JOURNAL OF APPLIED STATISTICS
2020, VOL. 47, NOS. 13–15, 2690–2710
https://doi.org/10.1080/02664763.2019.1668365

On a new type of Birnbaum-Saunders models and its inference and application to fatigue data

Jaime Arrué[a], Reinaldo B. Arellano-Valle[b], Héctor W. Gómez[a] and Víctor Leiva[c]

[a]Department of Mathematics, Universidad de Antofagasta, Antofagasta, Chile; [b]Department of Statistics, Pontificia Universidad Católica de Chile, Santiago, Chile; [c]School of Industrial Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile

ABSTRACT

The Birnbaum-Saunders distribution is a widely studied model with diverse applications. Its origins are in the modeling of lifetimes associated with material fatigue. By using a motivating example, we show that, even when lifetime data related to fatigue are modeled, the Birnbaum-Saunders distribution can be unsuitable to fit these data in the distribution tails. Based on the nice properties of the Birnbaum-Saunders model, in this work, we use a modified skew-normal distribution to construct such a model. This allows us to obtain flexibility in skewness and kurtosis, which is controlled by a shape parameter. We provide a mathematical characterization of this new type of Birnbaum-Saunders distribution and then its statistical characterization is derived by using the maximum-likelihood method, including the associated information matrices. In order to improve the inferential performance, we correct the bias of the corresponding estimators, which is supported by a simulation study. To conclude our investigation, we retake the motivating example based on fatigue life data to show the good agreement between the new type of Birnbaum-Saunders distribution proposed in this work and the data, reporting its potential applications.

ARTICLE HISTORY
Received 16 November 2018
Accepted 10 July 2019

KEYWORDS
Correction of bias; fatigue life data; maximum likelihood method; Monte Carlo simulation; R software; skew-normal distribution

1. Bibliographical review and motivating example

In this section, we provide an introduction to the topic accompanied by a state of the art about studies linked to the present research. In addition, a motivating example from material fatigue is discussed to justify the development of the proposed methodology.

1.1. Introduction

The Birnbaum-Saunders (BS) model is considered as a life distribution due to its origins describing fatigue of materials subject to stress. Then, the BS distribution assumes a prominent role in the areas of reliability and survival analysis, being a good alternative to standard distributions used in such areas. For more details, see [10, 26, pp. 651–663, 32].

CONTACT Víctor Leiva victorleivasanchez@gmail.com

The origins of the BS model lie in physical problems of vibrations in commercial aircrafts which provoke material fatigue. Note that fatigue is a structural damage that occurs when the material is exposed to stress and tension fluctuations. Motivated by these problems, Birnbaum and Saunders [10] proposed a family of distributions to model the failure time of materials and equipments subject to dynamic loads, where the failure stems from the development and growth of a dominant crack. Birnbaum and Saunders [11] formalized this distribution as a probabilistic model for fatigue life, developed estimation of its parameters using the maximum likelihood (ML) method and derived some of their mathematical and statistical properties.

Although its origins were in the statistical modeling of equipment lifetimes under dynamic loads, the BS distribution has been studied and applied in many fields including, but not restricted to, business and industry as well as earth, economic/financial/management, engineering and medical sciences, which have been conducted by an international, multidisciplinary research group during more than one decade; see [7–9,12,15,16,18–25,27,29–31,33–35,37–43,46,48–51], and references therein.

The BS distribution is asymmetric, unimodal and has two parameters modifying its shape and scale, this last one being besides its median. The BS distribution has attractive properties, being one of them its relation with the normal or Gaussian model; see [32]. Indeed, the BS distribution is similar to the normal distribution, but in an asymmetric setting, where the median, one of the BS parameters, is known to be a better centrality indicator than the mean. Therefore, the BS model has been quite effective to describe data following distributions skewed to the right. The BS model has good properties which have allowed it to be widely applied to different areas, as mentioned, but its main applications are found in problems of material fatigue. Consequently, in this kind of problems is where the BS distribution should provide a better fitting, due to its genesis from the cumulative damage law; see [32].

The main objective of this paper is to propose a methodology using a new type of BS distribution based on the modified skew-normal (MSN) distribution studied by Arrué *et al.* [3]. This allows us to obtain flexibility in skewness and kurtosis, which is controlled by an additional shape parameter. Thus, the new type of BS distribution has three parameters which modify its shape and scale. We estimate the three model parameters with the ML method. Performance of these estimators and their bias correction (reduction) is evaluated by Monte Carlo simulations. The numerical results obtained in this investigation were computed with `R`, an open source statistical software for graphics and data analysis that is widely used by the international scientific community; see www.R-project.org and R Core Team [44]. In the following subsection, we present an example from material fatigue to motivate our study.

1.2. Motivating example from material fatigue and description of contents

Aluminium is a widely used material in engineering. Some properties of aluminium 6061-T6 are: ductility, high strength, malleability, resistance to corrosion and tenacity, among others. These characteristics make this material to be commonly used for diverse purposes. For example, to built camera lenses; aircraft and spacecrafts components; electrical accessories; utensils for transportation; vehicle brakes components and transmission shafts. Fatigue is a feature that can cause failure of aluminium specimens. In many engineering studies, it is necessary to consider probability models for describing aluminium fatigue life data. An important aspect in parametric lifetime analysis is to identify the appropriate life distribution. Most of the probabilistic models used to describe lifetime data are chosen based on one or more of the following three aspects: (A1) a theoretical argument for the specimen's mechanism of failure; (A2) a model that has previously been used with a successful result; and (A3) an appropriate model that fits the data well. A frequent problem detected in (A3) is that life tests generate little data that represent the failure of the specimen. This is generally due to the time horizon of the life test, which is often short, or lack of financial resources to wait for all specimens to fail. In this type of data, it has been verified that some life distributions of two parameters, that have certain flexibility, fit reasonably well in the central region, because usually there is a suitable number of data in this region of the distribution. However, in lifetime data analysis sometimes the interest is in the low or high quantiles of the distribution, for example, to establish lethal doses, policies of maintenance, safe life, or warranties. Indeed, the fit of life distributions in their tails is often inaccurate. This is because of the lack of observations in the tails, which conducts to a wide discrepancy among probability models of similar characteristics. Such a situation makes it difficult to identify an appropriate life distribution through visual and goodness-of-fit tests, which is a problem because these methods are often considered to decide what distribution is going to be used. Thus, it is better to begin the selection of the candidate distribution with a theoretical argument. For example, an argument of 'fatigue' or 'cumulative damage' justifies the BS distribution. In spite of the aforementioned, we cannot only consider theoretical arguments, since the argued models can also be appropriate for other applications. In summary, when we must choose one or several probability models as candidates, the life distribution that is selected to solve the problem must be logical and, in addition, supported by graphical tools and statistical tests for fitting the data and/or criteria of model selection. It is important to have in mind that due to the time horizon above mentioned, censored data can also be present. We do not consider it in our study because it is beyond the objective of the present investigation. However, some comments about future studies related to censored data are mentioned in the final section.

The motivation for our study came from a real-world fatigue data set analyzed by Birnbaum and Saunders [11] and Martínez-Flórez *et al.* [41], corresponding to fatigue life in cycles $\times 10^{-3}$ of $n = 101$ aluminium specimens of 6061-T6 type, which were cut in parallel to the rotation direction at a speed of 18 cycles per second, with a maximum stress of 21,000 psi. We refer to these data as 'fatigue'. Table 1 reports basic statistics for the data under analysis related to minimum and maximum values, mean, median, standard deviation (SD) as well as coefficients of variation (CV), skewness (CS) and kurtosis (CK). Figure 1 shows the histogram and box-plots (standard and adjusted for asymmetry) using fatigue data. For details about the box-plot adjusted for asymmetric data, see [45].

First, it seems natural to model 'fatigue' data by the well-known normal distribution, because it is a widely studied distribution with good properties. In addition, the normal distribution of mean μ and variance σ^2 –denoted by $\mathrm{N}(\mu, \sigma^2)$– is supported by the

Table 1. descriptive summary for fatigue data.

n	Min	Max	Mean	Median	SD	CV	CS	CK
101	370	2440	1400.84	1401	391.01	27.912%	0.14	2.72

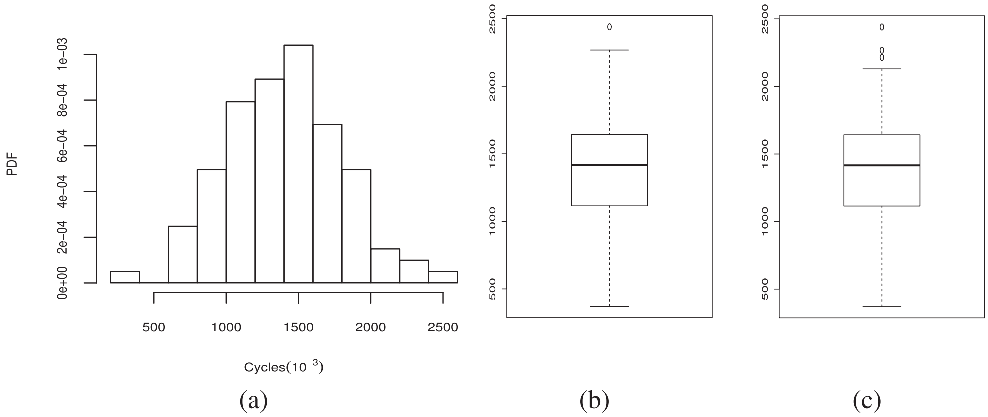


Figure 1. Histogram (a), standard box-plot (b) and adjusted box-plot (c) for fatigue data.

