldots, u_6$, represented by the coefficients, $c_2, \ldots, c_6$, while the coefficients $\delta_{j2}, \delta_{j3}, \delta_{j4}$, and δ_{j5} represent the breakdown of the effects of treatments at each time (days). Each dummy variable u_j represents the treatment as follows:

$$u_j = \begin{cases} 0, \text{ if } t \text{ for } \neq \text{j} \\ 1, \text{ if , } t \text{ for } = \text{j}. \end{cases}$$

Tabela 5.3 - MLEs to test the effects of the treatments (use of antioxidants in Bengal lychees) compared two-by-two.

Days	Hypothese	Estimate	SE	p-value	Days	Hypothese	Estimate	SE	p-value
0	$H_0 : \tau_2 - \tau_1 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_2 - \tau_1 = 0$	-3.571^{ns}	4.078	0.383
0	$H_0 : \tau_3 - \tau_1 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_3 - \tau_1 = 0$	-3.222^{ns}	4.240	0.449
0	$H_0 : \tau_4 - \tau_1 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_4 - \tau_1 = 0$	-9.730^{*}	4.825	0.046
0	$H_0 : \tau_5 - \tau_1 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_5 - \tau_1 = 0$	-10.985^{*}	4.085	0.008
0	$H_0 : \tau_3 - \tau_2 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_3 - \tau_2 = 0$	0.349^{*}	2.151	0.871
0	$H_0 : \tau_4 - \tau_2 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_4 - \tau_2 = 0$	-6.158^{ns}	3.151	0.053
0	$H_0 : \tau_5 - \tau_2 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_5 - \tau_2 = 0$	-7.413^{*}	1.826	0.000
0	$H_0 : \tau_4 - \tau_3 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_4 - \tau_3 = 0$	-6.507^{ns}	3.358	0.055
0	$H_0 : \tau_5 - \tau_3 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_5 - \tau_3 = 0$	-7.762^{*}	2.163	0.000
0	$H_0 : \tau_5 - \tau_4 = 0$	0.000^{ns}	2.274	1.000	09	$H_0 : \tau_5 - \tau_4 = 0$	-1.255^{ns}	3.159	0.692
3	$H_0 : \tau_2 - \tau_1 = 0$	-1.184^{ns}	4.359	0.786	12	$H_0 : \tau_2 - \tau_1 = 0$	-0.329^{ns}	2.102	0.875
3	$H_0 : \tau_3 - \tau_1 = 0$	4.515^{ns}	3.016	0.137	12	$H_0 : \tau_3 - \tau_1 = 0$	-0.534^{ns}	1.852	0.773
3	$H_0 : \tau_4 - \tau_1 = 0$	0.453^{ns}	3.078	0.883	12	$H_0 : \tau_4 - \tau_1 = 0$	-7.243^{*}	1.938	0.000
3	$H_0 : \tau_5 - \tau_1 = 0$	-3.458^{ns}	2.981	0.249	12	$H_0 : \tau_5 - \tau_1 = 0$	-9.443^{*}	1.966	0.000
3	$H_0 : \tau_3 - \tau_2 = 0$	5.700^{ns}	3.518	0.108	12	$H_0 : \tau_3 - \tau_2 = 0$	-0.205^{ns}	1.674	0.902
3	$H_0 : \tau_4 - \tau_2 = 0$	1.638^{ns}	3.571	0.647	12	$H_0 : \tau_4 - \tau_2 = 0$	-6.914^{*}	1.769	0.000
3	$H_0 : \tau_5 - \tau_2 = 0$	-2.273^{ns}	3.488	0.516	12	$H_0 : \tau_5 - \tau_2 = 0$	-9.114^{*}	1.800	0.000
3	$H_0 : \tau_4 - \tau_3 = 0$	-4.061^{*}	1.688	0.018	12	$H_0 : \tau_4 - \tau_3 = 0$	-6.708^{*}	1.464	0.000
3	$H_0 : \tau_5 - \tau_3 = 0$	-7.973^{ns}	1.505	0.000	12	$H_0 : \tau_5 - \tau_3 = 0$	-8.909^{*}	1.500	0.000
3	$H_0 : \tau_5 - \tau_4 = 0$	-3.912^{*}	1.626	0.018	12	$H_0 : \tau_5 - \tau_4 = 0$	-2.200^{ns}	1.605	0.174
6	$H_0 : \tau_2 - \tau_1 = 0$	-4.168^{ns}	2.300	0.073	15	$H_0 : \tau_2 - \tau_1 = 0$	-1.429^{ns}	1.797	0.428
6	$H_0 : \tau_3 - \tau_1 = 0$	1.449^{ns}	2.184	0.508	15	$H_0 : \tau_3 - \tau_1 = 0$	0.913^{ns}	1.823	0.617
6	$H_0 : \tau_4 - \tau_1 = 0$	-2.433^{ns}	2.838	0.393	15	$H_0 : \tau_4 - \tau_1 = 0$	-4.387^{*}	2.096	0.039
6	$H_0 : \tau_5 - \tau_1 = 0$	-4.435^{*}	2.010	0.029	15	$H_0 : \tau_5 - \tau_1 = 0$	-2.612^{ns}	2.428	0.285
6	$H_0 : \tau_3 - \tau_2 = 0$	5.617^{*}	2.496	0.026	15	$H_0 : \tau_3 - \tau_2 = 0$	2.342^{ns}	1.452	0.110
6	$H_0 : \tau_4 - \tau_2 = 0$	1.735^{ns}	3.085	0.575	15	$H_0 : \tau_4 - \tau_2 = 0$	-2.958^{ns}	1.783	0.100
6	$H_0 : \tau_5 - \tau_2 = 0$	-0.267^{ns}	2.345	0.909	15	$H_0 : \tau_5 - \tau_2 = 0$	-1.183^{ns}	2.164	0.585
6	$H_0 : \tau_4 - \tau_3 = 0$	-3.882^{ns}	2.999	0.198	15	$H_0 : \tau_4 - \tau_3 = 0$	-5.300^{*}	1.810	0.004
6	$H_0 : \tau_5 - \tau_3 = 0$	-5.885^{*}	2.231	0.009	15	$H_0 : \tau_5 - \tau_3 = 0$	-3.525^{ns}	2.186	0.110
6	$H_0 : \tau_5 - \tau_4 = 0$	-2.002^{ns}	2.875	0.487	15	$H_0 : \tau_{55} - \tau_4 = 0$	1.775^{ns}	2.418	0.464

Table 5.3 presents the MLEs for the contrasts between the treatments (Treat1, Treat2,

Treat3, Treat4 and Treat5) taken two-by-two for the model (5.22). The hypotheses were formulated to compare the treatments where one of them serves as a benchmark within each of the time intervals (0, 3, 6, 9, 12, 15 days). For example, hypothesis $H_0 : \tau_5 - \tau_3 = 0$ on day 6 indicates a significant difference exists (0.0098) between Treat 5 and Treat 3.

5.5.3 Analysis 3: the linear and quadratic OLLSN regression models

To study the linear and quadratic effects between the response variable ascorbic acid concentration (y_{ijk}) and the longitudinal factor t, the OLLSN regression model with random effect described in equation (5.7) assumes the following structure:

$$y_{ijk} = \begin{cases} \overbrace{\beta_0}^{intercept\ effect} + \overbrace{\beta_1 t}^{linear\ effect} + \mathrm{w}_i + \sigma z_{ijk} & \text{(a)} \\ \underbrace{\beta_0}_{intercept\ effect} + \beta_1 t + \underbrace{\beta_2 t^2}_{quadratic\ effect} + \mathrm{w}_i + \sigma z_{ijk} & \text{(b)} \end{cases} \tag{5.23}$$

Table 5.4 presents the MLEs of the parameters for location (β_0, β_1), scale (σ^2, $\sigma^2_{\beta_0}$) and shape (λ, α). The Figure 5.4 (a) represents the linear effects of the ascorbic acid concentration in function of time intervals, indicating the ascorbic acid level declines every three days by an average of -0.6307 for model 5.23 (a).

Tabela 5.4 - MLEs for the linear model 16(a) for the data on use of antioxidants on Bengal lychees.

Parameter	Estimates	Parameter	Estimates	Linear predictor
β_0	-0.6307	β_1	35.4324	1.8024
β_0	-0.6307	β_1	35.4324	0.5119
β_0	-0.6307	β_1	35.4324	2.6096
β_0	-0.6307	β_1	35.4324	-1.9408
β_0	-0.6307	β_1	35.4324	-2.9831
Parameter	Estimates	SE	z-valor	p-value
σ^2	1.0400	1.0198	-	-
$\sigma^2_{\beta_0}$	4.9741	2.2302	-	-
λ	0.0910	1.3632	0.0670	0.9470
α	0.3941	0.3665	2.5400	0.0125

Table 5.5 presents the MLEs of the parameters for location (β_0, β_1, β_2), scale (σ^2, $\sigma^2_{\beta_0}$) and shape (λ, α) for model 5.23 (b). The estimate $\hat{\beta}_2$ (0.0133) indicates that the quadratic effects of the ascorbic acid concentration in function of time intervals practically null.

