Brazilian Journal of Probability and Statistics
2016, Vol. 30, No. 2, 196–220
DOI: 10.1214/14-BJPS272

Log-symmetric distributions: Statistical properties and parameter estimation

Luis Hernando Vanegas[a,b] **and Gilberto A. Paula**[a]

[a]*Universidade de São Paulo*
[b]*Universidad Nacional de Colombia*

Abstract. In this paper, we study the main statistical properties of the class of log-symmetric distributions, which includes as special cases bimodal distributions as well as distributions that have heavier/lighter tails than those of the log-normal distribution. This family includes distributions such as the log-normal, log-Student-t, harmonic law, Birnbaum–Saunders, Birnbaum–Saunders-t and generalized Birnbaum–Saunders. We derive quantile-based measures of location, dispersion, skewness, relative dispersion and kurtosis for the log-symmetric class that are appropriate in the context of asymmetric and heavy-tailed distributions. Additionally, we discuss parameter estimation based on both classical and Bayesian approaches. The usefulness of the log-symmetric class is illustrated through a statistical analysis of a real dataset, in which the performance of the log-symmetric class is compared with that of some competitive and very flexible distributions.

1 Introduction

Data whose interest variable is continuous, strictly positive, and asymmetric and that may include outliers are commonly found in practice in various fields of knowledge. In fact, there is an extensive body of literature about distributions whose support is the interval $(0, \infty)$. Some of the more flexible distributions include the generalized modified Weibull (Carrasco, Ortega and Cordeiro (2008)), generalized Inverse Gaussian (see, e.g., Jørgensen, 1982), and generalized Gamma (Stacy, 1962) distributions. However, according to Limpert, Stahel and Abbt (2001), the log-normal distribution has been successfully applied in an enormous range of applications. Thus, to describe the behavior of strictly positive data, we consider the log-symmetric distribution class, which is a generalization of the log-normal distribution that is flexible enough to include as special cases bimodal distributions as well as distributions that have heavier/lighter tails than those of the log-normal distribution. Furthermore, the log-symmetric distributions are endowed with two interesting properties, closure under change of scale and closure under reciprocals, which, according to Puig (2008), are very desirable properties for distributions that are used to describe data with ratios of positive magnitudes.

Key words and phrases. Skewness, asymmetric distributions, Bayesian analysis, robust estimates, log-location-scale lifetime models.
Received March 2014; accepted November 2014.

The log-symmetric class also generalizes and makes more flexible the distributions that have been developed to describe lifetimes under the assumption of accumulated damage (e.g., the Birnbaum–Saunders (Birnbaum and Saunders, 1969), Birnbaum–Saunders-t (Barros, Paula and Leiva, 2008, Paula et al., 2012) and generalized Birnbaum–Saunders (Díaz-García and Leiva, 2005) distributions) by introducing therein an additional parameter that may be used to control the shape of the hazard function and to regulate the skewness and the relative dispersion. Furthermore, as demonstrated in this work, the log-symmetric class has several desirable statistical properties that may make it preferable to alternative distributions. For instance, the two parameters of the log-symmetric class are orthogonal and they may be interpreted directly as median and skewness (or relative dispersion), which are, in the context of asymmetric distributions, the most meaningful measures of location and shape, respectively. In addition, the extension of the log-symmetric class to the multivariate case is straightforward (Marchenko and Genton, 2010). This paper studies the main statistical properties of the class of log-symmetric distributions and addresses the key issues of parameter estimation based on both classical and Bayesian approaches.

The remainder of this paper is organized as follows: in Section 2, the log-symmetric class is characterized, and some of its main statistical properties are derived. Section 3 addresses the parameter estimation based on the maximum likelihood method by using the Fisher scoring and expectation-maximization (EM) algorithms. A joint iterative process for estimating the scale and power parameters is presented. Section 4 is devoted to a method for exact inferen ce provided by the Bayesian inference using Markov chain Monte Carlo (MCMC) methods. In Section 5, a practical use of the log-symmetric class of distributions is illustrated through an application to per capita gross domestic product data, in which the performance of the log-symmetric class is compared with that of some competitive and very flexible distributions such as the generalized modified Weibull, generalized Inverse Gaussian, generalized Gamma, log-skew-t (see, e.g., Azzalini, Dal Cappello and Kotz (2003)) and Box-cox-t (see, e.g., Rigby and Stasinopoulos (2006)) distributions.

2 Log-symmetric distributions

Let Y be a continuous and symmetric random variable whose distribution belongs to the symmetric class (see, e.g., Fang, Kotz and Ng, 1990) that has location parameter $-\infty < \mu < \infty$, dispersion parameter $\phi > 0$ and density generator $g(\cdot)$. We denote this class as $Y \sim \mathcal{S}(\mu, \phi, g(\cdot))$. Its probability density function is given by $f_Y(y) = g[(y-\mu)^2/\phi]/\sqrt{\phi}$ for $-\infty < y < \infty$ provided that $g(u) > 0$ for $u > 0$ and $\int_0^\infty u^{-1/2} g(u)\,\partial u = 1$. Then, by setting $T = \exp(Y)$, a new class of distributions, the so-called log-symmetric class, is obtained. We denote this class, whose support is the interval $(0, \infty)$, as $T \sim \mathcal{LS}(\eta, \phi, g(\cdot))$, where $\eta = \exp(\mu)$ and ϕ are its scale and power parameters, respectively (see, e.g., Marshall and Olkin, 2007, Chapters 7 and 12). The probability density function of T reduces to

$$f_T(t) = \frac{g(\tilde{t}^2)}{t\sqrt{\phi}}, \qquad t > 0, \tag{2.1}$$

where $\tilde{t} = \log[(t/\eta)^{1/\sqrt{\phi}}]$. The density generator $g(\cdot)$ may involve an extra parameter (or an extra parameter vector), which is denoted here as ζ; for convenience, we assume that this parameter is known. Members of the class of distributions characterized by (2.1) include the log-normal, log-Student-t, log-power-exponential, log-logistic type I and II, log-hyperbolic, log-slash, log-contaminated-normal, harmonic law (see Puig (2008) and references therein), Birnbaum–Saunders (Birnbaum and Saunders, 1969), Birnbaum–Saunders-t (see, e.g., Barros, Paula and Leiva, 2008) and generalized Birnbaum–Saunders (see, e.g. Díaz-García and Leiva, 2005, Leiva et al., 2008) distributions. It is noteworthy that the Birnbaum–Saunders, Birnbaum–Saunders-t and generalized Birnbaum–Saunders distributions cited above are special cases (in which $\phi = 4$) of the (extended) homonymous distributions that will be considered here. Similarly, the harmonic law cited above is a special case (in which $\phi = 1$) of the (extended) homonymous distribution that will be considered here.

Moreover, if $T \sim \mathcal{LS}(\eta, \phi, g(\cdot))$ then, one can verify the following properties, which are immediate consequences of the definition of the log-symmetric class:

(P1) The cumulative distribution function (c.d.f.) of T may be written as $F_T(t) = F_Z(\tilde{t})$, where $F_Z(\cdot)$ is the c.d.f. of $Z = (Y - \mu)/\sqrt{\phi} \sim \mathcal{S}(0, 1, g(\cdot))$.

(P2) $T^* = (T/\eta)^{1/\sqrt{\phi}} \sim \mathcal{LS}(1, 1, g(\cdot))$, that is, T^* follows standard log-symmetric distribution.

(P3) $cT \sim \mathcal{LS}(c\eta, \phi, g(\cdot))$ for all constant $c > 0$.

(P4) $T^c \sim \mathcal{LS}(\eta^c, c^2\phi, g(\cdot))$ for all constant $c \neq 0$.

(P5) (T/η) and (η/T) are random variables that are identically distributed.

(P6) If $\mathrm{E}(T^r)$ and $\mathrm{E}(Z)$ exist, then $\mathrm{E}(T^r) \geq \eta^r$.

(P7) If $\mathrm{M}_Y(r)$ exists, then $\mathrm{E}(T^r) = \mathrm{M}_Y(r)$, where $\mathrm{M}_Y(r)$ is the moment generating function of $Y = \log(T)$.

(P8) The quantile function of T is given by $\vartheta(q) = \eta \exp(\sqrt{\phi} Z_\zeta^{(q)})$, where $Z_\zeta^{(q)}$ is the $100(q)\%$ quantile of $Z = (Y - \mu)/\sqrt{\phi} \sim \mathcal{S}(0, 1, g(\cdot))$.

(P9) The Shannon entropy of T, which is denoted as $\mathrm{ET}(T)$, may be expressed as $\mathrm{ET}(T) = \log[\eta\sqrt{\phi}] + \mathrm{ET}(Z)$, provided that $\mathrm{ET}(Z)$ and $\mathrm{E}(Z)$ exist.

(P10) If the $\mathrm{W}(\cdot)$ function is such that $\log[\mathrm{W}(x)] = h[\log(x)]$, where $h : \mathbb{R} \to \mathbb{R}$ is an injective and differentiable odd function, then the distribution of $\overline{T} = \mathrm{W}(T^*)$ is $\mathcal{LS}(1, 1, \overline{g}(\cdot))$, where $\overline{g}(u) = g\{[h^{-1}(\sqrt{u})]^2\}/h'[h^{-1}(\sqrt{u})]$.

Let $\mathbf{t} = (t_1, t_2, \ldots, t_n)^\top$ be a random sample of size n from $T \sim \mathcal{LS}(\eta, \phi, g(\cdot))$. Then, the maximum likelihood estimates (MLEs) of η and ϕ (which are denoted here as $\hat{\eta}$ and $\hat{\phi}$, resp.) for known or fixed ζ may be written as

$$\hat{\eta} = \left\{ \prod_{k=1}^{n} t_k^{\mathrm{v}(\hat{\hat{t}}_k)} \right\}^{1/\sum_{k=1}^{n} \mathrm{v}(\hat{\hat{t}}_k)} \quad \text{and} \quad \hat{\phi} \propto \frac{\sum_{k=1}^{n} \mathrm{v}(\hat{\hat{t}}_k)[\log(t_k/\hat{\eta})]^2}{\sum_{k=1}^{n} \mathrm{v}(\hat{\hat{t}}_k)},$$

where $\hat{\hat{t}}_k = \log[(t_k/\hat{\eta})^{1/\sqrt{\hat{\phi}}}]$ and $\mathrm{v}(t) = -2g'(t^2)/g(t^2)$ is a weight function induced by $g(\cdot)$, and $\mathrm{v}(\hat{\hat{t}}_1), \ldots, \mathrm{v}(\hat{\hat{t}}_n)$ is a set of positive weights if the $g(u)$ function is monotonically decreasing for $u > 0$, with $g'(u) = \partial g(u)/\partial u$. Therefore, when $\mathrm{v}(\hat{\hat{t}}_k) > 0$ for $k = 1, \ldots, n$, the MLE of η , may be interpreted as a weighted geometric mean of $t_1, t_2, \ldots, t_n$, whereas the MLE of ϕ is proportional to a weighted arithmetic mean, for which $\mathrm{v}(\hat{\hat{t}}_1), \ldots, \mathrm{v}(\hat{\hat{t}}_n)$ are the individual-specific weights. Thus, the choice of $g(\cdot)$ may induce a $\mathrm{v}(\cdot)$ function that enables one to estimate the parameters using the maximum likelihood method in a manner that is robust to extreme or outlying observations (that is the $g(\cdot)$ function may induce an $\mathrm{v}(t)$ function whose value decreases as the t value departs from the *centre* of the T distribution).

