Revista Colombiana de Estadística - Theorical Statistics
July 2022, volume 45, issue 2, pp. 301–326
http://doi.org/10.15446/rce.v45n2.95530

Likelihood-Based Inference for the Asymmetric Exponentiated Bimodal Normal Model

Inferencia basada en verosimilitud para el modelo asimétrico bimodal normal exponenciado

Guillermo Martínez-Flórez[1,a], Mario Pacheco-López[2,b], Roger Tovar-Falón[1,c]

[1]Departamento de Matemáticas y Estadística, Facultad de Ciencias Básicas, Universidad de Córdoba, Montería, Colombia

[2]Facultad de Estadítica, Universidad Santo Tomás, Bogotá, Colombia

Abstract

Asymmetric probability distributions have been widely studied by various authors in recent decades. Special interest has been had families of flexible distributions with the capability to have into account degree of skewness and kurtosis greater than the classical distributions widely known in statistical theory. While, most of the new distributions fit unimodal data, and a few fit bimodal data, in the bimodal proposals, singularity problems have been found in the information matrices. Therefore, in this paper, extensions of the alpha-power family of distributions are developed, which have non-singular information matrix. The new proposals are based on the bimodal-normal and bimodal elliptical skew-normal distributions. These new extensions allow modeling asymmetric bimodal data, which are commonly found in several areas of scientific interest. The properties of these new distributions of probability are also studied in detail, and the statistical inference process is carried out to estimate the parameters of the proposed models. The stochastic convergence for the maximum likelihood estimator (MLE) vector can be found due to the non-singularity of the expected information matrix in the corresponding support. We also introduced extensions of the asymmetric bimodal normal and bimodal elliptical skew-normal models for the situations in which the data present censorship. A small simulation study to evaluate the properties of the MLE is also presented and, finally, two applications to real data set are presented for illustrative purposes.

***Key words*:** Alpha-Power distribution; Asymmetric models; Bimodal normal distribution; Censored data; Maximum likelihood estimation.

[a]Ph.D. E-mail: guillermomartinez@correo.unicordoba.edu.co
[b]Ph.D. E-mail: mariopacheco@usantotomas.edu.co
[c]Ph.D. E-mail: rjtovar@correo.unicordoba.edu.co

Resumen

Las distribuciones de probabilidad asimétricas han sido ampliamente estudiadas por diversos autores en las últimas décadas. Se ha tenido especial interés en familias de distribuciones flexibles con la capacidad de tener en cuenta grados de asimetría y curtosis mayores que las distribuciones clásicas ampliamente conocidas en teoría estadística. Si bien la mayoría de las nuevas distribuciones se ajustan a datos unimodales y unas pocas a datos bimodales, en las propuestas bimodales se han encontrado problemas de singularidad en las matrices de información. Por lo tanto, en este artículo se desarrollan extensiones de la familia de distribuciones alfa-potencia, que tienen matriz de información no singular. Las nuevas propuestas se basan en las distribuciones bimodal-normal y bimodal elíptica sesgada-normal. Estas nuevas extensiones permiten modelar datos bimodales asimétricos, que se encuentran comúnmente en varias áreas de interés científico. También se estudian en detalle las propiedades de estas nuevas distribuciones de probabilidad, y se realiza el proceso de inferencia estadística para estimar los parámetros de los modelos propuestos. La convergencia estocástica para el vector estimador de máxima verosimilitud (EMV) se puede encontrar debido a la no singularidad de la matriz de información esperada en el soporte correspondiente. También introdujimos extensiones de los modelos asimétrico bimodal normal y bimodal elíptico sesgado-normal para las situaciones en las que los datos presentan censura. También se presenta un pequeño estudio de simulación para evaluar las propiedades del EMV y, finalmente, se presentan dos aplicaciones a conjuntos de datos reales con fines ilustrativos.

***Palabras clave*:** Distribución normal bimodal; Distribución alfa-potencia; Datos censurados; Modelos asimétricos; Estimación por máxima verosimilitud.

1. Introduction

The normal distribution has been used to model a wide variety of variables of interest because it provides the basis for the classical statistical inference due to its relationship with the central limit theorem. In applications of the scientific area such as biology, economics and medicine, among others, it is common to suppose normality in the data set, however, this assumption many times is deviated from the reality, leading to errors in the estimation and inference process. One of these cases occurs when asymmetric or heavy-tailed data are being analyzed, and we reparameterize or transform the variables, which can lead to difficulties in the interpretation of the results.

The most recognized distributions to analyze data with high degree of asymmetry corresponds to the skew-normal (SN) proposed by Azzalini (1985). The SN distribution is an extension of the normal distribution for modeling the asymmetric structures present in the data; however, this model has a difficulty, since its information matrix is singular when the asymmetry parameter is close to zero. The probability density function (PDF) of the SN model is given by

$$f_{\mathrm{SN}}(z;\lambda) = 2\phi(z)\Phi(\lambda z), \qquad z \in \mathbb{R}, \tag{1}$$

where ϕ and Φ are the PDF and the cumulative distribution function (CDF), respectively, of the standard normal distribution; and λ is a parameter that controls the asymmetry of the data. The SN model is denoted by $Z \sim \mathrm{SN}(\lambda)$ and has been extensively studied by Azzalini (1985), Azzalini (1986), Henze (1986), Pewsey (2000), among others.

In the same sense of modeling asymmetric data, Durrans (1992) introduced the distribution of fractional order statistics, also called exponentiated model by Gupta & Gupta (2004), and power distribution by Pewsey et al. (2012). This model generates a good base for the creation of distributions that can fit asymmetric data, and it has served as fundamental support for the solution of the problem of the singularity presented by the information matrix of the SN model, since the information matrix of the exponentiated distribution for the normal case, which is called power-normal (PN) distribution, is non-singular for values close to one in the skewness parameter. The PN distribution denoted by $Z \sim \mathrm{PN}(\alpha)$, has the PDF given by

$$f_{\mathrm{PN}}(z;\alpha) = \alpha\phi(z)\{\Phi(z)\}^{\alpha-1}, \quad z \in \mathbb{R}, \quad \alpha \in \mathbb{R}^{+}, \tag{2}$$

where $\alpha > 0$ is a shape parameter. The location-scale extension of the random variable Z, that is, $X = \xi + \eta Z$, where $\xi \in \mathbb{R}$ is a location parameter, and $\eta \in R^{+}$ is a scale parameter, has the PDF given by

$$f_{\mathrm{PN}}(x;\xi,\eta,\alpha) = \frac{\alpha}{\eta}\phi\left(\frac{x-\xi}{\eta}\right)\left\{\Phi\left(\frac{x-\xi}{\eta}\right)\right\}^{\alpha-1}, \quad x \in \mathbb{R}, \tag{3}$$

and we denoted it as $X \sim \mathrm{PN}(\xi,\eta,\alpha)$. The properties of the PN model have been studied in detail by Gupta & Gupta (2008) and Pewsey et al. (2012).

An extension of the PN model for the case of the SN model was introduced by Martínez-Flórez et al. (2014), where the PDF and the CDF of the normal distribution were replaced by the respective functions of the SN model. This proposal turns out more flexible regarding skewness and kurtosis than the normal, SN and PN models.

Although the SN and PN models fit data presenting high (or low) skewness or kurtosis, they only fit data sets whose shape of the density is unimodal, not being efficient concerning bimodal data. To fit data sets that present bimodality, the mixture of normal distributions is frequently used; however, this option is not always the best since there are some problems in the estimation of the parameters, and therefore, the mixture of distributions becomes a controversial topic, see Marin et al. (2005), mainly because one has to deal with identifiability issues.

Proposals for modeling data sets with bimodal shape have been studied by many authors; Elal-Olivero (2010) for example, introduced the bimodal-normal (BN) model which provides an excellent tool for studying variables with two maximums in the support, even so, this model only fits symmetric bimodal structures. For the case of skewed data, Elal-Olivero, Gómez & Quintana (2009) proposed the bimodal version of the SN distribution and studied its properties; while Gómez et al. (2009) considered a class of flexible bimodal SN model. Other type of symmetric bimodal SN model was studied by Kim (2005), whereas, Arnold et al. (2009) extended the Kim's model to the SN asymmetric bimodal case. Additional works have been published by Elal-Olivero, Olivares-Pacheco, Gómez & Bolfarine (2009), who studied a new class of distributions for non-negative data; Bolfarine et al. (2018) introduced a bimodal extension of the PN model, and Martínez-Flórez et al. (2020) proposed a type of distribution capable to fit data with up to three modes.

Unfortunately, the asymmetric bimodal models defined from the SN distribution inherit the issue of singularity of the information matrix (even in the positive data case); conversely case to the bimodal PN model whose information matrix turned out non-singular. Taking advantage of the non-singularity of the information matrix of the alpha-power (AP) model and the flexibility to fit bimodal data from the BN and elliptical bimodal normal (ELBN) models; it is useful to study the behavior of the resulting distribution from the combination of the generic AP structure and the BN and ELBN distributions.

This paper is organized as follows: in Section 2 the BN distribution of Elal-Olivero (2010) and some of its main properties are featured. In Section 3, the exponentiated bimodal normal (EBN) model is introduced and its properties are studied in detail. The location-scale extension is considered, too. The maximum likelihood (ML) method to obtain the estimates of the parameter model is used and the information matrix is calculated, which is shown to be non-singular. In Section 4, the asymmetric exponentiated elliptical bimodal normal (EEBN) model is introduced. This proposal is an extension of the skew-elliptical bimodal family considered by Elal-Olivero, Gómez & Quintana (2009). The properties and inference process for the models is also presented. The version of the EBN and EEBN models for the situations of censored data are also considered here. The Section 5 presents a small Monte Carlo simulation and, finally, in Section 6 two real data applications are presented.

