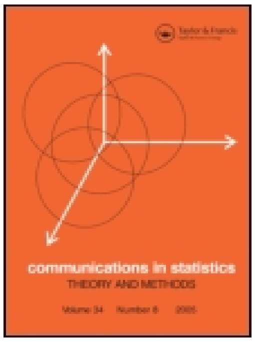

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: http://www.tandfonline.com/loi/lsta20

On a nonlinear Birnbaum–Saunders model based on a bivariate construction and its characteristics

Mohsen Khosravi, Víctor Leiva, Ahad Jamalizadeh & Emilio Porcu

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2016, VOL. 45, NO. 3, 772–793
http://dx.doi.org/10.1080/03610926.2013.851223

Taylor & Francis
Taylor & Francis Group

On a nonlinear Birnbaum–Saunders model based on a bivariate construction and its characteristics

Mohsen Khosravi[a], Víctor Leiva[b,c], Ahad Jamalizadeh[a], and Emilio Porcu[d]

[a]Department of Statistics, Shahid Bahonar University of Kerman, Iran; [b]Institute of Statistics, Universidad de Valparaíso, Chile; [c]Faculty of Engineering and Sciences, Universidad Adolfo Ibáñez, Chile; [d]Department of Mathematics, Universidad Federico Santa María, Chile

ABSTRACT

The Birnbaum-Saunders (BS) distribution is an asymmetric probability model that is receiving considerable attention. In this article, we propose a methodology based on a new class of BS models generated from the Student-*t* distribution. We obtain a recurrence relationship for a BS distribution based on a nonlinear skew–*t* distribution. Model parameters estimators are obtained by means of the maximum likelihood method, which are evaluated by Monte Carlo simulations. We illustrate the obtained results by analyzing two real data sets. These data analyses allow the adequacy of the proposed model to be shown and discussed by applying model selection tools.

ARTICLE HISTORY
Received 19 December 2012
Accepted 25 September 2013

KEYWORDS
Data analysis; Elliptically contoured distributions; Likelihood and Monte Carlo methods; Linear and nonlinear skew–elliptic distributions

MATHEMATICS SUBJECT CLASSIFICATION
Primary 60E05; Secondary 62–07

1. Introduction

The Birnbaum–Saunders (BS) distribution was originally introduced by Birnbaum and Saunders (1969) for modeling the fatigue lifetime of a specimen subjected to cyclic stress. For more details about the BS distribution, see Johnson et al. (1995, pp. 651–663). Although the BS distribution finds its genesis in engineering, it has been successful in modeling a wide variety of positively skewed, nonnegative data, including business, environmental, forest, human behavior, informatics, medical, mortality and toxicological applications; see, e.g., Balakrishnan et al. (2007, 2009a), Leiva et al. (2007, 2008a, 2009, 2010a, 2010b, 2011, 2012, 2014), Barros et al. (2008), Ahmed et al. (2010), Vilca et al. (2010, 2011), Paula et al. (2012), Ferreira et al. (2012), and Marchant et al. (2013a, b).

Considerable attention has been paid to the construction of flexible families of distributions, which are useful for diverse practical applications. For example, the normal distribution has been extended to more general families of symmetric and skewed distributions, all of them being often very flexible for describing diverse types of data. For more details about elliptically contoured (EC) distributions, that in the univariate case are simply symmetric distributions in $\mathbb{R}$, and about skewed distributions; see Fang et al. (1990), Gupta and Varga (1993), and Genton (2004).

Because a random variable (RV) following the BS distribution can be considered as a monotone transformation of a RV with normal distribution, generalizations of the BS distribution can be obtained by switching the distribution of the basis RV by using diverse arguments,

CONTACT Víctor Leiva victorleivasanchez@gmail.com www.victorleiva.cl

allowing us to construct more general classes of models. Several extensions and generalizations of the BS distribution have been proposed by a number of authors, including Díaz–García and Leiva (2005), Vilca and Leiva (2006), Sanhueza et al. (2008), Gómez et al. (2009), Guiraud et al. (2009), Balakrishnan et al. (2009b), Ahmed et al. (2010), Ferreira et al. (2012), and Santos–Neto et al. (2012), among others, which allow a major degree of flexibility to be obtained for this distribution.

Skew distributions have been widely studied; see, e.g., Azzalini (1985, 2005) and Genton (2004). When the skew–normal (SN) distribution is used in place of the normal distribution, outliers located at the left–tail are usually accommodated in a better way; see Vilca et al. (2011). Branco et al. (2012) analyzed skew–t distributions, where the linear and nonlinear terms associated with this type of distributions were used. Following this, we adopt the terms of linear and nonlinear skew–EC (SEC) distributions, which are based on a bivariate skew–elliptic distribution. Díaz–García and Leiva (2005) and Vilca and Leiva (2006) generated models with good properties for the BS distribution based on EC and SEC models, respectively.

The objectives of this article are: (i) to propose a new family of BS distributions, (ii) to find estimates of the parameters of this new distribution, and (iii) to apply the obtained results to real-world data. We focus on the Student–t (or simply t) distribution, due to its important properties and because this is often used as an alternative model to the normal distribution. We establish a recurrence relation for a BS distribution based on a nonlinear skew–t distribution. This relationship is valid for any positive value of the shape parameter of the t distribution, known as degrees of freedom (DF), and allows a recursive evaluation of the cumulative distribution function (CDF) for any other value of this parameter to be obtained. The paper is organized as follows. In Sec. 2, we provide a background about BS, EC and linear SEC (LSE) distributions. In Sec. 3, we consider nonlinear SEC (NSE) distributions and then propose the new model for the BS distribution termed as the nonlinear SEC BS distribution, in short NSEBS. In this section, the aforementioned recurrence relationships are established. In Sec. 4, we propose ML estimators for the parameters of the NSEBS distribution. In Sec. 5, we carry out the numerical part of this study by using simulated and real-world data. The performance of the ML estimators is evaluated by the Monte Carlo method, whereas potential applications of the NSEBS distribution are explored by two real data sets and the ML estimation method. In Sec. 6, we furnish some concluding remarks of this work.

2. Background

In this section, we provide some preliminary notions of BS, EC, and SEC distributions.

2.1. *The BS distribution*

The CDF and probability density function (PDF) of a BS RV, say $T_{\rm N} \sim {\rm BS}(\alpha, \beta)$, are

$$F_{T_{\rm N}}(t;\alpha,\beta) = \Phi\left(a(t;\alpha,\beta)\right) \quad \text{and} \quad f_{T_{\rm N}}(t;\alpha,\beta) = \phi\left(a(t;\alpha,\beta)\right) A(t;\alpha,\beta), \quad t > 0, \tag{1}$$

where $\alpha > 0$ and $\beta > 0$ are shape and scale parameters, and Φ and ϕ denote the standard normal CDF and PDF, respectively, i.e., $\Phi(z) = \int_{-\infty}^{z} \phi(x)\,{\rm d}x$, with $\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$, for $x \in \mathbb{R}$. In addition, $a(t;\alpha,\beta) = [1/\alpha][\sqrt{t/\beta} - \sqrt{\beta/t}]$ and $A(t;\alpha,\beta) = {\rm d}a(t;\alpha,\beta)/{\rm d}t = [t+\beta]/[2\alpha\sqrt{\beta}\sqrt{t^3}]$.

2.2. EC distributions

A RV X has a univariate EC distribution if its PDF is given by

$$f_{\rm EC}\big(x;\mu,\sigma^2,h^{(1)}\big)=\frac{c}{\sigma}h^{(1)}\big([x-\mu]^2/\sigma^2\big),\ x\in\mathbb{R},\ \mu\in\mathbb{R},\sigma>0,$$

where $h^{(1)}$ is the kernel of the PDF of X and c is its normalizing constant, which is denoted by $X\sim{\rm EC}(\mu,\sigma^2,h^{(1)})$. More specifically, $h^{(1)}$ is a mapping from $\mathbb{R}_+\cup\{0\}$ to $\mathbb{R}_+$, that is, $h^{(1)}\colon\mathbb{R}_+\cup\{0\}\longrightarrow\mathbb{R}_+$. Two particular EC distributions are the normal and $t(\nu)$ cases, where $h^{(1)}(u)=[1/(2\pi)^{1/2}]\exp(-u/2)$ and $h^{(1)}(u)=[\Gamma([\nu+1]/2)/\{\Gamma(\nu/2)(\nu\pi)^{1/2}\}][1+u/\nu]^{-[\nu+1]/2}$, respectively, for $u>0$. In general, for the p–variate normal and $t(\nu)$ distributions,

$$h^{(p)}(u)=\frac{1}{[2\pi]^{p/2}}\exp(-u/2)\quad\text{and}\quad h^{(p)}(u)=\frac{\Gamma([\nu+p]/2)}{\Gamma(\nu/2)\,[\nu\pi]^{p/2}}\left[1+\frac{u}{\nu}\right]^{-\frac{\nu+p}{2}},\quad u>0,\tag{2}$$

respectively, with $h^{(p)}\colon\mathbb{R}_+\cup\{0\}\longrightarrow\mathbb{R}_+$; see Jamalizadeh et al. (2009b).

2.3. ECBS distributions

Díaz–García and Leiva (2005) introduced a generalized BS distribution by replacing the standard normal CDF Φ in (1) with the CDF of an EC distribution, in short ECBS. A RV $T_{\rm EC}$ is said to have the ECBS distribution, with parameters α and β, and kernel $h^{(1)}$, if its CDF is given by

$$F_{T_{\rm EC}}\big(t;\alpha,\beta,h^{(1)}\big)=F_{\rm EC}\big(a(t;\alpha,\beta),h^{(1)}\big),$$

where $F_{\rm EC}(\cdot;\ h^{(1)})$ stands for the EC(0, 1, $h^{(1)}$) CDF, which is denoted by $T_{\rm EC}\sim{\rm ECBS}(\alpha,\beta,h^{(1)})$. Two particular cases of ECBS distributions are the BS and BS–$t(\nu)$ models. The CDF and PDF of a BS–$t(\nu)$ RV, say $T_t\sim$ BS–$t(\alpha,\beta,\nu)$, are

$$F_{T_t}(t;\alpha,\beta)=G(a(t;\alpha,\beta),\nu)\quad\text{and}\quad f_{T_t}(t;\alpha,\beta)=g(a(t;\alpha,\beta),\nu)\,A(t;\alpha,\beta),\quad t>0,$$

where $\nu>0$ are the DFs, and G, g denote the $t(\nu)$ CDF and PDF, respectively, given by

$$G(z;\nu)=\frac{1}{2}\left[1+I_{\frac{z^2}{z^2+\nu}}\left(\tfrac{1}{2},\tfrac{\nu}{2}\right)\right]\ \text{and}$$

$$g(z;\nu)=\frac{\rm d}{{\rm d}z}G(z;\nu)=\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\,\pi}\,\Gamma(\nu/2)}\left[1+z^2/\nu\right]^{-\frac{\nu+1}{2}},\tag{3}$$

for $z\in\mathbb{R}$ and $\nu>0$, where $I_x(a,b)=\int_0^x t^{a-1}[1-t]^{b-1}\,{\rm d}t/\int_0^1 t^{a-1}[1-t]^{b-1}\,{\rm d}t$ is the incomplete beta function ratio.

2.4. SEC and BS distributions

Let the RVs X_1 and X_2 be independent identically EC(0, 1, $h_{(1)}$) distributed. Then, the RV $X_{\lambda,h^{(1)}}$ is said to have a LSE distribution with parameter λ and kernel $h^{(1)}$, which is denoted by $X_{\lambda,h^{(1)}}\sim{\rm LSE}(\lambda,h^{(1)})$, if $X_{\lambda,h^{(1)}}$ has the same distribution as X_1 conditional to $(X_2<\lambda X_1)$, that is, $X_{\lambda,h^{(1)}}\overset{\rm d}{=}X_1\mid(X_2<\lambda X_1)$, where "$\overset{\rm d}{=}$" means equal in distribution. Hence, the PDF of $X_{\lambda,h^{(1)}}$ is

$$f_{\rm LSE}\big(x;\lambda,h^{(1)}\big)=2f_{\rm EC}\big(x;h^{(1)}\big)F_{\rm EC}\big(\lambda x;h^{(1)}\big),\quad x\in\mathbb{R},\lambda\in\mathbb{R},$$

where $f_{\rm EC}(\,\cdot\,;h^{(1)})$ is the EC(0, 1, $h^{(1)}$) PDF.