Below are listed some distributions of the log-symmetric class

- Log-normal(η, ϕ):
$$g(u) \propto \exp\left[-\frac{1}{2}u\right] \quad \text{and} \quad \mathrm{v}(t) = 1.$$
- Log-Student-$t(\eta, \phi, \zeta)$:
$$g(u) \propto \left[1 + \frac{u}{\zeta}\right]^{-(\zeta+1)/2}, \qquad \zeta > 0 \quad \text{and} \quad \mathrm{v}(t) = \frac{\zeta + 1}{\zeta + t^2}.$$
If $\phi > (\zeta + 1)^2/4\zeta$, then the function $f_T(t)$ is monotonically decreasing.
- Log-power-exponential(η, ϕ, ζ):
$$g(u) \propto \exp\left[-\frac{1}{2}u^{1/(1+\zeta)}\right], \qquad -1 < \zeta \leq 1 \quad \text{and} \quad \mathrm{v}(t) = \frac{|t|^{-(2\zeta)/(\zeta+1)}}{1 + \zeta}.$$
The log-normal$(\zeta = 0)$ and the log-Laplace$(\zeta = 1)$ distributions are special cases. If $\zeta = 1$ and $\phi > 1/4$, then $f_{T(t)}$ is a monotonically decreasing function because $f'_{T(t)} < 0$ for $t \in (0, \eta) \cup (\eta, \infty)$. Similarly, if $\zeta = 1$ and $\phi < 1/4$, then η is a mode of T because $f_T(t)/f_T(\eta) \leq 1$ for $t > 0$. The distribution of $Y = \log(T)$ is power exponential (Box and Tiao, 1973).
- Log-hyperbolic(η, ϕ, ζ):
$$g(u) \propto \exp[-\zeta\sqrt{1 + u}], \qquad \zeta > 0 \quad \text{and} \quad \mathrm{v}(t) = \frac{\zeta}{\sqrt{1 + t^2}}.$$
The log-normal (η, σ^2) distribution is a limiting case when $\phi \to \infty$ and $\phi/\zeta \to \sigma^2$. Similarly, the log-Laplace(η, σ^2) distribution is a limiting case when $\phi \to 0$

and $\phi/\zeta^2 \to 4\sigma^2$ (Fonseca, Migon Helio and Ferreira Marco (2012)). If $\phi \geq \zeta^2$, then the function $f_T(t)$ is monotonically decreasing. The distribution of $Y = \log(T)$ is symmetric hyperbolic (Barndoff-Nielsen, 1977). The moment generating function of Y allows to calculate the moments of T by using the following expression

$$\mathrm{E}(T^r) = \eta^r \frac{\mathrm{K}_1(\sqrt{\zeta^2 - \phi r^2})}{\mathrm{K}_1(\zeta)} \frac{\zeta}{\sqrt{\zeta^2 - \phi r^2}}, \qquad |r| < \zeta/\sqrt{\phi},$$

in which $\mathrm{K}_r(\zeta) = \frac{1}{2}\int_0^\infty x^{r-1} \exp[-\frac{\zeta}{2}(x + \frac{1}{x})]\,\partial x$ is the modified Bessel function of third-order and index r.

- Log-slash(η, ϕ, ζ):

$$g(u) \propto \mathrm{IGF}\left(\zeta + \frac{1}{2}, \frac{u}{2}\right) \quad \text{and} \quad \mathrm{v}(t) = \mathrm{IGF}\left(\zeta + \frac{3}{2}, \frac{t^2}{2}\right) \Big/ \mathrm{IGF}\left(\zeta + \frac{1}{2}, \frac{t^2}{2}\right),$$

where $\zeta > 0$, $\mathrm{IGF}(a, x) = \int_0^1 \exp(-tx) t^{a-1}\,\partial t$ is the incomplete gamma function for $a > 0$ and $x \geq 0$. The distribution of $Y = \log(T)$ is slash (Rogers and Tukey, 1972).

- Log-contaminated-normal$(\eta, \phi, \zeta = (\zeta_1, \zeta_2)^\top)$: The log-contaminated-normal distribution is a convex linear combination of the $T_1 \sim$ log-normal$(\eta, \phi/\zeta_2)$ and $T_2 \sim$ log-normal(η, ϕ) distributions, that is, $f_T(t) = \zeta_1 f_{T_1}(t) + (1 - \zeta_1) f_{T_2}(t)$. Therefore,

$$g(u) \propto \sqrt{\zeta_2} \exp\left[-\frac{1}{2}\zeta_2 u\right] + \frac{(1 - \zeta_1)}{\zeta_1} \exp\left[-\frac{1}{2}u\right], \qquad 0 < \zeta_1 < 1, 0 < \zeta_2 < 1$$

and

$$\mathrm{v}(t) = \frac{{\zeta_2}^{3/2}\zeta_1 \exp[(1 - \zeta_2)(t^2/2)] + (1 - \zeta_1)}{{\zeta_2}^{1/2}\zeta_1 \exp[(1 - \zeta_2)(t^2/2)] + (1 - \zeta_1)}.$$

The moments of T are given by

$$\mathrm{E}(T^r) = \eta^r \exp\left[\frac{1}{2}\phi r^2\right]\left\{\zeta_1 \exp\left[\frac{1 - \zeta_2}{2\zeta_2}\phi r^2\right] + (1 - \zeta_1)\right\}.$$

If the mode of T exists, then it is within the interval $(\eta \exp(-\phi/\zeta_2), \eta \exp(-\phi))$. The distribution of $Y = \log(T)$ is contaminated normal.

- (extended) Birnbaum–Saunders(η, ϕ, ζ):

$$g(u) \propto \cosh(u^{1/2}) \exp\left[-\frac{2}{\zeta^2}\sinh^2(u^{1/2})\right], \qquad \zeta > 0$$

and

$$\mathrm{v}(t) = \frac{\sinh(t)}{t}\left[\frac{4\cosh(t)}{\zeta^2} - \frac{1}{\cosh(t)}\right],$$

where $\sinh(\cdot)$ and $\cosh(\cdot)$ represent the hyperbolic sine and cosine functions, respectively. The weighting $\mathrm{v}(t)$ of the Birnbaum–Saunders distribution increases as $|t|$ also increases. In addition, the Birnbaum–Saunders distribution is bimodal if $\zeta > 2$ and $\phi < \varrho(\zeta)$, where

$$\varrho(\zeta) = (\sqrt{1+2\zeta^2} - 3) \times \left[\frac{1}{\sqrt{1+\sqrt{1+2\zeta^2}}} - \frac{\sqrt{1+\sqrt{1+2\zeta^2}}}{\zeta^2}\right]^2.$$

Note that $\varrho(\zeta) < 1$ for all $\zeta > 2$. The moments of T are given by (Rieck, 1999)

$$\mathrm{E}(T^r) = \eta^r \frac{\exp(1/\zeta^2)}{\zeta\sqrt{2\pi}}[\mathrm{K}_{r_1^*}(1/\zeta^2) + \mathrm{K}_{r_2^*}(1/\zeta^2)],$$

where $r_1^* = (r\sqrt{\phi}+1)/2$ and $r_2^* = (r\sqrt{\phi}-1)/2$. The distribution of $Y = \log(T)$ is sinh-normal (Rieck and Nedelman, 1991).

- (extended) Birnbaum–Saunders-$t\,(\eta, \phi, \zeta = (\zeta_1, \zeta_2)^\top)$:

$$g(u) \propto \cosh(u^{1/2})[\zeta_2\zeta_1^2 + 4\sinh^2(u^{1/2})]^{-(\zeta_2+1)/2}, \qquad \zeta_1 > 0, \zeta_2 > 0$$

and

$$\mathrm{v}(t) = \frac{\sinh(t)}{t}\left[\frac{4(\zeta_2+1)\cosh(t)}{\zeta_2\zeta_1^2 + 4\sinh^2(t)} - \frac{1}{\cosh(t)}\right].$$

The Birnbaum–Saunders-t distribution is bimodal if $\zeta_1 > 2\sqrt{1+1/\zeta_2}$ and $\phi < [\mathrm{v}(t_1)t_1]^2$, where $t_1 = \log[\sqrt{t_0^2-1} + t_0]$ and t_0 is given by

$$t_0 = \frac{1}{2}\left[(\zeta_2+1) + \frac{2\zeta_2}{\zeta_1^2\zeta_2 - 4}\right]^{-1/2} \times \left[(\zeta_2+3) + \sqrt{(\zeta_2+3)^2 + 2(\zeta_2+1)(\zeta_1^2\zeta_2-4) + 4\zeta_2}\right]^{1/2}.$$

The distribution of $Y = \log(T)$ is sinh-t (see, e.g., Paula et al., 2012).

- (extended) Generalized Birnbaum–Saunders$(\eta, \phi, \zeta = (\zeta_1, \zeta_2)^\top)$:

$$g(u) \propto \cosh(u^{1/2}) \times h_{\zeta_2}\left[\frac{4}{\zeta_1^2}\sinh^2(u^{1/2})\right], \qquad \zeta_1 > 0,$$

where $h_{\zeta_2}(\cdot)$ represents the kernel of the symmetric distribution (indexed by the extra parameter ζ_2) that describes the *cumulative damage*, where $h_{\zeta_2}(u) > 0$ for $u > 0$ and $\int_0^\infty u^{-1/2}h_{\zeta_2}(u)\,\partial u = 1$. The distribution of $Z^* = (2/\zeta_1)\sinh(\tilde{t})$ is given by $f_{Z^*}(z) = h_{\zeta_2}(z^2)$ (Leiva et al., 2008). The Birnbaum–Saunders, Birnbaum–Saunders-t and slash-Birnbaum–Saunders distributions are special cases (Balakrishnan et al., 2009).

