MATHEMATICAL COMMUNICATIONS
Math. Commun. **22**(2017), 53–73

A new extended mixture normal distribution

Božidar V. Popović[1,*], Gauss Cordeiro[2], Edwin M. Ortega[3] and Marcelino A. R. Pascoa[4]

[1] *Faculty of Philosophy, University of Montenegro, Danila Bojovića bb, 81 400 Nikšić, Montenegro*
[2] *Departamento de Estatística, Universidade Federal de Pernambuco, Av. Professor Moraes Rego, 1235-Cidade Universitária, Recife - PE, 50670-901, Brazil*
[3] *Departamento de Ciências Exatas, Universidade de São Paulo, Av. Pádua Dias, 11-Piracicaba/SP-CEP 13418-900, Brazil*
[4] *Departamento de Estatística, Universidade Federal de Mato Grosso, Av. Fernando Correa, 2 367-Boa Esperana, Cuiabá-MT 78.060-900, Brazil*

Received December 21, 2015; accepted May 23, 2016

Abstract. The normal distribution is the most important model in statistics for analysis of continuous data. We propose a new distribution, called the extended mixture normal distribution, based on a linear mixture model. We obtain explicit expressions for the ordinary and incomplete moments, generating and quantile functions, mean deviations and two measures of entropy. The maximum likelihood and Bayesian methods are used to estimate the model parameters. We prove empirically that the new distribution can be a better model than the normal and other classical distributions by means of an application to real data.
AMS subject classifications: 60E05, 62F15, 62P12

Key words: extended normal, generating function, mean deviation, mixture normal, moment, normal distribution, quantile expansion

1. Introduction

Let $\phi(x)$ be the standard normal (SN) probability density function (pdf) given by

$$\phi(x) = \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{x^2}{2}}, \tag{1}$$

for $x \in \mathbb{R}$. We define the standard *extended normal* (EN) density function (for $r = 0, 1, \ldots$) by

$$\phi_r(x) = c_r\, x^{2r}\, \phi(x), \tag{2}$$

where

$$c_r = \frac{\sqrt{2\pi}}{2^{(2r+1)/2}\, \Gamma\left(\frac{2r+1}{2}\right)}$$

*Corresponding author. *Email addresses:* `bozidarpopovic@gmail.com` (B. V. Popović), `gausscordeiro@uol.com.br` (G. Cordeiro), `edwin@usp.br` (E. M. Ortega), `marcelino.pascoa@gmail.com` (M. A. R. Pascoa)

and $\Gamma(\cdot)$ is the gamma function. It is easy to check that $\phi_r(x)$ is a genuine density function in $\mathbb{R}$. Here, $c_0 = 1$ and $\phi_0(x) = \phi(x)$ and $c_1 = 1$ and $\phi_1(x) = x^2\,\phi(x)$.

The new parameter r is really a shape parameter. The plots in Figure 1 reveal that the spread of the two modes of the EN pdf increases when r increases.

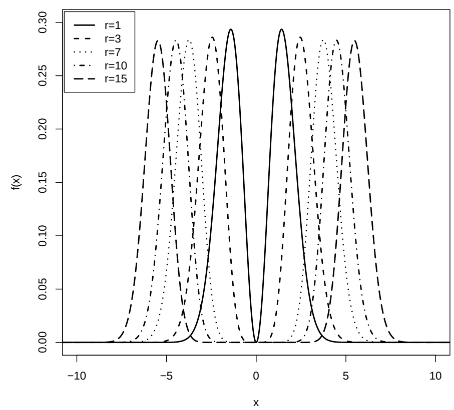


Figure 1: *Plots of the EN density function for some parameter values*

The EN model is not very flexible even with the additional parameter r, and then we construct a linear mixture distribution.

We define the (standard) *extended mixture normal* (EMN) density by

$$g(x; r, \alpha) = (1-\alpha)\,\phi(x) + \alpha\,\phi_r(x), \tag{3}$$

where $\alpha \in (0, 1)$ and $x \in \mathbb{R}$. Clearly, $g(x; r, \alpha)$ is a symmetric density function. For $r = 1$, equation (3) reduces to the *symmetric component normal* density function, namely $g(x; 1, \alpha) = [(1-\alpha) + \alpha\,x^2]\,\phi(x)$. Further, $g(x; r, 0) = \phi(x)$.

Yakowitz and Spragins (1968) demonstrated that a finite mixture is identifiable if a relation of the type $g(x; r_1, \alpha_1) = g(x; r_2, \alpha_2)$ implies $r_1 = r_2$ and $\alpha_1 = \alpha_2$. From the definition given by (3), it is easy to prove that model (3) is identifiable.

Hereafter, let $X \sim$EN(r, α) be a random variable having density function (3) and $Z \sim$N$(0, 1)$. The cumulative distribution function (cdf) of X is given by

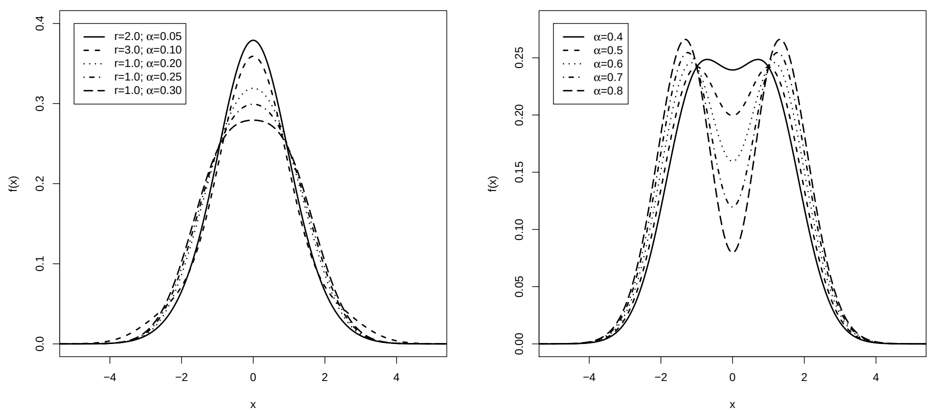
$$G(x; r, \alpha) = (1-\alpha)\Phi(x) + \frac{\alpha\,c_r}{2^{1-r}}\left[\Gamma\left(r + \frac{1}{2}\right) + \gamma\left(r + \frac{1}{2}, \frac{x^2}{2}\right)\right], \tag{4}$$

where $\Phi(x)$ is the standard normal cdf and $\gamma(a, z) = \int_0^z t^{a-1}\,\mathrm{e}^{-t}\,\mathrm{d}t$ is the incomplete gamma function. If $Y = \mu + \sigma X$, then Y has density

$$g(y; r, \alpha, \mu, \sigma) = \frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\left[(1-\alpha) + \alpha\,c_r\left(\frac{y-\mu}{\sigma}\right)^{2r}\right], \tag{5}$$

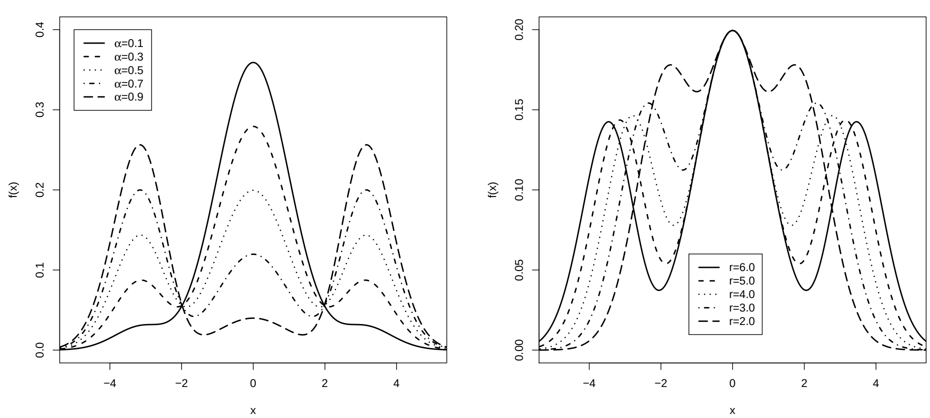
where $y \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, $r = 0, 1, 2, \ldots$ and $0 < \alpha < 1$ are shape parameters. A random variable Y having density function (5) is denoted by $Y \sim$ EMN(r, α, μ, σ). For $\mu = 0$ and $\sigma = 1$, we obtain (3).

Figures 2 and 3 display some plots of the EMN density for selected values of r and α with μ and σ fixed. Figure 2a reveals that this density function is unimodal when α increases ($\mu = 0$ and $\sigma = 1$). For lower values of α, the maxima of the EMN density function increases. For fixed values $r = 1$, $\mu = 0$ and $\sigma = 1$, this density function possesses bimodal characteristics (see Figure 2b).