A RV T_{LSE} is said to have a linear skew–EC BS distribution, in short LSEBS (see Vilca and Leiva, 2006), with parameters α, β, and λ, and kernel $h^{(1)}$, if its CDF is given by

$$F_{T_{\text{LSE}}}\big(t;\alpha,\beta,\lambda,h^{(1)}\big) = F_{\text{LSE}}\big(a(t;\alpha,\beta),\lambda,h^{(1)}\big),$$

where $F_{\text{LSE}}(\,\cdot\,;\lambda,h^{(1)})$ is the LSE$(\lambda,h^{(1)})$ CDF and $a(t;\alpha,\beta)$ is as given in (1), which is denoted by $T_{\text{LSE}} \sim \text{LSEBS}(\alpha,\beta,\lambda,h^{(1)})$. In the normal case, the PDF of T_{LSE} reduces to

$$f_{T_{\text{LSN}}}(t;\alpha,\beta,\lambda) = 2\phi(a(t;\alpha,\beta))\Phi(\lambda\, a(t;\alpha,\beta))\,A(t;\alpha,\beta), \quad t>0,$$

for $A(t;\alpha,\beta)$ given also in (1), whereas for the $t(\nu)$ case, we have

$$f_{T_{\text{LS-}t}}(t;\alpha,\beta,\lambda,\nu) = 2g(a(t;\alpha,\beta),\nu)\,G(\lambda a(t;\alpha,\beta),\nu)A(t;\alpha,\beta), \quad t>0,$$

where g and G are given in (3).

3. The nonlinear skew–elliptic BS distribution

In this section, we consider a NSE distribution and then we use it as kernel for constructing the NSEBS distribution.

3.1. *The nonlinear skew–elliptic distribution*

Let $\boldsymbol{X}$ be a bivariate random vector following a standard EC distribution, that is, $\boldsymbol{X} = (X_1,X_2)^\top \sim \text{EC}_2(\boldsymbol{0},\boldsymbol{I}_2,h^{(2)})$, where $h^{(2)}\colon \mathbb{R}_+ \cup \{0\} \longrightarrow \mathbb{R}_+$. Two particular cases of bivariate EC distributions are the normal and $t(\nu)$ models, where, according to expression given in (2), $h^{(2)}(u) = [1/2\pi]\exp(-u/2)$ and $h^{(2)}(u) = [\Gamma(\nu/2+1)/\{\Gamma(\nu/2)\,\nu\pi\}][1+u/\nu]^{-\nu/2-1}$, respectively, for $u>0$.

A RV $X_{\lambda,h^{(2)}}$ is said to have a NSE distribution with parameter $\lambda \in \mathbb{R}$ and kernel $h^{(2)}$, which is denoted by $X_{\lambda,h^{(2)}} \sim \text{NSE}(\lambda,h^{(2)})$, if

$$X_{\lambda,h^{(2)}} \stackrel{\text{d}}{=} X_1 \mid (X_2 < \lambda X_1).$$

We can easily show that the PDF of $X_{\lambda,h^{(2)}}$ is

$$f_{\text{NSE}}\big(x;\lambda,h^{(2)}\big) = 2 f_{\text{EC}}\big(x;h^{(1)}\big)F_{\text{EC}}\big(\lambda x;h^{(1)}_{x^2}\big), \quad x\in\mathbb{R},\ \lambda\in\mathbb{R},$$

where $f_{\text{EC}}(\,\cdot\,;h^{(1)})$ and $F_{\text{EC}}(\,\cdot\,;h^{(1)})$ are the EC$(0,1,h^{(1)})$ PDF and CDF, respectively, $F_{\text{EC}}(\cdot;h^{(1)}_{x^2})$ is the EC$(0,1,h^{(1)}_{x^2})$ CDF, and

$$h^{(1)}_{x^2}(u) = \frac{h^{(2)}(u+x^2)}{h^{(1)}(x^2)}, \quad u \geq 0, \tag{4}$$

for $h^{(1)}>0$, as mentioned. It is worth to stress that, in the normal case, the LSE and NSE distributions are identical and its PDF is $f_{\text{NLSN}}(x,\lambda) = f_{\text{LSN}}(x;\lambda) = 2\phi(x)\Phi(\lambda x)$, for $x\in\mathbb{R}$ and $\lambda\in\mathbb{R}$. However, in the $t(\nu)$ case, the PDFs of the LSE and NSE distributions are

$$f_{\text{LS-}t}(x;\nu,\lambda) = 2g(x;\nu)\,G(\lambda x;\nu), \quad x\in\mathbb{R}, \lambda\in\mathbb{R}, \tag{5}$$

and

$$f_{\text{NLS-}t}(x;\nu,\lambda) = 2g(x;\nu)\,G\left(\lambda x\sqrt{\frac{\nu+1}{\nu+x^2}};\nu+1\right), \quad x\in\mathbb{R}, \lambda\in\mathbb{R}, \tag{6}$$

where G and g are the $t(\nu)$ CDF and PDF, respectively; see, e.g., Azzalini and Capitanio (2003).

3.2. A BS distribution based on the NSE model

Following Vilca and Leiva (2006), we introduce the NSEBS distribution. The RV T_{NSE} is said to have the NSEBS distribution with parameters α, β, λ, and kernel $h^{(2)}$, if its CDF can be expressed as

$$F_{T_{\mathrm{NSE}}}\left(t;\alpha,\beta,\lambda,h^{(2)}\right)=F_{\mathrm{NSE}}\left(a(t;\alpha,\beta),\lambda,h^{(2)}\right),\quad t>0, \tag{7}$$

for $\alpha>0$, $\beta>0$ and $\lambda\in\mathbb{R}$, which is denoted by $T_{\mathrm{NSE}}\sim \mathrm{NSEBS}(\alpha,\beta,\lambda,h^{(2)})$. It is clear that

$$Z=\frac{1}{\alpha}\left[\sqrt{\frac{T_{\mathrm{NSE}}}{\beta}}-\sqrt{\frac{\beta}{T_{\mathrm{NSE}}}}\right]\sim \mathrm{NSE}\left(\lambda,h^{(2)}\right)\quad\text{and}\quad T_{\mathrm{NSE}}=\beta\left[\frac{\alpha Z}{2}+\sqrt{\left(\frac{\alpha Z}{2}\right)^2+1}\right]^2. \tag{8}$$

Thus, if $T_{\mathrm{NSE}}\sim \mathrm{NSEBS}(\alpha,\beta,\lambda,h^{(2)})$, then the PDF of T_{NSE} is

$$f_{T_{\mathrm{NSE}}}(t;\alpha,\beta,\lambda,h^{(2)})=2f_{\mathrm{EC}}(a(t;\alpha,\beta),h^{(1)})F_{\mathrm{EC}}(\lambda a(t;\alpha,\beta),h^{(1)}_{a^2(t;\alpha,\beta)})A(t;\alpha,\beta),\ t>0, \tag{9}$$

with $h^{(1)}_{a^2(t;\alpha,\beta)}$ being given in an analogous way as in (4). Let us now to introduce some properties of the distribution of the RV $T_{\mathrm{NSE}}\sim \mathrm{NSEBS}(\alpha,\beta,\lambda,h^{(2)})$.

Theorem 3.1. *Let $T_{NSE}\sim NSEBS(\alpha,\beta,\lambda,h^{(2)})$. Then:*
(i) $aT_{NSE}\sim NSEBS(\alpha,a\beta,\lambda,h^{(2)})$, for $a>0$, and
(ii) $1/T_{NSE}\sim NSEBS(\alpha,1/\beta,-\lambda,h^{(2)})$.

Proof. The proofs of (i)–(ii) follow directly by a change of variable. □

Theorem 3.2. *Let $T_{NSE}\sim NSEBS(\alpha,\beta,\lambda,h^{(2)})$ and $T_{EC}\sim ECBS(\alpha,\beta,h^{(1)})$. Then, their mean, variance, and coefficients of variation (CV), skewness (CS) and kurtosis (CK) satisfy the following relationships:*
(i) $\mathrm{E}[T_{\mathrm{NSE}}]=\mathrm{E}[T_{\mathrm{EC}}]+\frac{\alpha\beta}{2}\omega_1$;
(ii) $\mathrm{V}[T_{\mathrm{NSE}}]=\mathrm{V}[T_{\mathrm{EC}}]+[\frac{\alpha\beta}{2}]^2\alpha_\omega$;
(iii) $\mathrm{CV}[T_{\mathrm{NSE}}]=\mathrm{CV}[T_{\mathrm{EC}}]\frac{\sqrt{1+[\alpha\beta]^2\alpha_\omega/\mathrm{V}[2T_{\mathrm{EC}}]}}{1+\alpha\beta\omega_1/\mathrm{E}[2T_{\mathrm{EC}}]}$;
(iv) $\mathrm{CS}[T_{\mathrm{NSE}}]=\mathrm{CS}[T_{\mathrm{EC}}][\frac{4+5\alpha^2}{4+5\alpha^2+\alpha_\omega}]^{3/2}+\frac{2[a_0+a_1\alpha+a_2\alpha^2]}{[4+5\alpha^2+\alpha_\omega]^{3/2}}$;
(v) $\mathrm{CK}[T_{\mathrm{NSE}}]=\left[\mathrm{CK}[T_{\mathrm{EC}}]+\frac{(b_0+b_1\alpha+b_2\alpha^2+b_3\alpha^3)}{(4+5\alpha^2)^2}\right]\frac{[4+5\alpha^2]^2}{[4+5\alpha^2+\alpha_\omega]^2}$;

where $\alpha_\omega=2\alpha[\omega_3-\omega_1]-\omega_1^2$, $a_0=-6\omega_1+\omega_1^3+2\omega_3$, $a_1=3\omega_1^2-3\omega_1\omega_3$, $a_2=-6\omega_1-3\omega_3+2\omega_5$, $b_0=24\omega_1^2-3\omega_1^4-16\omega_1\omega_3$, $b_1=-[96\omega_1+12\omega_1^3+16\omega_3-12\omega_1^2\omega_3-16\omega_5]$, $b_2=18\omega_1^2+24\omega_1\omega_3-16\omega_1\omega_5$, *and* $b_3=8\omega_7-168\omega_1-1616\omega_5$, *with* $\omega_k=\mathrm{E}\left[X^k_{\lambda,h^{(2)}}\{(\alpha X_{\lambda,h^{(2)}})^2+4\}^{1/2}\right]$, *for* $k=1,3,5,7$, *and* $X_{\lambda,h^{(2)}}\sim \mathrm{NSE}(\lambda,h^{(2)})$. *In order to calculate values for ω_k, integrals involved must be solved using numerical methods.*

Proof. It can be easily obtained from (8). □

Remark 3.1. All the results in Theorem 3.2 are valid if we replace T_{NSE} by T_{LSE}. Of course, in this case, we have $\omega_k=\mathrm{E}[X^k_{\lambda,h^{(1)}}\{(\alpha X_{\lambda,h^{(1)}})^2+4\}^{1/2}]$, for $k=1,3,5,7$, with $X_{\lambda,h^{(1)}}\sim \mathrm{LSE}(\lambda,h^{(1)})$. Furthermore, from Theorem 3.2, as expected, when we set $\lambda=0$, we readily deduce the mean, variance, CV, CS, and CK of the ECBS distribution.

3.3. A BS–t distribution based on the NSE model

Now, we first present the LSEBS PDF with $t(\nu)$ kernel and then we derive the NSEBS PDF with $t(\nu)$ kernel, obtaining some useful results.

