

Research Article

A New Class of Distributions Generated by the Extended Bimodal-Normal Distribution

Milton A. Cortés, David Elal-Olivero, and Juan F. Olivares-Pacheco 

Departamento de Matemática, Facultad de Ingeniería, Universidad de Atacama, Copiapó, Chile

Correspondence should be addressed to Juan F. Olivares-Pacheco; jfolivar@uda.cl

Received 21 June 2018; Accepted 27 September 2018; Published 1 November 2018

Academic Editor: Steve Su

Copyright © 2018 Milton A. Cortés et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this study, we present a new family of distributions through generalization of the extended bimodal-normal distribution. This family includes several special cases, like the normal, Birnbaum-Saunders, Student's t , and Laplace distribution, that are developed and defined using stochastic representation. The theoretical properties are derived, and easily implemented Monte Carlo simulation schemes are presented. An inferential study is performed for the Laplace distribution. We end with an illustration of two real data sets.

1. Introduction

Although the normal distribution is the most popular probability model in statistics, several random phenomena in nature cannot be described by the normal distribution. In this regard, Azzalini [1] introduced an extension of the normal distribution called skew-normal distribution, where this model shares some properties with the standard normal model; it is mathematically tractable and it has a wide range of the coefficients of skewness and kurtosis. From this work, an important line of research focusing on finding new distributions that offered greater flexibility is generated.

More recently, Elal-Olivero [2] introduced a new class of skew-normal distribution called alpha-skew-normal distribution. In doing so, he first defined a new bimodal-symmetric normal distribution with probability density function given by

$$\zeta(x) = x^2 \phi(x), \quad x \in \mathbb{R}, \quad (1)$$

where $\phi(\cdot)$ is the standard normal density, which is defined as the bimodal-normal (BN) distribution. Furthermore, he studies some properties of this distribution and presents its stochastic representation as the product of two independent random variables \sqrt{T} and V , where $T \sim \chi_{(3)}^2$ and V is a discrete random variable such that $P(V = \pm 1) = 1/2$; that

is, $X = \sqrt{T}V$ has the distribution BN. On the other hand, an extension of the BN density is given by

$$\phi_{\xi}(x) = \left(\frac{1 + \xi x^2}{1 + \xi} \right) \phi(x), \quad x \in \mathbb{R}, \quad (2)$$

where $\xi \geq 0$ is the shape parameter. Note that this density function is symmetric and is characterized by incorporating bimodality into the normal distribution, which is controlled by the parameter ξ . Elal-Olivero [2] presents this extension as the symmetric-component of the alpha-skew-normal distribution. Furthermore, (2) also can be deduced from the model presented in Elal-Olivero et al. [3]. In this regard, Gui et al. [4] incorporated (2) into the slash distribution, developed its properties, and performed inferential studies, whereas Gómez and Guerrero [5] incorporated (2) into the Birnbaum-Saunders distribution, tested its bimodality, and demonstrated its principal properties.

The objective of this article is to present a new family of distributions through generalization of (2). This generalization can be applied to any density function, thereby producing a more flexible model incorporating a shape parameter. Depending on the density at which we apply this generalization, it is observed that the new model is flexible enough to support uni- and bimodal shapes. Furthermore,

Gui et al. [4] and Gómez and Guerrero [5] are particular cases of the generalization proposal.

This article is organized as follows. In Section 2, we present a generalization of (2) and review some particular cases (normal, Birnbaum-Saunders, Student's t , and Laplace distribution). In Section 3, we develop the basic properties of the cases from Section 2 and study the effects of this new generalization. In Section 4, we study some inferential aspects of the extended Laplace distribution using maximum likelihood estimation and perform a Monte Carlo simulation study. We conclude in Section 5 with a discussion.

2. A General Class of Distributions

This section describes a general class of distributions generated by (2), presents its basic properties, and derives explicit expressions for the normal, Birnbaum-Saunders, Student's t , and Laplace distribution.

2.1. Characterization and Properties

Theorem 1 (general class of distributions). *Let g be a probability density function and h a positive continuous function such that $E_g[h(Y)] = \kappa < \infty$, where $Y \sim g$. Then,*

$$f(x; \xi) = \left(\frac{1 + \xi h(x)}{1 + \kappa \xi} \right) g(x), \quad x \in \mathbb{R}, \quad (3)$$

is a probability density function with shape parameter $\xi \geq 0$.

Proof. If we note $f(x; \xi)$ can be represented as a mixture of two densities, then the result follows immediately; that is, $f(x; \xi) = pg(x) + (1 - p)(1/\kappa)h(x)g(x)$, where $p = 1/(1 + \kappa\xi)$. \square

Remark 2. On the basis of Theorem 1, we can make the following observations:

- (1) If $\xi = 0$, then $f(x; \xi) = g(x)$, $x \in \mathbb{R}$.
- (2) If $\xi \rightarrow \infty$, then $f(x; \xi) \rightarrow (1/\kappa)h(x)g(x)$, $x \in \mathbb{R}$.

Theorem 3 (stochastic representation). *Let $X_1 \sim g(x)$ and $X_2 \sim (1/\kappa)h(x)g(x)$ be independent random variables. If*

$$X = \begin{cases} X_1, & \text{with probability } \frac{1}{1 + \kappa\xi}, \\ X_2, & \text{with probability } \frac{\kappa\xi}{1 + \kappa\xi}, \end{cases} \quad (4)$$

then $X \sim f(x; \xi)$.

Proof. Since $f(x; \xi)$ can be represented as a mixture, the result follows immediately. \square

Remark 4. If $X \sim f(x; \xi)$, then

- (1) The cumulative distribution function is given by

$$F(x; \xi) = \frac{1}{1 + \kappa\xi} (F_1(x; \xi) + \kappa\xi F_2(x; \xi)), \quad (5)$$

where $F_1(x; \xi)$ and $F_2(x; \xi)$ are the cumulative distribution functions of $X_1 \sim g(x)$ and $X_2 \sim (1/\kappa)h(x)g(x)$, respectively.

- (2) The moment generating function is given by

$$M_X(t) = \frac{1}{1 + \kappa\xi} (M_1(t) + \kappa\xi M_2(t)), \quad (6)$$

where $M_1(t)$ and $M_2(t)$ are the moments generating functions of $X_1 \sim g(x)$ and $X_2 \sim (1/\kappa)h(x)g(x)$, respectively, if both exist.

- (3) The r -th moment of the random variable X is given by

$$E[X^r] = \frac{1}{1 + \kappa\xi} (E_g[X^r] + \xi E_g[X^r h(X)]), \quad (7)$$

$r = 1, 2, \dots$

2.2. Special Cases. In this section, explicit expressions are provided for the probability density function in (3) for the normal, Birnbaum-Saunders, Student's t , and Laplace distribution and different choices of h . These models are selected to show the benefits of the proposed extension, and the choice of the function $h(x)$ is conditioned upon a positive function with finite expectation.