(a) *α increasing, $\mu = 0$ and $\sigma = 1$* (b) *$r = 1$, $\mu = 0$ and $\sigma = 1$*

Figure 2: *Plots of EMN density functions*



(a) *α increasing, $r = 5$, $\mu = 0$ and $\sigma = 1$* (b) *r decreasing, $\alpha = 0.5$, $\mu = 0$ and $\sigma = 1$*

Figure 3: *Plots of the density functions*

The lower values of the EMN pdf are most influenced by the greater values of α. For some values of the parameters r, μ and σ, we note that this density possesses three modes (see Figures 3a and 3b). When the parameter α increases then the maxima that correspond to $\mu = 0$ decrease. On the other hand, when r decrease the maxima corresponding to $x = 0$ increase. At the end, we can conclude that the parameters r and α have a strong influence on the shape of the EMN density function.

We sometimes omit the dependence of the pdf and cdf of the EMN distribution on the parameters. It is worth to mention that this distribution belongs to the Gram-Charlier series type. One can verify that the density $g(x)$ is a specific rth order Gram-Charlier series, i.e., the series that represents $g(x)$ stops at degree r. The relationships between the coefficients of this series and the moments of random variables are discussed in Chapter 12, Eq. (38), p. 16 in [8].

We introduce a distribution that extends the normal distribution meaning that $|x|$ possess the chi-squared distribution with $\nu = 2r + 1$ degrees of freedom (see Chapter 17, Eq.(62) in [8]). In the same chapter, moment recursions and cumulants are given by Eqs. (63) and (64) for all integer values of r.

Siddiqui and Weiss (1963) and Krysicki (1963) studied some properties of mixtures of chi-squared distributions with one degree of freedom. Suppose that the random variable X possesses the EMN distribution and defines the random variable $Y = X^2$. The density of Y is given by

$$f(y) = \frac{d}{dy} P\left(-\sqrt{y} \le X \le \sqrt{y}\right).$$

After some calculation, we can write

$$f(y) = (1-\alpha)\,\pi_{\chi^2(1)}(y) + \alpha\,\pi_{\chi^2(2r+1)}(y). \tag{6}$$

Then, the pdf of Y can be expressed as a mixture of chi-square densities with one and $2r + 1$ degrees of freedom.

By using the characteristic function corresponding to (6), we can study the random variable $Z = \sum_{i=1}^{n} X_i$, where the X_i's are iid random variables with pdf (3). Some straightforward calculations lead to the following representation for the pdf of Z

$$f(z) = \sum_{k=0}^{n} \binom{n}{k} (1-\alpha)^k\,\alpha^{n-k}\,\pi_{\chi^2(n+2r(n-k))}(z).$$

So, the pdf of Z can be written as a linear combination of chi-square densities with $n + 2r(n-k)$ degrees of freedom (for $k = 0, 1, \ldots, n$).

The rest of the paper is organized as follows. A range of mathematical properties of the proposed mixture distribution is explored in Sections 2 to 9. Estimation of model parameters by the maximum likelihood method is addressed in Section 10. Bayesian analysis is investigated in Section 11. An application to a real data set is given in Section 12. Finally, some conclusions are given in Section 13.

2. Shape characteristics

We examine shape characteristics of the pdf of X. The first derivative of (3) is

$$g'(x; r, \alpha) = -\frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\, \mathrm{x}\, \left(\alpha\, \mathrm{c_r}\, \mathrm{x}^{2\mathrm{r}} - 2\mathrm{r}\alpha\, \mathrm{c_r}\, \mathrm{x}^{2\mathrm{r}-2} + 1 - \alpha\right). \tag{7}$$

There may be more than one root of equation (7). If $x = x_0$ is its root, then it corresponds to a local maximum, a local minimum or an inflexion point depending on whether $\lambda(x_0) < 0$, $\lambda(x_0) > 0$ or $\lambda(x_0) = 0$, where $\lambda(x) = \dfrac{d^2\, g(x; r, \alpha)}{dx^2}$ is given by

$$\lambda(x) = \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{-\frac{\mathrm{x}^2}{2}} \left\{\alpha\, \mathrm{c_r}\, \mathrm{x}^{2\mathrm{r}-2} \left[\mathrm{x}^4 - \mathrm{x}^2 + 2\mathrm{r}(2\mathrm{r}-1)\right] - (1-\alpha)(\mathrm{x}^2+1)\right\}.$$

3. Useful expansion

A power series expansion for the EMN cdf can be easily derived from the power series for the standard normal cdf and for the incomplete gamma function. For $n \geq 0$, let

$$a_n = \frac{(-1)^n\,(1-\alpha)}{2^{n+1}\,(2n+1)\,n!} \quad \text{and} \quad b_n = \frac{(-1)^n\,\alpha\, c_r}{(2n+2r+1)\,2^n\,\sqrt{2\pi}\,n!}.$$

We can rewrite $G(x; r, \alpha)$ as

$$G(x; r, \alpha) = \frac{1}{2} + \sum_{n\geq 0} a_n\; x^{2n+1} + \sum_{n\geq 0} b_n\; x^{2(n+r)+1}.$$

We can combine the two power series in just one power series, whose coefficients are conveniently defined by $d_n = a_n$ for $n = 0, 1, \ldots, r-1$ and $d_n = a_n + b_{n-r}$ for $n = r, r+1, r+2, \ldots$. Then,

$$G(x; r, \alpha) = \sum_{n\geq 0} e_n\; x^n, \tag{8}$$

where $e_0 = 1/2$, $e_1 = d_0$ and, for $n = 1, 2, 3, \ldots$: $e_{2n} = 0$ and $e_{2n+1} = d_n$. Hereafter, we can denote $G(x; r, \alpha)$ by $G(x)$. Equation (8) is the main result of this section.

4. Moments

The moments $E(X^\beta)$ (for $\beta > -1$) are given by

$$\mu'_\beta = E(X^\beta) = \frac{1}{\sqrt{2\pi}} \left\{(1-\alpha)\, d_1 + \alpha\, c_r\, d_2\right\}, \tag{9}$$

where

$$d_1 = \int_{-\infty}^{+\infty} y^\beta\, \mathrm{e}^{-y^2/2}\, \mathrm{d}y \quad \text{and} \quad d_2 = \int_{-\infty}^{+\infty} y^{\beta+2r}\, \mathrm{e}^{-y^2/2}\, \mathrm{d}y.$$

If β is an odd number, $d_1 = d_2 = 0$. If β is not an odd number, we use equation (3.462.3) by Gradshteyn and Ryzhik (2007) to obtain $d_1 = \sqrt{2\pi}\ \mathrm{i}^{-\beta}\, D_\beta(0)$ and $d_2 = \sqrt{2\pi}\ \mathrm{i}^{-\beta-2r}\, D_{\beta+2r}(0)$, where $\mathrm{i} = \sqrt{-1}$, $D_\nu(z)$ is the parabolic cylinder function defined by

$$\begin{aligned} D_\nu(z) = 2^{\nu/2}\, \mathrm{e}^{\mathrm{z}^2/4} \Bigg[& \frac{\sqrt{\pi}}{\Gamma[(1-\nu)/2]}\, {}_1\mathrm{F}_1\left(-\frac{\nu}{2}; \frac{1}{2}; \frac{\mathrm{z}^2}{2}\right) - \frac{\mathrm{z}\sqrt{2\pi}}{\Gamma(-\nu/2)} \\ & \times {}_1F_1\left(\frac{1-\nu}{2}; \frac{3}{2}; \frac{z^2}{2}\right)\Bigg], \end{aligned} \tag{10}$$

${}_1F_1(\cdot;\cdot;\cdot)$ is the confluent hypergeometric function of the first kind given by

$$ {}_1F_1(a;b;z) = \sum_{j\geq 0} \frac{(a)_j}{(b)_j} \frac{z^j}{j!}, $$

and $(a)_j = \Gamma(a+j)/\Gamma(a)$ denotes the Pochhammer symbol. It follows that $D_\nu(0) = 2^{\nu/2}\sqrt{\pi}/[\Gamma\left(\frac{1-\nu}{2}\right)]$. From the last equation, d_1 and d_2 can be expressed as

$$ d_1 = \frac{2^{\frac{\beta+1}{2}}\,\pi}{\mathrm{i}^\beta\,\Gamma\left(\frac{1-\beta}{2}\right)} \quad \text{and} \quad d_2 = \frac{2^{\frac{\beta+2r+1}{2}}\,\pi}{\mathrm{i}^{\beta+2r}\,\Gamma\left(\frac{1-\beta-2r}{2}\right)}. $$

