The Extended Birnbaum–Saunders Distribution Based on the Scale Shape Mixture of Skew Normal Distributions

Tahereh Poursadeghfard^{1,2} · Alireza Nematollahi¹ · Ahad Jamalizadeh³

Received: 23 April 2018 / Accepted: 23 September 2021 / Published online: 9 November 2021
© The Author(s) 2021

Abstract

In this article, a large class of univariate Birnbaum–Saunders distributions based on the scale shape mixture of skew normal distributions is introduced which generates suitable subclasses for modeling asymmetric data in a variety of settings. The moments and maximum likelihood estimation procedures are discussed via an ECM-algorithm. The observed information matrix to approximate the asymptotic covariance matrix of the parameter estimates is then derived in some subclasses. A simulation study is also performed to evaluate the finite sample properties of ML estimators and finally, a real data set is analyzed for illustrative purposes.

Keywords Scale shape mixtures of skew normal (SSMSN) · Scale shape mixtures of skew normal-Birnbaum Saunders (SSMSN-BS) · Skew t t-Birnbaum Saunders (ST_vT_v -BS) · Skew generalized Laplace normal-Birnbaum Saunders (SGLN-BS) · Skew slash normal-Birnbaum Saunders (SSN-BS)

Abbreviations

SSN-BS	Skew slash normal Birnbaum Saunders
BS	Birnbaum Saunders
SNN-BS	Skew slash normal Birnbaum Saunders
SN	Skew normal (SN)
SMSN	Scale mixture the skew normal (SMSN)
STN	Skew t normal
SSMSN	Scale shape mixture of skew normal
SSMSN-BS	Scale shape mixture of skew normal-Birnbaum Saunders

✉ Alireza Nematollahi
ar.nematollahi@shirazu.ac.ir

¹ Department of Statistics, College of Sciences, Shiraz University, Shiraz, Iran

² Department of Mathematics, Islamic Azad University, Firoozabad Branch, Firoozabad, Iran

³ Department of Statistics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran

$ST_v T_v$ -BS	Skew t t-Birnbaum Saunders
SGLN-BS	Skew generalized laplace normal-Birnbaum Saunders

1 Introduction

The two-parameter Birnbaum Saunders (*BS*) distribution, as a life type distribution, was originally introduced by Birnbaum and Saunders [8] in a failure model due to cracks, see Miner [31]. This distribution is related with normal distribution by means of a very simple functional relationship, and is based on a physical argument of cumulative damage that produces fatigue in the materials. The *BS* distribution, as a skew distribution, has been frequently applied in last few years; to biological model by Desmond [14], to medical field by Leiva et al. [26] and Barros et al. [5], to the forestry and environmental sciences by Podaski [32], Leiva et al. [28] and Vilca et al. [37], to fatigue life model by Cordeiro et al. [12], and to genetic model by Hasani et al. [20].

However, the *BS* model may suffer from a lack of robustness in the presence of extreme outlying observations. To allay this problem, several extensions of the *BS* distribution have been proposed in the literature, see e.g. Diaz-Garcia and Leiva [15], Sanhueza et al. [36], Leiva et al. [27], Gomez et al. [17], Vilca et al. [38], Hashemi et al. [25], Poursadeghfard et al. [33], Reyes et al. [34], and Mahbudi et al. [29], Benkhalifa [7],

The skew normal (*SN*) distribution proposed by Azzalini [2, 3] has been widely used in many applications to accommdate data with skewness. For more flexibilities, several extension of the *SN* model have been still considered. Branco and Dey [9] discussed the family of scale mixtures the skew-normal (*SMSN*) distributions. The random variable Y is said to have the univariate *SMSN* model when it has the following representation

$$Y \mid \tau \stackrel{d}{=} SN(\epsilon, \sigma^2 \tau^{-1}, \lambda),$$

where τ is a positive random variable with cumulative distribution function (cdf) $H(\tau; \nu)$. The cdf $H(\tau; \nu)$ is known as the mixing scale distribution. The univariate skew-t ([4]) and skew-slash ([39]) distributions are two special cases of *SMSN* distributions. Jammalizadeh et al. (2017) consider a more general class of scale-shape mixtures of skew normal (*SSMSN*) distributions and applied the EM-type algorithm for the computation of maximum likelihood estimates of parameters. This model includes *SMSN* model as a special case. Further, any model of the form $\frac{2}{\sigma} f\left(\frac{y-\epsilon}{\sigma}\right) G\left(\lambda \left(\frac{y-\epsilon}{\sigma}\right)\right)$, for which f and G are scale mixtures of the normal distributions, belongs to this class. For example, the skew t normal (*STN*) distribution discussed by Gomez et al. [18] and Cabral et al. [11] is a special case of *SSMSN* distributions.

In this paper, we introduce a new extension class of the *BS* distributions based on the *SSMSN* class, named the scale shape mixture of skew normal- Birnbaum Saunders (*SSMSN-BS*) class. It includes the *STN-BS* distributions in Poursadeghfard

et al. [33], the SNT-BS distributions in Hashemi et al. [25] and the SN-BS distributions in Vilca et al. (38) as the special cases. According to more flexibility of the class of SSMSN-BS distributions, they are attractive for modeling skewed and positive data sets in a much wider range, and we hope they have a better fit to compare the gamma, lognormal, weibull and exponential distributions as the lifetime distributions. However, a direct likelihood maximization for this class is difficult to compute due to the complexity of the likelihood function. To overcome this hurdle, we propose the ECM algorithm to estimate the parameters based on a convenient stochastic representation.

The rest of this paper is structured as follows. In Sect. 2, a useful stochastic representation of the SSMSN-BS class and its subclasses is presented, and basic properties of them are studied. In Sect. 3, the ECM-type algorithms for calculating ML estimates of parameters are provided. The information matrix for obtaining the asymptotic covariance matrix of the ML estimates are calculated in Sect. 4. Finally, a simulation study is presented in Sect. 5, where the proposed methodology through a real data set is also illustrated.

2 The Class of SSMSN-BS Distributions

In this section, we present the definition and some simple properties of the SSMSN-BS distributions. Some important asymmetric distributions generated from SSMSN-BS are also studied. Suppose Z_1 and Z_2 be two independent copies from $N(0, 1)$, and $\tau = (\tau_1, \tau_2)^T$ be a positive bivariate random variable, i.e. $P(\tau_1 > 0, \tau_2 > 0) = 1$, with probability density function (pdf) $h(\tau, \eta)$ where η is a vector of parameters. Further, we assume that both $\mathbf{Z} = (Z_1, Z_2)^T$ and $\tau = (\tau_1, \tau_2)^T$ are independent. A random variable Y is said to have the scale shape mixture of skew normal (SSMSN) distribution, denoted by $Y \sim \text{SSMSN}(\lambda, \eta)$, if

$$Y \mid (\tau_1, \tau_2) \stackrel{d}{=} SN\left(0, \tau_1^{-1}, \lambda \left(\frac{\tau_1}{\tau_2}\right)^{-\frac{1}{2}}\right). \quad (1)$$

By using the stochastic representation in (1) the random variable Y has following useful representation

$$Y \stackrel{d}{=} \left(\frac{\tau_1^{-\frac{1}{2}} f^{\frac{1}{2}}}{\sqrt{f + \lambda^2}} Z_2 + \frac{\lambda \tau_1^{-\frac{1}{2}}}{\sqrt{f + \lambda^2}} |Z_1| \right), \quad (2)$$

with $f = \frac{\tau_1}{\tau_2}$. Denoting by $\gamma = \sqrt{\tau_1^{-1}(f + \lambda^2)} |Z_1|$, a further hierarchical representation of the SSMSN distribution can be written as

$$\begin{aligned} Y \mid (\gamma, \tau_1, \tau_2) &\stackrel{d}{=} N\left(\frac{\lambda\gamma}{f + \lambda^2}, \frac{f}{\tau_1(f + \lambda^2)}\right), \\ \gamma \mid (\tau_1, \tau_2) &\stackrel{d}{=} TN\left(0, \frac{f + \lambda^2}{\tau_1}; (0, +\infty)\right), \end{aligned} \quad (3)$$

where $TN(\mu, \sigma^2; (a, b))$ represents the truncated normal distribution for $N(\mu, \sigma^2)$ lying within the truncated interval (a, b) , see Jammalizadeh and Lin (2017) for more details.

The cdf and pdf of a BS random variable, denoted by $T \sim BS(\alpha, \beta)$, are given by

$$\begin{aligned} F_{BS}(t; \alpha, \beta) &= \Phi(a(t; \alpha, \beta)), \\ f_{BS}(t; \alpha, \beta) &= \phi(a(t; \alpha, \beta))A(t; \alpha, \beta), t > 0, \end{aligned} \quad (4)$$

where

$$\begin{aligned} a(t; \alpha, \beta) &= \frac{1}{\alpha} \left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}} \right), \\ A(t; \alpha, \beta) &= \frac{da(t; \alpha, \beta)}{dt} = \frac{t + \beta}{2\alpha\sqrt{\beta}\sqrt{t^3}}, \end{aligned} \quad (5)$$

and $\alpha > 0$ and $\beta > 0$ are shape and scale parameters, and Φ and ϕ denote the cdf and pdf of the standard normal distribution, respectively.

The stochastic representation of T is

$$T = \frac{\beta}{4} \left[\alpha Z + \sqrt{(\alpha Z)^2 + 4} \right]^2, \quad Z \sim N(0, 1). \quad (6)$$

See Birnbaum and Saunders [8].

Definition 2.1 If $Z \sim SSMSN(\lambda, \eta)$ in (6), then T is said to have a BS distribution based on $SSMSN$ distribution with parameter $(\alpha, \beta, \lambda, \eta)$. It is denoted by $T \sim SSMSN-BS(\alpha, \beta, \lambda, \eta)$.

