

REVIEW

Open Access



A new extended normal regression model: simulations and applications

Maria C.S. Lima^{1*†} , Gauss M. Cordeiro^{1†}, Edwin M.M. Ortega^{2†} and Abraão D.C. Nascimento^{3†}

*Correspondence:

maria@de.ufpe.br

[†]Maria C.S. Lima, Gauss M. Cordeiro, Edwin M.M. Ortega, and Abraão D.C. Nascimento contributed equally to this work.

¹Departamento de Estatística, Universidade Federal de Pernambuco, PE 50740-540 Recife, Brazil

Full list of author information is available at the end of the article

Abstract

Various applications in natural science require models more accurate than well-known distributions. In this context, several generators of distributions have been recently proposed. We introduce a new four-parameter extended normal (EN) distribution, which can provide better fits than the skew-normal and beta normal distributions as proved empirically in two applications to real data. We present Monte Carlo simulations to investigate the effectiveness of the EN distribution using the Kullback-Leibler divergence criterion. The classical regression model is not recommended for most practical applications because it oversimplifies real world problems. We propose an EN regression model and show its usefulness in practice by comparing with other regression models. We adopt maximum likelihood method for estimating the model parameters of both proposed distribution and regression model.

Keywords: Kullback-Leibler divergence criterion, Maximum likelihood procedures, Monte Carlo simulation, Normal distribution, Regression

AMS Subject Classification: Primary 60E05; secondary 62N05; 62F10

1 Introduction

In recent years, several methods for generating new models from classic distributions have been proposed. A detailed study about “the evolution of methods for generalizing classic distributions” was made by Lee et al. (2013). A generalization of the standard normal distribution is sought because it can provide more accurate statistical models and inferential procedures. For instance, the beta normal distribution was pioneered by Eugene et al. (2002), who discussed some of its structural properties.

Additionally, the beta generalized normal (BGN) distribution was proposed by Cintra et al. (2013) to extend the beta normal distribution. They applied the BGN model to the synthetic aperture radar image processing. This paper presents a new *extended normal* (EN) distribution based on the family introduced by Cordeiro et al. (2013).

For any continuous cumulative distribution function (cdf) $G(x)$, Cordeiro et al. (2013) defined the cdf of the *exponentiated generalized* (EG) family by

$$F(x) = [1 - \{1 - G(x)\}^a]^b, \quad (1)$$

where $a > 0$ and $b > 0$ are two additional shape parameters whose role is to generate distributions with heavier/lighter tails and provide wider ranges for skewness and kurtosis. These parameters are sought as a manner to furnish a more flexible distribution.

Because of its tractable cdf (1), the EG family can be used quite effectively even if the data are censored. This family is capable to return univariate models for any type of support. Further, it allows for greater flexibility of its tails and can be widely applied in many areas such as engineering and biology.

Its probability density function (pdf) has a very simple form

$$f(x) = ab \{1 - G(x)\}^{a-1} [1 - \{1 - G(x)\}^a]^{b-1} g(x). \tag{2}$$

An important advantage of the density (2) is its ability of fitting skewed data that can not be often fitted by existing distributions. Based on the cdf $G(x)$ and pdf $g(x)$ of any baseline G distribution, we can associate the EG-G pdf (2) with two extra parameters. The EG family can be used for discriminating between the G and EG-G distributions.

The baseline distribution $G(x)$ is a special case of (2) when $a = b = 1$. For $a = 1$, it gives the *exponentiated-G* (“Exp-G”) class. If $b = 1$, we obtain the Lehmann type II-G (LTII-G) class. Eq. (2) generalizes both Lehmann types I and II alternative classes (Lehmann 1953). In fact, this equation can be defined as the exponentiated generator applied to the LTII-G class.

Note that even if $g(x)$ is a symmetric density, the density $f(x)$ will not be symmetric. The cdf (1) has tractable properties especially for simulations, since its quantile function (qf) has a simple form

$$x = Q_G \left\{ \left[1 - \left(1 - u^{\frac{1}{b}} \right)^{\frac{1}{a}} \right] \right\},$$

where $Q_G(u)$ is the baseline qf.

This paper is outlined as follows. In Section 2, we define the EN distribution and provide plots of its density function. A linear representation for the EN density function is derived in Section 3. We obtain an explicit expression for its moments in Section 4. In Section 5, we provide the maximum likelihood estimates (MLEs) of the parameters. In Section 6, we define the EN regression model and discuss the estimation of the model parameters. In Section 7, we perform some simulations and present three applications to real data sets. Finally, some concluding remarks are addressed in Section 8.

2 The EN distribution

Due to the analytical tractability of its pdf and its importance in asymptotic theory (such as the *central limit theorem* and *delta method*), the normal distribution is the most popular model distribution in applications to real data with support in \mathbb{R} .

When the number of observations is large, it can serve as an approximate distribution for several other models. The normal $N(\mu, \sigma)$ pdf (for $x \in \mathbb{R}$) is

$$g(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right), \tag{3}$$

where $\mu \in \mathbb{R}$ is a mean parameter, $\sigma > 0$ is a scale parameter and $\phi(x) = (2\pi)^{-1/2} e^{-x^2/2}$ is the standard normal pdf.

Its cdf has the form

$$G(x; \mu, \sigma) = \int_{-\infty}^x g(t; \mu, \sigma) dt = \Phi\left(\frac{x-\mu}{\sigma}\right), \tag{4}$$

where $\Phi(x) = \int_{-\infty}^x \phi(t) dt$ is the standard normal cdf.

By inserting (3) and (4) in Eqs. (1) and (2), the cdf and pdf of the EN distribution (for $x \in \mathbb{R}$) can be expressed, respectively, as

$$F(x) = \left[1 - \left\{ 1 - \Phi\left(\frac{x - \mu}{\sigma}\right) \right\}^a \right]^b \tag{5}$$

and

$$f(x) = \frac{ab}{\sigma} \left\{ 1 - \Phi\left(\frac{x - \mu}{\sigma}\right) \right\}^{a-1} \left[1 - \left\{ 1 - \Phi\left(\frac{x - \mu}{\sigma}\right) \right\}^a \right]^{b-1} \times \phi\left(\frac{x - \mu}{\sigma}\right). \tag{6}$$

Hereafter, a random variable X having density (6) is denoted by $X \sim EN(a, b, \mu, \sigma)$. Evidently, this density does not involve any complicated function and the normal distribution arises as the basic exemplar when $a = b = 1$. It is a positive point of the current generalization. Moreover, the qf of X is

$$Q_{EN}(p) = \mu + \sigma \Phi^{-1}\left(1 - \left[1 - p^{\frac{1}{b}}\right]^{\frac{1}{a}}\right). \\ m = E(X) = \sigma ab \mathcal{I}_{a,b} + \mu,$$

where

$$\mathcal{I}_{a,b} = \int_{-\infty}^{\infty} z \phi(z) [1 - \Phi(z)]^{a-1} \{1 - [1 - \Phi(z)]^a\}^{b-1} dz.$$