Based on these expressions, equation (9) reduces to

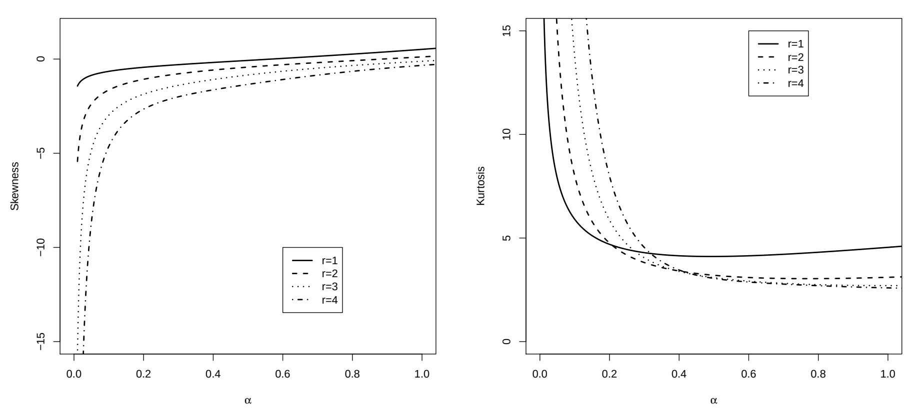
$$ \mu'_\beta = \frac{\sqrt{2^\beta\pi}}{\mathrm{i}^\beta}\left[\frac{1-\alpha}{\Gamma\left(\frac{1-\beta}{2}\right)} + \frac{\alpha\,2^r\,c_r}{\Gamma\left(\frac{1-\beta-2r}{2}\right)}\right]. \tag{11} $$

When $\alpha = 0$ in equation (11), the nth moment ($n \in \mathbb{N}$) of the SN distribution follows as a special case: $E(Z^n) = (2n-1)!!$ for n even and $E(Z^n) = 0$ for n odd, where $p!! = p(p-2)\ldots 31$ (for p odd).

If $n \in \mathbb{N}$, from (11) we obtain: $\mu'_n = E(X^n) = 0$ for n odd, and for n even

$$ \mu'_n = \frac{1}{\sqrt{2\pi}}\left[(1-\alpha)(2n-1)!! + \alpha\,c_r\,(2n+4r-1)!!\right]. \tag{12} $$

The skewness and kurtosis measures of X can be determined from ordinary moments using well-known relationships. Plots of these quantities for some choices of r as functions of α, by fixing $\mu = 0$, $\sigma = 1$, are displayed in Figure 4.



(a) (b)

Figure 4: *Skewness and kurtosis of X as functions of α for some values of r*

These plots reveal that the skewness and kurtosis of the EMN distribution are quite flexible.

For empirical purposes, the shapes of many distributions can be usefully described by incomplete moments. These moments play an important role in measuring inequality, for example, income quantiles and Lorenz and Bonferroni curves, which depend upon the first incomplete moment of the distribution. The nth incomplete moment of X is given by

$$m_n(z) = \frac{1-\alpha}{\sqrt{2\pi}} \int_{-\infty}^{z} x^n \, \mathrm{e}^{-\mathrm{x}^2/2} \, \mathrm{dx} + \frac{\alpha \, \mathrm{c_r}}{\sqrt{2\pi}} \int_{-\infty}^{\mathrm{z}} \mathrm{x}^{\mathrm{n+2r}} \, \mathrm{e}^{-\mathrm{x}^2/2} \, \mathrm{dx} \, .$$

Further, we can determine the integrals in this equation by setting $x^2/2 = u$. We can write

$$\begin{aligned} m_n(z) =& \frac{(1-\alpha)}{\sqrt{2\pi}} \, 2^{\frac{n-1}{2}} \left[\Gamma\left(\frac{n+1}{2}\right) + \gamma\left(\frac{n+1}{2}, \frac{z^2}{2}\right)\right] \\ &+ \frac{\alpha \, c_r}{\sqrt{2\pi}} \, 2^{\frac{n+2r-1}{2}} \left[\Gamma\left(\frac{n+1+2r}{2}\right) + \gamma\left(\frac{n+1+2r}{2}, \frac{x^2}{2}\right)\right]. \end{aligned} \tag{13}$$

Next, we obtain the probability weighted moments (PWMs) of X. They cover the summarization and description of theoretical probability distributions. These moments can estimate the parameters of a distribution whose inverse cannot be expressed explicitly. The $(s,p)th$ PWM of X is formally defined as $\tau_{s,p} = E[X^s \, G(X)^p] = \int_{-\infty}^{\infty} x^s \, G(x)^p \, g(x) \, \mathrm{d}x$. For calculating $\tau_{s,p}$ we use an equation of Gradshteyn and Ryzhik (2007, Section 0.314) for a power series raised to a positive integer power p (for $p = 1, 2, \ldots$)

$$\left(\sum_{n\geq 0} a_n \, x^n\right)^p = \sum_{n \geq 0} c_{p,n} \, x^n, \tag{14}$$

where the coefficients $c_{p,n}$ can be determined from the recurrence relation (for $k \geq 1$ with $c_{p,0} = a_0^p$)

$$c_{p,k} = (k \, a_0)^{-1} \sum_{j=1}^{m} [j(p+1) - k] \, a_k \, c_{p,k-m}.$$

Using (3), (8) and (14), we have $\tau_{s,p} = \sum_{n\geq 0} f_{p,n} \, (a_1 + a_2)$, where

$$a_1 = \frac{1-\alpha}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^{s+n} \, \mathrm{e}^{-\mathrm{x}^2/2} \, \mathrm{dx} \quad \text{and} \quad \mathrm{a}_2 = \frac{\alpha \, \mathrm{c_r}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{x}^{\mathrm{s+n+2r}} \, \mathrm{e}^{-\mathrm{x}^2/2} \, \mathrm{dx},$$

and the coefficients $f_{p,n}$ are given by $f_{p,0} = e_0^p = 2^{-p}$ and (for $n \geq 1$) $f_{p,n} = 2 \, n^{-1} \sum_{j=1}^{m} [j(p+1) - n] \, e_j \, f_{p,n-j}$.

Using parametric integration one can prove that (for $n \in \mathbb{N}$) holds

$$\int_{-\infty}^{+\infty} x^n \, \mathrm{e}^{-\frac{\mathrm{x}^2}{2}} \, \mathrm{dx} = 1 \cdot 3 \cdot 5 \ldots \cdot (\mathrm{n}-1) \sqrt{2\pi}.$$

So, using the last equation we obtain

$$a_1 + a_2 = 3 \times 5 \times \ldots \times (s+n-1)\,[1-\alpha+\alpha\, c_r(s-n)\ldots(s+n+2r-1)]\,.$$

Hence, $\tau_{s,p}$ can be expressed as

$$\tau_{s,p} = \sqrt{\pi} \sum_{n\geq 0} 3\times\ldots\times(s+n-1)\,[1-\alpha+\alpha\, c_r(s-n)\ldots(s+n+2r-1)]\; f_{p,n}\,. \quad (15)$$

Other kinds of moments such as the factorial and L-moments may also be obtained in closed form, but we consider only the previous moments for reasons of space. Equations (11), (12), (13) and (15) are the main results of this section.

5. Generating function

The moment generation function (mgf) of the random variable X is defined as $M(t) = E(\mathrm{e}^{\mathrm{t\,X}})$. Here, we provide two explicit expressions for $M(t)$. First, using a result in Gradshteyn and Ryzhik (2007, equation 3.462.3), we obtain

$$M(t) = (1-\alpha)\,\mathrm{e}^{-\mathrm{t}^2/2} + \frac{\alpha\,\mathrm{c_r}}{\sqrt{2\pi}}\,\mathrm{M_2(t)} = \mathrm{e}^{-\mathrm{t}^2/4}\left[(1-\alpha)\,\mathrm{e}^{-\mathrm{t}^2/4} + \alpha\,\mathrm{c_r}\,\mathrm{D_{2r}(it)}\right], \quad (16)$$

where $M_2(t) = \int_{-\infty}^{+\infty} x^{2r}\,\mathrm{e}^{-\frac{\mathrm{x}^2}{2}+\mathrm{tx}}\,\mathrm{d\,x}$ and $D_{2r}(it)$ is obtained from (10).

Second, setting $x - t = 2v$ and using the binomial expansion, $M_2(t)$ turns out to be

$$M_2(t) = \mathrm{e}^{\mathrm{t}^2/2} \sum_{\mathrm{k}=0}^{2\mathrm{r}} \binom{2\mathrm{r}}{\mathrm{k}} 2^{\mathrm{k}+1}\,\mathrm{t}^{2\mathrm{r}-\mathrm{k}} \int_{-\infty}^{+\infty} \mathrm{v}^{\mathrm{k}}\,\mathrm{e}^{-\mathrm{v}^2}\,\mathrm{dv}.$$

From equation (3.462.4) in Gradshteyn and Ryzhik (2007), we obtain $M_2(t)$ and then the last equation becomes

$$M(t) = (1-\alpha)\,\mathrm{e}^{-\mathrm{t}^2/2} + \alpha\,\sqrt{\pi}\,\mathrm{c_r}\,\mathrm{e}^{\mathrm{t}^2/2} \sum_{\mathrm{k}=0}^{2\mathrm{r}} \binom{2\mathrm{r}}{\mathrm{k}} \frac{\mathrm{t}^{2\mathrm{r}-\mathrm{k}}\,2^{\mathrm{k}+1/2}}{\mathrm{i}^{\mathrm{k}}\,\Gamma\left(\frac{1-\mathrm{k}}{2}\right)}. \quad (17)$$

Equations (16) and (17) are the main results of this section.

