Journal of Statistical Distributions and Applications
a SpringerOpen Journal

REVIEW

Open Access

CrossMark

Revisit of relationships and models for the Birnbaum-Saunders and inverse-Gaussian distributions

William Jason Owen[1][^] and Hon Keung Tony Ng[2][*]

*Correspondence:
ngh@mail.smu.edu
[^]Deceased
[2]Department of Statistical Science,
Southern Methodist University,
Dallas, TX, 75275 USA
Full list of author information is
available at the end of the article

Abstract

The Birnbaum-Saunders distribution was derived in 1969 as a lifetime model for a specimen subjected to cyclic patterns of stresses and strains, and the ultimate failure of the specimen is assumed to be due to the growth of a dominant crack in the material. The inverse Gaussian distribution is used to describe the first passage time for a particle (moving with constant velocity) that is subject to linear Brownian motion. These two models have a rich history, and they have been shown to be much related. In this article, these two models will be reviewed and comparisons will be made. Specifically, two moment-ratio diagrams will be presented that gives insight to the reason of both distributions often achieve similar fits to experimental data. Next, a generalized Birnbaum-Saunders distribution, will be presented and several properties will be derived. In particular, it will be shown that this generalized model can be expressed as a mixture of inverse Gaussian-type random variables (similar to the two-parameter Birnbaum-Saunders model). Estimation of the parameters in the generalized Birnbaum-Saunders distribution will be discussed. Lastly, some conclusions from this investigation are presented.

Keywords: Bimodality, Cycles to failure, Fatigue, Generalized Birnbaum-Saunders distribution, Moment-ratio diagram

1 Introduction

Birnbaum and Saunders (1969a) derived a lifetime distribution, known as the Birnbaum-Saunders distribution (abbreviated as B-S hereafter), which is founded on modeling the failure of a specimen that is subjected to a cyclic pattern stresses and strains. The ultimate failure is due to the growth of a dominant crack in the material. At each increment of load, this dominant crack extends by a random, non-negative amount. Another model, the inverse Gaussian distribution (abbreviated by I-G hereafter), was originally used to describe the first passage time in a Brownian motion. However, the inverse Gaussian distribution has been used quite frequently to model reliability data, and the relationship between the B-S and the I-G distributions was first established by Bhattacharyya and Fries (1982). Starting in Section 2, these two distributions will be presented, and the main results, extensions, and summaries from an extensive literature review will be provided. Following that section, the main purpose of this paper is two-fold and is introduced in the two following sections. In Section 3, comparisons of these two models will be made

by providing two moment-ratio diagrams for both the B-S and I-G distributions. The first diagram is a graph of the coefficient of variation (CV) versus skewness and the second diagram is a graph of the skewness versus the kurtosis. While functionally these standardized moments are very different, when they are graphed they prove to be very similar. This will further support the claims that the B-S and I-G distributions are nearly interchangeable. Following this, a three-parameter extension to the B-S distribution proposed by Díaz-García and Domínguez-Molina (2006) is presented in Section 4. Some very unique properties and summary values will be presented, and it is shown that the distribution has relationship to other models found in the literature. In addition, this model has the ability to exhibit a bimodal shape for certain ranges of parameter values. We also discuss different estimation methods for the parameters in the three-parameter B-S distribution and study their performances in Section 4. Two numerical examples are given in Section 5 to illustrate the usefulness of the generalized distribution and the estimation methods. Lastly, Section 6 presents some conclusions from this investigation.

2   The two-parameter B-S and I-G distributions

The precise derivation of the B-S model is given in Birnbaum and Saunders (1969a) and also summarized in Owen (2006). The cumulative distribution function (CDF) for a random variable S that follows the B-S distribution is given by

$$F(s) = \Pr(S \le s) = \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{s}{\beta}} - \sqrt{\frac{\beta}{s}}\right)\right], \;\; s > 0, \tag{1}$$

where $\alpha > 0$, $\beta > 0$, and Φ represents the standard normal CDF. To use shorthand, we describe the random variable S that has distribution function (1) as $S \sim$ B-S(α, β). The distribution (1) has some interesting properties given in Johnson et al. (1995b). The parameter β is a scale parameter since by definition $S/\beta \sim$ B-S$(\alpha, 1)$. In addition, β is the median for the distribution as from (1),

$$F(\beta) = \Pr(S \le \beta) = \Phi(0) = 0.5.$$

The parameter α is a shape parameter. The B-S distribution exhibits the well-known reciprocal property: the random variable $S^{-1} \sim$ B-S$(\alpha, 1/\beta)$, so it is in the same family of distributions (see Saunders 1974).

If Z is a standard normal random variable, then

$$S = \frac{\beta}{2}\left(2 + \alpha^2 Z^2 + \alpha Z\sqrt{\alpha^2 Z^2 + 4}\right) \tag{2}$$

has the B-S distribution in (1). This presentation is also given in Owen (2006) but it is slightly different from Birnbaum and Saunders (1969a). In addition, the expression in (2) is useful for random number generation and for deriving integer moments. The mean and variance for S are

$$E(S) = \beta\left(1 + \frac{\alpha^2}{2}\right) \tag{3}$$

$$\text{and } Var(S) = (\alpha\beta)^2\left(1 + \frac{5\alpha^2}{2}\right), \tag{4}$$

respectively. Higher order moments will come into play in the subsequent sections. Since these higher order moments cannot be found in the literature, they are provided here. The third and fourth moments (both about zero) are given by

$$E(S^3) = \beta^3 \left(1 + \frac{9\alpha^2}{2} + 9\alpha^4 + \frac{15\alpha^6}{2}\right) \tag{5}$$

$$\text{and } E(S^4) = \beta^4 \left(1 + 8\alpha^2 + 30\alpha^4 + 60\alpha^6 + \frac{105\alpha^8}{2}\right), \tag{6}$$

respectively. These can be expressed by using the relationship (2) above. The expected value for a standard normal variable that is raised to an integer power is given in Zacks (1992).

The probability density function (PDF) of the random variable S following the B-S distribution in (1) is given by

$$
\begin{aligned}
f(s) = F'(s) = {} & \frac{1}{2\alpha\beta}\sqrt{\frac{\beta}{s}}\left(1 + \frac{\beta}{s}\right) \\
& \times \frac{1}{\sqrt{2\pi}}\exp\left[-\frac{1}{2\alpha^2}\left(\frac{s}{\beta} - 2 + \frac{\beta}{s}\right)\right], \quad s > 0,
\end{aligned} \tag{7}
$$

where $\alpha > 0$, $\beta > 0$ and it is characteristically right-skewed when graphed. As α decreases, particularly for values less than unity, the density becomes nearly symmetric as the curve spread (variance) decreases. This two-parameter family of distributions has various applications in reliability and life testing. For example, Birnbaum and Saunders (1969b) consider an example to model the strength of aluminum coupons subjected to cyclic stresses and strains. Three data sets are considered, each having different levels of maximum stress per cycle. The parameters in (1) are estimated using maximum likelihood (ML) estimation. The likelihood equations are given in Birnbaum and Saunders (1969b) and the ML estimates for α and β need to be found using numerical techniques. Dupuis and Mills (1998) consider an alternative to ML estimation for the B-S distribution that is shown to be robust to the presence of contaminated data. Ng et al. (2006) considered modified moment techniques for estimation and bias-reduction methods for estimation of the parameters in the B-S model. Ashcar (1993) derived the approximate Fisher information matrix for the parameters α and β and considered Bayesian inference using non-informative and Jeffrey's priors. Other important efforts include a log-linear model for the B-S distribution, which was derived by Rieck and Nedleman (1991).

In the past decade, considerable research has been dedicated to the generalizations and applications of the B-S distribution. Numerous authors have investigated different aspects related to the B-S distribution. For instance, various generalizations and extensions of the B-S distribution are proposed and discussed; see, for example, Díaz-García and Leiva-Sánchez (2005), Díaz-García and Domínguez-Molina (2006), Owen (2006), Leiva et al. (2008), Sanhueza et al. (2008), Gómez et al. (2009), Guiraud et al. (2009), Leiva et al. (2009), Castillo et al. (2011), Cordeiro and Lemonte (2011), Genç (2013), Cordeiro and Lemonte (2014) and Cordeiro, et al.: A model with long-term survivors: Negative binomial Birnbaum-Saunders (to appear). These generalized B-S distributions have been applied to model data obtained from different disciplines including environmental sciences, reliability engineering and biomedical studies. For example, Leiva et al. (2008) used generalized B-S distributions to model air pollutant concentration, Leiva et al. (2009) proposed a length-biased B-S distribution for water quality data, Guiraud et al. (2009) used a non-central B-S distribution for reliability analysis and Cordeiro, et al.: A model with long-term survivors: Negative binomial Birnbaum-Saunders (to appear) proposed the negative

binomial B-S distribution as a cure rate survival model. Software packages written in R (R Core Team 2015) for B-S and generalized B-S distributions have been developed by Leiva et al. (2006) and Barros et al. (2009).

In particular, Owen (2006) derived a three-parameter model based on the B-S model in (7), and this was from a consequence of relaxing the assumption of independent crack extensions and viewing critical crack growth as a long-memory process. The three-parameter Birnbaum-Saunders distribution derived in Owen (2006) has PDF

$$f_1(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{a\sqrt{bt^\kappa}} \left(1 - \kappa + \frac{b\kappa}{t}\right) \exp\left[-\frac{(t-b)^2}{2a^2bt^{2\kappa}}\right], \quad t > 0, \tag{8}$$

where $a > 0, b > 0$ and $\kappa > 0$ are the parameters. To avoid confusion in a literature review, this model in (8) is referred to as the GB-S$_1$.

The I-G distribution, another important distribution in reliability analysis, has PDF

$$p(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right], \quad x > 0, \tag{9}$$

where $\mu > 0$ and $\lambda > 0$ are the parameters. We denote a random variable X having distribution (9) as $X \sim$ I-G(μ, λ). Both μ and λ behave as shape-type parameters. When graphed, (9) exhibits a right-skewed shape. Interested readers can refer to Johnson et al. (1995a) for a discussion on the I-G distribution and for more information, see Chhikara and Folks (1989). Of interest here is the scale-change property: if $X \sim$ I-G(μ, λ), then $aX \sim$ I-G$(a\mu, a\lambda)$ for a constant $a > 0$. In addition, the mean, variance, third and fourth moments (again, both about zero) are given by

$$E(X) = \mu, \tag{10}$$

$$Var(X) = \frac{\mu^3}{\lambda}, \tag{11}$$

$$E(X^3) = \mu^3 + \frac{3\mu^4}{\lambda} + \frac{3\mu^5}{\lambda^2}, \tag{12}$$

$$E(X^4) = \mu^4 + \frac{6\mu^5}{\lambda} + \frac{15\mu^6}{\lambda^2} + \frac{15\mu^7}{\lambda^3}, \tag{13}$$

respectively. The I-G distribution is a two-parameter exponential family (see Lehmann and Casella 1998), and this fact has often made the I-G distribution more attractive for the development of exact statistical procedures.