In next sections, other moment results are proved. Moreover, from the previous qf of the EN distribution, the associated median, say M , is

$$M = Q_{EN}(1/2) = \sigma z_{a,b} + \mu,$$

where $z_{a,b} = \Phi^{-1}\left(1 - \left[1 - 2^{-1/b}\right]^{1/a}\right)$ is the standard normal quantile at $1 - \left[1 - 2^{-1/b}\right]^{1/a}$. Thus, the next function suggests a symmetric discussion:

$$\begin{cases} \text{right asymmetry,} & \text{if } z_{a,b} > ab \mathcal{I}_{a,b} \\ \text{symmetry,} & \text{if } z_{a,b} = ab \mathcal{I}_{a,b} \\ \text{left asymmetry,} & \text{if } z_{a,b} < ab \mathcal{I}_{a,b}. \end{cases}$$

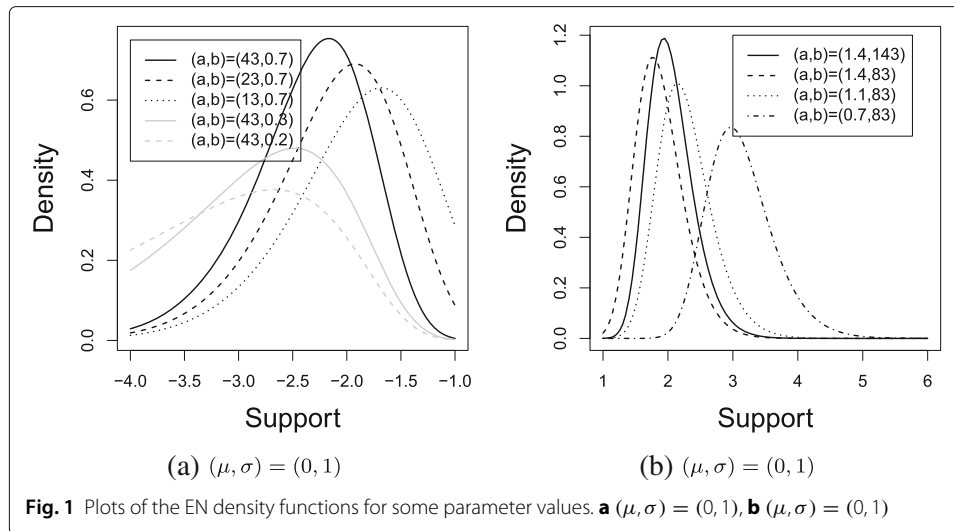
We motivate the paper by comparing the performances of the EN, normal, skew-normal (SN) and beta-normal (BN) models fitted to two real data sets. Figure 1 displays possible shapes of the density function (6) for some parameter values. We can note the flexibility of the EN distribution with respect to the normal distribution.

3 Linear representation

A useful linear representation for (2) can be derived using the concept of exponentiated distributions. For an arbitrary baseline cdf $G(x)$, a random variable T is said to have the exponentiated-G (Exp-G) distribution with power parameter $a > 0$, say $T \sim \text{Exp-G}(a)$, if its pdf and cdf are

$$H_a(x) = G^a(x) \quad \text{and} \quad h_a(x) = a g(x) G^{a-1}(x),$$

respectively. Several properties of the exponentiated distributions have been studied by some authors recently such as those for the exponentiated Weibull (Mudholkar and Srivastava 1993) and exponentiated generalized gamma (Cordeiro et al. 2013) distributions.



Theorem 1 Let $X \sim EN(a, b, \mu, \sigma)$. The pdf of X can be written as

$$f(x) = \sum_{j=0}^{\infty} w_{j+1} h_{j+1}(x), \tag{7}$$

where $h_{j+1}(x)$ is the exponentiated-normal (Exp-N) density with power parameter $j + 1$, say $Exp-N(\mu, \sigma, j + 1)$, namely

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

The proof of this theorem is given in Appendix A.

It is possible to verify using symbolic software (such as Maple) that $\sum_{j=0}^{\infty} w_{j+1} = 1$ as expected.

Equation (7) is the main result of this section. It reveals that the EN density is a linear combination of Exp-N densities. So, several mathematical properties of the proposed distribution can then be obtained from those of the Exp-N distribution using previous results given by Rêgo et al. (2012).

4 Moments

First, we determine the probability weighted moments (PWMs) of the standard normal distribution since they are required for the ordinary moments of the EN distribution. The standard normal PWMs are defined by

$$\tau_{n,j} = \int_{-\infty}^{\infty} z^n \Phi(z)^j \phi(z) dz,$$

for $n \geq 0$ and $j \geq 0$ integers.

The result holds

$$\Phi(z) = \frac{1}{2} \left\{ 1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right) \right\}, \quad z \in \mathbb{R}.$$

Applying the binomial expansion and interchanging terms gives

$$\tau_{n,j} = \frac{1}{2^j \sqrt{2\pi}} \sum_{m=0}^j \binom{j}{m} \int_{-\infty}^{\infty} z^n \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)^{j-m} \exp\left(-\frac{z^2}{2}\right) dz.$$

Based on the power series for the error function

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \sum_{r=0}^{\infty} \frac{(-1)^r z^{2r+1}}{(2r+1)r!},$$

we can obtain $\tau_{n,j}$ from Eqs. (9)–(11) given by Nadarajah (2008).

For $n + j - r$ even, we have

$$\tau_{n,j} = 2^{n/2} \pi^{-(j+1/2)} \sum_{\substack{r=0 \\ (n+j-r) \text{ even}}}^j \binom{j}{r} \left(\frac{\pi}{2}\right)^r \Gamma\left(\frac{n+j-r+1}{2}\right) \times F_A^{(j-r)}\left(\frac{n+j-r+1}{2}; \frac{1}{2}, \dots, \frac{1}{2}; \frac{3}{2}, \dots, \frac{3}{2}; -1, \dots, -1\right), \tag{8}$$

where $F_A^{(j-r)}(\cdot)$ is the Lauricella function of type A. See, for example, Exton (1978)¹. If $n + k - j$ is odd, the corresponding terms in $\tau_{n,j}$ vanish.

Corollary 1 *Suppose that $\mu'_n = E(X^n)$ exists. Then,*

$$\mu'_n = E(X^n) = \sum_{j=0}^{\infty} (j+1) w_{j+1} \tau_{n,j}, \tag{9}$$

where $\tau_{n,j}$ is given by (8).

The skewness and kurtosis of X can be computed from $Q_{EN}(p)$ using Bowley and Moors well-known quantities. Figure 2 displays plots of the skewness and kurtosis measures of X for selected values of a and b . We note that the skewness and kurtosis values for the normal distribution are obtained when values for (a, b) tend to $(1, 1)$.