Tabela 5.5 - MLEs for the linear model 5.23 (a) for the data on use of antioxidants on Bengal lychees.

Parameter	Estimates	Parameter	Estimates	Parameter	Estimates	Linear predictor
β_0	6.0434	β_1	-0.8303	β_2	0.0133	1.8611
β_0	6.0434	β_1	-0.8303	β_2	0.0133	0.4891
β_0	6.0434	β_1	-0.8303	β_2	0.0133	2.5727
β_0	6.0434	β_1	-0.8303	β_2	0.0133	-1.9120
β_0	6.0434	β_1	-0.8303	β_2	0.0133	-3.0110
Parameter	Estimates	SE	z-valor	p-value		
σ^2	1.040193	1.0198	-	-		
$\sigma^2_{\beta_0}$	4.9692	2.2291	-	-		
λ	0.0981	1.3045	0.075	0.94		
α	0.4118	0.3546	2.502	0.0138		

Table 5.6 presents the MLEs of the parameters for location (β_0, β_1, β_2), scale (σ^2, $\sigma^2_{\beta_0}$) and shape (λ, α) for the quadratic effects of the response to ascorbic acid concentration in function of time intervals for model 5.23 (b) for each of the treatments Treat 1,..., Treat 5. The Figures 5.4 (a) and 5.4 (b) show good prediction performance of the OLLSN regression model with random effect.

Tabela 5.6 - MLEs for the quadratic model 5.23 (b) for each treatment (use of antioxidants in Bengal lychees).

Treatments	Parameter	Estimates	SE	p-value
Treat 1	β_0	35.0877	1.0573	< 0.001
	β_1	0.0417	0.3640	0.910
	β_2	-0.0336	0.0223	0.148
Treat 2	β_0	35.1467	1.6868	< 0.001
	β_1	-0.3408	0.4496	0.457
	β_2	-0.0134	0.0255	0.604
Treat 3	β_0	35.5451	1.7194	< 0.001
	β_1	-0.2459	0.4488	0.590
	β_2	-0.0186	0.0267	0.492
Treat 4	β_0	36.0728	1.3524	< 0.001
	β_1	-1.3076	0.3723	< 0.001
	β_2	0.0260	0.0217	0.244
Treat 5	β_0	34.7695	1.4921	< 0.001
	β_1	-2.0608	0.4600	< 0.001
	β_2	0.0845	0.0297	< 0.001

(a) (b)

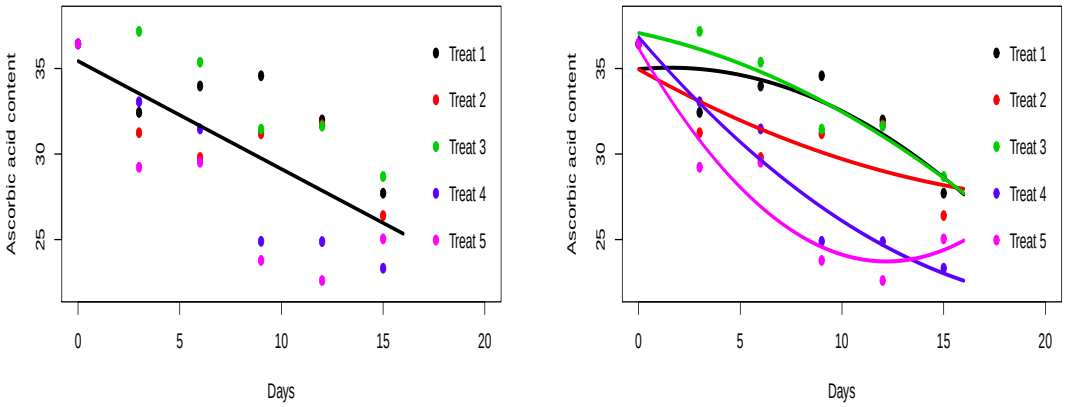


Figura 5.4 - Plots for model (5.23): (a) Represents the linear effects of the ascorbic acid concentration in function of times. (b) Represents the quadratic effect of the ascorbic acid concentration in function of the times.

5.5.4 Residual analysis

To detect the presence of discrepant observations as well as violation of the assumption of independence of the errors and homogeneity of the variances of the OLLSN model with random intercept, we analyzed the standardized marginal quantile residuals $\hat{rq}_{smarg(ijk)}$, the conditional quantile residuals $\hat{rq}_{sc(ijk)}$, then the simulated envelope for the marginal quantile residuals are built, as shown in Figure 5.5.

Figure 5.5(a) shows the marginal quantile residuals $\hat{rq}_{smarg(ijk)}$ in function of the days, indicating that the linear model is suitable to explain the relationship between the ascorbic acid concentration and time intervals (days). Figure 5.5(b) shows the conditional quantile residuals$\hat{rq}_{sc(ijk)}$ in function of the estimated values ($\hat{y}$), indicating that the adjusted OLLSN model with random intercept did not violate the assumption of homogeneity of variance.

Finally, we verify the quality of the adjustment range of the OLLSN regression model with effect random by constructing in Figure 5.5 (c) the normal probability plot for the component of the waste diversion with simulated envelope. This figure reveals that there is evidence of a good fit of the OLLSN regression model with random effect to the current data.

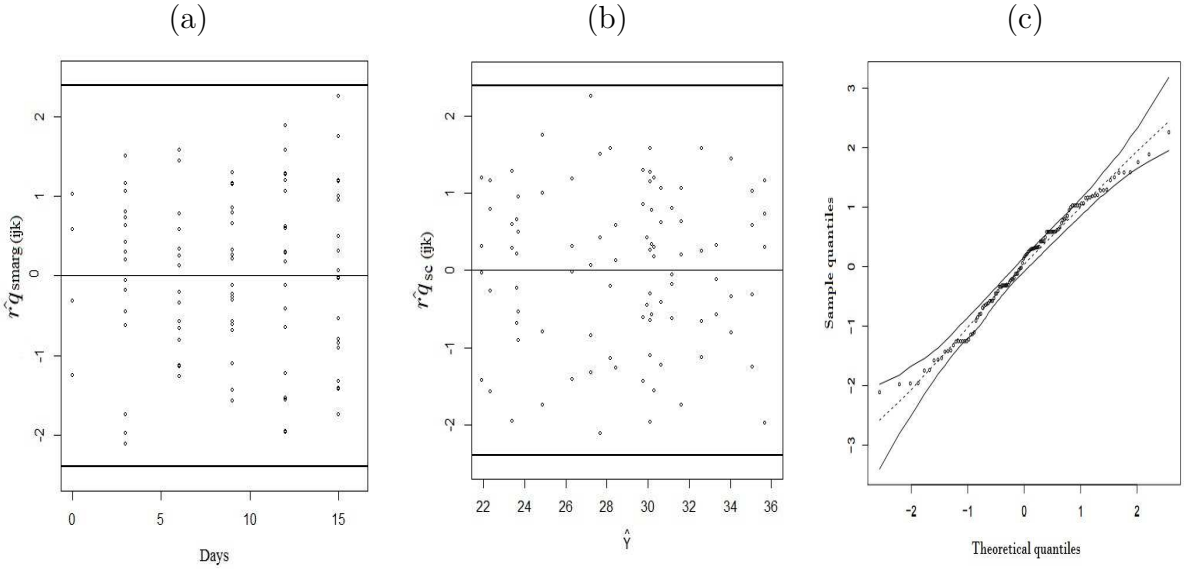


Figura 5.5 - Index plot of residuals on fitting the OLLSN regression model with random effect for for the use of antioxidants in the lychees "Bengal" data: (a) Marginal quantile residuals $\hat{rq}_{om(ijk)}$ versus days; (b) Conditional quantile residuals $versus\ \hat{Y}$; (c) Normal probability plot for the marginal quantile residuals with envelopes.

5.6 Concluding remarks

In this article we proposed a new odd log-logistic skew-normal (OLLSN) regression model with random effect. The proposed distribution is more versatile than the normal and skew-normal distributions, since it can be adjusted to bimodal data. The model parameters are estimated by the method of maximum likelihood. This extended regression model is very flexible and can be used in many practical situations. We performed an application non longitudinal data regarding the use of antioxidants on Bengal lychees, showing the flexibility of the new model in relation to the normal and skew-normal distributions. Besides this, we demonstrated that the model was well adjusted by analysis of residuals. Therefore, the new distribution proposed can be recommended to analyze data with repeated measures over time, especially when there are indications of asymmetry and bimodality.