3 Relationships between and comparisons of the B-S and I-G models

Of utmost importance here are the work published in two papers that relate the B-S and I-G models. In this section, relationships between these two models will be established.

3.1 Literature review

In Bhattacharyya and Fries (1982), the identical approach in Birnbaum and Saunders (1969a) was viewed as a Wiener process of accumulated fatigue in time (with positive drift parameter μ and diffusion constant τ^2), and this leads directly to an I-G distribution. From this derivation, it is shown that a B-S random variable or "event" is actually contained within the I-G model since only positive increments in the growth of the dominant crack are allowed in Birnbaum and Saunders (1969a) (a "negative growth" can be viewed as a repair in the dominant crack, but this event is rare whenever $\mu \gg \tau$). Therefore, the B-S distribution can be considered as an *approximation* of the I-G distribution, but Bhattacharyya and Fries (1982) argued that this approximation is not necessary since

the I-G distribution benefits from its exponential family structure. Still, the observation in Bhattacharyya and Fries (1982) has elucidated the observation that (7) and (9) can give similar shapes when fitting observed failure data. Following this result, Desmond (1986) demonstrated that the B-S distribution can be written as an equal mixture of an I-G distribution and the distribution of the reciprocal of an I-G random variable. That is, for the mutually statistically independent random variables X_1, X_2, and B, where $X_1 \sim$ I-G$(\beta, \beta/\alpha^2)$, $X_2^{-1} \sim$ I-G$(1/\beta, 1/(\alpha^2\beta))$, and B follows Bernoulli distribution with $\Pr(B = 0) = \Pr(B = 1) = 0.5$. Then,

$$BX_1 + (1 - B)X_2 = S \sim \text{B-S}(\alpha, \beta) \qquad (14)$$

Desmond (1986) argued that in a stochastic modeling point of view, the B-S distribution is often more appropriate but the I-G is preferred for statistical analysis (see, for example, Durham and Padgett 1997, Owen 2007). In addition, Desmond (1986) observed that the hazard functions for the B-S and I-G distributions are very similar, giving further evidence that the two models are nearly identical.

3.2 Alternative approach to relate the B-S and I-G distributions

Construction of moment ratio diagrams has provided alternative ways for comparing univariate distributions and illustrating their differences. By selecting two standardized moments for a host of distributions, these can be graphed and used to not only compare and contrast distribution qualities but also the diagrams give a means for selecting potential models based on sample data. The "standardized moments" that are being referred to are the coefficient of variation (CV) γ_2, coefficient of skewness (or third standardized moment) γ_3, and the kurtosis (or fourth standardized moment) γ_4. Following Cox and Oates (1984) (see also Johnson et al. 1995a; 1995b), the moment ratio diagrams are given by:

(I)   plotting γ_2 on the horizontal axis (abscissa) and γ_3 on the vertical axis (ordinate)
(II) plotting γ_3 on the horizontal axis and γ_4 on the vertical axis (the classical presentation of this graph is given upside down)

To identify a potential distribution to consider when modeling a dataset, sample standardized moments can be calculated and plotted as points in either (I) and/or (II) – thus, probability distributions that are "close" to the point estimates can be considered as candidates for probability models. A recent article by Vargo et al. (2010) revisited the moment ratio diagrams and presented comprehensive graphs of (I) and (II) above for over 30 commonly used univariate distributions. Limiting relationships between several well-known families (e.g., t and chi-square) were also considered. In addition, to identify potential distributions to model a dataset, the authors presented a novel application of bootstrapping. Therein, bootstrap samples are generated and the sample estimates of CV, skewness, and kurtosis (represented as $\hat{\gamma}_2$, $\hat{\gamma}_3$ and $\hat{\gamma}_4$, respectively) are calculated in order to generate the "concentration ellipse" to graph on the moment ratio diagrams. In this way, distributions that are close to the concentration ellipse should be considered as candidates. Since point estimates of higher moments can be highly variable, this bootstrap approach includes the sampling error.

Several distributions are presented in Vargo et al. (2010), but the B-S and I-G distributions are absent while the Wald distribution is included. The Wald distribution is an

ambiguous model; there are some references that state that the Wald distribution is identical to the I-G distribution, but other references claim that the Wald distribution is a special case of the I-G distribution with $\mu = 1$ (see, page 262 of Johnson et al. 1995a). Therefore, in this section the two moment-ratio diagrams (I) and (II) will be developed for both the two-parameter I-G and B-S distributions.

For a (general) random variable Y with mean μ and standard deviation σ, the three standardized moments (and their relationship to moments taken about zero) are given by

$$\gamma_2 = \frac{\sigma}{\mu}, \tag{15}$$

$$\gamma_3 = E\left[\left(\frac{Y-\mu}{\sigma}\right)^3\right] = \frac{E(Y^3) + 2\mu^3 - 3\mu(\mu^2 + \sigma^2)}{\sigma^3}, \tag{16}$$

$$\gamma_4 = E\left[\left(\frac{Y-\mu}{\sigma}\right)^4\right] = \frac{E(Y^4) - 4\mu E(Y^3) + 6\mu^2\sigma^2 - 3\mu^4}{\sigma^4}. \tag{17}$$

Taking these expressions along with the moments for B-S distribution presented in (3)–(6) and the moments for I-G distribution presented in (10)–(13), the results are summarized in Table 1.

The results in Table 1 correct the mistakes in the coefficient of skewness and kurtosis that appeared in the literature (see, for example, Ng et al. 2006, Balakrishnan et al. 2011, Lemonte and Ferrari 2011).

There are two remarks from Vargo et al. (2010) that are applicable here: (i) since β is a scale parameter in the B-S model, all of the standardized moments are free of β; and (ii) since μ and λ are both shape parameters in the I-G model, they both appear in the standardized moments. However, since all of the standardized moments are common functions of the quotient μ/λ, this will simplify the display of the moment-diagrams. Note that in Vargo et al. (2010), it is most typical that the plotted points seen in moment-ratio diagrams lie not far from the origin. Figures 1 and 2 are the moment ratio diagrams (I) and (II), respectively, for the B-S and I-G distributions. The similarity of the shapes provides more evidence that the two distributions are quite similar and even in many cases comparable.
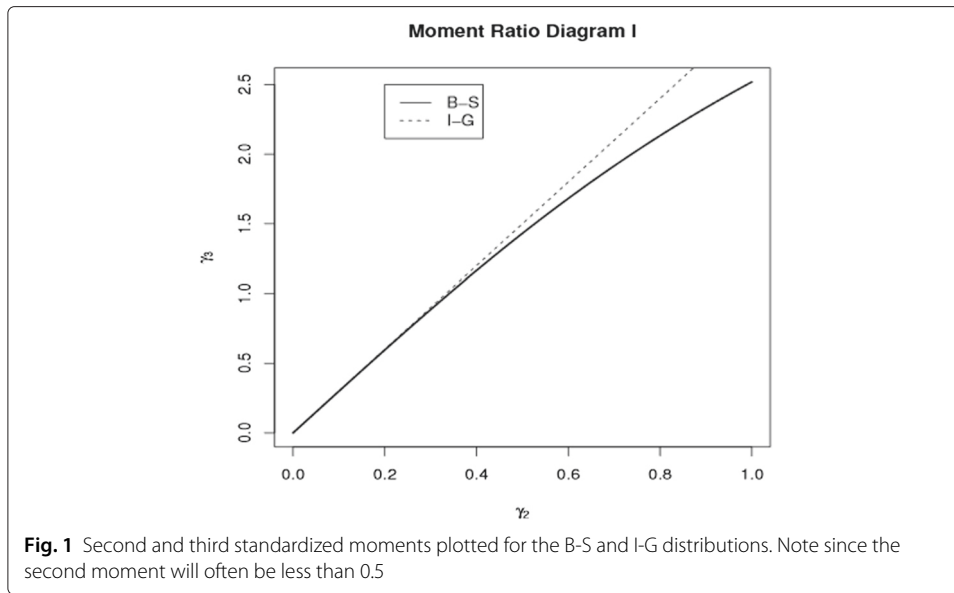
4 A generalized B-S distribution

The B-S distribution in (1) can be made more general by allowing the exponent (presently set to the value 1/2) to take on other values. In so doing, it will be shown that this extension exhibits some very interesting properties while still retaining a functional relationship to the I-G. We define the generalized Birnbaum-Saunders distribution by its CDF representation given by

$$F_2(t) = \Phi\left\{\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^\nu - \left(\frac{\beta}{t}\right)^\nu\right]\right\}, \tag{18}$$

Table 1 Coefficient of variation γ_2, coefficient of skewness γ_3, and kurtosis γ_4 for B-S and I-G distributions

Distribution	γ_2	γ_3	γ_4
B-S(α, β)	$\frac{\alpha\sqrt{4+5\alpha^2}}{2+\alpha^2}$	$\frac{4\alpha(6+11\alpha^2)}{(4+5\alpha^2)^{3/2}}$	$3 + \frac{6\alpha^2(40+93\alpha^2)}{(4+5\alpha^2)^2}$
I-G(μ, λ)	$\sqrt{\frac{\mu}{\lambda}}$	$3\sqrt{\frac{\mu}{\lambda}}$	$3 + 15\left(\frac{\mu}{\lambda}\right)$

Fig. 1 Second and third standardized moments plotted for the B-S and I-G distributions. Note since the second moment will often be less than 0.5

where $\alpha > 0$, $\beta > 0$, and $\nu > 0$ make this a valid CDF. This is refereed to as the "second" generalized B-S distribution as to not confuse with another, but very different, unrelated generalized B-S distribution in Owen (2006). We denote this generalized B-S distribution as GB-S$_2$. This generalization of B-S distribution was proposed by Díaz-García and Domínguez-Molina (2006) (see also Eq. (B9.2) of Sanhueza et al. 2008). This model includes (1) as a special case when $\nu = 0.5$. Here, we describe a random variable T that has distribution function (18) as $T \sim$ GB-S$_2(\alpha, \beta, \nu)$. The PDF for the GB-S$_2(\alpha, \beta, \nu)$ distribution is given by

$$f_2(t) = \frac{\nu}{t\alpha\sqrt{2\pi}} \left[\left(\frac{t}{\beta}\right)^\nu + \left(\frac{\beta}{t}\right)^\nu \right]$$
$$\times \exp\left\{ -\frac{1}{2\alpha^2} \left[\left(\frac{t}{\beta}\right)^{2\nu} + \left(\frac{\beta}{t}\right)^{2\nu} - 2 \right] \right\}, t > 0. \quad (19)$$