From (3), (5) and (6), the joint pdf T , γ and $\tau = (\tau_1, \tau_2)$ is given by

$$\begin{aligned} f(t, \gamma, \tau) &= \frac{1}{\pi} \exp\left(-\frac{1}{2}\tau_2\gamma^2 + \tau_2\lambda\gamma\tau a(t; \alpha, \beta) - \frac{1}{2}\tau_2\lambda^2a^2(t; \alpha, \beta) - \frac{1}{2}\tau_1a^2(t; \alpha, \beta)\right) \\ &\quad \times \sqrt{\tau_1\tau_2}h(\tau; \eta)A(t; \alpha, \beta). \end{aligned} \quad (7)$$

By integrating on γ in (7), we get

$$f(t, \tau) = 2\phi(\sqrt{\tau_1}a(t; \alpha, \beta))\Phi(\lambda\sqrt{\tau_2}a(t; \alpha, \beta))\sqrt{\tau_1}h(\tau; \eta). \quad (8)$$

The marginal density of T is then given by

$$f(t) = \int_0^\infty \int_0^\infty f(t, \tau_1, \tau_2) d\tau_1 d\tau_2. \quad (9)$$

Dividing (7) by (8) gives

$$f(\gamma | t, \tau_1, \tau_2) = \frac{\phi\left(\sqrt{\tau_2}(\gamma - \lambda a(t; \alpha, \beta))\right)}{\sqrt{\tau_2} \Phi\left(\lambda \sqrt{\tau_2} a(t; \alpha, \beta)\right)} = f(\gamma | t, \tau_2), \quad (10)$$

implying that γ and τ_1 are conditionally independent given $T = t$ and τ_2 . From (10), it is clear that

$$\gamma | (t, \tau_2) \sim TN\left(\lambda a(t; \alpha, \beta), \frac{1}{\tau_2}, (0, \infty)\right). \quad (11)$$

The conditional pdf of $\tau = (\tau_1, \tau_2)$ given $T = t$ is also given by

$$f(\tau_1, \tau_2 | t) = \frac{2\phi\left(\sqrt{\tau_1} a(t; \alpha, \beta)\right) \Phi\left(\lambda \sqrt{\tau_2} a(t; \alpha, \beta)\right) \sqrt{\tau_1} h(\tau; \eta)}{f(t)}. \quad (12)$$

Proposition 2.1 *Let $T \sim \text{SSMSN-BS}$, then the mean, variance, coefficient of variation, skewness and kurtosis of T are given by*

$$\begin{aligned} E[T] &= E[T_1] + \frac{\alpha\beta}{2}\omega_1, \\ V[T] &= V[T_1] + \left[\frac{\alpha\beta}{2}\right]^2 \alpha_\omega, \\ \gamma[T] &= \gamma[T_1] \frac{\sqrt{1 + \frac{\alpha^2\beta^2\alpha_\omega}{V(2T_1)}}}{1 + \frac{\alpha\beta\omega_1}{E[2T_1]}}, \\ \alpha_3[T] &= \alpha_3[T_1] \left[\frac{\alpha^2(2ET^{*4} - E^2T^{*2}) + 4ET^{*2}}{\alpha^2(2ET^{*4} - E^2T^{*2}) + 4ET^{*2} + \alpha\omega} \right]^{\frac{3}{2}} \\ &\quad + \frac{2[a_0 + a_1\alpha + a_2\alpha^2]}{[\alpha^2(2ET^{*4} - E^2T^{*2}) + 4ET^{*2} + \alpha\omega]^{\frac{3}{2}}}, \\ \alpha_4[T] &= \left[\alpha_4[T_1] + \frac{b_0 + b_1\alpha + b_2\alpha^2 + b_3\alpha^3}{(\alpha^2(2ET^{*4} - E^2T^{*2}) + 4ET^{*2})^2} \right] \\ &\quad \frac{[\alpha^2(2ET^{*4} - E^2T^{*2}) + 4ET^{*2}]^2}{[\alpha^2(2ET^{*4} - E^2T^{*2}) + 4ET^{*2} + \alpha\omega]^2}, \end{aligned} \quad (13)$$

respectively, where

$$\begin{aligned}
E[T_1] &= \frac{\beta}{2} \left[\alpha^2 ET^{*2} + 2 \right], \\
V[T_1] &= \frac{\beta^2 \alpha^2}{4} \left[\alpha^2 (2ET^{*4} - E^2 T^{*2}) + 4ET^{*2} \right], \\
\gamma[T_1] &= \frac{\alpha \sqrt{\alpha^2 (2ET^{*4} - E^2 T^{*2}) + 4ET^{*2}}}{\left[\alpha^2 ET^{*2} + 2 \right]}, \\
\alpha_3[T_1] &= \frac{1}{[V[T_1]]^{\frac{3}{2}}} \frac{\beta^3 \alpha^4}{8} \left[\alpha^2 (4ET^{*6} - 6ET^{*4} ET^{*2} + 2E^3 T^{*2}) + 12ET^{*4} - 12E^2 T^{*2} \right], \\
\alpha_4[T_1] &= \frac{1}{[V[T_1]]^2} \frac{\beta^4 \alpha^4}{8} \left[\alpha^4 (8ET^{*8} - 16ET^{*2} ET^{*6} + 12ET^{*4} E^2 T^{*2} - 3E^4 T^{*2}) \right. \\
&\quad \left. + \alpha^2 (32ET^{*6} - 48ET^{*4} ET^{*2} + 24E^3 T^{*3}) + 16ET^{*4} \right],
\end{aligned} \tag{14}$$

with $T^*(v_1) = \tau_1^{-\frac{1}{2}}(v_1)Z(0, 1)$, $ET^{*2} = E\tau_1^{-1}$, $ET^{*4} = 3E\tau_1^{-2}$, $ET^{*6} = 15E\tau_1^{-3}$, $ET^{*8} = 105E\tau_1^{-4}$, and

$$\begin{aligned}
\alpha_\omega &= 2\alpha(\omega_3 - \omega_1 ET^{*2}) - \omega_1^2, \\
a_0 &= \omega_1^3 - 6\omega_1 ET^{*2} + 2\omega_3, \\
a_1 &= 3\omega_1^2 ET^{*2} - 3\omega_1 \omega_3, \\
a_2 &= 2\omega_5 - 3\omega_3 ET^{*2} - 3\omega_1 ET^{*4} + 3\omega_1 E^2 T^{*2}, \\
b_0 &= -3\omega_1^4 - 16\omega_1 \omega_3 + 24\omega_1^2 ET^{*2}, \\
b_1 &= 16\omega_5 - 16\omega_3 ET^{*2} - 48\omega_1 ET^{*4} + 48\omega_1 E^2 T^{*2} + 12\omega_1^2 \omega_3 - 12\omega_1^3 ET^{*2}, \\
b_2 &= -16\omega_1 \omega_5 + 12\omega_1^2 ET^{*4} - 18\omega_1^2 E^2 T^{*2} + 24\omega_1 \omega_3 ET^{*2}, \\
b_3 &= 8\omega_7 - 16\omega_1 ET^{*6} - 16\omega_5 ET^{*2} + 12\omega_3 E^2 T^{*2} + 24\omega_1 ET^{*2} ET^{*4} - 12\omega_1 E^3 T^{*2}, \\
\omega_k &= E[Y^k \sqrt{\alpha^2 Y^2 + 4}], k = 1, 3, 5, 7, \quad Y \sim SSMSN(\lambda, \eta).
\end{aligned} \tag{15}$$

In the following subsections, some asymmetric distributions generated by SSMSN – BS are studied.

2.1 The Skew $t_1 t_2$ -BS Distribution

Let τ_1 and τ_2 be two independent gamma random variables with shape and rate parameters equal to $\frac{v_i}{2}$, namely $\Gamma\left(\frac{v_i}{2}, \frac{v_i}{2}\right)$ for $i = 1, 2$, with the following joint density

$$h(\tau; \eta) = f\left(\tau_1; \frac{v_1}{2}, \frac{v_1}{2}\right) f\left(\tau_2; \frac{v_2}{2}, \frac{v_2}{2}\right), \tag{16}$$

where $\eta = (v_1, v_2)$ and f denotes the pdf of the gamma distribution, then Y in (2) and (3), is said to have the skew $t_1 t_2$ distribution with parameter (λ, η) , and will be denoted by $Y \sim ST_{v_1 T_{v_2}}(\lambda, v_1, v_2)$. Now we define the skew

$t_1 t_2$ -BS distribution as follows:

Definition 2.2 If $Y \stackrel{d}{=} Z \sim ST_{v_1} T_{v_2}(\lambda, v_1, v_2)$ in (6), we say that the random variable T follows the $ST_{v_1} T_{v_2}$ -BS distribution with the following pdf

$$f_{ST_{v_1} T_{v_2}-BS}(t; \alpha, \beta, \lambda, v_1, v_2) = 2t(a(t; \alpha, \beta); v_1) T(\lambda a(t; \alpha, \beta); v_2) \times A(t; \alpha, \beta), \quad t, \alpha, \beta, v_1, v_2 > 0, \lambda \in R. \quad (17)$$

where $t(\cdot; v_1)$, $T(\cdot; v_2)$ denote the pdf and cdf of the Student-t distribution with degree of freedom v_1, v_2 respectively.

We shall write $T \sim ST_{v_1} T_{v_2}$ -BS($\alpha, \beta, \lambda, v_1, v_2$) to denote that T follows the $ST_{v_1} T_{v_2}$ -BS distribution with pdf in (17). The mean, variance and measures of skewness and kurtosis coefficients of T follows from the equations in (13)–(15), for which, we have

$$T^*(v_1) \stackrel{d}{=} \Gamma^{-\frac{1}{2}}\left(\frac{v_1}{2}, \frac{v_1}{2}\right) Z(0, 1), \quad (18)$$

with

$$\begin{aligned} ET^{*2} &= \frac{v_1}{v_1 - 1}, \quad v_1 > 2, \\ ET^{*4} &= \frac{3v_1^2}{(v_1 - 1)(v_1 - 4)}, \quad v_1 > 4, \\ ET^{*6} &= \frac{15v_1^3}{(v_1 - 1)(v_1 - 4)(v_1 - 6)}, \quad v_1 > 6, \\ ET^{*8} &= \frac{105v_1^4}{(v_1 - 1)(v_1 - 4)(v_1 - 6)(v_1 - 8)}, \quad v_1 > 8, \\ Y &\sim ST_{v_1} T_{v_2}(\lambda, v_1, v_2). \end{aligned}$$

From (12), the conditional pdf of τ_1 given $T = t$ is

$$(\tau_1 | T = t) \sim \Gamma\left(\frac{v_1 + 1}{2}, \frac{v_1 + a^2(t; \alpha, \beta)}{2}\right), \quad (19)$$

and the conditional pdf of τ_2 given $T = t$ is

$$f(\tau_2 | T = t) = \frac{\left(\frac{v_2}{2}\right)^{\frac{v_2}{2}} \tau_2^{\frac{v_2}{2}-1} \exp\left(-\frac{v_2}{2}\tau_2\right) \Phi\left(\lambda\sqrt{\tau_2}a(t; \alpha, \beta)\right)}{T(\lambda a(t; \alpha, \beta); v_2)}. \quad (20)$$

So, we have

$$\begin{aligned}
E(\tau_1 | T = t) &= E(\tau_1 | T = t) = \frac{v_1 + 1}{v_1 + a^2(t; \alpha, \beta)}, \\
E(\log \tau_1 | T = t) &= DG\left(\frac{v_1 + 1}{2}\right) - \log\left(\frac{v_1 + a^2(t; \alpha, \beta)}{2}\right), \\
E(\tau_2 | T = t) &= \frac{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v_2 + 2}{v_2}}; v_2 + 2\right)}{T(\lambda a(t; \alpha, \beta); v_2)}, \\
E(\log \tau_2 | T = t) &= \frac{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v_2 + 2}{v_2}}; v_2 + 2\right)}{T(\lambda a(t; \alpha, \beta); v_2)} - \log\left(\frac{v_2}{2}\right) - 1 + DG\left(\frac{v_2 + 1}{2}\right) \\
&\quad + \frac{1}{T(\lambda a(t; \alpha, \beta); v_2)} \int_{-\infty}^{\lambda a(t; \alpha, \beta)} \frac{x^2 - 1}{(v_2 + x)} t(x; v_2) dx. \\
E(\gamma \tau_2 | T = t) &= \frac{1}{T(\lambda a(t; \alpha, \beta); v_2)} \lambda a(t; \alpha, \beta) T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v_2 + 2}{v_2}}; v_2 + 2\right) \\
&\quad + \frac{1}{T(\lambda a(t; \alpha, \beta); v_2)} \frac{\Gamma\left(\frac{v_2 + 1}{2}\right) \left(\frac{v_2}{2}\right)^{\frac{v_2}{2}}}{\Gamma\left(\frac{v_2}{2}\right) (\lambda^2 a^2(t; \alpha, \beta) + v_2)^{\left(\frac{v_2 + 1}{2}\right)}},
\end{aligned} \tag{21}$$

where $DG(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ is the digamma function.

