

## A new extended Birnbaum-Saunders model with cure fraction: classical and Bayesian approach

Edwin M. M. Ortega<sup>a</sup>, Gauss M. Cordeiro<sup>b</sup>, Adriano K. Suzuki<sup>1,c</sup>, Thiago G. Ramires<sup>a</sup>

<sup>a</sup>Departamento de Ciências Exatas, Universidade de São Paulo, Brazil;

<sup>b</sup>Departamento de Estatística, Universidade Federal de Pernambuco, Brazil;

<sup>c</sup>Departamento de Matemática Aplicada e Estatística, Universidade de São Paulo, Brazil

---

### Abstract

A four-parameter extended fatigue lifetime model called the odd Birnbaum-Saunders geometric distribution is proposed. This model extends the odd Birnbaum-Saunders and Birnbaum-Saunders distributions. We derive some properties of the new distribution that include expressions for the ordinary moments and generating and quantile functions. The method of maximum likelihood and a Bayesian approach are adopted to estimate the model parameters; in addition, various simulations are performed for different parameter settings and sample sizes. We propose two new models with a cure rate called the odd Birnbaum-Saunders mixture and odd Birnbaum-Saunders geometric models by assuming that the number of competing causes for the event of interest has a geometric distribution. The applicability of the new models are illustrated by means of ethylene data and melanoma data with cure fraction.

Keywords: Bayesian estimation, Birnbaum-Saunders distribution, lifetime data distribution, maximum likelihood estimation, simulations

---

### 1. Introduction

The two-parameter Birnbaum-Saunders (BS) distribution is a popular distribution for modeling lifetime data and phenomenon with monotone failure rates that has been applied in several research areas such as engineering, hydrology and survival analysis. However, it does not provide a reasonable parametric fit for non-monotone failure rates such as bathtub shaped and unimodal failure rates, which are common in reliability and biological studies. Bathtub hazard curves have nearly flat middle portions and the associated densities have a positive anti-mode. Unimodal failure rates can be observed in the course of a disease whose mortality reaches a peak after a finite period and then declines gradually.

Many lifetime distributions have been constructed with a view for applications in several areas such as survival analysis, reliability engineering, demography and actuarial studies, and hydrology. The BS model, also known as the fatigue life distribution, is a very popular model which has been extensively used for modeling the failure times of fatiguing materials and lifetime data in the fields cited above. Many generalizations of the BS model have been recently proposed by introducing shape parameters to better explore the tails and other properties of the generated distributions. For example, Cordeiro and Lemonte (2011) discussed the  $\beta$ BS model, Castillo *et al.* (2011) introduced a new extension of the BS distribution based on the family of the epsilon-skew-symmetric distributions, Ortega

---

<sup>1</sup> Corresponding author: Departamento de Matemática Aplicada e Estatística, Universidade de São Paulo, Avenida Trabalhador São-carlense, 400 - Centro CEP: 13566-59, São Carlos-SP, Brasil. E-mail: [suzuki@icmc.usp.br](mailto:suzuki@icmc.usp.br)

*et al.* (2012) developed the censored log-BS regression model, Cordeiro *et al.* (2013) proposed the McDonald-BS class of distributions, Lemonte (2013) defined a new extension of the BS distribution and Cordeiro and Lemonte (2014) studied exponentiated generalized BS distribution.

A random variable  $W$  having the  $BS(\lambda, \beta)$  distribution with shape parameter  $\lambda > 0$  and scale parameter  $\beta > 0$  is defined by

$$W = \beta \left[ \frac{\lambda U}{2} + \left\{ \left( \frac{\lambda U}{2} \right)^2 + 1 \right\}^{\frac{1}{2}} \right]^2,$$

where  $U$  is a standard normal random variable. The cumulative distribution function (cdf) of  $W$  is given by