5 Estimation

Consider a random variable $X \sim \text{EN}(a, b, \mu, \sigma)$ and let $\theta = (a, b, \mu, \sigma)^\top$ be the model parameters, where $(\cdot)^\top$ is the transposition operator. Thus, the associated log-likelihood function for one observation x is

$$\ell(\theta; x) = \log(a) + \log(b) - \log(\sigma) + (a - 1) \log\left[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right] + (b - 1) \log\left\{1 - \left[1 - \Phi\left(\frac{x-\mu}{\sigma}\right)\right]^a\right\} + \log\left[\phi\left(\frac{x-\mu}{\sigma}\right)\right]. \tag{10}$$

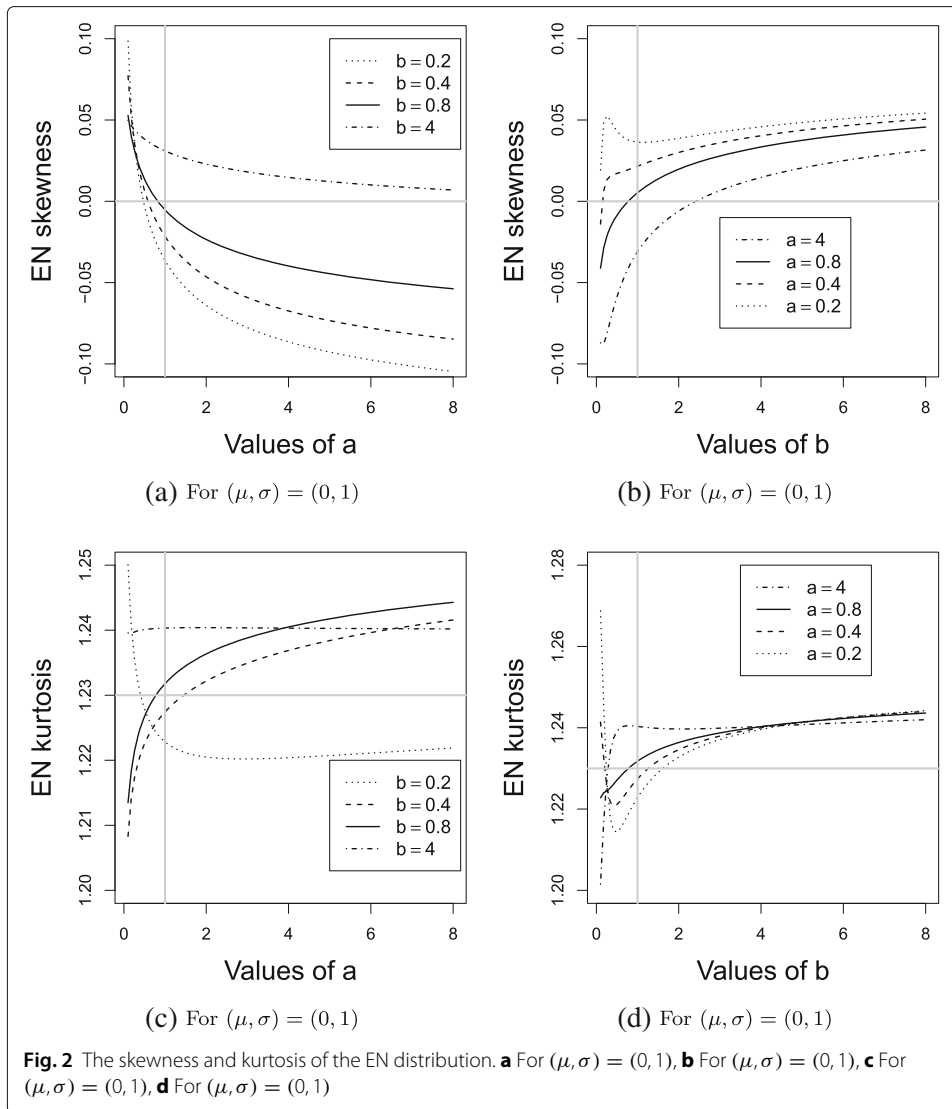
Given a data set x_1, \dots, x_n , the MLE of θ is determined by maximizing $\ell_n(\theta) = \sum_{i=1}^n \ell(\theta; x_i)$.

Based on Eq. (10), the score vector is

$$U_\theta = (U_a, U_b, U_\mu, U_\sigma)^\top = \left(\frac{\partial \ell_n(\theta)}{\partial a}, \frac{\partial \ell_n(\theta)}{\partial b}, \frac{\partial \ell_n(\theta)}{\partial \mu}, \frac{\partial \ell_n(\theta)}{\partial \sigma}\right)^\top,$$

whose components are

$$U_a = \frac{n}{a} + \sum_{i=1}^n \log\left[1 - \Phi\left(\frac{x_i - \mu}{\sigma}\right)\right] - (b - 1) \sum_{i=1}^n \left[\frac{\left[1 - \Phi\left(\frac{x_i - \mu}{\sigma}\right)\right]^a \log\left[1 - \Phi\left(\frac{x_i - \mu}{\sigma}\right)\right]}{1 - \left[1 - \Phi\left(\frac{x_i - \mu}{\sigma}\right)\right]^a}\right],$$



$$U_b = \frac{n}{b} + \sum_{i=1}^n \log \left\{ 1 - \left[1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right) \right]^a \right\},$$

$$U_\mu = \left(\frac{a-1}{\sigma} \right) \sum_{i=1}^n \left[\frac{\phi \left(\frac{x_i - \mu}{\sigma} \right)}{1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right)} \right] - \frac{1}{\sigma} \sum_{i=1}^n \left[\frac{\phi' \left(\frac{x_i - \mu}{\sigma} \right)}{\phi \left(\frac{x_i - \mu}{\sigma} \right)} \right]$$

$$- \frac{a(b-1)}{\sigma} \sum_{i=1}^n \left\{ \frac{\phi \left(\frac{x_i - \mu}{\sigma} \right) [1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right)]^{a-1}}{1 - [1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right)]^a} \right\}$$

and

$$U_\sigma = \frac{n}{\sigma} - \frac{\mu(a-1)}{\sigma^2} \sum_{i=1}^n \left[\frac{\phi \left(\frac{x_i - \mu}{\sigma} \right)}{1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right)} \right] + \frac{\mu}{\sigma^2} \sum_{i=1}^n \left[\frac{\phi' \left(\frac{x_i - \mu}{\sigma} \right)}{\phi \left(\frac{x_i - \mu}{\sigma} \right)} \right]$$

$$+ \frac{a\mu(b-1)}{\sigma^2} \sum_{i=1}^n \left\{ \frac{\phi \left(\frac{x_i - \mu}{\sigma} \right) [1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right)]^{a-1}}{1 - [1 - \Phi \left(\frac{x_i - \mu}{\sigma} \right)]^a} \right\}.$$

An advantage of the EN distribution is that the MLE \widehat{b} has a partially closed-form expression. Suppose that the observed information matrix is non-negative definite. The MLE of b can be expressed in terms of the MLEs \widehat{a} , $\widehat{\mu}$ and $\widehat{\sigma}$ as

$$\begin{aligned}\widehat{b} &= \varphi(\widehat{a}, \widehat{\mu}, \widehat{\sigma}, \{x_1, \dots, x_n\}) \\ &= \left(n^{-1} \sum_{i=1}^n \log \left\{ 1 - \left[1 - \Phi \left(\frac{x_i - \widehat{\mu}}{\widehat{\sigma}} \right) \right]^{\widehat{a}} \right\} \right)^{-1}.\end{aligned}\quad (11)$$

This fact is important at least for two reasons. The estimates become the solutions of a system with three equations and three variables (say “(3, 3) system”) instead of a (4, 4) system. Further, Eq. (11) clarifies the relationship of \widehat{b} with \widehat{a} , $\widehat{\mu}$ and $\widehat{\sigma}$. More details are described in the simulation section.