6. Quantile expansion

First, we invert $G(x) = G(x; r, \alpha)$ in (4) to obtain a power series expansion for the EMN quantile function (qf), say $x = Q(u)$. We shall use the Lagrange theorem to derive a power series for $Q(u)$. We assume that the power series expansion holds

$$w = G(x) = w_0 + \sum_{n=1}^{\infty} g_n\,(x-x_0)^n, \qquad g_1 = G'(x) \neq 0,$$

where $G(x)$ is analytic at a simple x_0-point. Then, the inverse function $x = Q(u) = G^{-1}(u)$ exists and it is single-valued in the neighborhood of the point $u = u_0$. The power series inverse $x = Q(u)$ is given by Markushevich (1965, vol. 2, p. 88)

$$x = Q(u) = x_0 + \sum_{n=1}^{\infty} h_n \, (u - u_0)^n,$$

where

$$h_n = \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left\{ [\psi(x)]^n \right\} \Big|_{x=x_0} \quad \text{and} \quad \psi(x) = \frac{x - x_0}{G(x) - x_0}.$$

From (8) we can write

$$G(x) = \frac{1}{2} + x \, \left(e_1 + e_2 x + e_3 x^2 + \ldots \right).$$

Setting $m_n = e_{n+1}$ for $n = 0, 1, 2, \ldots$, we obtain $G(x) = 0.5 + x \sum_{n=0}^{\infty} m_n \, x^n$, where $m_0 = d_0$, $m_1 = 0$, $m_2 = d_1$, $m_3 = 0$, $m_4 = d_2$, and so on. Setting $x_0 = 0$ and $u_0 = 1/2$, we define

$$\psi(x) = \frac{x}{G(x) - \frac{1}{2}} = \frac{1}{\sum_{n=0}^{\infty} m_n \, x^n}.$$

The inverse of the power series $\sum_{n=0}^{\infty} m_n \, x^n$ follows from a result by Gradshteyn and Ryzhik (2007, equation 0.313)

$$\psi(x) = \frac{1}{\sum_{n=0}^{\infty} m_n \, x^n} = \frac{1}{m_0} \sum_{n=0}^{\infty} p_n \, x^n,$$

where the coefficients p_n can be determined from $p_n = -m_0^{-1} \sum_{k=1}^{n} m_{k+1} \, p_{n-k}$, $n \geq 1$ with $p_0 = 1$. Then, $\psi(x)^n = \left(\frac{1}{m_0} \sum_{i=0}^{\infty} p_i \, x^i \right)^n$.

Using Equation (14), we can write $\psi(z)^n = \frac{1}{m_0^n} \sum_{i=0}^{\infty} q_{n,i} \, x^i$, where the coefficients $q_{n,i}$ (for $i = 1, 2, \ldots$) are given by $q_{n,i} = i^{-1} \sum_{m=1}^{i} [m\,(n+1) - i] \, p_m \, q_{n,i-m}$, and $q_{n,0} = p_0^n = 1$. The quantity $q_{n,i}$ can be determined from $q_{n,0}, \ldots, q_{n,i-1}$ and therefore from $p_0, \ldots, p_i$ by programming numerically our expansions in any algebraic or numerical software.

The derivative of order $(n-1)$ of $\psi(x)^n$ gives

$$h_n = \frac{1}{n!} \frac{d^{n-1}}{dx^{n-1}} \left\{ [\psi(x)]^n \right\} \Big|_{x=0} = \frac{q_{n,n-1}}{n \, m_0^n}.$$

Hence, a power series for the ENN qf reduces to

$$Q(u) = \sum_{n=1}^{\infty} b_n \, \left(u - \frac{1}{2} \right)^n, \tag{18}$$

where $b_n = q_{n,n-1}/(n \, m_0^n)$. Equation (18) can be used to derive alternative explicit expressions for the ordinary and incomplete moments, mgf and mean deviations of

X. For example, setting $c_n = b_{n+1} = q_{n+1,n}/[(n+1)\, m_0^{n+1}]$ for $n = 0, 1, \ldots$, we can obtain (for $p = 1, 2, ...$)

$$E(X^p) = \int_0^1 \left(u - \frac{1}{2}\right)^p \left[\sum_{n=0}^{\infty} c_n \left(u - \frac{1}{2}\right)^n\right]^p \mathrm{du}.$$

By using (14) and interchanging the integral with the sum

$$E(X^p) = \sum_{n=0}^{\infty} d_{p,n} \int_0^1 \left(u - \frac{1}{2}\right)^{n+p} \mathrm{du} = \sum_{\mathrm{n}=0}^{\infty} \frac{[1 - (-1)^{\mathrm{n+p+1}}]\, \mathrm{d_{p,n}}}{(\mathrm{n+p+1})\, 2^{\mathrm{n+p+1}}}, \tag{19}$$

where the quantities $d_{p,n}$ can be obtained from the recurrence relation (for $n \geq 1$ with $d_{p,0} = c_0^p$) $d_{p,n} = (n\, c_0)^{-1} \sum_{j=1}^{n} [j\,(p+1) - n]\, c_j\, d_{p,n-j}$. Equations (18) and (19) are the main results of this section.

7. Mean deviations

The mean deviations about the mean ($\delta_1 = E(|X - \mu_1'|)$) and the median ($\delta_2 = E(|X - M|)$) of X can be expressed as

$$\delta_1 = 2\mu_1'\, G(\mu_1') - 2m_1(\mu_1') \qquad \text{and} \qquad \delta_2 = \mu_1' - 2m_1(M), \tag{20}$$

respectively, where

$$\mu_1' = E(X) = \frac{1}{\sqrt{2\pi}} \left[(1-\alpha) + \alpha\, c_r\, (4r+1)!!\right],$$

$G(\cdot)$ is obtained from (4), M is the median determined by the nonlinear equation

$$(1-\alpha)\Phi(M) + \frac{\alpha\, c_r}{2^{1-r}} \left[\Gamma\left(r + \frac{1}{2}\right) + \gamma\left(r + \frac{1}{2}, \frac{M^2}{2}\right)\right] = 1/2$$

and using (13) with $n = 1$, we obtain

$$m_1(z) = \frac{(1-\alpha)}{\sqrt{2\pi}} \left[1 + \gamma\left(1, \frac{z^2}{2}\right)\right] + \frac{\alpha\, 2^r\, c_r}{\sqrt{2\pi}} \left[r! + \gamma\left(r+1, \frac{z^2}{2}\right)\right].$$

A useful application of mean deviations refers to the Lorenz and Bonferroni curves. They are important in fields like economics, reliability, demography, insurance and medicine. For a given probability π, they are defined by $L(\pi) = m_1(q)/\mu_1'$ and $B(\pi) = m_1(q)/(\pi\, \mu_1')$, respectively, where $q = Q(\pi) = F^{-1}(\pi)$ can be determined from (18).

8. Entropies

An entropy is a measure of variation or uncertainty of a random variable X. Two popular entropy measures are the Rényi and Shannon entropies (Shannon, 1951; Rényi, 1961). The Rényi entropy of a random variable with pdf $g(x)$ is defined by

$$I_R(\gamma) = \frac{1}{1-\gamma} \log\left(\int_{-\infty}^{\infty} g^{\gamma}(x) dx\right)$$

for $\gamma > 0$ and $\gamma \neq 1$.

