CHILEAN JOURNAL OF STATISTICS

Edited by Víctor Leiva and Carolina Marchant

A free open-access journal indexed by

Volume 12 Number 2
December 2021
ISSN: 0718-7912 (print)
ISSN: 0718-7920 (online)

Published by the
Chilean Statistical Society
SOCHE
SOCIEDAD CHILENA DE ESTADÍSTICA

Aims

The Chilean Journal of Statistics (ChJS) is an official publication of the Chilean Statistical Society (`www.soche.cl`). The ChJS takes the place of *Revista de la Sociedad Chilena de Estadística*, which was published from 1984 to 2000.

The ChJS covers a broad range of topics in statistics, as well as in artificial intelligence, big data, data science, and machine learning, focused mainly on research articles. However, review, survey, and teaching papers, as well as material for statistical discussion, could be also published exceptionally. Each paper published in the ChJS must consider, in addition to its theoretical and/or methodological novelty, simulations for validating its novel theoretical and/or methodological proposal, as well as an illustration/application with real data.

The ChJS editorial board plans to publish one volume per year, with two issues in each volume. On some occasions, certain events or topics may be published in one or more special issues prepared by a guest editor.

Chilean Journal of Statistics

Volume 12, Number 2
December 2021

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

http://www.soche.cl/chjs

Contents

Chilean Journal of Statistics
Vol. 12, No. 2, December 2021, 125–144

Distribution Theory
Research Paper

On some properties of the bimodal normal distribution and its bivariate version

Roberto Vila[1,*], Helton Saulo[1], and Jamer Roldan[2]

[1]Department of Statistics, Universidade de Brasília, Brasília, Brazil
[2]Department of Mathematics, Instituto Federal de Goiás, Goiás, Brazil

(Received: 01 June 2021 · Accepted in final form: 23 December 2021)

Abstract

In this work, we derive some novel properties of the bimodal normal distribution. Some of its mathematical properties are examined. We provide a formal proof for the bimodality, present a stochastic representation, and assess identifiability. We also provide a closed formula for the moments of the bimodal normal distribution. We then discuss the maximum likelihood estimates as well as the existence of these estimates, and also some asymptotic properties of the estimator of the parameter that controls the bimodality. A bivariate version of the bimodal normal distribution is derived and some characteristics such as covariance and correlation are analyzed. We study stationarity and ergodicity and a triangular array central limit theorem. Finally, a Monte Carlo study is carried out for evaluating the performance of the maximum likelihood estimators empirically.

Keywords: Bimodality · Bivariate distribution · Central limit theorem · Ergodicity · Identifiability · Maximum likelihood method · Stationarity.

Mathematics Subject Classification: Primary 60E99 · Secondary 62E99.

1. Introduction

Bimodal distributions play an important role in the applied statistical literature; see, for example, Eugene et al. (2002), Hassan and El-Bassiouni (2016), and Alizadeh et al. (2017). The use of mixture-free bimodal distributions is very important as often real-world data are better modeled by these models, and in general, mixtures of distributions may suffer from identifiability problems in the parameter estimation; see Vila et al. (2020).

Recently, Gómez-Déniz et al. (2021) introduced a family of continuous distributions appropriate to describe the behavior of bimodal data. This family can accommodate any symmetric distribution and includes the bimodal normal (BN) as a special case. Bivariate distributions are of interest; see, for example, Saulo et al. (2020).

In this work, we derive some novel properties of the BN distribution. Particularly, in Section 2, we describe some preliminary properties, including the behavior of the density and hazard functions, median, moment generating function, mean, variance, among others. In Section 3, we obtain some results on the bimodality property of the BN distribution,

*Corresponding author. Email: rovig161@gmail.com

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

http://www.soche.cl/chjs

and the stochastic representation and moments are derived in Section 4. Also, in this section, we study some aspects of identifiability. In Section 5, we discuss maximum likelihood (ML) estimation, existence of the ML estimates, and some asymptotic properties of the ML estimators. A bivariate version of the BN distribution is derived and some characteristics such as covariance and correlation are analyzed in Section 6. In Section 7, the concepts of stationarity and ergodicity of a BN random process are studied. Ergodicity is an important ingredient to study functions of the distributional characteristics of the process when we have one realization of it. We find out that the BN random process is non-stationary. This result allows us to study, in Section 8, the triangular array central limit theorem, which is of vital importance in statistics. All theoretical results in this paper are new in the literature. In Section 9, we carry out Monte Carlo simulations. Finally, in Section 10, we discuss conclusions.

2. Preliminary properties

The random variable X follows a BN distribution if its probability density function (PDF) is given by

$$f_{\alpha,\zeta}(x) = \sqrt{2\pi}\,\mathrm{sech}(\zeta\alpha)\phi(\alpha)\phi(x)\cosh[\alpha(x-\zeta)], \quad x \in \mathbb{R}, \tag{1}$$

where $\zeta \in \mathbb{R}$ and $\alpha \in \mathbb{R}$ are shape and location parameters, respectively, ϕ is the standard normal PDF, and $\mathrm{sech}(z) = 1/\cosh(z)$, with $\cosh(z) = [\exp(z)+\exp(-z)]/2$. The parameter ζ presented in Equation (1) controls the skewness and the parameter α is related to the bimodality; see Gómez-Déniz et al. (2021).

In this work, we derive some novel properties of a special case of Equation (1), more specifically when $\zeta = 0$. Then, we say that a real-valued random variable X has a BN distribution with parameter vector parameter $\boldsymbol{\theta} = (\mu, \sigma, \alpha)^\top$, with $\mu \in \mathbb{R}$, $\sigma > 0$, and $\alpha \in \mathbb{R}$, denoted by $X \sim \mathrm{BN}(\boldsymbol{\theta})$, if its PDF is expressed as

$$f(x;\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 - \frac{\alpha^2}{2}\right] \cosh\left[\alpha\left(\frac{x-\mu}{\sigma}\right)\right], \quad x \in \mathbb{R}, \tag{2}$$

where μ is a location parameter, σ is a scale parameter, and α is a parameter that controls the unimodality or bimodality of the distribution. When α approaches zero (that is, $|\alpha| \leq 1$) the distribution becomes unimodal and when α grows (that is, $|\alpha| > 1$) the bimodality becomes more accentuated. When $\alpha = 0$, we have the known normal distribution. For more details, see Theorem 3.1.

Let $X \sim \mathrm{BN}(\boldsymbol{\theta})$ with PDF $f(x;\boldsymbol{\theta})$ given in Equation (2). Then, the behavior of $f(x;\boldsymbol{\theta})$ with $x \to 0$ or $x \to \pm\infty$ is stated as

$$\lim_{x\to 0} f(x;\boldsymbol{\theta}) = \sqrt{2\pi}\,\phi_{\mu,\sigma^2}(0)\phi(\alpha)\cosh\left(\frac{\alpha\mu}{\sigma}\right) \quad \text{and} \quad \lim_{x\to\pm\infty} f(x;\boldsymbol{\theta}) = 0, \tag{3}$$

where $\phi_{\mu,\sigma^2}(x)$ is the PDF of the normal distribution with mean μ and variance σ^2, and then we denote it by $\phi(x)$ instead of $\phi_{0,1}(x)$.

Observe that the cumulative distribution function (CDF) of $X \sim \mathrm{BN}(\boldsymbol{\theta})$ is given by

$$F(x;\boldsymbol{\theta}) = \frac{1}{4}\left[2 + \mathrm{erf}\left(\frac{x-\mu-\alpha\sigma}{\sigma\sqrt{2}}\right) + \mathrm{erf}\left(\frac{x-\mu+\alpha\sigma}{\sigma\sqrt{2}}\right)\right], \tag{4}$$

where $\operatorname{erf}(x) = 2\int_0^x \exp(-t^2)\,\mathrm{d}t/\sqrt{\pi}$ is the error function. Note that $\lim_{\alpha\to 0} F(x;0,1,\alpha) = (1/2)[1+\operatorname{erf}(x/\sqrt{2})] = \Phi(x)$, with Φ being the CDF of the standard normal distribution.

The CDF presented in Equation (4) is a special case of the bimodal skewed symmetric distribution of Hassan and El-Bassiouni (2016).

The hazard function $h(x;\boldsymbol{\theta}) = f(x;\boldsymbol{\theta})/[1-F(x;\boldsymbol{\theta})]$ has the following behavior when $x \to 0$ or $x \to \pm\infty$: $\lim_{x\to-\infty} h(x;\boldsymbol{\theta}) = 0$, $\lim_{x\to+\infty} h(x;\boldsymbol{\theta}) = +\infty$ and

$$\lim_{x\to 0} h(x;\boldsymbol{\theta}) = \frac{4\sqrt{2\pi}\,\phi_{\mu,\sigma^2}(0)\phi(\alpha)\cosh(\alpha\mu/\sigma)}{2-\operatorname{erf}\left(\frac{-\mu-\alpha\sigma}{\sigma\sqrt{2}}\right)-\operatorname{erf}\left(\frac{-\mu+\alpha\sigma}{\sigma\sqrt{2}}\right)}.$$

From the above limits, it can be concluded that the hazard function is non-decreasing.