Corollary 5 (normal case). *If $h(x) = x^2$ and $g(x) = \phi(x)$, then X has the probability density function given by*

$$f(x; \xi) = \left(\frac{1 + \xi x^2}{1 + \xi} \right) \phi(x), \quad x \in \mathbb{R}, \quad (8)$$

and we say that X has an "extended normal distribution," which is denoted as $X \sim EN(\xi)$.

Corollary 6 (Birnbaum-Saunders case). *Let $a_x(\alpha, \beta) = (1/\alpha)(\sqrt{x}/\beta - \sqrt{\beta/x})$. If $h(x) = a_x^2(\alpha, \beta)$ and $g(x) = a'_x(\alpha, \beta)\phi(a_x(\alpha, \beta))$ where $a'_x(\alpha, \beta)$ is the derivative with respect to x , with $\alpha > 0$ and $\beta > 0$, then X has the probability density function given by*

$$f(x; \alpha, \beta, \xi) = \left(\frac{1 + \xi a_x^2(\alpha, \beta)}{1 + \xi} \right) a'_x(\alpha, \beta) \phi(a_x(\alpha, \beta)), \quad (9)$$

$x > 0$,

and we say that X has an "extended Birnbaum-Saunders distribution," which is denoted as $X \sim EBS(\alpha, \beta, \xi)$.

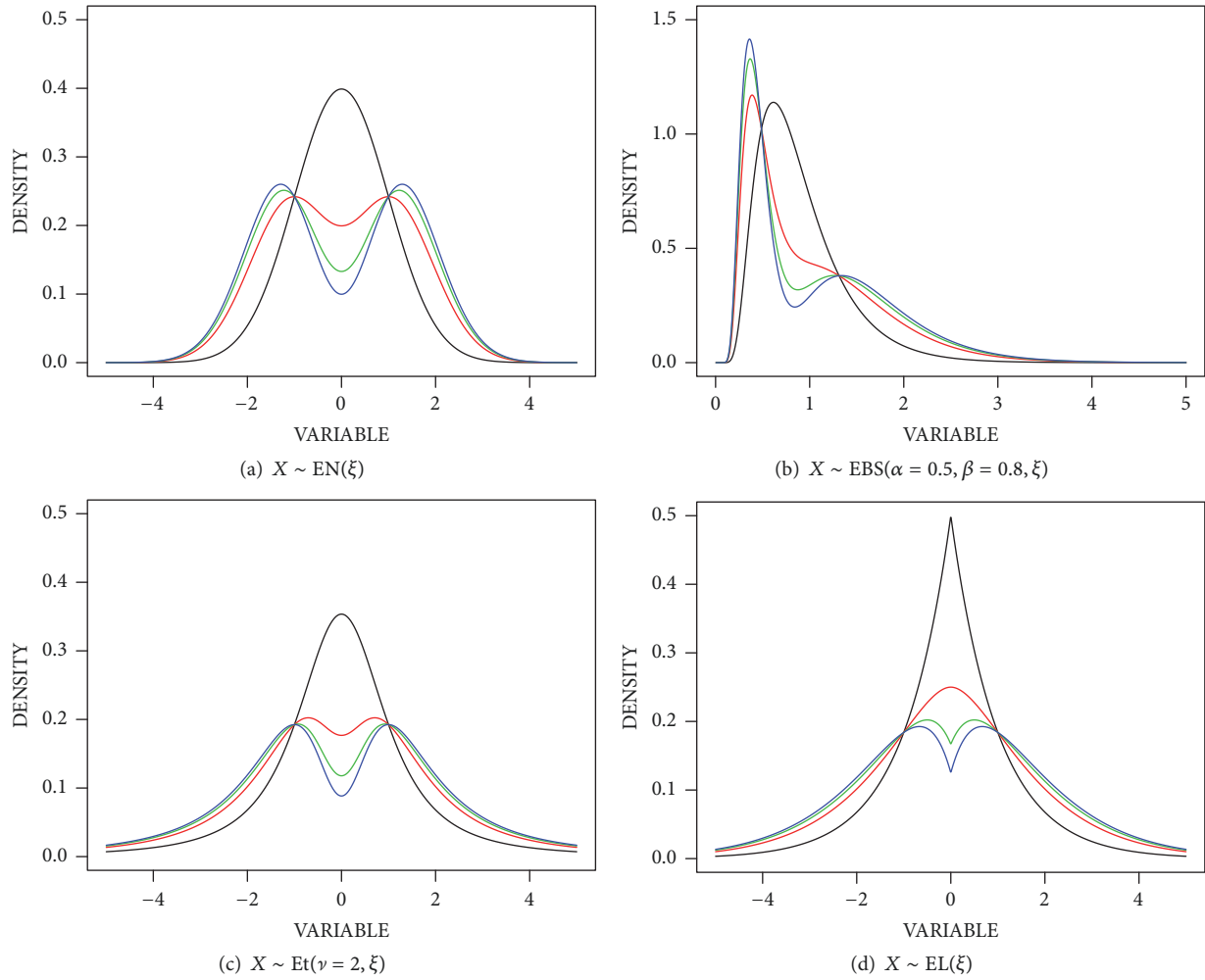


FIGURE 1: PDF's for the special cases defined in Corollaries 5–8 and several values of ξ . $\xi = 0$ (black line), $\xi = 1$ (red line), $\xi = 2$ (green line), and $\xi = 3$ (blue line).

Corollary 7 (Student's t case). If $h(x) = x^2(1 + \nu)/(x^2 + \nu)$ and $g(x) = (\Gamma((1 + \nu)/2)/\sqrt{\pi\nu}\Gamma(\nu/2))(1 + x^2/\nu)^{-(1+\nu)/2}$ where $\nu > 0$, then X has the probability density function given by

$$f(x; \nu, \xi) = \left(\frac{\nu + (1 + \xi + \xi\nu)x^2}{(\nu + x^2)(1 + \xi)} \right) \cdot \frac{\Gamma((1 + \nu)/2)}{\sqrt{\pi\nu}\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu} \right)^{-(1+\nu)/2}, \quad (10)$$

$x \in \mathbb{R}$,

and we say that X has an “extended Student's t distribution,” which is denoted as $X \sim Et(\nu, \xi)$.

Corollary 8 (Laplace case). If $h(x) = |x|$ and $g(x) = (1/2)e^{-|x|}$, then X has the probability density function given by

$$f(x; \xi) = \left(\frac{1 + \xi|x|}{2(1 + \xi)} \right) e^{-|x|}, \quad x \in \mathbb{R}, \quad (11)$$

and we say that X has an “extended Laplace distribution,” which is denoted as $X \sim EL(\xi)$.

As we notice, in the Corollaries 5–8 and Figure 1, when the function $g(x)$ is a symmetric density, the effect of the extension is that the model supports uni- and bimodal shapes. On the other hand, if the model has positive support, the bimodality depends on the choice of parameters, as seen in the Birnbaum-Saunders distribution case.