Assuming $\gamma = n = 2, 3, \ldots$ and using the binomial expansion, the last equation can be expressed as

$$\begin{aligned} I_R(n) &= \frac{1}{1-n} \log \left\{ \left(\frac{1}{\sqrt{2\pi}} \right)^n \sum_{k=0}^{n} \binom{n}{k} (\alpha\, c_r)^k (1-\alpha)^{n-k} \int_{-\infty}^{+\infty} x^{2kr}\, \mathrm{e}^{-\frac{\mathrm{n\,x}^2}{2}} \mathrm{dx} \right\} \\ &= \frac{1}{1-n} \left\{ J_n + \log \left[\sum_{k=0}^{n} \binom{n}{k} (\alpha c_r)^k (1-\alpha)^{n-k} \int_{0}^{+\infty} x^{2kr}\, \mathrm{e}^{-\frac{\mathrm{n\,x}^2}{2}} \mathrm{dx} \right] \right\}, \end{aligned}$$

where $J_n = -\frac{n}{2}\log(2\pi) + \log(2)$ and

$$\begin{aligned} I_R(n) = &\frac{1}{1-n} \Bigg\{ J_n + \log \Bigg[\sum_{k=0}^{n} \binom{n}{k} (\alpha c_r)^k (1-\alpha)^{n-k} \left(\frac{2}{n} \right)^{kr+\frac{1}{2}} \\ &\times \Gamma\left(kr + \frac{1}{2} \right) \Bigg] \Bigg\}. \end{aligned} \tag{21}$$

We can write $I_R(\gamma) = (1-\gamma)^{-1} E\{g(X)^{\gamma-1}\}$. Let $\delta = E(X)$. For γ real positive, we have

$$E\{g(X)^{\gamma-1}\} = \delta^{\gamma-1} E\left(\{1 + \theta\,[g(X) - \delta]\}^{\gamma-1} \right),$$

where $\theta = \delta^{-1}$. From the generalized binomial expansion, we can write

$$\{1 + \theta\,[g(X) - \delta]\}^{\gamma-1} = 1 + \sum_{n=1}^{\infty} \frac{\theta^n P_n}{n!} [g(X) - \delta]^n,$$

where $P_n = \prod_{j=0}^{n-1} (\gamma - 1 - j)$. Further, we have

$$E\{g(X)^{\gamma-1}\} = \delta^{\gamma-1} \left(1 + \sum_{n=2}^{\infty} \frac{\theta^n P_n}{n!} E\{[g(X) - \delta]^n\} \right). \tag{22}$$

We now have to determine $E\{[g(X)]^n\}$ for $n \geq 2$. From (3) and using the binomial expansion, we obtain

$$\rho_n = E\{[g(X)]^n\} = \sum_{m=0}^{n} \binom{n}{m} (\alpha\, c_r)^m (1-\alpha)^{n-m}\, \psi_{m,n},$$

where $\psi_{m,n} = E\{X^{2mr}\, \phi(X)^n\}$. Then,

$$\psi_{m,n} = 2 \int_0^{\infty} x^{2mr}\, \phi(x)^{n+1}\, \mathrm{dx}.$$

Setting $(n+1)x^2/2 = z$, we can easily write ρ_n as

$$\rho_n = \sum_{m=0}^{n} \frac{2^{mr-n/2}\,(1-\alpha)^{n-m}\,(\alpha\, c_r)^m}{\sqrt{\pi^{n+1}\,(n+1)^{2mr+1}}} \binom{n}{m} \Gamma\left(mr + \frac{1}{2} \right).$$

By expanding the binomial term in (22), an explicit expression for $I_R(\gamma)$ follows as

$$I_R(\gamma) = (1-\gamma)^{-1}\,\delta^{\gamma-1}\,\left(1+\sum_{n=2}^{\infty}\frac{\theta^n\,P_n}{n!}\,\sum_{j=0}^{n}\binom{n}{j}\,(-\delta)^{n-j}\,\rho_j\right), \tag{23}$$

which holds for any γ real positive and $\gamma \neq 1$, where ρ_j is given before.

Next, the Shannon entropy of a random variable X is defined by $E\{-\log[g(X)]\}$. It is a special case of the Rényi entropy when $\gamma \uparrow 1$. Equation (21) is very complicated for limiting, and then we derive an explicit expression for the Shannon entropy from its definition. We can write

$$\begin{aligned} E\big\{-\log[g(X)]\big\} &= -\int_{-\infty}^{+\infty}\big[(1-\alpha)\phi(x)+\alpha c_r x^{2r}\phi(x)\big]\,\log\{(1-\alpha)\phi(x)\\ &\quad + \alpha c_r x^{2r}\phi(x)\}\,\mathrm{d}x\\ &= -2\int_{0}^{+\infty}\big[(1-\alpha)\phi(x)+\alpha c_r x^{2r}\phi(x)\big]\,\log\{(1-\alpha)\phi(x)\\ &\quad + \alpha c_r x^{2r}\phi(x)\}\,\mathrm{d}x \end{aligned}$$

and then

$$\begin{aligned} E\{-\log[g(X)]\} = &-2\left\{\frac{1-\alpha}{\sqrt{2\pi}}\int_0^{\infty}\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\left[\log\left(\frac{1}{\sqrt{2\pi}}\right)-\frac{\mathrm{x}^2}{2}\right]\mathrm{dx}\right.\\ &+\frac{\alpha c_r}{\sqrt{2\pi}}\int_0^{+\infty}x^{2r}\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\left[\log\left(\frac{1}{\sqrt{2\pi}}\right)-\frac{\mathrm{x}^2}{2}\right]\mathrm{dx}\\ &+\frac{1-\alpha}{\sqrt{2\pi}}\int_0^{+\infty}\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\log\left[(1-\alpha)+\alpha\mathrm{c_r}\,\mathrm{x}^{2\mathrm{r}}\right]\mathrm{dx}\\ &\left.+\frac{\alpha c_r}{\sqrt{2\pi}}\int_0^{+\infty}x^{2r}\,\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\log\left[(1-\alpha)+\alpha\mathrm{c_r}\,\mathrm{x}^{2\mathrm{r}}\right]\mathrm{dx}\right\}. \end{aligned} \tag{24}$$

The calculations of the four integrals in (24) are given in Appendix A. Equations (21), (23) and (24) are the main results of this section.

9. Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1,\ldots,X_n$ is a random sample from the EMN distribution. Let $X_{i:n}$ denote the ith order statistic. The pdf of $X_{i:n}$ can be expressed as

$$g_{i:n}(x) = K\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}\,g(x)\,G(x)^{j+i-1}, \tag{25}$$

where $K = n!/[(i-1)!\,(n-i)!]$. From equations (8) and (14), we can write

$$G(x;r,\alpha)^{j+i-1} = \sum_{p\geq 0} f_{j+i-1,p}\;x^p, \tag{26}$$

where $f_{j+i-1,p}$ is defined in Section 4. Substituting (26) into (25), we have

$$g_{i:n}(x) = K \sum_{j=0}^{n-i} \sum_{p \geq 0} (-1)^j f_{j+i-1,p} \binom{n-i}{j} x^p \left[(1-\alpha)\,\phi(x) + \alpha\, c_r\, x^{2r}\, \phi(x)\right].$$

Next, we can easily obtain the moments of the order statistics from (11) as

$$E(X_{i:n}^{\beta}) = K \sum_{j=0}^{n-i} \sum_{p \geq 0} (-1)^j f_{j+i-1,p} \binom{n-i}{j} J(\beta, p, \alpha, r), \tag{27}$$

where

$$J(\beta, p, \alpha, r) = \frac{\sqrt{2^{\beta+p}\,\pi}}{\mathrm{i}^{\beta+p}} \left[\frac{1-\alpha}{\Gamma\left(\frac{1-\beta-p}{2}\right)} + \frac{\alpha\, 2^r\, c_r}{\mathrm{i}^{2r}\,\Gamma\left(\frac{1-\beta-p-2r}{2}\right)}\right].$$

Equation (27) is the main result of this section. Consider that $\beta = n$ is a positive integer. If $n+p$ is odd, $J(\beta, p, \alpha, r)$ vanishes, whereas if $n+p$ is even, it reduces to

$$J(\beta, p, \alpha, r) = \frac{1}{\sqrt{2\pi}} \left[(1-\alpha)(2n+2p-1)!! + \alpha\, c_r\, (2n+2p+4r-1)!!\right].$$

10. Maximum likelihood estimation

The parameters of the EMN distribution are estimated by maximum likelihood from complete samples only. Let $y_1, \ldots, y_n$ be a random sample of size n from the EMN(r, α, μ, σ) distribution. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (\alpha, \mu, \sigma)^T$ follows from (5) as

$$l(\boldsymbol{\theta}) = -n \log(\sigma) + \sum_{i=1}^{n} \log\left[\phi\left(\frac{y_i-\mu}{\sigma}\right)\right] + \sum_{i=1}^{n} \log\left[(1-\alpha) + \alpha\, c_r \left(\frac{y_i-\mu}{\sigma}\right)^{2r}\right].$$

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

$$\begin{aligned}
U_\alpha(\boldsymbol{\theta}) &= c_r \sum_{i=1}^{n} \frac{(\frac{y_i-\mu}{\sigma})^{2r} - 1}{(1-\alpha) + \alpha\, c_r (\frac{y_i-\mu}{\sigma})^{2r}},\\
U_\mu(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \left(\frac{y_i-\mu}{\sigma}\right) + \frac{\alpha\, c_r\, 2r}{\sigma} \sum_{i=1}^{n} \frac{(\frac{y_i-\mu}{\sigma})^{2r-1}}{(1-\alpha) + \alpha\, c_r (\frac{y_i-\mu}{\sigma})^{2r}},\\
U_\sigma(\boldsymbol{\theta}) &= -\frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^{n} \left(\frac{y_i-\mu}{\sigma}\right)^2 + \frac{\alpha\, c_r\, 2r}{\sigma} \sum_{i=1}^{n} \frac{(\frac{y_i-\mu}{\sigma})^{2r}}{(1-\alpha) + \alpha\, c_r (\frac{y_i-\mu}{\sigma})^{2r}},
\end{aligned}$$

where

$$c_r = \frac{\sqrt{2\pi}}{2^{(2r+1)/2}\,\Gamma\left(\frac{2r+1}{2}\right)},$$

and $\psi(\cdot)$ is the digamma function.