$$G_{\lambda, \beta}(w) = \Phi(v), \quad w > 0, \quad (1.1)$$

where  $v = \lambda^{-1} \rho(w/\beta)$ ,  $\rho(u) = u^{1/2} - u^{-1/2}$  and  $\Phi(\cdot)$  is the standard normal cumulative function. The parameter  $\beta$  is the median of the distribution:  $G(\beta) = \Phi(0) = 1/2$ . For any  $k > 0$ ,  $kW \sim BS(\lambda, k\beta)$ . Kundu *et al.* (2008) discussed the shape of its hazard function. The probability density function (pdf) corresponding to (1.1) is

$$g_{\lambda, \beta}(w) = \kappa(\lambda, \beta) w^{-\frac{3}{2}} (w + \beta) \exp \left\{ -\frac{\tau(w/\beta)}{2\lambda^2} \right\}, \quad w > 0, \quad (1.2)$$

where  $\kappa(\lambda, \beta) = \exp(\lambda^{-2}) / (2\lambda \sqrt{2\pi\beta})$  and  $\tau(u) = u + u^{-1}$ .

The cdf of the odd Birnbaum-Saunders ("OBS" for short) distribution with an additional shape parameter  $\alpha > 0$  is defined by

$$\Pi_{\alpha, \lambda, \beta}(w) = \int_0^{\frac{G_{\lambda, \beta}(w)}{\bar{G}_{\lambda, \beta}(w)}} \frac{\alpha x^{\alpha-1}}{(1+x^\alpha)^2} dx = \frac{G_{\lambda, \beta}(w)^\alpha}{G_{\lambda, \beta}(w)^\alpha + \bar{G}_{\lambda, \beta}(w)^\alpha}, \quad (1.3)$$

where  $\bar{G}_{\lambda, \beta}(w) = 1 - G_{\lambda, \beta}(w)$ .

The BS cdf  $G_{\lambda, \beta}(w)$  is clearly a special case of (1.3) when  $\alpha = 1$ . We note that there is no complicated function in equation (1.3) in contrast with the  $\beta$ BS model (Cordeiro and Lemonte, 2011), which involves the beta incomplete function and two extra parameters. The OBS density is given by

$$\pi_{\alpha, \lambda, \beta}(w) = \frac{\alpha g_{\lambda, \beta}(w) \{G_{\lambda, \beta}(w) [1 - G_{\lambda, \beta}(w)]\}^{\alpha-1}}{\{G_{\lambda, \beta}(w)^\alpha + [1 - G_{\lambda, \beta}(w)]^\alpha\}^2}. \quad (1.4)$$

Therefore, the parameter  $\alpha$  represents the quotient of the log odds ratio for the generated and baseline distributions.

We now consider that the random variable  $W$  has the OBS cdf (1.3) with parameters  $(\alpha, \lambda, \beta)^T$ , say  $W \sim OBS(\alpha, \lambda, \beta)$ .

The cdf and pdf of  $W$  are given by

$$\Pi_{\alpha, \lambda, \beta}(w) = \frac{\Phi(v)^\alpha}{\Phi(v)^\alpha + [1 - \Phi(v)]^\alpha}, \quad w > 0, \quad (1.5)$$

and

$$\pi_{\alpha,\lambda,\beta}(w) = \frac{\alpha\kappa(\lambda,\beta)w^{-\frac{3}{2}}(w+\beta)\exp\left\{-\frac{\tau(w/\beta)}{2\lambda^2}\right\}\{\Phi(v)[1-\Phi(v)]\}^{\alpha-1}}{\{\Phi(v)^\alpha + [1-\Phi(v)]^\alpha\}^2}, \tag{1.6}$$

respectively.

The rest of the paper is organized as follows. In Section 2, we define the odd Birnbaum Saunders geometric (OBSG) model by compounding OBS and geometric distributions. Parameter inference is investigated in Section 3, where two methods (maximum likelihood and Bayesian approach) for parameter estimation are discussed. Further, various simulations are performed for different parameter settings and sample sizes. In Section 4, we introduce two models with long-term survivals and estimate their parameters by maximum likelihood and Bayesian approaches. Applications to two real data sets are addressed in Section 5. The paper provides concludes remarks in Section 6.