Additionally, in order to make inference on the model parameters, the total observed information matrix is $J(\theta) = \{-U_{rs}\}$, where $U_{rs} = \partial^2 \ell(\theta) / \partial \theta_r \partial \theta_s$, for $r, s \in \{a, b, \mu, \sigma\}$. By differentiating the score function, we obtain the Hessian matrix elements U_{rs} given in Appendix B.

6 The EN regression model

The classical normal linear regression model is usually applied in science and engineering to describe symmetrical data for which linear functions of unknown parameters are used to explain the phenomena under study. However, it is well-known that several phenomena are not always in agreement with the classical regression model due to lack of symmetry and/or the presence of heavy and lightly tails in the empirical distribution. As an alternative to overcome this shortcoming, we propose a new regression model based on the EN distribution thus extending the normal linear regression.

Let $\mathbf{v}_i = (v_{i1}, \dots, v_{ip})^\top$ be the $p \times 1$ explanatory variable vector associated with the i th response variable x_i (for $i = 1, \dots, n$). Let X_i be a response variable having the EN distribution given by (6) re-parameterized as

$$X_i = \mathbf{v}_i^\top \boldsymbol{\beta} + \sigma Z_i, \quad (12)$$

where the random error $Z \sim EN(a, b, 0, 1)$ has the standardized EN distribution, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ is the unknown vector of coefficients, $\sigma > 0$ is an unknown dispersion parameter and \mathbf{v}_i is the explanatory vector modeling the location parameter $\mu_i = \mathbf{v}_i^\top \boldsymbol{\beta}$.

Hence, the location parameter vector $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$ of the EN regression model has the linear structure $\boldsymbol{\mu} = \mathbf{V}\boldsymbol{\beta}$, where $\mathbf{V} = [\mathbf{v}_1 | \dots | \mathbf{v}_n]^\top$ is a known model matrix.

The EN regression model (12) opens new possibilities for fitting many different types of data, since the EN distribution is much more flexible than the normal distribution. The most important special regressions are:

- For $a = 1$, it gives the exponentiated-normal (Exp-N) regression model, which has not been explored, but it can be understood as a regression under the power normal distribution pioneered by Kundu and Gupta (2013).
- For $b = 1$, it reduces to the LTII-normal (LTII-N) regression model defined as a linear model under the LTII-N distribution.
- If $a = b = 1$, it reduces to the normal linear regression.

For statistical inference on the EN regression model, we consider a sample $(X_1, \mathbf{v}_1), \dots, (X_n, \mathbf{v}_n)$ of n independent observations. The log-likelihood function for the vector of parameters $\boldsymbol{\eta} = (a, b, \sigma, \boldsymbol{\beta}^\top)^\top$ of model (12) is

$$\begin{aligned} \ell(\boldsymbol{\eta}) = & n \log\left(\frac{ab}{\sigma}\right) + \sum_{i=1}^n \log[\phi(z_i)] + (a-1) \sum_{i=1}^n \log[1 - \Phi(z_i)] \\ & + (b-1) \sum_{i=1}^n \log\{1 - [1 - \Phi(z_i)]^a\}, \end{aligned} \tag{13}$$

where $z_i = (x_i - \mathbf{v}_i^\top \boldsymbol{\beta})/\sigma$ and x_i is a possible outcome of X_i .

The components of the score vector $U(\boldsymbol{\eta})$ are

$$\frac{\partial \ell(\boldsymbol{\eta})}{\partial a} = \frac{n}{a} + \sum_{i=1}^n \log[1 - \Phi(z_i)] - (b-1) \sum_{i=1}^n \frac{[1 - \Phi(z_i)]^a \log[1 - \Phi(z_i)]}{\{1 - [1 - \Phi(z_i)]^a\}},$$

$$\frac{\partial \ell(\boldsymbol{\eta})}{\partial b} = \frac{n}{b} + \sum_{i=1}^n \log\{1 - [1 - \Phi(z_i)]^a\},$$

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\eta})}{\partial \sigma} = & -\frac{n}{\sigma} - \frac{2}{\sigma} \sum_{i=1}^n z_i^2 + \frac{(a-1)}{\sigma} \sum_{i=1}^n \frac{z_i \phi(z_i)}{[1 - \Phi(z_i)]} \\ & - \frac{a(b-1)}{\sigma} \sum_{i=1}^n \frac{z_i \phi(z_i) [1 - \Phi(z_i)]^{a-1}}{\{1 - [1 - \Phi(z_i)]^a\}}, \end{aligned}$$

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\eta})}{\partial \beta_j} = & -\frac{2}{\sigma} \sum_{i=1}^n v_{ij} z_i + \frac{(a-1)}{\sigma} \sum_{i=1}^n \frac{v_{ij} \phi(z_i)}{[1 - \Phi(z_i)]} \\ & - \frac{a(b-1)}{\sigma} \sum_{i=1}^n \frac{v_{ij} \phi(z_i) [1 - \Phi(z_i)]^{a-1}}{\{1 - [1 - \Phi(z_i)]^a\}}, \end{aligned}$$

where $j = 1, \dots, p$.

Note that a closed-form expression for the MLE $\hat{\boldsymbol{\eta}}$ is analytically intractable and, therefore, its computation has to be performed numerically by means of a nonlinear optimization algorithm.

We can maximize the log-likelihood function (13) based on the Newton-Raphson method. In particular, we use the matrix programming language **Ox** (MaxBFGS function) (see Doornik 2007) to calculate $\hat{\boldsymbol{\eta}}$. Initial values for $\boldsymbol{\beta}$ and σ can be taken from the fit of the classical regression model ($a = b = 1$).

Under general regularity conditions, the asymptotic distribution of $(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta})$ is multivariate normal $N_{p+3}(0, K(\boldsymbol{\eta})^{-1})$, where $K(\boldsymbol{\eta})$ is the expected information matrix. These conditions can be found in Cox and Hinkley's *Theoretical Statistics* book (1974). The asymptotic covariance matrix $K(\boldsymbol{\eta})^{-1}$ of $\hat{\boldsymbol{\eta}}$ can be approximated by the inverse of the $(p+3) \times (p+3)$ observed information matrix $J(\boldsymbol{\eta})$ and then the inference on the parameter vector $\boldsymbol{\eta}$ can be based on the normal approximation $N_{p+3}(0, J(\boldsymbol{\eta})^{-1})$ for $\hat{\boldsymbol{\eta}}$.

Besides estimation of the model parameters, hypotheses tests can be considered using likelihood ratio (LR) statistics.