A routine calculation shows that, if $X \sim \mathrm{BN}(\boldsymbol{\theta})$, then:

(P.1) (PDF) The random variable $Z = (X-\mu)/\sigma$, where $\mu \in \mathbb{R}$ and $\sigma > 0$, has PDF given by $f(z;0,1,\alpha) = (1/\sqrt{2\pi})\exp[-(z^2+\alpha^2)/2]\cosh(\alpha z)$, for $z \in \mathbb{R}$, that is, $Z \sim \mathrm{BN}(0,1,\alpha)$;

(P.2) If f is a Borel measurable function then

$$\mathbb{E}[f((X-\mu)/\sigma)] = \exp(-\alpha^2/2)\mathbb{E}_\Phi[f(Z)\cosh(\alpha Z)],$$

where $Z \sim \mathrm{N}(0,1)$ and $\mathbb{E}_\Phi$ denotes the expectation with respect to distribution function Φ;

(P.3) (Symmetry) $f(\mu-x;\boldsymbol{\theta}) = f(\mu+x;\boldsymbol{\theta})$ for all real numbers x;

(P.4) (Median) The median m satisfies that

$$\operatorname{erf}[(m-\mu-\alpha\sigma)/(\sigma\sqrt{2})] = \operatorname{erf}[(-m+\mu-\alpha\sigma)/(\sigma\sqrt{2})], \quad \text{so that} \quad m = \mu;$$

(P.5) (Moment generating function) $M_X(t) = \exp\left(\mu t + \sigma^2 t^2/2\right)\cosh(\alpha\sigma t)$, $t \in \mathbb{R}$;

(P.6) (Characteristic function) $\phi_X(t) = \exp(i\mu t - \sigma^2 t^2/2)\cosh(i\alpha\sigma t)$, for $t \in \mathbb{R}$;

(P.7) (Mean) $\mathbb{E}(X) = \mu$;

(P.8) (Variance) $\mathrm{Var}(X) = \sigma^2(1+\alpha^2)$;

(P.9) (Skewness) $v = 0$, that is, the distribution is approximately symmetrical;

(P.10) (Kurtosis) $\kappa = \alpha^2(\alpha^2+6)+3$;

(P.11) (Mean absolute deviation) $\mathrm{MAD} = [2\phi(\alpha) + \alpha\operatorname{erf}(\alpha/\sqrt{2})]\sigma$;

(P.12) (Shannon entropy) $\mathrm{SE} = \log(\sqrt{2\pi\sigma^2}) + 2\alpha^2 + 1/2 - \exp(-\alpha^2/2)[\exp(2\alpha^2)+1]/2$.

3. Unimodality and bimodality of the BN distribution

In this section, we provide unimodal and bimodal features of the BN distribution.

Theorem 3.1 The PDF of the BN distribution given in Equation (2) is unimodal when $|\alpha| \le 1$ and is bimodal when $|\alpha| > 1$.

Proof Let us suppose that $\alpha \neq 0$ because for the case $\alpha = 0$ the unimodality is well known. The derivative of $f(x;\boldsymbol{\theta})$ with respect to x is given by

$$f'(x;\boldsymbol{\theta}) = \frac{f(x;\boldsymbol{\theta})}{\sigma}\left\{\alpha\tanh\left[\alpha\left(\frac{x-\mu}{\sigma}\right)\right] - \left(\frac{x-\mu}{\sigma}\right)\right\}.$$

Then, $f'(x;\boldsymbol{\theta}) = 0$ if and only if

$$\tanh\left[\alpha\left(\frac{x-\mu}{\sigma}\right)\right] = \frac{x-\mu}{\alpha\sigma}. \tag{5}$$

Let $g(x;\boldsymbol{\theta}) = \tanh[\alpha(x-\mu)/\sigma] - (x-\mu)/(\alpha\sigma)$. Note that, for all $\alpha \neq 0$, $x = \mu$ is a root of $g(x;\boldsymbol{\theta})$. We divide the proof in the following two steps:

(i) First step: proving unimodality. Note that, $g'(x;\boldsymbol{\theta}) = (1/\sigma)\{\alpha\text{sech}^2[\alpha(x-\mu)/\sigma] - 1/\alpha\} < 0$ on $(-\infty, +\infty)$ when $0 < \alpha \leq 1$, and $g'(x;\boldsymbol{\theta}) > 0$ on $(-\infty, +\infty)$ when $-1 \leq \alpha < 0$, because $\text{sech}^2(x) \leq 1$. Since the function $g(x;\boldsymbol{\theta})$ has opposite signs at the extremes of the interval (that is, $\lim_{x\to-\infty} g(x;\boldsymbol{\theta}) = +\infty$, $\lim_{x\to+\infty} g(x;\boldsymbol{\theta}) = -\infty$ when $0 < \alpha \leq 1$, and $\lim_{x\to-\infty} g(x;\boldsymbol{\theta}) = -\infty$, $\lim_{x\to+\infty} g(x;\boldsymbol{\theta}) = +\infty$ when $-1 \leq \alpha < 0$) and it is monotonic, it will have a single zero at $x = \mu$. Then, since $\lim_{x\to\pm\infty} f(x;\boldsymbol{\theta}) \overset{(3)}{=} 0$, the unimodality of the BN distribution with PDF stated in Equation (2) is guaranteed.

(ii) Second step: proving bimodality. Without loss of generality, now we assume that $\alpha > 1$ because the other case $\alpha < -1$ is verified using similar arguments. For this case, note that $g(x;\boldsymbol{\theta}) > 0$ when $x \leq \mu - \sigma\alpha$ and $g(x;\boldsymbol{\theta}) < 0$ when $x \geq \mu + \sigma\alpha$. Then, there is no root of $g(x;\boldsymbol{\theta})$ outside of the interval $(\mu - \sigma\alpha, \mu + \sigma\alpha)$. Using Intermediate value theorem, $g(\mu - \sigma\alpha;\boldsymbol{\theta}) = 1 - \tanh(\alpha^2) > 0$, $\varepsilon^- = \lim_{x\to\mu^-} g(x;\boldsymbol{\theta}) < 0$, and $\varepsilon^+ = \lim_{x\to\mu^+} g(x;\boldsymbol{\theta}) > 0$, $g(\alpha;\boldsymbol{\theta}) = \tanh(\alpha^2) - 1 < 0$. Thus, there are $c_1 \in (\mu - \sigma\alpha, \varepsilon^-)$ and $c_3 \in (\varepsilon^+, \mu + \sigma\alpha)$: $g(c_i;\boldsymbol{\theta}) = 0$ for $i = 1, 3$. Now, we prove uniqueness of root on $(\mu - \sigma\alpha, \varepsilon^-)$. Indeed, assume that $g(x;\boldsymbol{\theta})$ has two solutions $g(a;\boldsymbol{\theta}) = g(b;\boldsymbol{\theta}) = 0$, $\mu - \sigma\alpha < a < b < \varepsilon^-$, then according to the Rolle theorem there is $c^* \in (a, b)$: $g'(c^*;\boldsymbol{\theta}) = 0$. But $g'(x;\boldsymbol{\theta}) = (1/\sigma)[\alpha\text{sech}^2(\alpha x) - 1/\alpha] < 0$ on $(\mu - \sigma\alpha, \varepsilon^-)$ with $\alpha > 1$, and has no solutions, contradiction. Therefore, $g(x;\boldsymbol{\theta})$ has exactly one real solution on $(\mu - \sigma\alpha, \varepsilon^-)$. Similarly, it is verified that on $(\varepsilon^+, \mu + \sigma\alpha)$, $g(x;\boldsymbol{\theta})$ has exactly one real solution. Thus, for $\alpha > 1$, $g(x;\boldsymbol{\theta})$ has exactly three real roots, denoted by x_1, x_2, x_3, such that $x_1 < x_2 = \mu < x_3$. Therefore, since $\lim_{x\to\pm\infty} f(x;\boldsymbol{\theta}) \overset{(3)}{=} 0$, the bimodality of the BN distribution expressed in (2) follows. ■

Remark 1 The modes of the BN distribution belong to the interval $(\mu - \sigma\alpha, \mu + \sigma\alpha)$. By symmetry, there is $\delta = \delta(\sigma, \alpha) \in (0, \sigma\alpha)$ so that $x_1 = \mu - \delta$ and $x_3 = \mu + \delta$. Moreover, when $|\alpha| > 1$ and $|x|$ is sufficiently large, the modes of the BN distribution are given by $x_1 \approx \mu - \sigma\alpha$ and $x_3 \approx \mu + \sigma\alpha$, because $\lim_{x\to\pm\infty} \tanh[\alpha(x-\mu)/\sigma] = \pm 1$.

Corollary 3.2 The modal point $x_0 = x_0(\boldsymbol{\theta})$ is a non-decreasing function of μ whenever $|\alpha| \leq 1$.

Proof By Equation (5), a modal point x_0 of the BN distribution satisfies

$$x_0 = \alpha\sigma \tanh\left[\alpha\left(\frac{x_0-\mu}{\sigma}\right)\right] + \mu. \tag{6}$$

Differentiating x_0 with respect to μ, we obtain

$$\frac{\partial x_0}{\partial \mu} = 1 - \alpha^2 \text{sech}^2\left[\alpha\left(\frac{x_0-\mu}{\sigma}\right)\right] \geq 0,$$

whenever $|\alpha| \leq 1$. Hence, x_0 is a non-decreasing function of μ. ■

Corollary 3.3 The modal point $x_0 = x_0(\boldsymbol{\theta})$ is a non-decreasing function of σ (respectively of α) whenever $x_0 \geq \mu$ and a non-increasing function of σ (respectively of α) whenever

$x_0 < \mu$.

Proof Differentiating x_0 in Equation (6) with respect to σ and α, we get

$$\frac{\partial x_0}{\partial \sigma} = \alpha \tanh\left[\alpha\left(\frac{x_0-\mu}{\sigma}\right)\right] - \alpha^2\left(\frac{x_0-\mu}{\sigma}\right)\operatorname{sech}^2\left[\alpha\left(\frac{x_0-\mu}{\sigma}\right)\right]$$

and

$$\frac{\partial x_0}{\partial \sigma} = \sigma\left\{\tanh\left[\alpha\left(\frac{x_0-\mu}{\sigma}\right)\right] + \alpha\left(\frac{x_0-\mu}{\sigma}\right)\operatorname{sech}^2\left[\alpha\left(\frac{x_0-\mu}{\sigma}\right)\right]\right\}.$$

From the above equations, it follows that $\partial x_0/\partial\sigma \geq 0$ (respectively $\partial x_0/\partial\alpha \geq 0$) whenever $x_0 \geq \mu$ and $\partial x_0/\partial\sigma < 0$ (respectively $\partial x_0/\partial\alpha < 0$) whenever $x_0 < \mu$. ■

4. Stochastic representation, moments, and identifiability

In this section, we provide the stochastic representation, moments, and identifiability of the BN distribution.

Proposition 4.1 Suppose Z_{μ,σ^2} has a normal distribution with expected value μ and variance σ^2. Let W have the Bernoulli distribution, so that $W = \alpha\sigma$ or $W = -\alpha\sigma$, each with probability $1/2$, and assume W is independent of Z_{μ,σ^2}. If $X = Z_{\mu,\sigma^2} + W$ then $X \sim \mathrm{BN}(\boldsymbol{\theta})$. Conversely, if $X \sim \mathrm{BN}(\boldsymbol{\theta})$, then $X = Z_{\mu,\sigma^2} + W$.