Fig. 2 Third and fourth standardized moments plotted for the B-S and I-G distributions
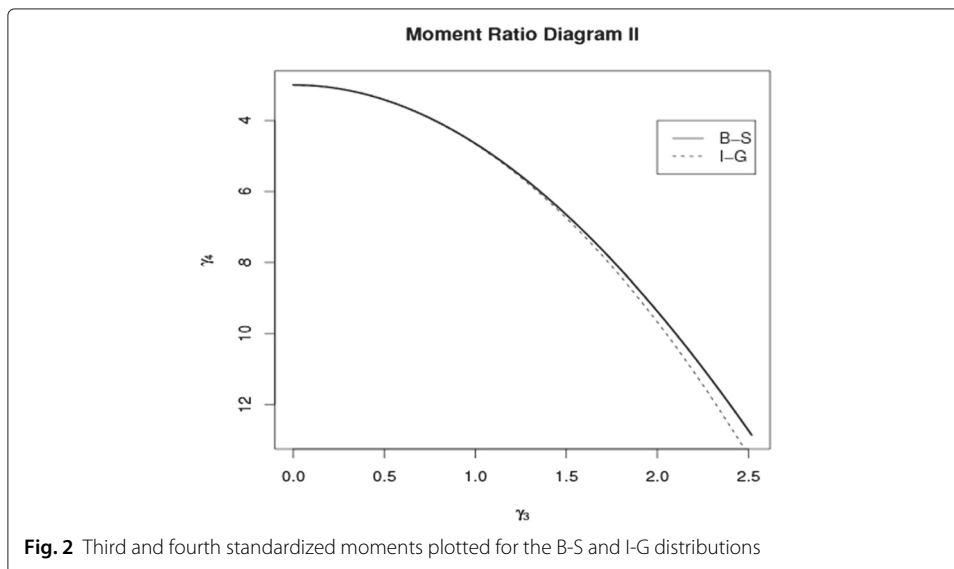
This generalized Birnbaum-Saunders distribution is a flexible three-parameter family of distributions, and the shape of the density varies widely with different values of the parameters. Interestingly, the density is *bimodal* if both $\alpha > 2$ and $\nu > 2$, and the major mode is always less than the minor mode. In Figs. 3 and 4, we present various graphs of the PDF in (19); since β is a scale parameter, without loss of generality, it is fixed at unity for all cases. For the reader's interest, we also provide the values of the mean and standard deviation for each distribution that were calculated using (24).

As it can be seen, the densities in Fig. 4 are interesting since that while the medians are equal the means and standard deviations are quite different. Lastly, the ability for (19) to achieve a bimodal shape truly expands the flexibility for the model; often, when dealing with a dataset with two modes a mixture model is the standard approach (see, for example, Chen et al. 2008).

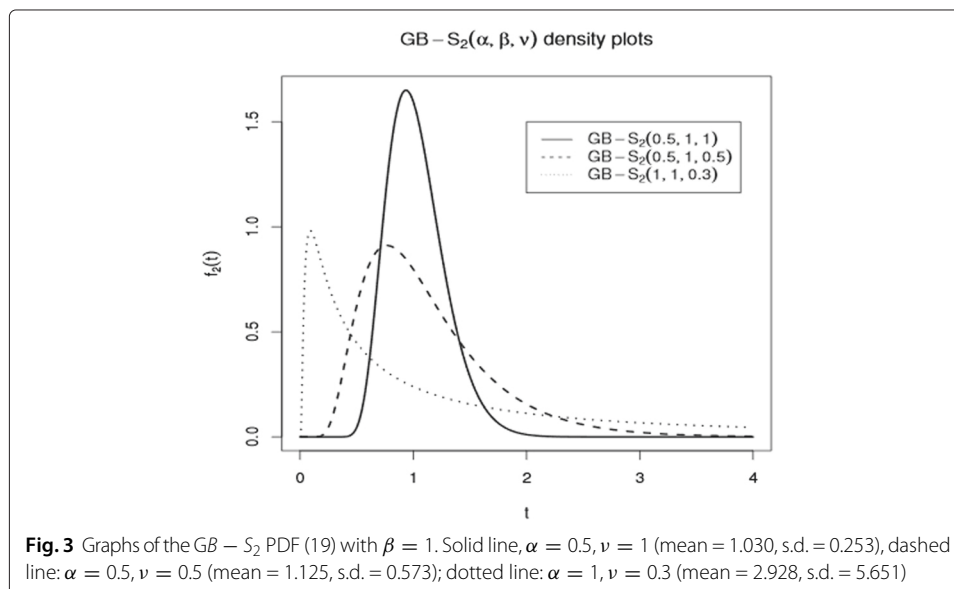### 4.1 Properties and related distributions

The distribution in (18) has some very interesting properties, and many similar properties of the two-parameter model (1) still hold for (18). Namely, β remains a scale parameter and is also the median for the distribution. Both α and ν are shape-type parameters. If Z is a standard normal variable, then the random variable

$$T = \beta \left[\frac{\alpha Z + \sqrt{\alpha^2 Z^2 + 4}}{2} \right]^{1/\nu} \tag{20}$$

follows GB-S$_2(\alpha, \beta, \nu)$. Thus, random variates from this generalized distribution can easily be simulated. Clearly, the reciprocal property is also preserved since a trivial calculation shows that

$$T^{-1} \sim \text{GB-S}_2(\alpha, 1/\beta, \nu).$$

In fact, (18) generally describes the family of distributions of power transformations for random variables distributed as (1); it is straightforward to show that if $S \sim \text{B-S}(\alpha, \beta)$,

Fig. 3 Graphs of the G$B - S_2$ PDF (19) with $\beta = 1$. Solid line, $\alpha = 0.5$, $\nu = 1$ (mean = 1.030, s.d. = 0.253), dashed line: $\alpha = 0.5$, $\nu = 0.5$ (mean = 1.125, s.d. = 0.573); dotted line: $\alpha = 1$, $\nu = 0.3$ (mean = 2.928, s.d. = 5.651)

Fig. 4 Graphs of the $GB - S_2$ PDF (19) with $\beta = 1$. Solid line: $\alpha = 8$, $\nu = 6$ (mean = 1.046, s.d. = 0.309), dashed line: $\alpha = 8$, $\nu = 2.5$ (mean = 1.277, s.d. = 1.038), dotted line: $\alpha = 0.25$, $\nu = 0.25$ (mean = 1.131, s.d. = 0.595)
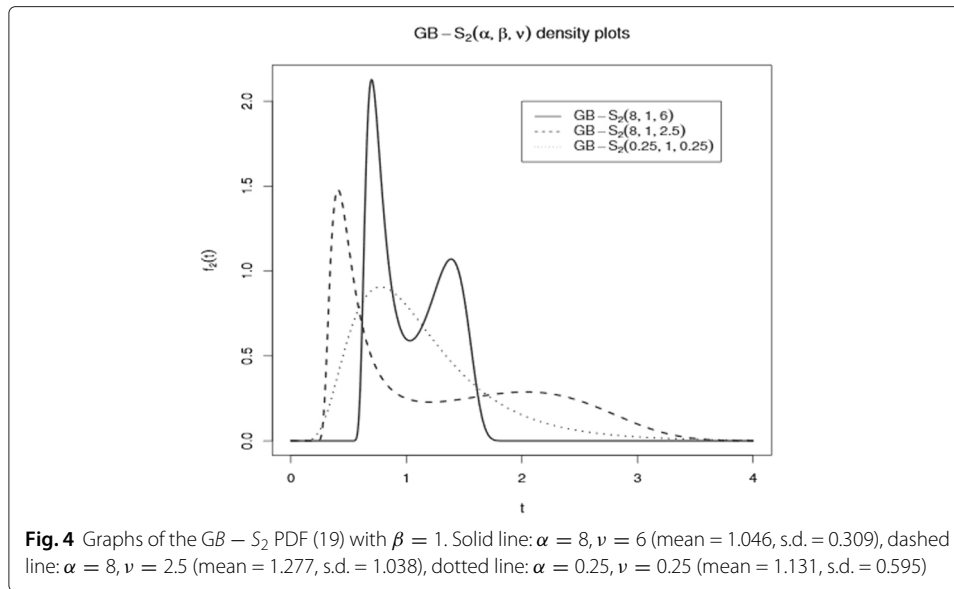
then for any nonzero real-valued constant k, $S^k \sim \text{GB-S}_2(\alpha, \beta^k, 0.5|k|^{-1})$, where $|\cdot|$ represents the absolute value function. In addition, when k is a positive integer, the GB-S_2 variable S^k can be expressed as a mixture of I-G-type random variables. Following the work of Desmond (1986) as described in Section 3.1, the binomial theorem can be applied. Since all cross-product terms are zero, we can obtain the following relationship:

$$S^k = [BX_1 + (1 - B)X_2]^k = BX_1^k + (1 - B)X_2^k, \tag{21}$$

where X_1, X_2, and B are as defined in Section 3.1. In (21), the random variables X_1 and X_2 are raised to the power k, which are called power I-G distribution. One may refer to Hossain et al. (1997) for a description of the power I-G distribution as well as properties of mixtures of I-G random variables.

Closed-form expressions of moments for (18) do not exist for all values of the parameters but they can be evaluated using the following relationship. If $T \sim \text{GB-S}_2(\alpha, \beta, \nu)$, then it can be shown that the transformation $U = \ln T$ has the following distribution function

$$G(u) = \Phi\left[\frac{2}{\alpha}\sinh\left(\frac{u - \delta}{\eta}\right)\right], \quad -\infty < u < \infty, \tag{22}$$

where $\delta = \ln\beta$ and $\eta = 1/\nu$. The distribution (22) is the three-parameter form of the sinh-normal (here, abbreviated by S-N) distribution (see, for example, Johnson et al. 1995b), and we denote a random variable U following distribution (22) as $U \sim$ S-N(α, δ, η). The parameter $\alpha > 0$ is a shape parameter, $-\infty < \delta < \infty$ is a location parameter (also the mean of the distribution), and $\eta > 0$ is a scale parameter. The PDF of the S-N distribution is always symmetric about δ and it is mound shaped if $\alpha > 2$. For values of $\alpha > 2$, the PDF is bimodal. This distribution is also referred to as the "central" S-N distribution. For a detail description of the S-N distribution and a normal approximation to the S-N distribution for small values of α, see Rieck (1999).

The relationship between the two-parameter B-S distribution and a two-parameter S-N distribution was established in Rieck and Nedleman (1991): if $S \sim$ B-S(α, β), then $U = \ln S \sim$ S-N$(\alpha, \ln\beta, 2)$. Later, Rieck (1999) obtained expressions for integer and fractional

moments for distribution (22) using the moment generating function (MGF) of the two-parameter S-N distribution via the relation

$$M_U(r) = E[\exp(Ur)] = E(S^r).$$

This method is applicable to the GB-S$_2$ distribution as well. Following Rieck (1999), the MGF for the three-parameter S-N distribution (21) is given by

$$M_U(r) = \exp(\delta r) \frac{K_{(\eta r+1)/2}(\alpha^{-2}) + K_{(\eta r-1)/2}(\alpha^{-2})}{2K_{1/2}(\alpha^{-2})} \tag{23}$$

where r is any real number and $K_\omega(z)$ is a modified Bessel function of the third kind of order ω (Watson 1995). This function $K_\omega(z)$ can be expressed in an integral form as

$$K_\omega(z) = \frac{1}{2} \int_{-\infty}^{\infty} \exp[-z\cosh(t) - \omega t]\, dt.$$

Numerous software packages (e.g., R, S-PLUS, EXCEL) can evaluate $K_\omega(z)$ for specific parameter values of ω and z. Of interest here is the result from Rieck (1999) that the denominator in (23) reduces to $\alpha\sqrt{2\pi}\exp(-\alpha^{-2})$. Using this result, we obtain the expression

$$E(T^r) = \beta^r \frac{\exp(\alpha^{-2})}{\alpha\sqrt{2\pi}} \left[K_{(r/\nu+1)/2}(\alpha^{-2}) + K_{(r/\nu-1)/2}(\alpha^{-2}) \right]. \tag{24}$$

Thus, (24) can be used to calculate the mean and variance for the GB-S$_2(\alpha, \beta, \nu)$ distribution analogous to the formulae provided in (3) and (4) for the B-S model.