Let $T_{\text{LSE}} \sim \text{LSEBS}(\alpha, \beta, \lambda, h^{(1)})$, with $h^{(1)}$ being the univariate $t(\nu)$ kernel, which is denoted by $T_{\text{LS}-t} \sim \text{LSBS}\text{–}t(\alpha, \beta, \lambda, \nu)$. From (5), the PDF of $T_{\text{LS-}t}$ is given by

$$f_{T_{\text{LS-}t}}(t; \alpha, \beta, \lambda, \nu) = 2g(a(t; \alpha, \beta), \nu)G(\lambda a(t; \alpha, \beta), \nu)\, A(t; \alpha, \beta), \qquad t > 0.$$

Analogously, let $T_{\text{NSE}} \sim \text{NSEB}(\alpha, \beta, \lambda, h^{(2)})$, with $h^{(2)}$ being the bivariate $t(\nu)$ kernel, which is denoted by $T_{\text{NLS-}t} \sim \text{NSBS}\text{–}t(\alpha, \beta, \lambda, \nu)$. From (6) and (9), the PDF of $T_{\text{NLS-}t}$ is given by

$$f_{T_{\text{NLS-}t}}(t; \alpha, \beta, \lambda, \nu) = 2g(a(t; \alpha, \beta), \nu)G\Big(\lambda a(t; \alpha, \beta)\sqrt{\tfrac{\nu+1}{\nu+a^2(t;\alpha,\beta)}}, \nu+1\Big)A(t; \alpha, \beta),\ t > 0. \quad (10)$$

Kundu et al. (2008) and Azevedo et al. (2012) analyzed the BS and BS–t hazard rates. Let $T_{\text{NLS-}t} \sim \text{NSBS}\text{–}t(\alpha, \beta, \lambda, \nu)$, with its PDF as in (10). Then, the hazard rate of the RV $T_{\text{NLS-}t}$ is

$$h_{T_{\text{NLS-}t}}(t; \alpha, \beta, \lambda, \nu) = \frac{2g(a(t; \alpha, \beta), \nu)\, G\left(\lambda a(t; \alpha, \beta)\sqrt{\tfrac{\nu+1}{\nu+a^2(t;\alpha,\beta)}}, \nu+1\right)}{1 - F_{T_{\text{NLS-}t}}(t; \alpha, \beta, \lambda, \nu)} A(t; \alpha, \beta), t > 0.$$

Graphical plots of this hazard rate, for some values of the parameters α, β, λ, and ν, are shown in Fig. 1. From this figure, note that such a rate is unimodal for the values of ν considered. A deeper study for this rate will be explored by the authors in future works, but, based on the article by Azevedo et al. (2012), we suspect that some other shapes could be obtained.

To compute the mean and variance of the RV $T_{\text{NLS-}t}$, we need to compute $\text{E}[Z^2]$, $\text{E}[Z^4]$, and then

$$\text{E}\left[Z\left(\alpha^2 Z^2 + 4\right)^{1/2}\right] \quad \text{and} \quad \text{E}\left[Z^3\left(\alpha^2 Z^2 + 4\right)^{1/2}\right], \quad (11)$$

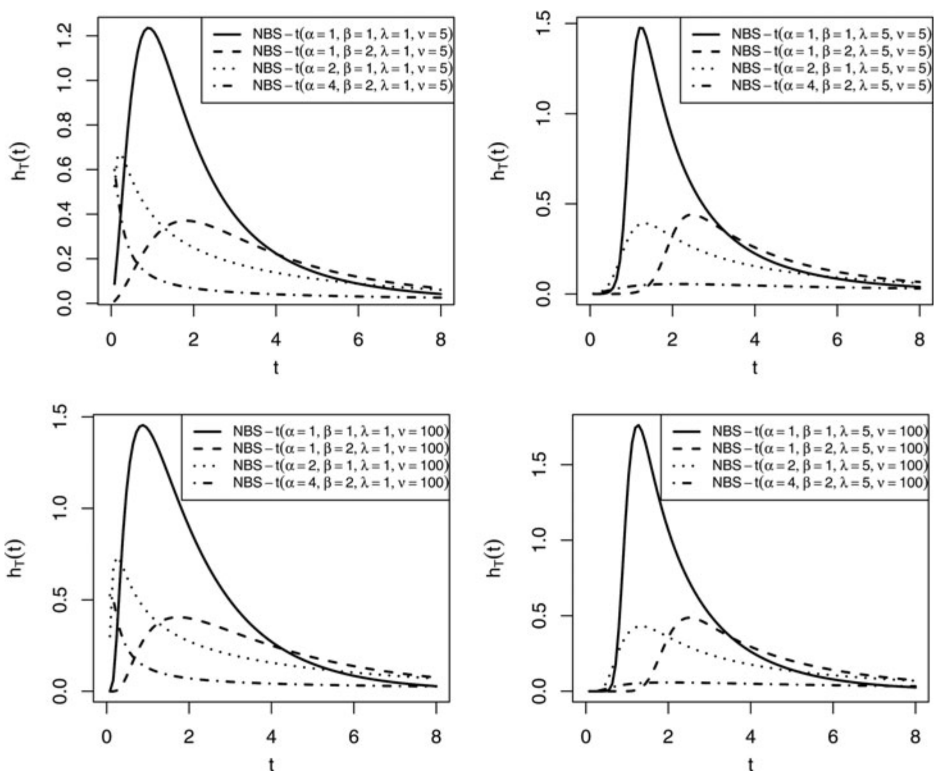

Figure 1. Plots of the NSBS–t hazard rate for the indicated values of the parameters α, β, λ, and ν.

where $Z = [1/\alpha][\sqrt{T_{\text{NLS-}t}/\beta} - \sqrt{\beta/T_{\text{NLS-}t}}]$. In order to obtain the values of the expectations given in (11), the integrals involved in their calculation must be solved using numerical methods. We employ the Monte Carlo method and the R software (www.r-project.org) for performing these computations. Specifically, we simulate a random sample of size 10^7 from the distribution of Z given in (11). Consequently, we can obtain $[1/\beta]\text{E}[T_{\text{NLS-}t}]$ and $[1/\beta^2]\text{V}[T_{\text{NLS-}t}]$ for some values of the parameters α and λ.

Theorem 3.3. *Let* $T_{\text{NLS-}t} \sim \text{NSBS–}t(\alpha, \beta, \lambda, \nu)$ *and* Z *as given in (11). Then,*

$$\frac{1}{\beta}\text{E}[T_{\text{NLS-}t}] = 1 + \frac{\alpha^2}{2}\left[\frac{\nu}{2(\nu-2)}\right] + \alpha\text{E}\left[Z(\alpha^2Z^2+4)^{1/2}\right] \quad \textit{and}$$

$$\frac{1}{\beta^2}\text{V}[T_{\text{NLS-}t}] = \alpha^2\left[\frac{\nu}{\nu-2}\right] - \frac{\alpha^2}{4}\left[\text{E}[Z\left(\alpha^2Z^2+4\right)^{1/2}]\right]^2 + \frac{\alpha^3}{2}\text{E}\left[Z^3(\alpha^2Z^2+4)^{1/2}\right]$$
$$-\frac{\alpha^3}{2}\left[\frac{\nu}{\nu-2}\right]\text{E}\left[Z(\alpha^2Z^2+4)^{1/2}\right] - \frac{\alpha^4}{4}\left[\frac{\nu}{\nu-2}\right]^2 + \frac{3}{2}\alpha^4\frac{\nu^2}{[\nu-2][\nu-4]}.$$

Proof. It can be easily obtained by using parts (i) and (ii) of Theorem 3.2. □

Values of $[1/\beta]\text{E}[T_{\text{NLS-}t}]$ and $[1/\beta^2]\text{V}[T_{\text{NLS-}t}]$, for several values of the parameters α and λ, and for $\nu = 5$, are given in Table 1.

Remark 3.2. When $\lambda = 0$, the values of $[1/\beta]\text{E}[T_{\text{NLS-}t}]$ and $[1/\beta^2]\text{V}[T_{\text{NLS-}t}]$, given in the first row of Table 1, coincide with the values of these quantities in the case of the BS–$t(\nu)$ distribution. In general, we see that these two quantities increase with α, which also occurs when λ decreases. It is worth to mention that the distribution of the RV Z given in (11) tends to the half–$t(\nu)$ distribution, when λ increases, and, in that case, we have a BS distribution with half–$t(\nu)$ kernel.

Table 1. Values of $[1/\beta]\text{E}[T_{\text{NS-}t}]$ and $[1/\beta^2]\text{V}[T_{\text{NS-}t}]$ with $\nu = 5$ and for the indicated values of λ and α.

$\lambda \setminus \alpha$	0.5	2	3.5	5	6.5	8	9.5
0	1.208	4.337	11.214	21.822	36.195	54.261	76.193
	1.145	190.162	1641.247	7253.996	21751.862	46729.694	95492.279
3	1.718	8.215	21.956	43.084	71.682	107.714	151.273
	1.679	339.225	3246.092	12970.776	38565.041	87383.219	182483.642
6	1.736	8.303	22.14	43.434	72.131	108.456	152.104
	1.655	348.739	3229.176	13545.049	38553.335	92044.642	170044.812
9	1.740	8.324	22.172	43.480	72.267	108.557	152.210
	1.658	342.347	3345.040	13393.849	38384.620	94634.190	191744.945
12	1.741	8.326	22.176	43.517	72.267	108.548	152.268
	1.667	341.340	3217.119	13686.092	37927.456	88541.029	169903.450
15	1.741	8.322	22.186	43.455	72.319	108.579	152.278
	1.669	359.303	3260.363	14340.800	38983.000	87670.763	175286.394
18	1.742	8.324	22.180	43.492	72.194	108.517	152.435
	1.634	356.849	3162.357	13697.336	37360.655	89327.588	182284.343
21	1.742	8.324	22.190	43.493	72.272	108.575	152.232
	1.641	339.941	3262.861	13959.965	39876.811	86820.889	175909.480

3.4. Recurrence relations for the NSBS–t distribution

Jamalizadeh et al. (2009a) obtained a recurrence relation for the CDF of the NLS–$t(\nu)$ distribution as follows:

$$F_{\text{NLS-}t}(t;\lambda,\nu+1)=F_{\text{NLS-}t}\left(\sqrt{\frac{\nu-1}{\nu+1}}t;\lambda,\nu-1\right)+\frac{\Gamma(\frac{\nu}{2})[\nu+1]^{\frac{\nu-1}{2}}}{\sqrt{\pi}\Gamma(\frac{\nu+1}{2})}\times\frac{t}{[\nu+1+t^2]^{\frac{\nu}{2}}}G\left(\frac{\sqrt{\nu}\lambda t}{\sqrt{\nu+1+t^2}},\nu\right),\quad t\in\mathbb{R},\nu>1. \tag{12}$$

From (7), we know that, in the case of the $t(\nu)$ kernel,

$$F_{T_{\text{NLS-}t}}(t;\alpha,\beta,\lambda,\nu)=F_{\text{NLS-}t}(a(t;\alpha,\beta),\lambda,\nu),\quad t>0. \tag{13}$$

Then, from (12) and (13), we obtain a recurrence relation for the NSBS–$t(\nu)$ distribution in the following theorem.