The maximum likelihood method is used since it is conceptually easy although the profile log-likelihood for r could be difficult to compute in some cases (see, for example, Lange *et al.*, 1989; Berkane *et al.*, 1994; Cordeiro and Andrade, 2009; Ortega *et al.*, 2009). The maximization of $l(\boldsymbol{\theta})$ follows the same two steps for obtaining the maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$. In the first step of the iterative process we fix a range of values for r. Then, we obtain the MLEs $\widetilde{\alpha}(r)$, $\widetilde{\mu}(r)$ and $\widetilde{\sigma}(r)$ conditioned on r fixed, and then the maximized log-likelihood function $\mathrm{l}_{\max}(r)$ is determined. In this step, we use the NLMixed procedure in SAS. In the second step, the log-likelihood $\mathrm{l}_{\max}(r)$ is maximized, and then $\widehat{r}$ is obtained. The MLEs of α, μ and σ are given by $\widehat{\alpha} = \widetilde{\alpha}(\widehat{r})$, $\widehat{\mu} = \widetilde{\mu}(\widehat{r})$ and $\widehat{\sigma} = \widetilde{\sigma}(\widehat{r})$, respectively. This procedure is performed by assuming r fixed. Initial values for μ and σ can be taken from the fit of the standard normal model with $\mu = 0$ and $\sigma = 1$. The parameter α is in $(0, 1)$ and then we take 0.5 as initial guess. For interval estimation and hypothesis tests on the model parameters, we require a 3×3 observed information matrix $J(\boldsymbol{\theta}) = -\{J_{rs}\}$, where $r, s = \alpha, \mu$ and σ. The elements J_{rs} are given in Appendix B. Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is $N_3(0, I(\boldsymbol{\theta})^{-1})$, where $I(\boldsymbol{\theta}) = E[J(\boldsymbol{\theta})]$. Based on the multivariate normal $N_3(0, J(\widehat{\boldsymbol{\theta}})^{-1})$ distribution, we can construct approximate confidence intervals for the parameters. We can evaluate the maximum values of the unrestricted and restricted log-likelihoods to obtain likelihood ratio (LR) statistics for testing some sub-models of the EMN distribution in the classical way.

11. Bayesian analysis

In the Bayesian approach, the information referring to the model parameters is obtained through a posterior marginal distribution. We use the simulation method of Markov Chain Monte Carlo (MCMC) such as the Metropolis-Hastings algorithm. Since we have no prior information from historical data or from the previous experiment, we assign conjugate but weakly informative prior distributions to the parameters. Since we assume an informative (but weakly) prior distribution, the posterior distribution is a well-defined proper distribution. Further, we assume that the elements of the parameter vector are independent and consider that the joint prior distribution of the unknown parameters has a density function given by

$$\pi(\alpha, \mu, \sigma) \propto \pi(\alpha) \times \pi(\mu) \times \pi(\sigma). \tag{28}$$

Here, $\alpha \sim \mathrm{Be}(a, b)$, $\mu \sim N(\mu_1, \sigma_1^2)$ and $\sigma \sim \Gamma(a_1, b_1)$, where $\mathrm{Be}(a, b)$ denotes a beta distribution with a density function given by

$$f(\upsilon; a, b) = \frac{1}{\mathrm{B}(a, b)} \upsilon^{a-1}(1 - \upsilon)^{b-1},$$

where $\upsilon \in (0, 1)$, $a > 0$ and $b > 0$, $N(\mu_1, \sigma_1^2)$ denotes a normal distribution with mean μ_1 and variance σ_1^2 and $\Gamma(a_1, b_1)$ denotes the gamma model with a density function given by

$$f(\nu; a_1, b_1) = \frac{b_1^{a_1} \nu^{a_1 - 1} \mathrm{e}^{-\nu \mathrm{b}_1}}{\Gamma(a_1)},$$

where $\nu > 0$, $a_1 > 0$ and $b_1 > 0$. All hyper-parameters are specified. Combining the likelihood function $l(\boldsymbol{\theta})$ and the prior distribution (28), the joint posterior distribution for μ, σ and α reduces to

$$\pi(\alpha,\mu,\sigma|y) \propto \left(\frac{1}{\sigma}\right)^n \prod_{i=1}^{n} \phi\left(\frac{y_i-\mu}{\sigma}\right) \prod_{i=1}^{n} \left[(1-\alpha)+\alpha\, c_r \left(\frac{y_i-\mu}{\sigma}\right)^{2r}\right] \times \pi(\alpha,\mu,\sigma).$$

The joint posterior density above is analytically intractable because the integration of the joint posterior density is not easy to perform. In this direction, we first obtain the full conditional distributions of the unknown parameters given by

$$\pi(\alpha|y,\mu,\sigma) \propto \prod_{i=1}^{n} \left[(1-\alpha)+\alpha\, c_r \left(\frac{y_i-\mu}{\sigma}\right)^{2r}\right] \times \pi(\alpha),$$
$$\pi(\mu|y,\alpha,\sigma) \propto \prod_{i=1}^{n} \phi\left(\frac{y_i-\mu}{\sigma}\right) \prod_{i=1}^{n} \left[(1-\alpha)+\alpha\, c_r \left(\frac{y_i-\mu}{\sigma}\right)^{2r}\right] \times \pi(\mu)$$

and

$$\pi(\sigma|y,\alpha,\mu) \propto \left(\frac{1}{\sigma}\right)^n \prod_{i=1}^{n} \phi\left(\frac{y_i-\mu}{\sigma}\right) \prod_{i=1}^{n} \left[(1-\alpha)+\alpha\, c_r \left(\frac{y_i-\mu}{\sigma}\right)^{2r}\right] \times \pi(\sigma).$$

Since the full conditional distributions for μ, σ and α do not have explicit expressions, we use the Metropolis-Hastings algorithm.

12. Application: minimum flow data

We model the lower discharge of at least seven consecutive days and a return period (time) of 10 years ($Q_{7,10}$) of the Cuiabá River, Cuiabá, Mato Grosso, Brazil. We consider the data given by Andrade *et al.* (2007). The calculation of the lower discharge for seven consecutive days and a return period (time) of 10 years ($Q_{7,10}$) is an important hydrological parameter with applications in the study planning and management of the use of water resources. This study aims to model the lower flood (discharge) of at least seven consecutive days and a return period (time) of 10 years ($Q_{7,10}$) in the Cuiabá River, part of the Brazilian Pantanal (Swamp), since the ecosystem is strongly influenced by the hydrological system. For determining $Q_{7,10}$, we use a data series from 38 years (January 1962 to October 1999) relating to lower flows of n^o 66260001 hydrological station, installed in the Cuiabá River in the city of Cuiabá, Mato Grosso, Brazil.

As mentioned in Section 10, the parameter r is assumed to be fixed in order to obtain the MLEs. We verify that the profile log-likelihood $l(\hat{\alpha}(r),\hat{\mu}(r),\hat{\sigma}(r))$ reaches its maximum value at $r = 1$. Hence, this value is taken for the MLE of r. All computations are performed using the NLMixed procedure in SAS.

An alternative approach to modeling these data can be provided by the normal distribution. It belongs to the class of symmetric best known distributions due to various interesting properties and theoretical development achieved over the years.

Further, the t-Student distribution is used to model the behavior of data that come from a distribution with tails heavier than normal, reducing the influence of aberrant observations. There are various extensions of this distribution; see, for example, the skew-normal distribution (Azzalini, 1985). We also compare the proposed model with a mixture of two normal distributions.

- **t-Student distribution**

$$f(y) = \frac{\nu^{\nu/2}}{B(1/2,\nu/2)\sqrt{\phi}}\left\{\nu + \left(\frac{y-\mu}{\sqrt{\phi}}\right)\right\}, \quad y \in \mathbb{R},$$

  where $B(a,b) = [\Gamma(a)\Gamma(b)]/\Gamma(a+b)$ is the beta function and $\nu > 0$ is the number of degrees of freedom.

- **skew-normal distribution**

$$f(y) = \frac{2}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)\Phi\left[\lambda\left(\frac{y-\mu}{\sigma}\right)\right], \quad y \in \mathbb{R}, \tag{29}$$

  where $\lambda \in \mathbb{R}$ is the parameter of asymmetry, $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal pdf and cdf, respectively. Density (29) holds for $y \in \mathbb{R}$ and it is symmetric if $\lambda = 0$ (Azzalini, 1985).