## 2. The odd Birnbaum-Saunders geometric model under activation scheme

Generalizing continuous univariate distributions by introducing additional shape parameters in a distribution is a useful practice to obtain multiple density and hazard rate shapes. Suppose that  $\{W_i\}_{i=1}^Z$  are independent and identically distributed (iid) random variables having the OBS distribution with cdf and pdf given by (1.5) and (1.6), respectively.

Let  $Z$  be a geometric random variable with probability mass function  $P(z; p) = (1 - p)p^{z-1}$  for  $z \in \mathbb{N}$  and  $p \in (0, 1)$ . We define the random variable  $T = \min(\{W_i\}_{i=1}^Z)$  having the OBSG distribution. We provide three motivations for the proposed distribution, where  $Z$  has a geometric distribution, which can be applied in some interesting situations such as: (i) Reliability: the new distribution can arise in series and parallel systems with identical components, which appear in many industrial applications and biological organisms; (ii) Times to the first and last failure: suppose the failure of a device occurs due to the presence of an unknown number  $Z$  of initial defects of the same kind, which can be identifiable only after causing failure and are repaired perfectly. Let  $W_i$  be the time to the failure of the device due to the  $i^{th}$  defect, for  $i \geq 1$ . Under the assumptions that the  $W_i$ 's are iid OBS random variables independent of  $Z$ , the OBSG distribution is appropriate for modeling the times to the first and last failures, respectively; (iii) Time to relapse of cancer under the first-activation scheme: suppose that an individual in the population is susceptible to a certain type of cancer. Let  $Z$  be the number of carcinogenic cells for an individual left active after the initial treatment and denote by  $W_i$  the time spent for the  $i^{th}$  carcinogenic cell to produce a detectable cancer mass, for  $i \geq 1$ . Under the assumptions that  $\{W_i\}_{i \geq 1}$  is a sequence of iid OBS random variables independent of  $Z$ , we conclude that the time to relapse of cancer of a susceptible individual has the OBSG distribution.

The conditional density function of  $T = \min(\{W_i\}_{i=1}^Z)$  given  $Z = z$  is

$$f(t|z; \alpha, \lambda, \beta) = \frac{z\alpha\kappa(\lambda,\beta)t^{-\frac{3}{2}}(t+\beta)\exp\left\{-\frac{\tau(t/\beta)}{2\lambda^2}\right\}\{\Phi(v)[1-\Phi(v)]\}^{\alpha-1}[1-\Phi(v)]^{\alpha(z-1)}}{\{\Phi(v)^\alpha + [1-\Phi(v)]^\alpha\}^2\{\Phi(v)^\alpha + [1-\Phi(v)]^\alpha\}^{z-1}},$$

where  $v = \lambda^{-1}\rho(t/\beta)$  and  $\rho(u) = u^{1/2} - u^{-1/2}$ .

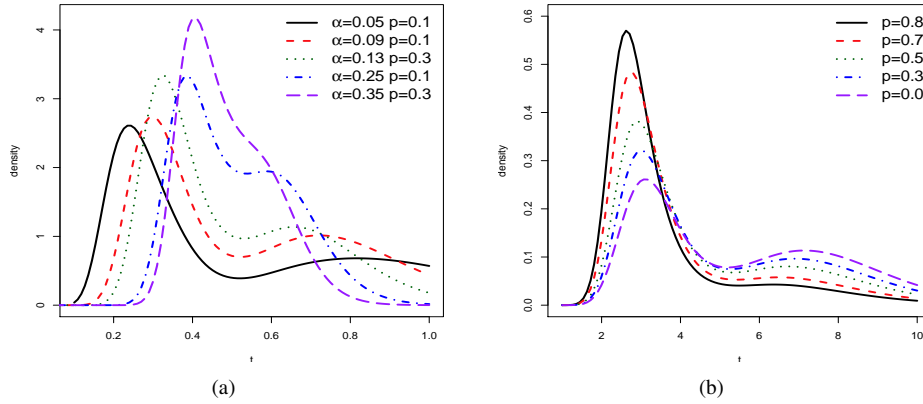


Figure 1: The odd Birnbaum-Saunders geometric density function for some parameter values: (a) For values  $\lambda = 0.1$  and  $\beta = 0.5$ , (b) For values  $\lambda = 0.1, \beta = 5.0$ , and  $\alpha = 0.1$ .