Proof By law of total probability and by independence, we get

$$\begin{aligned}
\mathbb{P}(X \leq x) &= \mathbb{P}(Z_{\mu,\sigma^2} + \alpha\sigma \leq x)\,\mathbb{P}(W = \alpha\sigma) + \mathbb{P}(Z_{\mu,\sigma^2} - \alpha\sigma \leq x)\,\mathbb{P}(W = -\alpha\sigma) \\
&= \mathbb{P}(Z_{\mu,\sigma^2} + \alpha\sigma \leq x)\,\frac{1}{2} + \mathbb{P}(Z_{\mu,\sigma^2} - \alpha\sigma \leq x)\,\frac{1}{2} \\
&= \Phi\left(\frac{x-\mu-\alpha\sigma}{\sigma}\right)\frac{1}{2} + \Phi\left(\frac{x-\mu+\alpha\sigma}{\sigma}\right)\frac{1}{2}.
\end{aligned}$$

By using the identity $\Phi(x) = (1/2)[1 + \operatorname{erf}(x/\sqrt{2})]$, the above expression is equal to

$$\frac{1}{4}\left[2 + \operatorname{erf}\left(\frac{x-\mu-\alpha\sigma}{\sigma\sqrt{2}}\right) + \operatorname{erf}\left(\frac{x-\mu+\alpha\sigma}{\sigma\sqrt{2}}\right)\right] \overset{(4)}{=} F(x;\boldsymbol{\theta}), \quad x \in \mathbb{R}.$$

Then, we have completed the proof. ■

Proposition 4.2 Let $X \sim \mathrm{BN}(\boldsymbol{\theta})$. Then,

$$\mathbb{E}(X^n) = \begin{cases} \sigma^n 2^{\frac{n-2}{2}} \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}} \left[{}_1F_1\left(-\frac{n}{2}, \frac{1}{2}; -\frac{\{\mu+\alpha\sigma\}^2}{2\sigma^2}\right) + {}_1F_1\left(-\frac{n}{2}, \frac{1}{2}; -\frac{\{\mu-\alpha\sigma\}^2}{2\sigma^2}\right)\right], & n \text{ even}, \\[2ex] \sigma^{n-1} 2^{\frac{n-1}{2}} \frac{\Gamma(\frac{n}{2}+1)}{\sqrt{\pi}} \left[(\mu+\alpha\sigma)\,{}_1F_1\left(\frac{1-n}{2}, \frac{3}{2}; -\frac{\{\mu+\alpha\sigma\}^2}{2\sigma^2}\right) + (\mu-\alpha\sigma)\,{}_1F_1\left(\frac{1-n}{2}, \frac{3}{2}; -\frac{\{\mu-\alpha\sigma\}^2}{2\sigma^2}\right)\right], & n \text{ odd}, \end{cases}$$

where ${}_1F_1(a,b;x) = [\Gamma(b)/\Gamma(a)]\sum_{k=0}^{\infty}[\Gamma(a+k)/\Gamma(b+k)](x^k/k!)$ is the Kummer confluent hypergeometric function; see Winkelbauer (2014).

Proof By Proposition 4.1, we have

$$\mathbb{E}(X^n) = \frac{1}{2}\left[\mathbb{E}_{\Phi_{\mu+\alpha\sigma,\sigma^2}}(X^n) + \mathbb{E}_{\Phi_{\mu-\alpha\sigma,\sigma^2}}(X^n)\right],$$

where $\mathbb{E}_{\Phi_{\mu+\alpha\sigma,\sigma^2}}$ denotes the expectation with respect to distribution function $\Phi_{\mu+\alpha\sigma,\sigma^2}$. By combining the above equality with the following known identity (Winkelbauer, 2014), for $Y \sim \mathrm{N}(\mu, \sigma^2)$,

$$\mathbb{E}(Y^n) = \begin{cases} \sigma^n 2^{n/2}\, \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}}\, {}_1F_1\big(-\frac{n}{2}, \frac{1}{2}; -\frac{\mu^2}{2\sigma^2}\big), & n \text{ even},\\ \mu\sigma^{n-1}2^{(n+1)/2}\, \frac{\Gamma(\frac{n}{2}+1)}{\sqrt{\pi}}\, {}_1F_1\big(\frac{1-n}{2}, \frac{3}{2}; -\frac{\mu^2}{2\sigma^2}\big), & n \text{ odd}, \end{cases}$$

the proof follows. ■

Proposition 4.3 Let $X \sim \mathrm{BN}(\boldsymbol{\theta})$. Then,

$$\mathbb{E}\left[\left(\frac{X-\mu}{\sqrt{\mathrm{Var}(X)}}\right)^n\right] = \begin{cases} \frac{1}{(1+\alpha^2)^{n/2}} \sum\limits_{\substack{0\leq k\leq n\\ k \text{ even}}} \binom{n}{k}\alpha^{n-k}2^{-\frac{k}{2}}\, \frac{k!}{(k/2)!}, & n \text{ even},\\ 0, & n \text{ odd}. \end{cases}$$

Proof By using Proposition 4.1 and that $\mathrm{Var}(X) = \sigma^2(1+\alpha^2)$, we get

$$\mathbb{E}\left[\left(\frac{X-\mu}{\sqrt{\mathrm{Var}(X)}}\right)^n\right] = \frac{1}{2(1+\alpha^2)^{n/2}}\left\{\mathbb{E}_{\Phi_{\mu+\alpha\sigma,\sigma^2}}\left[\left(\frac{X-\mu}{\sigma}\right)^n\right] + \mathbb{E}_{\Phi_{\mu-\alpha\sigma,\sigma^2}}\left[\left(\frac{X-\mu}{\sigma}\right)^n\right]\right\},$$

where $\mathbb{E}_{\Phi_{\mu+\alpha\sigma,\sigma^2}}$ denotes the expectation with respect to distribution function $\Phi_{\mu+\alpha\sigma,\sigma^2}$. Taking the changes of variable $z = (x-\mu)/\sigma$ and $\mathrm{d}z = \mathrm{d}x/\sigma$, and a binomial expansion, we have

$$\begin{aligned} \mathbb{E}\left[\left(\frac{X-\mu}{\sqrt{\mathrm{Var}(X)}}\right)^n\right] &= \frac{1}{2(1+\alpha^2)^{n/2}}\left\{\mathbb{E}_\Phi\left[(Z+\alpha)^n\right] + \mathbb{E}_\Phi\left[(Z-\alpha)^n\right]\right\}\\ &= \frac{1}{2(1+\alpha^2)^{n/2}} \sum_{k=0}^{n}\binom{n}{k}\left[1+(-1)^{n-k}\right]\alpha^{n-k}\mathbb{E}_\Phi(Z^k). \end{aligned} \tag{7}$$

A simple observation shows that, when n is even,

$$\frac{1}{2(1+\alpha^2)^{n/2}} \sum_{k=0}^{n}\binom{n}{k}\left[1+(-1)^{n-k}\right]\alpha^{n-k}\mathbb{E}_\Phi(Z^k) = \frac{1}{(1+\alpha^2)^{n/2}} \sum_{\substack{0\leq k\leq n\\ k \text{ even}}}\binom{n}{k}\alpha^{n-k}\mathbb{E}_\Phi(Z^k), \tag{8}$$

and, when n is odd,

$$\frac{1}{2(1+\alpha^2)^{n/2}} \sum_{k=0}^{n}\binom{n}{k}\left[1+(-1)^{n-k}\right]\alpha^{n-k}\mathbb{E}_\Phi(Z^k) = \frac{1}{(1+\alpha^2)^{n/2}} \sum_{\substack{0\leq k\leq n\\ k \text{ odd}}}\binom{n}{k}\alpha^{n-k}\mathbb{E}_\Phi(Z^k). \tag{9}$$

Thus, by combining the known identities, $\mathbb{E}_{\Phi}(Z^k) = 0$ for k odd, and

$$\mathbb{E}_{\Phi}(Z^k) = 2^{-\frac{k}{2}} \frac{k!}{(k/2)!},$$

for k even, considering Equations (7), (8) and (9), the proof follows. ■

As a consequence of Proposition 4.1, we know that the BN PDF $f(x;\boldsymbol{\theta})$ given in Equation (2), with parameter vector $\boldsymbol{\theta} = (\mu, \sigma, \alpha)^\top$, can be written as a finite mixture of two normal distributions with different location parameters, that is, given by

$$f(x;\boldsymbol{\theta}) = \frac{1}{2}\left[\phi_{\mu+\alpha\sigma,\sigma^2}(x) + \phi_{\mu-\alpha\sigma,\sigma^2}(x)\right]. \tag{10}$$

Let $\mathcal{N}$ be the family of normal distributions stated as

$$\mathcal{N} = \left\{F\colon F(x;\mu,\sigma) = \int_{-\infty}^{x} \phi_{\mu,\sigma^2}(y)\,\mathrm{d}y,\ \mu \in \mathbb{R}, \sigma > 0,\ x \in \mathbb{R}\right\}.$$

In addition, let $\mathcal{H}_{\mathcal{N}}$ be the class of all finite mixtures of $\mathcal{N}$. It is well-known that the class $\mathcal{H}_{\mathcal{N}}$ is identifiable; see Teicher (1963). The following result proves the identifiability of the BN distribution.

Proposition 4.4 The mapping $\boldsymbol{\theta} \longmapsto f(x;\boldsymbol{\theta})$, for all $x \in \mathbb{R}$, is one-to-one.