References

ATKINSON, A.C. (1985). **Plots, transformations, and regression**. Oxford: University Press, 282p.

AZAALINI, A. (1985). A class of distributions which includes the normal ones. **Scandinavian**

Journal of Statistics, **12**, 171-178.

AZAALINI, A.; CAPITANIO, A. (1999). Statistical application of the multivariate skew normal distribution. **Journal of the Royal Statistical Society**, Padova, v.61, n. 3, p. 579-602.

AZAALINI, A., Del Cappello, T., & Kotz, S. (2002). Log-skew-normal and log-skew-t distributions as models for family income data. **Journal of Income Distribution**, **11**(3-4), 2-2.

BONAT, W. H., KRAINSKI, E. T., RIBEIRO JR, P. J., ZEVIANI, W. M. (2012). **Métodos computacionais em inferência estatística**. Piracicaba: Região Brasileira da Sociedade Internacional de Biometria. 289p.

BOX, G.; COX, D. R. (1964). An analysis of transformations. **Journal of the Royal Society**, **26**, 211-252.

BRAGA, A. S.; CORDEIRO, G. M.; ORTEGA, E. M. M.; da CRUZ, J. N. (2016). The odd log-logistic normal distribution: Theory and applications in analysis os experiments. **Journal of Statistical Theory and Practice**, **10**, 311-335.

CROWDER, M. J., HAND, D. J. (1990). **Analysis of repeated measures**. [S.l.]: CRC Press, 1990. v. 41.

DIGGLE, P. J. (1988). An approach to the analysis of repeated measurements. **Biometrics**, JSTOR, p. 959-971.

FABIO, L. Z., PAULA, G. A., de CASTRO, M. A. (2012). A Poisson mixed model with nonnormal random effect distribution. **Computational Statistics and Data Analysis**, Amisterdam, v.56, p. 1499-1510.

FERREIRA, W. L., de MORAIS, A. R. (2013). Análise da Influência do Cáfé no Ganho de Peso de Animais (Ratos) por meio de modelo linear misto **Rev. Bras. Biom.**, **31**(4), 485-500.

GLEATON, J. U.; LYNCH, J. D. (2006). Properties of generalized log-logistic families of lifetime distributions. **Journal of Probability and Statistical Science**, **4**, 51-64.

HASHIMOTO, E. M. (2013). **Modelo de regressão gama-G em análise de sobrevivência**. Tese (Doutorado) - Escola Superior de Agricultura Luiz de Queiroz.

LESAFFRE, E.; SPIESSENS, B. (2001). On the effect of the number of quadrature points in a logistic random effects model: an example. **Journal of the Royal Statistical Society**: Series C (Applied Statistics), Wiley Online Library, v. 50, n. 3, p. 325-35.

LITTEL, R. C; HENRY, P. R.; AMMERMAN, C.B. (1998) Statistical analysis of repeated measures data using sas procedures. **Journal of animal science**, v. 76, n. 4, p. 1216-1231.

NOBRE, J.; SINGER, J. Residual analysis for linear mixed models. (2007). **Biometrical Journal**, Hoboken, v. 49, p. 863-875.

PINHEIRO, J. C., BATES, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. **Journal of Computational and Graphical Statistics**, Alexandria, v. 4, n. 1, p. 12-35.

SAHU, S. K.; DEY, D. K. & BRANCO, M. D. (2003). A new class of multivariate skew distributions with applications to bayesian regression models. **The Canadian Journal of Statistic**, **31**, 129-150.

SAHU, S. K.; DEY, D. K. (2004) On multivariate survival models with a skewed frailty and a correlated baseline hazard process. Skew-elliptical distributions and their applications: **a journey beyond normality**, 321-338.

VAIDA, FLORIN; X.U., RONGHUI. (2000). Proportional hazards model with random effects. **Statistics in medicine**, Citeseer, v. 19, n. 24, p. 3309-3324.

VERBEKE, G., & LESAFFRE, E. (1997). The effect of misspecifying the random-effects distribution in linear mixed models for longitudinal data. **Computational Statistics & Data Analysis**, **23**(4), 541-556.

VILLEGAS, C.; PAULA, G. A.; LEIVA, V. (2011). Birnbaum-saunders mixed models for censored reliability data analysis. **IEEE Transactions on Reliability**, v. 60, n. 4, p. 748-758.

WEST, B.; WELCH, K.B.; GALECKI, A.T. **Linear mixed models: a pratical guide using statistical software**. New York: Chapman & Hall, 2007. 376 p.

APPENDICES

Appendix A - Demonstration t-Student Distribution

In this demonstration we consider the random error variable defined as $t = (x - \mu)/\sigma$. Thus, the cumulative distribution function (cdf) of the t-Student distribution is given by,

$$F(t) = \frac{1}{2} + \frac{1}{\sqrt{\nu} B\left(\frac{1}{2}; \frac{\nu}{2}\right)} \underbrace{\int_{-\infty}^{T} \left[1 + \frac{s^2}{\nu}\right]^{-\frac{(\nu+1)}{2}} ds}_{M}. \tag{5.24}$$

Resolving the integral M,

$$\mathrm{M} = \int_{-\infty}^{T} \left[1 + \frac{s^2}{\nu}\right]^{-\frac{(\nu+1)}{2}} ds = \int_{-\infty}^{T} \left[\frac{\nu}{\nu + s^2}\right]^{\frac{(\nu+1)}{2}} ds.$$

Setting $t = \nu/(\nu + s^2)$ and making the following changes to $s \to 0 \Rightarrow t \to 1$ and to $s \to \nu/(\nu + s^2) \Rightarrow t \to \nu/(\nu + t^2)$ and still that $t = \frac{\nu}{\nu+s^2} \Rightarrow s^2 + \nu = \frac{\nu}{t} \Rightarrow s = \pm\sqrt{\frac{\nu}{t} - \nu} \Rightarrow s = \pm\sqrt{\frac{\nu(1-t)}{t}} \Rightarrow ds = \pm\frac{\left(\nu+s^2\right)^2}{2\nu s}\, dt$ and considering $s = \pm t^{-\frac{1}{2}} \nu^{\frac{1}{2}} (1-t)^{\frac{1}{2}}$. Thus, to $s = -t^{-\frac{1}{2}} \nu^{\frac{1}{2}} (1-t)^{\frac{1}{2}}$ we have that,

$$\begin{cases} \nu + s^2 = \frac{\nu}{t} \Rightarrow \left(\nu + s^2\right)^2 = \frac{\nu^2}{t^2} \\ -2\nu s = -2\nu \left[-\sqrt{\frac{\nu(1-t)}{t}}\right] = 2\frac{\nu^{\frac{3}{2}} (1-t)^{\frac{1}{2}}}{t^{\frac{1}{2}}} \end{cases}$$

The ratio of the two last equation and considering the following expression $\nu^{\frac{1}{2}} t^{-\frac{3}{2}} (1-t)^{-\frac{1}{2}} \big/ 2$. We have that the function M is given by,

$$\mathrm{M} = \int_{1}^{\frac{\nu}{\nu+t^2}} t^{\frac{(\nu+1)}{2}} \frac{\left(\nu + s^2\right)^2}{-2\nu s} dt = \frac{\nu^{\frac{1}{2}}}{2} \int_{1}^{\frac{\nu}{\nu+t^2}} t^{\frac{\nu}{2}-1} (1-t)^{\frac{1}{2}-1} dt. \tag{5.25}$$

Inserting (5.25) in (5.24), we obtain,

$$\begin{aligned} F(t) &= \frac{1}{2} + \frac{1}{\sqrt{\nu} B\left(\frac{1}{2}; \frac{\nu}{2}\right)} \frac{\nu^{\frac{1}{2}}}{2} \int_{1}^{\frac{\nu}{\nu+t^2}} y^{\frac{\nu}{2}-1} (1-y)^{\frac{1}{2}-1} dy \\ F(t) &= \frac{1}{2} - \frac{1}{2B\left(\frac{1}{2}; \frac{\nu}{2}\right)} \int_{\frac{\nu}{\nu+t^2}}^{1} y^{\frac{\nu}{2}-1} (1-y)^{\frac{1}{2}-1} dy \\ F(t) &= \frac{1}{2} - \frac{1}{2}\left[1 - \frac{1}{B\left(\frac{1}{2}; \frac{\nu}{2}\right)} \int_{0}^{\frac{\nu}{\nu+t^2}} y^{\frac{\nu}{2}-1} (1-y)^{\frac{1}{2}-1} dy\right] \end{aligned}$$

$$
\begin{aligned}
F(t) &= \frac{1}{2} - \frac{1}{2}\left[\frac{1}{B\left(\frac{1}{2};\frac{\nu}{2}\right)}\int_0^1 y^{\frac{\nu}{2}-1}(1-y)^{\frac{1}{2}-1}\,dy\right. \\
&\quad \left. - \frac{1}{B\left(\frac{1}{2};\frac{\nu}{2}\right)}\int_0^{\frac{\nu}{\nu+t^2}} y^{\frac{\nu}{2}-1}(1-y)^{\frac{1}{2}-1}\,dy\right] \\
F(t) &= \frac{1}{2} - \frac{1}{2}\left[I\left(1;\frac{\nu}{2};\frac{1}{2}\right) - I\left(\frac{\nu}{\nu+t^2};\frac{\nu}{2};\frac{1}{2}\right)\right] \blacksquare.
\end{aligned}
$$