3. Some Results of the Special Cases

In this section, we develop some properties associated with the models defined in Corollaries 5–8. The cumulative distribution function, moment, and stochastic representation will be presented when they correspond to the cases at hand. Some proofs are straightforward and are, therefore, omitted.

3.1. Extended Normal Distribution. The extended normal distribution is the basis for the development of the specific

TABLE 1: Alternative stochastic representation of special cases.

Case	Representation	Variables
EN	$\sqrt{\frac{1}{1+\xi}}Z + \sqrt{\frac{\xi}{1+\xi}}X$	$Z \sim N(0, 1)$ and $X \sim \text{BN}$
EBS	$\beta \left(\frac{\alpha}{2}Y + \sqrt{\left(\frac{\alpha}{2}Y\right)^2 + 1} \right)^2$	$Y \sim \text{EN}(\xi)$
Et	$\frac{Y}{\sqrt{V/\nu}}$	$Y \sim \text{EN}(\xi)$ and $V \sim \chi_{(\nu)}^2$
EL	$\sqrt{2WY}$	$W \sim \text{exp}(1)$ and $Y \sim \text{EN}(\xi)$

cases discussed previously. If $g(x) = \phi(x)$ and $h(x) = x^2$, then, from Theorem 3, we know that $X_1 \sim \phi(x)$ and $X_2 \sim x^2\phi(x)$. Note that the distribution of X_2 corresponds to the bimodal-normal distribution, for which the stochastic representation was presented in Section 1.

The stochastic representation of $X \sim \text{EN}(\xi)$ is obtained through Theorem 3. Table 1 shows an alternative way to generate random variables $X \sim \text{EN}(\xi)$. Furthermore, the stochastic representation has a form that is similar to the representation given in Henze [6] for the skew-normal distribution presented in Azzalini [1].

Remark 9. If $X \sim \text{EN}(\xi)$, then

- (1) The cumulative distribution function is given by

$$F(x; \xi) = \Phi(x) - \frac{\xi x}{1+\xi} \phi(x), \quad x \in \mathbb{R}. \quad (12)$$

- (2) The r -th moment of the random variable X is given by

$$\begin{aligned} E[X^r] \\ = (1 + (-1)^r) \frac{2^{-1+r/2} (1 + \xi + r\xi) \Gamma((1+r)/2)}{\sqrt{\pi} (1 + \xi)}, \quad (13) \\ r = 1, 2, \dots \end{aligned}$$

- (3) The expected value and variance of the random variable X is given by

$$\begin{aligned} E[X] &= 0, \\ V(X) &= \frac{1 + 3\xi}{1 + \xi}. \end{aligned} \quad (14)$$

3.2. Extended Birnbaum-Saunders distribution. The Birnbaum-Saunders (BS) distribution (see Birnbaum and Saunders [7, 8]) describes the lifetime of components exposed to fatigue caused by cyclical stress and tension. Since 1969, the number of studies that have investigated this distribution and discussed the development of both its theoretical properties and its applications has increased dramatically. Because of its significance, this distribution has been extended in a variety of manners to relax its behavior and thus make it applicable

to a wide range of situations. For example, see Birnbaum and Saunders [7, 8], Mann et al. [9], Desmond [10, 11], Chang and Tang [12], Díaz-García and Leiva-Sánchez [13], Gómez et al. [14], and Olmos et al. [15]. The BS distribution with parameters $\alpha > 0$ and $\beta > 0$ has density function given by

$$g(x) = a'_x(\alpha, \beta) \phi(a_x(\alpha, \beta)), \quad x > 0, \quad (15)$$

where $a_x(\alpha, \beta)$ is defined in Corollary 6 and is denoted as $\text{BS}(\alpha, \beta)$. If $X_1 \sim \text{BS}(\alpha, \beta)$ and $h(x) = a_x^2(\alpha, \beta)$, then

$$X_2 \sim \frac{1}{\kappa} h(x) g(x) = a_x^2(\alpha, \beta) a'_x(\alpha, \beta) \phi(a_x(\alpha, \beta)), \quad (16) \\ x > 0,$$

with a stochastic representation given by

$$X_2 = \beta \left(\frac{\alpha}{2}Y + \sqrt{\left(\frac{\alpha}{2}Y\right)^2 + 1} \right)^2, \quad (17)$$

where $Y \sim \text{BN}$. From Theorem 1, the extended Birnbaum-Saunders distribution has density (9) and from Theorem 3 we can generate random variables $X \sim \text{EBS}(\alpha, \beta, \xi)$. An alternative way to generate this random variable can be seen in Table 1.

Theorem 10. Let $X \sim \text{EBS}(\alpha, \beta, \xi)$ with $\alpha > 0$, $\beta > 0$ and $\xi \geq 0$. Then

- (1) $aX \sim \text{EBS}(\alpha, a\beta, \xi)$, for $a > 0$.
(2) $X^{-1} \sim \text{EBS}(\alpha, \beta^{-1}, \xi)$.

Proof. The proofs are immediate from the theorem of the change of variable. \square

Remark 11. Like the Birnbaum-Saunders distribution observes that the property (1) established in Theorem 10 implies that the EBS distribution belongs to the scale family, whilst the property (2) implies that it also belongs to the family of random variables closed under reciprocation; see Saunders [16]. Furthermore, based on properties (1) and (2), we can have the two-parameter EBS distribution: $\alpha X/\beta \sim \text{EBS}(\alpha, \alpha, \xi)$.

Remark 12. If $X \sim \text{EBS}(\alpha, \beta, \xi)$, then

(1) The cumulative distribution function is given by

$$F(x; \alpha, \beta, \xi) = F_{\text{BS}}(x; \alpha, \beta) - \frac{\xi a_x(\alpha, \beta)}{1 + \xi} \phi(a_x(\alpha, \beta)), \quad x > 0, \quad (18)$$

where $F_{\text{BS}}(\cdot)$ is the cumulative distribution function of the Birnbaum-Saunders distribution.

(2) The r -th moment of the random variable X is given by

$$\begin{aligned} E[X^r] &= \beta^r \sum_{j=0}^r \binom{2r}{2j} \sum_{m=0}^j \binom{j}{m} \left(\frac{\alpha}{2}\right)^{2(r+m-j)} E[Y^{2(r+m-j)}], \quad (19) \\ & \quad r = 1, 2, \dots, \end{aligned}$$

where $Y \sim \text{EN}(\xi)$.