2. Bimodal Models

The SN (Azzalini, 1985) and generalized gaussian (GG) (Durrans, 1992) models are characterized by being models for fitting unimodal data, that is, they must not be used in situations where the data present bimodality, as occurs in some areas such as economics, health, engineering, among others. The models of bimodal type have been studied by Kim (2005), who introduced the bimodal extension of the SN model, called "two-pieces skew-normal model (TN)", whose PDF is given by

$$f_{\mathrm{TN}}(u;\lambda) = c_\lambda \phi(u)\Phi\big(\lambda|u|\big), \quad u \in \mathbb{R}, \tag{4}$$

where λ is a real number, and $c_\lambda = 2\pi/(\pi + 2\arctan(\lambda))$ is a normalizing constant. We use the notation $\mathrm{TN}(\lambda)$. For $\lambda > 0$, Kim (2005) shown that the TN model is bimodal and symmetric around zero.

An asymmetric extension of the Kim's model was presented by Arnold et al. (2009), which is called "*the extended two-pieces skew-normal model* (ETN)", and whose PDF is given by

$$f_{\mathrm{ETN}}(u;\theta) = 2c_\lambda \phi(u)\Phi(\lambda|u|)\Phi(\beta u), \quad u \in \mathbb{R}, \tag{5}$$

where β and λ are real numbers and, c_λ is a normalizing constant. This model is denoted by $\mathrm{ETN}(\lambda, \beta)$. Gómez et al. (2009) defined an asymmetric bimodal model named flexible skew-normal distribution, which is denoted by $\mathrm{FSN}(\alpha, \delta)$, and has PDF is given by

$$f_{\mathrm{FSN}}(u;\theta) = c_\delta \phi(|u| + \delta)\Phi(\alpha u), \quad u \in \mathbb{R}, \tag{6}$$

where δ is a real number and $c_\delta = (1 - \Phi(\delta))^{-1}$ is a normalizing constant. For $\delta < 0$, Gómez et al. (2009) shown that the FSN model is bimodal and, for $\delta \geq 0$ the model is unimodal. Elal-Olivero (2010) studied the asymmetric bimodal alpha-skew-normal (ASN) model with PDF given by

$$f_{\mathrm{ASN}}(x;\alpha) = \frac{(1-\alpha x)^2 + 1}{2+\alpha}\phi(x), \quad x \in \mathbb{R}, \tag{7}$$

with $\alpha \in \mathbb{R}$. The model in equation (7) is denoted by $\mathrm{ASN}(\alpha)$, and the location-scale extension by $\mathrm{ASN}(\xi, \eta, \alpha)$, where $\xi \in \mathbb{R}$ is a location parameter and $\eta > 0$ is a scale parameter. The properties of the model (7) as well as the study of the estimation process of its parameters can be seen in Elal-Olivero (2010). On the other hand, Elal-Olivero, Olivares-Pacheco, Gómez & Bolfarine (2009) presented a type of general distribution which is possible to obtain a bimodal density, the PDF of this distribution is given by

$$f(z;\gamma) = \frac{1 + \gamma z^2}{1 + \gamma k} f_0(z), \quad z \in \mathbb{R}, \tag{8}$$

where $f_0(\cdot)$ is a symmetric unimodal PDF around zero, k is such that $k = \int_{-\infty}^{\infty} z^2 f_0(z) < \infty$ and, $\gamma \in \mathbb{R}^+ \cup \{0\}$ is a parameter that controls the bimodality of the model. For the case $f_0(z) = \phi(z)$, the PDF of the standard normal distribution, it has

$$f_{\mathrm{ELBN}}(z;\gamma) = \frac{1 + \gamma z^2}{1 + \gamma}\phi(z), \quad z \in \mathbb{R}, \tag{9}$$

which is called the elliptical bimodal normal distribution and, is denoted by $\mathrm{ELBN}(\gamma)$. The ELBN model is a symmetric bimodal model and, for $\gamma = 0$, the distribution reduces to the PDF of the normal distribution.

2.1. Bimodal Normal Model

A random variable Z is said to have a bimodal normal distribution (Elal-Olivero, 2010), if its PDF is given by

$$f_{\mathrm{BN}}(z) = z^2 \phi(z), \quad z \in \mathbb{R}. \tag{10}$$

This is denoted by $Z \sim$ BN. The CDF and survival function of the BN distribution are given, respectively by

$$\mathcal{F}_{\mathrm{BN}}(z) = \Phi(z) - z\phi(z), \quad \text{for } z \in \mathbb{R}, \tag{11}$$

and

$$\mathcal{S}_{\mathrm{BN}}(t) = 2\big(1 - \Phi(t) + t\phi(t)\big) = 2\big(\mathcal{S}_{\mathrm{N}}(t) + t\phi(t)\big), \quad \text{for } t > 0, \tag{12}$$

where $\Phi(\cdot)$, $\phi(\cdot)$ and $\mathcal{S}_{\mathrm{N}}(\cdot)$ are de PDF, CDF and survival function of the standard normal distribution, respectively. The Hazard function is given by

$$h_{\mathrm{BN}}(t) = \frac{t^2\phi(t)}{2\big(1 - \Phi(t) + t\phi(t)\big)}, \quad t > 0.$$

If $Z \sim$ BN, then $\mathbb{E}(Z) = 0$, $\mathbb{E}(Z^2) = 3$, $\mathbb{E}(Z^3) = 0$ and $\mathbb{E}(Z^4) = 15$. Therefore, the coefficients of asymmetry and kurtosis are given by $\beta_1 = 0$ and $\sqrt{\beta_2} = 1$, respectively. Similarly, the moment-generating function (MGF) can be expressed as

$$M_Z(t) = 2\exp\left(\frac{t^2}{2}\right).$$

The location and scale parameters are introduced by considering the random variable Z following a BN distribution. Thus, if $Z \sim$ BN, then the random variable $X = \xi + \eta Z$ is a BN with location parameter ξ and scale parameter η. Its PDF is given by

$$f_{\mathrm{BN}}(x; \xi, \eta) = \frac{1}{\eta}\left(\frac{x-\xi}{\eta}\right)^2 \phi\left(\frac{x-\xi}{\eta}\right), \quad x \in \mathbb{R},$$

which is denoted by $X \sim \mathrm{BN}(\xi, \eta)$. The MGF of the random variable $X \sim \mathrm{BN}(\xi, \eta)$ is given by

$$M_X(t) = \exp\left(\xi t + \frac{1}{2}\eta^2 t^2\right)\left(\eta^2 t^2 + 1\right).$$

2.1.1. Maximum Likelihood Estimation

Let $X_1, X_2, \ldots, X_n$ be a random sample such that, $X_i \sim \mathrm{BN}(\xi, \eta)$ for $i = 1, 2, \ldots, n$. The log-likelihood function for estimating the parameters vector $\boldsymbol{\theta} = (\xi, \eta)^\top$ is given by

$$\ell(\boldsymbol{\theta}; \mathbf{X}) = -n\log(\eta) + 2\sum_{i=1}^{n}\log(z_i) + \sum_{i=1}^{n}\log(\phi(z_i)). \tag{13}$$

The score functions, which are defined as the first partial derivative of the log-likelihood function regarding each of the parameters, are given by

$$\frac{\partial \ell}{\partial \xi} = -\frac{1}{\eta}\left\{2\sum_{i=1}^{n}\frac{1}{z_i} - \sum_{i=1}^{n} z_i\right\} \quad \text{and} \quad \frac{\partial \ell}{\partial \eta} = -\frac{1}{\eta}\left\{3n - \sum_{i=1}^{n} z_i^2\right\}.$$

By equating to zero the score functions $\frac{\partial \ell}{\partial \xi} = 0$ and $\frac{\partial \ell}{\partial \eta} = 0$, the score equations are obtained and then, we have the following system of equations

$$\eta^2 = \frac{1}{3n}\sum_{i=1}^{n}(x_i - \xi)^2 \quad \text{and} \quad 2\eta^2 \sum_{i=1}^{n} \frac{1}{x_i - \xi} = \sum_{i=1}^{n}(x_i - \xi). \tag{14}$$

The maximum likelihood estimates are obtained as the solution of the system of equations in (14), which can be solved by using iterative numerical methods such as Newton-Raphson or quase-Newton. The elements of the observed information matrix are defined as minus the second-order partial derivatives of the log-likelihood function regarding each of the parameters, ξ and η; that we denoted by $\psi_{\xi\xi}$, $\psi_{\xi\eta}$ and $\psi_{\eta\eta}$, can be expressed as

$$\psi_{\xi\xi} = \frac{1}{\eta^2}\left\{2\sum_{i=1}^{n}\frac{1}{z_i^2} + n\right\}, \quad \psi_{\xi\eta} = \frac{2}{\eta^2}\sum_{i=1}^{n} z_i \quad \text{and} \quad \psi_{\eta\eta} = \frac{3}{\eta^2}\left\{\sum_{i=1}^{n} z_i^2 - n\right\}.$$

To get the elements of the expected information matrix, we multiply by n^{-1} the elements of the observed information matrix, and we denoted them by $i_{\xi\xi}$, $i_{\xi\eta}$ and $i_{\eta\eta}$. These elements can be written as

$$i_{\xi\xi} = \frac{3}{\eta^2}, \quad i_{\xi\eta} = 0 \quad \text{and} \quad i_{\eta\eta} = \frac{3}{\eta^2}\mathbb{E}\left(Z_i^2\right) - \frac{3}{\eta^2} = \frac{9}{\eta^2} - \frac{3}{\eta^2} = \frac{6}{\eta^2}.$$