The same way using $s = t^{\frac{1}{2}}\nu^{\frac{1}{2}}(1-t)^{\frac{1}{2}}$ we can arrive in the result of the expression (5.24). Setting t $= (x-\mu)/\sigma$, ie,

$$
G(x;\mu,\sigma,\nu) = \frac{1}{2} + \frac{1}{2}\left\{I\left[1,\frac{\nu}{2},\frac{1}{2}\right] - I\left[\frac{\nu}{\nu+\left(\frac{x-\mu}{\sigma}\right)^2},\frac{\nu}{2},\frac{1}{2}\right]\right\}\operatorname{sgn}\left(\frac{x-\mu}{\sigma}\right),
$$

where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, ν is the number of degrees of freedom, $I(x;a,b)$ is the regularized beta function defined by $I(x;a,b) = B(x;a,b)/B(a,b)$ with $B(x;a,b)$ the incomplete beta function and $B(a,b)$ is the beta function and $\phi_{t_\nu}(\cdot)$ and $\Phi_{T_\nu}(\cdot)$ are the pdf and cdf of the standard Student's t distribution, respectively, and $\operatorname{sgn}(\cdot)$ is the sign function.

Appendix B - Mathematical properties

Power series methods are at the heart of many aspects of applied mathematics and statistics. We obtain some mathematical properties of the new distribution using a power series expansion for the qf of the Student-t distribution, say $Q_\nu(u) = \Phi_\nu^{-1}(u)$.

Moments

The function $Q_\nu(u)$ does not have a closed-form expression, but it can be written as a power series given by

$$
Q_\nu(u) = \sum_{k=0}^{\infty} a_k\, v^k, \tag{5.26}
$$

where $v = g_\nu\,(u-1/2)$, $g_\nu = \sqrt{\nu\pi}\frac{\Gamma(\nu)}{\Gamma(\frac{\nu+1}{2})}$ and the coefficients a_k's are defined from the following b_k's by: $a_k = 0$ for $k = 0,2,4,\ldots$ and $a_k = b_{(k-1)/2}$ for $k = 1,3,5,\ldots$, the quantities b_k (for $k \geq 1$) are given by cubic recurrence equation (Shaw, 2006)

$$
\begin{aligned}
b_k &= \frac{1}{2k(2k+1)}\sum_{r=0}^{k-1}\sum_{s=0}^{k-r-1} b_r\, b_s\, b_{k-r-s-1} \\
&\times \left\{(1+\nu^{-1})[(2s+1)(2k-2r-2s-1)] - 2r(2r+1)\nu^{-1}\right\}.
\end{aligned} \tag{5.27}
$$

and $b_0 = 1$. It is easily obtained the first coefficients as: $b_1 = (\nu+1)/(6\nu)$, $b_2 = (7\nu^2+8\nu+1)/(120\nu^2)$, $b_3 = (127\nu^3+135\nu^2+9\nu+1)/(5040\nu^3)\ldots$

We can rewrite (5.26) as

$$Q_\nu(u) = \sum_{k=0}^{\infty} a_k\, g_\nu^k\, (u-1/2)^k = \sum_{k=0}^{\infty} a_k\, g_\nu^k \sum_{i=0}^{k} \binom{k}{i} (-1/2)^{k-i}\, u^i,$$

and then by changing the sums $\sum_{k=0}^{\infty}\sum_{i=0}^{k}$ by $\sum_{i=0}^{\infty}\sum_{k=i}^{\infty}$, we have

$$Q_\nu(u) = \sum_{i=0}^{\infty} p_i\, u^i,$$

where $p_i = \sum_{k=i}^{\infty} a_k\, g_\nu^k \binom{k}{i} (-1/2)^{k-i}$ for $i \geq 0$.

Let V_{k+1} be a random variable having pdf $h_{k+1}(t) = (k+1)\,\sigma^{-1}\phi_\nu\left(\frac{t-\mu}{\sigma}\right)\Phi_\nu\left(\frac{t-\mu}{\sigma}\right)^k$ defined above.

Setting $(t-\mu)/\sigma = x$, we can determine the nth moment of V_{k+1} in terms of the qf of the Student's t distribution

$$\tau_{n,k+1} = E(V_{k+1}^n) = \sum_{j=0}^{\infty} \binom{n}{j} \mu^{n-j}\, \sigma^j\, I_\nu(j,k),$$

where $I_\nu(j,k) = \int_0^1 Q_\nu(u)^j\, u^k du$.

By application of an equation in Section 0.314 of (GRADSHTEYN; RYZHIK, 2000) for a power series raised to a positive integer power j

$$Q_\nu(u)^j = \left(\sum_{i=0}^{\infty} p_i\, u^i\right)^j = \sum_{i=0}^{\infty} c_{j,i}\, u^i. \tag{5.28}$$

Here, the coefficients $c_{n,i}$ (for $i = 1, 2, \ldots$) are easily obtained from the recurrence equation

$$c_{j,i} = (i\, p_0)^{-1} \sum_{m=1}^{i} [m(j+1) - i]\, p_m\, c_{j,i-m},$$

where $c_{j,0} = p_0^j$. The coefficient $c_{j,i}$ can be determined from $c_{j,0}, \ldots, c_{j,i-1}$ for programming numerically in any algebraic or numerical software.

We now obtain $I_\nu(j,k)$ by using (5.28)

$$I_\nu(j,k) = \int_0^1 \left(\sum_{i=0}^{\infty} c_{j,i}\, u^i\right) u^k du = \sum_{i=0}^{\infty} \frac{c_{j,i}}{i+k+1}$$

and then

$$\tau_{n,k+1} = \sum_{i,j=0}^{\infty} \binom{n}{j} \frac{\mu^{n-j}\,\sigma^{j}\,c_{j,i}}{i+k+1}. \tag{5.29}$$

By combining (3.14) and (5.29), we can determine the ordinary moments of T.

The effect of the additional shape parameter α on the kurtosis of the OLLS model can be considered based on quantile measures as given in (3.9). The shortcomings of the classical kurtosis measure are well-known. One of the earliest kurtosis measures to be suggested is the Bowley kurtosis. The Moors kurtosis is based on octiles and for the OLLS model is given by

$$K = \frac{Q_{T_\nu}(7/8) - Q_{T_\nu}(5/8) + Q_{T_\nu}(3/8) - Q_{T_\nu}(1/8)}{Q_{T_\nu}(6/8) - Q_{T_\nu}(2/8)}. \tag{5.30}$$

The statistic K is less sensitive to outliers and it exists even for distributions without moments. Plots of the kurtosis of T are displayed in Figures 5.6 and 5.7.

(a) (b)

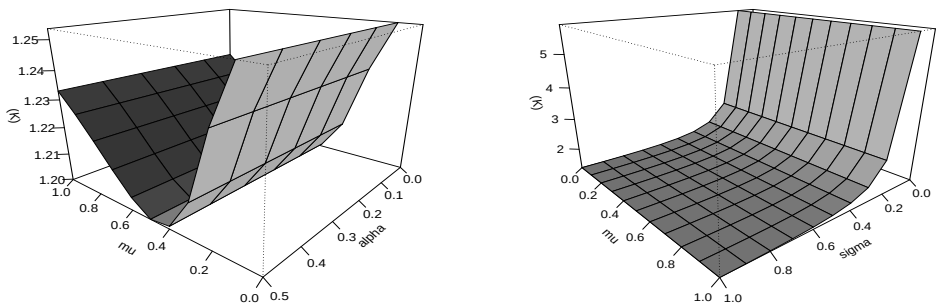


Figura 5.6 - Moors' kurtosis (K) for the OLLS distribution: plots a) as functions of $\mu \in [0,2]$ with $\alpha \in [1,2]$ and b) as functions of $\mu \in [0,1]$ with $\sigma \in [0,1]$.

Appendix C - The components of the score vector $U(\boldsymbol{\theta})$

The components of the score vector $U(\boldsymbol{\theta})$ for model (2.36) are given in this Appendix.