4.2 Estimation of parameters

Suppose $T_1, T_2, \ldots, T_n$ are independent and identically distributed random variables from the GB-S$_2$ distribution with CDF in (18) and PDF in (19). We denote the observed values of $T_1, T_2, \ldots, T_n$ by $t_1, t_2, \ldots, t_n$. The ML estimation method can be used here. The likelihood function can be expressed as

$$L(\alpha, \beta, \nu) \propto \frac{\nu^n}{\alpha^n (2\pi)^{n/2}} \prod_{i=1}^{n} \frac{1}{t_i} \left[\left(\frac{t_i}{\beta}\right)^\nu + \left(\frac{\beta}{t_i}\right)^\nu \right]$$

$$\times \exp\left\{ -\frac{1}{2\alpha^2} \left[\left(\frac{t_i}{\beta}\right)^{2\nu} - 2 + \left(\frac{\beta}{t_i}\right)^{2\nu} \right] \right\}.$$

Thus, the log-likelihood function is

$$\ln L(\alpha, \beta, \nu) = constant + n \ln \nu - n \ln \alpha$$

$$+ \sum_{i=1}^{n} \ln \left[\left(\frac{t_i}{\beta}\right)^\nu + \left(\frac{\beta}{t_i}\right)^\nu \right]$$

$$- \frac{1}{2\alpha^2} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta}\right)^{2\nu} - 2 + \left(\frac{\beta}{t_i}\right)^{2\nu} \right].$$

Taking derivatives of the log-likelihood function with respect to the parameters and setting these derivatives to zeros, we have the likelihood equations (see Appendix):

$$\alpha = \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta}\right)^{2\nu} - 2 + \left(\frac{\beta}{t_i}\right)^{2\nu} \right] \right\}^{1/2}, \tag{25}$$

$$\sum_{i=1}^{n} \frac{\left(\frac{t_i}{\beta}\right)^{\nu} - \left(\frac{\beta}{t_i}\right)^{\nu}}{\left(\frac{t_i}{\beta}\right)^{\nu} + \left(\frac{\beta}{t_i}\right)^{\nu}} = n \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta}\right)^{2\nu} - \left(\frac{\beta}{t_i}\right)^{2\nu} \right]$$

$$\times \left\{ \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta}\right)^{2\nu} - 2 + \left(\frac{\beta}{t_i}\right)^{2\nu} \right] \right\}^{-1}, \qquad (26)$$

$$\sum_{i=1}^{n} \frac{\left[\ln\left(\frac{t_i}{\beta}\right)\right]\left[\left(\frac{t_i}{\beta}\right)^{\nu} - \left(\frac{\beta}{t_i}\right)^{\nu}\right]}{\left(\frac{t_i}{\beta}\right)^{\nu} + \left(\frac{\beta}{t_i}\right)^{\nu}} = -\frac{n}{\nu} + n \left\{ \sum_{i=1}^{n} \left[\ln\left(\frac{t_i}{\beta}\right)\right]\left[\left(\frac{t_i}{\beta}\right)^{2\nu} - \left(\frac{\beta}{t_i}\right)^{2\nu}\right] \right\}$$

$$\times \left\{ \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta}\right)^{2\nu} - 2 + \left(\frac{\beta}{t_i}\right)^{2\nu} \right] \right\}^{-1}. \qquad (27)$$

The maximum likelihood estimators (MLEs) of β and ν, denoted by $\hat{\beta}$ and $\hat{\nu}$, respectively, can be obtained by solving Eqs. (26) and (27) simultaneously. Numerical methods for two-dimensional optimizations can be employed here. Then, the MLEs of α, $\hat{\alpha}$, can be obtained by Eq. (25). Then, the observed Fisher information matrix is given by

$$\mathbf{I}(\hat{\alpha}, \hat{\beta}, \hat{\nu}) = \begin{bmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{bmatrix}$$

$$= \begin{bmatrix} -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \alpha^2} & -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \alpha \partial \beta} & -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \alpha \partial \nu} \\ -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \alpha \partial \beta} & -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \beta^2} & -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \beta \partial \nu} \\ -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \alpha \partial \nu} & -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \beta \partial \nu} & -\frac{\partial \ln L(\alpha,\beta,\nu)}{\partial \nu^2} \end{bmatrix} \Bigg|_{(\alpha,\beta,\nu)=(\hat{\alpha},\hat{\beta},\hat{\nu})}$$

where

$$I_{11} = \frac{2n}{\hat{\alpha}^2},$$

$$I_{12} = I_{21} = \frac{2\hat{\nu}}{\hat{\beta}\hat{\alpha}^3} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\hat{\beta}}\right)^{2\hat{\nu}} - \left(\frac{\hat{\beta}}{t_i}\right)^{2\hat{\nu}} \right],$$

$$I_{13} = I_{31} = -\frac{2}{\hat{\alpha}^3} \sum_{i=1}^{n} \left[\ln\left(\frac{t_i}{\hat{\beta}}\right)\right]\left[\left(\frac{t_i}{\hat{\beta}}\right)^{2\hat{\nu}} - \left(\frac{\hat{\beta}}{t_i}\right)^{2\hat{\nu}} \right],$$

$$I_{22} = -\frac{4\hat{\nu}^2}{\hat{\beta}^2} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\hat{\beta}}\right)^{\hat{\nu}} - \left(\frac{\hat{\beta}}{t_i}\right)^{\hat{\nu}} \right]^{-2} + \frac{2n\hat{\nu}(\hat{\alpha}^2 + 2)}{\hat{\alpha}^2 \hat{\beta}^2},$$

$$I_{23} = I_{32} = \frac{4\hat{\nu}}{\hat{\beta}} \sum_{i=1}^{n} \left[\ln\left(\frac{t_i}{\hat{\beta}}\right)\right]\left[\left(\frac{t_i}{\hat{\beta}}\right)^{\hat{\nu}} + \left(\frac{\hat{\beta}}{t_i}\right)^{\hat{\nu}} \right]^{-2}$$

$$- \frac{2\hat{\nu}}{\hat{\beta}\hat{\alpha}^2} \sum_{i=1}^{n} \left[\ln\left(\frac{t_i}{\hat{\beta}}\right)\right]\left[\left(\frac{t_i}{\hat{\beta}}\right)^{2\hat{\nu}} + \left(\frac{\hat{\beta}}{t_i}\right)^{2\hat{\nu}} \right],$$

$$I_{33} = \frac{n}{\hat{v}^2} - 4\sum_{i=1}^{n}\left[\ln\left(\frac{t_i}{\hat{\beta}}\right)\right]^2\left[\left(\frac{t_i}{\hat{\beta}}\right)^{\hat{v}} + \left(\frac{\hat{\beta}}{t_i}\right)^{\hat{v}}\right]^{-2},$$

$$+\frac{2}{\hat{\alpha}^2}\sum_{i=1}^{n}\left[\ln\left(\frac{t_i}{\hat{\beta}}\right)\right]\left[\left(\frac{t_i}{\hat{\beta}}\right)^{2\hat{v}} + \left(\frac{\hat{\beta}}{t_i}\right)^{2\hat{v}}\right].$$

Hence, a local estimate of the asymptotic variance-covariance matrix of the MLEs can be obtained by inverting the observed Fisher information matrix as

$$\mathbf{I}^{-1}(\hat{\alpha},\hat{\beta},\hat{v}) = \begin{bmatrix} \widehat{Var}(\hat{\alpha}) & \widehat{Cov}(\hat{\alpha},\hat{\beta}) & \widehat{Cov}(\hat{\alpha},\hat{v}) \\ & \widehat{Var}(\hat{\beta}) & \widehat{Cov}(\hat{\beta},\hat{v}) \\ & & \widehat{Var}(\hat{v}) \end{bmatrix}.$$

Following the asymptotic theory of MLEs and the fact that α, β and v are positive parameters, use log transformations to obtain approximate confidence intervals for these parameters (see, for example, Meeker and Escobar 1998). Specifically, for parameter α, we can approximate the distribution of $\frac{\ln(\hat{\alpha})-\ln(\alpha)}{\sqrt{Var(\ln(\hat{\alpha}))}}$ by a standard normal distribution, where the variance of the log-transformed MLE, $Var(\ln(\hat{\alpha}))$, can be approximated by delta method as

$$\widehat{Var}(\ln(\hat{\alpha})) = \frac{\widehat{Var}(\hat{\alpha})}{\hat{\alpha}^2}.$$

A two-sided $100(1-\delta)\%$ normal-approximation confidence interval for α obtained in this manner is then given by

$$[\hat{\alpha}_l, \hat{\alpha}_u] = \left[\frac{\hat{\alpha}}{\exp\left(\frac{z_{1-\frac{\delta}{2}}\sqrt{\widehat{Var}(\hat{\alpha})}}{\hat{\alpha}}\right)}, \hat{\alpha}\cdot\exp\left(\frac{z_{1-\frac{\delta}{2}}\sqrt{\widehat{Var}(\hat{\alpha})}}{\hat{\alpha}}\right)\right], \tag{28}$$

where z_q is the q-th upper percentile of a standard normal distribution. Following the same procedure, normal-approximation confidence intervals for β and v can be constructed.

Beside the use of ML estimation, moment and quantile based estimators can also be considered. Based on the first moments of the random variable T and $1/T$, we have

$$\frac{E(T)}{E(1/T)} = \beta^2.$$

Therefore, the parameter β can be estimated by using the first sample moments of T and $1/T$, i.e.,

$$\tilde{\beta}_1 = \left(\frac{\sum_{i=1}^{n}t_i}{\sum_{i=1}^{n}1/t_i}\right)^{1/2} = \frac{2n}{\hat{\alpha}^2}.$$

Consider the monotone transformation

$$W = \frac{1}{2}\left[\left(\frac{T}{\beta}\right)^v - \left(\frac{\beta}{T}\right)^v\right],$$

which follows a normal distribution with mean 0 and variance $\alpha^2/4$. For known values of β and v, by equating the sample second moment of W to $\alpha^2/4$, an estimator of the parameter α can be expressed as

$$\tilde{\alpha} = \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta} \right)^{\nu} - \left(\frac{\beta}{t_i} \right)^{\nu} \right]^2 \right\}^{1/2}.$$

Note that this equation is equivalent to the estimate obtained based on the likelihood equation for α.