- (extended) Harmonic law(η, ϕ, ζ):

$$g(u) \propto \exp\big[-\zeta \cosh\big(u^{1/2}\big)\big], \qquad \zeta > 0 \quad \text{and} \quad \mathrm{v}(t) = \zeta \frac{\sinh(t)}{t}.$$

  The moments of T are given by

$$\mathrm{E}\big(T^r\big) = \eta^r \frac{\mathrm{K}_{r^*}(\zeta)}{\mathrm{K}_0(\zeta)},$$

  where $r^* = r\sqrt{\phi}$. The mode of T is $\eta[\zeta^{-1}(\sqrt{\phi + \zeta^2} - \sqrt{\phi})]^{\sqrt{\phi}}$. Some interesting properties of the harmonic law are described by Puig (2008).

The values of the individual-specific weights $\mathrm{v}(t)$ are strictly positive for the log-normal, log-Student-t, log-power-exponential, log-slash, log-hyperbolic, log-contaminated-normal, harmonic law, Birnbaum–Saunders (for $\zeta \leq 2$) and Birnbaum–Saunders-t (for $\zeta_1 \leq 2\sqrt{1 + 1/\zeta_2}$) distributions. Furthermore, for distributions with tails that are heavier than those of the log-normal distribution (e.g., the log-Student-t, log-power-exponential (for $0 < \zeta \leq 1$), log-slash, log-hyperbolic and log-contaminated-normal distributions) the individual-specific weights tend to be smaller as t departs from the "center" of the T distribution. Therefore, for distributions that have heavier tails, the MLEs of η and ϕ are less sensitive to extreme or outlying observations than for the log-normal distribution. Similar results hold for the distributions that have been developed to describe lifetimes under the assumption of cumulative damage. In fact, one can verify that the weights of extreme or outlying observations for the Birnbaum–Saunders-t distribution are (relatively) smaller than those for the Birnbaum–Saunders distribution. Figure 1 shows the probability density functions of some log-symmetric distributions. This figure illustrates the flexibility of the log-symmetric class.

2.1 Hazard function

The hazard function of T is given by $r_T(t) = f_T(t)/[1 - F_T(t)]$. The function $r_T(t)$ of the class of log-symmetric distributions is quite flexible and can take various shapes as illustrated in Figure 2. In fact, as demostrated by (Glaser, 1980), the following statements hold: (a) If $\delta(t) > 0$ for all $t \in \mathbb{R}$, then $r_T(t)$ is increasing. (b) If $\delta(t) < 0$ for all $t \in \mathbb{R}$, then $r_T(t)$ is decreasing. (c) If $\lim_{t\to 0} f_T(t) = \infty$ and there exists t_0 for which $\delta(t) > 0$ in $t < t_0$, $\delta(t_0) = 0$, and $\delta(t) < 0$ in $t > t_0$, then $r_T(t)$ is decreasing. (d) If $\lim_{t\to 0} f_T(t) = 0$ and there exists t_0 for which $\delta(t) > 0$ in $t < t_0$, $\delta(t_0) = 0$, and $\delta(t) < 0$ in $t > t_0$, then $r_T(t)$ is upside-down bathtub shaped. (e) If $\lim_{t\to 0} f_T(t) = \infty$ and there exists t_0 for which $\delta(t) < 0$ in $t < t_0$, $\delta(t_0) = 0$, and $\delta(t) > 0$ in $t > t_0$, then $r_T(t)$ is bathtub shaped. (f) If $\lim_{t\to 0} f_T(t) = 0$ and there exists t_0 for which $\delta(t) < 0$ in $t < t_0$, $\delta(t_0) = 0$, and $\delta(t) > 0$ in $t > t_0$, then $r_T(t)$ is increasing. The $\delta(t)$ function may be written as

$$\delta(t) = \mathrm{v}'(t)t + \mathrm{v}(t)[1 - t\sqrt{\phi}] - \phi.$$

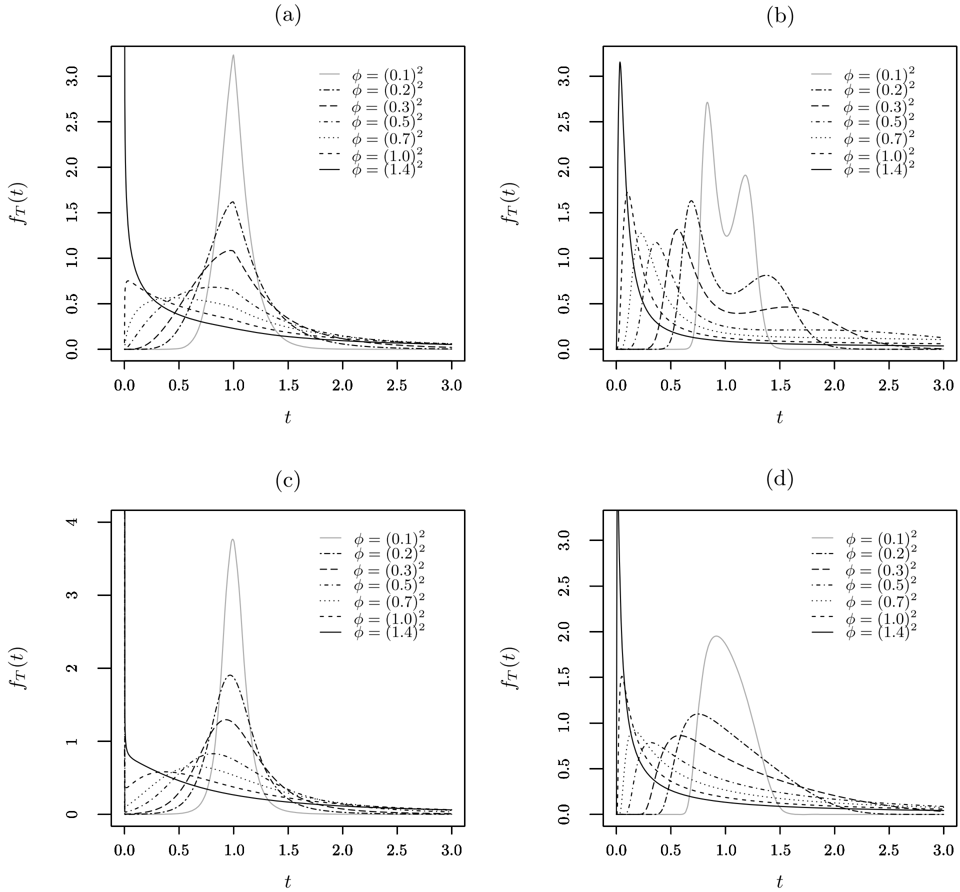

Figure 1 *Graph of the density function of the log-power-exponential*$(\eta = 1, \phi, \zeta = 0.5)$ (a), *Birnbaum–Saunders-t*$(\eta = 1, \phi, \zeta = (6, 4)^\top)$ (b), *log-Student-t*$(\eta = 1, \phi, \zeta = 4)$ (c), *and harmonic-law*$(\eta = 1, \phi, \zeta = 0.1)$ (d) *distributions*.

For example, the hazard function of the log-normal distribution is upside-down bathtub shaped. Moreover, if $\phi \geq 1$, then the hazard function of the harmonic law is upside-down bathtub shaped, and it is increasing if $\phi < 1$ and $\zeta > \phi/\sqrt{1-\phi}$. If $\phi \geq \zeta^2$ and $\phi \geq 2$, then the hazard function of the log-hyperbolic distribution is decreasing.

2.2 Summary of the shape

The measures that are most frequently used for assessing the location, dispersion, relative dispersion, skewness and kurtosis are based on moments. However, because of their derivation, such measures may be not appropriate in the context of asymmetric distributions. Furthermore, sometimes the moments are not finite or are quite difficult to calculate. Therefore, in this section, measures of the loca-

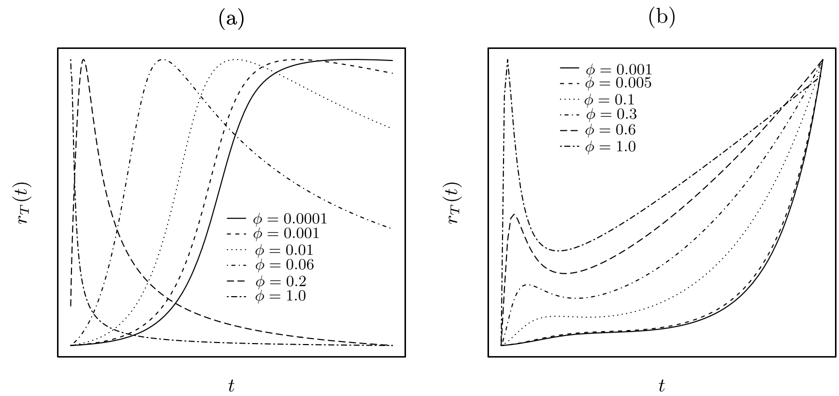

Figure 2 *Graph of the hazard function of the log-hyperbolic*$(\eta = 1, \phi, \zeta = 1)$ (a), *and Birnbaum–Saunders*$(\eta = 1, \phi, \zeta = 3)$ (b) *distributions.*

tion, dispersion, relative dispersion, skewness and kurtosis for the log-symmetric class are derived; these measures, exist even for distributions for which no moments exist, are appropriate in the context of asymmetric distributions, are easier to calculate and/or interprete than those based on moments, and are invariant under changes in the extreme tails of the distribution. Furthermore, some of these measures (that is those used to measure the relative dispersion, skewness and kurtosis) are invariant under location-scale transformations (see Groeneveld and Meeden, 1984).

2.2.1 *Location.* The median of $T \sim \mathcal{LS}(\eta, \phi, g(\cdot))$ is η, and the mode(s) of T may be written as $M_T = \eta \exp(t_T\sqrt{\phi})$ provided that $f_T(t)$ is a twice continuously differentiable function, in which t_T is(are) the solution(s) of

$$-\mathrm{v}(t)t = \sqrt{\phi} \qquad \text{restricted to } \mathrm{v}'(t)t^2\mathrm{sign}(t) > \sqrt{\phi}\,\mathrm{sign}(t),$$

where $\mathrm{v}'(t) = \partial \mathrm{v}(t)/\partial t$. In addition, it is possible to verify that $M_T < \eta$ if both $g(u)$ is monotonically decreasing for $u > 0$ and $f_T'(t)$ is continuous in M_T.