We use the following results

$$\frac{\partial \phi_{t_\nu}(z_{ij})}{\partial m} = \frac{\partial \phi_{t_\nu}(z_{ij})}{\partial \tau_i} = \frac{\partial \phi_{t_\nu}(z_{ij})}{\partial \beta_j} = \frac{1}{\sigma} z_{ij} \phi_{t_\nu}(z_{ij}), \qquad \frac{\partial \phi_{t_\nu}(z_{ij})}{\partial \sigma} = \frac{1}{\sigma} z_{ij}^2 \phi_{t_\nu}(z_{ij}),$$

(a) (b)

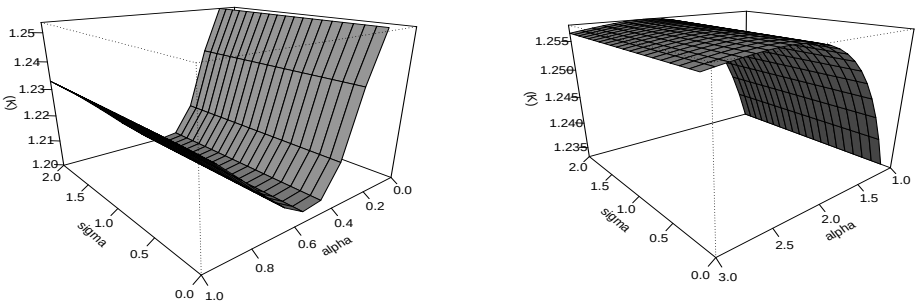


Figura 5.7 - Moors' kurtosis (K) for the OLLS distribution: plots a) as functions of $\sigma \in [0,1]$ with $\alpha \in [0,1]$ and b) as functions of $\sigma \in [0,2]$ with $\alpha \in [1,3]$.

$$\frac{\Phi_{T_\nu}(z_{ij})}{\partial\sigma} = \frac{1}{\sigma} z_{ij}\phi_{t_\nu}(z_{ij}), \qquad \frac{\partial\Phi_{T_\nu}(z_{ij})}{\partial m} = \frac{\partial\Phi_{T_\nu}(z_{ij})}{\partial\tau_i} = \frac{\partial\Phi_{T_\nu}(z_{ij})}{\partial\beta_j} = \frac{1}{\sigma}\phi_{t_\nu}(z_{ij}).$$

The components of the score vector $U(\boldsymbol{\theta})$ for model (2.36) are given by

$$\begin{aligned} U_\alpha(\boldsymbol{\theta}) &= \frac{n}{\alpha} - 2\sum_{i=1}^{I}\sum_{j=1}^{J} \frac{\Phi^{\alpha}_{T_\nu}(z_{ij})\log\left[\Phi_{T_\nu}(z_{ij})\right] + \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha}\log\left[1-\Phi_{T_\nu}(z_{ij})\right]}{\Phi^{\alpha}_{T_\nu}(z_{ij}) + \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha}} \\ &\quad + \sum_{i=1}^{I}\sum_{j=1}^{J}\log\left\{\Phi_{T_\nu}(z_{ij})\left[1-\Phi_{T_\nu}(z_{ij})\right]\right\}, \end{aligned}$$

$$\begin{aligned} U_m(\boldsymbol{\theta}) &= +\sum_{i=1}^{n}\frac{(\nu+1)\,z_i}{\sigma\left(\nu+z_i^2\right)} - \frac{2\alpha}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\phi_{t_\nu}(z_{ij})\left\{\Phi^{\alpha-1}_{T_\nu}(z_{ij}) - \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha-1}\right\}}{\Phi^{\alpha}_{T_\nu}(z_{ij}) + \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha}} \\ &\quad + \frac{(\alpha-1)}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{\phi_{t_\nu}(z_{ij})\left[1-2\Phi_{T_\nu}(z_{ij})\right]}{\Phi_{T_\nu}(z_{ij})\left[1-\Phi_{T_\nu}(z_{ij})\right]}, \end{aligned}$$

$$\begin{aligned} U_\sigma(\boldsymbol{\theta}) &= \frac{(\nu+1)}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{z_{ij}^2}{(\nu+z_{ij}^2)} - \frac{2\alpha}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{z_{ij}\phi_{t_\nu}(z_{ij})\left\{\Phi^{\alpha-1}_{T_\nu}(z_{ij}) - \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha-1}\right\}}{\Phi^{\alpha}_{T_\nu}(z_{ij}) + \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha}} \\ &\quad + \frac{(\alpha-1)}{\sigma}\sum_{i=1}^{I}\sum_{j=1}^{J}\frac{z_{ij}\phi_{t_\nu}(z_{ij})\left[1-2\Phi_{T_\nu}(z_{ij})\right]}{\Phi_{T_\nu}(z_{ij})\left[1-\Phi_{T_\nu}(z_{ij})\right]} - \frac{n}{\sigma}, \end{aligned}$$

$$U_{\tau_i}(\boldsymbol{\theta}) = \frac{1}{\sigma}\sum_{j=1}^{J} z_{ij} - \frac{2\alpha}{\sigma}\sum_{j=1}^{J}\frac{\phi_{t_\nu}(z_{ij})\left\{\Phi_{T_\nu}^{\alpha-1}(z_{ij}) - \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha-1}\right\}}{\Phi_{T_\nu}^{\alpha}(z_{ij}) + \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha}}$$
$$\frac{(\alpha-1)}{\sigma})\sum_{j=1}^{J}\frac{(z_{ij})\left[1-2\Phi_{T_\nu}(z_{ij})\right]}{\Phi_{T_\nu}(z_{ij})\left[1-\Phi_{T_\nu}(z_{ij})\right]},$$

$$U_{\beta_j}(\boldsymbol{\theta}) = \frac{1}{\sigma}\sum_{i=1}^{I} z_{ij} - \frac{2\alpha}{\sigma}\sum_{i=1}^{I}\frac{\phi_{t_\nu}(z_{ij})\left\{\Phi_{T_\nu}^{\alpha-1}(z_{ij}) - \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha-1}\right\}}{\Phi_{T_\nu}^{\alpha}(z_{ij}) + \left[1-\Phi_{T_\nu}(z_{ij})\right]^{\alpha}}$$
$$\frac{(\alpha-1)}{\sigma}\sum_{i=1}^{I}\frac{\phi_{t_\nu}(z_{ij})\left[1-2\Phi_{T_\nu}(z_{ij})\right]}{\Phi_{T_\nu}(z_{ij})\left[1-\Phi_{T_\nu}(z_{ij})\right]}.$$

Appendix D - Description of the function OLLSN

The OLLSN model is implemented in the gamlss function. We omitted several functions for the gamlss package and present only the functions related to the OLLSN distribution and its fit to a data set. This case, we presents the main functions related to the OLLSN model as well as basic functions to fit a OLLSN regression model.

Usage

OLLSN$(y, \mu, \sigma, \nu, \tau, \log = \text{FALSE})$

Function	**Description**
`dOLLSN(y,mu,sigma,lambda,alpha,log=FALSE)`	`the density function`
`pOLLSN(q,mu,sigma,lambda,alpha,log=FALSE)`	`the distribution function`
`qOLLSN(p,mu,sigma,lambda,alpha,log=FALSE)`	`the quantile function`
`rOLLSN(n,mu,sigma,lambda,alpha)`	`generates random deviates`

Arguments of gamlss $(\ldots, \text{family} = \text{"OLLSN"})$ **function**

```
            y   ~  explanatory variables related to mu
sigma.formula = ~  explanatory variables related to sigma
   nu.formula = ~  explanatory variables related to lambda
  tau.formula = ~  explanatory variables related to alpha
    alpha.fix = 0  for alpha = 0 we obtain the SN nested model
```

Extra arguments

```
 plot(model)           model diagnostic tools
 model$parameter.fv    extract the fitted value of parameter
 vcov(model)           to extract the variance-covariance matrix
 logLik(model)         extract the log likelihood.
```

To optimize the computational time, we can change the initial values of parameters using the `parameter.fix` function, otherwise we can increase the number of interactions using the `n.cyc` function. Fitting the model to censored data can be performed using the additional package gamlss.cens. The structure of the gamlss function is familiar to readers used to the R syntax (the glm function, in particular).