Proof Let us suppose that $f(x;\boldsymbol{\theta}) = f(x;\boldsymbol{\theta}')$ for all $x \in \mathbb{R}$. Thus, by Equation (10), we have that

$$\frac{1}{2}\left[\phi_{\mu+\alpha\sigma,\sigma^2}(x) + \phi_{\mu-\alpha\sigma,\sigma^2}(x)\right] = \frac{1}{2}\left[\phi_{\mu'+\alpha'\sigma',\sigma'^2}(x) + \phi_{\mu'-\alpha'\sigma',\sigma'^2}(x)\right].$$

Since $\mathcal{H}_{\mathcal{N}}$ is identifiable, we have $\mu \pm \alpha\sigma = \mu' \pm \alpha'\sigma'$ and $\sigma^2 = \sigma'^2$. From where immediately follows that $\mu = \mu'$, $\sigma = \sigma'$ and $\alpha = \alpha'$. Therefore, $\boldsymbol{\theta} = \boldsymbol{\theta}'$, and the identifiability of distribution follows. ■

5. Asymptotic properties

Let X be a random variable with BN distribution that depends on a parameter vector $\boldsymbol{\theta} = (\mu, \sigma, \alpha)^\top$, with $\boldsymbol{\theta}$ being an open subset of $\mathbb{R}^3$, where distinct values of $\boldsymbol{\theta}$ yield distinct distributions for X (see Section 4). Let $\boldsymbol{X} = (X_1, \ldots, X_n)^\top$ be a random sample of X. Then, the log-likelihood function for $\boldsymbol{\theta}$ is given by

$$l(\boldsymbol{\theta};\boldsymbol{X}) \propto -n\log(\sigma) - \frac{n\alpha^2}{2} - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 + \sum_{i=1}^{n}\log\left\{\cosh\left[\alpha\left(\frac{X_i-\mu}{\sigma}\right)\right]\right\}.$$

A simple computation shows that

$$\frac{\partial l(\boldsymbol{\theta};\boldsymbol{X})}{\partial\mu} = \frac{n}{\sigma}\left(\frac{\overline{X}-\mu}{\sigma}\right) - \frac{\alpha}{\sigma}\sum_{i=1}^{n}\tanh\left[\alpha\left(\frac{X_i-\mu}{\sigma}\right)\right], \tag{11}$$

$$\frac{\partial l(\boldsymbol{\theta};\boldsymbol{X})}{\partial\sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 - \frac{\alpha}{\sigma}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)\tanh\left[\alpha\left(\frac{X_i-\mu}{\sigma}\right)\right], \tag{12}$$

$$\frac{\partial l(\boldsymbol{\theta};\boldsymbol{X})}{\partial\alpha} = -\alpha n + \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)\tanh\left[\alpha\left(\frac{X_i-\mu}{\sigma}\right)\right]. \tag{13}$$

The log-likelihood equations for the estimators $\widehat{\mu}$, $\widehat{\sigma}$, $\widehat{\alpha}$ are given by

$$\widehat{\mu} = \overline{X} - \frac{\widehat{\alpha}}{n}\sum_{i=1}^{n}\tanh\left[\widehat{\alpha}\left(\frac{X_i-\widehat{\mu}}{\widehat{\sigma}}\right)\right],$$

$$\widehat{\sigma}^2 = \frac{1}{(1+\widehat{\alpha}^2)n}\sum_{i=1}^{n}(X_i-\widehat{\mu})^2,$$

$$\widehat{\alpha} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i-\widehat{\mu}}{\widehat{\sigma}}\right)\tanh\left[\widehat{\alpha}\left(\frac{X_i-\widehat{\mu}}{\widehat{\sigma}}\right)\right].$$

In the following two propositions, we study the existence of the ML estimator when two parameters are assumed to be known.

Proposition 5.1 If the parameters σ and α are known, then Equation (11) has at least one root on the interval $(-\infty,+\infty)$.

Proof One can readily verify that $\lim_{\mu\to\mp\infty}\partial l(\boldsymbol{\theta};\boldsymbol{X})/\partial\mu = \pm\infty$. Hence, by Intermediate value theorem, there exists at least one solution on the interval $(-\infty,+\infty)$. ■

Proposition 5.2 If the parameters μ and σ are known, then Equation (13) has at least one root on the interval $(-\infty,+\infty)$.

Proof Since $\lim_{\alpha\to\mp\infty}\partial l(\boldsymbol{\theta};\boldsymbol{X})/\partial\alpha = \pm\infty$, the proof follows the same reasoning as the proofs of Proposition 5.1. ■

Now, we calculate the expectation of the score defined by Equations (11), (12) and(13) when $n=1$. Indeed, by using the partial derivatives in Equations (11)-(13), with $n=1$, and the fact that $x\longmapsto x\cosh(\alpha x)$ and $x\longmapsto\sinh(\alpha x)$ are odd functions, we obtain

$$\begin{aligned}\mathbb{E}\left[\frac{\partial\log\{f(X;\boldsymbol{\theta})\}}{\partial\mu}\right] &= \frac{n}{\sigma}\,\mathbb{E}\left(\frac{X-\mu}{\sigma}\right) - \frac{\alpha}{\sigma}\,\mathbb{E}\left\{\tanh\left[\alpha\left(\frac{X-\mu}{\sigma}\right)\right]\right\}\\ &= \exp\left(-\frac{\alpha^2}{2}\right)\left\{\frac{n}{\sigma}\,\mathbb{E}_\Phi[Z\cosh(\alpha Z)] - \frac{\alpha}{\sigma}\,\mathbb{E}_\Phi[\sinh(\alpha Z)]\right\} = 0,\end{aligned}$$

where in the second line, the changes of variables $z=(x-\mu)/\sigma$, $\mathrm{d}z=\mathrm{d}x/\sigma$ were considered.

Analogously, since $\mathbb{E}_\Phi[Z^2\cosh(\alpha Z)] = (\alpha^2+1)\exp(\alpha^2/2)$ and $\mathbb{E}_\Phi[Z\sinh(\alpha Z)] =$

$\alpha \exp\left(\alpha^2/2\right)$, we get

$$\begin{aligned}
\mathbb{E}\left[\frac{\partial \log\{f(X;\boldsymbol{\theta})\}}{\partial \sigma}\right] &= -\frac{1}{\sigma} + \frac{1}{\sigma}\,\mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^2\right] - \frac{\alpha}{\sigma}\,\mathbb{E}\left\{\left(\frac{X-\mu}{\sigma}\right)\tanh\left[\alpha\left(\frac{X-\mu}{\sigma}\right)\right]\right\} \\
&= -\frac{1}{\sigma} + \frac{1}{\sigma}\,\exp\left(-\frac{\alpha^2}{2}\right)\mathbb{E}_\Phi\left[Z^2\cosh(\alpha Z)\right] - \frac{\alpha}{\sigma}\,\exp\left(-\frac{\alpha^2}{2}\right)\mathbb{E}_\Phi\left[Z\sinh(\alpha Z)\right] \\
&= 0
\end{aligned}$$

and let

$$\begin{aligned}
\mathbb{E}\left[\frac{\partial \log\{f(X;\boldsymbol{\theta})\}}{\partial \alpha}\right] &= \mathbb{E}\left\{\left(\frac{X-\mu}{\sigma}\right)\tanh\left[\alpha\left(\frac{X-\mu}{\sigma}\right)\right]\right\} - \alpha \\
&= \exp\left(-\frac{\alpha^2}{2}\right)\mathbb{E}_\Phi\left[Z\sinh(\alpha Z)\right] - \alpha \;=\; 0. \qquad (14)
\end{aligned}$$

In what remains of this section, for the sake of simplicity of presentation, we will assume that μ and σ are known parameters and α is unknown. We are interested in knowing the large sample properties of ML estimator $\widehat{\alpha}$ of the parameter α that generates uni- or bimodality in the BN distribution. We emphasize that similar results can be studied for μ and σ when the other parameters are known. Since

$$\frac{\partial^2 f(x;\boldsymbol{\theta})}{\partial \alpha^2} = \left[\alpha^2 + \left(\frac{x-\mu}{\sigma}\right)^2 - 1\right] f(x;\boldsymbol{\theta}) - 2\alpha\left(\frac{x-\mu}{\sigma}\right)\tanh\left[\alpha\left(\frac{x-\mu}{\sigma}\right)\right] f(x;\boldsymbol{\theta}),$$

$\mathbb{E}_\Phi\left[Z^2\cosh(\alpha Z)\right] = (\alpha^2+1)\exp\left(\alpha^2/2\right)$ and $\mathbb{E}_\Phi\left[Z\sinh(\alpha Z)\right] = \alpha\exp\left(\alpha^2/2\right)$, for $Z \sim \mathrm{N}(0,1)$, we have

$$\begin{aligned}
\int_{-\infty}^{+\infty} \frac{\partial^2 f(x;\boldsymbol{\theta})}{\partial \alpha^2}\,\mathrm{d}x &= \mathbb{E}\left[\alpha^2 + \left(\frac{X-\mu}{\sigma}\right)^2 - 1\right] - 2\alpha\mathbb{E}\left\{\left(\frac{X-\mu}{\sigma}\right)\tanh\left[\alpha\left(\frac{X-\mu}{\sigma}\right)\right]\right\} \\
&= \alpha^2 + \exp\left(-\frac{\alpha^2}{2}\right)\mathbb{E}_\Phi\left[Z^2\cosh(\alpha Z)\right] - 1 - 2\alpha\exp\left(-\frac{\alpha^2}{2}\right)\mathbb{E}_\Phi\left[Z\sinh(\alpha Z)\right] \\
&= 0, \qquad (15)
\end{aligned}$$

where in the second line, the changes of variables $z = (x-\mu)/\sigma$ and $\mathrm{d}z = \mathrm{d}x/\sigma$ were considered. In addition,

$$\frac{\partial^2 \log\{f(x;\boldsymbol{\theta})\}}{\partial \alpha^2} = \left(\frac{x-\mu}{\sigma}\right)^2 \mathrm{sech}^2\left[\alpha\left(\frac{x-\mu}{\sigma}\right)\right] - 1. \qquad (16)$$

Then, by Equations (15) and (16), the Fisher information may also be written as

$$\begin{aligned}\mathcal{I}(\alpha) = \mathbb{E}\Big[\frac{\partial \log\{f(X;\boldsymbol{\theta})\}}{\partial\alpha}\Big]^2 &= -\mathbb{E}\Big[\frac{\partial^2 \log\{f(X;\boldsymbol{\theta})\}}{\partial\alpha^2}\Big] + \int_{-\infty}^{+\infty}\frac{\partial^2 f(x;\boldsymbol{\theta})}{\partial\alpha^2}\,\mathrm{d}x \\ &= 1 - \mathbb{E}\Big\{\Big(\frac{X-\mu}{\sigma}\Big)^2 \operatorname{sech}^2\Big[\alpha\Big(\frac{X-\mu}{\sigma}\Big)\Big]\Big\} \\ &= 1 - \exp\Big(-\frac{\alpha^2}{2}\Big)\mathbb{E}_\Phi\big[Z^2\operatorname{sech}(\alpha Z)\big]. \end{aligned} \tag{17}$$

Theorem 5.3 Let us suppose that μ and σ are known parameters and α unknown, and $\Theta = \{\alpha \in \mathbb{R}\colon |\alpha| > 0\}$ be the parameter space. Then, with probability approaching one, as $n \to +\infty$, the log-likelihood equation $\partial l(\boldsymbol{\theta};\boldsymbol{X})/\partial\alpha = 0$ has a consistent solution, denoted by $\widehat{\alpha}$.