descriptive summary presented in Table 1 and the histogram and box-plots of Figure 1 for the fatigue data, which reveals us this distribution can be indeed suitable. The slight asymmetry observed in the histogram is confirmed by the CS reported in Table 1. The adjusted box-plot detects some atypical values that the standard box-plot does not detect. However, from Figure 2(a) and based on the theoretical quantile versus empirical quantile (QQ) plot, note that the normal distribution (with $\hat{\mu} = 1400.738, \hat{\sigma} = 389.135$) does not capture extreme data well, as it is known. Second, since the data represent material fatigue, it seems reasonable to model them with the BS distribution of shape parameter α and scale parameter β –denoted by BS(α, β)–; see details in [32]. The result of this fitting (with $\hat{\alpha} = 0.310, \hat{\beta} = 1336.563$) is presented in Figure 2(b), which is slightly better than the normal fitting mainly at the left tail. Nevertheless, the extreme data located at the right tail are not captured correctly by the BS distribution. Third, another alternative to improve the fitting is to use one of the most important asymmetric models in the literature, which is known as the skew-normal (SN) distribution of mean μ, variance σ^2 and asymmetry parameter λ_1 –denoted by SN$(\mu, \sigma^2, \lambda_1)$–, which was introduced by Azzalini [6]; see details about the SN distribution and its inference in Section 2.2. The new fitting

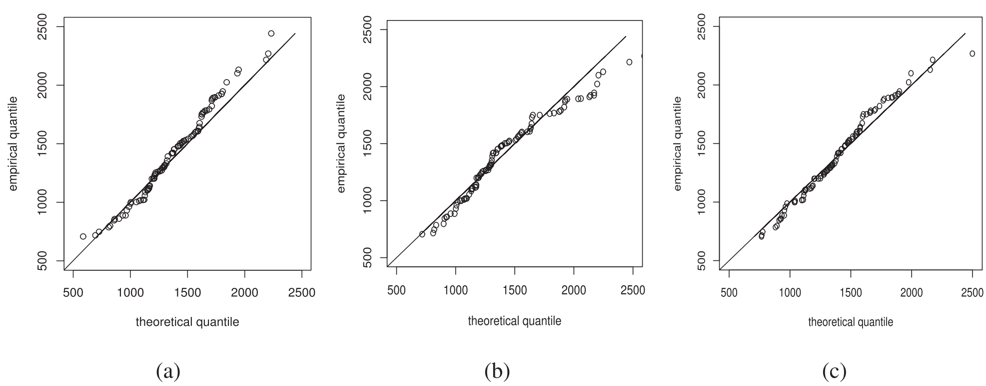


Figure 2. QQ plot for normal (a), BS (b) and SN (c) models with fatigue data.

using the SN distribution (with $\hat{\mu} = 1119.884$, $\hat{\sigma} = 479.982$, $\hat{\lambda}_1 = 1.077$) is presented in Figure 2(c), which once again is slightly better now at the right tail, but it is not totally satisfactory. In summary, it is necessary to consider a model to describe fatigue data which can accommodate asymmetry and kurtosis well. The three considered models are not totally suitable and their usage can distort the results obtained from the data analytics. Therefore, this example serves as a motivation to formulate a model that allows us to describe fatigue data adequately, which should be based on theoretical arguments useful in material fatigue, and to fit asymmetry and kurtosis well.

The rest of the paper proceeds as follows. Section 2 provides a background of BS, SN and MSN distributions. In Section 3, we formulate a new type of BS model, that we call the modified skew-normal Birnbaum-Saunders (MSNBS) distribution, state some of its mathematical properties, and conduct a shape analysis. Then, in Section 4, inference and estimation based on the ML method are derived for the parameters of the new model. In addition, in this section, we propose bias reduction on the corresponding ML estimators. Afterwards, in Section 5, the estimation methodology is evaluated through Monte Carlo simulations and the new model is illustrated with the fatigue data presented in the motivating example. Lastly, some concluding remarks and possible future research are mentioned in Section 6.

2. Preliminaries

In this section, we present a background related to the BS and MSN distributions.

2.1. The BS distribution

Let $X \sim \mathrm{BS}(\alpha, \beta)$. Then, the probability density function (PDF) and cumulative distribution function (CDF) of X are, respectively, given by

$$f_X(x; \alpha, \beta) = \phi(a_x(\alpha, \beta))\frac{x^{-3/2}(x+\beta)}{2\alpha\sqrt{\beta}}, \quad F_X(x; \alpha, \beta) = \Phi\left(a_x(\alpha, \beta)\right), \quad x > 0, \tag{1}$$

where $a_x(\alpha, \beta) = b(x)/\alpha$, with $b(x) = (\sqrt{x/\beta} - \sqrt{\beta/x})$, and ϕ, Φ are the standard normal PDF and CDF, respectively. The BS distribution admits the stochastic representation

$$X = \beta\left(\alpha Z/2 + \left((\alpha Z/2)^2 + 1\right)^{1/2}\right)^2, \tag{2}$$

where $Z \sim \mathrm{N}(0, 1)$. Note that the mode of X –denoted by x_{m}– is given by the solution of $(\beta - x_{\mathrm{m}})(x_{\mathrm{m}} + \beta)^2 = \alpha^2\beta x_{\mathrm{m}}(x_{\mathrm{m}} + 3\beta)$, which does not have a closed analytical form. Some properties of $X \sim \mathrm{BS}(\alpha, \beta)$ are: (i) $cX \sim \mathrm{BS}(\alpha, c\beta)$, for $c > 0$; and (ii) $1/X \sim \mathrm{BS}(\alpha, 1/\beta)$. The first property indicates that the BS distribution belongs to the scale family. The second property tells us that the BS distribution is closed under reciprocation. The survival function and hazard rate of the BS distribution are defined, in general, as $S_X(x) = 1 - F_X(x)$ and $r_X(x) = f_X(x)/S_X(x)$, for $x > 0$, respectively, where $f_T(t) = F_T'(t)$, that is, $f_T(t)$ is the derivative of $F_T(t)$ with respect to t. Observe that the BS hazard rate is upside-down bath shaped, with its change point –denoted by x_{c}– being a decreasing function of α and obtained from the solution of $\Phi(-a_x(\alpha, \beta))(\alpha^2 b''(x) -$

$(b'(x))^2 b(x)) + \alpha\phi(-a_x(\alpha,\beta))(b'(x))^2 = 0$, which does not have a closed analytical form, where $a_x(\alpha,\beta), b(x)$ are given from (1) and $b'(x), b''(x)$ are, respectively, the first and second derivatives of $b(x)$ with respect to x. In addition, $\lim_{x\to 0} r_X(x) = 0$ and $\lim_{x\to\infty} r_X(x) = 1/(2\alpha^2\beta)$. For more details about BS hazard rate, see [4,28].

Next, we propose a family of flexible distributions which can be useful for solving the problem presented in the motivating example of Section 1.2. This new family contains the BS distribution as a particular case and is based on the MSN distribution, which is detailed in the following subsection.

2.2. The modified skew-normal distribution

A random variable V is standard SN distributed with parameter λ_1 –denoted by $V \sim \mathrm{SN}(\lambda_1)$– if its PDF is given by

$$f_V(v;\lambda_1) = 2\phi(v)\Phi(\lambda_1 v), \quad v \in \mathbb{R},\ \lambda_1 \in \mathbb{R}, \tag{3}$$

where ϕ and Φ are, as mentioned, the standard normal PDF and CDF, respectively. When $\lambda_1 = 0$ in (3), the normal distribution is obtained. For values other than zero, the parameter λ_1 controls the asymmetry of the model, making it more flexible than the normal distribution. Indeed, in the motivating example of Section 1.2, the SN distribution shows a better performance than the normal model. However, the SN distribution does not model well some atypical data of this example presented in its right tail. Another aspect of the SN distribution to be taken into account, which can be considered as a disadvantage, is that its Fisher information matrix is non-singular only when the asymmetry parameter is $\lambda_1 \neq 0$. However, when $\lambda_1 = 0$, we are unable to perform standard asymptotic inference due to this disadvantage. Following Arellano-Valle *et al.* [1], a model with similar characteristics to the SN distribution was proposed by Arrué *et al.* [3], which corresponds to the MSN formulation. One of the advantages of this formulation is that the Fisher information matrix does not present the singularity problem that the SN distribution has; see also [2]. A random variable W follows an MSN distribution with parameter λ –denoted by $W \sim \mathrm{MSN}(\lambda)$– if its PDF is defined as

$$f_W(w,\lambda) = 2\phi(w)\Phi(\lambda u(w)), \quad w \in \mathbb{R},\ \lambda \in \mathbb{R}, \tag{4}$$

where $u(w) = w/\sqrt{1+w^2}$. Therefore, it is of interest to propose new models that can capture asymmetry and kurtosis in a flexible way, as well as being able to contain the BS model as a particular case and not as a limiting case. This allows the new model to be proposed in a context of material fatigue and to solve problems as that presented in the motivating example of Section 1.2. For the aforementioned, we propose in Section 3 the MSNBS distribution by considering in the stochastic representation for $X \sim \mathrm{BS}(\alpha,\beta)$ defined in (2) that $Z \sim \mathrm{N}(0,1)$ is replaced by $W \sim \mathrm{MSN}(\lambda)$.

3. The new type of BS distribution

In this section, we present some mathematical features and a shape analysis of the MSNBS distribution.

3.1. Probability functions

We state that a random variable T has an MSNBS distribution if we replace $X \sim \mathrm{BS}(\alpha, \beta)$ in the stochastic representation given in (2) by $T \sim \mathrm{MSNBS}(\alpha, \beta, \lambda)$, whereas $Z \sim \mathrm{N}(0, 1)$ is replaced by $W \sim \mathrm{MSN}(\lambda)$. Then, the PDF of T is expressed as

$$f_T(t; \alpha, \beta, \lambda) = 2\phi(a_t(\alpha, \beta))\Phi\left(\lambda u(a_t(\alpha, \beta))\right) \frac{t^{-3/2}}{2\alpha\sqrt{\beta}}(t + \beta),$$
$$t > 0, \quad \alpha > 0, \quad \beta > 0, \quad \lambda \in \mathbb{R}, \tag{5}$$

where $a_t(\alpha, \beta)$ is given analogously as in (1) and $u(w) = w/\sqrt{1 + w^2}$. The survival function and hazard rate of the MSNBS distribution are defined, respectively, as

$$S_T(t) = 1 - F_T(t), \quad r_T(t) = \frac{f_T(t)}{S_T(t)}, \quad t > 0,$$

where $F_T(t)$ is obtained in Proposition 3.2, whereas convergence of $r_T(t)$ is presented in Proposition 3.1. Both proofs are provided in Appendix 1.