(3) The expected value and variance of the random variable X is given by

$$\begin{aligned} E[X] &= \beta + \frac{\beta(1 + 3\xi)\alpha^2}{2(1 + \xi)}, \\ V(X) &= \frac{\alpha^2 \beta^2 (4(1 + \xi)(1 + 3\xi) + (5 + 3\xi(10 + 7\xi))\alpha^2)}{4(1 + \xi)^2}. \quad (20) \end{aligned}$$

3.3. Extended Student's t -Distribution. The Student's t -distribution serves as a robust alternative when it is desired to model data sets with atypical values and with a coefficient of kurtosis that is greater than of the normal distribution. The Student's t -distribution with parameter $\nu > 0$ has a density function given by

$$g(x) = \frac{\Gamma((1 + \nu)/2)}{\sqrt{\pi\nu}\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu}\right)^{-(1+\nu)/2}, \quad x \in \mathbb{R}. \quad (21)$$

If $X_1 \sim g(x)$ and $h(x) = x^2(1 + \nu)/(x^2 + \nu)$, then,

$$\begin{aligned} X_2 &\sim \frac{1}{\kappa} h(x) g(x) \\ &= \left(\frac{x^2(1 + \nu)}{x^2 + \nu}\right) \frac{\Gamma((1 + \nu)/2)}{\sqrt{\pi\nu}\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu}\right)^{-(1+\nu)/2}, \quad (22) \\ & \quad x \in \mathbb{R}, \end{aligned}$$

with a stochastic representation given by $X_2 = Y/\sqrt{W/\nu}$, where $Y \sim \text{BN}$ and $W \sim \chi_{(\nu)}^2$ are independent random variables. From Theorem 1, the extended Student's t -distribution has density (10) and from Theorem 3 we can generate random variables $X \sim \text{Et}(\nu, \xi)$. An alternative way to generate this random variable can be seen in Table 1.

Remark 13. If $X \sim \text{Et}(\nu, \xi)$, then

(1) The cumulative distribution function is given by

$$F(x; \nu, \xi) = F_T(x; \nu) - \frac{\xi x}{1 + \xi} f_T(x; \nu), \quad x \in \mathbb{R}, \quad (23)$$

where $F_T(\cdot)$ and $f_T(\cdot)$ are the cumulative distribution function and probability density function of the t -Student distribution, respectively.

(2) The r -th moment of the random variable X is given by

$$\begin{aligned} E[X^r] &= (1 + (-1)^r) \frac{\nu^{r/2} (1 + \xi + r\xi)}{2\sqrt{\pi} (1 + \xi) \Gamma(\nu/2)} \Gamma\left(\frac{1+r}{2}\right) \\ & \quad \cdot \Gamma\left(\frac{-r+\nu}{2}\right), \quad r = 1, 2, \dots, \quad (24) \end{aligned}$$

where $\nu > r$.

(3) The expected value and variance of the random variable X is given by

$$\begin{aligned} E[X] &= 0, \quad \nu > 1, \\ V(X) &= \frac{(1 + 3\xi)\nu}{(1 + \xi)(-2 + \nu)}, \quad \nu > 2. \quad (25) \end{aligned}$$

3.4. Extended Laplace Distribution. The Laplace (L) or double exponential distribution, which was originally published by Pierre Laplace in 1774, is a symmetric distribution with density function given by

$$g(x) = \frac{1}{2} e^{-|x|}, \quad x \in \mathbb{R}. \quad (26)$$

If $X_1 \sim L$ and $h(x) = |x|$, then,

$$X_2 \sim \frac{1}{\kappa} h(x) g(x) = \frac{|x|}{2} e^{-|x|}, \quad x \in \mathbb{R}, \quad (27)$$

with a stochastic representation given by $X_2 = \sqrt{2W}Y$, where $W \sim \exp(1)$ and $Y \sim \text{BN}$ are independent random variables. From Theorem 1, the extended Laplace distribution has density (11) and from Theorem 3 we can generate random variables $X \sim \text{EL}(\xi)$. An alternative way to generate this random variable can be seen in Table 1.

Remark 14. If $X \sim \text{EL}(\xi)$, then

(1) The cumulative distribution function is given by

$$F(x; \xi) = \begin{cases} \frac{1}{2} e^x \left(1 - \frac{\xi x}{1 + \xi}\right), & \text{if } x < 0, \\ 1 - \frac{1}{2} e^{-x} \left(1 + \frac{\xi x}{1 + \xi}\right), & \text{if } x \geq 0. \end{cases} \quad (28)$$

(2) The r -th moment of the random variable X is given by

$$\begin{aligned} E[X^r] &= (1 + (-1)^r) \frac{r(1 + \xi + r\xi)}{2(1 + \xi)} \Gamma(r), \quad (29) \\ & \quad r = 1, 2, \dots \end{aligned}$$

(3) The expected value and variance of the random variable X are given by

$$\begin{aligned} E[X] &= 0, \\ V(X) &= \frac{2(1+3\xi)}{1+\xi}. \end{aligned} \quad (30)$$

Table 1 shows an alternative way to generate random variables for the special cases defined in Corollaries 5–8. We can see that the extended normal distribution is the basis for the development of the specific cases discussed previously.

4. Inferential Aspects of the EL Distribution

In this section, we will study some inferential properties of the extended Laplace distribution defined in Corollary 8. We will explore maximum likelihood estimators and Monte Carlo simulation and will apply these results to two real data sets, comparing the fit with the Laplace distribution using the likelihood ratio and the Akaike Information Criterion (AIC).

4.1. Maximum Likelihood Estimator. In practice, it is common to work with a location and scale transformation $X = \mu + \sigma Y$, where $\mu \in \mathbb{R}$, $\sigma > 0$, and $Y \sim \text{EL}(\xi)$ with $\xi \geq 0$. Hence, the density for the random variable X , denoted as $X \sim \text{EL}(\mu, \sigma, \xi)$, is

$$f(x; \mu, \sigma, \xi) = \left(\frac{\sigma + \xi |x - \mu|}{2\sigma^2(1 + \xi)} \right) e^{-|x - \mu|/\sigma}, \quad x \in \mathbb{R}. \quad (31)$$

Assume that $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$ is a random sample of size n from an $\text{EL}(\mu, \sigma, \xi)$ distribution. From (31), the log-likelihood function is

$$\begin{aligned} \ell(\boldsymbol{\theta}) &\propto -2n \log(\sigma) - n \log(1 + \xi) \\ &\quad + \sum_{i=1}^n \log(\sigma + \xi |x_i - \mu|) - \frac{1}{\sigma} \sum_{i=1}^n |x_i - \mu|, \end{aligned} \quad (32)$$

where $\boldsymbol{\theta} = (\mu, \sigma, \xi)^\top$, which is a continuous function in each parameter, but it is not differentiable at $x_i = \mu$, for $i = 1, \dots, n$. Thus, by assuming $x_i \neq \mu$, for $i = 1, \dots, n$, we have that elements of the score vector are $\mathbf{S}(\boldsymbol{\theta}) = (S_\mu, S_\sigma, S_\xi)^\top$, where $S_a = \partial \ell / \partial a$, given by

$$\begin{aligned} S_\mu &= -\xi \sum_{i=1}^n \frac{\text{sgn}(x_i - \mu)}{\sigma + \xi |x_i - \mu|} + \frac{1}{\sigma} \sum_{i=1}^n \text{sgn}(x_i - \mu), \\ S_\sigma &= -\frac{2n}{\sigma} + \frac{1}{\sigma^2} \sum_{i=1}^n |x_i - \mu| + \sum_{i=1}^n \frac{1}{\sigma + \xi |x_i - \mu|}, \\ S_\xi &= -\frac{n}{1 + \xi} + \sum_{i=1}^n \frac{|x_i - \mu|}{\sigma + \xi |x_i - \mu|}, \end{aligned} \quad (33)$$

where $\text{sgn}(\cdot)$ denotes the sign function.