Proof Since $\partial \log\{f(x;\boldsymbol{\theta})\}/\partial\alpha$, $\partial^2 \log\{f(x;\boldsymbol{\theta})\}/\partial\alpha^2$, $\partial^3 \log\{f(x;\boldsymbol{\theta})\}/\partial\alpha^3]$ exist for all $\alpha \in \Theta$ and every x, by Cramér (1946) it is sufficient to prove that:

(i) $\mathbb{E}[\partial \log\{f(X;\boldsymbol{\theta})\}/\partial\alpha] = 0$ for all $\alpha \in \Theta$;
(ii) $-\infty < \mathbb{E}[\partial^2 \log\{f(X;\boldsymbol{\theta})\}/\partial\alpha^2] < 0$ for all $\alpha \in \Theta$;
(iii) There exists a function $H(x)$ such that for all $\alpha \in \Theta$, $l|\partial^3 \log\{f(x;\boldsymbol{\theta})\}/\partial\alpha^3| < H(x)$ and $\mathbb{E}[H(X)] < \infty$.

In what follows we show the validity of items (i), (ii) and (iii) above. By Equation (14), the statement of item (i) follows. In order to verify item (ii), note that $\exp(-\alpha^2/2)\mathbb{E}_\Phi[Z^2\operatorname{sech}(\alpha Z)] \le \mathbb{E}_\Phi(Z^2) = 1$ for all $\alpha \in \Theta$. Moreover, the two sides are equal if and only if $\alpha = 0$. Since $\alpha \in \Theta$ (that is, $\alpha \neq 0$), it follows that $\exp\big(-\alpha^2/2)\mathbb{E}_\Phi[Z^2\operatorname{sech}(\alpha Z)] < 1$. Hence,

$$-1 \le \mathbb{E}\Big[\frac{\partial^2 \log\{f(X;\boldsymbol{\theta})\}}{\partial\alpha^2}\Big] \overset{(17)}{=} \exp\Big(-\frac{\alpha^2}{2}\Big)\mathbb{E}_\Phi\big[Z^2\operatorname{sech}(\alpha Z)\big] - 1 < 0. \tag{18}$$

Then, item (ii) is valid. Thus, since $|\operatorname{sech}^2(\alpha x)| \le 1$ and $|\tanh(\alpha x)| \le 1$,

$$\begin{aligned}\Big|\frac{\partial^3 \log\{f(x;\boldsymbol{\theta})\}}{\partial\alpha^3}\Big| &= \Big|2\Big(\frac{X-\mu}{\sigma}\Big)^3\operatorname{sech}^2\Big[\alpha\Big(\frac{X-\mu}{\sigma}\Big)\Big]\tanh\Big[\alpha\Big(\frac{X-\mu}{\sigma}\Big)\Big]\Big| \\ &\le 2\Big|\Big(\frac{x-\mu}{\sigma}\Big)\Big|^3 = H(x),\end{aligned} \tag{19}$$

with $\mathbb{E}[H(X)] < \infty$. Therefore, we have completed the proof. ■

The following result supports the intuitive appeal of the ML estimator (Bahadur, 1971).

Proposition 5.4 Under hypothesis of Theorem 5.3, we have that

$$\lim_{n\to+\infty}\int_{\mathbb{R}^n}\mathbb{1}_{\{\boldsymbol{y}\in\mathbb{R}^n\,:\,\exp[l(\boldsymbol{\theta}';\boldsymbol{y})] > \exp[l(\boldsymbol{\theta};\boldsymbol{y})]\}}(\boldsymbol{x})\exp[l(\boldsymbol{\theta}';\boldsymbol{x})]\,\mathrm{d}\boldsymbol{x} = 1,$$

for any $\boldsymbol{\theta} = (\mu,\sigma,\alpha)^\top$, $\boldsymbol{\theta}' = (\mu,\sigma,\alpha')^\top \in \Theta$ with $\alpha \neq \alpha'$. Here, $\mathbb{1}_A(x)$ is the indicator function of a set A having the value 1 for all x in A and the value 0 for all x not in A.

Proof Since $\boldsymbol{X} = (X_1, \ldots, X_n)^\top$ is a random sample of $X \sim \mathrm{BN}(\boldsymbol{\theta})$, $X_1, \ldots, X_n$ are independent and identically distributed random variables with PDF $f(x; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta$. Therefore, as the BN distribution is identifiable (see Section 4), by Bahadur (1971), the proof follows. ■

Next, we state a central limit theorem for the ML estimator $\widehat{\alpha}$, which is important for studying confidence intervals and hypothesis tests, for example. Note that, under hypothesis of Theorem 5.3, the following conditions are satisfied:

(A.1) The mapping $\alpha \longmapsto f(x; \boldsymbol{\theta})$ is three times continuously differentiable on Θ, $\forall x \in \mathbb{R}$;
(A.2) By Equation (14), $\int_{-\infty}^{+\infty} \partial f(x; \boldsymbol{\theta})/\partial\alpha \, \mathrm{d}x = \mathbb{E}[\partial \log\{f(X; \boldsymbol{\theta})\}/\partial\alpha] = 0$ and, by Equation (15),
$\int_{-\infty}^{+\infty} \partial^2 f(x; \boldsymbol{\theta})/\partial\alpha^2 \, \mathrm{d}x = 0$;
(A.3) By Equations (17) and (18), $0 < \mathcal{I}(\alpha) = 1 - \mathbb{E}\big[X^2 \mathrm{sech}^2(\alpha X)\big] \leq 1$, $\forall \alpha \in \Theta$;
(A.4) By Equation (19), there exists a function $H(x)$ such that for all $\alpha \in \Theta$,

$$\left|\frac{\partial^3 \log\{f(x; \boldsymbol{\theta})\}}{\partial\alpha^3}\right| < H(x), \quad \mathbb{E}[H(X)] < \infty;$$

(A.5) By Theorem 5.3, the log-likelihood equation $\partial l(\boldsymbol{\theta}; \boldsymbol{X})/\partial\alpha = 0$ has a consistent solution $\widehat{\alpha}$.

Since conditions (A.1)-(A.5) are satisfied, by Cramér (1946), we have the following result.

Theorem 5.5 Under hypothesis of Theorem 5.3, it holds that, $\sqrt{n}(\widehat{\alpha} - \alpha)$ converges in distribution to $\mathrm{N}(0, 1/\mathcal{I}(\alpha))$ as $n \to +\infty$.

6. The bivariate BN distribution

We said that a real random vector $\boldsymbol{X} = (X_1, X_2)^\top$ has bivariate BN (BBN) distribution with parameter vector parameter $\boldsymbol{\psi} = (\mu_1, \mu_2, \sigma_1, \sigma_2, \alpha)^\top$, $\mu_i \in \mathbb{R}$, $\sigma_i > 0$, $\alpha \in \mathbb{R}$, denoted by $\boldsymbol{X} \sim \mathrm{BBN}(\boldsymbol{\psi})$, if its PDF is given by, for each $\mathbf{x} = (x_1, x_2)^\top \in \mathbb{R}^2$,

$$f(\mathbf{x}; \boldsymbol{\psi}) = \frac{\exp[\alpha^2(\rho^2 - 2)/2]}{\sigma_1\sigma_2}\, \phi\bigg(\frac{x_1 - \mu_1}{\sigma_1}, \frac{x_2 - \mu_2}{\sigma_2}; \rho\bigg) \cosh\bigg[\alpha\bigg(\frac{x_1 - \mu_1}{\sigma_1}\bigg) + \alpha(1 - \rho)\bigg(\frac{x_2 - \mu_2}{\sigma_2}\bigg)\bigg],$$

where $\rho \in (-1, 1)$ and

$$\phi(\mathbf{z}; \rho) = \frac{1}{2\pi\sqrt{1 - \rho^2}} \exp\bigg[-\frac{1}{2(1 - \rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2)\bigg], \quad \mathbf{z} = (z_1, z_2)^\top \in \mathbb{R}^2,$$

is the PDF of the standard bivariate normal distribution with correlation coefficient ρ.

A simple algebraic manipulation shows that

$$\int_{-\infty}^{\infty} f(\mathbf{x}; \boldsymbol{\psi}) \, \mathrm{d}x_1 = f(x_2; \boldsymbol{\theta}_2) \quad \text{and} \quad \int_{-\infty}^{\infty} f(\mathbf{x}; \boldsymbol{\psi}) \, \mathrm{d}x_2 = f(x_1; \boldsymbol{\theta}_1),$$

where $f(x_i; \boldsymbol{\theta}_i)$ is the PDF of the BN distribution stated in Equation (2) with parameter vector $\boldsymbol{\theta}_i = (\mu_i, \sigma_i, \alpha)^\top$, for $i = 1, 2$. Thus, if $\boldsymbol{X} = (X_1, X_2)^\top \sim \mathrm{BBN}(\boldsymbol{\psi})$ then $X_1 \sim \mathrm{BN}(\boldsymbol{\theta}_1)$ and $X_2 \sim \mathrm{BN}(\boldsymbol{\theta}_2)$.