Then, the four-parameter OBSG density function (for  $t > 0$ ) becomes

$$f(t; \alpha, \lambda, \beta, p) = \frac{(1 - p) \alpha \kappa(\lambda, \beta) t^{-\frac{3}{2}} (t + \beta) \exp\left\{-\frac{\tau(t/\beta)}{2\lambda^2}\right\} \{\Phi(v)[1 - \Phi(v)]\}^{\alpha-1}}{\{\Phi(v)^\alpha + [1 - \Phi(v)]^\alpha\}^2} \times \left\{1 - \frac{p[1 - \Phi(v)]^\alpha}{\Phi(v)^\alpha + [1 - \Phi(v)]^\alpha}\right\}^{-2}, \tag{2.1}$$

where  $\beta > 0$  is a scale parameter and  $\alpha > 0, \lambda > 0$  and  $p \in (0, 1)$  are shape parameters. The random variable  $T$  having density function (2.1) is denoted by  $T \sim \text{OBSG}(\alpha, \lambda, \beta, p)$ . We can omit sometimes the dependence on the vector of parameters and write  $f(t) = f(t; \alpha, \lambda, \beta, p)$ , etc. The cdf of  $T$  (for  $t > 0$ ) is

$$F(t) = \frac{\Phi(v)^\alpha}{\Phi(v)^\alpha + (1 - p)[1 - \Phi(v)]^\alpha}. \tag{2.2}$$

Clearly, the OBS distribution is a special case of (2.2) when  $p \rightarrow 0^+$ . The BS distribution can also be obtained when  $p \rightarrow 0^+$  and  $\alpha = 1$ . The Birnbaum-Saunders geometric (BSG) distribution is obtained when  $\alpha = 1$ .

The hazard rate function (hrf) of  $T$  is given by

$$S(t) = \frac{(1 - p)[1 - \Phi(v)]^\alpha}{\Phi^\alpha(v) + (1 - p)[1 - \Phi(v)]^\alpha}.$$

Plots of the OBSG density, survival and hazard functions for selected parameter values are displayed in Figures 1–3, respectively.

We show some structural properties of the OBSG distribution in the Appendix.

### 3. Inference

#### 3.1. Maximum likelihood estimation

Several approaches for parameter estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. The maximum likelihood estimators (MLEs) enjoy

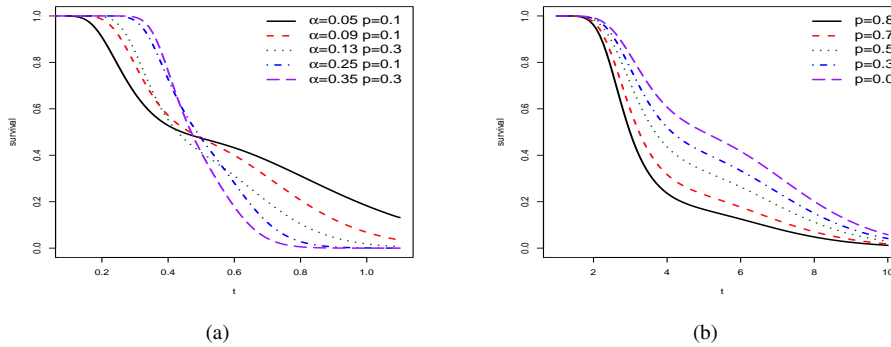


Figure 2: The odd Birnbaum-Saunders geometric survival function for some parameter values: (a) For values  $\beta = 3.0$  and  $p = 0.8$ , (b) For values  $\lambda = 0.15, \beta = 3.0$ , and  $\alpha = 0.3$ .