2.2.2 *Dispersion.* The interquartile range of $T \sim \mathcal{LS}(\eta, \phi, g(\cdot))$ may be expressed as

$$\varsigma = \vartheta(0.75) - \vartheta(0.25) = 2\eta \sinh\bigl(\sqrt{\phi} Z_\zeta^{(0.75)}\bigr), \qquad \varsigma \in (0, \infty).$$

2.2.3 *Relative dispersion.* The coefficient of quartile variation (Zwillinger and Kokoska (2000), page 17) is given by

$$\varpi = \frac{\vartheta(0.75) - \vartheta(0.25)}{\vartheta(0.75) + \vartheta(0.25)} = \tanh\bigl(\sqrt{\phi} Z_\zeta^{(0.75)}\bigr), \qquad \varpi \in (0, 1),$$

where $\tanh(\cdot)$ is the hyperbolic tangent function. Note that ϖ is a monotonically increasing function of ϕ for fixed ζ. Therefore, according to ϖ, ϕ may be interpreted as a measure of the relative dispersion of T for fixed ζ. According to Bonett (2006), ϖ may be preferable to the coefficient of variation for describing the relative dispersion in asymmetric distributions.

2.2.4 *Skewness.* A quantile-based measure of skewness (see, e.g., Hinkley, 1975, Groeneveld and Meeden, 1984) is given by

$$\varkappa(q) = \frac{\vartheta(q) + \vartheta(1-q) - 2\vartheta(1/2)}{\vartheta(1-q) - \vartheta(q)} = \operatorname{cosech}\big(\sqrt{\phi} Z_\zeta^{(q)}\big) - \operatorname{cotanh}\big(\sqrt{\phi} Z_\zeta^{(q)}\big),$$

where $\varkappa(q) \in (0, 1)$, $q \in (0, \frac{1}{2})$, and $\operatorname{cotanh}(\cdot)$ and $\operatorname{cosech}(\cdot)$ represent the hyperbolic cotangent and cosecant functions, respectively. A simple derivative reveals that for all $q \in (0, \frac{1}{2})$ the measure of skewness $\varkappa(q)$ is a monotonically increasing function of ϕ for fixed ζ. Therefore, ϕ may be interpreted as the skewness of T for fixed ζ.

2.2.5 *Kurtosis.* The kurtosis proposed by Moors (1988) reduces to

$$\varsigma = \frac{\vartheta(7/8) - \vartheta(5/8) + \vartheta(3/8) - \vartheta(1/8)}{\vartheta(6/8) - \vartheta(2/8)} = \frac{\sinh(\sqrt{\phi} Z_\zeta^{(7/8)}) - \sinh(\sqrt{\phi} Z_\zeta^{(5/8)})}{\sinh(\sqrt{\phi} Z_\zeta^{(6/8)})},$$

where $\varsigma \in [0, \infty)$.

The main conclusion of this section is that, irrespective of the value of ϕ, η is the median of the T distribution. Similarly, irrespective of the value of η, ϕ is a measure of the skewness (or the relative dispersion) of T for fixed ζ.

3 Maximum likelihood estimation

The log-likelihood function of the interest parameters can be written as

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2}\log(\phi) - \sum_{k=1}^{n}\log(t_k) + \sum_{k=1}^{n}\log[g(\tilde{t}_k^2)].$$

To calculate the maximum likelihood estimate of $\boldsymbol{\theta} = (\eta, \phi)^\top$, denoted as $\hat{\boldsymbol{\theta}}$, the system of equations given by $(\mathrm{U}_\eta(\hat{\boldsymbol{\theta}}), \mathrm{U}_\phi(\hat{\boldsymbol{\theta}})) = (\partial\ell(\hat{\boldsymbol{\theta}})/\partial\hat{\eta}, \partial\ell(\hat{\boldsymbol{\theta}})/\partial\hat{\phi}) = (0, 0)$ is solved using the Fisher scoring algorithm, where

$$\mathrm{U}_\eta(\boldsymbol{\theta}) = \frac{1}{\eta\phi}\log\left[\prod_{k=1}^{n}(t_k/\eta)^{\mathrm{v}(\tilde{t}_k)}\right]$$

and

$$\mathrm{U}_\phi(\boldsymbol{\theta}) = -\frac{n}{2\phi} + \frac{1}{2\phi}\sum_{k=1}^{n}\mathrm{v}(\tilde{t}_k)\tilde{t}_k^2.$$

The (expected) Fisher information matrix, which is denoted as $\mathbf{K}(\boldsymbol{\theta})$, is given by

$$n^{-1}\mathbf{K}(\boldsymbol{\theta}) = -n^{-1}\mathrm{E}\big[\partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^\top\big] = \begin{bmatrix} d_g(\zeta)/\phi\eta^2 & 0 \\ 0 & [f_g(\zeta)-1]/4\phi^2 \end{bmatrix},$$

where $d_g(\zeta) = \mathrm{E}[\mathrm{v}^2(Z)Z^2]$ and $f_g(\zeta) = \mathrm{E}[\mathrm{v}^2(Z)Z^4]$ for $Z \sim \mathcal{S}(0,1,g(\cdot))$. For instance, the quantity $d_g(\zeta)$ is equal to 1, $(\zeta+1)/(\zeta+3)$, $\{2^{1-\zeta}\Gamma[(3-\zeta)/2]\}/\{(1+\zeta)^2\Gamma[(1+\zeta)/2]\}$ and $2+\frac{4}{\zeta^2}-\frac{\sqrt{2\pi}}{\zeta}\{1-\mathrm{erf}(\frac{\sqrt{2}}{\zeta})\}\exp(\frac{2}{\zeta^2})$ when T is assumed to exhibit the log-normal, log-Student-t, log-power-exponential and Birnbaum–Saunders distributions, respectively, where $\Gamma(\cdot)$ represents the gamma function and $\mathrm{erf}(x) = (2/\sqrt{\pi})\int_0^x e^{-t^2}\,dt$. Similarly, the quantity $f_g(\zeta)$ is equal to 3, $3(\zeta+1)/(\zeta+3)$ and $(\zeta+3)/(\zeta+1)$ when T is assumed to exhibit the log-normal, log-Student-t and log-power-exponential distributions, respectively (see, e.g., Cordeiro et al. (2000); Villegas et al. (2013)). Then, the Fisher scoring algorithm becomes

Algorithm 3.1.

Step 1. Set the initial value of the parameter vector to $\boldsymbol{\theta}^{(0)}$.
Step 2. Initialize the counter of the algorithm as $l = 0$.
Step 3. Based on $\boldsymbol{\theta}^{(l)}$ calculate the following expressions:

$$\eta^{(l+1)} = \eta^{(l)}\left\{\prod_{k=1}^{n}[t_k/\eta^{(l)}]^{\rho(\tilde{t}_k^{(l)})}\right\}^{1/n}$$

and

$$\log[\phi^{(l+1)}] = \log[\phi^{(l)}] + \frac{2}{f_g(\zeta)-1}\left\{n^{-1}\sum_{k=1}^{n}\mathrm{v}(\tilde{t}_k^{(l)})[\tilde{t}_k^{(l)}]^2 - 1\right\},$$

where $\tilde{t}_k^{(l)} = \log[(t_k/\eta^{(l)})^{1/\sqrt{\phi^{(l)}}}]$ and $\rho(t) = \mathrm{v}(t)/d_g(\zeta)$ is a standardized version of the weight function.
Step 4. Update $l = (l+1)$ and $\boldsymbol{\theta}^{(l)}$.
Step 5. Repeat steps 3 and 4 until convergence of $\boldsymbol{\theta}^{(l)}$ is reached.

Because the MLEs of η and ϕ for the log-normal distribution have closed forms, they can be used as initial values for the iterative procedure for other log-symmetric distributions. Because some distributions, such as the log-Student-t, log-power-exponential (for $0 \le \zeta \le 1$), log-slash, log-hyperbolic and log-contaminated-normal distributions, may be obtained as a power mixture of log-normal distributions (see, e.g., Andrews and Mallows, 1974, West, 1987, Barndoff-Nielsen, 1977), the EM algorithm (Dempster, Laird and Rubin, 1977) can be used in those cases to develop an even more efficient iterative process for parameter estimation. Then, for these distributions the *step 3* of the Algorithm 3.1 reduces to

$$\eta^{(l+1)} = \left\{ \prod_{k=1}^{n} t_k^{\mathrm{v}(\tilde{t}_k^{(l)})} \right\}^{1/\sum_{k=1}^{n} \mathrm{v}(\tilde{t}_k^{(l)})} \qquad \text{and}$$

$$\phi^{(l+1)} = n^{-1} \sum_{k=1}^{n} \mathrm{v}(\tilde{t}_k^{(l)}) [\log(t_k/\eta^{(l)})]^2.$$

Moreover, according to Balakrishnan et al. (2009), some distributions of the generalized Birnbaum–Saunders class (e.g., the Birnbaum–Saunders-t and slash-Birnbaum–Saunders distributions) can be obtained as a mixture of the Birbaum–Saunders distribution. Thus, the EM algorithm can also be used in those cases to develop a more efficient iterative procedure of parameter estimation. For example, under the Birnbaum–Saunders-t distribution the joint iterative process for $\hat{\eta}$ and $\hat{\phi}$ becomes

Algorithm 3.2.