7 Numerical results

Three studies are presented in this section. First, we perform a Monte Carlo simulation study. Subsequently, two applications to real data show the potential uses of the new distribution. Third, the usefulness of the proposed regression model in Section 6 is proved empirically based on quality of life data.

7.1 Simulation study

Here, we provide a Monte Carlo simulation study in order to quantify the effectiveness of the EN distribution based on the symmetrized Kullback-Leibler divergence as a goodness-of-fit comparison criterion.

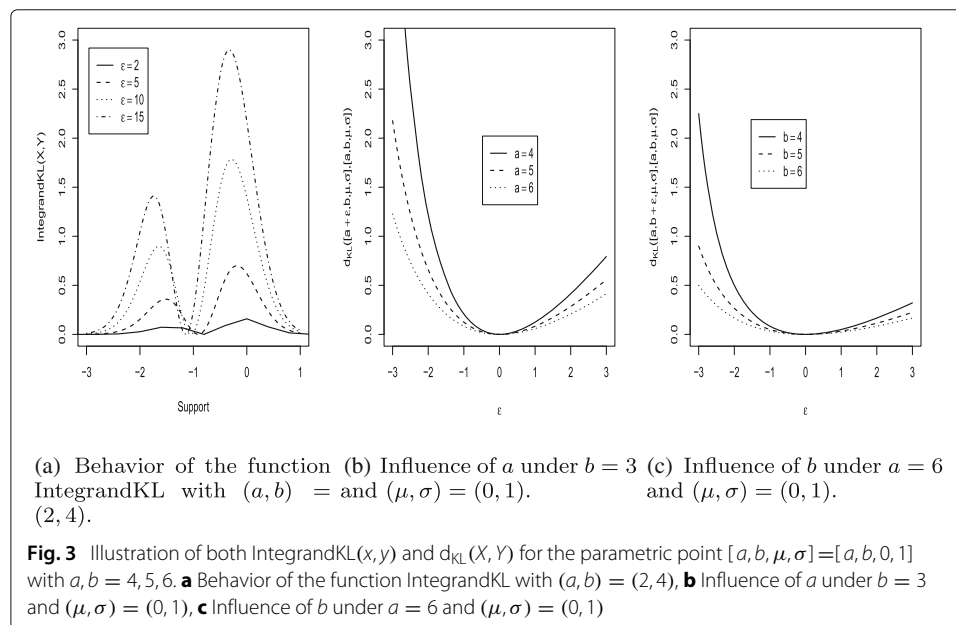
Initially, we provide a brief discussion on the Kullback-Leibler divergence. According to Cover and Thomas (1991), this measure is the quantification of the error by assuming that the Y model is true when the data follow the X distribution. For example, it has been proposed as essential parts of test statistics and strongly applied to contexts of radar synthetic aperture image processing in both univariate (Nascimento et al. 2010) and polarimetric (or multivariate) (Nascimento et al. 2014) perspectives.

In order to work with measures which satisfy non-negativity, symmetry and definiteness properties, Nascimento et al. (2010) considered the measure d_{KL} , namely

$$d_{KL}(X, Y) = \frac{1}{2} [D(X||Y) + D(Y||X)]$$

$$= \int_{\mathcal{D}} \underbrace{(f_X(x; [a_x, b_x, \mu_x, \sigma_x]) - f_Y(x; [a_y, b_y, \mu_y, \sigma_y])) \log \left(\frac{f_X(x; [a_x, b_x, \mu_x, \sigma_x])}{f_Y(x; [a_y, b_y, \mu_y, \sigma_y])} \right)}_{\equiv \text{IntegrandKL}(x,y)} dx.$$

Figure 3 displays both functions $\text{IntegrandKL}(x, y)$ and $d_{KL}(X, Y)$ at the parametric point $[a, b, \mu, \sigma] = [a, b, 0, 1]$ when $a, b = 4, 5, 6$. It is noticeable that this measure can be understood as a distance between the two points $\theta_1 = (a_1, b_1, \mu_2, \sigma_1)$ and $\theta_2 = (a_2, b_2, \mu_2, \sigma_2)$ —in the parametric space, say $d_{KL}(\theta_1, \theta_2)$.



For increasing values of ε , the $\text{Integrand}_{\text{KL}}(X, Y)$ has different forms. Further, $\text{Integrand}_{\text{KL}}(X, Y) \rightarrow 0$ when $\varepsilon \rightarrow 0$.

Figure 3b and c reveal the influence of a and b , respectively, when we employ a perturbation in each parameter under $(\mu, \sigma) = (0, 1)$. As expected, when the value of ε increases, the distance d_{KL} also increases in both cases. However, this distance is most evident when we take smaller negative values of ε .

Table 1 gives the asymptotic performance of the maximum likelihood procedure discussed in the previous section with respect to the Kullback-Leibler distance, where we identify critical scenarios under the parametric space, which can require a harder maximum likelihood estimation. The results support the fact: “when we wish to estimate one additional parameter (a or b) given that the MLE for the other parameter is known and higher than one, then the biases of the estimates tend to increase for high values of the parameter of interest.” In particular, at the MLE of b given \widehat{a} , the above information finds strong justification in Eq. (11). Based on this equation, when \widehat{a} takes high values, the MLE of b collapses for an indetermination algebraic.

7.2 Two applications to real data

Here, we perform two applications to real data sets. First, we consider the data the strengths of glass fibres analyzed by Jones and Faddy (2004). These data were obtained at the *National Physical Laboratory* (UK) to explain the breaking strength of sixty three glass fibres having length 1.5 cm.

As a second application, we consider the fatigue life data (Meeker and Escobar 1998) for sixty seven specimens of Alloy T7987 that failed before having accumulated three hundred thousand cycles of testing. The data set was rounded to the nearest thousand cycles.

We prove empirically the efficiency of the EN distribution versus the normal, skew-normal (SN) (Azzalini 1984) and beta normal (BN) (Eugene et al. 2002) distributions.

The SN density [$T \sim \text{SN}(a, \mu, \sigma)$] has the form (for $x, a, \mu \in \mathbb{R}$ and $\sigma > 0$)

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

and the BN density [$T \sim \text{BN}(\alpha, \beta, \mu, \sigma)$] is (for $x, \mu \in \mathbb{R}$ and $\alpha, \beta, \sigma > 0$)

$$f(x; \alpha, \beta, \mu, \sigma) = J \phi\left(\frac{x - \mu}{\sigma}\right) \left[\Phi\left(\frac{x - \mu}{\sigma}\right)\right]^{\alpha-1} \left[1 - \Phi\left(\frac{x - \mu}{\sigma}\right)\right]^{\beta-1},$$

where $J = \Gamma(\alpha + \beta) / [\Gamma(\alpha) \Gamma(\beta) \sigma]$.