Since the parameter β is the median of the distribution, therefore, the median of $T_1, T_2, \ldots, T_n$ can be used as an estimate of β. Let us denote this estimator as

$$\tilde{\beta}_2 = \text{Median}(T_1, T_2, \ldots, T_n).$$

After obtaining the estimate of β, we can consider the transformed data $x_i = t_i/\tilde{\beta}$, $i = 1, 2, \ldots, n$ and denote the corresponding order statistics as $x_{1:n} < x_{2:n} < \ldots < x_{n:n}$. Consider the following non-linear regression model,

$$y_i = g(x_i; \alpha, \nu) + \varepsilon_i, \quad i = 1, 2, \ldots, n,$$

where

$$
\begin{aligned}
y_i &= \Phi^{-1} \left[\frac{i - 0.5}{n} \right], \\
x_i &= x_{i:n}, \\
g(x; \alpha, \nu) &= \frac{1}{\alpha} (x^{\nu} - x^{-\nu}),
\end{aligned}
$$

and ε_i is the error term. Nonlinear least-squares estimates of the parameters α and ν of the above model can be obtained.

In order to evaluate the performance of different estimation procedures for the parameters of the GB-S$_2(\alpha, \beta, \nu)$ distribution, a Monte Carlo simulation study is conducted. We consider the following estimation procedures:

Method 1. Estimate β by $\tilde{\beta}_1$, then obtain an estimate of ν, say $\tilde{\nu}_1$ by Eq. (27) by substituting $\beta = \tilde{\beta}_1$. Use Eq. (25) to obtain an estimate of α, say $\tilde{\alpha}_1$ with $\beta = \tilde{\beta}_1$ and $\nu = \tilde{\nu}_1$.

Method 2. Estimate β by $\tilde{\beta}_2$, then obtain an estimate of ν, say $\tilde{\nu}_2$ by Eq. (27) by substituting $\beta = \tilde{\beta}_2$. Use Eq. (25) to obtain an estimate of α, say $\tilde{\alpha}_2$ with $\beta = \tilde{\beta}_2$ and $\nu = \tilde{\nu}_2$.

Method 3. Estimate β by $\tilde{\beta}_1$, then obtain estimates of ν and α, say $\tilde{\nu}_1^*$ and $\tilde{\alpha}_1^*$, by using the non-linear least-squares method.

Method 4. Estimate β by $\tilde{\beta}_2$, then obtain estimates of ν and α, say $\tilde{\nu}_2^*$ and $\tilde{\alpha}_2^*$, by using the non-linear least-squares method.

Method 5. Maximum likelihood estimation based on solving Eqs. (25) − (27).

In the simulation study, 1,000 simulations are used to estimate the biases and mean squared errors (MSEs) of these estimators. We consider the parameter settings $\alpha = 1, \nu = 0.5, \beta = 0.5, 1.0, 1.5$; $\alpha = 1, \nu = 0.9, \beta = 0.5, 1.0, 1.5$; $\alpha = 2, \nu = 0.5, \beta = 0.5, 1.0, 1.5$. The simulation results are presented in Table 2, Table 3 and Table 4 for sample sizes $n = 20$, 40 and 60.

From Table 2, Table 3 and Table 4, we observed that the ML estimators (Method 5) do not perform well in terms of MSE for small to moderate sample sizes (say, $n = 20$ and $n = 40$), especially for the estimation of the parameter α. Even for large sample size ($n = 60$), the MSEs of the ML estimators are larger than those obtained by other methods

Table 2 Simulated biases and MSEs for different estimation procedures for parameter setting $\alpha = 1.0, \nu = 0.5$

			β		α		ν	
			Bias	MSE	Bias	MSE	Bias	MSE
$n = 20$	$\beta = 1.0$	Method 1	0.0195	0.0425	0.9742	4.0217	0.2869	0.3524
		Method 2	0.0434	0.0813	0.4214	2.2961	0.0974	0.2553
		Method 3	0.0195	0.0425	0.1487	2.1084	−0.0322	0.2548
		Method 4	0.0434	0.0813	−0.4201	1.3537	−0.2534	0.2354
		Method 5	0.0307	0.0529	1.3228	7.3012	0.3447	0.4640
	$\beta = 1.5$	Method 1	0.0381	0.0932	0.9322	3.2543	0.2855	0.3089
		Method 2	0.0650	0.1731	0.3446	1.7079	0.0759	0.2060
		Method 3	0.0381	0.0932	0.0979	1.6965	−0.0443	0.2157
		Method 4	0.0650	0.1731	−0.4908	1.1100	−0.2836	0.2131
		Method 5	0.0408	0.1285	1.4301	7.0455	0.3891	0.4982
	$\beta = 0.5$	Method 1	0.0108	0.0102	0.9442	4.3086	0.2758	0.3304
		Method 2	0.0217	0.0216	0.4038	2.0323	0.0936	0.2282
		Method 3	0.0108	0.0102	0.1228	2.1729	−0.0445	0.2402
		Method 4	0.0217	0.0216	−0.4360	1.2374	−0.2616	0.2193
		Method 7	0.0095	0.0125	1.4200	7.2533	0.3673	0.4795
$n = 40$	$\beta = 1.0$	Method 1	0.0174	0.0208	0.4064	1.1396	0.1259	0.1394
		Method 2	0.0267	0.0434	0.1390	0.8548	0.0183	0.1206
		Method 3	0.0174	0.0208	−0.0899	0.9007	−0.0928	0.1441
		Method 4	0.0267	0.0434	−0.4459	0.8461	−0.2485	0.1733
		Method 5	0.0112	0.0220	0.5577	1.6258	0.1622	0.1789
	$\beta = 1.5$	Method 1	0.0130	0.0433	0.3903	1.1135	0.1147	0.1417
		Method 2	0.0244	0.0859	0.1178	0.7991	0.0066	0.1238
		Method 3	0.0130	0.0433	−0.0822	0.8952	−0.0911	0.1470
		Method 4	0.0244	0.0859	−0.4382	0.7508	−0.2415	0.1621
		Method 5	0.0140	0.0495	0.5002	1.5122	0.1441	0.1717
	$\beta = 0.5$	Method 1	0.0037	0.0050	0.3697	1.0955	0.1061	0.1347
		Method 2	0.0060	0.0093	0.1313	0.8476	0.0104	0.1214
		Method 3	0.0037	0.0050	−0.1212	0.8887	−0.1097	0.1481
		Method 4	0.0060	0.0093	−0.4268	0.8502	−0.2423	0.1716
		Method 5	−0.0001	0.0056	0.5676	1.5449	0.1693	0.1721
$n = 60$	$\beta = 1.0$	Method 1	−0.0023	0.0120	0.2280	0.6084	0.0645	0.0888
		Method 2	0.0046	0.0267	0.0521	0.5265	−0.0104	0.0873
		Method 3	−0.0023	0.0120	−0.1512	0.5816	−0.1065	0.1105
		Method 4	0.0046	0.0267	−0.4006	0.6432	−0.2208	0.1402
		Method 5	0.0132	0.0138	0.3006	0.7754	0.0893	0.1049
	$\beta = 1.5$	Method 1	0.0089	0.0309	0.2217	0.6390	0.0642	0.0901
		Method 2	0.0178	0.0562	0.0497	0.5494	−0.0094	0.0881
		Method 3	0.0089	0.0309	−0.1560	0.6107	−0.1074	0.1129
		Method 4	0.0178	0.0562	−0.3966	0.6474	−0.2171	0.1368
		Method 5	0.0044	0.0330	0.3469	0.8634	0.1039	0.1131
	$\beta = 0.5$	Method 1	0.0018	0.0032	0.2649	0.6466	0.0822	0.0934
		Method 2	0.0026	0.1010	0.0850	0.5690	0.0053	0.0911
		Method 3	0.0018	0.0032	−0.1172	0.5847	−0.0889	0.0145
		Method 4	0.0026	0.1010	−0.3739	0.6695	−0.2078	0.1414
		Method 5	0.0009	0.0033	0.3197	0.7679	0.0986	0.1013

Table 3 Simulated biases and MSEs for different estimation procedures for parameter setting $\alpha = 2.0, \nu = 0.5$

			β		α		ν	
			Bias	MSE	Bias	MSE	Bias	MSE
$n = 20$	$\beta = 1.0$	Method 1	0.0386	0.1125	1.3195	7.9982	0.1482	0.0990
		Method 2	0.1371	0.4192	0.0557	2.9805	−0.0424	0.0740
		Method 3	0.0386	0.1125	0.4002	4.6121	−0.0118	0.0910
		Method 4	0.1371	0.4192	−1.1470	3.0691	−0.3047	0.1604
		Method 5	0.0749	0.1563	1.6102	12.0105	0.1629	0.1341
	$\beta = 1.5$	Method 1	0.0741	0.2669	1.2255	6.2592	0.1490	0.0933
		Method 2	0.2483	0.9636	0.0169	2.5749	−0.0439	0.0760
		Method 3	0.0741	0.2669	0.3387	3.8775	−0.0086	0.0843
		Method 4	0.2483	0.9636	−1.1137	3.0016	−0.2946	0.1566
		Method 5	0.0736	0.2879	1.8437	14.1580	0.1863	0.1367
	$\beta = 0.5$	Method 1	0.0300	0.0311	1.1648	6.3143	0.1345	0.0929
		Method 2	0.0830	0.1013	0.0232	2.9223	−0.0484	0.0799
		Method 3	0.0300	0.0311	0.2871	3.8638	−0.0234	0.0886
		Method 4	0.0830	0.1013	−1.0694	3.2768	−0.2920	0.1621
		Method 5	0.0712	0.2610	1.7594	12.6750	0.1843	0.1315
$n = 40$	$\beta = 1.0$	Method 1	0.0283	0.0510	0.4013	1.8797	0.0427	0.0391
		Method 2	0.0675	0.1607	−0.1425	1.3239	−0.0564	0.0424
		Method 3	0.0283	0.0510	−0.0800	1.4982	−0.0522	0.0469
		Method 4	0.0675	0.1607	−0.9059	2.0129	−0.2316	0.1069
		Method 5	0.0253	0.0554	0.6419	2.4061	0.0782	0.0458
	$\beta = 1.5$	Method 1	0.0386	0.1025	0.5161	1.9371	0.0593	0.0384
		Method 2	0.1448	0.4378	−0.0506	1.2901	−0.0414	0.0429
		Method 3	0.0386	0.1025	0.0286	1.5041	−0.0350	0.0453
		Method 4	0.1448	0.4378	−0.8061	1.8517	−0.2089	0.0993
		Method 5	0.0369	0.1242	0.6034	2.3951	0.0697	0.0439
	$\beta = 0.5$	Method 1	0.0111	0.0126	0.5099	2.0122	0.0580	0.0370
		Method 2	0.0450	0.0478	−0.0599	1.4283	−0.0459	0.0443
		Method 3	0.0111	0.0126	0.0191	1.5585	−0.0372	0.0440
		Method 4	0.0450	0.0478	−0.8308	2.0662	−0.2190	0.1056
		Method 5	0.0121	0.0142	0.7498	2.9983	0.0894	0.0511
$n = 60$	$\beta = 1.0$	Method 1	0.0101	0.0312	0.3484	1.2761	0.0380	0.0300
		Method 2	0.0454	0.1185	−0.0580	0.9684	−0.0351	0.0315
		Method 3	0.0101	0.0312	0.0024	1.0757	−0.0282	0.0327
		Method 4	0.0454	0.1185	−0.6773	1.4934	−0.1730	0.0763
		Method 5	0.0117	0.0341	0.4055	1.3271	0.0490	0.0275
	$\beta = 1.5$	Method 1	0.0288	0.0703	0.2911	0.9903	0.0348	0.0239
		Method 2	0.0813	0.2419	−0.0925	0.7792	−0.0361	0.0269
		Method 3	0.0288	0.0703	−0.0473	0.8886	−0.0320	0.0278
		Method 4	0.0813	0.2419	−0.6837	1.3505	−0.1694	0.0701
		Method 5	0.0366	0.0835	0.3219	1.1432	0.0369	0.0272
	$\beta = 0.5$	Method 1	0.0120	0.0091	0.3565	1.1113	0.0423	0.0255
		Method 2	0.0365	0.0323	−0.0721	0.8640	−0.0365	0.0296
		Method 3	0.0120	0.0091	0.0150	0.9128	−0.0228	0.0269
		Method 4	0.0365	0.0323	−0.6760	1.3800	−0.1699	0.0720
		Method 5	0.0051	0.0081	0.3621	1.2234	0.0432	0.0260