Remark: If $v_1 = v_2 = v$, then $T \sim ST_v T_v \sim BS(\alpha, \beta, \lambda, v)$.

It should be noted that SSMSN-BS class includes some special distributions. In the following, some special cases are given.

(I) The skew t normal-BS (STN-BS) distribution formulated by taking $\tau_1 \sim \Gamma\left(\frac{v}{2}, \frac{v}{2}\right)$ and $\tau_2 = 1$ with probability 1. The pdf is

$$\begin{aligned}
f_{STN-BS}(t; \alpha, \beta, \lambda, v) &= 2t(a(t; \alpha, \beta); v) \Phi(\lambda a(t; \alpha, \beta)) A(t; \alpha, \beta) \\
t, \alpha, \beta, v &> 0, \lambda \in R.
\end{aligned} \tag{22}$$

See Poursadeghfar et al. (2016).

(II) The skew normal t-BS (SNT-BS) distribution formulated by $\tau_1 = 1$ with probability 1 and $\tau_2 \sim \Gamma\left(\frac{v}{2}, \frac{v}{2}\right)$. The pdf is

$$\begin{aligned}
f_{SNT-BS}(t; \alpha, \beta, \lambda, v) &= 2\phi(a(t; \alpha, \beta)) T(\lambda a(t; \alpha, \beta); v) A(t; \alpha, \beta) \\
t, \alpha, \beta, v &> 0, \lambda \in R.
\end{aligned} \tag{23}$$

See Hashemi et al. [25].

(III) The skew normal-BS (SN-BS) distribution obtained by taking $\tau_1 = \tau_2 = 1$ with probability 1. The pdf is

$$f_{SN-BS}(t; \alpha, \beta, \lambda) = 2\phi(a(t; \alpha, \beta)) A(t; \alpha, \beta) \quad t, \alpha, \beta > 0, \lambda \in R. \tag{24}$$

See Vilca et al. [38].

Fig. 1 The graph of densities, BS, SN-BS, STN-BS, SNT-BS, and $ST_v T_v$ -BS, for the selected of values of parameters

Figure 1 displays the graph of the densities, BS, SN-BS, STN-BS, SNT-BS, $ST_v T_v$ -BS with $\alpha = 0.5$, $\beta = 0.8$, $\lambda = 3$ and four different degrees of freedom $v = 1, 3, 5$ and 15 .

2.2 The Skew t-BS Distribution

Let $\tau_1 = \tau_2 = \tau$ with probability 1 be a one gamma random variable with shape and rate parameters equal to $\frac{v}{2}$, with the following pdf

$$h(\tau; v) = f\left(\tau; \frac{v}{2}, \frac{v}{2}\right). \quad (25)$$

Then Y in (2) is said to have the skew t distribution with parameter (λ, v) , and will be denoted by $Y \sim ST(\lambda, v)$.

Definition 2.3 If $Y \stackrel{d}{=} Z \sim ST(\lambda, v)$ in (6) we say that the random variable T follows the ST-BS distribution with pdf

$$f_{ST-BS}(t; \alpha, \beta, \lambda, v) = 2t(a(t; \alpha, \beta); v) T(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+1}{a^2(t; \alpha, \beta) + v}}, v+1) \times A(t; \alpha, \beta) \quad t, \alpha, \beta, v > 0, \lambda \in R. \quad (26)$$

we shall write $T \sim ST-BS(\alpha, \beta, \lambda, v)$ to denote that T follows the $ST-BS$ distribution.

The mean, variance and measures of skewness and kurtosis coefficients of T follows equations in (13)–(15), (18). In addition $Y \sim ST(\lambda, v)$.

From (12), the conditional pdf of τ given $T = t$ is

$$f(\tau | T = t) = \left(\frac{a^2(t; \alpha, \beta) + v}{2} \right)^{\left(\frac{v+1}{2} \right)} \frac{1}{\Gamma\left(\frac{v+1}{2}\right) T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+1}{a^2(t; \alpha, \beta) + v}}, v+1\right)} \times \exp\left(-\frac{\tau}{2}(a^2(t; \alpha, \beta) + v)\right) \Phi\left(\sqrt{\tau} \lambda a(t; \alpha, \beta)\right) \tau^{\frac{v-1}{2}}, \quad (27)$$

and so the conditional expectation of τ and $\log \tau$ given $T = t$ are

$$E(\tau | T = t) = \frac{v+1}{a^2(t; \alpha, \beta) + v} \frac{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+3}{a^2(t; \alpha, \beta) + v}}, v+3\right)}{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+1}{a^2(t; \alpha, \beta) + v}}, v+1\right)},$$

$$E(\log \tau | T = t) = DG\left(\frac{v+1}{2}\right) - \log\left(\frac{a^2(t; \alpha, \beta) + v}{2}\right) + \frac{v+1}{a^2(t; \alpha, \beta) + v} \left(\frac{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+3}{a^2(t; \alpha, \beta) + v}}, v+3\right)}{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+1}{a^2(t; \alpha, \beta) + v}}, v+1\right)} - 1 \right) + \lambda a(t; \alpha, \beta) \frac{a^2(t; \alpha, \beta) - 1}{\sqrt{(v+1)(a^2(t; \alpha, \beta) + v)}^3} + \frac{t\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+1}{a^2(t; \alpha, \beta) + v}}, v+1\right)}{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+1}{a^2(t; \alpha, \beta) + v}}, v+1\right)} + \frac{1}{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{v+1}{a^2(t; \alpha, \beta) + v}}, v+1\right)} \int_{-\infty}^{\lambda a(t; \alpha, \beta) \sqrt{\frac{v+1}{a^2(t; \alpha, \beta) + v}}} g(x; v) t(x; v+1) dx, \quad (28)$$

respectively, where $g(x; v) = DG\left(\frac{v+2}{2}\right) - DG\left(\frac{v+1}{2}\right) - \log\left(1 + \frac{x^2}{v+1}\right) + \frac{(v+1)x^2 - v - 1}{(v+1)(v+1+x^2)}$. Also

Fig. 2 The graph of densities, BS, ST-BS, for the selected of values of parameters

$$E(\gamma\tau \mid T = t) = \lambda a(t; \alpha, \beta) E(\tau \mid T = t) + \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)} \frac{1}{\sqrt{\pi}} \frac{1}{T\left(\lambda a(t; \alpha, \beta) \sqrt{\frac{\nu+1}{a^2(t; \alpha, \beta) + \nu}}; \nu + 1\right)} \frac{1}{(\lambda^2 a^2(t; \alpha, \beta) + a^2(t; \alpha, \beta) + \nu)^{\left(\frac{\nu+2}{2}\right)}}.$$

Note that in the case of $\tau = 1$ with probability 1, the SN-BS distribution is obtained. Figure 2 displays the graph of the densities, BS, ST-BS with $\alpha = 0.5$, $\beta = 0.8$, $\lambda = 3$ and four different degrees of freedom $\nu = 1, 3, 5, 15$.

2.3 The Skew Generalized Laplace Normal-BS Distribution

A random variable Y is said to have a generalized Laplace (GL) distribution, denote by $Y \sim GL(\nu)$, if it has the pdf

$$f_{GL}(y; \nu) = \frac{1}{\sqrt{2\pi} 2^{\nu-1} \Gamma(\nu)} |y|^{v-\frac{1}{2}} k_{s-\nu}(|y|), \quad (29)$$

where $k_r(x)$ denotes the modified Bessel function of third kind of order $r \in R$ defined by

$$k_r(x) = \frac{1}{2} \int_0^\infty y^{r-1} \exp\left(-\frac{x}{2}\left(y + \frac{1}{y}\right)\right) dy, \quad x > 0. \quad (30)$$

The first derivative of (30) with respect to r is

$$k'_r(x) = \frac{dk_r(x)}{dr} = \frac{1}{2} \int_0^\infty (\log y) y^{r-1} \exp\left(-\frac{x}{2}\left(y + \frac{1}{y}\right)\right) dy, \quad x > 0. \quad (31)$$

Let $\tau_1 \sim \frac{1}{\chi^2_{(2v)}}$, $\tau_2 = 1$ with probability 1 in (2), where $\chi^2_{(2v)}$ denotes the chi squared distribution with df $2v$, then Y is said to have the skew generalized Laplace normal distribution with pdf given by

$$f_{SGLN}(y; \lambda, v) = 2f_{GL}(y; v)\Phi(\lambda y), \quad (32)$$

and will be denoted by $Y \sim SGLN(\lambda, v)$. See Jammalizadeh and Lin (2016).

Definition 2.4 If $Y \stackrel{d}{=} Z \sim SGLN(\lambda, v)$ in (6), we say that the random variable T follows the SGLN-BS distribution with pdf

$$f_{SGLN-BS}(t; \alpha, \beta, \lambda, v) = 2f_{GL}(a(t; \alpha, \beta); v)\Phi(\lambda a(t; \alpha, \beta))A(t; \alpha, \beta) \quad t, \alpha, \beta, v > 0, \lambda \in R. \quad (33)$$

We write $T \sim SGLN-BS(\alpha, \beta, \lambda, v)$ to denote that T follows the *SGLN – BS* distribution. The mean, variance and measures of skewness and kurtosis coefficients of T follows from equation in (13)–(15), for which, we have

$$T^*(v) \stackrel{d}{=} \sqrt{\chi^2(2v)}Z(0, 1), \quad (34)$$

with

$$\begin{aligned} ET^{*2} &= 2v, \\ ET^{*4} &= 12v(v+1), \\ ET^{*6} &= 120v(v+1)(v+2), \\ ET^{*8} &= 1680v(v+1)(v+2)(v+3), \\ Y &\sim SGLN(\alpha, \beta, \lambda, v). \end{aligned}$$

From (12), we can show that the conditional pdf of τ given $T = t$ is

$$\tau \mid T = t \sim GIG\left(v = \frac{1}{2} - \alpha, \kappa = 1, \psi = a^2(t; \alpha, \beta)\right), \quad (35)$$

where GIG is the Generalized Inverse Gaussian distribution, see Jørgensen [21].

So we obtain

Fig. 3 The graph of densities, SGLN-BS, for the selected of values

$$E(\tau \mid T = t) = |a(t; \alpha, \beta)|^{-1} \frac{k_{v+1}(|a(t; \alpha, \beta)|)}{k_v(|a(t; \alpha, \beta)|)},$$

$$E(\log \tau \mid T = t) = -\log |a(t; \alpha, \beta)| + \frac{k'_v(|a(t; \alpha, \beta)|)}{k_v(|a(t; \alpha, \beta)|)}. \quad (36)$$

Figure 3 displays several density functions of T with $\alpha = 0.5$, $\beta = 0.8$, $\lambda = 3$ and four different $v = 1, 3, 5$ and 15 .