We compare the distributions using three goodness-of-fit (GoF) measures: Anderson-Darling (A^*), Cramer-Von Mises (W^*) and Kolmogorov-Smirnov (KS) statistics. We

Table 1 The KL distance between fitted and theoretical densities for $n = 100$ and different values for a and b

a	b					
	0.5	0.8	1.0	1.5	2.0	2.5
0.5	6.837	5.984	5.662	4.064	2.690	0.545
0.8	4.259	5.628	4.636	4.497	3.581	1.263
1.0	4.663	5.149	4.999	4.834	4.242	2.135
1.5	1.988	3.450	4.349	5.791	5.549	4.301
2.0	1.527	2.884	3.687	6.786	7.075	7.517
2.5	1.435	2.728	3.837	6.590	7.287	10.148

adopt the *goodness.fit* function from the **R** program through the BFGS method. According to detailed discussion in Quang (1989), these measures are more indicated than the Akaike information criterion (AIC) and Bayesian information criterion (BIC) or some of their variations, which are more useful for nested models. Table 2 gives the GoF measures for each fitted distribution with respect to both data sets.

The GoF's measures for the EN distribution correspond to the lowest values among the discrimination criteria (highlighted in Table 2). These results provide evidence that the EN distribution is the most suitable model (among those considered) to describe both data sets.

7.3 Application for regression models

We assess changes on the *oral health-related quality of life* (OHRQL) of schoolchildren. To that end, a follow-up exam of three years was made to evaluate the impact of caries incidence on the OHRQL of adolescents. The data were obtained from a study (for more details, see Paula et al. 2012) developed by the Department of Community Dentistry, Division of Health Education and Health Promotion, Piracicaba Dental School, University of Campinas-UNICAMP.

The variables employed are (for $i = 1, \dots, 291$):

- x_i : overall score of the OHRQL at time of follow up;
- v_{i1} : number of teeth decayed, missing and filled (TDMF) (0 = without TDMF increment; 1 = with TDMF increment).

We analyze these data based on the EN regression model

$$X_i = \beta_0 + \beta_1 v_{i1} + \sigma Z_i, \quad i = 1, \dots, 291,$$

where the errors Z_i 's are independent random variables having the EN($a, b, 0, 1$) distribution.

The gamma-normal (GN) (Lima et al. 2015) distribution extends the normal distribution and can be used to fit data that come from a distribution with heavy tails reducing the influence of aberrant observations. The GN density with location parameter $\mu \in \mathbb{R}$, dispersion parameter $\sigma > 0$ and shape parameter $a > 0$ takes the form

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

Further, the EN regression model is compared with the Exp-N, LTII-N, normal and GN regression models. Table 3 provides the MLEs of the parameters for the EN regression and these models.

Iterative maximization of the log-likelihood function (13) starts with initial values for β and σ taken from the fit of the classical regression model ($a = b = 1$). In general, all fitted regression models reveal that v_1 is significant at a 1% level of significance and that there is a significant difference between the levels of the numbers of teeth decayed, missing and filled. As expected, we find reciprocal relations between $\mu_i = \mathbb{E}(X_i)$ and v_{1i} in the EN, LTII-N, GN and normal regression models, except for the Exp-N regression (which-although well adjusted-does not seem to be a coherent model). On the other hand, based on the estimates of σ , the EN regression model reveals advantages in relation to the other models.

Table 2 The GoF measures of the fitted EN, SN, normal (N) and beta normal (BN) distributions to two real data sets

Measures	Models			
	EN	SN	N	BN
$\hat{\theta}_1$	<i>First application</i>			
Sd($\hat{\theta}$)	(0.2490, 3.2023, 1.7709, 0.9768)	(4.746, 1.109, 0.016)	(4.7619, 1.1094)	(991.2293, 54.4065, -23.4249, 17.3049)
W*	(< 0.0001, 0.0005, < 0.0001, < 0.0001)	(1.029, 0.039, 0.004)	(0.2420, 0.1711)	(1.1628, 1.3819, 0.4080, 0.1533)
A*	0.0630	1.5672	0.0800	0.4296
KS	0.3918	7.5418	0.4758	2.3953
	0.1484	1.0000	0.1452	1.0000
$\hat{\theta}_1$	<i>Second application</i>			
Sd($\hat{\theta}$)	(0.4804, 100.0000, -141.4568, 75.5640)	(0.0009, 166.0232, 32.7894)	(166.0746, 46.3746)	(1001.6514, 4.5535, -565.4361, 276.6320)
W*	(0.0126, 0.0005, 0.0002, 0.0001)	(0.1082, 0.0001, <0.001)	(5.6655, 4.0061)	(149.8685, 0.6375, 0.2980, 0.1719)
A*	0.0259	0.1037	0.1185	0.3018
KS	0.2428	0.8134	0.8942	1.8874
	0.0532	0.8519	0.0835	1.000

Table 3 MLEs, SEs in (·) and p -values between [-] for the EN, Exp-N, LTII-N, normal and GN regressions fitted to the OHRQL data

Model	Estimates				
	a	b	σ	β_0	β_1
EN	0.1058 (0.0156)	0.1624 (0.0139)	7.0973 (0.1076)	30.8602 (0.5989) [< 0.001]	-10.3707 (0.5728) [< 0.001]
Exp-N	1	0.0489 (0.0024)	9.6059 (0.1431)	64.4529 (2.2873) [< 0.001]	14.2817 (1.3420) [< 0.001]
LTII-N	0.1900 (0.0129)	1	10.5286 (0.2224)	4.6992 (1.5689) [0.0030]	-7.4972 (0.8623) [< 0.001]
Normal	1	1	19.3853 (0.8035)	25.6139 (1.9289) [< 0.0001]	-7.2402 (2.3872) [0.0026]
GN	5.1249 (2.9739)		23.5382 (1.4877)	-31.2214 (28.4676) [0.2737]	-7.0051 (2.2982) [0.0025]

The values of the AIC, Consistent Akaike Information Criterion (CAIC) and BIC to compare the fitted models are given in Table 4.

It is clear that the EN regression model outperforms the other regressions irrespective of the criteria and then we can conclude that the new regression model can be used effectively in the analysis of the current data set. A comparison of the proposed regression model with some of its sub-models using LR statistics is addressed in Table 5.

The figures in this table, specially the p -values, indicate that the EN regression model yields a better fit to these data than the other sub-models.

A graphical comparison among the fitted regression models is reported in Figure 4. The plots of these curves are the empirical cdf and the estimated cdf. Based on these plots, it is evident that the EN regression model provides a superior fit to the current data.

8 Conclusions

Flexible statistical distributions have been sought for describing data from practical situations in which the use of classical ones is not recommended. In this paper, we propose an extension of the normal distribution based on the exponentiated generalized family defined by Cordeiro et al. (2011), which adds two extra shape parameters to a baseline distribution. We provide some structural properties of the new *extended normal* (EN) distribution. The model parameters are estimated by maximum likelihood. The efficiency

Table 4 AIC, CAIC and BIC Statistics

Model	Statistic		
	AIC	CAIC	BIC
EN	2515.4	2515.6	2533.7
Exp-N	2725.8	2725.9	2740.5
LTII-N	2523.6	2523.8	2538.3
Normal	2557.2	2557.3	2568.3
GN	2545.0	2545.1	2559.7

Table 5 LR statistics

Model	Hypotheses	λ	p -value
EN vs Exp-N	$H_0 : a = 1$ vs $H_1 : H_0$ is false	212.4	< 0.001
EN vs LTII-N	$H_0 : b = 1$ vs $H_1 : H_0$ is false	10.2	0.0014
EN vs Normal	$H_0 : a = b = 1$ vs $H_1 : H_0$ is false	45.8	< 0.001

of this distribution is illustrated by means of two applications to real data sets. There is a clear evidence that the EN distribution outperforms the skew-normal distribution and can be a competitive alternative to the beta normal distribution. The classical regression model does not produce good results in many real problems, and for this reason several extensions have arisen in recent years. We propose a new regression model based on the EN distribution and prove its importance in real applications. This new regression model opens a wide range of research topics following the basic inference concepts of the normal linear regression model.