- **mixtures of normal distributions**

$$f(y) = \frac{\alpha}{\sigma_1\sqrt{2\pi}}\exp\left\{\frac{-(y-\mu_1)^2}{2\sigma_1^2}\right\} + \frac{(1-\alpha)}{\sigma_2\sqrt{2\pi}}\exp\left\{\frac{-(y-\mu_2)^2}{2\sigma_2^2}\right\}, \tag{30}$$

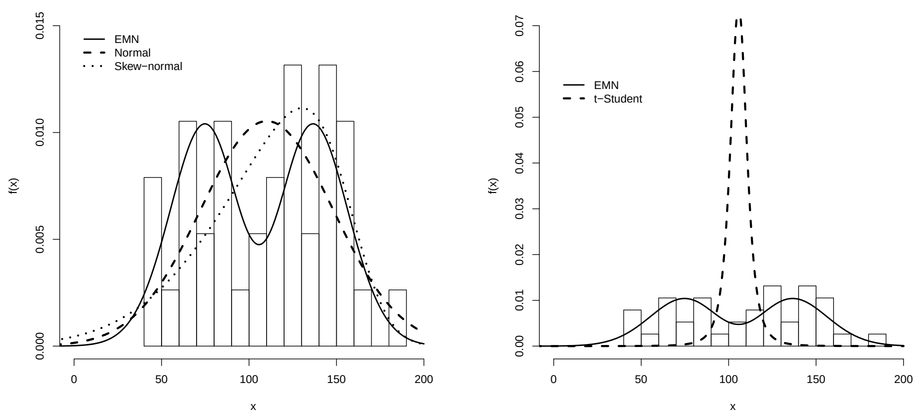
  where $\alpha \in (0,1)$, $\mu_1 \in \mathbb{R}$, $\mu_2 \in \mathbb{R}$, $\sigma_1 > 0$ and $\sigma_2 > 0$. Density (30) holds for $y \in \mathbb{R}$.

Model		r	α	μ	σ	AIC	GD	BIC
EMN		1	0.7077	105.71	24.541	383.4	375.4	388.4
			(0.159)	(4.214)	(2.278)			
				μ	σ			
Normal				110.21	37.8730	388.0	384.0	391.3
				(6.1438)	(4.3443)			
			ν	μ	ϕ			
t-Student			3	113.07	1098.18	395.2	389.2	398.4
				(7.0329)	(323.26)			
			λ	μ	σ			
Skew-normal			-3.3933	157.52	60.5994	388.9	382.9	393.8
			(2.0010)	(9.3845)	(10.0991)			
	α	μ_1	σ_1	μ_2	σ_2			
Mixture-normal	0.3962	70.1840	18.3674	136.4751	20.3236	386.3	376.3	394.6
	(0.1111)	(7.5343)	(5.2844)	(6.4182)	(4.8234)			

Table 1: *MLEs of model parameters for the minimum flow data, the corresponding SEs (given in parentheses) and the AIC, GD and BIC statistics*

Table 1 lists the MLEs (and the corresponding standard errors in parentheses) of the model parameters and the values of the Akaike Information Criterion (AIC), global deviance (GD) and the Bayesian Information Criterion (BIC) for some fitted models. These results indicate that the EMN model has the lowest AIC, GD and BIC values, and therefore it could be chosen as the best model.

Plots of the fitted EMN, normal, t-Student and skew-normal distributions over the histogram of the data are displayed in Figures 5a and 5b. They indicate that the EMN distribution provides the best fit to these data.



(a) *Estimated densities of the EMN, normal and skew-normal models for minimum flow data* (b) *Estimated densities of the EMN and t-Student models for minimum flow data*

Figure 5: *Estimated densities*

Bayesian analysis:

The following independent priors are considered to perform the Metropolis-Hastings algorithm: $\alpha \sim \text{Be}(0.5, 0.5)$, $\mu \sim N(0, 10)$ and $\sigma \sim \Gamma(0.01, 0.01)$, so that we have a vague prior distribution. We fix $r = 1$. Considering these prior density functions, we generate two parallel independent runs of the Metropolis-Hastings with size 150,000 for each parameter, disregarding the first 15,000 iterations to eliminate the effects of the initial values and to avoid correlation problems, we consider a spacing of size 10, obtaining a sample of size 13,500 from each chain. To monitor the convergence of the Metropolis-Hastings, we perform the methods suggested by Cowles and Carlin (1996). Further, we use the between and within sequence information following the approach developed by Gelman and Rubin (1992) to obtain the potential scale reduction, $\widehat{R}$. For all cases, these values are close to one, indicating the convergence of the chain. The approximate posterior marginal density functions for the parameters are presented in Figure 6.

In Table 2, we report posterior summaries for the parameters of the EMN model. We note that the values for the means a posteriori (Table 2) are quite close (as expected) to the MLEs given in Table 1. SD represents the standard deviation from the posterior distributions of the parameters and HPD represents the 95% highest posterior density (HPD) intervals.

Parameter	Median	SD	HPD (95%)	$\hat{R}$
α	0.7005	0.2170	(0.2229; 0.9883)	0.9997
μ	105.80	0.8059	(104.22; 107.38)	0.9999
σ	24.45	0.5040	(23.41; 25.39)	1.0024

Table 2: *Posterior summaries for the parameters from the EMN model for the minimum flow data*

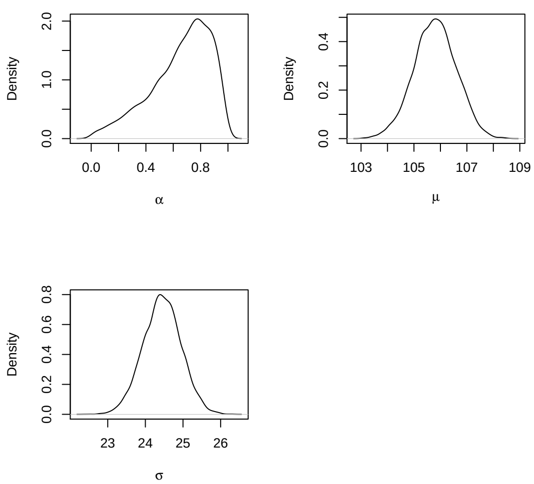


Figure 6: *Approximate posterior marginal densities for the parameters from the EMN model for the minimum flow data*

13. Conclusions

We define a new generalized normal model, the so-called extended mixture normal distribution, for analysis of symmetric real data. It includes special cases such as the normal and symmetric component models. We obtain some structural properties of the new distribution, including explicit expressions for the ordinary and incomplete moments, generating and quantile functions, mean deviations, two types of entropy and order statistics. The estimation of model parameters is performed using maximum likelihood and the Bayesian method. The observed information matrix is determined. We prove empirically that the new model can provide a better fit than the normal, t-Student and skew-normal distributions by means of a real data set.

Appendix A: Shannon entropy

Here, we provide the calculations of the four integrals in (24). Setting $x^2/2 = u$, we obtain the first two integrals as

$$\int_0^{+\infty} \mathrm{e}^{-\frac{\mathrm{x}^2}{2}} \left[\log\left(\frac{1}{\sqrt{2\pi}}\right) - \frac{\mathrm{x}^2}{2}\right] \mathrm{dx} = \sqrt{\tfrac{\pi}{2}}\log\left(\tfrac{1}{\sqrt{2\pi}}\right) - \tfrac{\sqrt{2\pi}}{4} \tag{31}$$

and

$$\int_0^{+\infty} x^{2r}\,\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\left[\log\left(\frac{1}{\sqrt{2\pi}}\right)-\frac{\mathrm{x}^2}{2}\right]\mathrm{dx}=2^{r-\frac{1}{2}}\left[\log\left(\frac{1}{\sqrt{2\pi}}\right)\Gamma\left(r+\frac{1}{2}\right)\right.$$
$$\left.-\,\Gamma\left(r+\frac{3}{2}\right)\right]. \tag{32}$$

For calculating the third integral in (24), we use a power series for the exponential function. Then,

$$\int_0^{+\infty}\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\log\left[(1-\alpha)+\alpha\mathrm{c_r}\mathrm{x}^{2\mathrm{r}}\right]\mathrm{dx}=\sqrt{\frac{\pi}{2}}\log(1-\alpha)+\sum_{k\geq 0}\frac{(-1)^k}{2^k\,k!}\int_0^{+\infty}x^{2k}$$
$$\times\log\left(1+\frac{\alpha c_r}{1-\alpha}\right)\mathrm{d}x.$$

Setting $\frac{\alpha\,c_r}{1-\alpha}\,x^{2r}=u$ in the last equation, we have

$$\int_0^{+\infty}\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\log\left[(1-\alpha)+\alpha\mathrm{c_r}\mathrm{x}^{2\mathrm{r}}\right]\mathrm{dx}=\sqrt{\frac{\pi}{2}}\log(1-\alpha)+\sum_{k\geq 0}\frac{(-1)^k}{2^{k+1}rk!}\left(\frac{1-\alpha}{\alpha c_r}\right)^{\frac{2k+1}{2r}}$$
$$\times\int_0^{+\infty}u^{\frac{2k+1}{2r}-1}\log(1+u)\mathrm{d}u.$$