Proposition 3.1: *Let $T \sim MSNBS(\alpha, \beta, \lambda)$. Then,*

$$\lim_{t \to \infty} r_T(t) = \frac{1}{2\alpha^2\beta}.$$

Figure 3(a)–(d) shows shapes of the MSNBS PDF for different values of α, β and λ. Note that the MSNBS PDF has high flexibility shapes when changing the values of α and λ. Figure 3(e,f) displays shapes of the MSNBS hazard rate with unimodal behavior for different values of α and λ. Observe that unimodality of the MSNBS PDF and hazard rate cannot be obtained in a closed analytical form, but their non-linear equations may be generated in a similar way as in the BS case presented in Section 2.1.

Remark 3.1: Note that if $\lambda = 0$, the MSNBS PDF given in (5) reduces to the BS PDF given in (1), that is, the MSNBS model contains the BS distribution as particular case when the parameter λ is zero, inheriting some of the properties and applications that the BS distribution has. Therefore, the MSNBS model is an extension of the BS model by incorporating the parameter λ, which in turn affects its asymmetry and kurtosis, making the BS model flexible. We can use the likelihood ratio test to detect significant differences between the BS and MSNBS models when fitting a data set, as well as performing asymptotic inference when $\lambda = 0$, since the Fisher information matrix is non-singular in that case.

3.2. Properties and moments

Some properties and moments of the MSNBS distribution are provided next.

Proposition 3.2: *Let $T \sim MSNBS(\alpha, \beta, \lambda)$ and $W \sim MSN(\lambda)$. Then, the CDF of T is $F_T(t; \alpha, \beta, \lambda) = F_W(a_t(\alpha, \beta); \lambda)$, for $t > 0$, where F_W is the MSN CDF obtained from the PDF given in* (4).

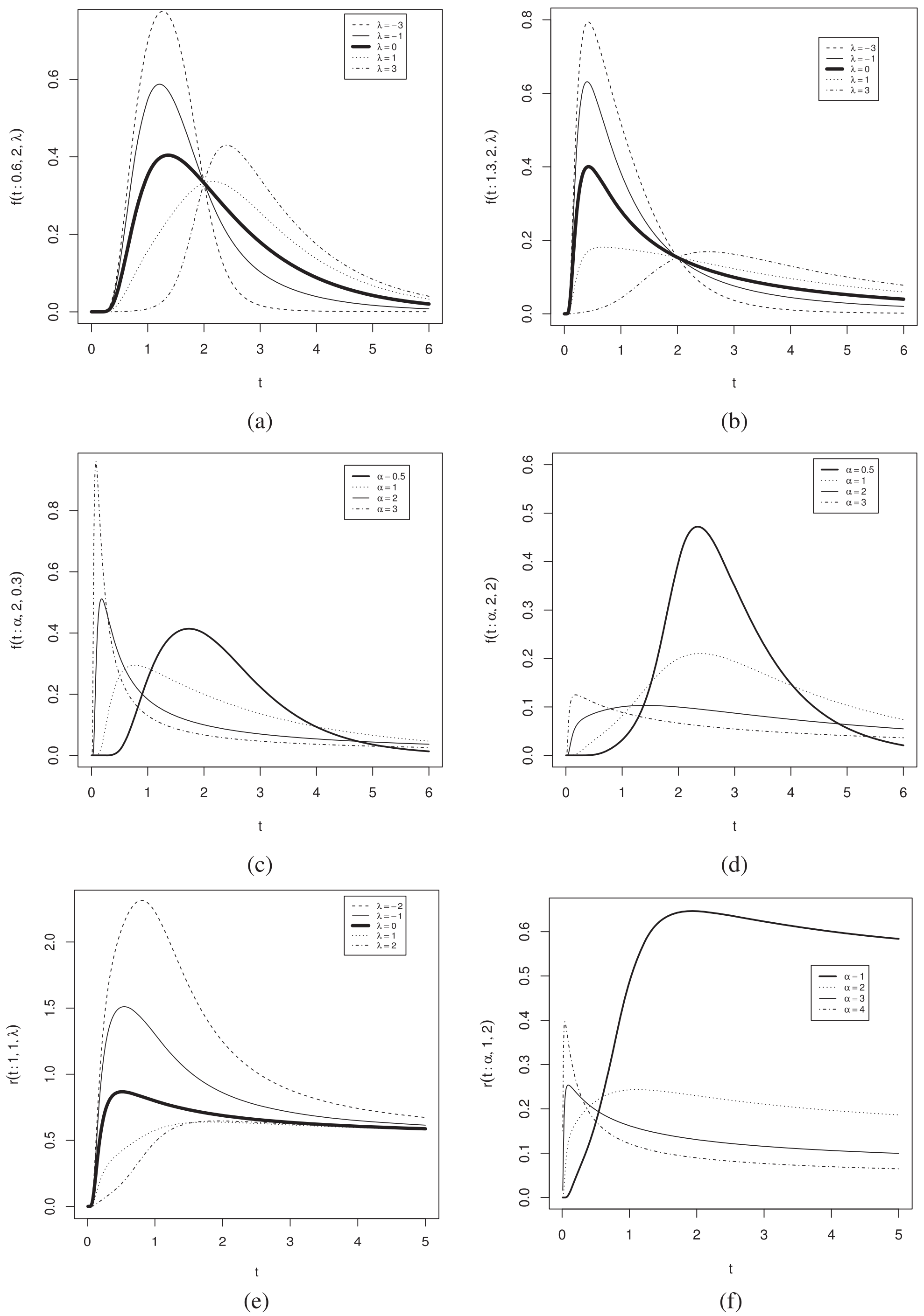
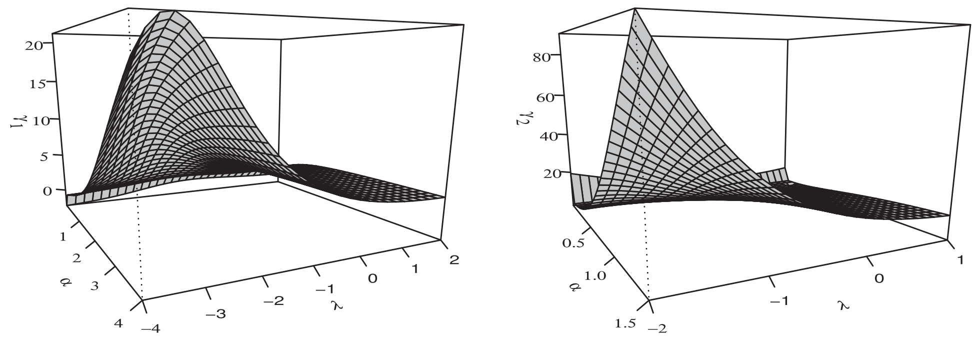


Figure 3. plots of the MSNBS PDF (a)–(d) and hazard rate (e,f).



(a) (b)

Figure 4. plots of the CS γ_1 (a) and CK γ_2 (b) of the MSNBS model.

Proposition 3.3: *Let $T \sim MSNBS(\alpha, \beta, \lambda)$ and $W \sim MSN(\lambda)$. Then, the pth percentile of T is*

$$t(p; \alpha, \beta, \lambda) = \beta \left(\alpha w_p/2 + \left((\alpha w_p/2)^2 + 1 \right)^{1/2} \right)^2,$$

where w_p is the pth percentile of the MSN distribution, given by $w_p = F_W^{-1}(p; \lambda)$, with F_W^{-1} being the inverse function of the MSN CDF F_W.

Proposition 3.4: *Let $T \sim MSNBS(\alpha, \beta, \lambda)$ and $W \sim MSN(\lambda)$. Then, the rth moment of T is*

$$\mathrm{E}[T^r] = 2\beta^r \sum_{i=0}^{r} C_{2i}^{2r} \sum_{k=0}^{i} C_k^i \left(\frac{\alpha}{2}\right)^{2(r-k)} (2(r-k)-1)!! + 2\beta^r \sum_{i=0}^{r-1} C_{2i+1}^{2r} H_i,$$

where C_m^n is the combinatorial number of n over m, $H_i = (1/2)\mathrm{E}[h_i(W)]$, with $h_i(z) = (\alpha z/2)^{2r-(2i+1)}((\alpha z/2)^2 + 1)^{(2i+1)/2}$, and $(2(r-k)-1)!! = (2(r-k)-1) \times (2(r-k) - 3) \times \cdots \times 1$.

In order to obtain the expressions of the CV and CK, we must calculate the central moments μ'_r using the non-central moments μ_r, for $r = 2, 3, 4$. To achieve this, we use the relations $\mu'_2 = \mu_2 - \mu_1^2$, $\mu'_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3$ and $\mu'_4 = \mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4$. These relations allow the CS and CK to be calculated, which are given by $\gamma_1 = \mu'_3/(\mu'_2)^{3/2}$ and $\gamma_2 = \mu'_4/(\mu'_2)^2 - 3$, respectively. From Figure 4(a), note that γ_1 converges to $2\sqrt{2}$ as $\lambda \to +\infty$ and $\alpha \to \infty$. However, when $\lambda \to -\infty$ and $\alpha \to +\infty$, γ_1 increases indefinitely. If $\lambda \to \pm\infty$ and $\alpha \to 0$, then $\gamma_1 \to -\infty$. Figure 4(b) shows that γ_2 converges to 15 as $\lambda \to +\infty$ and $\alpha \to \infty$. Nevertheless, as $\lambda \to -\infty$ and $\alpha \to \infty$, γ_1 increases indefinitely, similarly when $\lambda \to \pm\infty$ and $\alpha \to 0$. In contrast, the minimum value of γ_2 is 1.95.