Theorem 3.4. *Let $T_{\text{NLS-}t}\sim$ NSBS–$t(\alpha,\beta,\lambda,\nu)$, with $\nu>1$. Then,*

$$F_{T_{\text{NLS-}t}}(t;\alpha,\beta,\lambda,\nu+1)=F_{T_{\text{NLS-}t}}\left(t;\sqrt{\frac{\nu+1}{\nu-1}}\alpha,\beta,\lambda,\nu-1\right)+\frac{\Gamma(\frac{\nu}{2})[\nu+1]^{\frac{\nu-1}{2}}}{\sqrt{\pi}\,\Gamma(\frac{\nu+1}{2})}\times\frac{a(t;\alpha,\beta)}{[\nu+1+a^2(t;\alpha,\beta)]^{\frac{\nu}{2}}}G\left(\frac{\sqrt{\nu}\lambda a(t;\alpha,\beta)}{\sqrt{\nu+1+a^2(t;\alpha,\beta)}},\nu\right),\quad t>0.$$

Proof. It follows by replacing $F_{\text{NLS-}t}(t;\lambda,\nu+1)$ given in (12) in (13) with $\nu+1$ DFs. In the special case when $\lambda=0$, expression in Theorem 3.4 provides a recurrence relation for the CDF of the RV T_t with BS–$t(\nu)$ distribution as

$$F_{T_t}(t;\alpha,\beta,\nu+1)=F_{T_t}\left(t;\sqrt{\frac{\nu+1}{\nu-1}}\alpha,\beta,\nu-1\right)+\frac{\Gamma(\frac{\nu}{2})[\nu+1]^{\frac{\nu-1}{2}}}{2\sqrt{\pi}\Gamma(\frac{\nu+1}{2})}\frac{a(t;\alpha,\beta)}{[\nu+1+a^2(t;\alpha,\beta)]^{\frac{\nu}{2}}},$$

for $t>0$ and $\nu>1$. Now, by using the known expressions of

$$F_{T_t}(t;\alpha,\beta,1)=\frac{1}{2}+\frac{1}{\pi}\tan^{-1}(a(t;\alpha,\beta))\ \text{and}$$
$$F_{T_t}(t;\alpha,\beta,2)=\frac{1}{2}\left[1+\frac{a(t;\alpha,\beta)}{\sqrt{2+a^2(t;\alpha,\beta)}}\right], \tag{14}$$

for $t>0$, the relation in (14) can be recursively used to produce expressions for other integer values of ν. For example, we obtain it for $t>0$ as

$$F_{T_t}(t;\alpha,\beta,3)=\frac{1}{2}+\frac{1}{\pi}\tan^{-1}\left(\frac{a(t;\alpha,\beta)}{\sqrt{3}}\right)\ \text{and}$$
$$F_{T_t}(t;\alpha,\beta,4)=\frac{1}{2}+\frac{a^3(t;\alpha,\beta)+6a(t;\alpha,\beta)}{2(4+a^2(t;\alpha,\beta))^{3/2}},\ t>0. \tag{15}$$

For the skew–Cauchy kernel (see Behboodian et al., 2006), an expression for the respective CDF is

$$F_{T_{\text{NLS-}t}}(t;\alpha,\beta,\lambda,1)=\frac{1}{\pi}\left[\tan^{-1}(a(t;\alpha,\beta))+\cos^{-1}\left(\frac{\lambda}{\sqrt{[1+\lambda^2][1+a^2(t;\alpha,\beta)]}}\right)\right],\ t>0. \tag{16}$$

For $\nu = 2$, we have

$$F_{T_{\text{NLS-}t}}(t;\alpha,\beta,\lambda,2) = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}(\lambda) + \frac{a(t;\alpha,\beta)}{\sqrt{2+a^2(t;\alpha,\beta)}} \times \left[\frac{1}{2} + \frac{1}{\pi}\tan^{-1}\left(\lambda a(t;\alpha,\beta)/\sqrt{2+a^2(t;\alpha,\beta)}\right)\right], \quad t>0. \quad (17)$$

Now, upon setting $\nu = 2, 3, \ldots$, in the recurrence relation in Theorem 3.4, and using the expressions given in (16) and (17), and the expressions of $F_{T_t}(t;\alpha,\beta,1)$, $F_{T_t}(t;2,\alpha,\beta)$, $F_{T_t}(t;\alpha,\beta,3)$ and $F_{T_t}(t;\alpha,\beta,4)$ presented from (14) to (15), we can derive explicit expressions for $F_{T_{\text{NLS-}t}}(t;\alpha,\beta,\lambda,3)$, $F_{T_{\text{NLS-}t}}(t;\alpha,\beta,\lambda,4)$, and so on, in a simple recursive manner. □

4. Estimation, inference and model selection

In this section, we propose ML estimators for the parameters of the NSBS–$t(\nu)$ distribution. An optimal procedure for selecting the parameter ν and some inferential aspects are also tried here.

4.1. *Estimation*

Consider a random sample of size n, say $\boldsymbol{T} = (T_1, \ldots, T_n)^\top$, where $T_i \sim$ NSBS–$t(\alpha,\beta,\lambda,\nu)$, for $i = 1, \ldots, n$. The log–likelihood function for $\boldsymbol{\theta} = (\alpha,\beta,\lambda)^\top$ based on observed data $\boldsymbol{t} = (t_1, \ldots, t_n)^\top$ is given by $\ell_n = \ell_n(\boldsymbol{\theta};\boldsymbol{t}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta};t_i)$, where

$$\ell_i(\boldsymbol{\theta};t_i) = \log(2) + \log(A(t_i;\alpha,\beta)) + \log\left(g\left(a\left(t_i;\alpha,\beta\right),\nu\right)\right) + \log\left(G\left(\lambda a(t_i;\alpha,\beta)\sqrt{\frac{\nu+1}{\nu+a^2(t;\alpha,\beta)}}, \nu+1\right)\right).$$

The ML estimate $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by solving the system of nonlinear equations

$$\boldsymbol{U}_n = \left(\frac{\partial \ell_n}{\partial \alpha}, \frac{\partial \ell_n}{\partial \beta}, \frac{\partial \ell_n}{\partial \lambda}\right)^\top = \boldsymbol{0}$$

with a numerical algorithm, such as the Newton–Raphson or Nelder–Mead methods; see details in next section. Initial guesses for the nonlinear iterative procedure, that allow the ML estimates to be computed, can be obtained in a similar way to that proposed in Vilca et al. (2011).

4.2. *Selecting the ν parameter*

We must decide whether to estimate or to fix the value of the parameter ν of the NSBS–$t(\nu)$ distribution. We follow a similar procedure proposed in Barros et al. (2009) and Paula et al. (2012). Thus, in order to estimate the structural parameters α, β, and λ, we fix integer values for ν within the interval [1, 100], selecting the value of it that maximizes the likelihood function, searching its optimal value by means of Algorithm 1.

Algorithm 1 Search of the optimal value for ν.

1. For ν from $\nu = 1$ to $\nu = 100$ by 1;
   1.1: Estimate the parameters α, β, and λ of the NSBS–$t(\nu)$ model by using the profile ML method at the corresponding value of ν;
   1.2: Compute the value of the likelihood function evaluating it at the ML estimates of α, β and λ obtained in Step 1.1 and at the corresponding value of ν used in the profile ML method; and
2. Choose the value of ν that maximizes the likelihood function and then consider the ML estimates of α, β and λ as result.

4.3. Inference and Hessian matrix

By the asymptotic normality of the ML estimators, we have $\widehat{\boldsymbol{\theta}} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda})^\top \dot{\sim} \mathrm{N}_3(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}})$, where $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$ is the asymptotic covariance matrix of $\widehat{\boldsymbol{\theta}}$, which can be approximated by $-\boldsymbol{H}_n^{-1}(\boldsymbol{\theta})$, with $-\boldsymbol{H}_n(\boldsymbol{\theta})$ being the observed information matrix obtained from the Hessian matrix $\boldsymbol{H}_n(\boldsymbol{\theta}) = [\partial^2\ell(\boldsymbol{\theta})/\partial\theta_1\partial\theta_2]$, where $\theta_1, \theta_2 = \alpha,\beta$, or λ. Then, $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} \approx -\boldsymbol{H}_n^{-1}(\widehat{\boldsymbol{\theta}})$, and

$$-\boldsymbol{H}_n^{1/2}(\widehat{\boldsymbol{\theta}})[\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}]^\top \xrightarrow{\mathcal{D}} \mathrm{N}_3(\boldsymbol{0}, \boldsymbol{I}_3), \text{ as } n \longrightarrow \infty,$$

with "$\xrightarrow{\mathcal{D}}$" denoting convergence in distribution to. Thus, based on this asymptotic result, we can construct approximate confidence regions and intervals, and hypothesis tests for α, β, and λ.

The elements of the Hessian matrix are given as follows. The second derivatives of $\ell_i(\boldsymbol{\theta}) = \ell_i(\boldsymbol{\theta}; t_i)$ are

$$\begin{aligned}
&\frac{\partial^2\ell_i(\boldsymbol{\theta})}{\partial\theta_1\partial\theta_2} \\
&= -\frac{1}{A^2(t_i;\alpha,\beta)}\frac{\partial A(t_i;\alpha,\beta)}{\partial\theta_1}\frac{\partial A(t_i;\alpha,\beta)}{\partial\theta_2} + \frac{1}{A(t_i;\alpha,\beta)}\frac{\partial^2 A(t_i;\alpha,\beta)}{\partial\theta_1\partial\theta_2} \\
&\quad - \frac{\nu+1}{\nu+a^2(t_i;\alpha,\beta)}\left[a\,(t_i;\alpha,\beta)\frac{\partial^2 a(t_i;\alpha,\beta)}{\partial\theta_1\partial\theta_2} + \frac{\nu-a^2(t_i;\alpha,\beta)}{\nu+a^2(t_i;\alpha,\beta)}\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_1}\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_2}\right] \\
&\quad + \lambda\nu\sqrt{\nu+1}\left\{\frac{W_G\left(\lambda a(t_i;\alpha,\beta)\sqrt{\frac{\nu+1}{\nu+a^2(t_i;\alpha,\beta)}}\right)}{\sqrt{(\nu+a^2(t_i;\alpha,\beta))^3}}\left[\frac{\partial^2 a(t_i;\alpha,\beta)}{\partial\theta_1\partial\theta_2} - \frac{3a\,(t_i;\alpha,\beta)}{\nu+a^2(t_i;\alpha,\beta)}\right.\right. \\
&\quad \left.\times\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_1}\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_2}\right] + \frac{\lambda\nu\sqrt{\nu+1}}{[\nu+a^2(t_i;\alpha,\beta)]^3} \\
&\quad \left.\times W_G{}'\left(\lambda a(t_i;\alpha,\beta)\sqrt{\frac{\nu+1}{\nu+a^2(t_i;\alpha,\beta)}}\right)\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_1}\frac{\partial a(t_i;\alpha,\beta)}{\partial\theta_2}\right\}, \theta_1, \theta_2 = \alpha \text{ or } \beta,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2\ell_i(\boldsymbol{\theta})}{\partial\alpha\partial\lambda} &= \frac{\partial a(t_i;\alpha,\beta)}{\partial\alpha}\left[W_G{}'\left(\lambda a(t_i;\alpha,\beta)\sqrt{\frac{\nu+1}{\nu+a^2(t_i;\alpha,\beta)}}\right)\frac{\lambda\nu[\nu+1]a(t_i;\alpha,\beta)}{(\nu+a^2(t_i;\alpha,\beta))^2}\right. \\
&\quad \left. + \frac{\nu\sqrt{\nu+1}\,W_G\left(\lambda a(t_i;\alpha,\beta)\sqrt{\frac{\nu+1}{\nu+a^2(t_i;\alpha,\beta)}}\right)}{\sqrt{(\nu+a^2(t_i;\alpha,\beta))^3}}\right],
\end{aligned}$$

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial\beta\partial\lambda} = \frac{\partial a(t_i;\alpha,\beta)}{\partial\beta}\left[W_G{}'\left(\lambda a(t_i;\alpha,\beta)\sqrt{\frac{\nu+1}{\nu+a^2(t_i;\alpha,\beta)}}\right)\frac{\lambda\nu(\nu+1)a(t_i;\alpha,\beta)}{(\nu+a^2(t_i;\alpha,\beta))^2} + \frac{\nu\sqrt{\nu+1}\,W_G\left(\lambda a(t_i;\alpha,\beta)\sqrt{\frac{\nu+1}{\nu+a^2(t_i;\alpha,\beta)}}\right)}{\sqrt{(\nu+a^2(t_i;\alpha,\beta))^3}}\right],$$

$$\frac{\partial^2 \ell_i(\boldsymbol{\theta})}{\partial\lambda^2} = W_G{}'\left(\lambda a(t_i;\alpha,\beta)\sqrt{\frac{\nu+1}{\nu+a^2(t_i;\alpha,\beta)}}\right)\frac{(\nu+1)a^2(t_i;\alpha,\beta)}{[\nu+a^2(t_i;\alpha,\beta)]},$$

where $W_G(x) = g(x;\nu+1)/G(x;\nu+1)$ and $W_G{}'(x) = -W_G(x)[(\nu+2)\,x/(\nu+1+x^2) + W_G(x)]$.

4.4. Model selection

Model selection criteria based on loss of information, such as Akaike (AIC), Schwarz's Bayesian (BIC) and Hannan-Quinn (HQIC) information criteria, can be used to compare models for the same data set. In order to detect the magnitude of the differences between the BIC values, we can use the Bayes factor (BF). In order to define the BF, we assume that the data D have arisen from one of two hypothetical models. Thus, H_1 (model with a smaller BIC value) is contrasted to H_2 (model compared to H_1), according to $P(D|H_1)$ or $P(D|H_2)$; see Vilca et al. (2011). An interpretation of the BF is displayed in Table 2.