Step 1. Set the initial value of the parameter vector to $\boldsymbol{\theta}^{(0)}$.
Step 2. Initialize the counter of the EM algorithm as $m = 0$.
Step 3. Calculate $\mathbf{u}^{(m)} = (u_1^{(m)}, \ldots, u_n^{(m)})^\top$ based on $\boldsymbol{\theta}^{(m)}$ as follows:

$$u_k^{(m)} = \frac{\zeta_1^2(\zeta_2 + 1)}{\zeta_1^2 \zeta_2 + [2\sinh(\tilde{t}_k^{(m)})]^2} \qquad \text{for } k = 1, \ldots, n.$$

Step 4. Calculate $d_g^*(\zeta_1/[u_1^{(m)}]^{1/2}), \ldots, d_g^*(\zeta_1/[u_n^{(m)}]^{1/2})$, where

$$d_g^*(\zeta) = 2 + \frac{4}{\zeta^2} - \frac{\sqrt{2\pi}}{\zeta} \left\{ 1 - \frac{2}{\sqrt{\pi}} \int_0^{\sqrt{2}/\zeta} \exp(-t^2)\, \partial t \right\} \exp\left(\frac{2}{\zeta^2} \right).$$

Step 5. Calculate $f_g^*(\zeta_1/[u_1^{(m)}]^{1/2}), \ldots, f_g^*(\zeta_1/[u_n^{(m)}]^{1/2})$, where

$$f_g^*(\zeta) = \mathrm{E}\big[\big(4\sinh(Z)\cosh(Z)Z/\zeta^2 - \tanh(Z)Z\big)^2\big]$$

and where Z exhibits a standard sinh-normal(ζ) distribution.
Step 6. Initialize the counter of the Fisher scoring algorithm as $l = 0$.
Step 7. Set the initial value of the parameter vector to $\boldsymbol{\theta}_*^{(0)} = \boldsymbol{\theta}^{(m)}$.
Step 8. Perform the following algorithm based on $\boldsymbol{\theta}_*^{(l)}$:
(A) Compute the following expressions:

$$\eta^{(l+1)} = \eta^{(l)} \left\{ \prod_{k=1}^{n} [t_k/\eta^{(l)}]^{\rho^*(\tilde{t}_k^{(l)}, u_k^{(m)})} \right\} \quad \text{and}$$

$$\log[\phi^{(l+1)}] = \log[\phi^{(l)}] + \frac{2\sum_{k=1}^{n} \{\mathrm{v}^*(\tilde{t}_k^{(l)}, u_k^{(m)})[\tilde{t}_k^{(l)}]^2 - 1\}}{\sum_{k=1}^{n} \{f_g^*(\zeta_1/[u_k^{(m)}]^{1/2}) - 1\}},$$

where

$$\rho^*(\tilde{t}_k^{(l)}, u_k^{(m)}) = \mathrm{v}^*(\tilde{t}_k^{(l)}, u_k^{(m)}) \Big/ \sum_{i=1}^{n} d_g^*(\zeta_1/[u_i^{(m)}]^{1/2}),$$

$$k = 1, \ldots, n$$

and

$$\mathrm{v}^*(\tilde{t}, u) = \frac{\sinh(\tilde{t})}{\tilde{t}} \left[\frac{4\cosh(\tilde{t})u}{\zeta_1^2} - \frac{1}{\cosh(\tilde{t})} \right].$$

(B) Update $l = (l+1)$ and $\boldsymbol{\theta}_*^{(l)}$.
(C) Repeat steps (A) and (B) until convergence of $\boldsymbol{\theta}_*^{(l)}$ is reached.

Step 9. Update $m = (m+1)$ and $\boldsymbol{\theta}^{(m)} = \boldsymbol{\theta}_*^{(l)}$.

Step 10. Repeat steps 3, 4, 5, 6, 7, 8 and 9 until convergence of $\boldsymbol{\theta}^{(m)}$ is reached.

The usual regularity conditions of large sample theory are fulfilled by all of the log-symmetric distributions listed above except for the log-Laplace distribution (i.e., the log-power-exponential distribution for $\zeta = 1$) (see Cordeiro et al. (2000)). Thus, the asymptotic distribution of the maximum likelihood estimator of $\boldsymbol{\theta}$ is the following:

$$\sqrt{n}\begin{pmatrix} \hat{\eta} - \eta \\ \hat{\phi} - \phi \end{pmatrix} \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}_2\left(\mathbf{0};\ \begin{bmatrix} \phi\eta^2/d_g(\zeta) & 0 \\ 0 & 4\phi^2/[f_g(\zeta)-1] \end{bmatrix}\right).$$

Hence, $\hat{\eta}$ and $\hat{\phi}$ are asymptotically independent.

3.1 Measuring goodness-of-fit

The goodness-of-fit is quantified using the following statistic, which is quite intuitive and has the advantage of graphical representation:

$$\Upsilon_\zeta = n^{-1} \sum_{k=1}^{n} |\Phi^{-1}[F_T(\hat{\tilde{t}}^{(k)})] - \upsilon^{(k)}|,$$

where $F_T(\cdot)$ is the cumulative distribution function of T, $\hat{\tilde{t}}^{(k)}$ is the kth order statistic of $\hat{\tilde{t}}$, $\upsilon^{(k)}$ is the expectation of the kth order statistic for a random sample of size n of a standard normal distribution and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution. Note that $\Phi^{-1}[F_T(\tilde{t}^{(1)})], \ldots, \Phi^{-1}[F_T(\tilde{t}^{(n)})]$ represent an ordered random sample from a standard normal distribution and $\hat{\eta}$ and $\hat{\phi}$ are a consistent estimators. Then, smaller values of Υ_ζ indicate better goodness-of-fit. Graphically, the criterion Υ_ζ indicates that the smaller the difference between the normal Q–Q plot of $\Phi^{-1}[F_T(\hat{\tilde{t}}^{(1)})], \ldots, \Phi^{-1}[F_T(\hat{\tilde{t}}^{(n)})]$ and a straight line (with zero intercept and unit slope), the better the goodness-of-fit. One advantage of this criterion is that it allows for graphically evaluating the appropriateness or agreement with the data of the tails (heavier or lighter) and/or the unimoda lity/bimodality of the distribution postulated for T.

3.2 Choosing the extra parameter value

The density generator $g(\cdot)$ considered in the T distribution involves the extra parameter ζ, which is assumed to be known or fixed in the estimation process described above. This assumption ensures easy calculation of confidence regions and hypothesis testing for η and ϕ using Wald and Rao statistics because the Fisher information matrix is diagonal. The motivation for this assumption also comes from the paper by Lucas (1997), which demonstrated that the robustness against outlying observations of Student-t models remains only if the degrees of freedom are fixed instead of estimated using the maximum likelihood method. In addition, Kano, Berkane and Bentler (1993) (and references therein) reported difficulties in calculating the extra parameter using the maximum likelihood method for the power exponential and contaminated normal distributions. Thus, to consider a unified approach for log-symmetric distributions, we propose choosing the extra parameter value by minimizing the Υ_ζ statistic. In fact, if the estimator $\hat{\zeta} = \operatorname{argmin} \Upsilon_\zeta$ is consistent; and $d_g(\cdot)$ and $f_g(\cdot)$ are continuous functions, then the multivariate Slutsky's theorem allows one to demonstrate that the ζ value may be replaced with the value obtained by minimizing the Υ_ζ statistic without changing the asymptotic distribution of $\hat{\eta}$ and $\hat{\phi}$.

To investigate the performance of the proposed criterion, a simu lation study is performed. First, a sample of size n is generated from a standard log-symmetric distribution. The resulting sample is used to estimate η and ϕ using the maximum likelihood method and to choose the extra parameter value by minimizing the Υ_ζ statistic. This process is replicated $R = 5000$ times. To consider di fferent simulation scenarios, different log-symmetric distributions are used (i.e., the log-Student-t, log-power-exponential, log-hyperbolic, log-slash and Birnbaum–Saunders distributions) and the sample size is modified by considering $n = 50$, 100, 200, 400 and 800. As summary measures, the median (M) and interquartile range (IR) of the R chosen extra parameter values are considered. The results are presented in Table 1. It can be observed that in all scenarios, the median of the extra parameter values yielded by the proposed method tends to the true value as the size of the sample increases. Similarly, the variability around the median decreases as the size of the sample increases. These results indicate that for large sample sizes, the difference between the extra parameter value yielded by the proposed method and the true parameter value is *small*. Therefore, for large sample sizes, the inference on η and ϕ could be based on the asymptotic distribution described above even when the extra parameter value is unknown but has been chosen using the proposed method.

4 Bayesian inference

The method for inference about the interest parameters using the classical approach is based on the asymptotic properties of the maximum likelihood estimator.

Table 1 *The median (M) and the interquartile range (IR) of the $R = 5000$ chosen extra parameter values by minimizing the Υ_ζ statistic*

		$n = 50$		$n = 100$		$n = 200$		$n = 400$		$n = 800$	
Distribution	ζ	M	IR	M	IR	M	IR	M	IR	M	IR
Log-Student-t	2	1.80	1.10	1.90	0.70	1.95	0.50	2.00	0.40	2.00	0.20
	4	3.30	3.10	3.60	2.30	3.80	1.80	3.90	1.23	3.95	0.90
	6	4.60	6.20	5.20	4.80	5.50	3.70	5.80	2.60	5.90	2.60
	8	5.40	9.50	6.40	7.90	7.00	6.10	7.60	4.50	7.85	3.20
Log-power-exp.	0.1	0.17	0.39	0.14	0.33	0.12	0.23	0.11	0.17	0.10	0.12
	0.2	0.26	0.46	0.23	0.35	0.22	0.25	0.21	0.17	0.20	0.13
	0.3	0.32	0.46	0.31	0.34	0.31	0.26	0.30	0.19	0.30	0.13
	0.4	0.39	0.44	0.39	0.38	0.40	0.26	0.40	0.19	0.40	0.14
Log-hyperbolic	1.0	0.95	2.00	1.00	1.60	1.00	1.10	1.00	0.80	1.00	0.50
	1.1	1.01	2.20	1.03	1.50	1.08	1.10	1.10	0.80	1.10	0.60
	1.2	1.05	2.20	1.10	1.60	1.13	1.20	1.18	0.80	1.20	0.60
	1.3	1.10	2.20	1.15	1.70	1.20	1.10	1.25	0.90	1.30	0.60
Log-slash	0.8	0.75	0.30	0.78	0.21	0.79	0.16	0.81	0.12	0.80	0.08
	0.9	0.80	0.34	0.85	0.26	0.89	0.20	0.89	0.14	0.90	0.10
	1.0	0.89	0.39	0.92	0.33	0.98	0.24	0.99	0.16	1.00	0.13
	1.1	0.97	0.41	1.03	0.36	1.06	0.25	1.08	0.20	1.10	0.15
Birnbaum–Saunders	1.5	1.20	1.80	1.30	0.90	1.40	0.60	1.40	0.50	1.50	0.30
	2.0	1.70	1.70	1.80	1.10	1.90	0.80	2.00	0.50	2.00	0.30
	2.5	2.10	1.90	2.30	1.30	2.40	0.80	2.40	0.50	2.50	0.40
	3.0	2.75	2.20	2.85	1.40	2.95	1.00	2.95	0.60	3.00	0.40

Therefore, for small sample sizes, such inference may be inadequate. Furthermore, some distributions, such as log-Laplace distribution, do not satisfy the usual regularity conditions; consequently, for such distributions, parameter inference cannot be performed using the standard method. Thus, this section considers inference using a Bayesian approach based on Markov chain Monte Carlo (MCMC) methods. One of the main advantages of Bayesian inference is that it is exact and available for any parametric model. For simplicity, it is supposed that η and ϕ are independent and have the following prior distributions:

$$\eta \sim \text{log-normal}(a_\eta, b_\eta) \quad \text{and} \quad \phi^{-1} \sim \text{Gamma}(c_\phi, d_\phi),$$

where $a_\eta > 0$, $b_\eta > 0$, $c_\phi > 0$ and $d_\phi > 0$ are assumed to be known. Next, we describe how samples are drawn from the posterior distribution of $\boldsymbol{\theta}$.