Hence, the maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ solves the score equations $\mathbf{S}(\boldsymbol{\theta}) = \mathbf{0}$. Which must be obtained through a numerical method. A lot of software, including optimization

toolbox, can be used for obtaining the maximum likelihood estimates. To achieve the maximization of log-likelihood function, we used the function `optim` on R (see R Core Team [17]), the specific method being `Nelder-Mead` (see Nelder and Mead [18]), that uses only function values and is robust but relatively slow. It will work reasonably well for nondifferentiable functions.

For obtaining the standard errors of the maximum likelihood estimates one should compute the information matrix $\mathbf{F}(\boldsymbol{\theta})$. It is well known that the elements of $\mathbf{F}(\boldsymbol{\theta})$ are given by

$$\begin{aligned} \mathbf{F}(\boldsymbol{\theta}) &= -E \left[\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j} \right], \\ i, j &= 1, 2, 3 \text{ and } \boldsymbol{\theta} = (\mu, \sigma, \xi)^\top. \end{aligned} \quad (34)$$

Since expectation over EL distribution is not straightforward, numerical methods should be performed to obtain the explicit form of the information matrix. This matrix can be approximated by the observed information matrix $\mathbf{I}(\hat{\boldsymbol{\theta}})$, which is defined as minus the Hessian matrix evaluated at $\hat{\boldsymbol{\theta}}$; that is

$$\mathbf{I}(\hat{\boldsymbol{\theta}}) \approx - \left(\frac{\partial^2 \ell}{\partial \theta_i \partial \theta_j} \right)_{i,j}, \quad (35)$$

where the second derivatives are given below:

$$\begin{aligned} \frac{\partial^2 \ell}{\partial \mu^2} &= -\xi^2 \sum_{i=1}^n \frac{1}{(\sigma + \xi |x_i - \mu|)^2}, \\ \frac{\partial^2 \ell}{\partial \mu \partial \sigma} &= -\frac{1}{\sigma^2} \sum_{i=1}^n \text{sgn}(x_i - \mu) + \xi \sum_{i=1}^n \frac{\text{sgn}(x_i - \mu)}{(\sigma + \xi |x_i - \mu|)^2}, \\ \frac{\partial^2 \ell}{\partial \mu \partial \xi} &= -\sigma \sum_{i=1}^n \frac{\text{sgn}(x_i - \mu)}{(\sigma + \xi |x_i - \mu|)^2}, \\ \frac{\partial^2 \ell}{\partial \sigma^2} &= \frac{2n}{\sigma^2} - \sum_{i=1}^n \frac{1}{(\sigma + \xi |x_i - \mu|)^2} - \frac{2}{\sigma^3} \sum_{i=1}^n |x_i - \mu|, \\ \frac{\partial^2 \ell}{\partial \sigma \partial \xi} &= -\sum_{i=1}^n \frac{|x_i - \mu|}{(\sigma + \xi |x_i - \mu|)^2}, \\ \frac{\partial^2 \ell}{\partial \xi^2} &= \frac{n}{(1 + \xi)^2} - \sum_{i=1}^n \frac{(x_i - \mu)^2}{(\sigma + \xi |x_i - \mu|)^2}. \end{aligned} \quad (36)$$

Thus, we use the observed information matrix for computing the standard errors in the rest of the paper. Note that this approximation of the observed information matrix is obtained under a less stringent supposition, this is, assuming that the density function is absolutely continuous, as is the case with the Laplace distribution (see Kotz et al. [19], remark 2.6.1).

4.2. Numerical Study. We shall use Monte Carlo simulation to evaluate the finite sample performance of the maximum

TABLE 2: Empirical means and standard error for maximum likelihood estimators.

n	Parameters			Estimations		
	μ	σ	ξ	$\hat{\mu}$ (SE)	$\hat{\sigma}$ (SE)	$\hat{\xi}$ (SE)
50	0	1	0.5	0.0204 (0.4820)	1.0389 (0.3676)	0.5309 (1.0881)
			1	0.0242 (0.3654)	1.0280 (0.2816)	1.0336 (1.5642)
			3	0.0427 (0.1831)	0.9960 (0.1650)	3.1850 (3.7918)
	5	2	0.5	5.1040 (0.8495)	2.0082 (0.6631)	0.6770 (1.1768)
			1	5.1001 (0.6149)	2.0170 (0.5393)	1.2138 (1.7345)
			3	5.0971 (0.3587)	1.9945 (0.3280)	3.2898 (3.9105)
100	0	1	0.5	0.0204 (0.2839)	1.0221 (0.2745)	0.5318 (0.8100)
			1	0.0252 (0.2173)	1.0146 (0.1887)	1.0345 (1.0549)
			3	0.0248 (0.1206)	0.9968 (0.1203)	3.1760 (2.8356)
	5	2	0.5	5.0733 (0.5707)	1.9895 (0.4675)	0.6398 (0.7921)
			1	5.0679 (0.3964)	1.9881 (0.3951)	1.1650 (1.2693)
			3	5.0555 (0.2397)	1.9874 (0.2444)	3.2888 (2.9978)
250	0	1	0.5	0.0145 (0.1416)	1.0051 (0.1676)	0.5347 (0.5004)
			1	0.0213 (0.1290)	1.0011 (0.1171)	1.0335 (0.6633)
			3	0.0065 (0.0699)	0.9963 (0.0732)	3.1530 (1.6987)
	5	2	0.5	5.0293 (0.2936)	1.9586 (0.2806)	0.6131 (0.4630)
			1	5.0441 (0.2255)	1.9774 (0.2034)	1.1208 (0.6059)
			3	5.0255 (0.1387)	1.9922 (0.1469)	3.2357 (1.7630)
500	0	1	0.5	0.0107 (0.0806)	0.9960 (0.1048)	0.5341 (0.3108)
			1	0.0149 (0.0856)	0.9992 (0.0721)	1.0362 (0.3982)
			3	0.0067 (0.0488)	0.9944 (0.0497)	3.1315 (1.1299)
	5	2	0.5	5.0058 (0.1794)	1.9592 (0.1822)	0.5861 (0.2886)
			1	5.0205 (0.1587)	1.9802 (0.1348)	1.0922 (0.3904)
			3	5.0165 (0.0954)	1.9908 (0.0985)	3.2127 (1.1524)

likelihood estimator. The number of Monte Carlo replications was 1000 from simulated samples of the EL distribution for several samples sizes. Each sample was generated using the stochastic representation of the EL distribution, described above. For each generated sample, we obtain the maximum likelihood estimates using the function `optim` on R, the specific method being Nelder-Mead.