Appendix E - Program OLLSN for package gamlss

```
require(numDeriv)

OLLSN <- function (mu.link = "identity",sigma.link="log",nu.link =
                   "identity",tau.link = "log"){
mstats <- checklink( "mu.link", "odd log logistic skew normal",
                        substitute(mu.link),
mstats <- checklink( "mu.link", "odd log logistic skew normal",
                       substitute(mu.link),
                       c("inverse", "log", "identity", "own"))
dstats <- - checklink("sigma.link", "odd log logistic skew normal",
                         substitute(sigma.link),
```

```
                        c("inverse", "log", "identity", "own"))
vstats <- checklink("nu.link", "odd log logistic skew normal",
                      substitute(nu.link),
                      c("1/nu**2", "log", "identity", "own"))
tstats <- checklink(  "tau.link", "odd log logistic skew normal",
                        substitute(tau.link),
                        c("1/tau**2", "log", "identity", "own"))

      structure(
list(family = c("OLLSN", "odd log logistic skew normal"),
          parameters = list(mu=TRUE, sigma=TRUE, nu=TRUE, tau=TRUE),
               nopar = 4,
                type = "Continuous",
             mu.link = as.character(substitute(mu.link)),
          sigma.link = as.character(substitute(sigma.link)),
             nu.link = as.character(substitute(nu.link)),
            tau.link = as.character(substitute(tau.link)),
          mu.linkfun = mstats$linkfun,
       sigma.linkfun = dstats$linkfun,
          nu.linkfun = vstats$linkfun,
         tau.linkfun = tstats$linkfun,
          mu.linkinv = mstats$linkinv,
       sigma.linkinv = dstats$linkinv,
          nu.linkinv = vstats$linkinv,
         tau.linkinv = tstats$linkinv,
               mu.dr = mstats$mu.eta,
            sigma.dr = dstats$mu.eta,
               nu.dr = vstats$mu.eta,
              tau.dr = tstats$mu.eta,

=====================================================


dldm <- function(y,mu,sigma,nu,tau){
lpdf <- function(t,x,sigma,nu,tau){log(dauxiOLLSN(t,x,sigma,nu,tau))}
```

```
dldm <- grad(func=lpdf,t=y,x=mu,sigma=sigma,nu=nu,tau=tau,method='simple')
dldm
           },

=====================================================


d2ldm2 <- function(y,mu,sigma,nu,tau){
  lpdf <- function(t,x,sigma,nu,tau){log(dauxiOLLSN(t,x,sigma,nu,tau))}
  dldm <- grad(func=lpdf,t=y,x=mu,sigma=sigma,nu=nu,tau=tau,method='simple')
d2ldm2 <-  -dldm * dldm d2ldm2 <- ifelse(d2ldm2 = 1e-15,
d2ldm2,-1e-15) d2ldm2
            },

=====================================================


dldd <- function(y,mu,sigma,nu,tau){ lpdf <-
function(t,mu,x,nu,tau){log(dauxiOLLSN(t,mu,x,nu,tau))} dldd <-
grad(func=lpdf,t=y,mu=mu,x=sigma,nu=nu,tau=tau,method='simple') dldd
           } ,

=====================================================


d2ldd2 <- function(y,mu,sigma,nu,tau){
 lpdf  <- function(t,mu,x,nu,tau){log(dauxiOLLSN(t,mu,x,nu,tau))}
 dldd  <- grad(func=lpdf,t=y,mu=mu,x=sigma,nu=nu,tau=tau,method='simple')
d2ldd2 <- -dldd*dldd d2ldd2 <- ifelse(d2ldd2 < -1e-15,
d2ldd2,-1e-15) d2ldd2
           },

=====================================================


dldv <- function(y,mu,sigma,nu,tau){ lpdf <-
```

```
function(t,mu,sigma,x,tau){log(dauxiOLLSN(t,mu,sigma,x,tau))} dldv
<-
grad(func=lpdf,t=y,mu=mu,sigma=sigma,x=nu,tau=tau,method='simple')
dldv
          },

=====================================================


 d2ldv2 <- function(y,mu,sigma,nu,tau){
   lpdf <- function(t,mu,sigma,x,tau){log(dauxiOLLSN(t,mu,sigma,x,tau))}
   dldv <- grad(func=lpdf,t=y,mu=mu,sigma=sigma,x=nu,tau=tau,method='simple')
 d2ldv2 <- -dldv * dldv
d2ldv2  <- ifelse(d2ldv2 < -1e-15, d2ldv2,-1e-15) d2ldv2
           },

=====================================================


dldt <- function(y,mu,sigma,nu,tau){ lpdf <-
function(t,mu,sigma,nu,x){log(dauxiOLLSN(t,mu,sigma,nu,x))} dldt <-
grad(func=lpdf,t=y,mu=mu,sigma=sigma,nu=nu,x=tau) dldt
           } ,

=====================================================


d2ldt2 <- function(y,mu,sigma,nu,tau){
  lpdf <- function(t,mu,sigma,nu,x){log(dauxiOLLSN(t,mu,sigma,nu,x))}
  dldt <- grad(func=lpdf,t=y,mu=mu,sigma=sigma,nu=nu,x=tau,method='simple')
d2ldt2 <- -dldt * dldt d2ldt2 <- ifelse(d2ldt2 < -1e-15,
d2ldt2,-1e-15) d2ldt2
           },

=====================================================
```

```
d2ldmdd <- function(y,mu,sigma,nu,tau){\\
   lpdf <- function(t,x,sigma,nu,tau){log(dauxiOLLSN(t,x,sigma,nu,tau))}
   dldm <- grad(func=lpdf,t=y,x=mu,sigma=sigma,nu=nu,tau=tau,method='simple')
   lpdf <- function(t,mu,x,nu,tau){log(dauxiOLLSN(t,mu,x,nu,tau))}
   dldd <- grad(func=lpdf,t=y,mu=mu,x=sigma,nu=nu,tau=tau)
d2ldmdd <-- (dldm * dldd) d2ldmdd <-
ifelse(is.na(d2ldmdd)==TRUE,0,d2ldmdd) d2ldmdd
           },

=====================================================


d2ldmdv <- function(y,mu,sigma,nu,tau){
   lpdf <- function(t,x,sigma,nu,tau){log(dauxiOLLSN(t,x,sigma,nu,tau))}
   dldm <- grad(func=lpdf,t=y,x=mu,sigma=sigma,nu=nu,tau=tau,method='simple')
   lpdf <- -function(t,mu,sigma,x,tau){log(dauxiOLLSN(t,mu,sigma,x,tau))}
   dldv <- -grad(func=lpdf,t=y,mu=mu,sigma=sigma,x=nu,tau=tau)
d2ldmdv <- -(dldm * dldv) d2ldmdv
           },

=====================================================


 d2ldmdt <- function(y,mu,sigma,nu,tau){
    lpdf <- function(t,x,sigma,nu,tau){log(dauxiOLLSN(t,x,sigma,nu,tau))}
    dldm <- grad(func=lpdf,t=y,x=mu,sigma=sigma,nu=nu,tau=tau,method='simple')
    lpdf <- function(t,mu,sigma,nu,x){log(dauxiOLLSN(t,mu,sigma,nu,x))}
    dldt <- grad(func=lpdf,t=y,mu=mu,sigma=sigma,nu=nu,x=tau)
 d2ldmdt <- -(dldm*dldt)
 d2ldmdt
          },

=====================================================


d2ldddv <- function(y,mu,sigma,nu,tau){
```

```
   lpdf <- function(t,mu,x,nu,tau){log(dauxiOLLSN(t,mu,x,nu,tau))}\\
   dldd <- grad(func=lpdf,t=y,mu=mu,x=sigma,nu=nu,tau=tau,method='simple')
   lpdf <- function(t,mu,sigma,x,tau){log(dauxiOLLSN(t,mu,sigma,x,tau))}
   dldv <- grad(func=lpdf,t=y,mu=mu,sigma=sigma,x=nu,tau=tau)
d2ldddv <- -(dldd * dldv) d2ldddv
           },

=================================================


d2ldddt <- function(y,mu,sigma,nu,tau){
   lpdf <- function(t,mu,x,nu,tau){log(dauxiOLLSN(t,mu,x,nu,tau))}
   dldd <- grad(func=lpdf,t=y,mu=mu,x=sigma,nu=nu,tau=tau)\\
   lpdf <- function(t,mu,sigma,nu,x){log(dauxiOLLSN(t,mu,sigma,nu,x))}
   dldt <- grad(func=lpdf,t=y,mu=mu,sigma=sigma,nu=nu,x=tau)
d2ldddt <- -(dldd*dldt) d2ldddt
           },

=================================================


d2ldvdt <- function(y,mu,sigma,nu,tau){
   lpdf <- function(t,mu,sigma,x,tau){log(dauxiOLLSN(t,mu,sigma,x,tau))}
   dldv <- grad(func=lpdf,t=y,mu=mu,sigma=sigma,x=nu,tau=tau)\\
   lpdf <- function(t,mu,sigma,nu,x){log(dauxiOLLSN(t,mu,sigma,nu,x))}
   dldt <- grad(func=lpdf,t=y,mu=mu,sigma=sigma,nu=nu,x=tau)
d2ldvdt <- -(dldv*dldt) d2ldvdt
           },

=================================================


G.dev.incr <- function(y,mu,sigma,nu,tau,...)
                        {
                        -2*dOLLSN(y,mu,sigma,nu,tau,log=TRUE)
                         } ,
```

```
     rqres <- expression(
                rqres(pfun="pOLLSN", type="Continuous", y=y, mu=mu,
                sigma=sigma, nu=nu, tau=tau)) ,
   mu.initial <- expression(mu <- (y+mean(y))/2),
sigma.initial <- expression(sigma <- rep(sd(y), length(y))),
   nu.initial <- expression(nu <- rep(1, length(y))),
  tau.initial <- expression(tau <-rep(1, length(y))),
     mu.valid <- function(mu) TRUE,
  sigma.valid <- function(sigma) all(sigma > 0),
     nu.valid <- function(nu) TRUE,
    tau.valid <- function(tau) all(tau > 0),
      y.valid <- function(y)  TRUE
         ),
           class = c("gamlss.family","family"))
                       }
```