4.1 Log-normal, log-Student-*t*, log-slash, log-contaminated-normal, log-hyperbolic and log-Laplace distributions

One can sample from a joint posterior distribution of η and ϕ using Gibbs sampling (see, e.g., Gelfand and Smith, 1990), which involves successive sampling from the complete conditional densities. Because some distributions, such as the log-Student-t, log-slash, log-contaminated-normal, log-hyperbolic and log-Laplace distributions, may be obtained as a shape mixture of log-normal distributions, the algorithm can be easily developed using a data augmentation scheme, in which the complete conditional densities are known distributions. Next, a simple algorithm is described.

Algorithm 4.1.

Step 1. Set the initial value of the parameter vector to $\boldsymbol{\theta}^{(0)}$.

Step 2. Based on $\boldsymbol{\theta}^{(l)}$ sample $\mathbf{u}^{(l+1)} = (u_1^{(l+1)}, \ldots, u_n^{(l+1)})$ independent as follows:

(A) Log-normal: $\mathrm{P}[u_k^{(l+1)} = 1] = 1$.

(B) Log-Student-t:

$$u_k|\eta^{(l)}, \phi^{(l)} \sim \text{Gamma}\left(\frac{\zeta+1}{2}, \frac{1}{2}([\tilde{t}_k^{(l)}]^2 + \zeta)\right),$$

where $u \sim \text{Gamma}(a, b)$ represents a random variable with pro babil-ity density function given by $f(u) \propto u^{a-1}\exp(-bu)$.

(C) Log-slash:

$$u_k|\eta^{(l)}, \phi^{(l)} \sim \text{TGamma}\left(\frac{2\zeta+1}{2}, \frac{1}{2}[\tilde{t}_k^{(l)}]^2; (0, 1)\right),$$

where TGamma$(\cdot, \cdot; (0, 1))$ represents a random variable with truncated gamma distribution within the interval $(0, 1)$ (see, e.g., Nadarajah and Kotz, 2006).

(D) Log-contaminated-normal:

$$u_k|\eta^{(l)}, \phi^{(l)} \sim \begin{cases} \zeta_2, & \text{with probability } p \propto \zeta_2^{1/2}\zeta_1 \exp\left(-\frac{\zeta_2}{2}[\tilde{t}_k^{(l)}]^2\right), \\ 1, & \text{with probability } q \propto (1-\zeta_1)\exp\left(-\frac{1}{2}[\tilde{t}_k^{(l)}]^2\right), \end{cases}$$

(E) Log-Laplace:

$$u_k|\eta^{(l)}, \phi^{(l)} \sim \text{GIG}\left(\frac{1}{2}, [\tilde{t}_k^{(l)}]^2, \frac{1}{4}\right),$$

where $u \sim \text{GIG}(a, b, c)$ is a random variable with generali zed Inverse Gaussian distribution (see, e.g., Hörmann and Leydold, 2015) and density function given by

$$f(u) \propto u^{a-1}\exp\left(-\frac{1}{2}[b/u + cu]\right).$$

(F) Log-hyperbolic:

$$u_k|\eta^{(l)}, \phi^{(l)} \sim \text{GIG}\Big(\frac{1}{2}, [\tilde{t}_k^{(l)}]^2 + 1, \zeta^2\Big).$$

Step 3. Based on $\mathbf{u}^{(l+1)}$ and $\phi^{(l)}$, sample $\eta^{(l+1)}$ as follows:

$$\eta|\mathbf{u}^{(l+1)}, \phi^{(l)} \sim \text{log-normal}\big(a^{(l+1)}, b^{(l+1)}\big) \qquad \text{where}$$

(A) Log-normal, log-Student-t, log-slash and log-contaminated-normal:

$$b^{(l+1)} = \left(\sum_{k=1}^{n} \frac{u_k^{(l+1)}}{\phi^{(l)}} + \frac{1}{b_\eta}\right)^{-1} \quad \text{and}$$

$$a^{(l+1)} = \left[a_\eta^{\phi^{(l)}/b_\eta}\left(\prod_{k=1}^{n} t_k^{u_k^{(l+1)}}\right)\right]^{b^{(l+1)}/\phi^{(l)}}.$$

(B) Log-Laplace and log-hyperbolic:

$$b^{(l+1)} = \left(\sum_{k=1}^{n} \frac{1/u_k^{(l+1)}}{\phi^{(l)}} + \frac{1}{b_\eta}\right)^{-1} \quad \text{and}$$

$$a^{(l+1)} = \left[a_\eta^{\phi^{(l)}/b_\eta}\left(\prod_{k=1}^{n} t_k^{1/u_k^{(l+1)}}\right)\right]^{b^{(l+1)}/\phi^{(l)}}.$$

Step 4. Based on $\mathbf{u}^{(l+1)}$ and $\eta^{(l+1)}$, sample $\phi^{(l+1)}$ as follows:

$$\phi^{-1}|\mathbf{u}^{(l+1)}, \eta^{(l+1)} \sim \text{Gamma}\Big(\frac{n}{2} + c_\phi, d^{(l+1)}\Big) \qquad \text{where}$$

(A) Log-normal, log-Student-t, log-slash and log-contaminated-normal:

$$d^{(l+1)} = \frac{1}{2}\sum_{k=1}^{n}\left[\log\Big(\frac{t_k}{\eta^{(l+1)}}\Big)\right]^2 u_k^{(l+1)} + d_\phi.$$

(B) Log-Laplace and log-hyperbolic:

$$d^{(l+1)} = \frac{1}{2}\sum_{k=1}^{n}\frac{1}{u_k^{(l+1)}}\left[\log\Big(\frac{t_k}{\eta^{(l+1)}}\Big)\right]^2 + d_\phi.$$

Step 5. Repeat steps 2, 3 and 4 until convergence is reached.

4.2 Harmonic law, Log-power-exponential ($-1 < \zeta < 1$), Birnbaum–Saunders and Birnbaum–Saunders-*t* distributions

For these distributions, the posterior conditional densities of η given ϕ and ϕ given η are unknown. Therefore, samples from the complete conditional densities are drawn using the Metropolis–Hastings method (see, e.g., Chib and Greenberg, 1995). Next, a simple algorithm is presented.

Algorithm 4.2.

Step 1. Set the initial value of the parameter vector to $\boldsymbol{\theta}^{(0)}$.

Step 2. Based on $\boldsymbol{\theta}^{(l)}$ sample η^* from the log-normal$(a^{(l+1)}, b^{(l+1)})$ distribution, where

$$b^{(l+1)} = \left(\frac{n}{\lambda\phi^{(l)}} + \frac{1}{b_\eta}\right)^{-1} \quad \text{and} \quad a^{(l+1)} = \left[a_\eta^{\phi^{(l)}/b_\eta}\left(\prod_{k=1}^{n} t_k\right)\right]^{b^{(l+1)}/\lambda\phi^{(l)}},$$

where $\lambda > 0$ is a tunning parameter. Then, a new value $\eta^{(l+1)} = \eta^*$ is accepted with probability

$$\min\left\{1, \frac{f_\eta(\eta^*|\phi^{(l)})}{f_\eta(\eta^{(l)}|\phi^{(l)})}\right\},$$

$$\text{where } f_\eta(\eta|\phi) \propto \eta^{-1}\left[\prod_{k=1}^{n} g(\tilde{t}_k^2)\right] \exp\left\{-\frac{1}{2b_\eta}\left[\log\left(\frac{\eta}{a_\eta}\right)\right]^2\right\}.$$

Step 3. Based on $\eta^{(l+1)}$ and $\phi^{(l)}$ sample $1/\phi^*$ from the Gamma$(c^{(l+1)}, d^{(l+1)})$ distribution, where

$$c^{(l+1)} = \frac{n}{2} + c_\phi \quad \text{and} \quad d^{(l+1)} = \frac{1}{2\lambda}\sum_{k=1}^{n}\left[\log\left(\frac{t_k}{\eta^{(l+1)}}\right)\right]^2 + d_\phi.$$

Then, a new value $\phi^{(l+1)} = \phi^*$ is accepted with probability

$$\min\left\{1, \frac{f_\phi(\phi^*|\eta^{(l+1)})}{f_\phi(\phi^{(l)}|\eta^{(l+1)})}\right\},$$

$$\text{where } f_\phi(\phi|\eta) \propto \left[\prod_{k=1}^{n} g(\tilde{t}_k^2)\right](1/\phi)^{n/2+c_\phi-1}\exp\left(-\frac{d_\phi}{\phi}\right).$$

Step 4. Repeat steps 2 and 3 until convergence is reached.

The value of the tuning parameter in the Algorithm 4.2 may be set to $\lambda = \text{Var}(\log(T^*))$, where T^* exhibits a standard log-symmetric distribution.

4.3 Unknown extra parameter

When the extra parameter ζ is unknown, an additional *step* may be included in the algorithms described above to draw samples from the posterior conditional distribution of ζ given $\boldsymbol{\theta}$; it is assumed that ζ and $\boldsymbol{\theta}$ have independent prior distributions. For instance,

(A) Log-hyperbolic distribution: ζ is assumed to exhibit the log-normal(c_ζ, d_ζ) distribution. Thus, ζ^* is sampled from the log-normal $(\zeta^{(l)}, \lambda)$ distribution, where $\lambda > 0$ is a tunning parameter. A new value $\zeta^{(l+1)} = \zeta^*$ is accepted with probability given by

$$\min\left\{1, \frac{f_\zeta(\zeta^*|\mathbf{u}^{(l+1)})}{f_\zeta(\zeta^{(l)}|\mathbf{u}^{(l+1)})}\right\},$$

$$\text{where } f_\zeta(\zeta|\mathbf{u}) \propto \left[\mathrm{K}_1(\sqrt{\zeta})\right]^{-n} \exp\left(-\zeta\left[\frac{1}{2}\sum_{k=1}^{n} u_k\right] - \frac{1}{2d_\zeta}\left[\log\left(\frac{\zeta}{c_\zeta}\right)\right]^2\right).$$

(B) Log-slash distribution: ζ is assumed to exhibit the Gamma(c_ζ, d_ζ) distribution. Then, based on $\mathbf{u}^{(l+1)}$, $\zeta^{(l+1)}$ is sampled from the Gamma$(c_\zeta^{(l+1)}, d_\zeta^{(l+1)})$ distribution, where $c_\zeta^{(l+1)} = c_\zeta + n$ and $d_\zeta^{(l+1)} = -\sum_{k=1}^n \log(u_k^{(l+1)}) + d_\zeta$.