Table 4 Simulated biases and MSEs for different estimation procedures for parameter setting $\alpha = 1.0, \nu = 0.9$

			β		α		ν	
			Bias	MSE	Bias	MSE	Bias	MSE
$n = 20$	$\beta = 1.0$	Method 1	0.0098	0.0132	0.8859	4.1859	0.4528	1.0031
		Method 2	0.0126	0.0243	0.4174	2.6917	0.1578	0.7571
		Method 3	0.0098	0.0132	0.0205	2.1971	−0.1634	0.7715
		Method 4	0.0126	0.0243	−0.4101	1.4524	−0.4618	0.7328
		Method 5	0.0083	0.0153	1.4731	7.7400	0.6839	1.6105
	$\beta = 1.5$	Method 1	0.0130	0.0309	0.8992	3.9446	0.4732	1.0613
		Method 2	0.0220	0.0545	0.4182	2.1596	0.1773	0.7605
		Method 3	0.0130	0.0309	0.0643	2.2646	−0.1191	0.7975
		Method 4	0.0220	0.0545	−0.4120	1.3491	−0.4478	0.7151
		Method 5	0.0174	0.0324	1.3639	6.1369	0.6680	1.4928
	$\beta = 0.5$	Method 1	0.0016	0.0033	0.8285	3.4397	0.4223	0.9522
		Method 2	0.0022	0.0060	0.3769	1.9818	0.1481	0.6989
		Method 3	0.0016	0.0033	−0.0263	1.8751	−0.1866	0.7413
		Method 4	0.0022	0.0060	−0.4537	1.2252	−0.4826	0.6884
		Method 5	0.0041	0.0039	1.3170	6.0487	0.6453	1.4899
$n = 40$	$\beta = 1.0$	Method 1	0.0038	0.0066	0.3101	0.9566	0.1507	0.4135
		Method 2	0.0058	0.0128	0.1039	0.7441	0.0009	0.3774
		Method 3	0.0038	0.0066	−0.1887	0.7752	−0.2426	0.4574
		Method 4	0.0058	0.0128	−0.4548	0.7507	−0.4519	0.5265
		Method 5	0.0029	0.0070	0.5476	1.5316	0.2904	0.5658
	$\beta = 1.5$	Method 1	0.0002	0.0144	0.3060	1.0136	0.1518	0.4263
		Method 2	0.0002	0.0270	0.1067	0.8040	0.0045	0.3930
		Method 3	0.0002	0.0144	−0.2057	0.8406	−0.2554	0.4841
		Method 4	0.0002	0.0270	−0.4648	0.8255	−0.4624	0.5604
		Method 5	0.0102	0.0149	0.5592	1.5119	0.3069	0.5646
	$\beta = 0.5$	Method 1	−0.0010	0.0015	0.3810	1.0063	0.2074	0.4302
		Method 2	−0.0004	0.0029	0.1759	0.7747	0.0596	0.3760
		Method 3	−0.0010	0.0015	−0.1378	0.7647	−0.1964	0.4470
		Method 4	−0.0004	0.0029	−0.4162	0.7652	−0.4181	0.5223
		Method 5	0.0015	0.0017	0.5205	1.5198	0.2768	0.5461
$n = 60$	$\beta = 1.0$	Method 1	0.0012	0.0042	0.1727	0.5760	0.0825	0.2853
		Method 2	0.0017	0.0076	0.0400	0.5214	−0.0212	0.2874
		Method 3	0.0012	0.0042	−0.2268	0.5858	−0.2465	0.3796
		Method 4	0.0017	0.0076	−0.4125	0.6330	−0.4002	0.4547
		Method 5	0.0013	0.0041	0.3112	0.7285	0.1738	0.3287
	$\beta = 1.5$	Method 1	−0.0019	0.0094	0.2492	0.6194	0.1329	0.2882
		Method 2	−0.1581	0.0429	0.1031	0.5236	0.0221	0.2745
		Method 3	−0.0019	0.0094	−0.0013	0.5552	−0.0017	0.3187
		Method 4	−0.1581	0.0429	−0.3658	0.5916	−0.3655	0.4219
		Method 5	0.0072	0.0104	0.3172	0.7224	0.1742	0.3264
	$\beta = 0.5$	Method 1	0.0017	0.0010	0.1963	0.5976	0.0976	0.2855
		Method 2	0.0032	0.0020	0.0502	0.5026	−0.0130	0.2738
		Method 3	0.0017	0.0010	−0.2188	0.5991	−0.2424	0.3867
		Method 4	0.0032	0.0020	−0.4192	0.6177	−0.4049	0.4456
		Method 5	0.0010	0.0011	0.3104	0.8103	0.1656	0.3594

in most cases. Therefore, we would not consider the ML estimators in the subsequence comparisons.

Based on the simulation results, for parameter β, the estimator based on the first sample moments of T and $1/T$, i.e., $\tilde{\beta}_1$ (Method 1 and Method 3) gives smallest MSEs in most situations. However, the estimators for α and ν based on the value of $\tilde{\beta}_1$ are not the best among all the methods considered here. For estimation of parameters α and ν with small sample sizes ($n = 20$), we observed that Method 4 performs better when the true value of $\alpha = 1.0$ and Method 2 performs better when the true value of $\alpha = 2.0$. It is interesting to point out that even Method 4 does not perform as well as Method 2 when the true value of $\alpha = 2.0$, the variances of the estimators from Method 4 are much smaller than that of Method 2. For moderate and large sample sizes ($n = 40$ and 60), Method 2 gives the smaller MSEs in most cases.

Overall speaking, we would recommend the use of moment-based estimator $\tilde{\beta}_1$ for the estimation of the parameter β for any sample sizes. For small sample sizes, we would suggest the use of non-linear least-squared method by setting $\beta = \tilde{\beta}_2$ (i.e., Method 4) to estimate the parameters α and ν. For moderate to large sample sizes, the use of likelihood equations by setting $\beta = \tilde{\beta}_2$ (i.e., Method 2) is recommended.

5 Illustrative examples

5.1 Example 1: simulated data from two-fold Weibull mixture

In this subsection, we use a simulated dataset to illustrate the usefulness of the generalized Birnbaun-Saunders distribution in modeling bimodal data and the estimation procedures studied in Section 4. We consider a two-fold Weibull mixture model which is commonly used to model two-fold competing risk failure mechanism involving two failure modes in reliability engineering. The PDF of the two-fold Weibull mixture is given by (see, for example, Murthy et al. 2004, Razal and Salih 2009):

$$f_{MW}(x) = w f_W(x; a_1, b_1) + (1 - w) f_W(x, a_2, b_2),\qquad(29)$$

where

$$f_W(x; a, b) = \frac{a}{b} \left(\frac{x}{b}\right)^{(a-1)} \exp\left[-\left(\frac{x}{b}\right)^a\right], \ \ x > 0,$$

is the density of the Weibull distribution with shape parameter $a > 0$ and scale parameter $b > 0$, and $0 < w < 1$ is the mixing parameter. A sample of size 50 is generated from the two-fold Weibull mixture in (29) with $w = 0.6, a_1 = 6, b_1 = 1, a_2 = 6, b_2 = 1$. The dataset is presented in Table 5 and the GS-B$_2$ distribution is used to model this dataset. The estimates of the model parameters based on different methods studied in Section 4 are presented in Table 6 and the histogram of the dataset with the fitted probability density functions are plotted in Fig. 5. We also presented the 95 % confidence intervals for the model parameters based on normal-approximation of the log-transformed MLEs that are discussed in Section 4.

Here, we also compare the model studied in Section 4 to the three-parameter Birnbaum-Saunders distribution studied in Owen (2006) with PDF in (8). For the dataset presented in Table 5, the maximum likelihood estimates of the parameters a, b and κ are 0.3856, 1.1336 and 0.4536, respectively, which give the maximum log-likelihood to

Table 5 Simulated dataset from a two-folded Weibull mixture model

0.5446	0.5934	0.5958	0.6106	0.6335	0.6945	0.7097	0.7431	0.7587	0.7851
0.8162	0.8450	0.8502	0.8503	0.8549	0.8743	0.8801	0.9335	0.9833	0.9836
1.0268	1.0394	1.0412	1.0472	1.0732	1.0849	1.0948	1.1129	1.1673	1.1824
1.3267	1.4230	1.4706	1.4995	1.5185	1.5287	1.5353	1.5362	1.5797	1.5995
1.7477	1.7985	1.8231	1.8341	1.8701	1.8793	1.9394	1.9522	2.0401	2.1118

be -63.80. Comparing to the maximum log-likelihood based on the proposed three-parameter GS-B_2 distribution which is -23.69, the proposed model clearly provides a better fit for this dataset.

Although the five-parameter model in (29) is more flexible than the three-parameter GS-B_2 distribution, from this example, we can see that the GS-B_2 distribution can be a simpler and effective alternative to model the bimodal behavior in the density function.

5.2 Example 2: spot exchange rate of euro into sterling pound

To further illustrate the usefulness of the GS-B_2 distribution in modeling bimodal data, we consider a real data example of spot exchange rate of the Euro into sterling pound. The data for this example were downloaded from the Bank of England Statistical Interactive Database (http://www.bankofengland.co.uk/boeapps/iadb) which consist of 3,786 daily observations on the spot exchange rate of the Euro into sterling pound during the period August 29, 2000 to August 15, 2015. A random sample of size 100 of these 3,786 daily spot exchange rate observations are presented in Table 7 and the histogram of the data is presented in Fig. 6, where it is readily seen that the distribution is bimodal. The GS-B_2 distribution is used to model this dataset. The estimates of the model parameters based on different methods studied in Section 4 are presented in Table 8 and the fitted probability density functions are plotted in Fig. 6. We also presented the 95 % confidence intervals for the model parameters based on normal-approximation of the log-transformed MLEs that are discussed in Section 4.