By using previous results, a laborious algebraic calculation gives

$$\mathbb{E}(X_1|X_2=x_2)=\mu_1+\rho\sigma_1\Big(\frac{x_2-\mu_2}{\sigma_2}\Big)+(1-\rho^2)\sigma_1\tanh\Big[\alpha\Big(\frac{x_2-\mu_2}{\sigma_2}\Big)\Big],$$

that is,

$$\mathbb{E}(X_1|X_2)=\mu_1+\rho\sigma_1\Big(\frac{X_2-\mu_2}{\sigma_2}\Big)+(1-\rho^2)\sigma_1\tanh\Big[\alpha\Big(\frac{X_2-\mu_2}{\sigma_2}\Big)\Big] \quad \text{almost sure.}$$

In consequence,

$$\begin{aligned}\mathbb{E}(X_1X_2)&=\mathbb{E}[X_2\mathbb{E}(X_1|X_2)]\\ &=\mu_1\mathbb{E}(X_2)+\rho\sigma_1\mathbb{E}\Big[X_2\Big(\frac{X_2-\mu_2}{\sigma_2}\Big)\Big]+(1-\rho^2)\sigma_1\mathbb{E}\Big\{X_2\tanh\Big[\alpha\Big(\frac{X_2-\mu_2}{\sigma_2}\Big)\Big]\Big\}.\end{aligned}$$

Since $X_2\sim \mathrm{BN}(\boldsymbol{\theta}_2)$, we get

$$\mathbb{E}(X_1X_2)=\mu_1\mu_2+\rho\sigma_1\sigma_2(1+\alpha^2)+(1-\rho^2)\sigma_1\sigma_2\alpha.$$

Hence, as $\mathbb{E}(X_i)=\mu_i$ and $\mathrm{Var}(X_i)=\sigma_i^2(1+\alpha^2)$ (see properties P.7 and P.8 in Section 2),

$$\mathrm{Cov}(X_1,X_2)=\sigma_1\sigma_2[\rho(1+\alpha^2)+(1-\rho^2)\alpha]; \tag{20}$$

$$\rho(X_1,X_2)=\frac{\rho(1+\alpha^2)+(1-\rho^2)\alpha}{(1+\alpha^2)}.$$

The covariance matrix is given by

$$\boldsymbol{\Sigma}=\begin{bmatrix}\sigma_1^2(1+\alpha^2) & \sigma_1\sigma_2[\rho(1+\alpha^2)+(1-\rho^2)\alpha]\\ \sigma_1\sigma_2[\rho(1+\alpha^2)+(1-\rho^2)\alpha] & \sigma_2^2(1+\alpha^2)\end{bmatrix}.$$

Some immediate observations are the following:

- When $\alpha=0$, we have the following known facts corresponding to bivariate normal distribution: $\mathrm{Cov}(X_1,X_2)=\rho\sigma_1\sigma_2$ and $\rho(X_1,X_2)=\rho$.
- When $\rho=0$, we have $\mathrm{Cov}(X_1,X_2)=\sigma_1\sigma_2\alpha$ and $\rho(X_1,X_2)=\alpha/(1+\alpha^2)$.
- When $\rho=\alpha=0$, X_1 and X_2 are independent.

7. Stationarity and ergodicity

In this section, we provide stationarity and ergodicity properties of the BN distribution.

Definition 7.1 A process X_t is strict-sense stationary (SSS) if its finite-dimensional distributions at times $t_1<\cdots<t_n$, $\forall n\in\mathbb{N}$, are the same after any time interval of length time interval of length t_0. Thus, for each $n\in\mathbb{N}$ and $t_1<\cdots<t_n$ and $(x_1,\ldots,x_n)^\top\in\mathbb{R}^n$ we have

$$\mathbb{P}(X_{t_1+t_0}\leq x_1,\ldots,X_{t_n+t_0}\leq x_n)=\mathbb{P}(X_{t_1}\leq x_1,\ldots,X_{t_n}\leq x_n),$$

for any time t_0.

We say that a process X_t is a BN random process if $X_t \sim \mathrm{BN}(\boldsymbol{\theta}_t)$, where $\boldsymbol{\theta}_t = (\mu_t, \sigma_t, \alpha)^\top$, with $\mu_t \in \mathbb{R}$, $\sigma_t > 0$ and $\alpha \in \mathbb{R}$.

PROPOSITION 7.2 The BN random process is not SSS when μ_t and σ_t are not independent of time.

PROOF If a random process is SSS, then all expected values of functions of the random process must also be stationary. Since $\mathbb{E}(X_t) = \mu_t$ and $\mathrm{Var}(X_t) = \sigma_t^2(1+\alpha^2)$ (see properties P.7 and P.8 in Section 2) change depend on t, we have that the PDF changes with time. Then, the non-stationarity of random process follows. ■

DEFINITION 7.3 A process X_t is weak-sense stationary (WSS) if:

- $\mathbb{E}(X_t) = \mu$ is independent over time;
- $\mathbb{E}(X_t^2) < \infty$;
- $C_X(t,s) = \mathrm{Cov}(X_t, X_s)$ only depends on the distance between the times considered.

If X_t is a BN random process, it is known that $\mathbb{E}(X_t) = \mu_t$, $\mathbb{E}(X_t^2) = \sigma_t^2(1+\alpha^2) + \mu_t^2$ (see Section 2) and that $C_X(t,s) \overset{(20)}{=} \sigma_t \sigma_s[\rho(1+\alpha^2) + (1-\rho^2)\alpha]$. Then, the next result follows.

PROPOSITION 7.4 The BN random process is not WSS when μ_t and σ_t are not independent of time.

Remark 1 In the case that μ_t and σ_t [or $\rho = \alpha = 0$] are independent of time, it is clear that the BN process is SSS and WSS.

In many real-life situations, it is not always possible to have many realizations of the random process available to estimate a population parameter (for example, the mean, variance and covariance function of process), as in classical estimation, but rather a single one. In this case, in order to study the process, we calculate the temporal characteristic of the process.

DEFINITION 7.5 Let X_t be a random process. Then, we define the temporal mean of X_t as

$$\langle m_X \rangle_T = \frac{1}{2T} \int_{-T}^{T} X_t \, \mathrm{d}t, \quad T > 0.$$

DEFINITION 7.6 A process X_t with mean μ independent of time is mean ergodic if

$$\lim_{T\to\infty} \mathrm{Var}(\langle m_X \rangle_T) = \lim_{T\to\infty} \mathbb{E}(\langle m_X \rangle_T - \mu)^2 = 0.$$

PROPOSITION 7.7 The BN random process with mean μ independent of time is mean ergodic whenever

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \sigma_t \, \mathrm{d}t = 0. \tag{21}$$

For example, we can take $\sigma_t = \exp(-t^2)$.

PROOF A simple calculus shows that

$$\mathrm{Var}(\langle m_X \rangle_T) = \frac{1}{4T^2} \int_{-T}^{T} \int_{-T}^{T} C_X(t,t') \, \mathrm{d}t' \mathrm{d}t.$$

Since $C_X(t,s) \overset{(20)}{=} \sigma_t\sigma_{t'}[\rho(1+\alpha^2)+(1-\rho^2)\alpha]$, it follows that

$$\mathrm{Var}(\langle m_X\rangle_T) = [\rho(1+\alpha^2)+(1-\rho^2)\alpha]\left(\frac{1}{2T}\int_{-T}^{T}\sigma_t\,\mathrm{d}t\right)^2.$$

Letting $T\to\infty$ in the above equality, from condition stated in Equation (21), the proof follows. ■

Definition 7.8 A WSS process X_t is covariance-ergodic if

$$\lim_{T\to\infty}\mathrm{Var}\left[\frac{1}{2T}\int_{-T}^{T}(X_t-\mu)(X_{t+s}-\mu)\,\mathrm{d}t\right]=0.$$

When $s=0$ the WSS process is called variance ergodic.

In general, the BN random process X_t is not a WSS process (see Proposition 7.4). Then, it is clear that X_t is not a covariance ergodic process.

Proposition 7.9 When μ_t is independent of time and $\rho=\alpha=0$, the BN process is variance ergodic whenever

$$\lim_{T\to\infty}\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T}\mathrm{Cov}(X_t^2,X_{t'}^2)\,\mathrm{d}t'\mathrm{d}t=\lim_{T\to\infty}\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T}\mathrm{Cov}(X_t^2,X_{t'})\,\mathrm{d}t'\mathrm{d}t=0. \tag{22}$$

Proof When $\rho=\alpha=0$, $C_X(t,t')=0$. A simple calculus shows that

$$\mathrm{Var}\left[\frac{1}{2T}\int_{-T}^{T}(X_t-\mu)^2\,\mathrm{d}t\right]=\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T}\mathrm{Cov}[(X_t-\mu)^2,(X_{t'}-u)^2]\,\mathrm{d}t'\mathrm{d}t.$$

Since $C_X(t,t')=0$, the above expression is given by

$$=\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T}[\mathrm{Cov}(X_t^2,X_{t'}^2)-2\mu\mathrm{Cov}(X_t^2,X_{t'})-2\mu\mathrm{Cov}(X_t,X_{t'}^2)]\,\mathrm{d}t'\mathrm{d}t.$$

By using condition stated in Equation (22), the proof follows. ■

8. A triangular array central limit theorem

In this section, we provide a triangular array central limit theorem for the BN distribution.

Definition 8.1 Two random variables X and Y are said to be positively quadrant dependent (PQD) if, for all $x,y\in\mathbb{R}$,

$$G(x,y)=\mathbb{P}(X>x,Y>y)-\mathbb{P}(X>x)\mathbb{P}(Y>y)\geq 0.$$

It is usual to rewrite $G(x,y)$ using CDFs as

$$G(x,y)=\mathbb{P}(X\leq x,Y\leq y)-\mathbb{P}(X\leq x)\mathbb{P}(Y\leq y). \tag{23}$$

Remark 1 If F is a CDF, for all $x, y \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$, then

$$F(\min\{x,y\}-\alpha)+F(\min\{x,y\}+\alpha) \geq \frac{1}{2}\left[F(x-\alpha)+F(x+\alpha)\right]\left[F(y-\alpha)+F(y+\alpha)\right].$$

Indeed, without loss of generality, assume that $x < y$. Thus,

$$\begin{aligned} F(\min\{x,y\}-\alpha)+F(\min\{x,y\}+\alpha) &= F(x-\alpha)+F(x+\alpha) \\ &\geq \frac{1}{2}\left[F(x-\alpha)+F(x+\alpha)\right]\left[F(y-\alpha)+F(y+\alpha)\right], \end{aligned}$$

because $0 \leq F(y-\alpha)+F(y+\alpha) \leq 2$.