4. Estimation and inference

In this section, we perform inference and estimation based on the ML method for the MSNBS model parameters. In addition, we propose bias reduction of the corresponding estimators.

4.1. Estimation and information matrices

Let $T_1, \ldots, T_n$ be a random sample of size n from $T \sim \mathrm{MSNBS}(\boldsymbol{\theta})$, with $\boldsymbol{\theta} = (\alpha, \beta, \lambda)$. The log-likelihood function for $\boldsymbol{\theta}$ is given by

$$\ell(\boldsymbol{\theta}) = c_0 - n\log(\alpha) - \frac{n}{2}\log(\beta) - \frac{1}{2}\sum_{i=1}^{n} a_i^2 - \frac{3}{2}\sum_{i=1}^{n}\log(t_i) + \sum_{i=1}^{n}\log(t_i+\beta) + \sum_{i=1}^{n}\log(\Phi(\lambda u(a_i))),$$

where c_0 is a constant which does not depend on $\boldsymbol{\theta}$, $a_i = a_{t_i}(\alpha,\beta) = (1/\alpha)(\sqrt{t_i/\beta} - \sqrt{\beta/t_i})$, with $t_i > 0$, for $i = 1, \ldots, n$, $\alpha > 0$, $\beta > 0$ and $\lambda \in \mathbb{R}$. Then, the corresponding score vector has elements defined as

$$S_\alpha = \frac{n}{\alpha}\left(\overline{a^2} - 1 - \lambda\bar{\rho}_{103}\right), \quad S_\lambda = n\bar{\rho}_{101},$$

$$S_\beta = n\left(-\frac{1}{2\beta} + \frac{1}{2n\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2} - \frac{1}{t_i}\right) + \frac{1}{n}\sum_{i=1}^{n}\frac{1}{t_i+\beta} - \frac{\lambda}{2\alpha\beta}\bar{\rho}_{013}\right),$$

where

$$\bar{\rho}_{npm} = \frac{1}{n}\sum_{i=1}^{n}\frac{a_i^n \widetilde{w}_i^p \zeta_i}{(1+a_i^2)^{m/2}}, \quad \overline{a^2} = \frac{1}{n}\sum_{i=1}^{n} a_i^2,$$

with

$$\zeta_i = \zeta(\lambda u(a_i)) = \frac{\phi(\lambda u(a_i))}{\Phi(\lambda u(a_i))}, \quad \widetilde{w}_i = \sqrt{\frac{t_i}{\beta}} + \sqrt{\frac{\beta}{t_i}}, \quad u(x) = \frac{x}{(1+x^2)^{1/2}}.$$

Resolving the system given by the score functions after equating them to zero, we obtain the simplified equation system given by

$$-\frac{1}{2\beta} + \frac{1}{2n\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2} - \frac{1}{t_i}\right) + \frac{1}{n}\sum_{i=1}^{n}\frac{1}{t_i+\beta} = \frac{\hat{\lambda}}{2\alpha\beta}\bar{\rho}_{013},$$

where $\overline{a^2} - 1 = \lambda\bar{\rho}_{103}$ and $\bar{\rho}_{101} = 0$, whose solution, based on numerical methods, gives us the ML estimates of the model parameters. The elements of the observed information matrix, $J_{\theta_i\theta_j} = -\partial^2\ell(\boldsymbol{\theta})/\partial\theta_i\partial\theta_j$ (see Appendix 2), are given by

$$J_{\alpha\alpha} = \frac{n}{\alpha^2}\left(1 - 3\overline{a^2} + \lambda(\bar{\rho}_{103} + \bar{\rho}_{105} - 2\bar{\rho}_{305}) - \lambda^3\bar{\rho}_{307} - \lambda^2\bar{\eta}_{203}\right),$$

$$J_{\beta\alpha} = \frac{n}{2\alpha^2\beta}\left(-\frac{2}{n\alpha}\sum_{i=1}^{n}\left(\frac{t_i}{\beta} - \frac{\beta}{t_i}\right) + \lambda(\bar{\rho}_{015} - 2\bar{\rho}_{215}) - \lambda^3\bar{\rho}_{217} - \lambda^2\bar{\eta}_{113}\right),$$

$$J_{\alpha\lambda} = \frac{n}{\alpha}\left(-\bar{\rho}_{103} + \lambda^2\bar{\rho}_{305} + \lambda\bar{\eta}_{202}\right),$$

$$J_{\beta\beta} = n\left(\frac{1}{2\beta^2} - \frac{1}{n\alpha^2\beta^3}\sum_{i=1}^{n} t_i - \frac{1}{n}\sum_{i=1}^{n}\frac{1}{(t_i+\beta)^2} + \frac{1}{4\alpha^2\beta^2}\left(\lambda(2\alpha\bar{\rho}_{013} + \alpha^2\bar{\rho}_{103})\right.\right.$$
$$\left.\left. - \lambda^3(\alpha^2\bar{\rho}_{307} + 4\bar{\rho}_{107}) - \lambda^2(\alpha^2\bar{\eta}_{203} + 4\bar{\eta}_{003}) - 3\lambda(\alpha^2\bar{\rho}_{305} + 4\bar{\rho}_{105})\right)\right),$$

$$J_{\lambda\beta} = \frac{n}{2\alpha\beta}\left(-\bar{\rho}_{013} + \lambda^2\bar{\rho}_{215} + \lambda\bar{\eta}_{112}\right), J_{\lambda\lambda} = -n(\lambda\bar{\rho}_{303} + \bar{\eta}_{201}),$$

where $\bar{\eta}_{npm} = (1/n)\sum_{i=1}^{n}(a_i^n\widetilde{w}_i^p\zeta_i^2)/(1+a_i^2)^m$. The elements of the expected Fisher information matrix, $I_{\theta_i\theta_j} = -\mathrm{E}[S_iS_j]$ (see Appendix 2), are obtained as

$$I_{\alpha\alpha} = \frac{1}{\alpha^2}\left(2 + \lambda^2\eta_{203}\right), \quad I_{\lambda\beta} = -\frac{1}{2\alpha\beta}(-\rho_{013} + \lambda^2\rho_{215} + \lambda\eta_{112}),$$
$$I_{\alpha\lambda} = -\frac{1}{\alpha}\lambda\eta_{202}, \quad I_{\lambda\lambda} = \eta_{201},$$
$$I_{\beta\alpha} = -\frac{1}{2\alpha^2\beta}\left(-\frac{2}{\alpha}\left(\frac{\mathrm{E}[T]}{\beta} - \beta\mathrm{E}\left[\frac{1}{T}\right]\right) + \lambda(\rho_{015} - 2\rho_{215}) - \lambda^3\rho_{217} - \lambda^2\eta_{113}\right),$$
$$I_{\beta\beta} = -\frac{1}{2\beta^2} + \frac{\mathrm{E}[T]}{\alpha^2\beta^3} + \mathrm{E}\left[\frac{1}{(T+\beta)^2}\right] + \frac{1}{4\alpha^2\beta^2}(\lambda^2(\alpha^2\eta_{203} - 2\lambda\alpha\rho_{013} + 4\eta_{003})),$$

where $\rho_{npm} = \mathrm{E}[(a^n\tilde{w}^p\zeta)/(1+a^2)^{m/2}]$ and $\eta_{npm} = \mathrm{E}[(a^n\tilde{w}^p\zeta^2)/(1+a^2)^m]$, with $\rho_{npm} = 0$ if n is odd and $p = 0$. For the case $\lambda = 0$, we have

$$\boldsymbol{I}(\alpha,\beta,0) = \begin{pmatrix} \frac{2}{\alpha^2} & 0 & 0 \\ 0 & d_{22} & d_{23} \\ 0 & d_{23} & d_{33} \end{pmatrix}, \tag{6}$$

where

$$d_{23} = 2\sqrt{\frac{2}{\pi}}\int_0^\infty \frac{\sqrt{1+(\frac{\alpha z}{2})^2}}{(1+z^2)^{3/2}}\phi(z)\,\mathrm{d}z, \quad d_{33} = \frac{2}{\pi}(1 - (2\pi)^{1/2}\exp(1/2)\Phi(-1)),$$
$$d_{22} = \frac{1}{\alpha^2\beta^2} + \frac{1}{\sqrt{2\pi}\beta^2}\left(\sqrt{\frac{\pi}{2}} - \frac{\pi\exp(2/\alpha^2)(1-\Phi\left(\frac{2}{\alpha}\right))}{\alpha}\right).$$

Note that the matrix $\boldsymbol{I}(\alpha,\beta,0)$ defined in (6) is non-singular, which allows us to apply the standard asymptotic theory for testing the hypotheses H_0: $\lambda = 0$ (BS model) vs. H_1: $\lambda \neq 0$ (MSBBS model).