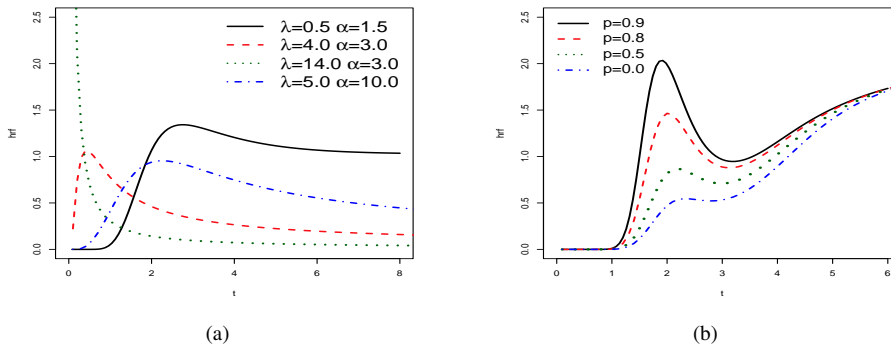


Figure 3: The odd Birnbaum-Saunders geometric hazard rate function for some parameter values: (a) For values  $\beta = 3.0$  and  $p = 0.8$ , (b) For values  $\lambda = 0.15, \beta = 3.0$ , and  $\alpha = 0.3$ .

desirable properties and can be used when constructing confidence intervals and regions as well as in test statistics. The normal approximation for these estimators in asymptotic theory is easily handled either analytically or numerically. Therefore, we consider the estimation of the unknown parameters for the OBSG model by maximum likelihood. Let  $T_i$  be a random variable following (2.1) with the vector of parameters  $\theta = (\alpha, \lambda, \beta, p)^T$ . Data encountered in survival analysis and reliability studies are often censored. A very simple random censoring mechanism that is often realistic is one in which each individual  $i$  is assumed to have a lifetime  $T_i$  and a censoring time  $C_i$ , where  $T_i$  and  $C_i$  are independent random variables. Suppose the data consist of  $n$  independent observations  $t_i = \min(T_i, C_i)$  for  $i = 1, \dots, n$ . The distribution of  $C_i$  does not depend on any unknown parameters of  $T_i$ . Parametric inference for such data is usually based on likelihood methods and their asymptotic theory. The censored log-likelihood  $l(\theta)$  for the model parameters is given by

$$\begin{aligned}
 l(\theta) = & n \log(1 - p) + r \log[\alpha \kappa(\lambda, \beta)] - \frac{3}{2} \sum_{i \in F} \log(t_i) + \sum_{i \in F} \log(t_i + \beta) - \frac{1}{2\lambda^2} \sum_{i \in F} \tau \left( \frac{t_i}{\beta} \right) \\
 & + (\alpha - 1) \sum_{i \in F} \log\{\Phi(v_i)[1 - \Phi(v_i)]\} - 2 \sum_{i \in F} \log\{\Phi(v_i)^\alpha + (1 - p)[1 - \Phi(v_i)]^\alpha\} \\
 & + \alpha \sum_{i \in C} \log[1 - \Phi(v_i)] - \sum_{i \in C} \log\{\Phi(v_i)^\alpha + (1 - p)[1 - \Phi(v_i)]^\alpha\},
 \end{aligned}$$

where  $v = \lambda^{-1}\rho(t_i/\beta)$ ,  $\rho(z) = z^{1/2} - z^{-1/2}$ ,  $r$  is the number of failures and  $F$  and  $C$  denote uncensored and censored sets of observations, respectively.