It is easy to see that, the determinant of the information matrix is $\det(\mathbf{I}(\hat{\boldsymbol{\theta}})) = 18/\eta^4 \neq 0$, and therefore, the information matrix is non-singular, guaranteeing the existence of the covariance matrix of the MLEs. The Fisher information matrix is given by

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}) = \mathbf{I}^{-1}(\boldsymbol{\theta}) = \begin{pmatrix} \frac{\eta^2}{3} & 0 \\ 0 & \frac{\eta^2}{6} \end{pmatrix}. \tag{15}$$

The existence of $\mathbf{I}(\boldsymbol{\theta})^{-1}$ also guarantees the asymptotic properties of the MLE $\hat{\boldsymbol{\theta}} = (\hat{\xi}, \hat{\eta})^\top$, so that, the $\hat{\boldsymbol{\theta}}$ follows asymptotic normal distribution when n value increase.

$$(\hat{\xi}, \hat{\eta})^\top \xrightarrow{D} \mathrm{N}_2\big((\xi, \eta)^\top, \mathbf{I}(\boldsymbol{\theta})^{-1}\big),$$

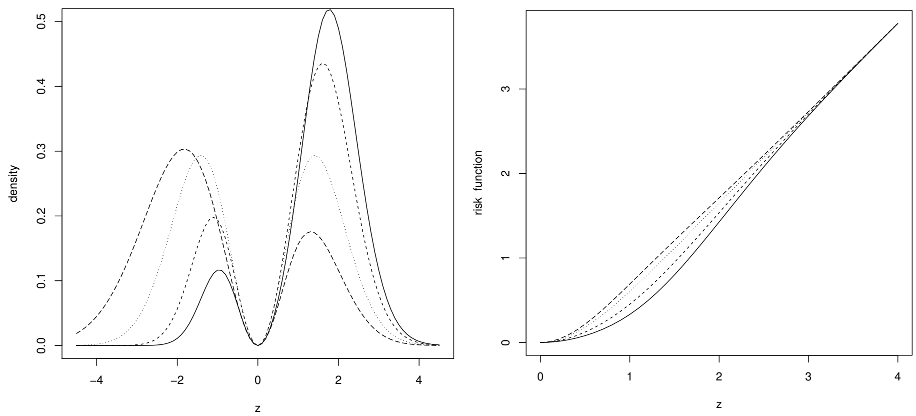
that is, the maximum likelihood estimators of the model parameters are consistent and asymptotically follow a normal distribution with a covariance matrix, the inverse of the Fisher information matrix.

3. Exponentiated Bimodal Normal Model

In this section, we introduce the exponentiated bimodal normal model which is an extension of the BN model to the case of the AP family of distributions and has the PDF given by

$$f_{\mathrm{EBN}}(z;\alpha) = \alpha z^2 \phi(z)\left[\Phi(z) - z\phi(z)\right]^{\alpha-1}, \quad z \in \mathbb{R}, \tag{16}$$

where $\alpha \in \mathbb{R}^+$. We use the notation EBN(α). One can see that, for $\alpha = 1$ the BN model is obtained and the graph is symmetric around zero in this standard case; for other values of α, the model is asymmetric and bimodal. Then, it is concluded that the parameter α explains the asymmetry of the model. Figure 1(a) shows the behavior of the EBN(α) model for some selected values of the α parameter.



(a) (b)

FIGURE 1: PDF and $h_{\mathrm{EBN}}(t)$ for $\alpha = 3$ (solid line), 2 (dashed line), 1 (dotted line) and 0.5 (dotted-dashed line). (a) PDF and (b) $h_{\mathrm{EBN}}(t)$.

As an inherited property of the AP family of distributions, one can note that, the distribution function of this model corresponds to the power of the distribution function of the BN model of Elal-Olivero (2010). Thus, the CDF of the EBN model is given by

$$\mathcal{F}_{\mathrm{EBN}}(z;\alpha) = [\Phi(z) - z\phi(z)]^{\alpha} = [\mathcal{F}_{\mathrm{BN}}(z)]^{\alpha}, \quad z \in \mathbb{R}, \tag{17}$$

where $\mathcal{F}_{\mathrm{BN}}(\cdot)$ is the CDF of the BN model of Elal-Olivero (2010). It follows that the survival and Hazard functions of the EBN model are given by

$$\mathcal{S}_{\mathrm{EBN}}(t) = 1 - [\Phi(t) - t\phi(t)]^{\alpha} \quad \text{and} \quad h_{\mathrm{EBN}}(t) = \alpha t^2 \phi(t) \frac{[\Phi(t) - t\phi(t)]^{\alpha-1}}{1 - [\Phi(t) - t\phi(t)]^{\alpha}},$$

respectively. The behavior of the Hazard function for $t > 0$ can be seen in Figure 1-(b) which is non-decreasing (strictly) and convergent.

3.1. Moments for the Standard Case

The moments of the EBN model do not have a particular shape since they cannot be calculated explicitly, and they can only be calculated numerically. In general, the r-th moment of the random variable Z with EBN distribution is given by

$$\mu_r = \mathbb{E}(Z^r) = \alpha \int_{-\infty}^{\infty} z^{r+2}\phi(z)\,[\Phi(z) - z\phi(z)]^{\alpha-1}\,dz. \tag{18}$$

3.2. Location-Scale Extension

If $Z \sim \text{EBN}(\alpha)$, the location-scale extension of the EBN model is obtained from the transformation $X = \xi + \eta Z$, where $\xi \in \mathbb{R}$ and $\eta \in \mathbb{R}^{+}$. The PDF of X is given by

$$f_{\text{EBN}}(x; \xi, \eta, \alpha) = \frac{\alpha}{\eta} z^2 \phi(z) \left[\Phi(z) - z\phi(z)\right]^{\alpha - 1}$$

where $z = (x - \xi)/\eta$. We denote this extension by $\text{EBN}(\xi, \eta, \alpha)$ and its CDF for $\mathcal{F}_{\text{EBN}}(x; \boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\xi, \eta, \alpha)^{\top}$.

3.3. Moments and Moment Generating Function for Location-scale Model

For the location-scale version of the $\text{EBN}(\xi, \eta, \alpha)$ model, the r-th moment is obtained from the expression

$$\mathbb{E}(X^r) = \sum_{l=0}^{r} \binom{r}{l} \mu^l \sigma^{r-l} \mathbb{E}(Z^{r-l}),$$

where $Z \sim \text{EBN}(\alpha)$. The central moments, $\mu_r' = \mathbb{E}\big(X - \mathbb{E}(X)\big)^r$, for $r = 2, 3, 4$ can be calculated by using the expressions

$$\mu_2' = \mu_2 - \mu_1^2, \quad \mu_3' = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 \quad \text{and} \quad \mu_4' = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4.$$

Therefore, the variance, the coefficient of variation, and the asymmetry and kurtosis coefficients can be calculated as

$$\sigma^2 = \mu_2', \quad \text{CV} = \frac{\sqrt{\sigma^2}}{\mu_1}, \quad \sqrt{\beta_1} = \frac{\mu_3'}{[\mu_2']^{3/2}} \quad \text{and} \quad \beta_2 = \frac{\mu_4'}{[\mu_2']^2},$$

respectively. The MGF of the EBN distribution does not have special shape and is calculated from the general formula

$$M_X(t) = \alpha e^{\xi t + \frac{\eta^2 t^2}{2}} \int_{-\infty}^{\infty} z^2 \phi(z - \eta t) \left[\Phi(z) - z\phi(z)\right]^{\alpha - 1} dz. \tag{19}$$

3.4. Maximum Likelihood Estimation

For a random sample of size n, $\mathbf{X} = (x_1, x_2, \ldots, x_n)^{\top}$, with $X_i \sim \text{EBN}(\xi, \eta, \alpha)$, for $i = 1, \ldots, n$, the log-likelihood function is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}; \mathbf{X}) = n\log(\alpha) - n\log(\eta) + \sum_{i=1}^{n} \log\left(z_i^2\right) - \frac{1}{2}\sum_{i=1}^{n} z_i^2 \\ + (\alpha - 1)\sum_{i=1}^{n} \log\left[\Phi(z_i) - z_i\phi(z_i)\right],\end{aligned}$$

where $\boldsymbol{\theta} = (\xi, \eta, \alpha)^\top$ and $z_i = (x_i - \xi)/\eta$, for $i = 1, \ldots, n$. After some calculations, we obtain the following score functions

$$\frac{\partial l}{\partial \xi} = -\frac{1}{\eta}\left[\sum_{i=1}^{n} \frac{2}{z_i} - \sum_{i=1}^{n} z_i + (\alpha - 1)\sum_{i=1}^{n} \frac{z_i^2 \phi(z_i)}{\mathcal{F}_{\text{BN}}(z_i)}\right],$$
$$\frac{\partial l}{\partial \eta} = -\frac{1}{\eta}\left[3n - \sum_{i=1}^{n} z_i^2 + (\alpha - 1)\sum_{i=1}^{n} \frac{z_i^3 \phi(z_i)}{\mathcal{F}_{\text{BN}}(z_i)}\right],$$
$$\frac{\partial l}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \log\left[\Phi(z_i) - z_i\phi(z_i)\right],$$

where $\mathcal{F}_{\text{BN}}(z_i)$ is the CDF of the BN model given in (11). Taking the second partial derivative to the log-likelihood function, the following elements of the observed information matrix are obtained