Remark: Some special cases are as follows:

(I) If $\tau_1 = 1$ with probability 1 and $\tau_2 \sim \frac{1}{\chi^2_{(2v)}}$, then $T \sim \text{SNGL-BS}(\alpha, \beta, \lambda, v)$ with density function

$$f_{\text{SNGL-BS}}(t) = 2\phi(a(t; \alpha, \beta))F_{GL}(\lambda a(t; \alpha, \beta))A(t; \alpha, \beta),$$

(II) If $\tau_1 \sim \frac{1}{\chi^2_{(2v_1)}}$, $\tau_2 \sim \Gamma\left(\frac{v_2}{2}, \frac{v_2}{2}\right)$, then $T \sim \text{SGLT-BS}(\alpha, \beta, \lambda, v_1, v_2)$ with density function

$$f_{\text{SGLT-BS}}(t) = 2f_{GL}(a(t; \alpha, \beta); v_1)T(\lambda a(t; \alpha, \beta); v_2)A(t; \alpha, \beta),$$

(III) If $\tau_1 \sim \Gamma\left(\frac{v_1}{2}, \frac{v_1}{2}\right)$, $\tau_2 \sim \frac{1}{\chi^2_{(2v_1)}}$, then $T \sim \text{STGL-BS}(\alpha, \beta, \lambda, v_1, v_2)$ with density function

$$f_{\text{STGL-BS}}(t) = 2t(a(t; \alpha, \beta); v_1)F_{GL}(\lambda a(t; \alpha, \beta); v_2)A(t; \alpha, \beta).$$

where $F_{GL}(\cdot; v_2)$ is the cdf of the generalized Laplace distribution.

2.4 The Skew Slash Normal-BS Distribution

Let $\tau_1 \sim \text{Beta}\left(\frac{v}{2}, 1\right)$, $\tau_2 = 1$ with probability in (2), then we obtain the pdf of the skew slash normal distribution with shape parameter v and skewness parameter λ , denoted by $Y \sim \text{SSN}(v, \lambda)$. The pdf is given by

$$f_{\text{SSN}}(y; v, \lambda) = 2f_s(y; v)\Phi(\lambda y), \quad (37)$$

where $f_s(\cdot; v)$ is the pdf of the slash distribution with shape parameter v .

$$f_s(y; v) = \begin{cases} \frac{v\Gamma\left(\frac{v+1}{2}\right)2^{\frac{v}{2}-1}G\left(\frac{y^2}{2}, \frac{v+1}{2}\right)}{\sqrt{\pi}|y|^{\frac{v+1}{2}}} & y \neq 0, \\ \frac{v}{\sqrt{2\pi(v+1)}} & y = 0, \end{cases} \quad (38)$$

where $G(\cdot; \alpha)$ denotes the cdf of gamma distribution with scale parameter 1 and shape parameter α . See Rogers and Tukey [35].

Definition 2.5 If $Y \stackrel{d}{=} Z \sim \text{SSN}(v, \lambda)$ in (6), we say that the random variable T follows the SSN-BS with pdf

$$f(t; \alpha, \beta, \lambda, v) = 2f_s(a(t; \alpha, \beta); v)\Phi(\lambda a(t; \alpha, \beta))A(t; \alpha, \beta) \quad t, \alpha, \beta, v_2 > 0, \lambda \in \mathbb{R}. \quad (39)$$

We write $T \sim \text{SSN-BS}(\alpha, \beta, \lambda, v)$ to denote that T follows the SSN-BS distribution.

Figure 4 displays several density functions of $T \sim \text{SSN-BS}(\alpha, \beta, \lambda, v)$ with $\alpha = 0.5, \beta = 0.8, \lambda = 3$ and four different $v = 1, 3, 5, 15$.

Fig. 4 The graph of the densities, SSN-BS, for selected values of parameters

The mean, variance and measures of skewness and kurtosis coefficients of T follows equations in (13), (14), (15), for which we have

$$T^*(v) \stackrel{d}{=} \text{Beta}^{-\frac{1}{2}}\left(\frac{v}{2}, 1\right) Z(0, 1), \quad (40)$$

$$\begin{aligned} ET^{*2} &= \frac{v}{v-2}, \\ ET^{*4} &= 3 \frac{v}{v-4}, \\ ET^{*6} &= 15 \frac{v}{v-6}, \\ ET^{*8} &= 105 \frac{v}{v-8}, Y \sim SSN(v, \lambda). \end{aligned}$$

From (12), we can show that the conditional pdf of τ given $T = t$ is

$$f(\tau | T = t) = \begin{cases} \frac{|a(t; \alpha, \beta)|^{v+1} \tau^{\frac{v-1}{2}} \exp\left(-\frac{1}{2}a^2(t; \alpha, \beta)\right)}{\Gamma\left(\frac{v+1}{2}\right) 2^{\frac{v+1}{2}} G\left(\frac{a^2(t; \alpha, \beta)}{2}; \frac{v+1}{2}\right)} & t \neq \beta, \\ \frac{v+1}{2} \tau^{\frac{v+1}{2}-1} & t = \beta \end{cases}, \quad (41)$$

and the conditional expectation of τ and $\log(\tau)$ given $T = t$ is

$$\begin{aligned} E(\tau | T = t) &= \begin{cases} \frac{(v+1)G\left(a^2(t; \alpha, \beta); \frac{v+3}{2}\right)}{a^2(t; \alpha, \beta)G\left(\frac{a^2(t; \alpha, \beta)}{2}; \frac{v+1}{2}\right)} & t \neq \beta, \\ \frac{v+1}{v+3} & t = \beta \end{cases}, \\ E(\log \tau | T = t) &= \begin{cases} \log \frac{2}{a^2(t; \alpha, \beta)} + \frac{\int_0^{\frac{a^2(t; \alpha, \beta)}{2}} (\log x) x^{\frac{v-1}{2}} \exp(-x) dx}{\Gamma\left(\frac{v+1}{2}\right) G\left(\frac{a^2(t; \alpha, \beta)}{2}; \frac{v+1}{2}\right)} & t \neq \beta, \\ -\frac{2}{v+1} & t = \beta \end{cases}. \end{aligned} \quad (42)$$

3 Maximum Likelihood Estimation

In this section, we derive the ML estimation parameters of SSMSN-BS distributions via modification of the EM-algorithm (ECM-algorithm). The ECM algorithm modifies the EM Algorithm by replacing its Maximization step by a sequence of conditional maximization steps. For more details about EM and ECM algorithms, see Dempster et al. [13] and Meng and Rubin [30].

3.1 The General Case

Let $T_1, T_2, \dots, T_n$ are random samples from $SSMSN - BS(\alpha, \beta, \lambda, \eta)$ with the following hierarchical formulation

$$f(T_i = t_i | (\gamma_i, \tau_{1i}, \tau_{2i})) = \phi\left(\sqrt{\tau_{2i}(f_i + \lambda^2)}(a(t; \alpha, \beta)) - \frac{\sqrt{\tau_{2i}}\lambda\gamma}{\sqrt{f_i + \lambda^2}}\right) \times \sqrt{\tau_{2i}(f_i + \lambda^2)}A(t; \alpha, \beta), \gamma_i, | (\tau_{1i}, \tau_{2i}) \sim TN\left(0, \frac{f_i + \lambda^2}{\tau_{1i}}; (0, \infty)\right), \quad (43)$$

where $\gamma_i = \sqrt{\tau_{1i}(f_i + \lambda^2)}|Z_1|$, $f_i = \frac{\tau_{1i}}{\tau_{2i}}$, and $\tau_i = (\tau_{1i}, \tau_{2i})^T$ are n bivariate positive random samples with pdf $h(\tau_{1i}, \tau_{2i}; \eta)$.

Set the observed data by $\mathbf{t} = (t_1, \dots, t_n)^T$, the missing data $\gamma = (\gamma_1, \dots, \gamma_n)^T$ and $\tau = (\tau_1, \dots, \tau_n)^T$ and the complete data by $\mathbf{t}^c = (\mathbf{t}^T, \gamma^T, \tau^T)^T$. Then from (43), and $h(\tau_i, \eta)$ we can construct the complete data log-likelihood function of $\theta = (\alpha, \beta, \lambda, \eta)$ given complete data $\mathbf{t}^c = (t, \gamma, \tau_1, \tau_2)$. By ignoring the additive constant terms, the log-likelihood function is given by

$$\begin{aligned} \ell^{(c)}(\theta | \mathbf{t}^{(c)}) &= \sum_{i=1}^n \ell_i^{(c)}(\theta | (t_i, \tau_{1i}, \tau_{2i}, \gamma_i)) \\ &= -\frac{1}{2} \sum_{i=1}^n \tau_{2i} \gamma_i^2 + \frac{\lambda}{\alpha} \sum_{i=1}^n \tau_{2i} \gamma_i \varepsilon(t_i; \beta) \\ &\quad - \frac{1}{2} \frac{\lambda^2}{\alpha^2} \sum_{i=1}^n \tau_{2i} \varepsilon^2(t_i; \beta) - \frac{1}{2\alpha^2} \sum_{i=1}^n \tau_{1i} \varepsilon^2(t_i; \beta) \\ &\quad + \sum_{i=1}^n \log \frac{t_i + \beta}{\sqrt{\beta}} - n \log(\alpha) + \sum_{i=1}^n \log(\sqrt{\tau_{1i} \tau_{2i}} h(\tau_{1i}, \tau_{2i}; \eta)), \end{aligned} \quad (44)$$

where $\varepsilon(t_i; \beta) = \sqrt{\frac{t_i}{\beta}} - \sqrt{\frac{\beta}{t_i}}$. Suppose $\hat{\theta}^{(r)} = (\hat{\alpha}^{(r)}, \hat{\beta}^{(r)}, \hat{\lambda}^{(r)}, \hat{\eta}^{(r)})$ is the current estimate (in the r th iteration) of θ . Based on the ECM algorithm principle, in the E-step, we should first form the following conditional expectation

$$\begin{aligned} Q(\theta | \hat{\theta}^{(r)}) &= E(\ell^{(c)}(\theta | \mathbf{t}^{(c)})) \\ &= \frac{\lambda}{\alpha} \sum_{i=1}^n \hat{S}_{3i}^{(r)} \varepsilon(t_i; \beta) - \frac{1}{2} \frac{\lambda^2}{\alpha^2} \sum_{i=1}^n \hat{S}_{2i}^{(r)} \varepsilon^2(t_i; \beta) \\ &\quad - \frac{1}{2\alpha^2} \sum_{i=1}^n \hat{S}_{1i}^{(r)} \varepsilon^2(t_i; \beta) + \sum_{i=1}^n \log \frac{t_i + \beta}{\sqrt{\beta}} - n \log(\alpha) \\ &\quad + \sum_{i=1}^n \left[E\left(\log(\sqrt{\tau_{1i} \tau_{2i}} h(\tau_{1i}, \tau_{2i}; \eta)) | t_i, \hat{\theta}^{(r)}) \right] \end{aligned} \quad (45)$$

where

$$\begin{aligned}\hat{S}_{1i}^{(r)} &= E(\tau_{1i} \mid t_i, \hat{\boldsymbol{\theta}}^{(r)}), \\ \hat{S}_{2i}^{(r)} &= E(\tau_{2i} \mid t_i, \hat{\boldsymbol{\theta}}^{(r)}), \\ \hat{S}_{3i}^{(r)} &= E(\tau_{2i} \gamma_i \mid t_i, \hat{\boldsymbol{\theta}}^{(r)}).\end{aligned}\quad (46)$$

Then the ECM algorithm is done as follows:

E-step: Given $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}^{(r)}$, compute $\hat{S}_{1i}^{(r)}, \hat{S}_{2i}^{(r)}, \hat{S}_{3i}^{(r)}$, for $i = 1, \dots, n$.