This log-likelihood can be maximized either directly by using the R (optim function; <https://www.r-project.org/>), SAS (NLMIXED procedure; SAS Institute Inc., Cary, NC, USA), or Ox (MaxBFGS function; <http://www.doornik.com/ox/>), or by solving the nonlinear likelihood equations by its differentiation. The score functions for the parameters  $\alpha$ ,  $\lambda$ ,  $\beta$ , and  $p$  are given by

$$\begin{aligned}
 U_\alpha(\boldsymbol{\theta}) &= \frac{r}{\alpha} + \sum_{i \in F} \log \{\Phi(v_i) [1 - \Phi(v_i)]\} + \sum_{i \in C} \log [1 - \Phi(v_i)] \\
 &\quad - 2 \sum_{i \in F} \frac{\Phi(v_i)^\alpha \log [\Phi(v_i)] + (1-p) [1 - \Phi(v_i)]^\alpha \log [1 - \Phi(v_i)]}{\Phi(v_i)^\alpha + (1-p) [1 - \Phi(v_i)]^\alpha} \\
 &\quad - \sum_{i \in C} \frac{\Phi(v_i)^\alpha \log [\Phi(v_i)] + (1-p) [1 - \Phi(v_i)]^\alpha \log [1 - \Phi(v_i)]}{\Phi(v_i)^\alpha + (1-p) [1 - \Phi(v_i)]^\alpha}, \\
 U_\lambda(\boldsymbol{\theta}) &= r \frac{\kappa_\lambda(\lambda, \beta)}{\kappa(\lambda, \beta)} + \frac{1}{\lambda^3} \sum_{i \in F} \tau \left( \frac{t_i}{\beta} \right) + (\alpha - 1) \sum_{i \in F} \phi(v_i) v_{i(\lambda)} \left\{ \frac{1 - 2\Phi(v_i)}{\Phi(v_i) [1 - \Phi(v_i)]} \right\} \\
 &\quad - \alpha \sum_{i \in C} \frac{\phi(v_i) v_{i(\lambda)}}{1 - \Phi(v_i)} - 2\alpha \sum_{i \in F} \phi(v_i) v_{i(\lambda)} \left\{ \frac{\Phi(v_i)^{\alpha-1} - (1-p) [1 - \Phi(v_i)]^{\alpha-1}}{\Phi(v_i)^\alpha + (1-p) [1 - \Phi(v_i)]^\alpha} \right\} \\
 &\quad - \alpha \sum_{i \in C} \phi(v_i) v_{i(\lambda)} \left\{ \frac{\Phi(v_i)^{\alpha-1} - (1-p) [1 - \Phi(v_i)]^{\alpha-1}}{\Phi(v_i)^\alpha + (1-p) [1 - \Phi(v_i)]^\alpha} \right\}, \\
 U_\beta(\boldsymbol{\theta}) &= \frac{r\kappa_\beta(\lambda, \beta)}{\kappa(\lambda, \beta)} + \sum_{i \in F} \frac{1}{t_i + \beta} - \frac{1}{2\lambda^2} \sum_{i \in F} \left\{ \frac{1}{t_i} - \frac{t_i}{\beta^2} \right\} + (\alpha - 1) \sum_{i \in F} \phi(v_i) v_{i(\beta)} \left\{ \frac{1 - 2\Phi(v_i)}{\Phi(v_i) [1 - \Phi(v_i)]} \right\} \\
 &\quad - 2\alpha \sum_{i \in F} v_{i(\beta)} \phi(v_i) \left\{ \frac{\Phi^{\alpha-1}(v_i) - (1-p) [1 - \Phi(v_i)]^{\alpha-1}}{\Phi^\alpha(v_i) + (1-p) [1 - \Phi(v_i)]^\alpha} \right\} - \alpha \sum_{i \in C} \frac{v_{i(\beta)} \phi(v_i)}{1 - \Phi(v_i)} \\
 &\quad - \alpha \sum_{i \in C} v_{i(\beta)} \phi(v_i) \left\{ \frac{\Phi^{\alpha-1}(v_i) - (1-p) [1 - \Phi(v_i)]^{\alpha-1}}{\Phi^\alpha(v_i) + (1-p) [1 - \Phi(v_i)]^\alpha} \right\}, \\
 U_p(\boldsymbol{\theta}) &= \frac{-n}{1-p} + 2 \sum_{i \in F} \frac{[1 - \Phi(v_i)]^\alpha}{\Phi(v_i)^\alpha + (1-p) [1 - \Phi(v_i)]^\alpha} + \sum_{i \in C} \frac{[1 - \Phi(v_i)]^\alpha}{\Phi^\alpha(v_i) + (1-p) [1 - \Phi(v_i)]^\alpha},
 \end{aligned}$$