$$j_{\xi\xi} = \frac{1}{\eta^2}\left[\sum_{i=1}^{n} \frac{2}{z_i^2} + n + (\alpha - 1)\sum_{i=1}^{n} \frac{z_i\phi(z_i)}{\mathcal{F}_{\text{BN}}^2(z_i)}\left[\left(z_i^2 - 2\right)\mathcal{F}_{\text{BN}}(z_i) + z_i^3\phi(z_i)\right]\right],$$
$$j_{\xi\eta} = \frac{1}{\eta^2}\left[2\sum_{i=1}^{n} z_i + (\alpha - 1)\sum_{i=1}^{n} \frac{z_i^2\phi(z_i)}{\mathcal{F}_{\text{BN}}^2(z_i)}\left[\left(z_i^2 - 3\right)\mathcal{F}_{\text{BN}}(z_i) + z_i^3\phi(z_i)\right]\right],$$
$$j_{\eta\eta} = \frac{1}{\eta^2}\left[-3n + 3\sum_{i=1}^{n} z_i^2 + (\alpha - 1)\sum_{i=1}^{n} \frac{z_i^3\phi(z_i)}{\mathcal{F}_{\text{BN}}^2(z_i)}\left[\left(z_i^2 - 4\right)\mathcal{F}_{\text{BN}}(z_i) + z_i^3\phi(z_i)\right]\right],$$

$$j_{\xi\alpha} = \frac{1}{\eta}\sum_{i=1}^{n} \frac{z_i^2\phi(z_i)}{\mathcal{F}_{\text{BN}}(z_i)}, \qquad j_{\eta\alpha} = \frac{1}{\eta}\sum_{i=1}^{n} \frac{z_i^3\phi(z_i)}{\mathcal{F}_{\text{BN}}(z_i)}, \qquad j_{\alpha\alpha} = \frac{n}{\alpha^2}.$$

Taking the expected value to the previous expressions, we obtain the following elements of the Fisher information matrix $\mathbf{I}(\boldsymbol{\theta})$

$$i_{\xi\xi} = \frac{1}{\eta^2}\left[1 + \mathbb{E}(Z^{-2}) + (\alpha - 1)b_{12}\right], \quad i_{\xi\eta} = \frac{1}{\eta^2}\left[2\mathbb{E}(Z) + (\alpha - 1)b_{23}\right],$$
$$i_{\eta\eta} = \frac{1}{\eta^2}\left[-3 + 3\mathbb{E}(Z^2) + (\alpha - 1)b_{34}\right], \qquad i_{\xi\alpha} = \frac{a_2}{\eta}, \quad i_{\eta\alpha} = \frac{a_3}{\eta}, \quad i_{\alpha\alpha} = \frac{1}{\alpha^2},$$

where $a_j = \mathbb{E}\left[\frac{Z^j\phi(z)}{F(z)}\right]$ and $b_j = \mathbb{E}\left[a_j\left(\frac{Z^2-k}{F(z)} + a_3\right)\right]$. When $\alpha = 1$ is obtained using numerical calculations

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{pmatrix} \dfrac{3}{\eta^2} & 0 & \dfrac{0.7663}{\eta} \\ 0 & \dfrac{6}{\eta^2} & -\dfrac{1.0103}{\eta} \\ \dfrac{0.7663}{\eta} & -\dfrac{1.0103}{\eta} & 1 \end{pmatrix},$$

whose determinant is

$$\det(\mathbf{I}(\boldsymbol{\theta})) = \frac{11.4165}{\eta^4} \neq 0.$$

Then, the information matrix of the model is non-singular at the boundary of $\alpha = 1$. Here, the regularity conditions are generally satisfied and the usual $\sqrt{n}$ property for the MLEs is valid for all ξ, η and α. Therefore, for large sample sizes it follows that

$$(\hat{\xi}, \hat{\eta}, \hat{\alpha})^\top \xrightarrow{D} \mathrm{N}_3\big((\xi, \eta, \alpha)^\top, \mathbf{I}(\boldsymbol{\theta})^{-1}\big),$$

then, the MLEs are consistent and have asymptotic normal distribution with covariance matrix the inverse of the Fisher information matrix.

4. Asymmetric Exponentiated Elliptical Bimodal Normal Model

According to Elal-Olivero, Gómez & Quintana (2009), within the skew-elliptical bimodal family of ditributions, we find the elliptical bimodal normal model with PDF given by

$$f_{\mathrm{ELBN}}(z;\gamma) = \left(\frac{1+\gamma z^2}{1+\gamma}\right)\phi(z), \quad z \in \mathbb{R}, \tag{20}$$

where $\gamma \geq 0$ is a shape parameter. This model is denoted by ELBN(γ) and is unimodal for $\gamma = 0$ (normal case) and symmetric bimodal in other cases ($\gamma > 0$). The CDF is given by

$$\mathcal{F}_{\mathrm{ELBN}}(z;\gamma) = \Phi(z) - \frac{\gamma}{1+\gamma} z\phi(z), \quad z \in \mathbb{R}.$$

Now, we propose an extension of the exponentiated type for the ELBN(γ) model whose PDF is given by

$$f_{\mathrm{EEBN}}(z;\gamma,\alpha) = \alpha\left(\frac{1+\gamma z^2}{1+\gamma}\right)\phi(z)\left[\Phi(z) - \frac{\gamma}{1+\gamma} z\phi(z)\right]^{\alpha-1}, \tag{21}$$

where $z \in \mathbb{R}$, $\gamma \geq 0$ and $\alpha \in \mathbb{R}^+$. Here α is an asymmetry parameter. The model in (21) is called exponentiated elliptical bimodal normal (EEBN) model and we denoted it by EEBN(γ, α). One can be seen that, for $\alpha = 1$ the ELBN(γ) model is obtained, for $\gamma = 0$, the PN(α) model is followed, and if $\gamma \to \infty$, the EBN model is obtained; likewise, for $\gamma \to \infty$ and $\alpha = 1$ the BN model follows, and when $\gamma = 0$ and $\alpha = 1$ then the standard normal is obtained. Thus, EEBN model is more flexible than BN, PN, ELBN, EBN and normal distributions.

For the EEBN model can be shown that, when α tends to 1, the mode of the distribution approaches to $\pm\sqrt{2}$ and the distribution tends to be symmetric. In the standard case, the distribution has one mode for values of the random variable less than zero, and the other for values greater than zero. In addition, when $\alpha < 1$ the maximum value of the PDF is found for the mode that takes the smallest (negative) value, while for $\alpha > 1$, the maximum value of the PDF is found for the mode which takes the highest (positive) value.

The graphs in Figure 2(a), (b) and (c) show the behavior of the EEBN distribution for different values of γ and α. One can observe from the figure that the γ parameter determines the bimodality or unimodality of the distribution, while the α parameter explains the asymmetry of the model.

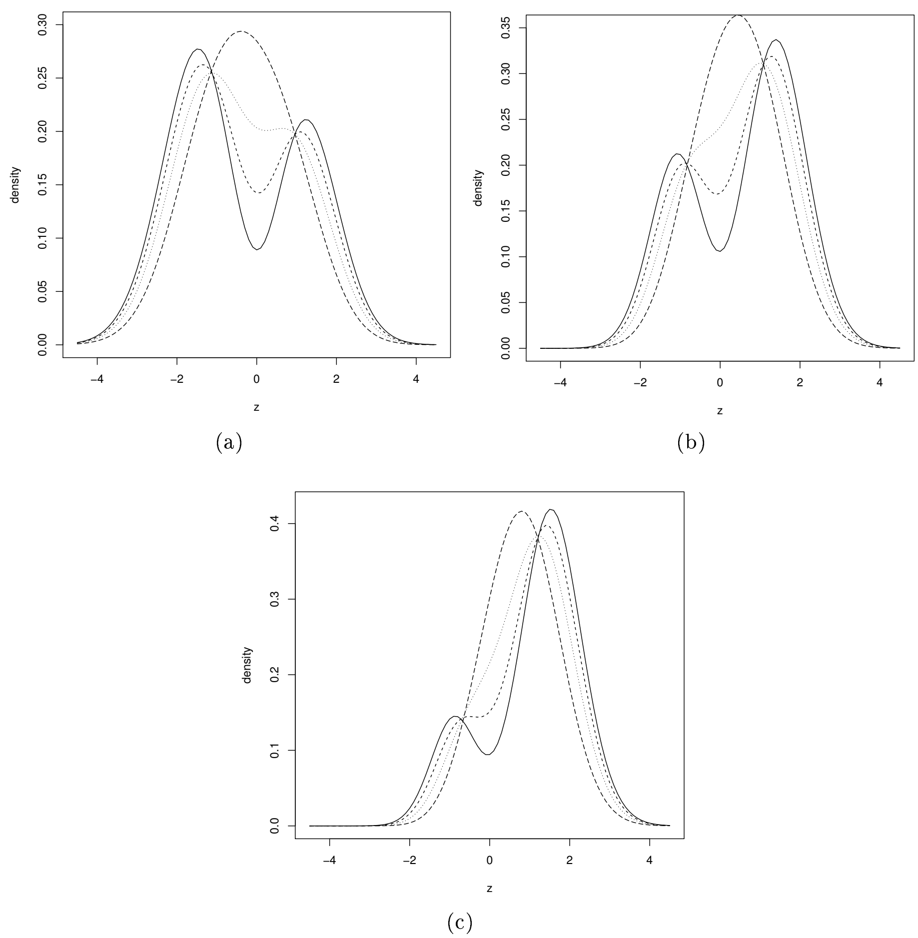


Figure 2: EEBN distribution for $\gamma = 0.25$ (dotted-dashed line), $\gamma = 0.75$ (dotted line), $\gamma = 1.5$ (dashed line) and $\gamma = 3$ (solid line). (a) $\alpha = 0.75$, (b) $\alpha = 1.5$ and (c) $\alpha = 2.25$.