CM-step 1: Fix $\beta = \hat{\beta}^{(r)}$ and update $\hat{\alpha}^{(r)}, \hat{\lambda}^{(r)}$ by maximizing (45) over α and λ , which leads to

$$\begin{aligned}\hat{\lambda}^{(r+1)} &= \frac{\hat{\alpha}^{(r+1)} \sum_{i=1}^n \epsilon(t_i; \hat{\beta}^{(r)}) \hat{S}_{3i}^{(r)}}{\sum_{i=1}^n \epsilon^2(t_i; \hat{\beta}^{(r)}) \hat{S}_{2i}^{(r)}}, \\ \hat{\alpha}^{2(r+1)} &= \frac{1}{n} \sum_{i=1}^n \epsilon^2(t_i; \hat{\beta}^{(r)}) \hat{S}_{1i}^{(r)}.\end{aligned}\quad (47)$$

CM-step 2: Updating $\hat{\boldsymbol{\eta}}^{(r)}$ is strongly related to form of the $h(\tau_{1i}, \tau_{2i}; \boldsymbol{\eta})$ said to the next section.

CM-step 3: Fix $\alpha = \hat{\alpha}^{(r+1)}, \lambda = \hat{\lambda}^{(r+1)}, \boldsymbol{\eta} = \hat{\boldsymbol{\eta}}^{(r+1)}$ and update $\hat{\beta}^{(r)}$ using

$$\hat{\beta}^{(r+1)} = \arg \max Q(\hat{\alpha}^{(r+1)}, \beta, \hat{\lambda}^{(r+1)}, \hat{\boldsymbol{\eta}}^{(r+1)} \mid \hat{\boldsymbol{\theta}}^{(r)}). \quad (48)$$

Note that the CM-step 3 requires a one-dimensional search for the root of β , which can be easily obtained by using the *optimize function* in the statistical software *R*.

3.2 ECM-Estimation for the $ST_{v_1} T_{v_2}$ -BS Distribution

For the $ST_{v_1} T_{v_2}$ -BS distribution, we recall equation (16)

$$h(\tau_i; \boldsymbol{\eta}) = f\left(\tau_{1i}; \frac{v_1}{2}, \frac{v_1}{2}\right) f\left(\tau_{2i}; \frac{v_2}{2}, \frac{v_2}{2}\right) \quad \boldsymbol{\eta} = (v_1, v_2), \quad (49)$$

from (45), the E-step involves the calculation of two additional conditional expectations, given by

$$\begin{aligned}\hat{S}_{4i}^{(r)} &= E(\log \tau_{1i} \mid t_i, \hat{\boldsymbol{\theta}}^{(r)}), \\ \hat{S}_{5i}^{(r)} &= E(\log \tau_{2i} \mid t_i, \hat{\boldsymbol{\theta}}^{(r)}),\end{aligned}\quad (50)$$

which can be calculated by using (21). Now, rewrite

CM-step 2: Fix $\alpha = \hat{\alpha}^{(r+1)}, \lambda = \hat{\lambda}^{(r+1)}, \beta = \hat{\beta}^{(r)}$ and update $\hat{v}_1^{(r)}$ and $\hat{v}_2^{(r)}$ by maximizing (45) over v_1, v_2 which leads to solve the root of the following equations

$$\begin{aligned}\log \frac{v_1}{2} - DG\left(\frac{v_1}{2}\right) + 1 + \frac{1}{n} \sum_{i=1}^n \left(\widehat{S}_{3i}^{(r)} - \widehat{S}_{1i}^{(r)}\right) &= 0, \\ \log \frac{v_2}{2} - DG\left(\frac{v_2}{2}\right) + 1 + \frac{1}{n} \sum_{i=1}^n \left(\widehat{S}_{3i}^{(r)} - \widehat{S}_{1i}^{(r)}\right) &= 0.\end{aligned}\quad (51)$$

3.3 ECM-Estimation for the ST-BS Distribution

For the ST-BS distribution, we recall equation (25)

$$h(\tau_i; v) = f\left(\tau_i; \frac{v}{2}, \frac{v}{2}\right), \quad (52)$$

The E-step (46) converts to

$$\begin{aligned}\widehat{S}_{1i}^{(r)} &= \widehat{S}_{2i}^{(r)} = \widehat{S}_i^{(r)} = E(\tau_i \mid t_i, \widehat{\boldsymbol{\theta}}^{(r)}), \\ \widehat{S}_{3i}^{(r)} &= E(\gamma_i \mid t_i, \widehat{\boldsymbol{\theta}}^{(r)}),\end{aligned}\quad (53)$$

and one additional conditional expectation, given by

$$\widehat{S}_{4i}^{(r)} = E(\log \tau_i \mid t_i, \widehat{\boldsymbol{\theta}}^{(r)}), \quad (54)$$

which can be calculated by using (26). Now, we rewrite

CM-step 2: Updating $\widehat{v}^{(r)}$ leads to solve the root of the following equation

$$\log \frac{v}{2} - DG\left(\frac{v}{2}\right) + 1 + \frac{1}{n} \sum_{i=1}^n \left(\widehat{S}_{4i}^{(r)} - \widehat{S}_{1i}^{(r)}\right) = 0. \quad (55)$$

3.4 ECM-Estimation for the SGLN-BS Distribution

As mentioned in Sect. 2.3, the SGLN-BS distribution can be generated from SSMSN-BS by taking $\tau_{1i} = \tau_i = \frac{d}{\chi_{(2v)}^2}$ and $\tau_{2i} = 1$ with probability 1 for $i = 1, \dots, n$, and so we have

$$h(\boldsymbol{\tau}_i; \boldsymbol{\eta}) = h(\tau_i; v) = \frac{1}{2^v \Gamma(v)} \tau_i^{-(v+1)} \exp\left(-\frac{1}{2\tau_i}\right). \quad (56)$$

Then the E-step (46) converts to

$$\begin{aligned}\widehat{S}_{1i}^{(r)} &= E(\tau_i \mid t_i, \widehat{\boldsymbol{\theta}}^{(r)}), \\ \widehat{S}_{2i}^{(r)} &= 1, \\ \widehat{S}_{3i}^{(r)} &= E(\gamma_i \mid t_i, \widehat{\boldsymbol{\theta}}^{(r)}),\end{aligned}\quad (57)$$

and additional conditional expectation

$$\hat{S}_{4i}^{(r)} = E(\log \tau_i \mid t_i, \hat{\theta}^{(r)}), \quad (58)$$

which can be calculated via (36). Now we rewrite

CM-step 2: Updating $\hat{v}^{(r)}$ leads to solve the root of the following equation

$$\log 2 + DG(v) + \frac{1}{n} \sum_{i=1}^n \hat{S}_{4i}^{(r)} = 0. \quad (59)$$

3.5 ECM-Estimation for the SSN-BS Distribution

As shown in Sect. 2.4, the SSN can be considered as a special case of SSMSN-BS by taking

$\tau_{1i} = \tau_i \stackrel{d}{=} \text{Beta}(v, 1)$ and $\tau_{2i} = 1$ with probability 1 for $i = 1, \dots, n$. So we have

$$h(\tau_i; \eta) = h(\tau_i; v) = v\tau_i^{v-1}. \quad (60)$$

The E-step (46) converts to

$$\begin{aligned} \hat{S}_{1i}^{(r)} &= E(\tau_i \mid t_i, \hat{\theta}^{(r)}), \\ \hat{S}_{2i}^{(r)} &= 1, \\ \hat{S}_{3i}^{(r)} &= E(\gamma_i \mid t_i, \hat{\theta}^{(r)}), \end{aligned} \quad (61)$$

and additional conditional expectation

$$\hat{S}_{4i}^{(r)} = E(\log \tau_i \mid t_i, \hat{\theta}^{(r)}). \quad (62)$$

which can be calculated via (42). Now we rewrite

CM-step 2: Update $\hat{v}^{(r)}$ by

$$\hat{v}^{(r+1)} = -\frac{n}{\hat{S}_{4i}^{(r)}}. \quad (63)$$

3.6 The Information Matrix

Under some regularity condition, the asymptotic covariance matrix of the ML estimates can be approximated by inverse of the observed information matrix, given by

$$I_0(\hat{\theta} \mid \mathbf{t}) = \frac{-\partial^2 \ell(\theta \mid \mathbf{t})}{\partial \theta \partial \theta^T} \Big|_{\theta=\hat{\theta}},$$

where $\ell(\theta \mid \mathbf{t})$ is the observed log-likelihood function on the basis of observations $T = (t_1, \dots, t_n)^T$.

Following Basford et al. [6], we obtain

$$I_0(\hat{\theta} | \mathbf{t}) = \sum_{i=1}^n \hat{\mathbf{S}}_i \hat{\mathbf{S}}_i^T,$$

where

$$\hat{\mathbf{S}}_i = \frac{\partial \ell_i(\boldsymbol{\theta} | t_i)}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = (\hat{\mathbf{S}}_{i,\alpha}, \hat{\mathbf{S}}_{i,\beta}, \hat{\mathbf{S}}_{i,\lambda}, \hat{\mathbf{S}}_{i,v})^T,$$

is the individual score statistic corresponding to the single observation t_i ($i = 1, \dots, n$). Explicit expression for the elements of $\hat{\mathbf{S}}_i$ are summarized below.