(C) Log-contaminated-normal distribution: $\zeta_1 \sim$ Beta$(c_{\zeta_1}, d_{\zeta_1})$ and $\zeta_2 \sim$ TGamma$(c_{\zeta_2}, d_{\zeta_2}; (0, 1))$ are independent. Then, based on $\mathbf{u}^{(l+1)}$, $\eta^{(l+1)}$ and $\phi^{(l+1)}$, $\zeta_1^{(l+1)}$ and $\zeta_2^{(l+1)}$ are sampled from the Beta$(c_{\zeta_1}^{(l+1)}, d_{\zeta_1}^{(l+1)})$ and TGamma$(c_{\zeta_2}^{(l+1)}, d_{\zeta_2}^{(l+1)}; (0, 1))$ distributions, where $c_{\zeta_1}^{(l+1)} = c_{\zeta_1} + \sum_{k=1}^n \mathrm{I}(u_k^{(l+1)} = \zeta_2)$, $d_{\zeta_1}^{(l+1)} = d_{\zeta_1} + \sum_{k=1}^n \mathrm{I}(u_k^{(l+1)} = 1)$, $c_{\zeta_2}^{(l+1)} = c_{\zeta_2} + \frac{1}{2}\sum_{k=1}^n \mathrm{I}(u_k^{(l+1)} = \zeta_2)$ and $d_{\zeta_2}^{(l+1)} = d_{\zeta_2} + \frac{1}{2}\sum_{k=1}^n [\tilde{t}_k^{(l+1)}]^2 \mathrm{I}(u_k^{(l+1)} = \zeta_2)$.

(D) Birnbaum–Saunders distribution: $1/\zeta^2$ is assumed to exhibit the Gamma(c_ζ, d_ζ) distribution. Based on $\eta^{(l+1)}$ and $\phi^{(l+1)}$, $1/[\zeta^{(l+1)}]^2$ is sampled from the Gamma$(c_\zeta^{(l+1)}, d_\zeta^{(l+1)})$ distribution, where $c_\zeta^{(l+1)} = c_\zeta + \frac{n}{2}$ and $d_\zeta^{(l+1)} = 2\sum_{k=1}^n \sinh^2(\tilde{t}_k^{(l+1)}) + d_\zeta$.

Maximum likelihood estimates for the same family of distributions may be used as initial values for Algorithms 4.1 and 4.2. Inferences about the parameters or functions of them are available from the approximate posterior marginal density. For example, we can summarize the simulated posterior distribution of η and ϕ by computing the summary statistics (i.e., the posterior means, medians, and standard deviations) and credible intervals. In the case of non-informative priors, comparisons with the maximum likelihood approach may be performed.

Table 2 *Values of $-2\ell(\hat{\boldsymbol{\theta}})$, AIC and BIC for the fitted distributions to the GDP data*

	Log-normal	Birnbaum–Saunders	log-skew-t	Box-cox-t	Generalized modified Weibull	Generalized Gamma	Generalized inverse Gaussian
$-2\ell(\hat{\boldsymbol{\theta}})$	3919.61	**3902.82**	3919.58	3919.58	3920.06	3921.62	3905.74
AIC	3923.60	**3906.82**	3927.58	3927.58	3928.06	3927.62	3911.75
BIC	3930.10	**3913.31**	3940.57	3940.57	3941.05	3937.36	3921.49

5 Application

Gross domestic product (GDP) divided by midyear population is known as the per capita GDP. The per capita GDP is most likely the best measure of a country's overall well being. The GDP is the sum of the gross value added by all resident producers in the economy and any product taxes and minus any subsidies that are not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. The dataset considered in this paper corresponds to the per capita GDP (current US$) of 190 countries during 2010, and it was downloaded from the World Bank's DataBank website (http://databank.worldbank.org/data/). All of the computations were performed using the R software package (R Core Team, 2013).

5.1 Maximum likelihood estimation

Table 2 lists the goodness-of-fit statistics $-2\ell(\hat{\boldsymbol{\theta}})$, AIC (Akaike, 1973) and BIC (Schwarz, 1978) for the log-normal, Birnbaum–Saunders($\zeta = 2.2$), log-skew-t, Box-cox-t, generalized modified Weibull, generalized Gamma and generalized Inverse Gaussian distributions fitted to the GDP data. The extra parameter ζ of the Birnbaum–Saunders distribution was chosen by minimizing the criterion Υ_ζ, as illustrated in Figure 3(a).

The Birnbaum–Saunders($\zeta = 2.2$) distribution has the lowest $-2\ell(\hat{\boldsymbol{\theta}})$, AIC and BIC values among all the fitted models; thus, it could be considered to be the best model. Figure 3(b) shows a plot of $\Phi^{-1}[F_T(\hat{\tilde{t}}^{(k)})]$ versus $\upsilon^{(k)}$ for the fitted Birnbaum–Saunders distribution; this plot indicates that the distribution describes the data adequately. A plot of the Birnbaum–Saunders($\zeta = 2.2$) density distribution is shown in Figure 4(a) (together with the data histogram). Similarly, Figure 4(b) presents the empirical cumulative distribution function of the per capita GDP and the cumulative distribution function of the Birnbaum–Saunders($\zeta = 2.2$) model. The MLEs (and the corresponding standard errors, which are given in parentheses) of the model parameters of the fitted Birnbaum–Saunders($\zeta = 2.2$) distribution are

$$\hat{\eta} = 4891.135(427.07) \quad \text{and} \quad \hat{\phi} = 3.187(0.21).$$

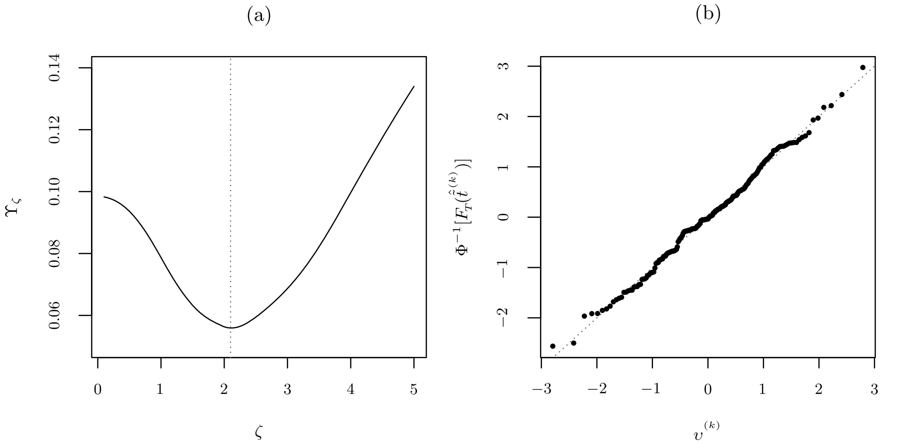

Figure 3 (a) *Graph of* Υ_ζ *under the Birnbaum–Saunders distribution*; (b) *plot of* $\Phi^{-1}[F_T(\hat{\tilde{t}}^{(k)})]$ *versus* $\upsilon^{(k)}$ *under the Birnbaum–Saunders*($\zeta = 2.2$) *distribution fitted to the GDP data.*

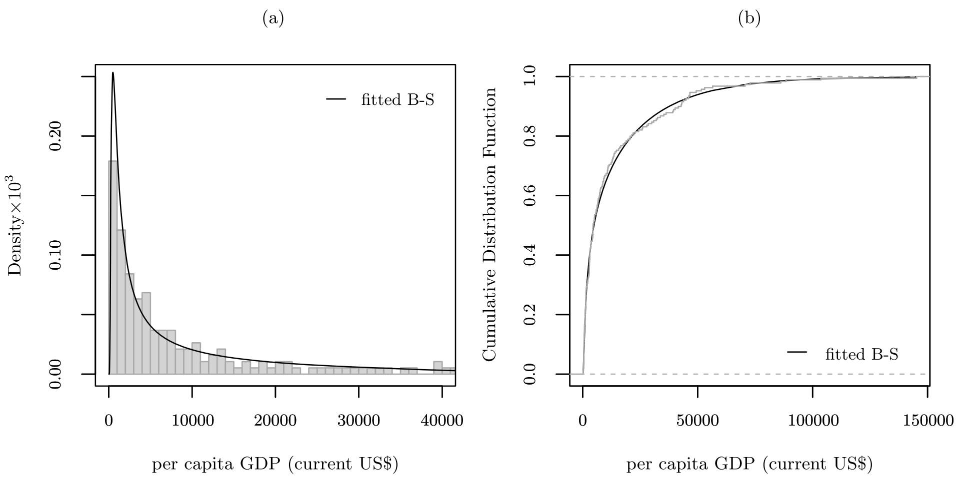

Figure 4 (a) *Histogram and* (b) *empirical cumulative distribution function for per capita GDP* (*current US$*) *of* 190 *countries during* 2010.

Because the Birnbaum–Saunders($\zeta = 2.2$) distribution was identified as the best model, and from the properties described in Section 2, one can conclude that the probability distribution of any macroeconomic indicator that can be expressed as $c_1 T^{c_2}$ also belongs to the log-symmetric class, where T re presents the per capita GDP during 2010 and $c_1 > 0$ and $c_2 \neq 0$ are known constants. The Birnbaum–Saunders($\zeta = 2.2$) distribution was also fitted to the per capita GDP for 2009; $\hat{\eta} = 4823.88(424.96)$ and $\hat{\phi} = 3.313(0.22)$ were obtained. Then, ignoring the variability associated with the point estimates of η and ϕ it may be concluded that the

Table 3 *Posterior mean, median, standard deviation (SD) and 95% credible interval for parameters of the log-normal and Birnbaum–Saunders distributions fitted to GDP data*

	Mean	Median	SD	2.5%	97.5%	DIC
			Log-normal			
η	4843.55	4808.20	541.03	3870.75	5978.52	3923.61
ϕ	2.35	2.33	0.25	1.92	2.88	
			Birnbaum–Saunders			
η	4837.86	4821.61	362.48	4181.02	5573.93	3910.80
ϕ	3.25	3.23	0.30	2.70	3.90	
ζ	2.20	2.19	0.19	1.84	2.60	

median of the per capita GDP distribution increased in 2010, whereas in the same year, both the skewness and relative dispersion of the per capita GDP distribution decreased. Similarly, the modes of the per capita GDP distributions were US$ 466.056 and US$ 500.174 in 2009 and 2010, respectively.