4.1. Properties of the EEBN Model

As an inherited property of the AP family of distributions, it can be seen that the distribution function of this model corresponds to the power of the distribution function of the ELBN model, thus, the CDF of the EEBN(γ, α) is

$$\mathcal{F}_{\text{EEBN}}(z;\gamma,\alpha) = \left[\Phi(z) - \frac{\gamma}{1+\gamma} z\phi(z)\right]^{\alpha}, \quad z \in \mathbb{R}.$$

Therefore, the survival and hazard functions are given by

$$\mathcal{S}_{\text{EEBN}}(t) = 1 - \left[\Phi(t) - \frac{\gamma}{1+\gamma} t\phi(t)\right]^{\alpha}, \quad t \in \mathbb{R},$$

and

$$h_{\text{EEBN}}(t) = \alpha \left(\frac{1+\gamma t^2}{1+\gamma}\right) \phi(t) \frac{\left[\Phi(t) - \frac{\gamma}{1+\gamma} t\phi(t)\right]^{\alpha-1}}{1 - \left[\Phi(t) - \frac{\gamma}{1+\gamma} t\phi(t)\right]^{\alpha}}, \quad t \in \mathbb{R},$$

respectively. The moments of the EEBN distribution do not have a specific shape and are obtained numerically from the general definition, as well as for the calculation of the variance and the asymmetry and kurtosis coefficients.

The location-scale version of the $Z \sim$ EEBN(γ, α) model, follows from the linear transformation $X = \xi + \eta Z$, where $\xi \in \mathbb{R}$ and $\eta \in \mathbb{R}^{+}$. Then, letting $z = (x - \xi)/\eta$, the PDF of X is given by

$$f_{\text{EEBN}}(x;\alpha,\xi,\eta) = \frac{\alpha}{\eta} \left(\frac{1+\gamma z^2}{1+\gamma}\right) \phi(z) \left[\Phi(z) - \frac{\gamma}{1+\gamma} z\phi(z)\right]^{\alpha-1}. \tag{22}$$

4.2. Parameters Estimation in EEBN Model

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n such that, $X_i \sim$ EEBN$(\xi, \eta, \gamma, \alpha)$ for $i = 1, 2, \ldots, n$. The log-likelihood function for estimating the parameters vector $\boldsymbol{\theta} = (\xi, \eta, \gamma, \alpha)^{\top}$ is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}; \mathbf{X}) = {} & n\log(\alpha) - n\log(\eta) + \sum_{i=1}^{n} \log\left(1 + \gamma z_i^2\right) - n\log(1+\gamma) - \frac{1}{2}\sum_{i=1}^{n} z_i^2 \\ & + (\alpha - 1)\sum_{i=1}^{n} \log\left[\Phi(z_i) - \frac{\gamma}{1+\gamma} z_i\phi(z_i)\right], \end{aligned} \tag{23}$$

where $z_i = (x_i - \xi)/\eta$, for $i = 1, \ldots, n$. After some calculations, the elements of the score function are given by

$$\begin{aligned} \frac{\partial \ell}{\partial \xi} &= -\frac{1}{\eta}\left[\sum_{i=1}^{n} \frac{2\gamma z_i}{1+\gamma z_i^2} - \sum_{i=1}^{n} z_i + \frac{\alpha-1}{1+\gamma}\sum_{i=1}^{n} \frac{(1+\gamma z_i^2)\phi(z_i)}{\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})}\right], \\ \frac{\partial \ell}{\partial \eta} &= -\frac{1}{\eta}\left[n + \sum_{i=1}^{n} \frac{2\gamma z_i^2}{1+\gamma z_i^2} - \sum_{i=1}^{n} z_i^2 + \frac{\alpha-1}{1+\gamma}\sum_{i=1}^{n} \frac{z_i(1+\gamma z_i^2)\phi(z_i)}{\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})}\right], \\ \frac{\partial \ell}{\partial \gamma} &= \sum_{i=1}^{n} \frac{z_i^2}{1+\gamma z_i^2} - \frac{n}{1+\gamma} - \frac{\alpha-1}{(1+\gamma)^2}\sum_{i=1}^{n} \frac{z_i\phi(z_i)}{\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})}, \\ \frac{\partial \ell}{\partial \alpha} &= \frac{n}{\alpha} + \sum_{i=1}^{n} \log\left[\Phi(z_i) - \frac{\gamma}{1+\gamma} z_i\phi(z_i)\right], \end{aligned}$$

where $\boldsymbol{\vartheta} = (\xi, \eta, \gamma)^\top$. The elements of the observed information matrix can be written as

$$\imath_{\xi\xi} = \frac{1}{\eta^2}\left[n - 2\gamma\sum_{i=1}^{n}\frac{1}{\left(1+\gamma z_i^2\right)^2} + 2\gamma^2\sum_{i=1}^{n}\frac{z_i^2}{\left(1+\gamma z_i^2\right)^2}\right] + \frac{\alpha-1}{\eta^2(1+\gamma)^2}\sum_{i=1}^{n}\frac{(1+\gamma)z_i\phi(z_i)\left[(1-2\gamma)+\gamma z_i^2\right]\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta}) + (1+\gamma z_i^2)\phi^2(z_i)}{\mathcal{F}^2_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})},$$

$$\imath_{\xi\eta} = \frac{1}{\eta^2}\left[-2\gamma\sum_{i=1}^{n}\frac{z_i}{\left(1+\gamma z_i^2\right)^2} + 2\gamma^2\sum_{i=1}^{n}\frac{z_i^3}{\left(1+\gamma z_i^2\right)^2} + \sum_{i=1}^{n} z_i\right] + \frac{\alpha-1}{\eta^2(1+\gamma)^2}\sum_{i=1}^{n}\frac{(1+\gamma)z_i^2\phi(z_i)\left[(1-2\gamma)+\gamma z_i^2\right]\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta}) + z_i(1+\gamma z_i^2)\phi^2(z_i)}{\mathcal{F}^2_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})},$$

$$\imath_{\eta\eta} = \frac{1}{\eta^2}\left[-n - 2\gamma\sum_{i=1}^{n}\frac{z_i^2(3+\gamma z_i^2)}{\left(1+\gamma z_i^2\right)^2} + 3\sum_{i=1}^{n} z_i^2\right] + \frac{\alpha-1}{\eta^2(1+\gamma)^2}\sum_{i=1}^{n}\frac{(1+\gamma)z_i\phi(z_i)\left[-2+(1-4\gamma)z_i^2+\gamma z_i^4\right]\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta}) + z_i^2(1+\gamma z_i^2)^2\phi^2(z_i)}{\mathcal{F}^2_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})},$$

$$\imath_{\xi\gamma} = \frac{2}{\eta}\sum_{i=1}^{n}\frac{z_i}{\left(1+\gamma z_i^2\right)^2} + \frac{\alpha-1}{\eta(1+\gamma)^3}\sum_{i=1}^{n} z_i\frac{(1+\gamma)(z_i^2-1)\phi(z_i)\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta}) + z_i(1+\gamma z_i^2)\phi^2(z_i)}{\mathcal{F}^2_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})},$$

$$\imath_{\eta\gamma} = \frac{2}{\eta}\sum_{i=1}^{n}\frac{z_i^2\left(1-\gamma z_i+\gamma z_i^2\right)}{\left(1+\gamma z_i^2\right)^2} + \frac{\alpha-1}{\eta(1+\gamma)^3}\sum_{i=1}^{n}\frac{(1+\gamma)(z_i^2-1)\phi(z_i)\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta}) + z_i(1+\gamma z_i^2)\phi^2(z_i)}{\mathcal{F}^2_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})},$$

$$\imath_{\xi\alpha} = \frac{1}{\eta(1+\gamma)}\sum_{i=1}^{n}\frac{(1+\gamma z_i^2)\phi(z_i)}{\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})}, \qquad \imath_{\eta\alpha} = \frac{1}{\eta(1+\gamma)}\sum_{i=1}^{n}\frac{z_i(1+\gamma z_i^2)\phi(z_i)}{\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})}$$

$$\imath_{\gamma\alpha} = \frac{1}{(1+\gamma)^2}\sum_{i=1}^{n}\frac{z_i\phi(z_i)}{\mathcal{F}_{\text{ELBN}}(z_i;\boldsymbol{\vartheta})}, \qquad \imath_{\alpha\alpha} = \frac{n}{\alpha^2}.$$

Then, the elements of the expected information matrix $\kappa_{\theta_j\theta_{j'}}$, are obtained by taking the expected value of the elements of the observed information matrix and by the structure of these elements, these cannot be found explicitly, so numerical methods must be used to find the respective expected values for each value of ξ, η, γ and α.

Letting $\kappa_{\theta_j\theta_{j'}} = \mathbb{E}(\iota_{\theta_j\theta_{j'}})$, then, the expected information matrix is $\mathbf{J}(\boldsymbol{\theta}) = (\kappa_{\theta_j\theta_{j'}})$ where $\theta_j, \theta_{j'}$ are in $\boldsymbol{\theta} = (\xi, \eta, \gamma, \alpha)^\top$. Since the observed information matrix converges asymptotically to the expected information matrix, for $\gamma \neq 0$ and large sample sizes, we have

$$(\hat{\xi}, \hat{\eta}, \hat{\gamma}, \hat{\alpha})^\top \xrightarrow{D} \mathrm{N}_4\big((\xi, \eta, \gamma, \alpha)^\top, \mathbf{J}(\boldsymbol{\theta})^{-1}\big),$$

Therefore, the MLE is consistent and have asymptotic normal distribution, with the covariance matrix being the inverse of the observed information matrix.