4.2. *Bias reduction*

In many cases, the ML estimate of λ overestimates its true value and, in some of these cases, it can be infinite with some probability. Particularly, when all terms $a_{t_i} > 0$, meaning $\min(t_i) > \hat{\beta}$, the profile of the log-likelihood function, defined by $\ell_{\text{profile}} = \ell(\hat{\alpha},\hat{\beta},\lambda)$, where $\hat{\alpha}$, $\hat{\beta}$ are the ML estimates with the value of λ being fixed, is increasing monotone. Therefore, we obtain an infinite ML estimate; see [36]. However, the probability of infinite estimates decreases rapidly as the sample size increases. In addition, as $\min(t_i) > \hat{\beta}$ is equivalent to $a_{t_i} > 0$, with $i = 1, \ldots, n$, and $W \sim \text{MSN}(\lambda)$, which implies that $-W \sim \text{MSN}(-\lambda)$, we focus our study only on positive values for λ. In this case, the bias of the ML estimator of λ is positive, and the ML estimators of the parameters α and β have small biases. We apply here the method proposed by Firth [17] to reduce the bias of the parameter λ, used by Sartori [47] for the SN distribution, and by Arrué *et al.* [3] for the MSN distribution. The bias of the ML estimator $\hat{\lambda}$ has order $O(n^{-1})$; see [14]. However, the bias of the modified estimator $\widehat{\lambda^\star}$ has order $O(n^2)$; see [17]. The solution to the equation of modified profile log-likelihood function is given by

$$U^\star_{\text{profile}}(\lambda) = U_{\text{profile}}(\lambda) + M(\lambda) = 0, \tag{7}$$

where U_{profile} is the profile of the score function defined as $U_{\text{profile}}(\lambda) = \sum_{i=1}^{n} u(\tilde{a}_i)\, \zeta(\lambda u(\tilde{a}_i))$, with $\tilde{a}_i = (1/\hat{\alpha})((t_i/\hat{\beta})^{1/2} - (\hat{\beta}/t_i)^{1/2})$, where $M(\lambda) = -\lambda a_{42}(\lambda)/(2a_{22}(\lambda))$ is a modified function, with $a_{kh}(\lambda) = E_\lambda(u(\tilde{a}_i)^k \zeta^h(\lambda u(\tilde{a}_i)))$ and $\zeta(x) = \phi(x)/\Phi(x)$, where $a_{k1} = 0$ when k is odd and $a_{kh} \geq 0$ when k, h are even. As $\hat{\lambda}^\star$ follows an asymptotically normal distribution with variance equal to $1/I_{\lambda\lambda}$, obtaining a symmetric confidence interval (CI) around λ, it might possibly contain negative values of λ, which does not always adjust to reality. In order to better reflect the symmetry for the CI, we introduce the concept of profile quasi-likelihood function associated with $U^\star_{\text{profile}}(\lambda)$ given in (7) and defined as

$$\ell^\star_{\text{profile}}(\lambda) = \int_{c_2}^{\lambda} U^\star_{\text{profile}}(t)\, \mathrm{d}t = \ell_{\text{profile}}(\lambda) - \ell_{\text{profile}}(c) + \int_{c_2}^{\lambda} M(t)\, \mathrm{d}t,$$

where c_2 is a real arbitrary value. Since $M(\lambda)$ is of order $O(1)$, hence $\ell^\star(\theta)$ is a penalized likelihood function, whose penalization is bounded. Thus, the log-likelihood ratio statistic based on $R^\star_{\text{profile}}(\lambda) = 2(\ell^\star_{\text{profile}}(\hat{\lambda}^\star) - \ell^\star_{\text{profile}}(\hat{\lambda}))$ maintains the usual asymptotic $\chi^2(1)$ distribution and it can be employed to calculate CIs for λ. We have that $M(\lambda)$ is bounded, which allows us to prove the existence of the estimator $\hat{\lambda}^\star$. This proof can be obtained almost directly from Arrué *et al.* [3].

5. Numerical evaluations

In this section, we evaluate the estimation methodology with Monte Carlo simulations and illustrate the new model with fatigue data presented in the motivating example (Section 1.2).

5.1. Monte Carlo simulations

We simulate data from the MSNBS(α, β, λ) distribution for different sample sizes and parameter values. Data are obtained by the stochastic representation given in (2) with $W \sim \text{MSN}(\lambda)$, where $W \mid S = s \sim \text{SN}(s)$ and $S \sim \text{N}(\lambda, 1)$; see [1]. From Table 2, note that λ is overestimated and also there are cases in which its estimate is infinity. This depends on the sample size and on the true value of the parameter. However, the estimation of parameters α and β is very good, since they are always finite and their biases are quite small. Firth

Table 2. Bias of the indicated estimator, empirical CP of 95% CIs based on $U^{\star}_{\text{profile}}(\lambda)$ and empirical percentage of cases when $\hat{\lambda}$ is finite, based on Monte Carlo simulations.

n	α	β	λ	Bias($\hat{\alpha}$)	Bias($\hat{\beta}$)	Bias($\hat{\lambda}$)^a	Bias($\hat{\lambda}^{\star}$)	CP	%($\hat{\lambda} < \infty$)
50	0.5	1	5	−0.0037	0.0035	1.8412	−0.7553	94.5	87.2
100	0.5	1	5	−0.0020	0.0017	1.2481	−0.3145	94.0	98.1
200	0.5	1	5	−0.0010	0.0005	0.5644	−0.0940	94.5	99.9
50	0.5	1	10	−0.0087	0.0082	1.6201	−3.4943	88.3	65.6
100	0.5	1	10	−0.0039	0.0026	3.4054	−1.4599	91.7	88.2
200	0.5	1	10	−0.0010	0.0006	2.8928	−0.3672	93.3	98.3
50	1	1	5	−0.0065	0.0145	2.1261	−0.8364	93.5	87.6
100	1	1	5	−0.0001	0.0033	1.5285	−0.2849	94.2	97.5
200	1	1	5	0.0012	0.0014	0.6865	−0.0948	94.7	100.
50	1	1	10	−0.0169	0.0173	1.4553	−3.6286	87.9	64.1
100	1	1	10	−0.0076	0.0050	3.3726	−1.5270	92.1	88.3
200	1	1	10	−0.0032	0.0019	2.9408	−0.4791	93.4	98.4
50	2	1	5	−0.0092	0.0788	2.0509	−1.2465	85.2	84.6
100	2	1	5	0.0085	0.0274	1.8715	−0.4878	92.1	96.9
200	2	1	5	0.0089	0.0045	0.8351	−0.1385	94.6	99.9
50	2	1	10	−0.0484	0.0550	1.4590	−4.1786	85.7	63.0
100	2	1	10	−0.0060	0.0134	3.7671	−1.6031	91.7	86.4
200	2	1	10	−0.0042	0.0069	2.9325	−0.6352	92.8	98.0

[a]Indicates this is computed when $\hat{\lambda} < \infty$.

[17] proposed to reduce the bias only for the parameter λ. The estimator of $\lambda^{\star}$ always exists and is finite. The bias reduction of λ is quite effective, considering the fact that this reduction applies when $\hat{\lambda}$ is finite and/or infinite. The empirical coverage probability (CP) of the corresponding CI is close to the nominal value (95%) and slightly less when the sample size is small, which is due to the high percentage of infinite values for $\hat{\lambda}$.

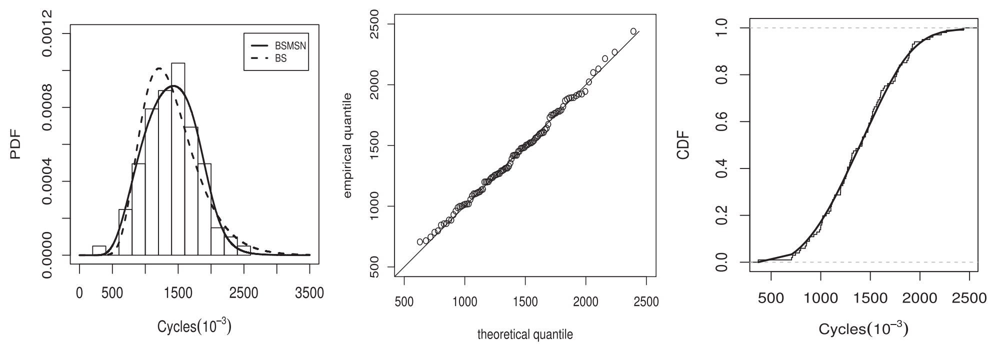
From Table 2, note that the biases of $\hat{\beta}$ and $\hat{\lambda}$ are always positive, which indicates that the corresponding log-likelihood function is increasing over (β, λ). However, although the bias of $\hat{\beta}$ is always positive, this bias is negligible even in small samples, so that estimates of β can be considered empirically unbiased. On the one hand, in the case of the bias of $\hat{\lambda}$, which is always positive as well, this sign for the bias can be justified on the fact that, in the underlying model, its parameters are generally overestimated by the ML method. On the other hand, the bias of $\hat{\lambda}^{\star}$ is negative due to this tends to underestimate the true value of λ, possibly explained by a high percentage of ML estimates of λ with infinite values. However, this bias approaches zero as the percentage of ML estimates of λ with infinite values decreases. When λ is zero (results omitted here for reasons of space), the bias of its estimator is very small (confirming that overestimation occurs for large values of λ) and fluctuating between positive and negative values, with a similar behavior for β. More simulation scenarios were considered but the results were similar so that we do not report them.

5.2. Motivating example (continuation)

Table 3 reports the ML estimates for the BS and MSNBS distribution parameters. The values in parentheses correspond to estimated asymptotic standard errors. Values of the log-likelihood function and Akaike information criterion (AIC) are also reported indicating an excellent agreement of the proposed model to the fatigue data; see Figure 5. In order

Table 3. ML estimates of indicated model parameters (with estimated standard errors in parenthesis) as well as values for the log-likelihood function and AIC with fatigue data.