In general, the BF is informative, because it presents ranges of values in which the degree of outperforming of one model with respect to another can be quantified. This factor can be obtained by using the approximation

$$2\log(B_{12}) \approx 2[\log(P(D \mid \widehat{\boldsymbol{\theta}}_1, H_1)) - \log(P(D \mid \widehat{\boldsymbol{\theta}}_2, H_2))] - [d_1 - d_2]\log(n), \tag{18}$$

where $P(D \mid \widehat{\boldsymbol{\theta}}_r, H_r) = P(D \mid H_r)$, with $\widehat{\boldsymbol{\theta}}_r$ being the ML estimate of $\boldsymbol{\theta}_r$ under the model in H_r, d_r is the dimension of $\boldsymbol{\theta}_r$, for $r = 1, 2$, and n is the sample size. Now, if the model H_1 is nested within the model H_2, then the approximation given in (18) can be written as

$$2\log(B_{12}) \approx \text{LR} - [d_1 - d_2]\log(n),$$

where LR is the standard likelihood ratio (LR) statistic to test H_1 against H_2, which is asymptotically distributed as a χ^2 distribution with $d_1 - d_2$ DFs, assuming H_1 is true, for $d_1 > d_2$.

5. Numerical applications

In this section, we carry out the numerical applications of this work by using simulated and real data. The performance of the ML estimators is evaluated by the Monte Carlo method. We

Table 2. Interpretation of the $\log(B_{12})$ related to the BF.

B_{12}	$2\log(B_{12})$	Evidence in favor of H_1
<1	<0	Negative (H_2 is accepted)
[1, 3]	[0, 2]	Weak
[3, 20]	[2, 6]	Positive
[20, 150]	[6, 10]	Strong
>150	>10	Very strong

explore the flexibility and potentiality of the NSEBS distribution by using two real-world data sets and the ML estimation method. This distribution is compared to other distributions by means of statistical tools and its adequacy to model these data is detected.

5.1. *Simulated data*

Here, we conduct some simulation experiments to evaluate the performance of the proposed ML estimators. We use a well–known stochastic representation for the NLS–$t(\nu)$ model given by

$$X \stackrel{\text{d}}{=} \delta\,|X_1| + \sqrt{1-\delta^2}X_2, \tag{19}$$

where $\delta = \lambda/\sqrt{1+\lambda^2}$ and $(X_1, X_2)^\top \sim t_2\,(\mathbf{0}, \boldsymbol{I}_2, \nu)$. Now, from equation

$$T_{\text{NLS-}t} \stackrel{\text{d}}{=} \frac{\beta}{4}\left[\alpha X + \sqrt{4+(\alpha X)^2}\right]^2$$

and from (19), we readily obtain the stochastic representation

$$T_{\text{NLS-}t} \stackrel{\text{d}}{=} \frac{\beta}{4}\left[\alpha(\delta\,|X_1| + \sqrt{1-\delta^2}X_2) + \sqrt{4+\{\alpha(\delta\,|X_1| + \sqrt{1-\delta^2}X_2)\}^2}\right]^2. \tag{20}$$

The data are simulated by means of the stochastic representation given in (20). We use Monte Carlo simulations to evaluate the finite–sample performance of the ML estimators of the NLS–$t(\nu)$ model parameters, whose estimates are determined by the Newton–Raphson procedure described in Sec. 4.1.

The scenario of the simulation is as follows: small, median, and large sample sizes are considered as $n \in \{10, 25, 100\}$, respectively; true values of the parameters considered are $\lambda \in \{-0.05, -2\}$ and $\nu \in \{2, 8, 50\}$, while the scale parameter is fixed at $\beta = 1$, without loss of generality. The number of Monte Carlo replications is taken as $B = 2{,}000$.

The samples are generated from the NLS–$t(\nu)$ model, called true distribution, and the estimation of parameters is computed from samples obtained by using the same or another distribution, called assumed distribution. The empirical relative bias (RB) in absolute value and the empirical mean square error (MSE) are the average from 2,000 simulated samples for each combination of n, α, and ν.

The results of the simulations are presented in Tables 3, 4, and 5, for the ML estimates of α, β, and λ, respectively. The estimation method for the parameters of the NLS–$t(\nu)$ distribution is implemented in the `R` software based on the direct maximization of the log–likelihood function for α, β and λ, by the linearly constrained optimization method (`constrOptim` function of the `R` software) of Nelder–Mead. From Tables 3, 4, and 5, note that the empirical RB and MSE of the estimators of α, β and λ decrease as the sample size n increases, as expected. Observe that, when the true and assumed models are the same, that is, ν is fixed, we get the expected results by analysing the empirical RB and MSE, in almost all of the considered values. Also, it is clear that, if the difference between ν of the true and assumed distributions increases, then the empirical MSE (for all of the estimators of α, β, and λ) increases as well.

5.2. *Real-world data*

Now, the flexibility and potentiality of the NSBS–$t(\nu)$ distribution are examined using two real-world data sets. We select these data to show the ability of our distribution in data modeling. For each data set, we obtain the ML estimates of the parameters of the NSBS–$t(\nu)$ distribution and fit it to the data. In addition, we compare the new model with BS and SN–BS

Table 3. Empirical RB and MSE of the estimator of α for the indicated values, distribution and λ, with $\beta = 1$.

			True distribution					
			NSBS–t(2)		NSBS–t(8)		NSBS–t(50)	
α	n	Assumed distribution	RB	MSE	RB	MSE	RB	MSE
		$\lambda = -0.05$						
0.2	10	NSBS–t(2)	0.0484	0.0068	0.1740	0.0039	0.1917	0.0040
		NSBS–t(8)	0.7995	0.0640	0.0657	0.0037	0.0698	0.0032
		NSBS–t(50)	1.1150	0.1246	0.2174	0.0057	0.0521	0.0028
	25	NSBS–t(2)	0.0631	0.0027	0.1566	0.0022	0.1694	0.0025
		NSBS–t(8)	0.7888	0.0361	0.1096	0.0020	0.0845	0.0018
		NSBS–t(50)	1.3547	0.1470	0.3236	0.0061	0.1235	0.0018
	100	NSBS–t(2)	0.0315	0.0007	0.1780	0.0016	0.2187	0.0023
		NSBS–t(8)	0.6790	0.0213	0.0874	0.0008	0.1406	0.0012
		NSBS–t(50)	1.2768	0.0885	0.3496	0.0058	0.0288	0.0007
0.5	10	NSBS–t(2)	0.0496	0.0436	0.1685	0.0239	0.1824	0.0272
		NSBS–t(8)	0.7158	0.2759	0.0649	0.0234	0.0895	0.0227
		NSBS–t(50)	0.9217	0.4512	0.0885	0.0322	0.0431	0.0161
	25	NSBS–t(2)	0.0750	0.0171	0.1385	0.0137	0.1470	0.0155
		NSBS–t(8)	0.7553	0.2178	0.1242	0.0134	0.1356	0.0160
		NSBS–t(50)	1.3629	0.6666	0.1609	0.0214	0.0152	0.0071
	100	NSBS–t(2)	0.0488	0.0047	0.1626	0.0093	0.1946	0.0132
		NSBS–t(8)	0.6890	0.1360	0.1317	0.0079	0.0943	0.0077
		NSBS–t(50)	1.7482	0.8864	0.2159	0.0173	0.0355	0.0022
1.0	10	NSBS–t(2)	0.0132	0.1048	0.1769	0.0987	0.1652	0.1013
		NSBS–t(8)	0.4170	0.3012	0.0171	0.0756	0.0281	0.0619
		NSBS–t(50)	0.6021	0.5457	0.1133	0.0969	0.0030	0.0629
	25	NSBS–t(2)	0.0721	0.0606	0.1365	0.0537	0.0755	0.0546
		NSBS–t(8)	0.5607	0.3784	0.0687	0.0473	0.0281	0.0283
		NSBS–t(50)	0.8135	0.7355	0.2390	0.1002	0.0664	0.0255
	100	NSBS–t(2)	0.0766	0.0266	0.1153	0.0304	0.0376	0.0326
		NSBS–t(8)	0.6569	0.4501	0.0902	0.0213	0.0363	0.0189
		NSBS–t(50)	0.9442	0.9092	0.2872	0.1096	0.0809	0.0175
		$\lambda = -2$						
0.2	10	NSBS–t(2)	0.0446	0.0068	0.3212	0.0062	0.3830	0.0074
		NSBS–t(8)	0.5357	0.0348	0.0989	0.0035	0.1928	0.0035
		NSBS–t(50)	1.0057	0.1140	0.0011	0.0055	0.1136	0.0030
	25	NSBS–t(2)	0.0174	0.0026	0.3018	0.0046	0.3679	0.0061
		NSBS–t(8)	0.5695	0.0222	0.0482	0.0016	0.1585	0.0021
		NSBS–t(50)	1.2035	0.1087	0.0965	0.0031	0.0500	0.0014
	100	NSBS–t(2)	0.0087	0.0008	0.3107	0.0042	0.3790	0.0059
		NSBS–t(8)	0.6259	0.0179	0.0092	0.0006	0.1526	0.0014
		NSBS–t(50)	1.2682	0.0755	0.1705	0.0021	0.0241	0.0005
0.5	10	NSBS–t(2)	0.0465	0.0425	0.3329	0.0410	0.3878	0.0468
		NSBS–t(8)	0.5089	0.1770	0.0727	0.0224	0.1919	0.0224
		NSBS–t(50)	0.7623	0.2832	0.0256	0.0329	0.1052	0.0186
	25	NSBS–t(2)	0.0147	0.0168	0.3012	0.0291	0.3697	0.0388
		NSBS–t(8)	0.5652	0.1311	0.0384	0.0107	0.1621	0.0139
		NSBS–t(50)	1.0035	0.3476	0.1005	0.0219	0.0511	0.0094
	100	NSBS–t(2)	0.0072	0.0056	0.3129	0.0265	0.3794	0.0373
		NSBS–t(8)	0.6275	0.1130	0.0102	0.0039	0.1553	0.0091
		NSBS–t(50)	1.2577	0.4414	0.1673	0.0132	0.0182	0.0036
1.0	10	NSBS–t(2)	0.0887	0.1178	0.3363	0.1654	0.3901	0.1953
		NSBS–t(8)	0.2379	0.1460	0.1007	0.0853	0.2058	0.1023
		NSBS–t(50)	0.2973	0.1674	0.0123	0.1432	0.1229	0.0776
	25	NSBS–t(2)	0.0475	0.0643	0.2992	0.1186	0.3671	0.1562
		NSBS–t(8)	0.3598	0.1711	0.0388	0.0471	0.1703	0.0630
		NSBS–t(50)	0.4462	0.2163	0.0879	0.1018	0.0630	0.0418
	100	NSBS–t(2)	0.0023	0.0269	0.3189	0.1121	0.3880	0.1574
		NSBS–t(8)	0.4648	0.2230	0.0163	0.0213	0.1641	0.0438
		NSBS–t(50)	0.4986	0.2489	0.1848	0.0653	0.0250	0.0192

Table 4. Empirical RB and MSE of the estimator of β for the indicated values, distribution and λ, with $\beta = 1$.