where

$$\begin{aligned}
 \kappa_\lambda(\lambda, \beta) &= \frac{\exp(\lambda^{-2})(-\lambda^2 - 2)}{2\lambda^4 \sqrt{2\pi\beta}}, & \kappa_\beta(\lambda, \beta) &= \frac{-\exp(\lambda^{-2})}{2\lambda \sqrt{2\pi}} \left( \frac{\beta^{-\frac{3}{2}}}{2} \right), \\
 v_{i(\lambda)} &= \frac{1}{\lambda^2} \left[ \left( \frac{t_i}{\beta} \right)^{-\frac{1}{2}} - \left( \frac{t_i}{\beta} \right)^{\frac{1}{2}} \right], & v_{i(\beta)} &= \frac{-\lambda^{-1}}{2} \left[ \left( \frac{t_i}{\beta^3} \right)^{\frac{1}{2}} + \left( \frac{1}{\beta t_i} \right)^{\frac{1}{2}} \right].
 \end{aligned}$$

The MLE  $\hat{\boldsymbol{\theta}}$  of  $\boldsymbol{\theta}$  can be obtained by solving the nonlinear likelihood equations  $U_\alpha(\boldsymbol{\theta}) = 0$ ,  $U_\lambda(\boldsymbol{\theta}) = 0$ ,  $U_\beta(\boldsymbol{\theta}) = 0$ , and  $U_p(\boldsymbol{\theta}) = 0$ . We employ the NLMIXED procedure in SAS because these equations cannot be solved analytically.

For interval estimation of  $(\alpha, \lambda, \beta, p)$  and tests of hypotheses on these parameters we obtain the observed information matrix since the expected information matrix requires numerical integration.

The elements of the  $4 \times 4$  observed information matrix  $J(\boldsymbol{\theta})$ , namely  $J_{\alpha\alpha}$ ,  $J_{\alpha\lambda}$ ,  $J_{\alpha\beta}$ ,  $J_{\alpha p}$ ,  $J_{\lambda\lambda}$ ,  $J_{\lambda\beta}$ ,  $J_{\lambda p}$ ,  $J_{\beta\beta}$ ,  $J_{\beta p}$ , and  $J_{pp}$ , can be calculated numerically.

Under standard regularity conditions, the asymptotic distribution of  $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$  is  $N_4(\mathbf{0}, I(\boldsymbol{\theta})^{-1})$ , where  $I(\boldsymbol{\theta})$  is the expected information matrix. This asymptotic behavior is valid if  $I(\boldsymbol{\theta})$  is replaced by  $J(\hat{\boldsymbol{\theta}})$ , i.e., the observed information matrix evaluated at  $\hat{\boldsymbol{\theta}}$ . The multivariate normal  $N_4(\mathbf{0}, J(\hat{\boldsymbol{\theta}})^{-1})$  distribution can then be used to construct approximate confidence intervals for the individual parameters.

Further, we can compute the maximum values of the unrestricted and restricted log-likelihoods to obtain the likelihood ratio (LR) statistics to test some sub-models of the OMSG distribution. For example, the test of  $H_0 : \alpha = 1$  versus  $H_1 : \alpha \neq 1$  is equivalent to compare the BSG distribution with the OMSG distribution. In this case, the LR statistic is

$$w = 2 \left\{ l(\hat{\alpha}, \hat{\lambda}, \hat{\beta}, \hat{p}) - l(1, \tilde{\lambda}, \tilde{\beta}, \tilde{p}) \right\},$$

where  $\hat{\alpha}$ ,  $\hat{\lambda}$ ,  $\hat{\beta}$ , and  $\hat{p}$  are the MLEs under  $H_1$  and  $\tilde{\lambda}$ ,  $\tilde{\beta}$ , and  $\tilde{p}$  are the estimates under  $H_0$ .