Estimate	BS	MSNBS
$\hat{\alpha}$	0.310(0.035)	0.498(0.049)
$\hat{\beta}$	1336.563(93.860)	1951.042(84.172)
$\hat{\lambda}$	–	−4.046(1.190)
log-likelihood	−751.332	−746.574
AIC	1510.664	1505.149



(a) (b) (c)

Figure 5. histogram with fitted BS and MSNBS distributions (a), QQ-plot of the MSNBS model (b) and the empirical and theoretical MSNBS CDF (c) for fatigue data.

to confirm this, we test the hypotheses H_0: $\lambda = 0$ versus H_1: $\lambda \neq 0$. Then, using the log-likelihood ratio statistic given by $R = 2(\ell_{BS(\hat{\alpha},\hat{\beta})} - \ell_{MSNBS(\hat{\alpha},\hat{\beta},\hat{\lambda})})$, and replacing the ML estimates given in Table 3, we have $R = 9.516$, with p-value $= 0.002$. Therefore, we reject the null hypothesis at 5% of significance and we can say that the MSNBS distribution is a more adequate model for describing the fatigue data than the BS distribution. Figure 5 shows the data histogram with fitted BS and MSNBS PDFs (a), the QQ-plot of the MSNBS model (b), and the empirical and theoretical MSNBS CDF (c), obtained by the parameter estimates of every model. All of these graphical plots visually confirm the excellent fitting of the MSNBS distribution to the fatigue data. Table 4 provides the ML estimates $\hat{\mu}$, $\hat{\sigma}$, $\hat{\lambda}$ and the modified ML estimate $\hat{\lambda}^{\star}$, with their corresponding asymptotic standard errors (in parentheses) obtained from the associated Fisher information matrix, as well as the CIs of λ and $\lambda^{\star}$. We use the asymptotic $N_3(\boldsymbol{\theta}, \boldsymbol{I}(\boldsymbol{\theta})^{-1}/n)$ distribution of $\hat{\boldsymbol{\theta}}$ for obtaining these inferential results, where $\boldsymbol{I}(\boldsymbol{\theta})$ is given as in (6) and $\boldsymbol{\theta} = (\mu, \sigma, \lambda)$ or $\boldsymbol{\theta} = (\mu, \sigma, \lambda^{\star})$. Note that the CI of λ is less accurate than in the case of not applying bias reduction.

Table 4. ML estimate of the indicated parameter with standard errors in parentheses and log-likelihood value for fatigue data.

$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\lambda}^{\star}$	$\ell(\hat{\mu},\hat{\sigma},\hat{\lambda})$	$\ell(\hat{\mu},\hat{\sigma},\hat{\lambda}^{\star})$
0.498(0.047)	1951.236(80.390)	−4.043(1.131)	–	−746.574	–
0.498(0.050)	1951.236(90.789)	–	−3.342(0.913)	–	−746.880

6. Conclusions and future work

We derived a new type of Birnbaum-Saunders model based on the modified skew-normal distribution, which resulted to be widely flexible. Fisher information matrices associated with this new model were obtained, which showed to be non-singular when the parameter takes the value zero. This new model inherits the problem of overestimation of the asymmetry parameter from its source distribution, which was solved by a modified maximum likelihood estimator that always exists and is bounded, providing finite estimates. Simulations were conducted to show the overestimation of this parameter, as well as the performance of the bias reduction of its estimator. This reduction was not applied to the estimators of shape and scale parameters because their biases are quite small. In the simulations, we detected that the bias of the estimator of the asymmetry parameter is quite reduced. The empirical coverage probability of the corresponding confidence interval is very close to the nominal value, which is due to the high percentage of estimates of the asymmetry parameter equal to infinite. We used a profile quasi-likelihood function, so that the associated confidence intervals can better reflect the asymmetry of the log-likelihood function. An application with real-world fatigue data was conducted and employed to motivate this study. The new model and the bias reduction allowed us to show the potential applications of this research and its benefit, when analyzing fatigue data, by means of fitting suitable models to this kind of data.

Based on our results, new avenues are open to continue with other studies. For example, the analysis of censored data with and without covariates by regression can considered. Also, note that the Birnbaum-Saunders distribution is generated from a physical law, which allows us to obtain an approximation based on the central limit theorem generating this distribution from the standard normal model. Each Birnbaum-Saunders parameter is associated with parameters of this physical law; see details in [32]. Adding more parameters, for example when considering a normal distribution of mean μ and variance σ^2 instead of the standard normal model, can make the new distribution even more flexible. Observe that having more parameters (five in this case instead of three in the case of our new model) may improve the model fitting, because always a model with more parameters fits a curve better than one with less parameters. However, problems in the estimation might be introduced. Our conjecture is that biases of $\hat{\mu}$ and $\hat{\sigma}^2$ should be small, as in the case of the MSN(μ, σ^2, λ) model. Ideas on Birnbaum-Saunders distributions with more parameters than its original version were explored by Athayde *et al.* [5], but no inferential issues were analyzed there, so that it remains as an open problem to be explored. Furthermore, observe that the work of Cox and Reid [13] proposes diagonalizing the information matrix to obtain parameter orthogonality. Then, another work to be explored is applying this method to obtain orthogonal reparameterizations, which could conduct to improve results for the studied model in this research.

Acknowledgments

The authors would like to thank the editors and reviewers for their constructive comments on an earlier version of this manuscript which resulted in this improved version.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The research of J. Arrué and H.W. Gómez was supported by Grant SEMILLERO-UA-2019. The research of Víctor Leiva was supported partially by grant 'Fondo Nacional de Desarrollo Científico y Tecnológico (Fondecyt) 1160868' from the National Commission for Scientific and Technological.

ORCID

Víctor Leiva http://orcid.org/0000-0003-4755-3270

References

[1] R.B. Arellano-Valle, H.W. Gómez, and F.A. Quintana, *A new class of skew-normal distributions*, Comm. Statist. Theory Methods 33 (2004), pp. 1465–1480.
[2] R.B. Arellano-Valle, H.W. Gómez, and H.S. Salinas, *A note on the Fisher information matrix for the skew-generalized-normal model*, SORT 37 (2013), pp. 19–28.
[3] J. Arrué, R.B. Arellano-Valle, and H.W. Gómez, *Bias reduction of maximum likelihood estimates for a modified skew normal distribution*, J. Stat. Comput. Simul. 86 (2016), pp. 2967–2984.
[4] E. Athayde, A. Azevedo, M. Barros, and V. Leiva, *Failure rate of Birnbaum-Saunders distributions: Shape, change-point, estimation and robustness*, Braz. J. Probab. Stat. 33 (2019), pp. 301–328.
[5] E. Athayde, A. Azevedo, V. Leiva, and A. Sanhueza, *About Birnbaum-Saunders distributions based on the Johnson system*, Comm. Statist. Theory Methods 41 (2012), pp. 2061–2079.
[6] A. Azzalini, *A class of distributions which includes the normal ones*, Scand. J. Statist. 11 (1985), pp. 171–178.
[7] N. Balakrishnan, R. Gupta, D. Kundu, V. Leiva, and A. Sanhueza, *On some mixture models based on the Birnbaum-Saunders distribution and associated inference*, J. Stat. Plan. Inference 141 (2011), pp. 2175–2190.
[8] S. Balamurali, P. Jeyadurga, and M. Usha, *Optimal design of repetitive group sampling plans for Weibull and gamma distributions with applications and comparison to the Birnbaum-Saunders distribution*, J. Appl. Stat. 45 (2018), pp. 2499–2520.
[9] C. Bhatti, *The Birnbaum-Saunders autoregressive conditional duration model*, Math. Comput. Simul. 80 (2010), pp. 2062–2078.
[10] Z.W. Birnbaum and S.C. Saunders, *A new family of life distributions*, J. Appl. Probab. 6 (1969), pp. 319–327.
[11] Z.W. Birnbaum and S.C. Saunders, *Estimation for a family of life distributions with applications to fatigue*, J. Appl. Probab. 6 (1969), pp. 328–347.
[12] N.O. Castillo, H.W. Gómez, and H. Bolfarine, *Epsilon Birnbaum-Saunders distribution family: Properties and inference*, Statist. Papers 52 (2011), pp. 871–883.
[13] D.R. Cox and N. Reid, *Parameter orthogonality and approximate conditional inference*, J. R. Stat. Soc. B 49 (1987), pp. 1–39.
[14] D.R. Cox and E. Snell, *A general definition of residuals*, J. R. Stat. Soc. B 2 (1968), pp. 248–275.
[15] M.F. Desousa, H. Saulo, V. Leiva, and P. Scalco, *On a tobit-Birnbaum-Saunders model with an application to antibody response to vaccine*, J. Appl. Stat. 45 (2018), pp. 932–955.
[16] M. Ferreira, M.I. Gomes, and V. Leiva, *On an extreme value version of the Birnbaum-Saunders distribution*, Revstat Stat. J. 10 (2012), pp. 181–210.
[17] D. Firth, *Bias reduction of maximum likelihood estimates*, Biometrika 80 (1993), pp. 27–38.
[18] R.V. Fonseca and F. Cribari-Neto, *Bimodal Birnbaum-Saunders generalized autoregressive score model*, J. Appl. Stat. 45 (2018), pp. 2585–2606.
[19] F. Garcia-Papani, V. Leiva, F. Ruggeri, and M.A. Uribe-Opazo, *Kriging with external drift in a Birnbaum-Saunders geostatistical model*, Stoch. Environ. Res. Risk Assess. 32 (2018), pp. 1517–1530.

[20] F. Garcia-Papani, V. Leiva, M.A. Uribe-Opazo, and R.G. Aykroyd, *Birnbaum-Saunders spatial regression models: Diagnostics and application to chemical data*, Chemometr. Intell. Lab. Syst. 177 (2018), pp. 114–128.

[21] F. Garcia-Papani, M.A. Uribe-Opazo, V. Leiva, and R.G. Aykroyd, *Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data*, Stoch. Environ. Res. Risk Assess. 31 (2017), pp. 105–124.

[22] H.W. Gómez, J.F. Olivares-Pacheco, and H. Bolfarine, *An extension of the generalized Birnbaum-Saunders distribution*, Statist. Probab. Lett. 79 (2009), pp. 331–338.

[23] X. Guo, H. Wu, G. Li, and Q. Li, *Inference for the common mean of several Birnbaum-Saunders populations*, J. Appl. Stat. 44 (2017), pp. 941–954.

[24] M. Huerta, V. Leiva, S. Liu, M. Rodriguez, and D. Villegas, *On a partial least squares regression model for asymmetric data with a chemical application in minings*, Chemometr. Intell. Lab. Syst. 190 (2019), pp. 55–68.

[25] X. Jin and J. Kawczak, *Birnbaum-Saunders and lognormal kernel estimators for modelling durations in high frequency financial data*, Ann. Econ. Finance 4 (2003), pp. 103–124.

[26] N.L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, Vol. 2, Wiley, New York, US, 1995.

[27] S. Kotz, V. Leiva, and A. Sanhueza, *Two new mixture models related to the inverse Gaussian distribution*, Methodol. Comput. Appl. Probab. 12 (2010), pp. 199–212.