			True distribution					
			NSBS–t(2)		NSBS–t(8)		NSBS–t(50)	
α	n	Assumed distribution	RB	MSE	RB	MSE	RB	MSE
				$\lambda = -0.05$				
0.2	10	NSBS–t(2)	0.0038	0.0229	0.0013	0.0245	0.0043	0.0295
		NSBS–t(8)	0.0248	0.0516	0.0033	0.0202	0.0038	0.0294
		NSBS–t(50)	0.0292	0.0950	0.0050	0.0226	0.0003	0.0157
	25	NSBS–t(2)	0.0103	0.0130	0.0000	0.0157	0.0065	0.0230
		NSBS–t(8)	0.0184	0.0402	0.0021	0.0151	0.0035	0.0222
		NSBS–t(50)	0.0346	0.0998	0.0069	0.0234	0.0036	0.0159
	100	NSBS–t(2)	0.0036	0.0053	0.0042	0.0098	0.0052	0.0068
		NSBS–t(8)	0.0200	0.0259	0.0046	0.0102	0.0008	0.0126
		NSBS–t(50)	0.0042	0.0502	0.0076	0.0235	0.0010	0.0134
0.5	10	NSBS–t(2)	0.0138	0.1064	0.0340	0.1230	0.0813	0.2212
		NSBS–t(8)	0.0176	0.2330	0.0263	0.0980	0.0748	0.2029
		NSBS–t(50)	0.0483	0.3466	0.0133	0.0830	0.0042	0.0420
	25	NSBS–t(2)	0.0251	0.0866	0.0208	0.0997	0.0692	0.1838
		NSBS–t(8)	0.0802	0.2313	0.0378	0.0979	0.0562	0.1570
		NSBS–t(50)	0.1040	0.4807	0.0137	0.0675	0.0041	0.0308
	100	NSBS–t(2)	0.0221	0.0436	0.0008	0.0529	0.0105	0.0690
		NSBS–t(8)	0.0880	0.1717	0.0321	0.0833	0.0391	0.1162
		NSBS–t(50)	0.1319	0.6197	0.0056	0.0637	0.0034	0.0241
1.0	10	NSBS–t(2)	0.0224	0.2017	0.0468	0.2392	0.2600	0.8278
		NSBS–t(8)	0.0188	0.3300	0.0123	0.1508	0.0668	0.2411
		NSBS–t(50)	0.0959	0.5071	0.1007	0.2921	0.0047	0.1422
	25	NSBS–t(2)	0.0160	0.1984	0.0158	0.2343	0.2465	0.8320
		NSBS–t(8)	0.0521	0.3799	0.0009	0.1459	0.0521	0.1640
		NSBS–t(50)	0.1063	0.5277	0.0913	0.3212	0.0279	0.1407
	100	NSBS–t(2)	0.0136	0.1735	0.0163	0.2182	0.1834	0.7386
		NSBS–t(8)	0.0323	0.4088	0.0108	0.1512	0.0238	0.1488
		NSBS–t(50)	0.0577	0.4009	0.1163	0.3391	0.0227	0.1202
				$\lambda = -2$				
0.2	10	NSBS–t(2)	0.0062	0.0047	0.0253	0.0043	0.0303	0.0046
		NSBS–t(8)	0.0216	0.0064	0.0138	0.0032	0.0222	0.0036
		NSBS–t(50)	0.0380	0.0134	0.0179	0.0052	0.0164	0.0032
	25	NSBS–t(2)	0.0080	0.0020	0.0281	0.0026	0.0345	0.0030
		NSBS–t(8)	0.0316	0.0037	0.0085	0.0019	0.0190	0.0022
		NSBS–t(50)	0.0657	0.0108	0.0005	0.0029	0.0098	0.0017
	100	NSBS–t(2)	0.0021	0.0007	0.0337	0.0014	0.0421	0.0024
		NSBS–t(8)	0.0445	0.0026	0.0034	0.0007	0.0210	0.0012
		NSBS–t(50)	0.0837	0.0088	0.0155	0.0011	0.0055	0.0007
0.5	10	NSBS–t(2)	0.0133	0.0295	0.0600	0.0270	0.0708	0.0266
		NSBS–t(8)	0.0402	0.0345	0.0287	0.0208	0.0459	0.0214
		NSBS–t(50)	0.0208	0.0527	0.0466	0.0315	0.0312	0.0172
	25	NSBS–t(2)	0.0129	0.0129	0.0624	0.0157	0.0811	0.0169
		NSBS–t(8)	0.0808	0.0227	0.0141	0.0112	0.0436	0.0132
		NSBS–t(50)	0.0927	0.0371	0.0011	0.0181	0.0197	0.0099
	100	NSBS–t(2)	0.0043	0.0046	0.0809	0.0110	0.0993	0.0130
		NSBS–t(8)	0.1042	0.0155	0.0055	0.0045	0.0474	0.0076
		NSBS–t(50)	0.1849	0.0426	0.0371	0.0072	0.0093	0.0045
1.0	10	NSBS–t(2)	0.0291	0.0817	0.0960	0.0853	0.0995	0.0844
		NSBS–t(8)	0.0602	0.0874	0.0447	0.0680	0.0660	0.0729
		NSBS–t(50)	0.1792	0.1458	0.0377	0.1104	0.0521	0.0615
	25	NSBS–t(2)	0.0205	0.0470	0.0962	0.0574	0.1236	0.0582
		NSBS–t(8)	0.0031	0.0472	0.0274	0.0410	0.0701	0.0509
		NSBS–t(50)	0.1902	0.1285	0.0058	0.0726	0.0357	0.0409
	100	NSBS–t(2)	0.0044	0.0207	0.1419	0.0382	0.1808	0.0455
		NSBS–t(8)	0.0801	0.0228	0.0131	0.0209	0.0901	0.0300
		NSBS–t(50)	0.2470	0.1046	0.0853	0.0313	0.0207	0.0207

Table 5. Empirical RB and MSE of the estimator of λ for the indicated values, distribution and λ, with $\beta = 1$.

			True distribution					
			NSBS–t(2)		NSBS–t(8)		NSBS–t(50)	
α	n	Assumed distribution	RB	MSE	RB	MSE	RB	MSE
			$\lambda = -0.05$					
0.2	10	NSBS–t(2)	0.8585	0.5943	1.2243	1.0816	0.8335	1.9926
		NSBS–t(8)	0.0085	1.2669	1.5018	0.7366	1.6136	1.8942
		NSBS–t(50)	0.3040	1.5609	0.8654	1.2689	1.0761	0.7307
	25	NSBS–t(2)	0.5666	0.4287	1.1280	0.7622	0.0385	1.5702
		NSBS–t(8)	0.0565	1.1185	1.1854	0.6452	1.9230	1.4679
		NSBS–t(50)	0.7390	0.9095	0.4232	1.2690	0.7488	0.7067
	100	NSBS–t(2)	0.2873	0.1745	0.2365	0.2779	1.1269	0.3710
		NSBS–t(8)	0.8770	0.6853	0.9533	0.4499	0.6397	0.7150
		NSBS–t(50)	0.9577	0.3057	0.3244	1.1442	0.9944	0.6558
0.5	10	NSBS–t(2)	1.3368	0.4935	1.8557	1.0284	0.3943	2.0053
		NSBS–t(8)	2.8756	0.9187	1.6163	0.6990	0.8212	1.9782
		NSBS–t(50)	3.7410	0.4991	1.3442	0.2188	1.3398	0.1645
	25	NSBS–t(2)	0.7407	0.4659	1.7884	0.8523	0.8613	1.8706
		NSBS–t(8)	0.3089	0.8139	0.0669	0.4856	1.0556	1.7488
		NSBS–t(50)	3.9934	0.8162	1.0135	0.6050	0.9137	0.1402
	100	NSBS–t(2)	0.1532	0.2226	1.1952	0.4188	1.1893	0.6870
		NSBS–t(8)	0.3403	0.5846	0.3969	0.4129	0.8006	1.1187
		NSBS–t(50)	5.4609	1.4360	1.3584	0.4300	0.8044	0.1543
1.0	10	NSBS–t(2)	3.5602	0.3942	7.0228	0.9454	3.6433	2.0177
		NSBS–t(8)	3.5517	0.4530	3.0366	0.3489	1.8041	0.7097
		NSBS–t(50)	3.7958	0.4767	0.9245	0.3948	2.4191	0.3379
	25	NSBS–t(2)	2.6552	0.4212	5.4130	0.8237	3.5329	1.5879
		NSBS–t(8)	3.2934	0.4285	2.5123	0.2974	0.6097	0.3795
		NSBS–t(50)	3.9270	0.4809	1.4940	0.5201	1.2875	0.3561
	100	NSBS–t(2)	2.0638	0.3148	4.8965	0.6843	4.0200	1.8152
		NSBS–t(8)	4.4350	0.5244	1.9480	0.3141	0.6165	0.2874
		NSBS–t(50)	2.8759	0.3283	0.9700	0.5950	1.0841	0.2906
			$\lambda = -2$					
0.2	10	NSBS–t(2)	0.1167	0.7158	0.0046	0.8903	0.0446	0.9012
		NSBS–t(8)	0.2616	0.9149	0.0650	0.7373	0.0368	0.8881
		NSBS–t(50)	0.2832	0.9277	0.1847	0.9539	0.0369	0.7333
	25	NSBS–t(2)	0.0634	0.5965	0.1141	0.7632	0.2021	0.8238
		NSBS–t(8)	0.2843	0.8329	0.0596	0.6332	0.0497	0.7800
		NSBS–t(50)	0.3441	0.8737	0.2442	1.0807	0.0465	0.6419
	100	NSBS–t(2)	0.0376	0.3812	0.2893	0.6452	0.3774	0.7902
		NSBS–t(8)	0.3725	0.7525	0.0372	0.4149	0.1634	0.5558
		NSBS–t(50)	0.4098	0.8641	0.3048	0.9858	0.0152	0.4164
0.5	10	NSBS–t(2)	0.1075	0.7134	0.0048	0.8956	0.0543	0.8984
		NSBS–t(8)	0.2360	0.9126	0.0693	0.7377	0.0323	0.9175
		NSBS–t(50)	0.2279	0.9258	0.1337	0.9750	0.0575	0.7413
	25	NSBS–t(2)	0.0603	0.6097	0.1352	0.7843	0.1985	0.8133
		NSBS–t(8)	0.2816	0.8222	0.0615	0.6433	0.0386	0.7875
		NSBS–t(50)	0.3237	0.8776	0.2401	1.0031	0.0385	0.6562
	100	NSBS–t(2)	0.0460	0.4010	0.2859	0.6580	0.3748	0.7816
		NSBS–t(8)	0.3633	0.7523	0.0388	0.4220	0.1506	0.5507
		NSBS–t(50)	0.4165	0.8731	0.2886	0.9876	0.0308	0.4389
1.0	10	NSBS–t(2)	0.0659	0.7041	0.0066	0.8805	0.0364	0.9140
		NSBS–t(8)	0.1052	0.8819	0.0672	0.7377	0.0134	0.9042
		NSBS–t(50)	0.0567	0.9014	0.2014	0.9843	0.0443	0.7313
	25	NSBS–t(2)	0.0556	0.6269	0.0893	0.8079	0.1585	0.8429
		NSBS–t(8)	0.1210	0.7492	0.0353	0.6647	0.0468	0.8190
		NSBS–t(50)	0.0992	0.8270	0.2039	1.1708	0.0286	0.6770
	100	NSBS–t(2)	0.0446	0.4403	0.2651	0.6699	0.3702	0.8089
		NSBS–t(8)	0.1910	0.4958	0.0324	0.4843	0.1578	0.6469
		NSBS–t(50)	0.2957	0.7466	0.3212	1.0850	0.0147	0.5088

Table 6. Daily ozone level measurements (in ppb = ppm × 1000) in New York, May–September, 1973.

41	36	12	18	28	23	19	8	7	16	11	14	18	14	34	6	30
1	11	4	32	23	45	115	37	29	71	39	64	40	77	97	97	85
10	27	7	48	35	61	79	63	23	21	37	20	12	13	49	32	16
108	20	52	82	50	64	59	39	9	16	78	35	66	122	89	110	44
65	22	59	23	31	44	21	9	45	168	73	76	118	84	85	96	78
91	47	32	20	23	21	24	44	21	28	9	13	46	18	13	24	16
23	36	7	14	30	14	18	20	11	135	80	28	73	13			

distributions using graphical methods, maximized log–likelihood ($\ell(\widehat{\boldsymbol{\theta}})$ say), and AIC, BIC, and HQIC.

5.2.1. *Ozone data*

Data correspond to daily ozone concentrations in New York since May until September, 1973. These data were provided by the New York State Department of Conservation and are shown in Table 6; see Vilca et al. (2011) and Ferreira et al. (2012).

Vilca et al. (2011) showed that the SN–BS distribution gives a better fit than the usual BS distribution for these ozone data. Table 7 presents a descriptive summary of such data, including location and dispersion, as well as the standard deviation (SD), CV, CS, and CK. An exploratory data analysis based on this table and Fig. 2 indicates a positively skewed distribution with a kurtosis greater than three. Thus, we think the NSEBS model can be suitable for these data.