(i) For the $ST_{v_1} T_{v_2} - BS$ distribution, the elements of $\hat{\mathbf{S}}_i$ contain

$$\begin{aligned} \hat{\mathbf{S}}_{i,\alpha} &= - \left(\frac{\hat{v}_1 + 1}{\hat{v}_1} \right) \left(1 + \frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{\hat{v}_1} \right)^{-1} a(t_i; \hat{\alpha}, \hat{\beta}) \frac{\partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta}) \\ &\quad + \frac{t(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}), \hat{v}_2)}{T(\hat{\lambda}^{(r)} a(t_i; \hat{\alpha}^{(r)}, \hat{\beta}^{(r)}), \hat{v}_2)} \frac{\hat{\lambda} \partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta}) + \frac{\frac{\partial}{\partial \alpha} A(t_i, \hat{\alpha}, \hat{\beta})}{A(t_i, \hat{\alpha}, \hat{\beta})}, \\ \hat{\mathbf{S}}_{i,\beta} &= - \left(\frac{\hat{v}_1 + 1}{\hat{v}_1} \right) \left(1 + \frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{\hat{v}_1} \right)^{-1} a(t_i; \hat{\alpha}, \hat{\beta}) \frac{\partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta}) \\ &\quad + \frac{t(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}), \hat{v}_2)}{T(\hat{\lambda}^{(r)} a(t_i; \hat{\alpha}^{(r)}, \hat{\beta}^{(r)}), \hat{v}_2)} \frac{\hat{\lambda} \partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta}) + \frac{\frac{\partial}{\partial \beta} A(t_i, \hat{\alpha}, \hat{\beta})}{A(t_i, \hat{\alpha}, \hat{\beta})}, \\ \hat{\mathbf{S}}_{i,\lambda} &= \frac{t(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}), \hat{v}_2)}{T(\hat{\lambda}^{(r)} a(t_i; \hat{\alpha}^{(r)}, \hat{\beta}^{(r)}), \hat{v}_2)} a(t_i; \hat{\alpha}, \hat{\beta}), \\ \hat{\mathbf{S}}_{i,v_1} &= \frac{1}{2} DG \left(\frac{\hat{v}_1 + 1}{2} \right) - \frac{1}{2} DG \left(\frac{\hat{v}_1}{2} \right) - \frac{1}{2} \hat{v}_1^{-1} - \frac{1}{2} \ln \left(1 + \frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{\hat{v}_1} \right) \\ &\quad + \frac{a^2(t_i; \hat{\alpha}, \hat{\beta}) (\hat{v}_1 + 1)}{2 \hat{v}_1 (\hat{v}_1 + a^2(t_i; \hat{\alpha}, \hat{\beta}))}, \\ \hat{\mathbf{S}}_{i,v_2} &= \frac{1}{2} DG \left(\frac{\hat{v}_1 + 1}{2} \right) - \frac{1}{2} DG \left(\frac{\hat{v}_1}{2} \right) - \frac{1}{2} \hat{v}_1^{-1} + \frac{1}{2} \int_{-\infty}^{\lambda a(t_i; \hat{\alpha}, \hat{\beta})} (-\ln \left(1 + \frac{x^2}{\hat{v}_2} \right) \\ &\quad + \frac{x^2 (\hat{v}_2 + 1)}{2 \hat{v}_2 (\hat{v}_2 + x^2)}) t(x; \hat{v}_2) dx. \end{aligned}$$

In the case of $v_1 = v_2 = v$, we have $\hat{\mathbf{S}}_{i,v} = \hat{\mathbf{S}}_{i,v_1} + \hat{\mathbf{S}}_{i,v_2}$.

(ii) For the ST-BS distribution, the elements of $\hat{\mathbf{S}}_i$ are obtained by

$$\begin{aligned}
\hat{\mathbf{S}}_{i,\alpha} &= - \left(\frac{\hat{v}+1}{\hat{v}} \right) \left(1 + \frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{\hat{v}} \right)^{-1} a(t_i; \hat{\alpha}, \hat{\beta}) \frac{\partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta}) \\
&\quad + \frac{t(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}; \hat{v}+1) \hat{\lambda} \frac{\partial a(t_i; \hat{\alpha}, \hat{\beta})}{\partial \alpha}}{T(\hat{\lambda}^{(r)} a(t_i; \hat{\alpha}^{(r)}, \hat{\beta}^{(r)}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}, \hat{v}+1)} \left(\sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}} \right. \\
&\quad \left. - a^2(t_i; \hat{\alpha}, \hat{\beta}) \sqrt{\frac{\hat{v}+1}{(a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v})^3}} \right) + \frac{\frac{\partial}{\partial \alpha} A(t_i, \hat{\alpha}, \hat{\beta})}{A(t_i, \hat{\alpha}, \hat{\beta})}, \\
\hat{\mathbf{S}}_{i,\beta} &= - \left(\frac{\hat{v}+1}{\hat{v}} \right) \left(1 + \frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{\hat{v}} \right)^{-1} a(t_i; \hat{\alpha}, \hat{\beta}) \frac{\partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta}) \\
&\quad + \frac{t(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}; \hat{v}) \hat{\lambda} \frac{\partial a(t_i; \hat{\alpha}, \hat{\beta})}{\partial \beta}}{T(\hat{\lambda}^{(r)} a(t_i; \hat{\alpha}^{(r)}, \hat{\beta}^{(r)}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}, \hat{v}+1)} \left(\sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}} \right. \\
&\quad \left. - a^2(t_i; \hat{\alpha}, \hat{\beta}) \sqrt{\frac{\hat{v}+1}{(a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v})^3}} \right) + \frac{\frac{\partial}{\partial \beta} A(t_i, \hat{\alpha}, \hat{\beta})}{A(t_i, \hat{\alpha}, \hat{\beta})}, \\
\hat{\mathbf{S}}_{i,\lambda} &= \frac{t(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}; \hat{v}+1)}{T(\hat{\lambda}^{(r)} a(t_i; \hat{\alpha}^{(r)}, \hat{\beta}^{(r)}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}, \hat{v}+1)} a(t_i; \hat{\alpha}, \hat{\beta}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}, \\
\hat{\mathbf{S}}_{i,v} &= \frac{1}{2} DG \left(\frac{\hat{v}+1}{2} \right) - \frac{1}{2} DG \left(\frac{\hat{v}}{2} \right) - \frac{1}{2} \hat{v}^{-1} - \frac{1}{2} \ln \left(1 + \frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{\hat{v}} \right) \\
&\quad + \frac{a^2(t_i; \hat{\alpha}, \hat{\beta}) (\hat{v}+1)}{2 \hat{v} (\hat{v} + a^2(t_i; \hat{\alpha}, \hat{\beta}))} \\
&\quad + \frac{1}{2} DG \left(\frac{\hat{v}+2}{2} \right) - \frac{1}{2} DG \left(\frac{\hat{v}+1}{2} \right) \\
&\quad - \frac{\frac{1}{2} \int_{-\infty}^{\lambda a(t_i; \hat{\alpha}, \hat{\beta}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}}{\frac{1}{2} T \left(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}) \sqrt{\frac{\hat{v}+1}{a^2(t_i; \hat{\alpha}, \hat{\beta}) + \hat{v}}}, \hat{v}+1 \right)} \left(-\ln \left(1 + \frac{x^2}{\hat{v}+1} \right) + \frac{x^2 (\hat{v}+2)}{(\hat{v}+1)(\hat{v}+1+x^2)} \right) t(x, \hat{v}+1) dx.
\end{aligned}$$

(iii) For the SGLN-BS distribution, the elements of $\hat{\mathbf{S}}_i$ are obtained by

$$\begin{aligned}
\hat{\mathbf{S}}_{i,\alpha} &= \left(\hat{v} - \frac{1}{2}\right) \frac{\frac{\partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta})}{a(t_i; \hat{\alpha}, \hat{\beta})} + \frac{k'_{0.5-\hat{v}} \left(\left| a(t_i; \hat{\alpha}, \hat{\beta}) \right| \right)}{k_{0.5-\hat{v}} \left(\left| a(t_i; \hat{\alpha}, \hat{\beta}) \right| \right)} \frac{a(t_i; \hat{\alpha}, \hat{\beta}) \frac{\partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta})}{\left| a(t_i; \hat{\alpha}, \hat{\beta}) \right|} \\
&\quad + \frac{\phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))}{\Phi(\hat{\lambda}^{(r)} a(t_i; \hat{\alpha}^{(r)}, \hat{\beta}^{(r)}))} \frac{\hat{\lambda} \partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta}) + \frac{\frac{\partial}{\partial \alpha} A(t_i, \hat{\alpha}, \hat{\beta})}{A(t_i, \hat{\alpha}, \hat{\beta})}, \\
\hat{\mathbf{S}}_{i,\beta} &= \left(\hat{v} - \frac{1}{2}\right) \frac{\frac{\partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta})}{a(t_i; \hat{\alpha}, \hat{\beta})} + \frac{k'_{0.5-\hat{v}} \left(\left| a(t_i; \hat{\alpha}, \hat{\beta}) \right| \right)}{k_{0.5-\hat{v}} \left(\left| a(t_i; \hat{\alpha}, \hat{\beta}) \right| \right)} \frac{a(t_i; \hat{\alpha}, \hat{\beta}) \frac{\partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta})}{\left| a(t_i; \hat{\alpha}, \hat{\beta}) \right|} \\
&\quad + \frac{\phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))}{\Phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))} \frac{\hat{\lambda} \partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta}) + \frac{\frac{\partial}{\partial \beta} A(t_i, \hat{\alpha}, \hat{\beta})}{A(t_i, \hat{\alpha}, \hat{\beta})}, \\
\hat{\mathbf{S}}_{i,\lambda} &= \frac{\phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))}{\Phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))} a(t_i; \hat{\alpha}, \hat{\beta}), \\
\hat{\mathbf{S}}_{i,v} &= -\ln 2 - DG(\hat{v}) + \ln \left| a(t_i, \hat{\alpha}, \hat{\beta}) \right| - \frac{k'_{0.5-\hat{v}} \left(\left| a(t_i; \hat{\alpha}, \hat{\beta}) \right| \right)}{k_{0.5-\hat{v}} \left(\left| a(t_i; \hat{\alpha}, \hat{\beta}) \right| \right)}.
\end{aligned}$$

(iv) For the SSN-BS distribution, the elements of $\hat{\mathbf{S}}_i$ are obtained by

$$\begin{aligned}
\hat{\mathbf{S}}_{i,\alpha} &= \frac{\left(a(t_i; \hat{\alpha}, \hat{\beta}) \right)^{\hat{v}} \exp \left(-a^2(t_i; \hat{\alpha}, \hat{\beta}) \right) \frac{\partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta})}{\Gamma \left(\frac{\hat{v}+1}{2} \right) G \left(\frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{2}, \frac{\hat{v}+1}{2} \right)} - \frac{(\hat{v}+1) \frac{\partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta})}{a(t_i; \hat{\alpha}, \hat{\beta})} \\
&\quad + \frac{\phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))}{\Phi(\hat{\lambda}^{(r)} a(t_i; \hat{\alpha}^{(r)}, \hat{\beta}^{(r)}))} \frac{\hat{\lambda} \partial}{\partial \alpha} a(t_i; \hat{\alpha}, \hat{\beta}) + \frac{\frac{\partial}{\partial \alpha} A(t_i, \hat{\alpha}, \hat{\beta})}{A(t_i, \hat{\alpha}, \hat{\beta})}, \\
\hat{\mathbf{S}}_{i,\beta} &= \frac{\left(a(t_i; \hat{\alpha}, \hat{\beta}) \right)^{\hat{v}} \exp \left(-a^2(t_i; \hat{\alpha}, \hat{\beta}) \right) \frac{\partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta})}{\Gamma \left(\frac{\hat{v}+1}{2} \right) G \left(\frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{2}, \frac{\hat{v}+1}{2} \right)} - \frac{(\hat{v}+1) \frac{\partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta})}{a(t_i; \hat{\alpha}, \hat{\beta})} \\
&\quad + \frac{\phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))}{\Phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))} \frac{\hat{\lambda} \partial}{\partial \beta} a(t_i; \hat{\alpha}, \hat{\beta}) + \frac{\frac{\partial}{\partial \beta} A(t_i, \hat{\alpha}, \hat{\beta})}{A(t_i, \hat{\alpha}, \hat{\beta})}, \\
\hat{\mathbf{S}}_{i,\lambda} &= \frac{\phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))}{\Phi(\hat{\lambda} a(t_i; \hat{\alpha}, \hat{\beta}))} a(t_i; \hat{\alpha}, \hat{\beta}), \\
\hat{\mathbf{S}}_{i,v} &= \frac{1}{\hat{v}} + \frac{\log 2}{2} + \frac{\int_0^{\frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{2}} (\log x) x^{\frac{\hat{v}+1}{2}-1} \exp(-x) dx}{2\Gamma \left(\frac{\hat{v}+1}{2} \right) G \left(\frac{a^2(t_i; \hat{\alpha}, \hat{\beta})}{2}, \frac{\hat{v}+1}{2} \right)} - \log \left| a(t_i; \hat{\alpha}, \hat{\beta}) \right|.
\end{aligned}$$

The covariance matrix can be useful for studying the asymptotic behavior of $\hat{\theta} = [\hat{\alpha}, \hat{\beta}, \hat{\lambda}, \hat{v}]$ by its asymptotic normality. The statistical inference about $\theta = [\alpha, \beta, \lambda, v]$ can be then made.