[28] D. Kundu, N. Kannan, and N. Balakrishnan, *On the hazard function of Birnbaum-Saunders distribution and associated inference*, Comput. Stat. Data Anal. 52 (2008), pp. 2692–2702.

[29] J. Leão, V. Leiva, H. Saulo, and V. Tomazella, *Birnbaum-Saunders frailty regression models: Diagnostics and application to medical data*, Biom. J. 59 (2017), pp. 291–314.

[30] J. Leão, V. Leiva, H. Saulo, and V. Tomazella, *A survival model with Birnbaum-Saunders frailty for uncensored and censored cancer data*, Braz. J. Probab. Statist. 32 (2018), pp. 707–729.

[31] J. Leão, V. Leiva, H. Saulo, and V. Tomazella, *Incorporation of frailties into a cure rate regression model and its diagnostics and application to melanoma data*, Stat. Med. 37 (2018), pp. 4421–4440.

[32] V. Leiva, *The Birnbaum-Saunders Distribution*, Academic Press, New York, US, 2016.

[33] V. Leiva, E. Athayde, C. Azevedo, and C. Marchant, *Modeling wind energy flux by a Birnbaum-Saunders distribution with unknown shift parameter*, J. Appl. Stat. 38 (2011), pp. 2819–2838.

[34] V. Leiva, C. Marchant, H. Saulo, M. Aslam, and F. Rojas, *Capability indices for Birnbaum-Saunders processes applied to electronic and food industries*, J. Appl. Stat. 41 (2014), pp. 1881–1902.

[35] C. Lillo, V. Leiva, O. Nicolis, and R.G. Aykroyd, *L-moments of the Birnbaum-Saunders distribution and its extreme value version: Estimation, goodness of fit and application to earthquake data*, J. Appl. Stat. 45 (2018), pp. 187–209.

[36] B. Liseo, *The skew-normal class of densities: Inferential aspects from a Bayesian viewpoint*, Statistica 50 (1990), pp. 59–70.

[37] C. Marchant, V. Leiva, and F.J.A. Cysneiros, *A multivariate log-linear model for Birnbaum-Saunders distributions*, IEEE Trans. Reliab. 65 (2016), pp. 816–827.

[38] C. Marchant, V. Leiva, G. Christakos, and M.A. Cavieres, *Monitoring urban environmental pollution by bivariate control charts: New methodology and case study in Santiago, Chile*, Environmetrics 30 (2019), p. e2551.

[39] C. Marchant, V. Leiva, F.J.A. Cysneiros, and S. Liu, *Robust multivariate control charts based on Birnbaum-Saunders distributions*, J. Stat. Comput. Simul. 88 (2018), pp. 182–202.

[40] C. Marchant, V. Leiva, F.J.A. Cysneiros, and J.F. Vivanco, *Diagnostics in multivariate generalized Birnbaum-Saunders regression models*, J. Appl. Stat. 43 (2016), pp. 2829–2849.

[41] G. Martínez-Flórez, H. Bolfarine, and H.W. Gómez, *An alpha-power extension for the Birnbaum-Saunders distribution*, Statistics 48 (2014), pp. 896–912.

[42] N.M. Olmos, G. Martínez-Flórez, and H. Bolfarine, *Bimodal Birnbaum-Saunders distribution with applications to non negative measurements*, Comm. Statist. Theory Methods 46 (2017), pp. 6240–6257.

[43] G.A. Paula, V. Leiva, M. Barros, and S. Liu, *Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance*, Appl. Stoch. Models Bus. Ind. 28 (2012), pp. 16–34.
[44] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017.
[45] P.J. Rousseeuw, C. Croux, V. Todorov, A. Ruckstuhl, M. Salibian-Barrera, T. Verbeke, M. Koller, and M. Maechler, *robustbase: Basic robust statistics*. R package version 0.92-6 (2016).
[46] M. Santos-Neto, F.J.A. Cysneiros, V. Leiva, and M. Barros, *Reparameterized Birnbaum-Saunders regression models with varying precision*, Electron. J. Stat. 10 (2016), pp. 2825–2855.
[47] N. Sartori, *Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions*, J. Stat. Plan. Inference 136 (2006), pp. 4259–4275.
[48] H. Saulo, J. Leao, V. Leiva, and R.G. Aykroyd, *Birnbaum-Saunders autoregressive conditional duration models applied to high-frequency financial data*, Statist. Papers (2019). Available at http://dx.doi.org/10.1007/s00362-017-0888-6.
[49] H. Saulo, V. Leiva, F.A. Ziegelmann, and C. Marchant, *A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data*, Stoch. Environ. Res. Risk Assess. 27 (2013), pp. 1479–1491.
[50] C. Villegas, G.A. Paula, and V. Leiva, *Birnbaum-Saunders mixed models for censored reliability data analysis*, IEEE Trans. Reliab. 60 (2011), pp. 748–758.
[51] P. Wanke and V. Leiva, *Exploring the potential use of the Birnbaum-Saunders distribution in inventory management*, Math. Probl. Eng. 2015(2015), pp. 1–9. Article ID 827246.

Appendix 1. Proofs

***Proof of Proposition* 3.1:** Let $T \sim \text{MSNBS}(\alpha, \beta, \lambda)$, with f_T, F_T being the PDF and CDF of T, respectively. Then, by applying L'Hopital rule, we obtain

$$\lim_{t\to\infty} h_T(t) = \lim_{t\to\infty} -\frac{f_T'(t)}{f_T(t)} = \frac{1}{2\alpha^2\beta}.$$

■

***Proof of Proposition* 3.2:** By considering

$$\begin{aligned} F_T(t;\alpha,\beta,\lambda) &= \int_0^t 2\phi(a_v(\alpha,\beta))\Phi\left(\lambda u(a_v(\alpha,\beta))\right)\frac{\mathrm{d}}{\mathrm{d}v}a_v(\alpha,\beta)\,\mathrm{d}v \\ &= \int_0^{a_t(\alpha,\beta)} 2\phi(w)\Phi(\lambda u(w))\,\mathrm{d}w = F_W(a_t(\alpha,\beta);\lambda), \end{aligned}$$

the proof is completed. ■

***Proof of Proposition* 3.3:** Let $T \sim \text{MSNBS}(\alpha, \beta, \lambda)$ and $W \sim \text{MSN}(\lambda)$. Then, the pth percentile of T is obtained by means of the stochastic representation given in (2) as

$$t(p;\alpha,\beta,\lambda) = \beta\left(\alpha w_p/2 + \left((\alpha w_p/2)^2 + 1\right)^{1/2}\right)^2,$$

where $w_p = F_W^{-1}(p;\lambda)$ is the pth percentile of the MSN distribution, with F_W^{-1} being the inverse function of the MSN CDF F_W. ■

Proof of Proposition 3.4**:** Let $T \sim \text{MSNBS}(\alpha,\beta,\lambda)$ and $W \sim \text{MSN}(\lambda)$. Then, by considering

$$\begin{aligned}
\mathrm{E}[T^r] &= \int_0^\infty t^r 2\phi(a_t(\alpha,\beta))\Phi\left(\lambda u(a_t(\alpha,\beta))\right)\frac{t^{-3/2}}{2\alpha\sqrt{\beta}}(t+\beta)\,\mathrm{d}t\\
&= \int_{-\infty}^{\infty} 2\beta^r\left(\frac{\alpha w}{2}+\sqrt{\left(\frac{\alpha w}{2}\right)^2+1}\right)^{2r}\phi(w)\Phi\left(\lambda u(w)\right)\mathrm{d}w\\
&= 2\beta^r\int_{-\infty}^{\infty}\left(\sum_{i=0}^{r}C_{2i}^{2r}\sum_{k=0}^{i}C_k^i\left(\frac{\alpha w}{2}\right)^{2(r-k)}+\sum_{i=0}^{r-1}C_{2i+1}^{2r}h_i(w)\right)\phi(w)\Phi(\lambda u(w))\,\mathrm{d}w\\
&= 2\beta^r\sum_{i=0}^{r}C_{2i}^{2r}\sum_{k=0}^{i}C_k^i\left(\frac{\alpha}{2}\right)^{2(r-k)}\mathrm{E}[W^{2(r-k)}]\\
&\quad + 2\beta^r\sum_{i=0}^{r-1}C_{2i+1}^{2r}\int_{-\infty}^{\infty}h_i(w)\phi(w)\Phi\left(\lambda u(w)\right)\mathrm{d}w\\
&= 2\beta^r\sum_{i=0}^{r}C_{2i}^{2r}\sum_{k=0}^{i}C_k^i\left(\frac{\alpha}{2}\right)^{2(r-k)}(2(r-k)-1)!!+2\beta^r\sum_{i=0}^{r-1}C_{2i+1}^{2r}H_i,
\end{aligned}$$

the proof is completed, where $h_i(w) = (\alpha w/2)^{2r-(2i+1)}((\alpha w/2)^2+1)^{(2i+1)/2}$. Note that $H_i = \int_{-\infty}^{\infty} h_i(w)\phi(w)\Phi(\lambda u(w))\,\mathrm{d}w$ must be computed numerically, and $\mathrm{E}[W^{2(r-k)}] = (2(r-k)-1)!!$ coincides with the corresponding even moments of the standard normal distribution for $r > k$. ■