By using the proposed estimation method, we observe that, for ozone data, the maximized log–likelihood values increase first and then decrease, as ν increase. The maximized log–likelihood values, computed as a function of ν, through the discrete profile search, are presented in Table 8. These values are plotted in Fig. 3. We observe that the maximum occurs at $\nu = 5$, with the associated log–likelihood function value being -541.0650. This is a value that is often obtained in the literature, when data following a heavy–tailed distribution are analyzed. It is worth to mention here that the selection of the best $t(\nu)$ kernel, through the maximized log–likelihood value, is equivalent to selecting the best model by the AIC or BIC, because the number of model parameters remains the same when ν varies.

We fit the NSBS–$t(\nu)$ distribution with $\nu = 5$ to ozone data and compare it to the BS and SN–BS distributions. Estimation and model checking are provided in Table 9, which consists of the ML estimates, $\ell(\widehat{\boldsymbol{\theta}})$, AICs, BICs and HQICs. Considering these values, we find that the NSBS–$t(\nu)$ model with $\nu = 5$ provides a better fit than the BS and SN–BS models, which is corroborated by Fig. 2, where a satisfactory fit is observed for ozone data. For these data, we use the BF (approximated by the BIC) to contrast H_1: NSBS–$t(\nu = 5)$ model versus H_2: SN–BS model, which has a value of $2\log(B_{12}) = 6.326$, indicating a strong evidence in favor of H_1; see Table 2. In addition, using the LR test, we contrast the null hypothesis H_0: SN–BS model vs. H_1: NSBS–$t(\nu = 5)$ model or equivalently, H_0: $\nu = \infty$ vs. H_1: $\nu = 5$. The value of the LR test statistic and the corresponding p–value are 11.08 and 0.0009, respectively. This is consistent with the results obtained through the BF, AIC, BIC and HQIC. Hence, we conclude that the NSBS–$t(\nu = 5)$ distribution provides a much better fit than the SN–BS distribution for these ozone data.

Table 7. Descriptive statistics for ozone data.

Median	Mean	SD	CV	CS	CK	Range	Min.	Max.	*n*
31.5	42.13	32.99	78.30%	1.21	4.11	167	1	168	116

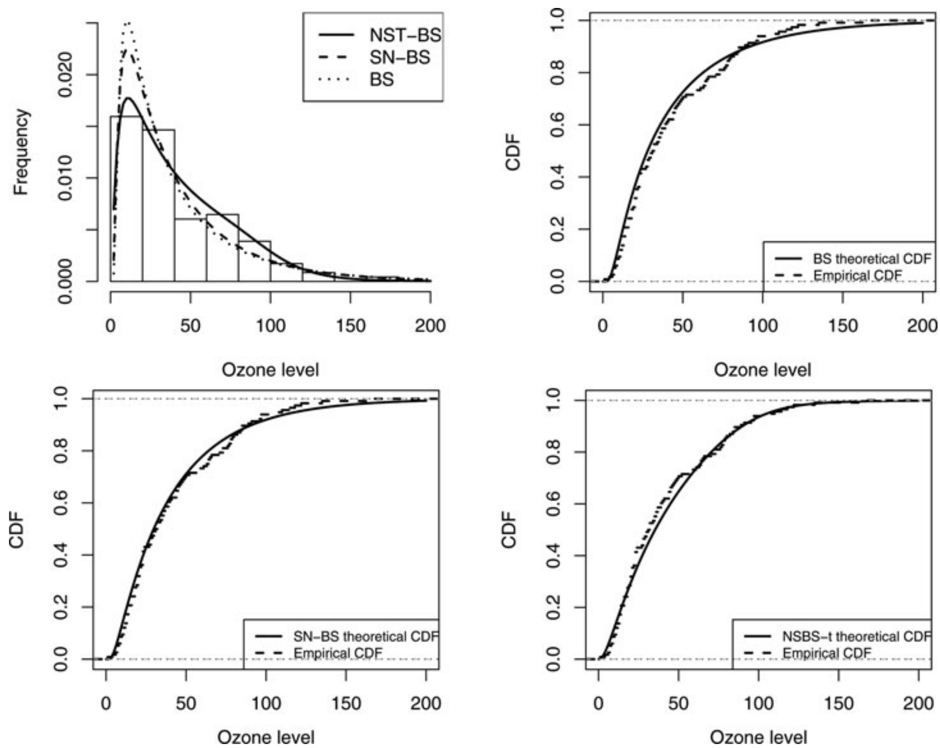

Figure 2. Histogram with the indicated estimated PDF (first panel left) and empirical CDF plot with indicated estimated theoretical CDFs (first panel right and second panel) for ozone data.

The estimated asymptotic covariance matrix of the ML estimators for the NSBS–$t(\nu)$ model parameters is

$$\widehat{\boldsymbol{\Sigma}}_{\widehat{\theta}} \approx -\boldsymbol{H}_n^{-1}(\widehat{\boldsymbol{\theta}}) = \begin{bmatrix} 0.0401 & 2.0561 & -0.5160 \\ 2.0561 & 174.5450 & -37.6750 \\ -0.5160 & -37.6750 & 12.3445 \end{bmatrix}.$$

Then, 95% confidence intervals for the parameters are given by $\alpha \in [1.597 \pm 0.392]$, $\beta \in [106.497 \pm 25.895]$ and $\lambda \in [-6.617 \pm 6.886]$, with $\nu = 5$.

5.2.2. *Fatigue data*

This data set was provided by Birnbaum and Saunders (1969) and corresponds to fatigue life of 6061–T6 aluminum coupons cut parallel to the direction of rolling and oscillated at 18 cycles per second. Data are displayed in Table 10, whereas Table 11 presents a descriptive summary of them.

Table 8. Maximized log–likelihood values versus $\nu = 1(1)100$ for ozone data.

ν	$\ell(\widehat{\boldsymbol{\theta}})$	ν	$\ell(\widehat{\boldsymbol{\theta}})$	ν	$\ell(\widehat{\boldsymbol{\theta}})$	ν	$\ell(\widehat{\boldsymbol{\theta}})$
1	−562.7759	2	−547.0028	3	−543.2078	4	−541.7778
5	−541.0650	6	−543.4296	7	−543.3661	8	−543.3229
9	−543.3278	10	−543.3609	11	−543.4108	12	−543.4694
13	−543.5329	14	−543.5977	15	−543.6625	16	−543.7260
17	−543.7877	18	−543.8471	19	−543.9041	20	−543.9584
21	−544.0102	22	−544.0595	23	−544.1063	24	−544.1508
25	−544.1933	25	−544.2337	26	−544.2717		

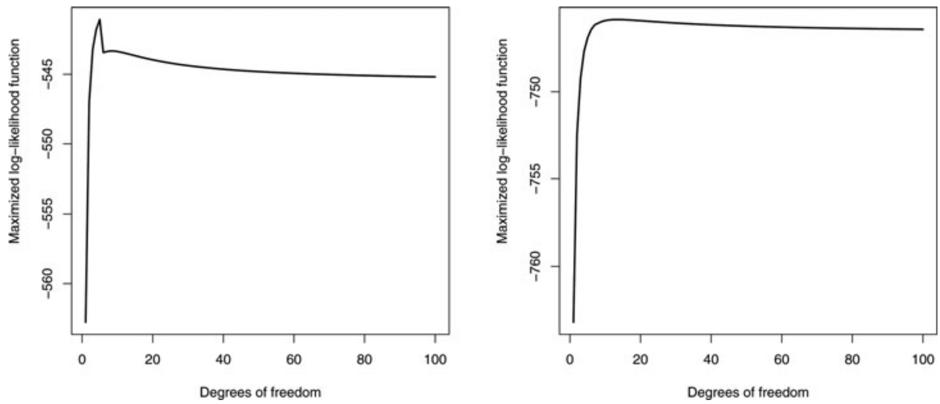

Figure 3. Plot of maximized log–likelihood function vs. ν for ozone (left) and fatigue (right) data.

We use the same procedure employed with ozone data for determining the value of ν. Table 12 and Fig. 3 present the optimum value of ν. We detect that the maximum value occurs at $\nu = 13$, with the associated log–likelihood function value being -745.8494.

We fit the NSBS–$t(\nu)$ distribution with $\nu = 13$ to fatigue data and compare it to the BS and SN–BS distributions. Estimation and model checking are provided in Table 13, which has the same configuration as Table 9. Regarding the values of the $\ell(\widehat{\boldsymbol{\theta}})$, AIC, BIC, and HQIC given in this table, we find that the NSBS–$t(\nu)$ model, now with $\nu = 13$, once again provides a better fit

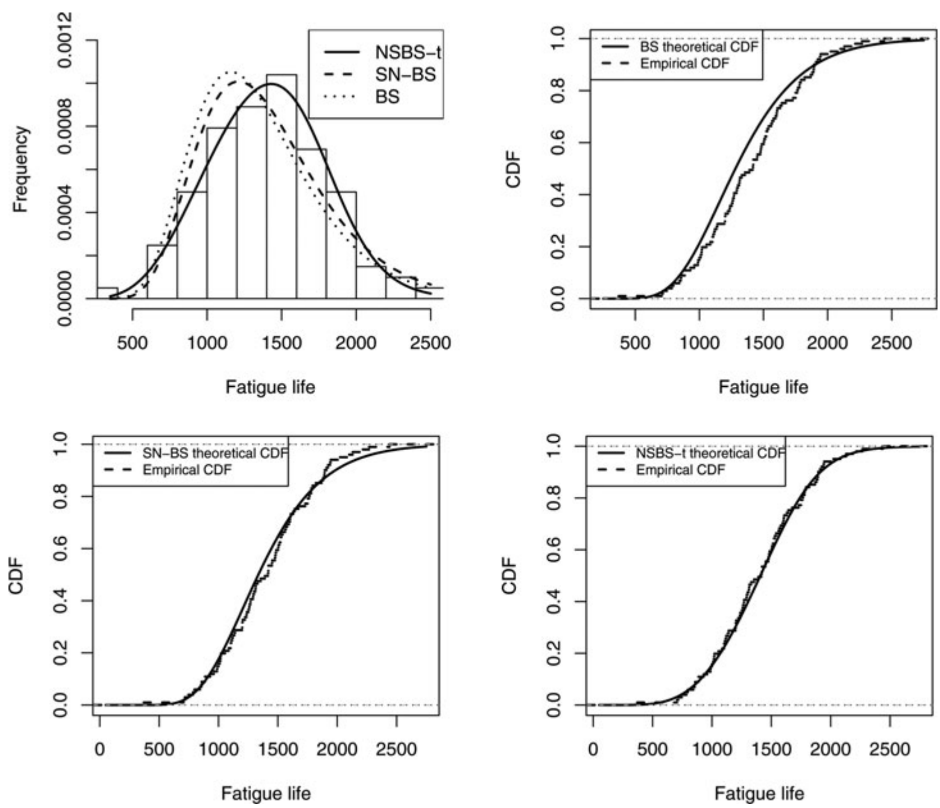

Figure 4. Histogram with the indicated estimated PDF (first panel left) and empirical CDF plot with indicated estimated theoretical CDFs (first panel right and second panel) for fatigue data.

Table 9. ML estimates and indicated criteria and distributions for ozone data.

Distribution	$\widehat{\boldsymbol{\theta}}$	$-\ell(\widehat{\boldsymbol{\theta}})$	AIC	BIC	HQIC
NSBS-$t(\nu=5)$	$(\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda}) = (1.597, 106.497, -6.617)$	541.065	1088.13	1096.391	1091.483
SN-BS	$(\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda}) = (1.270, 14.833, 1.067)$	545.605	1097.211	1105.471	1100.563
BS	$(\widehat{\alpha}, \widehat{\beta}) = (0.982, 28.023)$	549.097	1102.194	1107.701	1104.430

Table 10. Fatigue lifetimes (in cycles) of aluminum specimens exposed to a stress level of 21,000 psi.

370	706	716	746	758	797	844	855	858	886	886	930	960
988	990	1000	1010	1016	1018	1020	1055	1085	1102	1102	1108	1115
1120	1134	1140	1199	1200	1200	1203	1222	1235	1238	1252	1258	1226
1269	1270	1290	1293	1300	1310	1313	1315	1330	1355	1390	1416	1419
1420	1420	1450	1452	1475	1478	1481	1485	1502	1505	1513	1522	1522
1530	1540	1560	1567	1578	1594	1602	1604	1608	1630	1642	1674	1730
1750	1750	1763	1768	1781	1782	1792	1820	1868	1881	1890	1893	1895
1910	1923	1940	1945	2023	2100	2130	2215	2268	2440			

Table 11. Descriptive statistics for fatigue data.