Endnote

¹Exton H. Handbook of hypergeometric integrals: theory, applications, tables, computer programs, 1978

Appendix A: Proof for the Theorem 3.1

We consider the power series

$$(1 - z)^b = \sum_{k=0}^{\infty} (-1)^k \binom{b}{k} z^k,$$

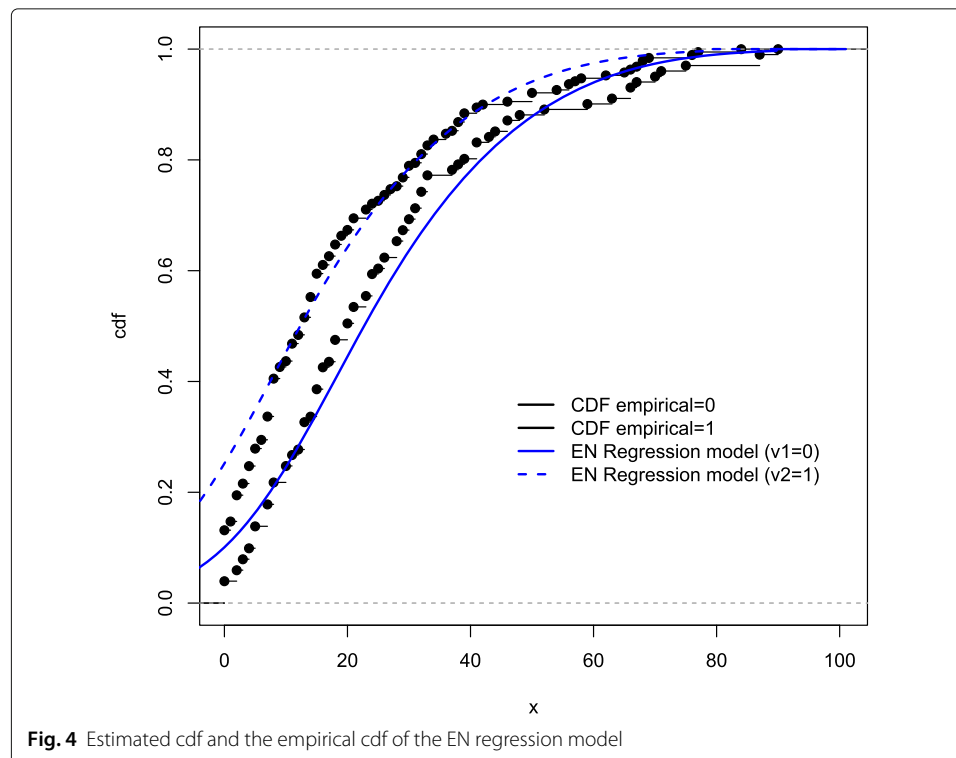


Fig. 4 Estimated cdf and the empirical cdf of the EN regression model

which holds for any real non-integer b and $|z| < 1$. Using this generalized binomial expansion twice in Eq. (1), we can write the EG-G cumulative distribution as

$$F(x) = \sum_{j=0}^{\infty} w_{j+1} H_{j+1}(x),$$

where $w_{j+1} = \sum_{m=1}^{\infty} (-1)^{j+m+1} \binom{b}{m} \binom{m}{j+1} a$ and $H_{j+1}(x)$ is the Exp-G cdf with power parameter $j + 1$. By differentiating the last equation, we obtain (7).

Appendix B: The Hessian matrix

The elements of the Hessian matrix are:

$$\begin{aligned}
 U_{aa} &= -\frac{n}{a^2} - (b-1) \sum_{i=1}^n \left\{ \frac{[1 - \Phi(\frac{x_i - \mu}{\sigma})]^a \log^2 [1 - \Phi(\frac{x_i - \mu}{\sigma})]}{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a} \right\} \\
 &\quad + (b-1) \sum_{i=1}^n \left\{ \frac{[1 - \Phi(\frac{x_i - \mu}{\sigma})]^{2a} \log^2 [1 - \Phi(\frac{x_i - \mu}{\sigma})]}{\{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a\}^2} \right\}, \\
 U_{ab} &= -\sum_{i=1}^n \left\{ \frac{[1 - \Phi(\frac{x_i - \mu}{\sigma})]^a \log [1 - \Phi(\frac{x_i - \mu}{\sigma})]}{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a} \right\}, \\
 U_{a\mu} &= \frac{1}{\sigma} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma})}{1 - \Phi(\frac{x_i - \mu}{\sigma})} \right\} \\
 &\quad - \frac{(b-1)}{\sigma} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma}) [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a}{[1 - \Phi(\frac{x_i - \mu}{\sigma})] \{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a\}} \right\} \\
 &\quad - \frac{a(b-1)}{\sigma} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma}) [1 - \Phi(\frac{x_i - \mu}{\sigma})]^{a-1} \log [1 - \Phi(\frac{x_i - \mu}{\sigma})]}{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a} \right\} \\
 &\quad + \frac{a(b-1)}{\sigma} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma}) [1 - \Phi(\frac{x_i - \mu}{\sigma})]^{2a-1} \log [1 - \Phi(\frac{x_i - \mu}{\sigma})]}{\{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a\}^2} \right\}, \\
 U_{a\sigma} &= -\frac{\mu}{\sigma^2} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma})}{1 - \Phi(\frac{x_i - \mu}{\sigma})} \right\} \\
 &\quad + \frac{\mu(b-1)}{\sigma^2} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma}) [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a}{[1 - \Phi(\frac{x_i - \mu}{\sigma})] \{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a\}} \right\} \\
 &\quad + \frac{a(b-1)\mu}{\sigma^2} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma}) [1 - \Phi(\frac{x_i - \mu}{\sigma})]^{a-1} \log [1 - \Phi(\frac{x_i - \mu}{\sigma})]}{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a} \right\} \\
 &\quad - \frac{a(b-1)\mu}{\sigma^2} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma}) [1 - \Phi(\frac{x_i - \mu}{\sigma})]^{2a-1} \log [1 - \Phi(\frac{x_i - \mu}{\sigma})]}{\{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a\}^2} \right\}, \\
 U_{bb} &= \frac{n}{b^2}, \\
 U_{b\mu} &= -\frac{a}{\sigma} \sum_{i=1}^n \left\{ \frac{\phi(\frac{x_i - \mu}{\sigma}) [1 - \Phi(\frac{x_i - \mu}{\sigma})]^{a-1}}{1 - [1 - \Phi(\frac{x_i - \mu}{\sigma})]^a} \right\}
 \end{aligned}$$