5.2 Bayesian inference

We now consider the prior distributions described in Section 4 with hyperparameters fixed as follows: $a_\eta = 1$, $b_\eta = 10{,}000$, $c_\phi = 0.0001$, $d_\phi = 0.0001$, $c_\zeta = 0.0001$, and $d_\zeta = 0.0001$. This setup allows for comparisons with the maximum likelihood approach. One chain of size 110,000 for each parameter was simulated, and the first 10,000 iterations were discarded to eliminate the effect of initial values. To avoid correlation, a spacing of size 10 was used, thereby obtaining an effective sample of size 10,000. Table 3 lists the summary statistics of the posterior distribution and the 95% credible interval for the parameters of the log-normal and Birnbaum–Saunders models fitted to the GDP data. The statistic DIC (see, e.g., Gelman et al., 2004) presented in Table 3 indicates that the Birnbaum–Saunders model describes the data better than the log-normal model. The inferential results are very similar to the results obtained using the maximum likelihood approach. Figure 5 displays the history of the chains and the approximate posterior marginal densities of the parameters η and ϕ for the Birnbaum–Saunders model.

Acknowledgments

The authors are grateful to two reviewers for their helpful comments and suggestions. This work receive financial support from CAPES, CNPq and FAPESP, Brazil.

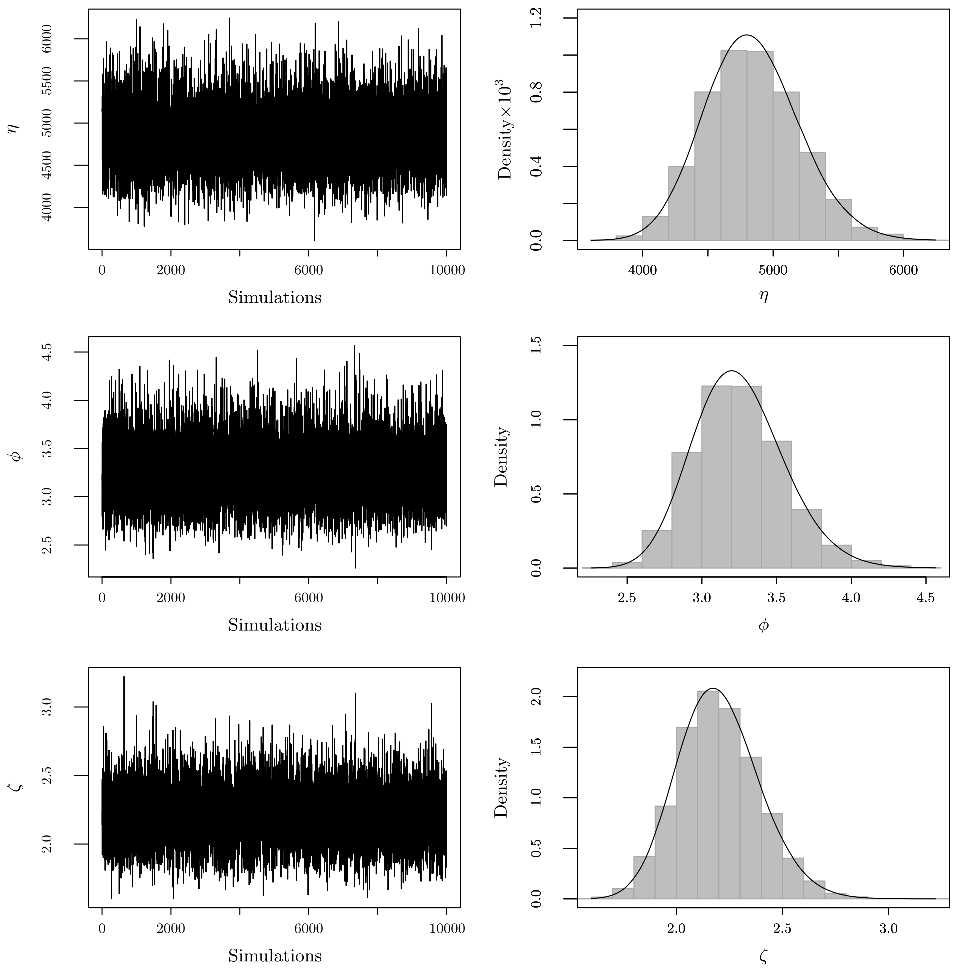

Figure 5 *History of the chains and the approximate posterior marginal densities of* η, ϕ *and* ζ *for the Birnbaum–Saunders distribution fitted to GDP data.*

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *International Symposium on Information Theory* (B. N. Petrov and F. Csaki, eds.) 267–281. Budapest, Hungary: Akademiai Kiado. MR0483125

Andrews, D. R. and Mallows, C. L. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B* **36**, 99–102. MR0359122

Azzalini, A., dal Cappello, T. and Kotz, S. (2003). Log-skew-normal and log-skew-*t* distributions as model for family income data. *Journal of Income Distribution* **11**, 12–20.

Balakrishnan, N., Leiva, V., Sanhueza, A. and Vilca, F. (2009). Estimation in the Birnbaum–Saunders distributions based on scale-mixture of normals and the EM-algorithm. *Statistics and Operations Research Transactions (SORT)* **33**, 171–192. MR2643505

Barndoff-Nielsen, O. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **353**, 401–419.

Barros, M., Paula, G. A. and Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: Robustness and diagnostics. *Lifetime Data Analysis* **14**, 316–332. MR2516848

Birnbaum, Z. W. and Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability* **6**, 319–327. MR0253493

Bonett, D. G. (2006). Confidence interval for a coefficient of quartile variation. *Computational Statistics & Data Analysis* **50**, 2953–2957. MR2239648

Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. Reading, MA: Addison-Wesely Publishing Company. MR0418321

Carrasco, J. M. F., Ortega, E. M. M. and Cordeiro, G. M. (2008). A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics & Data Analysis* **53**, 450–462. MR2649099

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis–Hastings algorithm. *The American Statistician* **49**, 327–335.

Cordeiro, G. M., Ferrari, S. L. P., Uribe-Opazo, M. A. and Vasconcellos, K. L. P. (2000). Corrected maximum-likelihood estimation in a class of symmetric nonlinear regression models. *Statistics & Probability Letters* **46**, 317–328. MR1743990

Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* **39**, 1–38. MR0501537

Díaz-García, J. A. and Leiva, V. (2005). A new family of life distributions based on elliptically contoured distributions. *Journal of Statistical Planning and Inference* **128**, 445–457. MR2102769

Fang, K. T., Kotz, S. and Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*. London: Chapman & Hall. MR1071174

Fonseca, T. C. O., Migon Helio, S. and Ferreira Marco, A. R. (2012). Bayesian analysis based on the Jeffreys prior for the hyperbolic distribution. *Brazilian Journal of Probability and Statistics* **26**, 327–343. MR2949082

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409. MR1141740

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2004). *Bayesian Data Analysis*. Boca Raton: Chapman & Hall. MR2027492

Glaser, R. E. (1980). Bathtub and related failure rate characterizations. *Journal of the American Statistical Association* **75**, 667–672. MR0590699

Groeneveld, R. A. and Meeden, G. (1984). Measuring skewness and kurtosis. *The Statistician* **33**, 391–399.

Hinkley, D. V. (1975). On power transformations to symmetry. *Biometrika* **62**, 101–111. MR0378205

Hörmann, W. and Leydold, J. (2015). Generating generalized inverse Gaussian random variates. *Statistics and Computing* **24**, 547–557. MR3223540

Jørgensen, B. (1982). *Statistical Properties of the Generalized Inverse Gaussian Distribution*. New York: Springer. MR0648107

Kano, Y., Berkane, M. and Bentler, P. (1993). Statistical inference based on pseudo-maximum likelihood estimators in elliptical populations. *Journal of the American Statistical Association* **88**, 135–143. MR1212483

Leiva, V., Riquelme, M., Balakrishnan, N. and Sanhueza, A. (2008). Lifetime analysis based on the generalized Birnbaum–Saunders distribution. *Computational Statistics & Data Analysis* **52**, 2079–2097. MR2418490

Limpert, E., Stahel, W. A. and Abbt, M. (2001). Log-normal distributions across the sciences: Key and clues. *BioScience* **51**, 341–352.

Lucas, A. (1997). Robustness of the Student-t based M-estimator. *Communications in Statistics, Theory and Methods* **26**, 1165–1182. MR1450228
Marchenko, Y. V. and Genton, M. G. (2010). Multivariate log-skew-elliptical distributions with applications to precipitation data. *Environmetrics* **21**, 318–340. MR2842246
Marshall, A. W. and Olkin, I. (2007). *Life Distributions*. New York: Springer. MR2344835
Moors, J. J. A. (1988). A quantile alternative for kurtosis. *The Statistician* **37**, 25–32.
Nadarajah, S. and Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematicae* **92**, 97–111. MR2265333
Paula, G. A., Leiva, V., Barros, M. and Liu, S. (2012). Robust statistical modeling using the Birnbaum–Saunders-t distribution applied to insurance distribution. *Applied Stochastic Model in Business and Industry* **28**, 16–34. MR2898899
Puig, P. (2008). A note on the harmonic law: A two-parameter family of distributions for ratios. *Statistics & Probability Letters* **78**, 320–326. MR2387553
R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at http://www.r-project.org.
Rieck, J. R. (1999). A moment-generating function with application to the Birnbaum–Saunders distribution. *Communications in Statistics—Theory and Methods* **28**, 2213–2222.
Rieck, J. R. and Nedelman, J. R. (1991). A log-linear model for the Birnbaum–Saunders distribution. *Technometrics* **33**, 51–60.
Rigby, R. A. and Stasinopoulos, D. M. (2006). Using the box-cox t distribution in GAMLSS to model skewness and kurtosis. *Statistical Modelling* **6**, 209–229. MR2252360
Rogers, W. H. and Tukey, J. W. (1972). Understanding some long-tailed symmetrical distributions. *Statistica Neerlandica* **26**, 211–226. MR0383605
Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464. MR0468014
Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics* **33**, 1187–1192. MR0143277
Villegas, C., Paula, G. A., Cysneiros, F. J. A. and Galea, M. (2013). Influence diagnostics in generalized symmetric linear models. *Computational Statistics & Data Analysis* **59**, 161–170. MR3000049
West, M. (1987). On scale mixtures of normal distributions. *Biometrika* **74**, 646–648. MR0909372
Zwillinger, D. and Kokoska, S. (2000). *Standard Probability and Statistical Tables and Formula*. Boca Raton: Chapman & Hall. MR1769133

Instituto de Matemática e Estatística
Universidade de São Paulo
São Paulo
Brazil
and
Departamento de Estadística
Universidad Nacional de Colombia
Bogotá
Colombia
E-mail: hvanegasp@gmail.com

Instituto de Matemática e Estatística
Universidade de São Paulo
São Paulo
Brazil
E-mail: giapaula@ime.usp.br