==========

```
dOLLSN <- function(x, mu = 0, sigma = 1, nu = 1, tau = .5, log =
FALSE){
          if (any(tau < 0))stop(paste("tau must be positive", "\ n", ""))
 pdfsn <- (2)*dnorm(x,mu,sigma)*pnorm(nu*((x-mu)/sigma))
 cdfsn <- pnorm((x-mu)/sigma)-2*owen((x-mu)/sigma,nu)\\
   fy1 <- (tau*pdfsn*(cdfsn*(1-cdfsn))**(tau-1))/((cdfsn**tau+(1-cdfsn)
           **tau)**2)
     if(log==FALSE) fy<-fy1 else fy<-log(fy1)
  fy
    }
```

==========

```
pOLLSN <- function(q, mu = 0, sigma = 1, nu = 1, tau = .5,
lower.tail
```

```
                    = TRUE, log.p = FALSE){
         if (any(tau < 0))  stop(paste("tau must be positive", "\ n", ""))
 cdfsn <- pnorm((q-mu)/sigma)-2*owen((q-mu)/sigma,nu)
  cdf1 <- (cdfsn**tau)/(cdfsn**tau+(1-cdfsn)**tau)
if(lower.tail==TRUE) cdf<-cdf1 else cdf<- 1-cdf1 if(log.p==FALSE)
cdf<- cdf else cdf<- log(cdf) cdf
 }
```

==========

```
qOLLSN <-  function(p, mu=0, sigma=1, nu=1, tau=.5, lower.tail =
                   TRUE, log.p = FALSE){
if (any(sigma < 0))  stop(paste("sigma must be positive", "\ n",
""))
if (any(tau < 0))    stop(paste("tau must be positive", "\ n", ""))\\
if (any(p < 0)|any(p > 1))  stop(paste("p must be between 0 and 1",
"\ n", "")) if (log.p==TRUE) p <- exp(p) else p <- p if
(lower.tail==TRUE) p <- p else p <- 1-p u <-
(p**(1/tau))/((1-p)**(1/tau)+p**(1/tau)) q <- mu + sigma*qsn(p=u,
xi=0 , omega=1, alpha=nu) q
 }
```

==========

```
rOLLSN <- function(n, mu=0, sigma=1, nu=1, tau=.5){ if (any(tau <
0))  stop(paste("tau must be positive", "\ n", ""))
    uni<- runif(n = n,0,1)
r <- qOLLSN(uni,mu =mu, sigma =sigma, nu=nu, tau=tau) r
 }
```

==========

```
owen <- function(h,a){ func <-
function(x,h)((exp(-0.5*h**2*(1+x**2)))/(1+x**2))*(1/(2*pi))
  temp1<-c()
  if(length(h) > 1  length(a)>1){ for(i in 1:length(h)){
    int <- integrate(f=func,lower =0,upper=a[i],h=h[i])
  temp1 <- c(temp1,intvalue)
  } }
  if(length(h) > 1  length(a)==1){ for(i in 1:length(h))\{
      int <- integrate(f=func,lower =0,upper=a,h=h[i])
    temp1 <- c(temp1,intvalue)
  } }
  if(length(h)==1  length(a)==1){
      int <- integrate(f=func,lower =0,upper=a,h=h)
    temp1 <- c(temp1,intvalue)
  }
return(temp1)}

=================================================


dauxiOLLSN <- function(t,mu,sigma,nu,tau){
     pdfsn <- (2)*dnorm(t,mu,sigma)*pnorm(nu*((t-mu)/sigma))
     cdfsn <- pnorm((t-mu)/sigma)-2*owen((t-mu)/sigma,nu)
       fy1 <- (tau*pdfsn*(cdfsn*(1-cdfsn))**(tau-1))/((cdfsn**tau +
              (1-cdfsn)**tau)**2)
       fy1
  }
```

Appendix F - The program model mixed

```
 rm(list=ls())
 require(sn)
 require(gamlss
```

```
 require(gmodels)
 require(sn)
 require(fGarch)
 require(GHQp)
 require(VGAM)
```

==========

Log - likelihood function fixed effects

```
llF      <- function(theta, Y, X1){
  nbeta1 <- ncol(X1)
  beta1  <- theta[1:nbeta1]
  sigma  <- theta[nbeta1+1]
  mu     <- X1%*%beta1
  lv     <- -sum(dnorm(Y, mean = mu, sd = exp(sigma), log = TRUE))
  return(lv)
    }
```

==========

Log - likelihood function mixed effects

```
llM <- function(theta, Y, X1, subject, quad) {
  N <- nlevels(subject)
  nbeta1 <- ncol(X1)
  beta1  <- theta[1:nbeta1]
  sigma  <- theta[nbeta1+1]
   t1    <- exp(theta[nbeta1+2])
  i      <- 1:N
  ll     <- lapply(i, llind, Y, X1, sigma, subject, beta1, t1)
  -sum(log(unlist(ll)))
  }
llind <- function(i, Y, X1, sigma, subject, beta1,t1) {
```

```
  y   <-  Y[subject==i]
  x1  <- X1[subject==i,]
  opt <- try(optim(par=0, fn=integrando,
                   y=y, x1=x1,
                   beta1=beta1, sigma=sigma, t1=t1,
                   log=TRUE,
                   hessian=TRUE, method="BFGS",
                   control=list(fnscale=-1)))
  if(class(opt) != "try-error"){
    x.hat <- opt$par
    Q   <- solve(-opt$hessian)
    Q12 <- chol(Q)
    Z   <- x.hat + sqrt(2) *as.numeric(Q12)*quad$nodes
    norma <- exp(-rowSums(quad$nodes^2))
    temp <- integrando(Z, y=y, x1=x1, beta1=beta1, sigma=sigma,
                       t1=t1, log=FALSE)
    integral <- sqrt(2) * det(Q) * sum(quad$product * temp / norma)
    return(integral)
  }
}
 integrando <- function(u,y,x1,beta1,sigma,t1,log=TRUE) {
               if(class(dim(u)) == "NULL"){u <- matrix(u,nrow=1,ncol=1)}
        ll <- apply(u,1,function(ui){
               if(ncol(X1)==1){
        mu <- x1*beta1 + ui
                            } else  mu <- x1%*%beta1 + x1[,1]*ui
    lv      <- sum(dnorm(y, mean = mu, sd = sigma, log = TRUE))
    temp1   <- lv
    temp2   <- dnorm(ui,0,t1, log=TRUE)
    temp1 + temp2  } )
  if(log == FALSE) ll <- exp(ll)
  return(ll)
}
```

==========

Orthodont data

```
Y        <- Orthodont$distance
Age      <- Orthodont$age
Subject  <- Orthodont$Subject
Subject  <- gl(27,4)
A        <- aov(Y ~ 1)
X1       <- model.matrix(A)
n.points <- 10
quad     <- GHQ(n=n.points, ndim=1, pruning=FALSE)
summary.lm(A)
```

==========

Fitting of the fixed effects to obtain start values

```
aux <- optim(c(24.02,2.929),  gr = NULL,llF, Y=Y, X1=X1, method = "BFGS",
       hessian = TRUE)
c(summary.lm(A)[4]$coefficients[,1],summary.lm(A)[6]$sigma);c(aux$par[1],
exp(aux$par[2]))
```

==========

Fitting of the mixed effects

```
n.points <- 10
quad     <- GHQ(n=n.points, ndim=2, pruning=FALSE)
aux1 <- optim(c(24.02,log(2.22),log(1.93)), llM, gr=NULL,
        Y=Y, X1=X1,subject=Subject,quad=quad,method = "BFGS",
              hessian = TRUE)
lmer(Y ~  (1|Subject))
c(aux1$par[1],exp(aux1$par[2:3]))
```

```
=================================================


Linear mixed model fit by REML ['lmerMod']
Formula: Y ~ (1 | Subject)
REML criterion at convergence: 515.3618
Random effects:
 Groups   Name        Std.Dev.
 Subject  (Intercept) 1.937
 Residual             2.220
Number of obs: 108, groups:  Subject, 27
Fixed Effects:
(Intercept)
      24.02
c(aux1$par[1],exp(aux1$par[2:3]))
[1] 24.032989  2.326995  1.836093
```

Appendix G - Adaptative Gauss-Hermite quadrature

The adaptative of Gauss-Hermite quadrature is important to resolve integrals of the type,

$$\int_{-\infty}^{\infty} g\left(x\right)dx = \int_{-\infty}^{\infty} f\left(x\right)v(x)dx$$

where $v(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$.