4 Simulation Study and an Illustrative Example

4.1 Simulation Study

We perform a simulation study to evaluate the finite sample properties of ML estimators described in Sect. 3. To conduct the experimental study, we generate different sample of sizes $n = 50, 100, 200, 400, 800$ from the STN-BS, SGLN-BS, SSN-BS models. The true parameters are supposed to be $\theta = [\alpha, \beta, \lambda, v] = [1, 1, 2, 2]$ and $\theta = [\alpha, \beta, \lambda, v] = [0.5, 1, 2, 2]$. Each simulated data set was fitted via the ECM algorithm under the same generated model using $M = 5000$ replications. To examine the performance of the ML estimates, for each sample size and for each estimate, $\hat{\theta} = [\hat{\alpha}, \hat{\beta}, \hat{\lambda}, \hat{v}]$, we compute the mean, $E[\hat{\theta}_i]$, the root mean square error, $\sqrt{MSE}$, and the relative bias in absolute value, RB , defined as

$$\sqrt{MSE} = \sqrt{E(\hat{\theta}_i - \theta_i)^2},$$

$$RB = \left| \frac{E(\hat{\theta}_i) - \theta_i}{\theta_i} \right|,$$

respectively, such that $\theta_1 = \alpha, \theta_2 = \beta, \theta_3 = \lambda, \theta_4 = v$. Tables 1, 2, 3 present the mean, $\%RB$, and $\sqrt{MSE}$ of the ML estimates of the parameters θ_i $i = 1, 2, 3, 4$ for STN-B, SGLN-BS and SSN-BS models.

Figures 5, 6, 7 and 8 show the graphical representations of RB and $\sqrt{MSE}$ of the four parameter estimates as a function of sample size n , respectively. Clearly, the $\sqrt{MSE}$ and RB values converge to zero when n increases.

4.2 An Illustrative Example

In this section, we consider the results in preceding sections by analyzing a data set originally reported by Bjerkedal [10] and analyzed by Kundu et al. [23]. Table 4 displays the summary of these data including sample median, mean, standard deviation (SD), coefficient of variation (CV), coefficient of skewness (CS), coefficient of kurtosis (CK), range, minimum, maximum and the sample size (n). As it is observed, the data comes from a positively skewed distribution with a kurtosis greater than three. So we propose that these data set can be suitable to modeling some kind of SSMSN-BS distributions listed in Table 5.

Table 1 Parameter estimates for STN-BS model, when $\beta = 1, \lambda = 2, v = 2$

n	α	Estimates of α			Estimates of β			Estimates of λ			Estimates of v		
		Mean	RB	$\sqrt{MSE}$	Mean	RB	$\sqrt{MSE}$	Mean	RB	$\sqrt{MSE}$	Mean	RB	$\sqrt{MSE}$
50	0.50	0.2504	0.4992	0.2496	1.3142	0.3142	0.3142	0.3090	0.8455	1.6910	1.2058	0.3971	0.7942
	1.00	1.0654	0.0653	0.4626	1.1240	0.1240	0.3072	2.5477	0.2739	2.5546	2.3454	0.1727	1.1204
100	0.50	0.3158	0.3683	0.1843	1.1900	.1900	0.1901	0.5059	0.7470	1.4947	1.4362	0.2819	0.5971
	1.00	1.0377	0.0377	0.27511	1.0504	0.0504	0.2091	2.5159	0.2580	1.5419	2.3321	0.1661	1.0208
200	0.50	0.3766	0.2469	0.1234	1.1002	0.1001	0.1002	1.1768	0.4116	0.8232	1.3086	0.3457	0.6914
	1.00	1.0352	0.0352	0.2017	0.9958	0.0042	0.1515	2.3871	0.1936	1.2068	2.0742	0.0371	0.6420
400	0.50	0.5487	0.0975	0.0581	0.9572	0.0428	0.0485	2.5467	0.2734	0.3533	2.2287	0.1144	0.3008
	1.00	1.0127	0.0127	0.1643	1.0113	0.0113	0.1320	2.1742	0.0871	0.8484	2.0694	0.0347	0.3565
800	0.50	0.4866	0.02686	0.0510	1.0137	0.0137	0.0480	1.8695	0.0652	0.4802	1.9814	0.0093	0.2612
	1.00	1.0039	0.0039	0.1199	1.0086	0.0086	0.0948	2.0753	0.0376	0.5706	2.0307	0.0153	0.2344

Table 2 Parameter estimates for SGLN-BS model, when $\beta = 1, \lambda = 2, v = 2$

n	α	Estimates of α			Estimates of β			Estimates of λ			Estimates of v		
		Mean	$\frac{RB}{\sqrt{MSE}}$	Mean	$\frac{RB}{\sqrt{MSE}}$	Mean	$\frac{RB}{\sqrt{MSE}}$	$\frac{RB}{\sqrt{MSE}}$	Mean	$\frac{RB}{\sqrt{MSE}}$	$\frac{RB}{\sqrt{MSE}}$	Mean	$\frac{RB}{\sqrt{MSE}}$
50	0.50	0.7332	0.4664	0.5831	0.0826	0.2413	0.4356	2.3801	2.8711	0.4356	4.3050	10.6099	20.8276
	1.00	1.0637	1.0205	1.0102	0.1260	0.4163	0.1502	2.1967	2.3004	0.1502	7.6242	17.2483	25.7804
100	0.50	0.4565	0.0870	0.1998	0.0600	0.1490	0.0448	1.2135	1.9105	0.0448	1.7750	5.5500	10.5625
	1.00	0.9407	0.0593	0.4659	0.0701	0.2411	0.0845	1.4416	2.1690	0.0845	2.3022	6.6045	14.0323
200	0.50	0.4586	0.08287	0.1662	0.0098	0.01090	0.04602	0.7397	1.9080	0.04602	2.0076	6.0153	12.3688
	1.00	0.9390	0.0610	0.3210	0.0421	0.1871	0.0484	0.7022	1.9032	0.0484	1.2363	4.4727	7.9253
400	0.50	0.4788	0.0424	0.1254	0.0087	0.0724	0.0477	0.5186	1.9046	0.0477	0.6512	3.3023	5.4609
	1.00	0.9639	0.0361	0.2042	0.0417	0.1823	0.0284	0.5658	1.9431	0.0284	0.2760	2.5519	1.8848
800	0.50	0.4896	0.0208	0.0832	0.0055	0.0551	0.0221	0.1063	1.9559	0.0221	0.1597	2.3194	1.1749
	1.00	0.9831	0.0169	0.1403	0.0202	0.1351	0.0117	0.3852	1.9767	0.0117	0.1154	2.2308	1.0099

Table 3 Parameter estimates for SSN-BS model, when $\beta = 1, \lambda = 2, v = 2$

n	α	Estimates of α			Estimates of β			Estimates of λ			Estimates of v		
		Mean	RB	$\sqrt{MSE}$	Mean	RB	$\sqrt{MSE}$	Mean	RB	$\sqrt{MSE}$	Mean	RB	$\sqrt{MSE}$
50	0.50	0.5455	0.0907	0.2283	1.0325	0.0325	0.17100	3.3249	0.6625	3.1448	5.2916	1.6458	11.2378
	1.00	1.1312	0.1312	0.4715	1.0868	0.0868	0.4125	3.5246	0.7623	3.2999	3.9504	0.9752	6.7937
100	0.50	0.5135	0.0271	0.1391	1.0219	0.0219	0.1239	2.4973	0.2486	1.6335	2.7839	0.3919	4.8319
	1.00	1.0323	0.0323	0.3043	1.0738	0.0738	0.3173	2.6611	0.3305	2.0328	2.3774	0.1887	1.7517
200	0.50	0.5077	0.0153	0.0939	1.0064	0.0064	0.0791	2.2536	0.1268	0.9806	2.1170	0.0585	0.5028
	1.00	1.0114	0.0114	0.2193	1.0401	0.0401	0.2179	2.2397	0.1198	1.0911	2.1250	0.0625	0.6821
400	0.50	0.5043	0.0085	0.0654	1.0044	0.0044	0.0543	2.1091	0.0546	0.6021	2.0571	0.0286	0.3053
	1.00	1.0021	0.0021	0.1519	1.0216	0.0216	0.1430	2.1102	0.0550	0.6983	2.0523	0.0262	0.2998
800	0.50	0.5014	0.0029	0.0465	1.0021	0.0021	0.0373	2.0499	0.0249	0.4856	2.0266	0.1330	0.2038
	1.00	0.9995	0.0005	0.1094	1.0117	.0117	0.0957	2.0439	0.0220	0.4806	2.0242	0.0121	0.2010

Fig. 5 RB values in parameter estimates when $\alpha = 0.5$, $\beta = 1$, $\lambda = 2$, $\nu = 2$

Estimation and model checking are provided in Table 5. We first obtain the ML estimation of the parameters via ECM algorithm described in Sect. 3. Then to assess the fitting performance, we use the maximized log-likelihood $\ell(\hat{\theta})$, the Akaike (AIC) and AIC with a correction ($AICc$) information criteria which defined as

$$AIC = 2m - 2\ell(\hat{\theta}),$$

$$AICc = AIC + \frac{2m(m+1)}{n-m-1},$$

where m is the number of the model parameters.

According to the AIC and $AICc$ we find that the $ST_{v_1}T_{v_2} - BS$ distribution provides the best fit.

Fig. 6 MSE root values in parameter estimates when $\alpha = 0.5$, $\beta = 1$, $\lambda = 2$, $\nu = 2$

The *PP* plots and empirical and theoretical cdf plots given in Figures 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 confirm again the appropriateness of the $ST_{v_1}T_{v_2} - BS$ distribution.

Also we use known distributions such as gamma, lognormal, weibull, exponential to compare the competition models. For the ML estimates we use *fitdistr* function in package *MASS* in the statistical software *R*. see Table 6. Again $ST_{v_1}T_{v_2} - BS$ distribution seems to be the best.