In order to analyze the point estimation results, we computed, for each sample size and for each estimator, the standard error from the observed information matrix defined in (35). The result can be seen in Table 2. From the results, we can see that the estimates are quite stable and estimates are asymptotically unbiased as expected, that is, it is observed that the bias becomes smaller as the sample size n increases.

4.3. Data Illustration. In this section we shall examine the application of the EL distribution to two real data sets. The first data set is related to the project WHO MONICA (World Health Organization Multinational Monitoring of Trends and Determinants in Cardiovascular Disease). This data set has been previously analyzed and studied in Kuulasmaa et al. [20], Kulathinal et al. [21], and de Castro et al. [22] and corresponds to the average annual rate of occurrence of cardiovascular mortality or the presence of coronary disease. The data are as follows: -5.6, -3.1, 0.3, -1.9, -1.6, 2.0, -0.9, -3.1, -7.4, -7.2, -4.6, 1.2, -2.3, -3.5, -4.9, -1.6, -3.5, -0.8, -4.1, -2.2, 0.9, 1.8, -6.7, -1.0, -2.3, 0.6, -1.8, 0.8, 2.2, -1.3,

-4.9, -5.4, 0.7, -2.9, -6.1, -0.6, -5.9, 2.3, -5.8, -4.3, -4.7, -0.4, 4.0, -2.9, 4.9, -0.9, -5.4, -8.7, -1.6, -5.9, -3.9, -7.0, -0.1, -5.9, -0.3, 2.4, 1.4, -5.1, 8.4, 4.8, -6.1, -1.1, 0.0, -8.5, -1.7, 7.5, 0.5, -0.1, 0.4, -5.4, -4.0, -3.3, -1.8, 4.3. The second data set consists of the heights in inches of 126 students from University of Pennsylvania. This data set has been previously analyzed and studied by Hassan and Hijazi [23] and Gui et al. [4]. The data are as follows: 55.00, 60.00, 60.25, 61.00, 61.75, 62.25, 62.25, 62.63, 62.75, 63.00, 63.25, 63.25, 63.25, 63.25, 63.38, 64.00, 64.25, 64.25, 64.50, 64.75, 64.75, 65.00, 65.00, 65.13, 65.13, 65.17, 65.25, 65.25, 65.25, 65.25, 65.25, 65.50, 65.75, 66.00, 66.00, 66.25, 66.25, 66.25, 66.50, 66.75, 66.75, 66.75, 67.00, 67.13, 67.25, 67.38, 67.50, 67.50, 67.75, 67.75, 67.75, 68.13, 68.75, 69.00, 69.00, 69.25, 69.50, 69.88, 70.00, 70.00, 70.25, 70.38, 71.00, 71.00, 71.25, 71.75, 65.50, 65.75, 66.00, 66.25, 66.75, 67.00, 67.13, 67.25, 68.00, 68.13, 68.75, 68.75, 69.13, 69.25, 69.25, 69.38, 69.75, 70.00, 70.13, 70.50, 70.75, 70.75, 71.00, 71.00, 71.50, 71.50, 71.63, 71.75, 71.75, 72.00, 72.00, 72.00, 72.00, 72.00, 72.00, 72.25, 72.25, 72.25, 72.38, 72.50, 72.50, 72.88, 73.00, 73.00, 73.00, 73.25, 73.25, 73.63, 74.00, 74.63, 75.00, 75.00, 75.05, 76.00, 77.00, 77.13, 77.50, 77.50, 78.50. Table 3 shows a descriptive summary of the data sets analyzed.

From these data sets, the maximum likelihood estimators of the parameters associated with the Laplace ($L(\mu, \sigma)$) and the extended Laplace ($EL(\mu, \sigma, \xi)$) distributions are obtained; the results of the comparison are summarized in Table 4 and

TABLE 3: Descriptive summary.

Data set	n	\bar{x}	s_x	Asymmetry	Kurtosis
WHO MONICA	74	-1.9014	3.6037	0.4724	3.0893
HEIGHT	126	68.5462	4.1578	-0.0497	3.0528

TABLE 4: Maximum likelihood estimation and standard error.

Parameter estimates	WHO MONICA		HEIGHT	
	L	EL	L	EL
μ	-1.7000 (0.0580)	-2.1310 (0.1889)	68.7500 (0.0353)	68.4236 (0.1469)
σ	3.3319 (0.4563)	1.6204 (0.1876)	3.7162 (0.3628)	1.7988 (0.1308)
ξ	-	3.5025 (2.9071)	-	7.9689 (5.1001)
Log-likelihood	-204.0124	-199.5330	-368.2153	-354.6919
AIC	412.0249	405.0661	740.4306	715.3838

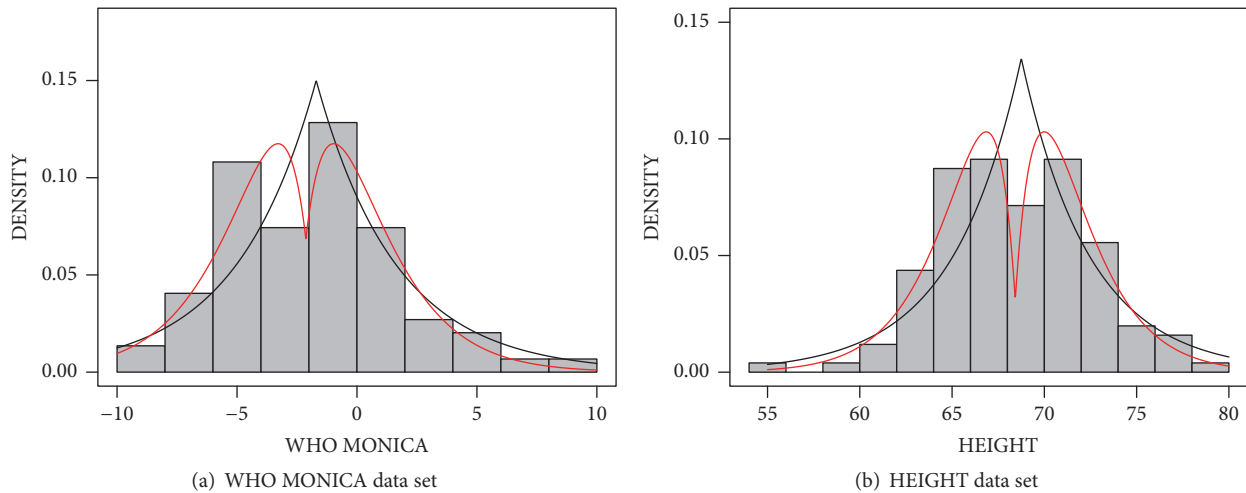


FIGURE 2: Histograms and the fitted curves. Extended Laplace distribution (red line) and Laplace distribution (black line).