Thus, choosing μ and σ^2 so that $g(x)$ be sampled in an appropriate region. In specific, consider $\hat{\mu}$ and $\hat{\sigma}^2$ given by,

$$\hat{\mu} = \arg\underbrace{\max g(x)}_{x} \qquad \text{e} \qquad \hat{\sigma}^2 = \left[-\frac{d^2}{dx^2}\log\left(g(x)\right)\right]_{x=\hat{\mu}}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ are estimators of the mean and variance of the normal distribution. If we define the function $g(x)$ as the product $g(x) = h(x)v(x;\hat{\mu},\hat{\sigma}^2)$, then we can write,

$$\int_{-\infty}^{\infty} g\left(x\right)dx = \int_{-\infty}^{\infty} h\left(x\right)\frac{1}{\hat{\sigma}\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\hat{\mu}}{\hat{\sigma}}\right)^2\right\}dx. \tag{5.31}$$

If we define $z = (x-\hat{\mu})\Big/\sqrt{2\hat{\sigma}^2}$, whereas: if x $\rightarrow \infty \Rightarrow$ x $\rightarrow \infty$; if x $\rightarrow -\infty \Rightarrow$ x $\rightarrow$

$-\infty$ *and* $dz = \frac{1}{\sqrt{2\hat{\sigma}^2}}$. Then, applying the Gauss-Hermite quadrature in the equation (5.31) we have to,

$$\int_{-\infty}^{\infty} g\left(x\right)dx = \int_{-\infty}^{\infty} h\left(\hat{\sigma}z\sqrt{2}+\hat{\mu}\right)\frac{e^{-z^2}}{\sqrt{\pi}}dz = \sum_{p=1}^{q}\frac{v_k}{\sqrt{\pi}}h\left(\hat{\sigma}s_k\sqrt{2}+\hat{\mu}\right) = \hat{\sigma}\sqrt{2}\sum_{k=1}^{q}v_k^{+}g\left(s_k^{+}\right)$$

with $w_k^{+} = w_k.\exp\left\{s_k^2\right\}$ and $s_k^{+} = \sqrt{2}\hat{\sigma}s_k + \hat{\mu}$, where q is the number of points used for approximation, $w(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$, s_k are the quadrature points $\{s_1, s_2, \cdots, s_k\}$ corresponds to the roots of the Hermite polynomial $H_q(x_k)(k = 1, ..., q)$, expressed by:

$$H_q(x) = (-1)^q \exp\left\{x^2\right\}\frac{d^q}{dx^q}\left\{\exp\left\{-x^2\right\}\right\},$$

and the associated weights w_k are given by:

$$w_k = \frac{2^{q-1}q!\sqrt{\pi}}{q^2\left[H_{q-1}(s_k)\right]^2}.$$

The function *gauss.quad* $(\cdot)$ of the package statmod of the statistical software R provides Gauss-Hermite the points and the weights. For example, with this instruction we can obtain the quadrature points and weights with q = 5.

```
quad <- gauss.quad(n=5, kind="hermite")
quad
```

```
$nodes [1] -2.020183e+00 -9.585725e-01  2.402579e-16 9.585725e-01
            2.020183e+00

$weights [1] 0.01995324 0.39361932 0.94530872 0.39361932 0.01995324
```

Defining the fuction:

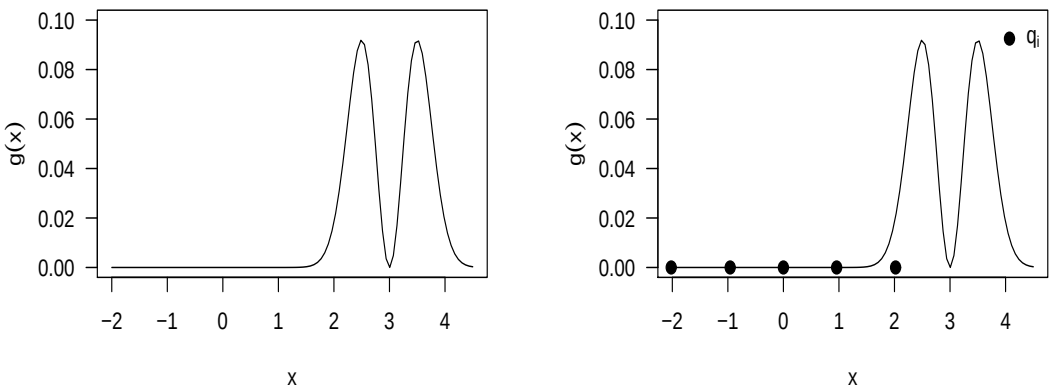
```
     g <- function(x) (x-3)^2*exp(-((x-3)/0.5)^2)
 log.g <- function(x) 2*log(x-3)-(x-3)^2
```

To obtain figure and adding the quadrature points:

```
curve(g, -2, 4.5, ylim=c(0,0.1), ylab=expression(g(x)), las=1)
```

Approximating the integral of g(x) by expression Gauss-Hermite quadrature

(a) (b)




```
sum(quad$weights * g(quad$nodes)*exp(quad$nodes^2))
0.0243759
integrate(f=g, lower=-Inf, upper=Inf)
0.1107784 with absolute error 9.9e-05
```

Using the adaptive version

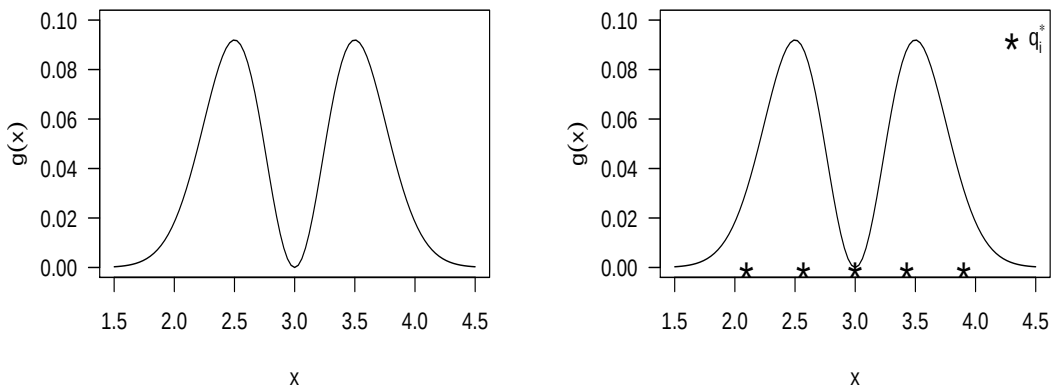
```
      opt <- optim(fn=log.g3,par=0,control=list(fnscale=-1),hessian=T,
                   method='CG')
    x.hat <- opt$par
sigma.hat <- as.numeric(sqrt(-1/opt$hessian))
       xi <- sqrt(2)*sigma.hat*quad$nodes + x.hat
       wi <- quad$weights

curve(g3, 1.5, 4.5, ylim=c(0,0.1), ylab=expression(g(x)), las=1)
points(x=xi, y=rep(0,5), pch='*', cex=2) legend('topright', bty='n',
legend=expression(q[i]^symbol("*")), pch='*', pt.cex=2)
```

Plotting g(x) with new quadrature points,

```
curve(g3, 1.5, 4.5, ylim=c(0,0.1), ylab=expression(g(x)), las=1)
points(x=xi, y=rep(0,5), pch='*', cex=2) legend('topright', bty='n',
legend=expression(q[i]^symbol("*")), pch='*', pt.cex=2)
```

(a) (b)




```
Approximating the integral of g(x) by Quadrature Gauss Hermite
Adaptative

sqrt(2)*sigma.hat*sum(wi*g3(xi)*exp(quad$nodes^2))
0.110702

True value with integrate function,

integrate(f=g, lower= - Inf, upper = Inf) 0.1107784 with absolute
error < 9.9e-05 .

==================================================
```