4.3. Asymmetric Censored Exponentiated Bimodal Normal Model

This section presents the extension of the EBN model to the case of censored data. Suppose a latent random variable

$$y_i = \begin{cases} y_i^*, & \text{if } y_i^* > c, \\ c, & \text{otherwise} \end{cases} \tag{24}$$

where Y_i^* follows an asymmetric exponentiated bimodal normal $\mathrm{EBN}(\xi, \eta, \alpha)$ distribution, for $i = 1, 2, ..., n$, then the resulting variable Y is said to have an asymmetric censored EBN distribution which is denoted by $\mathrm{CEBN}(\xi, \eta, \alpha)$. For getting the estimates of the parameters vector $\boldsymbol{\theta} = (\xi, \eta, \alpha)^\top$, the maximum likelihood method is used. Thus, for a random sample $\mathbf{X} = (X_1, X_2, \dots, X_n)^\top$, with $X_i \sim \mathrm{CEBN}(\xi, \eta, \alpha)$, the log-likelihood function has the form

$$\ell(\boldsymbol{\theta}; \mathbf{X}) = \alpha \sum_0 \log\left[\Phi(z_{0i}) - z_{0i}\phi(z_{0i})\right] + \sum_1 \left[\log(\alpha) - \log(\eta) + \log\left(z_i^2\right) - \frac{1}{2}z_i^2 + (\alpha - 1)\log\left[\Phi(z_i) - z_i\phi(z_i)\right]\right],$$

where $\sum_0$ and $\sum_1$ refer to the sum over the censored and uncensored observations, respectively; and $z_{0i} = (c - \xi)/\eta$. The maximization of this function regarding the vector of parameters leads to the MLE of the parameters of the model.

4.4. Asymmetric Censored Exponentiated Elliptical Bimodal Normal

We consider a latent random variable as in (24) and we change the assumption of EBN distribution by $Y_i^* \sim \mathrm{EEBN}(\xi, \eta, \gamma, \alpha)$ for $i = 1, 2, \dots, n$. In this case, the random variable Y follows an asymmetric censored exponentiated elliptical bimodal normal distribution and we use the notation $\mathrm{CEEBN}(\xi, \eta, \gamma, \alpha)$.

In the same way as in CEBN model, for obtaining the estimates of the parameters vector $\boldsymbol{\theta} = (\xi, \eta, \gamma, \alpha)^\top$, the ML method is used, thus, for a random sample $\mathbf{X} = (X_1, X_2, \ldots, X_n)^\top$, with $X_i \sim \text{CEEBN}(\xi, \eta, \gamma, \alpha)$, the log-likelihood function is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}; \mathbf{X}) = {} & \alpha \sum_0 \log \left[\Phi(z_{0i}) - \frac{\gamma}{1+\gamma} z_{0i}\phi(z_{0i})\right] \\ & + \sum_1 \left[\log(\alpha) - \log(\eta) + \log\left(1 + \gamma z_i^2\right)\right] \\ & + \sum_1 \left[-\log(1+\gamma) - \frac{1}{2}z_i^2 + (\alpha - 1)\log\left(\Phi(z_i) - \frac{\gamma}{1+\gamma} z_i \phi(z_i)\right)\right],\end{aligned}$$

where $\boldsymbol{\theta} = (\xi, \eta, \gamma, \alpha)^\top)$, $\sum_0$ and $\sum_1$ refer to the sum over the censored and uncensored observations, respectively, and $Z_{0i} = (c - \xi)/\eta$.

5. Simulation Study

5.1. Asymptotic Properties

In order to study the performance of the maximum likelihood estimator $\hat{\boldsymbol{\theta}} = (\hat{\xi}, \hat{\eta}, \hat{\alpha})^\top$ of the parameter $\boldsymbol{\theta} = (\xi, \eta, \alpha)^\top$ in the EBN model, a small Monte Carlo simulation study was carried out with sample sizes $n = 20, 30, 60$ and 90. The true values of the parameters were taken as $\xi = 0$, $\eta = 1$ and $\alpha = 1.25, 1.5, 1.75$. For each combination of parameters $10^4 = 10,000$ samples of the EBN model were generated. To evaluate the performance of the estimators, the absolute value of the bias (AVB) and the root of the mean square error (MSE) were considered, they are given by

$$\text{AVB}(\hat{\theta}_i) = \frac{1}{10.000}\left|\sum_{j=1}^{10.000}(\hat{\theta}_i^{(j)} - \theta_i)\right|; \quad \text{RMSE}(\hat{\theta}_i) = \frac{1}{100}\sqrt{\sum_{j=1}^{10.000}(\hat{\theta}_i^{(j)} - \theta_i)^2},$$

respectively, where $\hat{\theta}_i$ is the estimator of θ_i for the j-th sample, for $\hat{\theta}_i \in (\xi, \eta, \alpha)^\top$. The ML estimates of the parameters were calculated by using the *optim* function of R Development Core Team (2021). The optimization of the likelihood function was done by using iterative methods based on the Newton-Rapshon algorithm.

One can be seen from the Table 1 that the bias is very close to zero and tends to decrease when the value of n increases. The RMSE also tends to decrease when the value of n increases, indicating that estimates based on the ML method have good asymptotic properties. That pattern is the same for all different scenarios under consideration. Likewise, it can be concluded that the ML estimates are asymptotically consistent.

Table 1: Asymptotic behavior of the MLE for the parameters of the EBN model.

		$\hat{\xi}$		$\hat{\eta}$		$\hat{\alpha}$	
α	n	AVB	RMSE	AVB	RMSE	AVB	RMSE
	20	0.0545	0.2202	0.3321	0.0968	0.1336	0.5461
	30	0.0538	0.1971	0.3273	0.0786	0.0993	0.4138
1.25	60	0.0537	0.1717	0.3227	0.0589	0.0660	0.3183
	90	0.0524	0.1593	0.3209	0.0503	0.0631	0.2772
	120	0.0512	0.1519	0.3199	0.0460	0.0597	0.2591
	20	0.0570	0.2699	0.3053	0.1040	0.3863	1.2141
	30	0.0529	0.2139	0.3022	0.0814	0.3033	0.5320
1.50	60	0.0522	0.1827	0.2980	0.0623	0.2612	0.4017
	90	0.0502	0.1690	0.2965	0.0549	0.2461	0.3470
	120	0.0490	0.1573	0.2952	0.0486	0.2447	0.3162
	20	0.0472	0.4547	0.2804	0.1430	1.2301	7.9987
	30	0.0460	0.2656	0.2786	0.0945	0.5556	2.0027
1.75	60	0.0499	0.1958	0.2759	0.0686	0.4581	0.4745
	90	0.0499	0.1759	0.2737	0.0598	0.4406	0.4042
	120	0.0482	0.1638	0.2725	0.0538	0.4317	0.3663

5.1.1. Coverage Probability for the Standard Asymptotic Confidence Intervals

The likelihood function for a random sample of size n, $\mathbf{X} = (X_1, \ldots, X_n)^\top$, from the distribution $\text{EBN}(\xi, \eta, \alpha)$ is given by

$$\begin{aligned}\ell(\boldsymbol{\theta}; \mathbf{X}) = {} & n\log(\alpha) - n\log(\eta) + \sum_{i=1}^{n} \log\left(z_i^2\right) - \frac{1}{2}\sum_{i=1}^{n} z_i^2 \\ & + (\alpha - 1)\sum_{i=1}^{n} \log\left[\Phi(z_i) - z_i\phi(z_i)\right],\end{aligned}$$

with $z_i = (x_i - \xi)/\eta$, $i = 1, \ldots, n$. Therefore, the MLE of α can be obtained as

$$\hat{\alpha} = -\frac{n}{\sum_{i=1}^{n} \log\left[\Phi(z_i) - z_i\phi(z_i)\right]},$$

which is a complete and sufficient statistic for the parameter α,

$$T(\mathbf{X}) = \sum_{i=1}^{n} \log\left[\Phi(z_i) - z_i\phi(z_i)\right].$$

Therefore,

$$\frac{\partial^2 \ell(\alpha; \mathbf{y})}{\partial \alpha^2} = -\frac{n}{\alpha^2}.$$

It can be shown that the asymptotic variance of $\hat{\alpha}$ is given by α^2/n, which agrees with the asymptotic variance in the EBN model.

Since the distribution function of a random variable with EBN distribution follows uniform distribution on the interval $(0, 1)$ and, minus the logarithm of a random variable has uniform distribution, has an exponential distribution with parameter 1, then it follows that the random variable $W = -\sum_{i=1}^{n} \log\left[\Phi(Z_i) - Z_i\phi(Z_i)\right]$ has Gamma$(n, 1)$ distribution. Hence, $\hat{\alpha}$ follows a distribution according to $n\alpha/W$, for $n > 2$. Therefore,

$$\mathbb{E}(\hat{\alpha}) = n\alpha\mathbb{E}\left(\frac{1}{W}\right) = \frac{n}{n-1}\alpha.$$

In addition, one can show that

$$\mathbb{E}_{\alpha}(\hat{\alpha} - \alpha)^2 = \frac{n+2}{(n-1)(n-2)}\alpha^2.$$

From Chebyshev's inequality, it follows that

$$\text{Prob}_{\alpha}(|\hat{\alpha} - \alpha| \geq \epsilon) \leq \frac{\mathbb{E}_{\alpha}(\hat{\alpha} - \alpha)^2}{\epsilon^2} = \frac{n+2}{(n-1)(n-2)\epsilon^2}\alpha^2,$$

therefore, when n tends to ∞, it follows that $\text{Prob}_{\alpha}(|\hat{\alpha} - \alpha| \geq \epsilon)$ tends to zero, and we conclude $\hat{\alpha}$ is a consistent estimator for the parameter α.