Figure 2 shows the histograms and the fitted curves for the data set.

The AIC is used to compare the estimated models. As one can see, our model with the smallest values of AIC is preferable. In addition, we can use the likelihood ratio (LR) test statistic to confirm our claim. To do this, we consider the following hypotheses:

$$H_0 : \xi = 0 \text{ (L}(\mu, \sigma)) \text{ vs. } H_1 : \xi \neq 0 \text{ (EL}(\mu, \sigma, \xi)) \quad (37)$$

The value of the LR test statistic for the data WHO MONICA is 7.22 and for the data set of the height of students is 26.96 and comparing this quantity with $\chi_{(1)}^2 = 3.83$, the null hypothesis is rejected.

Figure 3 shows the graphs of the QQ-plots for the WHO MONICA and HEIGHT data sets calculated with the Laplace and extended Laplace models fitted with the maximum likelihood estimates of the parameters. These also show the good agreement of the EL distribution for the two data sets.

5. Conclusion and Final Comments

We have presented a generalization of the extended bimodal-normal distribution that depends on a shape parameter that controls the effect of bimodality when the density is symmetric. But, generally speaking, it produces a more flexible model in terms of asymmetry and kurtosis coefficients. Additionally, we have demonstrated its basic properties and stochastic representation, the latter of which played a significant role in the development of this work. The family of distributions includes a large number of distributions. For example, four of them were presented as corollaries, leaving the Laplace distribution for last, and used to develop some inferential aspects and Monte Carlo simulation schemes, which were facilitated by the use of stochastic representation in the generation of random variables. Finally, using two real data sets, we demonstrated that the proposed model resulted in better behavior relative to the standard Laplace model. Moreover, in the statistical literature, there are a variety of extensions of the Laplace distribution, in order to achieve greater flexibility, but without the effect of bimodality that

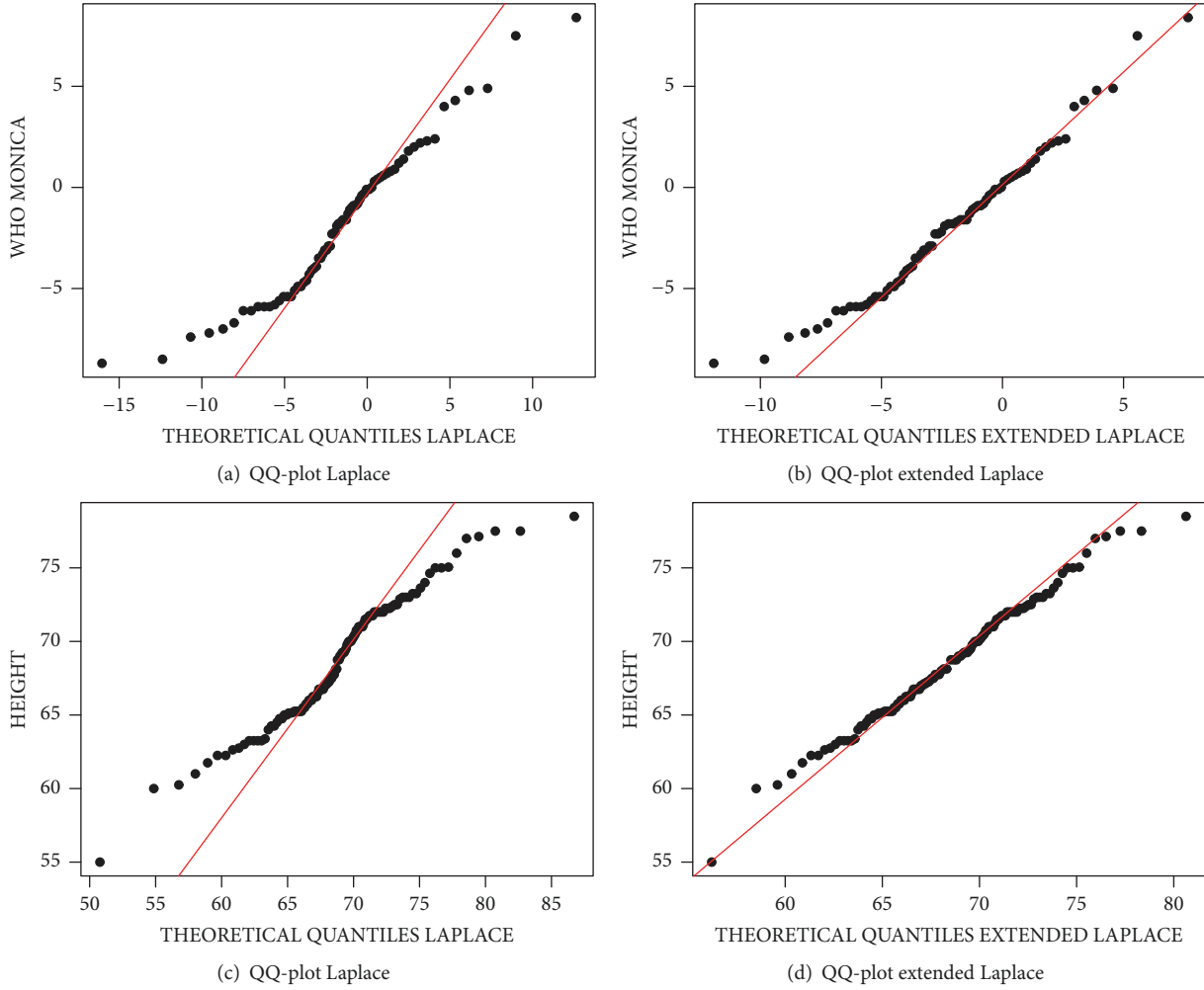


FIGURE 3: QQ-plot for the WHO MONICA and HEIGHT data sets.

fitted the data analyzed. Although bimodality can be achieved through a mixture of distributions, the proposed model is more parsimonious in terms of the number of parameters. It is important to emphasize that Theorem 1 can be extended, as demonstrated below.

Theorem 15. Let g_1 and g_2 be a probability density functions and h a positive continuous function such that $E_{g_2}[h(Y)] = \kappa < \infty$, where $Y \sim g_2$. Then

$$f(x; \xi) = \left(\frac{1}{1 + \kappa\xi} \right) g_1(x) + \left(\frac{\xi h(x)}{1 + \kappa\xi} \right) g_2(x), \quad (38)$$

is a probability density function with shape parameter $\xi \geq 0$.

Note that when $g_1(x) = g_2(x) = g(x)$ we have Theorem 1. Furthermore, this new extension has stochastic representation given by the following theorem.