$$\begin{aligned}
 U_{b\sigma} &= -\frac{a\mu}{\sigma^2} \sum_{i=1}^n \left\{ \frac{\phi\left(\frac{x_i-\mu}{\sigma}\right) [1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^{a-1}}{1 - [1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^a} \right\}, \\
 U_{\mu\mu} &= -\left(\frac{a-1}{\sigma^2}\right) \sum_{i=1}^n \left\{ \frac{\phi'\left(\frac{x_i-\mu}{\sigma}\right)}{1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)} + \frac{\phi^2\left(\frac{x_i-\mu}{\sigma}\right)}{[1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^2} \right\} \\
 &\quad + \frac{1}{\sigma^2} \sum_{i=1}^n \left\{ \frac{\phi\left(\frac{x_i-\mu}{\sigma}\right) \phi''\left(\frac{x_i-\mu}{\sigma}\right) - \phi'^2\left(\frac{x_i-\mu}{\sigma}\right)}{\phi^2\left(\frac{x_i-\mu}{\sigma}\right)} \right\} \\
 &\quad - \frac{a(b-1)}{\sigma} \sum_{i=1}^n \left\{ \frac{1}{\sigma} \frac{\phi'\left(\frac{x_i-\mu}{\sigma}\right) [1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^{a-1}}{1 - [1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^a} \right. \\
 &\quad + \left. \left(\frac{a-1}{\sigma}\right) \frac{\phi^2\left(\frac{x_i-\mu}{\sigma}\right) [1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^{a-2}}{1 - [1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^a} \right. \\
 &\quad \left. - \frac{a}{\sigma} \frac{\phi^2\left(\frac{x_i-\mu}{\sigma}\right) [1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^{2a-2}}{\left\{1 - [1 - \Phi\left(\frac{x_i-\mu}{\sigma}\right)]^a\right\}^2} \right\}, \\
 U_{\mu\sigma} &= -\frac{1}{\sigma} U_{\mu} - \frac{\mu}{\sigma} U_{\mu\mu} \quad \text{and} \quad U_{\sigma\sigma} = -\frac{n}{\sigma^2} + \frac{\mu}{\sigma^2} U_{\mu} - \frac{\mu}{\sigma} U_{\mu\sigma}.
 \end{aligned}$$

Abbreviations

AIC: Akaike information criterion; BGN: Beta generalized normal; BIC: Bayesian information criterion; cdf: Cumulative distribution function; CAIC: Consistent Akaike information criterion; EG: Exponentiated generalized; Exp-G: Exponentiated-G; EN: Extended normal; Exp-N: Exponentiated-normal; GoF: Goodness-of-fit; GN: Gamma-normal; LTII-G: Lehmann type II-G; LTII-N: LTII-normal; LR: Likelihood ratio; MLEs: Maximum likelihood estimates; N: Normal; OHRQL: Oral health-related quality of life; pdf: Probability density function; SN: Skew-normal; TDMF: Teeth decayed, missing and filled

Acknowledgements

The authors would like to thank the financial support of CNPq and FACEPE, Brazil.

Authors' contributions

The authors, viz MCSL, GMC, EMMO and ADCN with the consultation of each other carried out this work and drafted the manuscript together. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Possible interested readers can contact authors.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Departamento de Estatística, Universidade Federal de Pernambuco, PE 50740-540 Recife, Brazil. ²Departamento de Ciências Exatas, ESALQ, Universidade de São Paulo, Piracicaba/SP, Brazil. ³Departamento de Estatística, Universidade Federal de Pernambuco, PE 50740-540 Recife, Brazil.

Received: 30 January 2019 Accepted: 28 May 2019

Published online: 08 June 2019

References

- Azzalini, A: A class of distributions which includes the normal ones. *Scand. J. Stat.* **12**, 171–178 (1984)
- Cintra, RJ, Cordeiro, GM, Nascimento, ADC: Beta generalized normal distribution with an application for SAR. *Image Process.* **48**, 1–16 (2013)
- Cordeiro, GM, Cunha, DCC, Ortega, EMM: The exponentiated generalized class of distributions. *J. Data Sci.* **11**, 777–803 (2013)
- Cordeiro, GM, Ortega, EMM, Silva, GO: The exponentiated generalized gamma distribution with application to lifetime data. *J. Stat. Comput. Simul.* **81**, 827–842 (2011)
- Cover, TM, Thomas, JA, Ortega, EMM: *Elements of Information Theory*. Wiley-Interscience, New York (1991)
- Doornik, JA: *An Object-Oriented Matrix Language Ox 5*. Timberlake Consultants Press, London (2007)
- Eugene, N, Lee, C, Famoye, F: Beta-normal distribution and its applications. *Commun. Stat.-Theory Methods.* **31**, 497–512 (2002)

- Frery, AC, Nascimento, ADC, Cintra, RJ: Analytic Expressions for Stochastic Distances Between Relaxed Complex Wishart Distributions. *IEEE Trans. Geosci. Remote Sens.* **52**, 1213–1226 (2014)
- Jones, M, Faddy, MJ: A skew extension of the t -distribution, with applications. *Biom. J.* **65**, 159–174 (2004)
- Lee, C, Famoye, F, Alzaatreh, AY: Methods for generating families of univariate continuous distributions in the recent decades. *Wiley Interdiscip. Rev. Comput. Stat.* **5**, 219–238 (2013)
- Lehmann, EL: The power of rank tests. *Ann. Math. Statist.* **24**, 23–43 (1953)
- Lima, MCS, Cordeiro, GM, Ortega, EMM: A new extension of the normal distribution. *J. Data Sci.* **3**, 385–408 (2015)
- Meeker, WQ, Escobar, L: *Statistical Methods for Reliability Data*. Wiley, New York (1998)
- Mudholkar, GS, Srivastava, DK: Exponentiated Weibull family for analyzing bathtub failure-real data. *IEEE Trans. Reliab.* **42**, 299–302 (1993)
- Nadarajah, S: Explicit expressions for moments of order statistics. *Statistics and Probability Letters.* **78**, 196–205 (2008)
- Nascimento, ADC, Cintra, RJ, Frery, AC: Hypothesis Testing in Speckled Data with Stochastic Distances. *IEEE Trans. Geosci. Remote Sens.* **48**, 373–385 (2010)
- Paula, JS, Oliveira, M, Soares, MSP, Chaves, MGAM, Mialhe, FL: Perfil Epidemiológico dos Pacientes Atendidos no Pronto Atendimento da Faculdade de Odontologia da Universidade Federal de Juiz de Fora. *Arquivos em Odontologia (UFMG)*. **48**, 257–262 (2012)
- Quang, HV: Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica.* **57**, 307–333 (1989)
- Rêgo, LC, Cintra, RJ, Cordeiro, GM: On some properties of the beta normal distribution. *Commun. Stat. - Theory Methods.* **41**, 3722–3738 (2012)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