It can be shown that, a pivotal quantity for α is given by

$$-2\alpha \sum_{i=1}^{n} \log\left[\Phi(z_i) - z_i\phi(z_i)\right],$$

then, a confidence interval for α with level of confidence $100(1-\rho)\%$ is given by

$$\left(\frac{\hat{\alpha}\chi^2_{(2n,\rho/2)}}{2n}, \frac{\hat{\alpha}\chi^2_{(2n,1-\rho/2)}}{2n}\right).$$

Note that, as n increases the length (L) of the interval is smaller, so as $n \to \infty$, then $L \to 0$ is expected, thus, as n increases the probability of parameter coverage may decrease.

For this case we implement a small Monte Carlo simulation study, with $10,000 = 10^4$ iterations for the sample sizes of $n = 20, 30, 60, 90$ and 120, and α values of 1.25 , $1.50, 1.75$ and 2.25. We calculate the coverage rate of the parameter α for a level confidence of 95%. One can see that, (Table 2) as α is higher, the coverage percentage has at true level of trust.

6. Illustrations

In this section, two illustrations are presented where, it can be seen that the EBN and EEBN models are new alternatives to fitting bimodal data with high (or low) asymmetry and/or kurtosis coefficients.

Table 2: Coverage probability for the standard asymptotic confidence intervals of the EBN model.

α	n	PC(%)
	20	98.22
	30	95.46
1.25	60	78.86
	90	57.01
	120	35.5
	20	99.13
	30	98.05
1.50	60	90.59
	90	77.01
	120	60.35
	20	99.47
	30	98.47
1.75	60	95.10
	90	87.90
	120	78.86
	20	99.25
	30	99.00
2.25	60	97.15
	90	94.63
	120	91.32

6.1. Illustration 1

In the first illustration, the geyser data set, which is available online in the (R Development Core Team, 2021) was used. This data set can be downloaded in the web site `https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/faithful.html`, and consists of 272 observations about the waiting times between eruptions (measured in minutes) of the Old Faithful geyser in the Yellowstone National Park, Wyoming, U.S. Additional information about the data can be found in Azzalini & Bowman (1990).

The Table 3 shows the descriptive statistics for the data, where $\sqrt{b_1}$ and b_2 represent the asymmetry and kurtosis coefficients respectively. The previous results show that the data present a negative skewness and kurtosis below of the usual normal model.

Table 3: Statistical summary for the data.

$\bar{x}$	s_x^2	$\sqrt{b_1}$	b_2
70.897	13.594	-0.414	1.843

The bimodality test of Hartigan & Hartigan (1985) yields the value of the test statistic $D = 0.0414$, with p-value of 0.00181, indicating that the data have a bimodal distribution. In addition, the histogram in Figure 3(a) shows this bimodal behavior.

We fit the proposed asymmetric exponentiated bimodal normal (EBN) and the exponentiated elliptical bimodal normal (EEBN) models. We also fit the asymmetric bimodal normal (BN) model of Elal-Olivero (2010) and the symmetric bimodal model of Arnold et al. (2009) (ETN). Furthermore, we used the maximum likelihood (ML) method to obtain the estimates of the parameters in the fitted models. The ML estimates are presented in the Table 4. The standard errors of the estimates were obtained by using the observed information matrix, and to compare models, we also present the AIC (Akaike, 1974) BIC (Schwarz, 1978) and CAIC (Bozdogan, 1987) criteria, which are defined as follows:

$$AIC = -2\ell(\hat{\boldsymbol{\theta}}) + 2p, \quad BIC = -2\ell(\hat{\boldsymbol{\theta}}) + p\log(n)$$

and

$$CAIC = -2\ell(\hat{\boldsymbol{\theta}}) + p(\log(n) + 1),$$

where p is the number of estimated parameters in the fitted model. The best model is the one with the smallest AIC, BIC or CAIC. According to either of the considered criteria, the EEBN model presented the best fit to the data, followed by the BPN and the BN.

TABLE 4: Estimated parameters (standard errors) for the fitted models.

Parameters	EBN	BN	ETN	EEBN
ξ	66.676(0.158)	67.228(0.242)	66.899(0.916)	66.045(0.636)
η	8.109(0.212)	8.109(0.204)	13.036(0.589)	8.930(0.295)
γ	–	25.111(8.807)	1.547(0.336)	7.642(2.816)
α	1.358(0.088)	–	0.337(0.107)	1.456(0.128)
AIC	2106.75	2127.04	2156.89	2083.50
BIC	2117.57	2137.86	2171.31	2097.92
CAIC	2120.57	2140.86	2175.31	2101.92

A more specific justification for using the EBN model is by hypothesis testing under the assumption of a bimodal normal distribution, that is, hypothesis test

$$H_0 : \alpha = 1 \quad \text{versus} \quad H_1 : \alpha \neq 1,$$

which use the likelihood-ratio statistic given by

$$\Lambda = \frac{\mathcal{L}_{\text{BN}}(\hat{\boldsymbol{\theta}})}{\mathcal{L}_{\text{EBN}}(\hat{\boldsymbol{\theta}})},$$

where $\mathcal{L}$ denotes the likelihood function. For the considered data is obtained

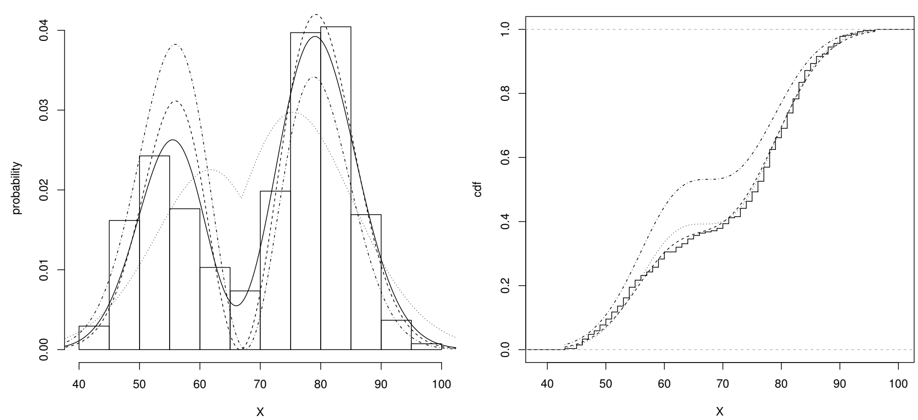
$$-2\log(\Lambda) = 2(1056.26 - 1050.37) = 11.78,$$

so, the $p-\text{value} = \text{Prob}(\chi^2_1 > 11.78) < 0.05$ gives strong evidence for the rejection of the null hypothesis, that is, the data set are not symmetric bimodal, and therefore, an asymmetric bimodal model must be fitted in accordance with the results of the bimodality test of Hartigan & Hartigan (1985) and Hartigan (1985).

To assess the fit of our model, we apply some normality fit tests. First, we apply the modified Lilliefors goodness-of-fit test by Sulewski (2019) and the normality test based on the empirical function proposed by Torabi et al. (2016). In Second place, we also performed the Kolmogorov-Smirnov goodness-of-fit test. The modified Lilliefors test produced the statistic $D = 0.1553$ with $p-$ value $= 0.001$; therefore, we reject the normality hypothesis. Similarly, the results of the normality test based on the empirical distribution produced the statistic $\mathrm{H}_n = 0.2462$ which is greater than the critical value at 5%, of significance, $\mathrm{H}_{106,5\%}$ (see Table 1 Torabi et al., 2016), and again the hypothesis of normality in the data is rejected.

On the other hand, we perform again the Kolmogorov-Smirnov goodness-of-fit test to evaluate our EEBN model and we obtained $D = 0.12661$ with $p-$value $= 0.1704$. Hence, the hypothesis of good fit to the distribution on EEBN of the set of observations is not rejected. Thus, we conclude that the EEBN model best fits the geyser dataset.

Finally, the Figures 3(a) and (b) show the behaviour of the fitted models and the empirical CDF for the fitted models BN, EBN and EEBN. It can be seen that the EEBN model has the best fit compared to the EBN, BN and ETN models.



(a) (b)

FIGURE 3: (a) Histogram for the variable under study. Models: EEBN (solid line), EBN (dotted line), BN (dashed line) and ETN (dotted-dashed line), (b) empirical CDF for EEBN (dotted line), EBN (dashed line) and, BN (dotted-dashed line).

6.2. Illustration 2

For the second application, we consider a data set corresponding to 1.275 HIV-infected people who have been reported in Surveillance and Epidemiology Service of the city of Bucaramanga, Colombia. Among other variables observed in this data set, we find the age, sex, date of admission to the SIVIGILA system, the presence or absence of HAART treatment, highly active antiretroviral therapy, CD-4 count and viral load measured as HIV-1 RNA copy number. This data set was provided by the Secretary of Departmental Health of Santander, Colombia and maintains absolute confidentiality and reserve on the identification of patients.

The database contains patients who are in different stages of treatment, we illustrate the application of the censored bimodal model with the viral load data for the 106 women who have been treated for at least one year with HAART drugs. The proportion of women below the detection limit of 50 copies per millimeter on a base 10 logarithmic scale, $\log_{10}(50)$, is 34.90%, additionally for the Hartigan & Hartigan (1985) bimodality test $D = 0.0686$ with $p-$value $= 0.0133$. rejecting the hypothesis of uniform distribution of viral load. The histogram shown in Figure 4(a) illustrates the existence of bimodality in the distribution of the data set. The Table 5 contains the descriptive statistics for the uncensored data, where it is observed an important positive asymmetry.

Table 5: Statistics summary of $\log_{10}$ HIV-1-RNA for the 69 female observations for the uncensored data.