The study of Boothe and Glassman (1987) showed that the mixture of two normal distributions is one of the best models to describe exchange rate data, hence we fit the exchange rate data in Table 7 with a mixture of two normal distributions with PDF
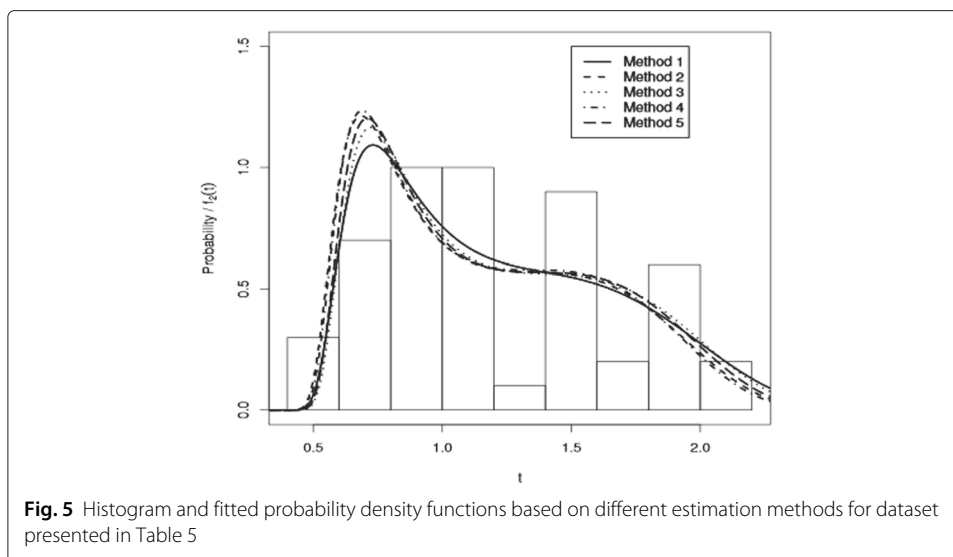
$$f_{MN}(x) = \lambda f_N(x; \mu_1, \sigma_1) + (1 - \lambda) f_N(x, \mu_2, \sigma_2), \tag{30}$$

where $0 \le \lambda \le 1$ and

$$f_N(x; \mu_j, \sigma_j) = \frac{1}{\sigma_j \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x - \mu_j}{\sigma_j} \right)^2 \right], -\infty < x < \infty,$$

Table 6 Estimates of the parameters in GS-B_2 for dataset presented in Table 5

	Estimate of		
	α	β	ν
Method 1	2.9565	1.1256	2.7307
Method 2	3.5729	1.0790	3.0254
Method 3	3.4970	1.1256	3.0461
Method 4	3.6660	1.0790	3.1143
Method 5	3.6325	1.1065	3.1135
95 % CI based on MLE (Method 5)	(1.5859, 8.3186)	(0.9193, 1.3319)	(2.0135, 4.8140)

Fig. 5 Histogram and fitted probability density functions based on different estimation methods for dataset presented in Table 5
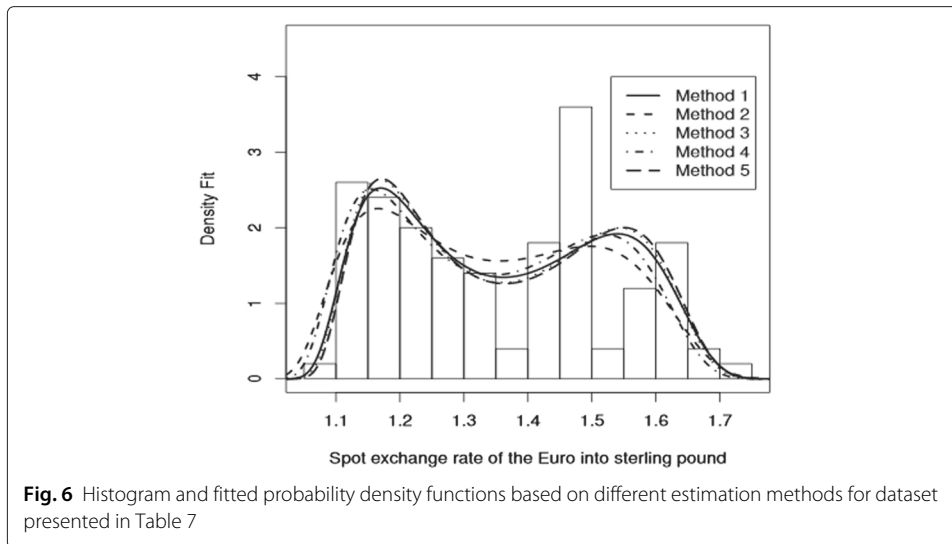
$-\infty < \mu_j < \infty$ and $\sigma_j > 0$ for $j = 1, 2$. The above normal mixture model is fitted by using the R package `normalmixEM` (Benaglia et al. 2009). The MLEs of the model parameters based on the exchange rate data in Table 7 are $\hat{\lambda} = 0.4681$, $\hat{\mu}_1 = 1.2002$, $\hat{\mu}_2 = 1.4992$, $\hat{\sigma}_1 = 0.0623$ and $\hat{\sigma}_2 = 0.0951$ with the maximum log-likelihood value to be -52.2517. Comparing to the maximum log-likelihood based on the proposed three-parameter GS-B$_2$ distribution which is -52.5557, the GS-B$_2$ model provides similar maximum likelihood in fitting this dataset using three parameters instead of five parameters. Note that the Akaike information criterion (AIC) of the GS-B$_2$ distribution is 94.5034 for fitting the data set in Table 7, while the AIC of the mixture of two normal distributions is 99.1114. Therefore, the generalized Birnbaum-Saunders GS-B$_2$ distribution could be chosen as a better model based on AIC.

6 Conclusions

The Birnbaum-Saunders and inverse-Gaussian distributions have a long, rich history in statistical literature. They have often deemed as interchangeable, and this article brought to light more comparisons of their utility and similarities. The moment-ratio diagrams showed another way that the densities and higher moments are quite similar.

Table 7 Random sample of size 100 from daily observations on the spot exchange rate of the Euro into sterling pound during the period August 29, 2000 to August 15, 2015

1.4404	1.1897	1.3923	1.6333	1.2857	1.1726	1.2251	1.4574	1.4757	1.5566
1.1298	1.1787	1.2004	1.2074	1.4873	1.4725	1.4547	1.2513	1.1411	1.3908
1.5754	1.5698	1.1461	1.6241	1.4463	1.4910	1.4633	1.4837	1.4326	1.2799
1.4767	1.4536	1.2907	1.1647	1.1254	1.4639	1.1938	1.2684	1.4454	1.4352
1.2393	1.6595	1.4127	1.1448	1.4688	1.4623	1.1615	1.2122	1.1044	1.5635
1.2000	1.4872	1.2342	1.1908	1.4856	1.4803	1.6295	1.6159	1.6568	1.1092
1.2565	1.2472	1.4729	1.4348	1.1293	1.3284	1.1209	1.6139	1.5138	1.4868
1.2270	1.5058	1.3918	1.1451	1.3256	1.4337	1.2467	1.5673	1.1771	1.2635
1.1433	1.3280	1.4605	1.4453	1.6184	1.0856	1.2671	1.1414	1.3252	1.5695
1.1762	1.3044	1.7034	1.6238	1.5790	1.4554	1.2554	1.3363	1.2449	1.6246
1.1963	1.2469	1.6014	1.1681	1.1846	1.3171	1.1761	1.1405	1.6248	1.4294

Fig. 6 Histogram and fitted probability density functions based on different estimation methods for dataset presented in Table 7

A generalized Birnbaun-Saunders distribution, the GB-S_2 model, is a three-parameter distribution that not only includes the usual two-parameter B-S distribution as a special case but also shares some unique relationships with the I-G model. Lastly, the fact that the GB-S_2 model can exhibit bimodality with certain parameter values makes it a very flexible distribution. It is hoped that this paper generates increased interests in the Birnbaum-Saunders, inverse-Gaussian and the three-parameter generalized Birnbaum-Saunders GB-S_2 models.

7　Appendix: likelihood equations for GB-S_2 distribution

Suppose $t_1, t_2, \ldots, t_n$ is a random sample of size n from the GB-S_2 distribution with CDF in (18) and PDF in (19). The log-likelihood function is

$$\ln L(\alpha, \beta, \nu) = constant + n \ln \nu - n \ln \alpha$$

$$+ \sum_{i=1}^{n} \ln \left[\left(\frac{t_i}{\beta} \right)^{\nu} + \left(\frac{\beta}{t_i} \right)^{\nu} \right]$$

$$- \frac{1}{2\alpha^2} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - 2 + \left(\frac{\beta}{t_i} \right)^{2\nu} \right].$$

Table 8 Estimates of the parameters in GS-B_2 for dataset presented in Table 7

	Estimate of		
	α	β	ν
Method 1	4.9277	1.3487	11.2675
Method 2	3.5225	1.3324	9.2726
Method 3	5.5220	1.3487	11.9603
Method 4	4.7636	1.3324	11.0038
Method 5	5.6042	1.3529	12.0478
95 % CI based on MLE (Method 5)	(3.2669, 9.6136)	(1.2985, 1.4095)	(9.4995, 15.2797)

Taking derivatives of the log-likelihood function with respect to the parameters, we have

$$\frac{\partial \ln L(\alpha, \beta, \nu)}{\partial \alpha} = -\frac{n}{\alpha} + \frac{1}{\alpha^3} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - 2 + \left(\frac{\beta}{t_i} \right)^{2\nu} \right],$$

$$\frac{\partial \ln L(\alpha, \beta, \nu)}{\partial \beta} = -\frac{\nu}{\beta} \sum_{i=1}^{n} \frac{\left(\frac{t_i}{\beta} \right)^{\nu} - \left(\frac{\beta}{t_i} \right)^{\nu}}{\left(\frac{t_i}{\beta} \right)^{\nu} + \left(\frac{\beta}{t_i} \right)^{\nu}} + \frac{\nu}{\beta \alpha^2} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - \left(\frac{\beta}{t_i} \right)^{2\nu} \right],$$

$$\frac{\partial \ln L(\alpha, \beta, \nu)}{\partial \nu} = \frac{n}{\nu} + \sum_{i=1}^{n} \frac{\left[\ln \left(\frac{t_i}{\beta} \right) \right] \left[\left(\frac{t_i}{\beta} \right)^{\nu} - \left(\frac{\beta}{t_i} \right)^{\nu} \right]}{\left(\frac{t_i}{\beta} \right)^{\nu} + \left(\frac{\beta}{t_i} \right)^{\nu}}$$

$$- \frac{1}{\alpha^2} \sum_{i=1}^{n} \left[\ln \left(\frac{t_i}{\beta} \right) \right] \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - \left(\frac{\beta}{t_i} \right)^{2\nu} \right].$$