By the stochastic representation of Proposition 4.1, if $X_j \sim \mathrm{BN}(\boldsymbol{\theta}_j)$, there are $Z_j \sim \mathrm{N}(0,1)$ and $A_j \sim \text{Bernoulli}(1/2)$, with $A_j \in \{\pm\alpha\}$, so that $X_j = \sigma_j(Z_j + A_j) + \mu_j$. From now on, in this section, we assume that variables Z_j and A_j are independent of j, that is, we have

$$X_j = \sigma_j(Z+A)+\mu_j. \tag{24}$$

Proposition 8.2 The random variables $X \sim \mathrm{BN}(\boldsymbol{\theta}_X)$ and $Y \sim \mathrm{BN}(\boldsymbol{\theta}_Y)$ are PQD, where $\boldsymbol{\theta}_X = (\mu_X, \sigma_X, \alpha)$, $\mu_X \in \mathbb{R}$, $\sigma_X > 0$ and $\alpha \in \mathbb{R}$.

Proof By Equation (24), $X = \sigma_X(Z+A)+\mu_X$ and $Y = \sigma_Y(Z+A)+\mu_Y$. Then, we get

$$\begin{aligned} \mathbb{P}(X \leq x, Y \leq y) &= \mathbb{P}\Big(Z \leq \frac{x-\mu_X}{\sigma_X} - A, Z \leq \frac{y-\mu_Y}{\sigma_Y} - A\Big) \\ &= \mathbb{P}\Big(Z \leq \min\Big\{\frac{x-\mu_X}{\sigma_X}, \frac{y-\mu_Y}{\sigma_Y}\Big\} - A\Big) = \mathbb{P}(Z \leq \varphi_{n;t_0}(A)). \end{aligned}$$

Let $\widehat{\mathbb{E}}$ be the expectation over A. By the Fubini theorem, we have

$$\begin{aligned} \mathbb{P}(Z \leq \varphi_{n;t_0}(A)) &= \widehat{\mathbb{E}}\Big[\Phi\Big(\min\Big\{\frac{x-\mu_X}{\sigma_X}, \frac{y-\mu_Y}{\sigma_Y}\Big\} - A\Big)\Big] \\ &= \frac{1}{2}\Big[\Phi\Big(\min\Big\{\frac{x-\mu_X}{\sigma_X}, \frac{y-\mu_Y}{\sigma_Y}\Big\} - \alpha\Big) + \Phi\Big(\min\Big\{\frac{x-\mu_X}{\sigma_X}, \frac{y-\mu_Y}{\sigma_Y}\Big\} + \alpha\Big)\Big]. \end{aligned}$$

Therefore,

$$\mathbb{P}(X \leq x, Y \leq y) = \frac{1}{2}\Big[\Phi\Big(\min\Big\{\frac{x-\mu_X}{\sigma_X}, \frac{y-\mu_Y}{\sigma_Y}\Big\} - \alpha\Big) + \Phi\Big(\min\Big\{\frac{x-\mu_X}{\sigma_X}, \frac{y-\mu_Y}{\sigma_Y}\Big\} + \alpha\Big)\Big].$$

Now, by using the identity $\mathrm{erf}(x/\sqrt{2}) = 2\Phi(x)-1$, the CDF stated in Equation (4) of $X \sim \mathrm{BN}(\boldsymbol{\theta}_X)$ is written as

$$\mathbb{P}(X \leq x) = \frac{1}{2}\Big[\Phi\Big(\frac{x-\mu_X}{\sigma_X} - \alpha\Big) + \Phi\Big(\frac{x-\mu_X}{\sigma_X} + \alpha\Big)\Big].$$

Hence, by Remark 1, we get

$$
\begin{aligned}
G(x,y) &\overset{(23)}{=} \mathbb{P}(X \leq x, Y \leq y) - \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) \\
&= \frac{1}{2}\left[\Phi\left(\min\left\{\frac{x-\mu_X}{\sigma_X}, \frac{y-\mu_Y}{\sigma_Y}\right\} - \alpha\right) + \Phi\left(\min\left\{\frac{x-\mu_X}{\sigma_X}, \frac{y-\mu_Y}{\sigma_Y}\right\} + \alpha\right)\right] \\
&\quad - \frac{1}{4}\left[\Phi\left(\frac{x-\mu_X}{\sigma_X} - \alpha\right) + \Phi\left(\frac{x-\mu_X}{\sigma_X} + \alpha\right)\right]\left[\Phi\left(\frac{y-\mu_Y}{\sigma_Y} - \alpha\right) + \Phi\left(\frac{y-\mu_Y}{\sigma_Y} + \alpha\right)\right] \\
&\geq 0.
\end{aligned}
$$

This completes the proof. ■

Definition 8.3 We define a sequence of random variables $\{X_j\}$ to be linearly positive quadrant dependent (LPQD) if for any disjoint A, B and positive $\{\lambda_j\}$, $\sum_{k\in A}\lambda_k X_k$ and $\sum_{l\in B}\lambda_l X_l$ are PQD.

A reasoning similar to the proof of Proposition 8.2 gives the following result.

Proposition 8.4 The sequence of random variables $\{X_j\}$, with $X_j \sim \mathrm{BN}(\boldsymbol{\theta}_j)$, is LPQD, where $\boldsymbol{\theta}_j = (\mu_j, \sigma_j, \alpha)$, $\mu_j \in \mathbb{R}$, $\sigma_j > 0$ and $\alpha \in \mathbb{R}$.

Theorem 8.5 Let $S_n = \sum_{j=1}^{M_n}[X_{n,j} - \mathbb{E}(X_{n,j})]$ where for each n, $X_{n,j} \sim \mathrm{BN}(\boldsymbol{\theta}_{n,j})$, with $\boldsymbol{\theta}_{n,j} = (\mu_{n,j}, \sigma_{n,j}, \alpha)$, $\mu_{n,j} \in \mathbb{R}$, $\sigma_{n,j} > 0$ and $\alpha \in \mathbb{R}$. Suppose there exist $c_1, c_2, c_3 \in (0,\infty)$ and a sequence $u_l \to 0$ so that, for all n, j, l, we have that

$$\sigma_{n,j}^2 \geq c_1, \ \sigma_{n,j}^3 \leq c_2; \tag{25}$$

$$\sum_{k=1}^{M_n} \sigma_{n,j}\sigma_{n,k} \leq c_3; \tag{26}$$

$$\sum_{\substack{k=1 \\ |k-j|\geq l}}^{M_n} \sigma_{n,j}\sigma_{n,k} \leq u_l. \tag{27}$$

Then,

$$\lim_{n\to\infty} \mathbb{P}\big([\mathrm{Var}(S_n)]^{-1/2} S_n \leq x\big) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} \exp(-x^2/2)\,\mathrm{d}x, \quad \forall x \in \mathbb{R}.$$

Proof Since for each n, $\{X_{n,j}\}$ is LPQD (see Proposition 8.4) but not SSS (see Proposition 7.2), by Cox and Grimmett (1984), it is enough to verify that

$$\operatorname{Var}(X_{n,j}) \geq \widetilde{c}_1, \ \mathbb{E}|X_{n,j} - \mathbb{E}(X_{n,j})|^3 \leq \widetilde{c}_2; \tag{28}$$

$$\sum_{k=1}^{M_n} \operatorname{Cov}(X_{n,j}, X_{n,k}) \leq \widetilde{c}_3; \tag{29}$$

$$\sum_{\substack{k=1 \\ |k-j|\geq l}}^{M_n} \operatorname{Cov}(X_{n,j}, X_{n,k}) \leq \widetilde{u}_l; \tag{30}$$

where $\widetilde{u}_l \to 0$. Indeed, since, by Equation (25), $\sigma_{n,j}^2 \geq c_1$ and $\operatorname{Var}(X_{n,j}) = \sigma_{n,j}^2(1+\alpha^2)$ (see property P.8 in Section 2), we have that $\operatorname{Var}(X_{n,j}) \geq \sigma_{n,j}^2 \geq c_1 = \widetilde{c}_1$. Moreover, using the representation in Equation (24) and the condition given in Equation (25), we obtain

$$\mathbb{E}|X_{n,j} - \mathbb{E}(X_{n,j})|^3 = \sigma_{n,j}^3 \mathbb{E}|Z + A|^3 \leq 6\sqrt{2/\pi}\, \sigma_{n,j}^3 \leq 5c_2 = \widetilde{c}_2,$$

that is, Equation (28) is satisfied.

Now, since $\operatorname{Cov}(X_{n,j}, X_{n,k}) \overset{(20)}{=} \sigma_{n,j}\sigma_{n,k}[\rho(1+\alpha^2) + (1-\rho^2)\alpha]$, by conditions given in Equations (26) and (27), the statements in Equations (29) and (30) follow by taking $\widetilde{c}_3 = c_3[\rho(1+\alpha^2) + (1-\rho^2)\alpha]$ and $\widetilde{u}_l = [\rho(1+\alpha^2) + (1-\rho^2)\alpha]u_l$, respectively. ■

Remark 2 The set of $\sigma_{n,k}$ satisfying conditions stated in Equations (25), (26) and (27) is non-empty. Indeed, let us take $M_n = n$ and $\sigma_{n,k} = r^{-k}$, with $k \geq 1$ and $r > 1$, for all n. Immediately, we have $\sigma_{n,k} > 0$ and $\sigma_{n,k} \leq 1$, that is, Equation (25) is valid. Moreover,

$$\sum_{k=1}^{n} \sigma_{n,j}\sigma_{n,k} \leq \sum_{k=1}^{n} \sigma_{n,k} \leq \sum_{k=1}^{\infty} \frac{1}{r^k} = \frac{1}{r-1}, \quad r > 1.$$