Changing variable $1+u=v^{-1}$ and using the binomial expansion, we obtain

$$\int_0^{+\infty}\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\;\log\left[(1-\alpha)+\alpha\mathrm{c_r}\,\mathrm{x}^{2\mathrm{r}}\right]\mathrm{dx}=\sqrt{\frac{\pi}{2}}\;\log(1-\alpha)$$
$$+\sum_{k,\;i\geq 0}\binom{\frac{2k-2r+1}{2r}}{i}\frac{(-1)^{k+i}}{2^{k+1}\,r\,k!}\left(\frac{1-\alpha}{\alpha c_r}\right)^{\frac{2k+1}{2r}}$$
$$\times\int_0^1 v^{i-\frac{2k+1}{2r}-1}\;\log\left(\frac{1}{v}\right)\;\mathrm{d}v.$$

For $i>\frac{2k+1}{2r}$, the last equation becomes

$$\int_0^{+\infty}\mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\;\log\left[(1-\alpha)+\alpha\mathrm{c_r}\,\mathrm{x}^{2\mathrm{r}}\right]\mathrm{dx}$$
$$=\sqrt{\frac{\pi}{2}}\;\log(1-\alpha)+\sum_{k,\;i\geq 0}\binom{\frac{2k-2r+1}{2r}}{i}\frac{(-1)^{k+i}}{2^{k+1}\,r\,k!}\left(\frac{1-\alpha}{\alpha c_r}\right)^{\frac{2k+1}{2r}}\left(i-\frac{2k+1}{2r}\right)^{-2}. \tag{33}$$

Following the same algebra of the previous case, the fourth integral reduces to

$$\begin{aligned}
&\int_0^{+\infty} x^{2r}\, \mathrm{e}^{-\frac{\mathrm{x}^2}{2}}\ \log\left[(1-\alpha)+\alpha \mathrm{c_r}\, \mathrm{x}^{2\mathrm{r}}\right] \mathrm{dx}\\
&\quad = 2^{r-\frac{1}{2}}\,\Gamma\left(r+\frac{1}{2}\right)\ \log(1-\alpha) \qquad (34)\\
&\qquad + \sum_{k,\ i\geq 0} \binom{\frac{2k+1}{2r}}{i} \frac{(-1)^{k+i}}{2^{k+1}\, r\, k!} \left(\frac{1-\alpha}{\alpha c_r}\right)^{\frac{2k+1}{2r}} \left(i-\frac{2k+2r+1}{2r}\right)^{-2}.
\end{aligned}$$

Substituting (31)–(34) into (24) gives an explicit expression for the Shannon entropy.

Appendix B: observed information matrix

The elements of the observed information matrix $J(\boldsymbol{\theta})$ for the parameters $(\alpha,\mu,\sigma)^T$ are given by:

$$J_{\alpha\alpha}(\boldsymbol{\theta}) = c_r \sum_{i=1}^{n} \frac{\left[(\frac{y_i-\mu}{\sigma})^{2r}-1\right]\left[1-c_r(\frac{y_i-\mu}{\sigma})^{2r}\right]}{\left[(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}\right]^2},$$

$$J_{\alpha\mu}(\boldsymbol{\theta}) = -\frac{2rc_r}{\sigma}\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2r-1}}{(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}} + \frac{2\alpha r c_r}{\sigma}\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2r-1}\left[(\frac{y_i-\mu}{\sigma})^{2r}-1\right]}{\left[(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}\right]^2},$$

$$J_{\alpha\sigma}(\boldsymbol{\theta}) = -\frac{2rc_r}{\sigma}\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2r}}{(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}} + \frac{2\alpha r c_r^2}{\sigma}\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2r}\left[(\frac{y_i-\mu}{\sigma})^{2r}-1\right]}{\left[(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}\right]^2},$$

$$\begin{aligned}
J_{\mu\mu}(\boldsymbol{\theta}) &= -\frac{n}{\sigma} - \frac{2\alpha r c_r(2r-1)}{\sigma}\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2(r-1)}}{(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}}\\
&\quad + \left(\frac{2\alpha r c_r}{\sigma}\right)^2\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2(2r-1)}}{\left[(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}\right]^2},
\end{aligned}$$

$$\begin{aligned}
J_{\mu\sigma}(\boldsymbol{\theta}) &= -\frac{1}{\sigma}\sum_{i=1}^{n}\left(\frac{y_i-\mu}{\sigma}\right) - \frac{2\alpha r c_r(2r-1)}{\sigma^2}\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2(r-1)}}{(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}}\\
&\quad + \left(\frac{2\alpha r c_r}{\sigma}\right)^2\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2r}\left[(\frac{y_i-\mu}{\sigma})^{2r}-1\right]}{\left[(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}\right]^2},
\end{aligned}$$

$$\begin{aligned}
J_{\sigma\sigma}(\boldsymbol{\theta}) &= \frac{n}{\sigma^2} - \frac{3}{\sigma^2}\sum_{i=1}^{n}\left(\frac{y_i-\mu}{\sigma}\right)^2 - \frac{4\alpha r^2 c_r}{\sigma^2}\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2r}}{(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}}\\
&\quad - \frac{2\alpha r c_r}{\sigma^2}\sum_{i=1}^{n}\frac{(\frac{y_i-\mu}{\sigma})^{2r}\left[(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}-2\alpha r c_r(\frac{y_i-\mu}{\sigma})^{2r}\right]}{\left[(1-\alpha)+\alpha c_r(\frac{y_i-\mu}{\sigma})^{2r}\right]^2},
\end{aligned}$$

where $c_r = \frac{\sqrt{2\pi}}{2^{(2r+1)/2}\,\Gamma\left(\frac{2r+1}{2}\right)}$.

References

[1] N. L. R. ANDRADE, R. M. P. MOURA, A. SILVEIRA, *Determinação da $Q_{7,10}$ para o Rio Cuiabá, Mato Grosso, Brasil e comparação com a vazão regularizada após a implantação do reservatório de aproveitamento múltiplo de manso*, 24° *Congresso Brasileiro de engenharia sanitária e ambiental*, Belo Horizonte, Minas Gerais, Brasil, 2007.

[2] A. AZZALINI, *A class of distributions which includes the normal ones*, Scand. J. Stat. **12**(1985), 171–178.

[3] M. BERKANE, Y. KANO, P. M. BENTLER, *Pseudo maximum likelihood estimation in elliptical theory: effects of misspecification*, Comput. Statist. Data Anal. **18**(1994), 255–267.

[4] G. M. CORDEIRO, M. A. ANDRADE, *Transformed generalized linear models*, J. Statist. Plann. Inference **139**(2009), 2970–2987.

[5] M. K. COWLES, B. P. CARLIN, *Markov chain Monte Carlo convergence diagnostics: a comparative review*, J. Amer. Statist. Assoc. **91**(1996), 133–169.

[6] A. GELMAN, D. B. RUBIN, *Inference from iterative simulation using multiple sequences (with discussion)*, Statist. Sci. **7**(1992), 457–472.

[7] I. S. GRADSHTEYN, I. M. RYZHIK, *Table of integrals, series, and products*, Academic Press, San Diego, 2007.

[8] N. L. JOHNSON, S. KOTZ, N. BALAKRISHNAN, *Continuous univariate distributions*, Volume 1, John Wiley & Sons, New York, 1994.

[9] W. KRYSICKI, *Application de la mCthode des moments a I'estimation des parametres d'un melange de dew distributions de Rayleigh*, Rev. Stat. Appl. **11**(1963), 25–45.

[10] K. L. LANGE, J. A. LITTLE, M. G. TAYLOR, *Robust statistical modelling using the t distribution,* J. Amer. Statist. Assoc. **84**(1989), 881–896.

[11] A. I. MARKUSHEVICH, *Theory of functions of a complex variable*, Volume II, Prentice-Hall, New Jersey, 1965.

[12] E. M. M. ORTEGA, V. G. CANCHO, G. A. PAULA, *Generalized log-gamma regression models with cure fraction*, Lifetime Data Anal. **15**(2009), 79–106.

[13] A. RÉNYI, *On measures of entropy and information*, in: *Proceedings of the 4th Berkeley symposium on mathematical statistics and probability*, (J. Neyman, Ed.), University of California Press, 1961, 547–561.

[14] M. M. SIDDIQUI, G. H. WEISS, *Families of distributions for hourly median power and instantaneous power of received radio signals*, J. Res. Natl. Bur. Stand. **67D**(1963), 753–762.

[15] C. E. SHANNON, *Prediction and entropy of printed English*, Bell Syst. Tech. J. **30**(1951), 50–64.

[16] S. J. YAKOWITZ, J. D. SPRAGINS, *On the identifiability of finite mixtures*, Ann. Math. Stat. **39**(1968), 209–214.