By setting these derivatives to zeros, we have the following likelihood equations

$$\frac{\partial \ln L(\alpha, \beta, \nu)}{\partial \alpha} = 0$$

$$\Rightarrow \alpha = \left\{ \frac{1}{n} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - 2 + \left(\frac{\beta}{t_i} \right)^{2\nu} \right] \right\}^{1/2},$$

$$\frac{\partial \ln L(\alpha, \beta, \nu)}{\partial \beta} = 0$$

$$\Rightarrow -\sum_{i=1}^{n} \frac{\left(\frac{t_i}{\beta} \right)^{\nu} - \left(\frac{\beta}{t_i} \right)^{\nu}}{\left(\frac{t_i}{\beta} \right)^{\nu} + \left(\frac{\beta}{t_i} \right)^{\nu}} + \frac{1}{\alpha^2} \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - \left(\frac{\beta}{t_i} \right)^{2\nu} \right] = 0$$

$$\Rightarrow \frac{1}{n} \sum_{i=1}^{n} \frac{\left(\frac{t_i}{\beta} \right)^{\nu} - \left(\frac{\beta}{t_i} \right)^{\nu}}{\left(\frac{t_i}{\beta} \right)^{\nu} + \left(\frac{\beta}{t_i} \right)^{\nu}}$$

$$= \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - \left(\frac{\beta}{t_i} \right)^{2\nu} \right] \left\{ \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - 2 + \left(\frac{\beta}{t_i} \right)^{2\nu} \right] \right\}^{-1},$$

$$\frac{\partial \ln L(\alpha, \beta, \nu)}{\partial \nu} = 0$$

$$\Rightarrow \frac{n}{\nu} + \sum_{i=1}^{n} \frac{\left[\ln \left(\frac{t_i}{\beta} \right) \right] \left[\left(\frac{t_i}{\beta} \right)^{\nu} - \left(\frac{\beta}{t_i} \right)^{\nu} \right]}{\left(\frac{t_i}{\beta} \right)^{\nu} + \left(\frac{\beta}{t_i} \right)^{\nu}}$$

$$= \frac{1}{\alpha^2} \sum_{i=1}^{n} \left[\ln \left(\frac{t_i}{\beta} \right) \right] \left[\left(\frac{t_i}{\beta} \right)^{2\nu} - \left(\frac{\beta}{t_i} \right)^{2\nu} \right]$$

$$\Rightarrow \frac{1}{\nu} + \frac{1}{n} \sum_{i=1}^{n} \frac{\left[\ln \left(\frac{t_i}{\beta} \right) \right] \left[\left(\frac{t_i}{\beta} \right)^{\nu} - \left(\frac{\beta}{t_i} \right)^{\nu} \right]}{\left(\frac{t_i}{\beta} \right)^{\nu} + \left(\frac{\beta}{t_i} \right)^{\nu}}$$

$$= \left\{ \sum_{i=1}^{n} \left[\ln\left(\frac{t_i}{\beta}\right) \right] \left[\left(\frac{t_i}{\beta}\right)^{2\nu} - \left(\frac{\beta}{t_i}\right)^{2\nu} \right] \right\}$$
$$\times \left\{ \sum_{i=1}^{n} \left[\left(\frac{t_i}{\beta}\right)^{2\nu} - 2 + \left(\frac{\beta}{t_i}\right)^{2\nu} \right] \right\}^{-1}.$$

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

HKTN and WJO contributed equally to the manuscript. HKTN is the only author who read and approved the final manuscript since WJO passed away before the manuscript was submitted.

Author details

[1]Department of Mathematics and Computer Science, University of Richmond, Richmond, VA, 23173 USA. [2]Department of Statistical Science, Southern Methodist University, Dallas, TX, 75275 USA.

References

Ashcar, JA: Inferences for the Birnbaum-Saunders fatigue life model using Bayesian methods. Comput. Stat. Data Anal. **15**, 367–380 (1993)

Balakrishnan, N, Gupta, R, Kundu, D, Leiva, V, Sanhueza, A: On Some Mixture Models Based on the Birnbaum-Saunders Distribution and Associated Inference. J. Stat. Plan. Infer. **141**, 2175–2190 (2011)

Barros, M, Paula, GA, Leiva, V: An R implementation for generalized Birnbaum-Saunders distributions. Comput. Stat. Data Anal. **53**, 1511–1528 (2009)

Benaglia, T, Chauveau, D, Hunter, DR: Young, DS: mixtools: An R Package for Analyzing Finite Mixture Models. J. Stat. Softw. **32**, 1–29 (2009). http://www.jstatsoft.org/v32/i06/

Bhattacharyya, GK, Fries, A: Fatigue failure models – Birnbaum-Saunders vs. inverse Gaussian. IEEE Trans. Reliab. **31**, 439–441 (1982)

Birnbaum, ZW, Saunders, SC: A new family of life distributions. J. Appl. Probab. **6**, 319–327 (1969a)

Birnbaum, Z W, Saunders, SC: Estimation for a family of life distributions with applications to fatigue. J. Appl. Probab. **6**, 328–347 (1969b)

Boothe, P, Glassman, D: The Statistical Distribution Of Exchange Rates. J. Int. Econ. **22**, 297–319 (1987)

Castillo, NO, Gómez, HW, Bolfarine, H: Epsilon Birnbaum-Saunders distribution family: Properties and inference. Stat. Pap. **52**, 871–883 (2011)

Chen, M, Ibeahim, JG, Chi, Y-Y: A new class of mixture models for differential gene expression in DNA microarray data. Journal of Statistical Planning and Inference. **138**, 387–404 (2008)

Chhikara, RS, Folks, L: The Inverse Gaussian Distribution: Theory, Methodology, and Applications. Marcel Dekker, New York (1989)

Cordeiro, GM, Cancho, VG, Ortega, EMM, Barriga, GDC: A model with long-term survivors: Negative binomial Birnbaum-Saunders. Communications in Statistics – Theory and Methods (to appear)

Cordeiro, GM, Lemonte, AJ: The β-Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling. Comput. Stat. Data Anal. **55**, 1445–1461 (2011)

Cordeiro, G M, Lemonte, AJ: The exponentiated generalized Birnbaum-Saunders distribution. Appl. Math. Comput. **247**, 762–779 (2014)

Cox, DR, Oates, D: Analysis of Survival Data. Chapman and Hall, New York (1984)

Desmond, AF: Stochastic models of failure in random environments. Can. J. Stat. **13**, 171–183 (1985)

Desmond, A F: On the relationship between two fatigue-life models. IEEE Trans. Reliab. **35**, 167–169 (1986)

Díaz-García, JA, Leiva-Sánchez, V: A new family of life distributions based on the elliptically contoured distributions. J Stat. Plan. Infer. **128**, 445–457 (2005)

Díaz-García, JA, Domínguez-Molina, JR: Some generalizations of Birnbaum-Saunders and sinh-normal distributions. Int. Math. Forum. **1**, 1709–1727 (2006)

Dupuis, DJ, Mills, JE: Robust Estimation of the Birnbaum-Saunders Distribution. IEEE Trans. Reliab. **47**, 88–95 (1998)

Durham, SD, Padgett, WJ: A cumulative damage model for system failure with application to carbon fibers and composites. Technometrics. **39**, 34–44 (1997)

Engelhardt, M, Bain, LJ, Wright, FT: Inferences on the parameters of the Birnbaum-Saunders fatigue life distribution based on maximum likelihood estimation. Technometrics. **23**, 251–256 (1981)

Genç, AI: The generalized T Birnbaum-Saunders family. Statistics. **47**, 613–625 (2013)

Gómez, HW, Olivares-Pacheco, JF, Bolfarine, H: An extension of the generalized Birnbaum-Saunders distribution. Stat. Probab. Lett. **79**, 331–338 (2009)

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

Hossain, MF, Kashiwagi, N, Hirano, K: A Generalization of the Power Inverse Gaussian Distribution and Some of its Properties. In: Johnson, NL, Balakrishnan, N (eds.) Advances in the Theory and Practice of Statistics: A Volume in Honor of Samuel Kotz. John Wiley & Sons. Sons, New York, (1997)

Johnson, NL, Kotz, S, Balakrishnan, N: Continuous Univariate Distributions, Vol. 1. John Wiley & Sons, New York (1995a)

Johnson, N L, Kotz, S, Balakrishnan, N: Continuous Univariate Distributions, Vol. 2. John Wiley & Sons, New York (1995b)

Lawless, JF: Statistical Models and Methods for Lifetime Data. John Wiley & Sons, New York (1982)

Lehmann, EL, Casella, G: Theory of Point Estimation. 2nd ed. Springer-Verlag, New York (1998)

Lemonte, AJ, Ferrari, SLP: Testing Hypothesis in the Birnbaum-Saunders Distribution Under Type-II Censored Samples. Comput. Stat. Data Anal. **55**, 2388–2399 (2011)

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

Leiva, V, Hernández, H, Riquelme, M: A new package for the Birnbaum-Saunders distribution. R Journal. **6**, 35–40 (2006)

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

Meeker, WQ, Escobar, LA: Statistical Methods for Reliability Data. John Wiley & Sons, New York (1998)

Miner, MA: Cumulative Damage in Fatigue. J. Appl. Mech. Trans. ASME. **67**, 159–164 (1945)

Murthy, DNP, Xie, M, Jiang, R: Weibull Models. John Wiley & Sons, New York (2004)

Ng, HKT, Kundu, D, Balakrishnan, N: Point and Interval Estimation for the Two-parameter Birnbaum-Saunders Distribution Based on Type-II Censored Samples. Comput. Stat. Data Anal. **50**, 3222–3242 (2006)

Owen, WJ: A new three-parameter extension to the Birnbaum-Saunders distribution. IEEE Trans. Reliab. **55**, 475–479 (2006)

Owen, W J: An exponential damage model for strength of fibrous composite materials. IEEE Trans. Reliab. **56**, 459–463 (2007)

Owen, WJ, Padgett, WJ: A Birnbaum-Saunders Accelerated Life Model. IEEE Trans. Reliab. **49**, 224–229 (2000)

Core Team, R: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria (2015)

Razal, AM, Salih, AA: Combining Two Weibull Distributions Using a Mixing Parameter. Eur. J. Sci. Res. **31**, 296–305 (2009)

Rieck, JR: Statistical Analysis for the Birnbaum-Saunders Fatigue Life Distribution. Ph. D. Thesis, Clemson University (1989)

Rieck, J R: A Moment-Generating Function with Application to the Birnbaum-Saunders Distribution. Communications in Statistics - Theory and Methods. **28**, 2213–2222 (1999)

Rieck, JR, Nedleman, J: A Log-Linear Model for the Birnbaum-Saunders Distribution. Technometrics. **33**, 51–60 (1991)

Sanhueza, A, Leiva, V, Balakrishnan, N: The Generalized Birnbaum-Saunders Distribution and Its Theory, Methodology, and Application. Communications in Statistics – Theory and Methods. **37**, 645–670 (2008)

Saunders, SC: A Family of random variables closed under reciprocation. J. Am. Stat. Assoc. **69**, 553–539 (1974)

Vargo, E, Pasupathy, R, Leemis, L: Moment-Ratio Diagrams for Univariate Distributions. J. Qual. Technol. **42**, 276–286 (2010)

Watson, GN: A Treatise on the Theory of Bessel Functions. Cambridge University Press, New York (1995)

Zacks, S: Introduction to Reliability Analysis. Springer-Verlag, New York (1992)