Median	Mean	SD	CV	CS	CK	Range	Min.	Max.	*n*
1416	1401	391.32	0.28	0.14	2.71	2070	370	2440	101

Table 12. Maximized log–likelihood values versus $\nu = 1(1)100$ for fatigue data.

ν	$\ell(\widehat{\boldsymbol{\theta}})$	ν	$\ell(\widehat{\boldsymbol{\theta}})$	ν	$\ell(\widehat{\boldsymbol{\theta}})$	ν	$\ell(\widehat{\boldsymbol{\theta}})$
1	−763.2098	2	−752.4563	3	−749.2208	4	−747.6780
5	−746.8990	6	−746.4282	7	−746.1561	8	−746.0586
9	−745.9642	10	−745.9069	11	−745.8737	12	−745.8562
13	−745.8494	14	−745.8498	15	−745.8551	16	−745.8638
17	−745.8748	18	−745.8873	19	−745.9008	20	−745.9148
21	−745.9291	22	−745.9435	23	−745.9578	24	−745.9720
25	−745.9859	26	−745.9995	27	−746.0128		

Table 13. ML estimates and indicated criteria and distributions for fatigue data.

Distribution	$\widehat{\boldsymbol{\theta}}$	$-\ell(\widehat{\boldsymbol{\theta}})$	AIC	BIC	HQC
NSBS–$t(\nu=13)$	$(\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda}) = (0.423, 1879.996, -2.728)$	745.849	1497.699	1505.545	1500.876
SN–BS	$(\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda}) = (0.324, 1222.002, 0.369)$	751.225	1508.451	1516.296	1511.627
BS	$(\widehat{\alpha}, \widehat{\beta}) = (0.311, 1280.834)$	752.356	1508.712	1513.942	1510.829

than the SN–BS and BS models, which is corroborated in Fig. 4. Analogously to ozone data, we use the BF in order to highlight the differences between the values of the criteria presented in Table 13. In order to contrast H_1: NSBS–$t(\nu = 13)$ model vs. H_2: SN–BS model, we obtain a value of $2\log(B_{12}) = 6.137$, indicating a strong evidence in favor of H_1; see Table 2. Hence, we test H_0: SN–BS model vs. H_1: NSBS–$t(\nu = 13)$ model or, equivalently, H_0: $\nu = \infty$ vs. H_1: ν

= 13. The LR test statistic value is 10.752 and its corresponding p–value is 0.001. Therefore, H_0 should be rejected in favor of H_1 for any usual significance level.

In the case of fatigue data, the estimated asymptotic covariance matrix of the ML estimators for the NSBS–$t(\nu)$ model parameters is given by

$$\widehat{\boldsymbol{\Sigma}}_{\widehat{\boldsymbol{\theta}}} \approx -\boldsymbol{H}_n^{-1}(\widehat{\boldsymbol{\theta}}) = \begin{bmatrix} 0.0032 & 4.8125 & -0.0521 \\ 4.8125 & 11140.2170 & -104.8150 \\ -0.0521 & -104.8150 & 1.3461 \end{bmatrix}$$

and then 95% confidence intervals for the parameters are given by $\alpha \in [0.423 \pm 0.111]$, $\beta \in [1879.996 \pm 206.873]$, and $\lambda \in [-2.728 \pm 2.274]$, with $\nu = 13$.

6. Concluding remarks

In this article, we have introduced a methodology based on a class of Birnbaum–Saunders models. We have considered a special case of this class, which is called the nonlinear skew–Student–t Birnbaum–Saunders distribution. We have obtained a recurrence relationship for its cumulative distribution function, that is valid for any positive value of the degrees of freedom of the Student–t model. This skewed distribution extends the nonlinear skew–normal Birnbaum–Saunders distribution, allowing the extreme percentiles to be predicted very well and better than the skew–normal Birnbaum–Saunders model. We have provided estimators of the distribution parameters and evaluated their performance by the Monte Carlo simulation method. This simulation study showed the good performance of these estimators. The characteristics of the nonlinear skew–normal Birnbaum–Saunders–t model have been illustrated by means two real-world data sets, which have shown the potential applications of the proposed methodology. Some issues to be considered in futures studies about this new model are related to regression models, censored data and multivariate versions.

Acknowledgments

The authors thank the Editors and anonymous referees for their constructive comments on an earlier version of this article.

Funding

Mohsen Khosravi and Ahad Jamalizadeh thank the Mahani Mathematical Research Center (Kerman, Iran) for its support. Victor Leiva and Emilio Porcu were supported by FONDECYT 1120879 and 1130647 grants of the Comisión Nacional de Investigación Científica y Tecnológica from the Chilean government.

References

Ahmed, S. E., Castro–Kuriss, C., Flores, E., Leiva, V., Sanhueza, A. (2010). A truncated version of the Birnbaum–Saunders distribution with an application in financial risk. *Pak. J. Statist.* 26:293–311.

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scand. J. Statist.* 12:171–178.

Azzalini, A., Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew–t distribution. *J. Roy. Statist. Soc. B* 65:367–389.

Azzalini, A. (2005). The skew–normal distribution and related multivariate families. *Scand. J. Statist.* 32:159–188.

Azevedo, C., Leiva, V., Athayde, E., Balakrishnan, N. (2012). Shape and change point analyses of the Birnbaum–Saunders–t hazard rate and associated estimation. *Comp. Statist. Data Anal.* 56:3887–3897.

Balakrishnan, N., Leiva, V., Lopez, J. (2007). Acceptance sampling plans from truncated life tests from generalized Birnbaum–Saunders distribution. *Commun. Statist. Simul. Comp.* 36:643–656.

Balakrishnan, N., Leiva, V., Sanhueza, A., Cabrera, E. (2009a). Mixture inverse Gaussian distribution and its transformations, moments and applications. *Statistics* 43:91–104.

Balakrishnan, N., Leiva, V., Sanhueza, A., Vilca, F. (2009b). Estimation in the Birnbaum–Saunders distribution based on scale–mixture of normals and the EM algorithm. *Statist. Oper. Res. Trans.* 33:171–192.

Barros, M., Paula, G. A., Leiva, V. (2008). A new class of survival regression models with heavy–tailed errors: robustness and diagnostics. *Lifetime Data Anal.* 14:316–332.

Barros, M., Paula, G. A., Leiva, V. (2009). An R implementation for generalized Birnbaum–Saunders distributions. *Comp. Statist. Data Anal.* 53:1511–1528.

Behboodian, J., Jamalizadeh, A., Balakrishnan, N. (2006). A new class of skew–Cauchy distributions. *Statist. Prob. Lett.* 76:1488–1493.

Birnbaum, Z., Saunders, S. (1969). A new family of life distributions. *J. Appl. Probab.* 6:637–652.

Branco, M. D., Genton, M. G., Liseo, B. (2012). Objective Bayesian analysis of skew–*t* distributions. *Scand. J. Statist.* 40:63–85.

Díaz–García, J., Leiva, V. (2005). A new family of life distributions based on the contoured ellipticly distributions. *J. Statist. Plann. Infer.* 128:445–457.

Fang, K., Kotz, S., Ng, K. (1990). *Symmetric Multivariate and Related Distribution*. London: Chapman & Hall.

Ferreira, M., Gomes, M. I., Leiva, V. (2012). On an extreme value version of the Birnbaum–Saunders distribution. *Revstat Statist. J.* 10:181–210.

Genton, M. G. (2004). *Skew–Elliptical Distributions and Their Applications: A Journey Beyond Normality*. Boca Raton, FL: Chapman and Hall/CRC Press.

Gómez, H. W., Olivares, J., Bolfarine, H. (2009). An extension of the generalized Birnbaum–Saunders distribution. *Statist. Prob. Lett.* 79:331–338.

Guiraud, P., Leiva, V., Fierro, R. (2009). A non–central version of the Birnbaum–Saunders distribution for reliability analysis. *IEEE Trans. Rel.* 58:152–160.

Gupta, A. K., Varga, T. (1993). *Elliptically Contoured Models in Statistics*. Dordrecht: Kluwer.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1995). *Continuous Univariate Distributions*. New York: Wiley.

Jamalizadeh, A., Khosravi, M., Balakrishnan, N. (2009a). Recurrence relations for distributions of a skew–*t* and a linear combination of order statistics from a bivariate *t*. *Comp. Statist. Data Anal.* 53:847–852.

Jamalizadeh, A., Mahmoodian, H., Balakrishnan, N. (2009b). Exact distribution of a linear combination of a variable and order statistics from the other two variables of a trivariate elliptic random vector as a mixture of skew–elliptic distributions. *Statist. Meth.* 6:634–644.

Kundu, D., Kannan, N., Balakrishnan, N. (2008). On the hazard function of Birnbaum–Saunders distribution and associated inference. *Comp. Statist. Data Anal.* 52:2692–2702.

Leiva, V., Barros, M., Paula, G. A., Galea, M. (2007). Influence diagnostics in log–Birnbaum–Saunders regression models with censored data. *Comp. Statist. Data Anal.* 51:5694–5707.

Leiva, V., Barros, M., Paula, G. A., Sanhueza, A. (2008a). Generalized Birnbaum–Saunders distributions applied to air pollutant concentration. *Environmetrics* 19:235–249.

Leiva, V., Sanhueza, A., Angulo, J. M. (2009). A length–biased version of the Birnbaum–Saunders distribution with application in water quality. *Stoch. Environ. Res. Risk. Assess.* 23:299–307.

Leiva, V., Sanhueza, A. Kotz, S., Araneda, N. (2010a). A unified mixture model based on the inverse gaussian distribution. *Pak. J. Statist.* 26:445–460.

Leiva, V., Vilca, F., Balakrishnan, N., Sanhueza, A. (2010b). A skewed sinh–normal distribution and its properties and application to air pollution. *Commun. Statist. Theor. Meth.* 39:426–443.

Leiva, V., Athayde, E., Azevedo, C., Marchant, C. (2011). Modeling wind energy flux by a Birnbaum–Saunders distribution with unknown shift parameter. *J. Appl. Statist.* 38:2819–2838.

Leiva, V., Ponce, M. G., Marchant, C., Bustos, O. (2012). Fatigue statistical distributions useful for modeling diameter and mortality of trees. *Col. J. Statist.* 35:349–367.

Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., Barros, M. (2014). Birnbaum–Saunders statistical modelling: a new approach. *Statist. Mod.* 14:21–48.

Marchant, C., Leiva, V., Cavieres, M. F., Sanhueza, A. (2013a). Air contaminant statistical distributions with application to PM10 in Santiago, Chile. *Rev. Environ. Contam. Tox.* 223:1–31.

Marchant, C., Bertin, K., Leiva, V., Saulo, H. (2013b). Generalized Birnbaum-Saunders kernel density estimators and an analysis of financial data. *Comp. Statist. Data Anal.* 63:1–15.

Paula, G. A., Leiva, V., Barros, M., Liu, S. (2012). Robust statistical modeling using the Birnbaum–Saunders–*t* distribution applied to insurance. *Appl. Stoch. Mod. Bus. Ind.* 28:16–34.

Sanhueza, A., Leiva, V., Balakrishnan, N. (2008). The generalized Birnbaum–Saunders distribution and its theory, methodology and application. *Commun. Statist. Theor. Meth.* 37:645–670.

Santos–Neto, M., Cysneiros, F. J. A. Leiva, V., Ahmed, S. E. (2012). On new parameterizations of the Birnbaum–Saunders distribution. *Pak. J. Statist.* 28:1–26.

Vilca, F., Leiva, V. (2006). A new fatigue life model based on the family of skew–elliptical distributions. *Commum. Statist. Theor. Meth.* 35:229–244.

Vilca, F., Sanhueza, A., Leiva, V., Christakos, G. (2010). An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile. *Stoch. Environ. Res. Risk. Assess.* 24:771–782.

Vilca, F., Santana, L., Leiva, V., Balakrishnan, N. (2011). Estimation of extreme percentiles in Birnbaum–Saunders distributions. *Comp. Statist. Data Anal.* 55:1665–1678.