$\bar{y}$	s_y^2	$\sqrt{b_1}$	b_2
3.410	1.424	0.362	2.042

Given the results of the bimodality test of Hartigan & Hartigan (1985) and the skewness presented in the uncensored data, we fit the asymmetric censored ETN model, see Martínez-Flórez et al. (2018), censored EBN and EEBN models to explain the viral load for the 106 women under study, we also fitted the censored BN model.

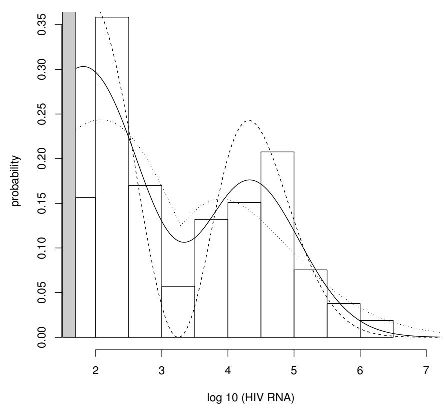
For fitting the CETN, CEBN and CEEBN models, the MLEs are obtained using the optimization method of the R Development Core Team (2021). The obtained estimates (standard errors in parentheses) are found in Table 6, where it can be observed that according to the AIC, BIC and CAIC criteria, the asymmetric censored CETN, CEBN and CEEBN models fit better than the CBN model, which was to be expected given that the CBN model is bimodal symmetric. The estimated proportion of censorship for the CEBN and CEEBN models is 33.44% and 33.89% respectively, illustrating the good fit of these models.

For this data set, the modified Lilliefors test produced the value $D = 0.1318$ with $p-$value $= 0.01$, while for the Torabi et al. (2016) normality test, the value of the statistic was $\mathrm{H}_n = 0.3792 > \mathrm{H}_{106.5\%}$; therefore both tests reject the hypothesis of normality of the data.

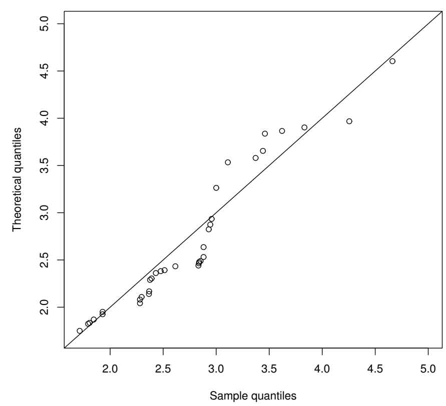
Regarding the Kolmogorov-Smirnov test to assess the hypothesis of good fit of the EEBN distribution to the HIV data set, we obtained $D = 0.4717$, with $p -$ value $= 0.3577$; therefore we do not reject the null hypothesis and conclude that the EEBN model fits better. The graphs in Figure 4(b) and (c) show the QQplot for the EBN and EEBN models of the estimated models, the good fit of the CEBN and CEEBN models can be observed.

Table 6: Estimates parameters (standard errors) for fitted models

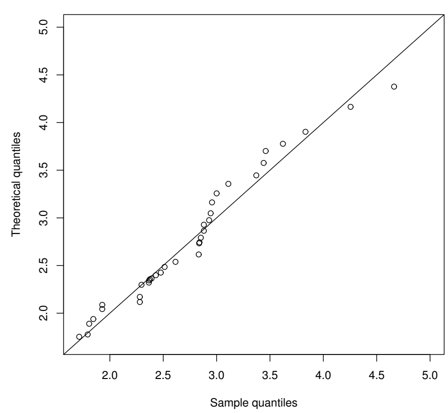
Parameters	CBN	CETN	CEBN	CEEBN
μ	1.485(0.040)	1.587(0.160)	1.550(0.044)	1.604(0.166)
η	0.889(0.042)	1.840(0.213)	0.804(0.044)	0.907(0.075)
γ	–	2.261(1.508)	–	2.090(1.047)
α	–	-0.588(0.199)	0.571(0.082)	0.547(0.116)
AIC	353.95	328.69	335.00	325.02
BIC	364.25	349.29	351.45	345.62
CAIC	366.25	353.24	354.45	349.62



(a)



(b)



(c)

Figure 4: (a) Histogram for the variable under study viral load and fitted models CETN(dotted line), CEBN(dashed line) CEEBN(solid line); (b) QQplot CEBN and (c) QQplot CEEBN.

7. Conclusions

In this paper, new families of distributions were introduced to model asymmetric bimodal data. The new proposals arise from the extension of the bimodal-normal model to the case of the alpha-power family of distributions. The case of models for censored data is also considered. The main properties of the proposed models are studied in detail and the inference process in the models was carried out using the maximum likelihood method. Among the main results of these proposals, one can note that, the information matrices of the studied models are non-singular, which is an advantage over other models support on the skew-normal distribution. Two applications with data sets allow to illustrate the applicability of the studied models, showing very good results compared to other existing models in the literature.

Funding

The research of G. Martínez-Flórez and R. Tovar-Falón was supported by project: Familia bimodal de distribuciones de probabilidad skew-normal alpha-potencia: FCB-18-15.

Acknowledgments

The authors are very thankful for the constructive comments made by the anonymous reviewers. The authors G. Martínez-Flórez and R. Tovar-Falón acknowledges the support given by Universidad de Córdoba, Monteria, Colombia.

[Received: June 2021 — Accepted: February 2022]

References

Akaike, H. (1974), 'A new look at the statistical model identification', *IEEE Transactions on Automatic Control* **19**(6), 716–723.

Arnold, B. C., Gómez, H. & Salinas, H. (2009), 'On multiple contraint skewed models', *Scandinavian Journal of Statistics* **43**(3), 279–293.

Azzalini, A. (1985), 'A class of distributions which includes the normal ones', *Scandinavian Journal of Statistics* **12**(2), 171–178.

Azzalini, A. (1986), 'Further results on a class of distributions which includes the normal ones', *Statistical* **46**(2), 199–208.

Azzalini, A. & Bowman, A. (1990), 'A look at some data on the old faithful geyser', *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **39**(3), 357–365.

Bolfarine, H., Martínez-Flórez, G. & Salinas, H. (2018), 'Bimodal symmetric-asymmetric power-normal families', *Communications in Statistics-Theory and Methods* **47**(2), 259–276.

Bozdogan, H. (1987), 'Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions', *Psychometrika* **52**(3), 345–370.

Durrans, S. (1992), 'Distributions of fractional order statistics in hydrology', *Water Resources Research* **28**(6), 1649–1655.

Elal-Olivero, D. (2010), 'Alpha-skew-normal distribution', *Proyecciones Journal of Mathematics* **29**(3), 224–240.

Elal-Olivero, D., Gómez, H. W. & Quintana, F. A. (2009), 'Bayesian modeling using a class of bimodal skew-elliptical distributions', *Journal of Statistical Planning and Inference* **139**(4), 1484–1492.

Elal-Olivero, D., Olivares-Pacheco, J., Gómez, H. & Bolfarine, H. (2009), 'A new class of non negative distributions generated by symmetric distributions', *Communications in Statistics - Theory and Methods* **38**(7), 993–1008.

Gómez, H., Olivero, D., Salinas, H. & Bolfarine, H. (2009), 'Bimodal extension based on the skew-normal distribution with application to pollen data', *Environmetrics* **22**(1), 50–62.

Gupta, R. & Gupta, R. (2004), 'Generalized skew normal model', *Test* **13**, 501–524.

Gupta, R. & Gupta, R. (2008), 'Analyzing skewed data by power normal model', *Test* **17**, 197–210.

Hartigan, J. & Hartigan, P. (1985), 'The dip test of unimodality', *The Annals of Statistics* **13**(1), 70–84.

Hartigan, P. (1985), 'Algorithm AS 217: Computation of the dip statistics to test for unimodality', *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **34**(1), 320–325.

Henze, N. (1986), 'A probabilistic representation of the skew-normal distribution', *Scandinavian Journal of Statistics* **13**(4), 271–275.

Kim, H. (2005), 'On a class of two-piece skew-normal distribution', *Statistics* **39**(6), 537–553.

Marin, J., Mengersen, K. & Robert, C. (2005), 'Bayesian modelling and inference on mixtures of distributions', *Handbook of Statistics* **25**, 459–507.

Martínez-Flórez, G., Bolfarine, H. & Gómez, H. (2014), 'Skew-normal alpha-power model', *Statistics* **48**(6), 1414–1428.

Martínez-Flórez, G., Bolfarine, H. & Gómez, H. (2018), 'Censored bimodal symmetric-asymmetric families', *Statistics and Its Interface* **11**(2), 237–249.

Martínez-Flórez, G., Tovar-Falón, R. & Jiménez-Narváez, M. (2020), 'Likelihood-Based Inference for the Asymmetric Beta-Skew Alpha-Power Distribution', *Symmetry* **12**(4), 613.

Pewsey, A. (2000), 'Problems of inference for Azzalini's skew-normal distribution', *Journal of Applied Statistics* **27**(7), 859–870.

Pewsey, A., Gómez, H. & Bolfarine, H. (2012), 'Likelihood-based inference for power distributions', *Test* **21**, 775–789.

R Development Core Team (2021), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org

Schwarz, G. E. (1978), 'Estimating the dimension of a model', *Annals of Statistics* **6**(2), 461–464.

Sulewski, P. (2019), 'Modified Lilliefors goodness-of-fit test for normality', *Communications in Statistics-Simulation and Computation* . doi.org/10.1080/03610918.2019.1664580

Torabi, H., Montazeri, N. H. & Grané, A. (2016), 'A test for normality based on the empirical distribution function', *SORT* **40**(1), 55–88.