Theorem 16. Let $X_1 \sim g_1(x)$ and $X_2 \sim (1/\kappa)h(x)g_2(x)$ be independent random variables. If

$$X = \begin{cases} X_1, & \text{with probability } \frac{1}{1 + \kappa\xi}, \\ X_2, & \text{with probability } \frac{\kappa\xi}{1 + \kappa\xi}, \end{cases} \quad (39)$$

then $X \sim f(x; \xi)$.

Proof. Since $f(x; \xi)$ is a mixture, the result follows immediately. \square

This new extension includes, for examples, the case of the slash distribution Rogers and Tukey [24], Mosteller and Tukey [25], and Kafadar [26], if we consider the following definitions: (i) $g_1(x) = q \int_0^1 u^q \phi(xu) du$, (ii) $g_2 = (q + 2) \int_0^1 u^{q+2} \phi(xu) du$, and (iii) $h(x) = qx^2/(q + 2)$, with $x \in \mathbb{R}$ and $q > 2$, with density given by

$$f(x; \xi) = \left(\frac{1}{1+\xi} \right) q \int_0^1 u^q \phi(xu) du + \left(\frac{\xi}{1+\xi} \right) qx^2 \int_0^1 u^{2+q} \phi(xu) du. \quad (40)$$

Data Availability

The WHO MONICA and HEIGHT data sets used to support the findings of this study are included within the article.

Disclosure

Preliminary results of this manuscript were presented as a poster in “Flexible Statistical Models for a Skewed World of Data: A Workshop in Honor of Reinaldo B. Arellano-Valle’s 65th Birthday” 2017.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Funding

The research of M. A. Cortés was supported in part by Grant DIUDA 221230. The research of D. Elal-Olivero was supported in part by Grant DIUDA 22287. The research of J. F. Olivares-Pacheco was supported in part by Grant DIUDA 22257.

References

- [1] A. Azzalini, “A class of distributions which includes the normal ones,” *Scandinavian Journal of Statistics*, vol. 12, no. 2, pp. 171–178, 1985.
- [2] D. Elal-Olivero, “Alpha-skew-normal distribution,” *Proyecciones Journal of Mathematics*, vol. 29, no. 3, pp. 224–240, 2010.
- [3] D. Elal-Olivero, H. W. Gómez, and F. A. Quintana, “Bayesian modeling using a class of bimodal skew-elliptical distributions,” *Journal of Statistical Planning and Inference*, vol. 139, no. 4, pp. 1484–1492, 2009.
- [4] W. Gui, P.-H. Chen, and H. Wu, “A symmetric component alpha normal slash distribution: properties and inferences,” *Journal of Statistical Theory and Applications*, vol. 12, no. 1, pp. 55–66, 2013.
- [5] H. W. Gómez and G. A. Guerrero, *Bimodal Birnbaum-Saunders Distribution*, COMCA 2014: Congreso de Matemática Capricornio, Copiapó, Chile, 2014.
- [6] N. Henze, “A probabilistic representation of the ‘skew-normal’ distribution,” *Scandinavian Journal of Statistics*, vol. 13, no. 4, pp. 271–275, 1986.
- [7] Z. W. Birnbaum and S. C. Saunders, “A new family of life distributions,” *Journal of Applied Probability*, vol. 6, no. 2, pp. 319–327, 1969.
- [8] Z. W. Birnbaum and S. C. Saunders, “Estimation for a family of life distribution with applications to fatigue,” *Journal of Applied Probability*, vol. 6, no. 2, pp. 328–347, 1969.
- [9] N. R. Mann, R. E. Schafer, and N. D. Singpurwalla, *Methods for Statistical Analysis of Reliability and Lifetime Data*, John Wiley & Sons, New York, NY, USA, 1974.
- [10] A. Desmond, “Stochastic models of failure in random environments,” *Canadian Journal of Statistics*, vol. 13, no. 2, pp. 171–183, 1985.
- [11] A. F. Desmond, “On the Relationship Between Two Fatigue-Life Models,” *IEEE Transactions on Reliability*, vol. 35, no. 2, pp. 167–169, 1986.
- [12] D. S. Chang and L. C. Tang, “Graphical analysis for Birnbaum-Saunders distribution,” *Microelectronics Reliability*, vol. 34, no. 1, pp. 17–22, 1994.
- [13] J. A. Díaz-García and V. Leiva-Sánchez, “A new family of life distributions based on the elliptically contoured distributions,” *Journal of Statistical Planning and Inference*, vol. 128, no. 2, pp. 445–457, 2005.
- [14] H. W. Gómez, J. F. Olivares-Pacheco, and H. Bolfarine, “An extension of the generalized Birnbaum-Saunders distribution,” *Statistics & Probability Letters*, vol. 79, no. 3, pp. 331–338, 2009.
- [15] N. M. Olmos, G. Martínez-Flórez, and H. Bolfarine, “Bimodal Birnbaum-Saunders distribution with applications to non negative measurements,” *Communications in Statistics—Theory and Methods*, vol. 46, no. 13, pp. 6240–6257, 2017.
- [16] S. C. Saunders, “A family of random variables closed under reciprocation,” *Journal of the American Statistical Association*, vol. 69, no. 346, pp. 533–539, 1974.
- [17] R Core Team, “R: A Language and Environment for Statistical Computing,” R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>, 2018.
- [18] J. A. Nelder and R. Mead, “A simplex algorithm for function minimization,” *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.
- [19] S. Kotz, T. J. Kozubowski, and K. Podgórski, *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economic, Engineering and Finance*, Birkhäuser-Springer, Boston, Mass, USA, 2001.
- [20] K. Kuulasmaa, H. Tunstall-Pedoe, A. Dobson et al., “Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations,” *The Lancet*, vol. 355, no. 9205, pp. 675–687, 2000.
- [21] S. B. Kulathinal, K. Kuulasmaa, and D. Gasbarra, “Estimation of an errors-in-variables regression model when the variances of the measurement errors vary between the observations,” *Statistics in Medicine*, vol. 21, no. 8, pp. 1089–1101, 2002.
- [22] M. de Castro, M. Galea, and H. Bolfarine, “Hypothesis testing in an errors-in-variables model with heteroscedastic measurement errors,” *Statistics in Medicine*, vol. 27, no. 25, pp. 5217–5234, 2008.
- [23] M. Y. Hassan and R. H. Hijazi, “A bimodal exponential power distribution,” *Pakistan Journal of Statistics*, vol. 26, no. 2, pp. 379–396, 2010.
- [24] W. H. Rogers and J. W. Tukey, “Understanding some long-tailed symmetrical distributions,” *Statistica Neerlandica. Journal of the Netherlands Society for Statistics and Operations Research*, vol. 26, no. 3, pp. 211–226, 1972.
- [25] F. Mosteller and J. Tukey, *Data Analysis and Regression: A Second Course in Statistics*, Addison-Wesley Pub. Co., 1977.
- [26] K. Kafadar, “A biweight approach to the one-sample problem,” *Journal of the American Statistical Association*, vol. 77, no. 378, pp. 416–424, 1982.