Fig. 7 RB values in parameter estimates when $\alpha = 1$, $\beta = 1$, $\lambda = 2$, $\nu = 2$

5 Concluding Remarks

In this paper, an extension of the Birnbaum-Saunders distribution, called scale shape mixture of skew normal Birnbaum–Saunders (SSMSN-BS) distribution and its subclasses, are introduced. The parameters estimation via ECM algorithm are discussed and the utility of models is shown by means of simulation and real data set. The SSMSN-BS distributions can be in modeling different types of data due to their properties and their flexibility.

Fig. 8 MSE root values in parameter estimates when $\alpha = 1, \beta = 1, \lambda = 2, \nu = 2$

Table 4 Descriptive statistics for the data

Median	Mean	SD	CV	CS	CK	Range	Min.	Max.	n
70	99.819	80.553	80.699	1.796	5.614	364	12	376	72

Table 5 The ML estimates and information criteria based on the SSMSN-BS distributions

Model	α	$\hat{\beta}$	$\hat{\lambda}$	$\hat{v}_1$	$\hat{v}_2$	$\ell(\hat{\theta})$	AIC	AIC _c
ST _{v₁} T _{v₂} -BS	0.695	51.781	42.320	4.530	0.191	386.499	782.998	783.907
ST _{v₂} T _{v₁} -BS	0.556	71.568	0.0428	2.854	2.854	392.212	792.424	793.021
STN-BS	0.557	71.994	0.0312	2.876		392.212	792.424	793.021
SNT-BS	0.760	77.535	0.200		4.548e-08	390.917	789.835	790.432
SN-BS	1.170	37.020	1.334			395.182	796.378	796.731
ST-BS	0.646	52.353	0.840	3.533		389.506	787.012	787.609
SSN-BS	0.7279	118.690	-0.686	5.614		390.427	788.853	789.450
SGLN-BS	0.704	55.000	0.538	0.816		387.807	783.614	784.211
T-BS	0.608	75.588		5		390.054	784.108	784.461
BS	0.760	77.534				390.917	785.834	786.008

Fig. 9 The PP plot (left) and empirical versus theoretical c.d.f. (right) for the BS model **Fig. 10** PP plot (left) and empirical versus theoretical c.d.f. (right) for the T-BS model

Fig. 11 *PP* plot (left) and empirical versus theoretical c.d.f. (right) for the SGLNT-BS model

Fig. 12 *PP* plot (left) and empirical versus theoretical c.d.f. (right) for the SSN-BS model

Fig. 13 *PP* plot (left) and empirical versus theoretical c.d.f. (right) for the ST-BS model

Fig. 14 *PP* plot (left) and empirical versus theoretical c.d.f. (right) for the SN-BS model

Fig. 15 PP plot (left) and empirical versus theoretical c.d.f. (right) for the SNT-BS model

Fig. 16 PP plot (left) and empirical versus theoretical c.d.f. (right) for the STN-BS model

The study of the bivariate SSMSN-BS class (see [22]), or more general, the study of the multivariate SSMSN – BS class (see [1]) are interesting problems that remain to be studied. These works are currently under progress and we hope to report our findings in a future paper.

Fig. 17 PP plot (left) and empirical versus theoretical c.d.f. (right) for the $ST_v T_v$ -BS model

Fig. 18 PP plot (left) and empirical versus theoretical c.d.f. (right) for the $ST_{v_1} T_{v_2}$ -BS model

Table 6 The ML estimates and information criteria on the four distributions

Model	Parameter	Parameter	$\ell(\hat{\theta})$	AIC	AICc
Gamma	Shape = 2.081 (0.3220)	Rate = 0.021 (0.0036)	394.248	792.495	792.669
Lognormal	Mean log = 4.344 (0.0837)	sd log = 0.710 (0.0592)	390.335	784.671	784.845
Weibull	Shape = 1.393 (0.1184)	Scale = 110.458 (9.924)	397.148	798.295	798.469
Exponential	Rate = 0.010 (0.0012)		403.442	808.884	808.941

Acknowledgements The authors would like to express their sincere thanks to the editor and referees for their encouragements, suggestions and valuable comments. They gratefully acknowledge the support of this work by Research Council of Shiraz University

Author Contributions The order of authorship is assigned based on contributions of authors. All authors read and approved the final manuscript.

Funding The authors received no financial support for the research, authorship, and/or publication of this article.

Declarations

Conflict of interest The authors declare they have no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Arellano-Valle, R.B., Ferreira, C.S., Genton, M.G.: Scale and shape mixtures of multivariate skew-normal. *J. Multivar. Anal.* **166**, 98–110 (2018)
2. Azzalini, A.: A class of distribution which includes the normal ones. *Scand. J. Stat.* **12**, 171–178 (1985)
3. Azzalini, A.: Further results on a class of distribution which includes the normal ones. *Statistica* **46**, 199–208 (1986)
4. Azzalini, A., Capitanio, A.: Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *J. R. Stat. Soc. Ser. B.* **65**, 367–389 (2003)
5. Barros, M., Paula, G.A., Leiva, V.: A new class of survival regression models with heavy tailed errors: robustness and diagnostics. *Lifetime Data Anal.* **14**, 316–332 (2008)
6. Basford, K.E., Greenway, D.R., McLachlan, G.J., Peel, D.: Standard error of fitted means under normal mixture. *Comput. Stat.* **12**, 1–17 (1997)
7. Benkhalifa, L.: The Weibull Birnbaum–Saunders distribution and its applications. *Stat. Optim. Inform. Comput.* **9**, 61–81 (2021)
8. Birnbaum, Z.W., Saunders, S.C.: A new family of life distribution. *J. Appl. Probab.* **6**, 319–327 (1969)
9. Branco, M.D., Dey, D.K.: A general class of multivariate skew-elliptical distribution. *J. Multivar. Anal.* **79**, 99–113 (2001)
10. Bjerkedal, T.: Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Am. J. Hyg.* **72**, 130–148 (1960)
11. Cabral, C.R.B., Bolfarine, H., Pereira, J.R.G.: Bayesian density estimation using skew student-t-normal mixtures. *Comput. Stat. Data Anal.* **52**, 5075–5090 (2008)
12. Cordeiro, G.M., Lemonte, A.J.: The β Birnbaum–Saunders distribution: an improved distribution for fatigue life modeling. *Comput. Stat. Data Anal.* **55**, 1445–1461 (2011)
13. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. R. Stat. Soc. Ser. B* **39**, 1–38 (1977)

14. Desmond, A.: Stochastic models of failure in random environments. *Can. J. Stat.* **13**, 171–183 (1985)
15. Diaz-Garcia, J.A., Leiva-Sanchez, V.: A new family of life distribution based on the elliptically contoured distributions. *J. Stat. Plan. Inference* **128**, 445–457 (2005)
16. Gokhale, S., Khare, M.: Statistical behavior of carbon monoxide from vehicular exhausts in urban environments. *Environ. Model. Softw.* **22**, 526–535 (2007)
17. Gomez, H.W., Olinares, J., Bolfarine, H.: An extension of the generalized Birnbaum–Saunders distribution. *Stat. Probab. Lett.* **79**, 331–338 (2009)
18. Gomez, H.W., Venegas, O., Bolfarine, H.: Skew-symmetric distributions generated by the distribution function of the normal distribution. *Environmetrics* **18**, 395–407 (2007)
19. Gomez, H.W., Olivares, J., Bolfarine, H.: An extension of the generalized Birnbaum–Saunders distribution. *Stat. Probab. Lett.* **79**, 331–338 (2009)
20. Jamlizadeh, A., Lin, T.I.: A general class of scale-shape mixtures of skew-normal distributions: properties and estimation. *Comput. Stat.* **32**, 451–474 (2017)
21. Jørgensen, S.: Statistical properties of the generalized inverse Gaussian distribution. Springer-Verlag, New York (1982)
22. Kundu, D., Balakrishnan, N., Jamalizadeh, A.: Bivariate Birnbaum–Saunders distribution and associate inference. *J. Multivar. Anal.* **101**, 113–125 (2010)
23. Kundu, D., Kannan, N., Balakrishnan, N.: On the hazard function of Birnbaum–Saunders distribution and associate inference. *Comput. Stat. Data Anal.* **52**, 2692–2702 (2008)
24. Hassani, H., Kalantari, M., Entezarian, M.R.: A new five-parameter Birnbaum–Saunders distribution for modeling bicoid gene expression data. *Math. Biosci.* **319**, 108–275 (2020)
25. Hashemi, F., Amirzadeh, V., Jamalizadeh, A.: An extension of the Birnbaum–Saunders distribution based on skew-normal-t distribution. *Stat. Res. Train. Cent.* **12**, 1–37 (2015)
26. Leiva, V., Barros, M., Paula, G.A., Galea, M.: Influence diagnostics in log-Birnbaum–Saunders regression models with censored data. *Comput. Stat. Data Anal.* **51**, 5694–5707 (2007)
27. Leiva, V., Riquelme, M., Balakrishnan, N., Sanhueza, A.: Lifetime analysis based on the generalized Birnbaum–Saunders. *Comput. Stat. Data Anal.* **52**, 2079–2097 (2008)
28. Leiva, V., Vilca, F., Balakrishnan, N., Sanhueza, A.: A skewed Sinh-normal distribution and its properties and application to air pollution. *Commun. Stat. Theory Methods* **39**, 426–443 (2010)
29. Mahbudi, S., Jamalizadeh, A., Farnoosh, R.: Use of finite mixture models with skew-t-normal Birnbaum–Saunders components in the analysis of wind speed: case studies in Ontario, Canada. *Renew. Energy* **162**, 196–211 (2020)
30. Meng, X.-L., Rubin, D.B.: Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika* **80**, 267–278 (1993)
31. Miner, M.A.: Cumulative damage in fatigue. *J. Appl. Mech.* **12**, 159–164 (1945)
32. Podaski, R.: Characterization of diameter data in near-natural forests using the Birnbaum–Saunders distribution. *Can. J. For. Res.* **38**, 518–527 (2008)
33. Poursadeghfard, T., Jamalizadeh, A., Nematollahi, A.: On the extended Birnbaum–Saunders distribution based on the skew-t-normal distribution. *Iran. J. Sci. Technol. Trans. A Sci.* **43**, 1689–1703 (2018)
34. Reyes, J., Barranco-Chamorro, I., Gallardo, D., Gomez, H.: Generalized modified slash Birnbaum–Saunders distribution. *Symmetry* **10**, 724–742 (2018)
35. Rogers, W.H., Tukey, J.W.: Understanding some long-tailed symmetrical distribution. *Stat. Neerl.* **26**, 211–226 (1972)
36. Sanhueza, A., Leiva, V., Balakrishnan, N.: The generalized Birnbaum–Saunders distribution and its theory, methodology and application. *Commun. Stat. Theory Methods* **37**, 645–670 (2008)
37. Vilca, F., Sanhueza, A., Leiva, V., Christakos, G.: An extended Birnbaum–Saunders model ad it application in the study of environmental quality quality in Santiago, Chile. *Stoch. Environ. Res. Risk Assess.* **24**, 771–782 (2010)
38. Vilca, F., Santana, L., Leiva, V., Balakrishnan, N.: Estimation of extreme percentiles in Birnbaum–Saunders distribution. *Comput. Stat. Data Anal.* **55**, 1665–1678 (2011)
39. Wang, J., Genton, M.G.: The multivariate skew-slash distribution. *J. Stat. Plan. Inference* **136**, 209–220 (2006)