Proof of conjectures obtained from Figure 4**:** Let $T \sim \text{MSNBS}(\alpha,\beta,\lambda)$ and $W \sim \text{MSN}(\lambda)$. Then, by considering $\beta = 1$ without loss of generality, we have

$$\begin{aligned}
\mathrm{E}[T^r] &= \int_0^\infty 2t^r\phi(a_t(\alpha,1))\Phi\left(\lambda u(a_t(\alpha,1))\right)\frac{t^{-3/2}}{2\alpha}(t+1)\,\mathrm{d}t\\
&= \int_{-\infty}^{\infty} 2g_\alpha(w)^r\phi(w)\Phi\left(\lambda u(a(w))\right)\mathrm{d}w,
\end{aligned}$$

where $g_\alpha(w) = (\alpha w/2+\sqrt{(\alpha w/2)^2+1})^2$. For the CS γ_1, the following limiting cases are obtained: (i) for $\alpha\to\infty$, $g_\alpha(w)\approx\alpha^2w^2$, whereas for $\lambda\to\infty$, $\Phi(-\lambda u(a(w)))\approx 0, \Phi(\lambda u(a(w)))\approx 1$, with $\alpha^{2r}\int_0^\infty w^{2r}\phi(w)\,\mathrm{d}w = \alpha^{2r}(2r-1)!!$, and then, by applying L'Hopital rule with respect to α, we have $\lim_{\alpha\to\infty}\lim_{\lambda\to\infty}\gamma_1 = 2\sqrt{2}$; (ii) when $\alpha\to\infty$ and $\lambda\to-\infty$, γ_1 is obtained by using the trial and error method; and (iii) for $\alpha\to 0$, $g_\alpha(w)\approx 1$, whereas for $\lambda\to\pm\infty$, $\Phi(-\lambda u(a(w)))\approx 0, \Phi(\lambda u(a(w)))\approx 1$, and then by using $\mathrm{E}[T^r]$ and change of variables, as well as by applying L'Hopital rule with respect to that change of variables, we have $\lim_{\alpha\to\infty}\lim_{\lambda\to\infty}\gamma_1 = -\infty$. For the CK γ_2, the following limiting cases are obtained: (i) for $\lambda\to+\infty$ and $a\to\infty$, we have a similar case to (i) of the CS, with $\lim_{\alpha\to\infty}\lim_{\lambda\to\infty}\gamma_2 = 15$; (ii) when $\alpha\to\infty$ and $\lambda\to-\infty$, γ_2 is obtained by using the trial and error method; (iii) for $\alpha\to 0$ and $\lambda\to\pm\infty$, we have a similar case to (iii) of the CS, with $\lim_{\alpha\to\infty}\lim_{\lambda\to\infty}\gamma_2 = -\infty$; and (iv) the minimum value of $\gamma_2 = 1.95$, which is obtained using the trial and error method. ■

Appendix 2. Score vector and Fisher information matrix

Considering the notations $\boldsymbol{\theta} = (\alpha, \beta)$, $u = u(a) = a/(1+a^2)^{1/2}$, $\zeta = \zeta(\lambda u(a)) = \phi(\lambda u(a))/\Phi(\lambda u(a))$, $a = a_t(\alpha, \beta) = w/\alpha$, $w = \sqrt{t/\beta} - \sqrt{\beta/t}$, $\tilde{w} = \sqrt{t/\beta} + \sqrt{\beta/t}$ and the derivatives

$$\frac{\partial a}{\partial \alpha} = -\frac{1}{\alpha}, \quad \frac{\partial a}{\partial \beta} = -\frac{\tilde{w}}{2\alpha\beta}, \quad \frac{\partial u(a)}{\partial \boldsymbol{\theta}} = \frac{1}{(1+a^2)^{3/2}} \frac{\partial a}{\partial \boldsymbol{\theta}},$$

$$\frac{\partial}{\partial \boldsymbol{\theta}} \left(\frac{a}{(1+a^2)^{3/2}} \right) = \frac{1-2a^2}{(1+a^2)^{5/2}} \frac{\partial a}{\partial \boldsymbol{\theta}},$$

$$\frac{\partial \zeta}{\partial \boldsymbol{\theta}} = \left(-\frac{\lambda^2 a \zeta}{(1+a^2)^2} - \frac{\lambda \zeta^2}{(1+a^2)^{3/2}} \right) \frac{\partial a}{\partial \boldsymbol{\theta}}, \frac{\partial \zeta}{\partial \lambda}$$
$$= -\frac{\lambda a^2 \zeta}{(1+a^2)} - \frac{a \zeta^2}{(1+a^2)^{1/2}},$$

we find that score functions of α, β and λ of the MSNBS model as

$$S_\alpha = \frac{1}{\alpha} \left(-1 + a^2 - \frac{\lambda a \zeta}{(1+a^2)^{3/2}} \right), \quad S_\lambda = \frac{a\zeta}{(1+a^2)^{1/2}},$$
$$S_\beta = -\frac{1}{2\beta} + \frac{1}{2\alpha^2} \left(\frac{t}{\beta^2} - \frac{1}{t} \right) + \frac{1}{t+\beta} - \frac{\lambda}{2\alpha\beta} \frac{\tilde{w}\zeta}{(1+a^2)^{3/2}}$$

and the elements of the Fisher information matrix as

$$\ell_{\alpha\alpha} = -\mathrm{E}[S_\alpha S_\alpha] = -\frac{1}{\alpha^2} \mathrm{E} \left[1 - 3a^2 + \lambda \frac{a\zeta}{(1+a^2)^{3/2}} + \lambda \frac{(a-2a^3)\zeta}{(1+a^2)^{5/2}} \right.$$
$$\left. -\lambda^3 \frac{a^3 \zeta}{(1+a^2)^{7/2}} - \lambda^2 \frac{a^2 \zeta^2}{(1+a^2)^3} \right]$$
$$= -\frac{1}{\alpha^2} \left(-2 + \lambda(\rho_{103} + \rho_{105} - 2\rho_{305}) - \lambda^3 \rho_{307} - \lambda^2 \eta_{203} \right) = \frac{1}{\alpha^2} \left(2 + \lambda^2 \eta_{203} \right),$$

$$\ell_{\beta\alpha} = -\mathrm{E}[S_\beta S_\alpha]$$
$$= -\mathrm{E} \left[-\frac{1}{\alpha^3 \beta} w\tilde{w} + \frac{1}{2\alpha^2 \beta} \left[\frac{\lambda(1-2a^2)\tilde{w}\zeta}{(1+a^2)^{5/2}} - \frac{\lambda^3 a^2 \tilde{w} \zeta}{(1+a^2)^{7/2}} - \frac{\lambda^2 a \tilde{w} \zeta^2}{(1+a^2)^3} \right] \right]$$
$$= -\frac{1}{2\alpha^2 \beta} \left(-\frac{2}{\alpha} \left[\frac{\mathrm{E}[T]}{\beta} - \beta \mathrm{E}\left[\frac{1}{T} \right] \right] + \lambda(\rho_{015} - 2\rho_{215}) - \lambda^3 \rho_{217} - \lambda^2 \eta_{113} \right),$$

$$\ell_{\alpha\lambda} = -\mathrm{E}[S_\alpha S_\lambda] = -\mathrm{E} \left[-\frac{1}{\alpha} \frac{a}{(1+a^2)^{3/2}} \left(\zeta - \frac{\lambda^2 a^2 \zeta}{(1+a^2)} - \frac{\lambda a \zeta^2}{(1+a^2)^{1/2}} \right) \right)$$
$$= -\frac{1}{\alpha} \left(-\rho_{103} + \lambda^2 \rho_{305} + \lambda \eta_{202} \right) = -\frac{1}{\alpha} \lambda \eta_{202},$$

$$
\begin{aligned}
\ell_{\beta\beta} &= -\mathrm{E}[S_\beta S_\beta] \\
&= -\mathrm{E}\left[\frac{1}{2\beta^2} - \frac{t}{\alpha^2\beta^3} - \frac{1}{(t+\beta)^2} + \frac{\lambda(2\alpha\tilde{w} + \alpha^2 a)\zeta}{4\alpha^2\beta^2(1+a^2)^{3/2}}\right. \\
&\quad \left. - \frac{\lambda^3(\alpha^2 a^3 + 4a)\zeta}{4\alpha^2\beta^2(1+a^2)^{7/2}} - \frac{\lambda^2(\alpha^2 a^2 + 4)\zeta^2}{4\alpha^2\beta^2(1+a^2)^3} - \frac{3\lambda(\alpha^2 a^3 + 4a)\zeta}{4\alpha^2\beta^2(1+a^2)^{5/2}}\right] \\
&= -\frac{1}{2\beta^2} + \frac{\mathrm{E}[T]}{\alpha^2\beta^3} + \mathrm{E}\left[\frac{1}{(T+\beta)^2}\right] \\
&\quad + \frac{1}{4\alpha^2\beta^2}(-\lambda(2\alpha\rho_{013} + \alpha^2\rho_{103}) + \lambda^3(\alpha^2\rho_{307} + 4\rho_{107}) \\
&\quad + \lambda^2(\alpha^2\eta_{203} + 4\eta_{003}) + 3\lambda(\alpha^2\rho_{305} + 4\rho_{105})) \\
&= -\frac{1}{2\beta^2} + \frac{\mathrm{E}[T]}{\alpha^2\beta^3} + \mathrm{E}\left[\frac{1}{(T+\beta)^2}\right] + \frac{1}{4\alpha^2\beta^2}(-2\lambda\alpha\rho_{013} + \lambda^2(\alpha^2\eta_{203} + 4\eta_{003})), \\
\ell_{\lambda\beta} &= -\mathrm{E}[S_\lambda S_\beta] = -\frac{1}{2\alpha\beta}\mathrm{E}\left[-\frac{\tilde{w}\zeta}{(1+a^2)^{3/2}} + \frac{\lambda^2 a^2\tilde{w}\zeta}{(1+a^2)^{5/2}} + \frac{\lambda a\tilde{w}\zeta^2}{(1+a^2)^2}\right] \\
&= -\frac{1}{2\alpha\beta}(-\rho_{013} + \lambda^2\rho_{215} + \lambda\eta_{112}), \\
\ell_{\lambda\lambda} &= -\mathrm{E}[S_\lambda S_\lambda] = -\mathrm{E}\left[-\frac{\lambda a^3\zeta}{(1+a^2)^{3/2}} - \frac{a^2\zeta^2}{1+a^2}\right] = \lambda\rho_{303} + \eta_{201} = \eta_{201}.
\end{aligned}
$$