Then, Equation (26) is satisfied. Thus, since $r^{|k-j|} \leq r^{j+k}$ for $r > 1$, we have $\sigma_{n,j}\sigma_{n,k} = r^{-(j+k)} \leq r^{-|k-j|}$. Therefore,

$$\sum_{\substack{k=1 \\ |k-j|\geq l}}^{n} \sigma_{n,j}\sigma_{n,k} \leq \sum_{\substack{k=1 \\ |k-j|\geq l}}^{n} \frac{1}{r^{|k-j|}} \leq \sum_{\substack{k=1 \\ |k-j|\geq l}}^{\infty} \frac{1}{r^{|k-j|}} = \sum_{m=l}^{\infty} \frac{1}{r^m} \left[\sum_{\substack{k=1 \\ |k-j|=m}}^{\infty} 1 \right],$$

where in the last equality we rearrange the summation terms. Since $\left[\sum_{k:|k-j|=m} 1\right]$ is the number of vertices at the boundary of the one-dimensional ball of radius m centered at j, there is $C > 0$ independent of j such that $\left[\sum_{k:|k-j|=m} 1\right] = C$. Hence,

$$\sum_{\substack{k=1 \\ |k-j|\geq l}}^{n} \sigma_{n,j}\sigma_{n,k} \leq C \sum_{m=l}^{\infty} \frac{1}{r^m} = u_l.$$

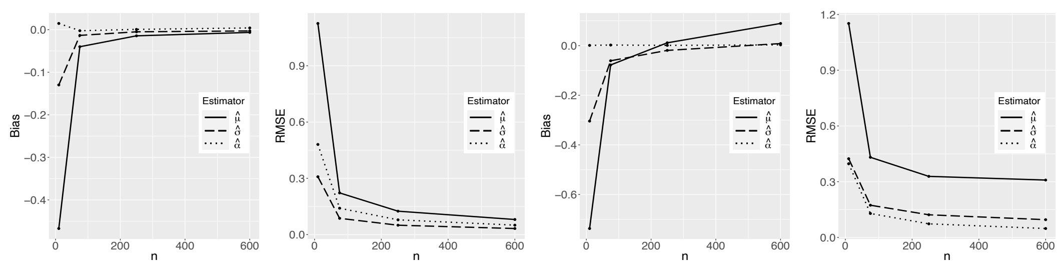
As $\sum_{m=0}^{\infty} r^{-m} = r(r-1)^{-1} < \infty$, for $r > 1$, it follows that $u_l \to 0$, when $l \to \infty$. Then, Equation (27) follows.

9. Numerical evaluation

In this section, a Monte Carlo simulation study was carried out to evaluate the performance of the ML estimators of the BN model; see Section 5. All numerical evaluations were done in the `R` software; see R-Team (2020).

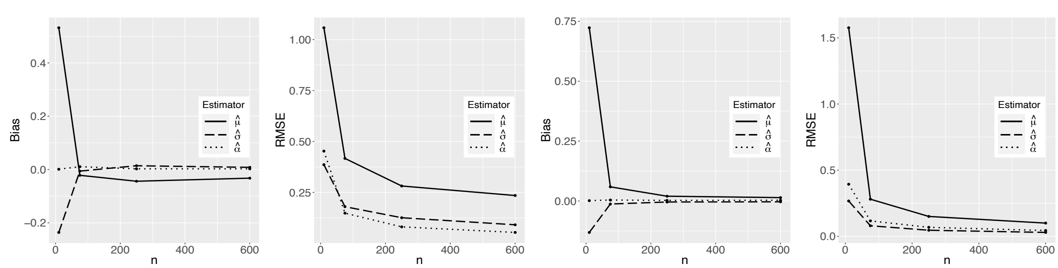
The simulation scenario considers sample size $n \in \{10, 75, 250, 600\}$, location parameter $\mu = 0.5$, scale parameter $\sigma = 1.0$, location parameter $\alpha \in \{-2.0, -0.5, 0.8, 3.0\}$, with 1,000 Monte Carlo replications for each combination of above-given parameters and sample size. The values of the location parameter α have been chosen in order to study the performance under unimodality and bimodality.

The ML estimation results for the considered BN model are presented in Figures 1-2. The empirical bias and root mean squared error (RMSE) are reported. A look at the results in Figures 1-2 allows us to conclude that, as the sample size increases, the empirical bias and RMSE both decrease, as expected. Moreover, we note that the performance of the estimate of μ is better when $|\alpha| > 1$ under bimodality.



(a) $\alpha = -2.0$ (b) $\alpha = -2.0$ (c) $\alpha = -0.5$ (d) $\alpha = -0.5$

Figure 1. Empirical bias and RMSE from simulated data for the indicated ML estimates of the listed BN model parameters, n and α.



(a) $\alpha = 0.8$ (b) $\alpha = 0.8$ (c) $\alpha = 3.0$ (d) $\alpha = 3.0$

Figure 2. Empirical bias and RMSE from simulated data for the indicated ML estimates of the listed BN model parameters, n and α.

10. Concluding remarks

We have stated novel properties of the bimodal normal distribution and discussed some mathematical properties, as well as proven its bimodality and identifiability. We have also analyzed some aspects related to the maximum likelihood estimation and its associated asymptotic properties. We have derived a bivariate version of the bimodal normal distribution and studied some of its characteristics such as covariance and correlation. We have considered stationarity and ergodicity as well as a triangular array central limit theorem. Finally, we have carried out Monte Carlo simulations to evaluate the behavior of the maximum likelihood estimators.

A possible limitation of our proposal might be associated with the moments. In this work, we have only derived the raw moments (moments of positive integer order). It would be interesting, if possible, to find the real moments. In addition, this work has studied consistency and a central limit theorem for one of the model parameters (since the others are known). Note that this parameter generates bimodality. It would be interesting that, using a more elaborate approach, considering an unknown parameter vector. As further research, one might explore the multivariate case and then obtain ergodicity and stationarity results.

Author contributions Conceptualization, R.V., H.S., J.R.; methodology, R.V., H.S., J.R.; software, R.V., H.S., J.R.; validation, R.V., H.S., J.R.; formal analysis, R.V., H.S., J.R.; investigation, R.V., H.S., J.R.; data curation, R.V., H.S., J.R.; writing-original draft preparation, R.V., H.S., J.R.; writing-review and editing, R.V., H.S., J.R.; visualization, R.V., H.S., J.R.; supervision, R.V., H.S., J.R. All authors have read and agreed the published version of the manuscript.

Acknowledgements The authors would like to thank the reviewers for helpful comments on our manuscript.

Funding The authors thank CNPq for the financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) (Finance Code 001).

Conflicts of interest The authors declare no conflict of interest.

References

Alizadeh, M., MirMostafaee, S.M.T.K, and Ghosh, I., 2017. A new extension of power Lindley distribution for analyzing bimodal data. Chilean Journal of Statistics, 8, 67–86.

Bahadur, R.R., 1971. Some Limit Theorems in Statistics. SIAM, Philadelphia.

Cox, J.T. and Grimmett, G., 1984. Central limit theorems for associated random variables and the percolation model. The Annals of Probability, 12, 514–528.

Cramér, H., 1946. Mathematical Methods of Statistics. Princeton University Press, New Jersey.

Eugene, N., Lee, C., and Famoye, F., 2002. Beta-normal distribution and its applications. Communications in Statistics: Theory and Methods, 31, 497–512.

Gómez-Déniz, E., Sarabia, J.M., and Calderón-Ojeda, E., 2021. Bimodal normal distribution: Extensions and applications. Journal of Computational and Applied Mathematics, 388, 113292.

Hassan, M.Y. and El-Bassiouni, M.Y., 2016. Bimodal skew-symmetric normal distribution. Communications in Statistics: Theory and Methods, 45, 1527-1541.

R-Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Saulo, H., Leao, J., Vila, R., Leiva, V., and Tomazella, V., 2020. On mean-based bivariate Birnbaum-Saunders distributions: properties, inference and application. Communications in Statistics: Theory and Methods 49, 6032–6056.

Teicher, H., 1963. Identifiability of finite mixtures. The Annals of Mathematical Statistics, 34, 1265–1269.

Vila, R., Leão, J., Saulo, H., Shahzad, M.N., and Santos-Neto, M., 2020. On a bimodal Birnbaum-Saunders distribution with applications to lifetime data. Brazilian Journal of Probability and Statistics, 34, 495–518.

Winkelbauer, A., 2014. Moments and absolute moments of the normal distribution. Preprint. ArXiv:1209.4340.

Information for authors

The editorial board of the Chilean Journal of Statistics (ChJS) is seeking papers, which will be refereed. We encourage the authors to submit a PDF electronic version of the manuscript in a free format to the Editors-in-Chief of the ChJS (E-mail: chilean.journal.of.statistics@gmail.com). Submitted manuscripts must be written in English and contain the name and affiliation of each author followed by a leading abstract and keywords. The authors must include a"cover letter" presenting their manuscript and mentioning: "We confirm that this manuscript has been read and approved by all named authors. In addition, we declare that the manuscript is original and it is not being published or submitted for publication elsewhere".

Preparation of accepted manuscripts

Manuscripts accepted in the ChJS must be prepared in Latex using the ChJS format. The Latex template and ChJS class files for preparation of accepted manuscripts are available at http://soche.cl/chjs/files/ChJS.zip. Such as its submitted version, manuscripts accepted in the ChJS must be written in English and contain the name and affiliation of each author, followed by a leading abstract and keywords, but now mathematics subject classification (primary and secondary) are required. AMS classification is available at http://www.ams.org/mathscinet/msc/. Sections must be numbered 1, 2, etc., where Section 1 is the introduction part. References must be collected at the end of the manuscript in alphabetical order as in the following examples:

Arellano-Valle, R., 1994. Elliptical Distributions: Properties, Inference and Applications in Regression Models. Unpublished Ph.D. Thesis. Department of Statistics, University of São Paulo, Brazil.

Cook, R.D., 1997. Local influence. In Kotz, S., Read, C.B., and Banks, D.L. (Eds.), Encyclopedia of Statistical Sciences, Vol. 1., Wiley, New York, pp. 380-385.

Rukhin, A.L., 2009. Identities for negative moments of quadratic forms in normal variables. Statistics and Probability Letters, 79, 1004-1007.

Stein, M.L., 1999. Statistical Interpolation of Spatial Data: Some Theory for Kriging. Springer, New York.

Tsay, R.S., Peña, D., and Pankratz, A.E., 2000. Outliers in multivariate time series. Biometrika, 87, 789-804.

References in the text must be given by the author's name and year of publication, e.g., Gelfand and Smith (1990). In the case of more than two authors, the citation must be written as Tsay et al. (2000).

Copyright