### 3.2. Bayesian estimation

An alternative statistical approach is to use the Bayesian method that allows the incorporation of previous knowledge of the parameters through an informative prior distribution. A non-informative prior structure may be considered when previous knowledge is unavailable. In this work we adopt proper prior distributions according to variations of the parametric space, but ensuring non-informativeness according to the fixed hyper-parameters that lead to such a situation. We assume  $\alpha$ ,  $\lambda$ ,  $\beta$ , and  $p$  are mutually independent, that is, we consider a prior distribution of  $\boldsymbol{\theta} = (\alpha, \lambda, \beta, p)$  given by  $\pi(\boldsymbol{\theta}) = \pi(\alpha)\pi(\lambda)\pi(\beta)\pi(p)$ , where  $\alpha \sim \text{Gamma}(0.1, 0.01)$ ,  $\lambda \sim \text{Gamma}(0.1, 0.01)$ ,  $\beta \sim \text{Gamma}(0.1, 0.01)$  and  $p \sim \text{Beta}(1, 1)$ , i.e., all the hyper-parameters have been specified to express non-informative priors. In order to generate our samples we then implement an Metropolis-Hasting algorithm within Gibbs iterations (Chib and Greenberg, 1995). Simulations were performed using the OpenBUGS software (Spiegelhalter *et al.*, 2007). In all computations we simulate two chains of size 60,000 for each parameter, disregarding the first 20,000 iterations to eliminate the effect of the initial values and to avoid autocorrelation problems. We then consider a spacing of size 10 and obtain an effective sample of size 8,000, upon which the posterior inference is based on. The MCMC convergence was monitored according to the methods recommended by Cowles and Carlin (1996), as well as trace plots.

#### 3.2.1. Model comparison criteria

In the literature, there are various methodologies which intend to analyze the suitability of distribution, as well as select the best fit from among a collection of distributions. One of the most used approaches is derived from the conditional predictive ordinate (CPO) statistic. For a detailed discussion on the CPO statistic and its applications to model selection, see Gelfand *et al.* (1992).

Let  $\mathcal{D}$  denote the full data and  $\mathcal{D}^{(-i)}$  denote the data with the deleted  $i^{\text{th}}$  observation. We denote the posterior density of  $\boldsymbol{\theta}$  given  $\mathcal{D}^{(-i)}$  by  $\pi(\boldsymbol{\theta}|\mathcal{D}^{(-i)})$ ,  $i = 1, \dots, n$ . For the  $i^{\text{th}}$  observation,  $\text{CPO}_i$  can be written as

$$\text{CPO}_i = \int_{\boldsymbol{\theta} \in \Theta} f(t_i|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\mathcal{D}^{(-i)}) d\boldsymbol{\theta} = \left\{ \int_{\boldsymbol{\theta}} \frac{\pi(\boldsymbol{\theta}|\mathcal{D})}{f(t_i|\boldsymbol{\theta})} d\boldsymbol{\theta} \right\}^{-1}. \quad (3.1)$$

The  $\text{CPO}_i$  can be interpreted as the height of the marginal density of the time for an event at  $t_i$ . Therefore, high  $\text{CPO}_i$  implies a better fit of the model. For model comparisons we use the log pseudo

Table 1: The mean, biases, and MSEs of the OMSG distribution for  $\alpha = 0.7$ ,  $\lambda = 0.8$ ,  $\beta = 2$ , and  $p = 0.1$ , and  $n = 100, 200, 300$ , and  $500$

Sample size	Parameter (true value)	Classical estimation			Bayesian estimation		
		Mean	Bias	MSE	Mean	Bias	MSE
N = 100	$p$ (0.1)	0.1238	0.0238	0.0220	0.1227	0.0227	0.0270
	$\alpha$ (0.7)	0.6562	-0.0438	0.0713	0.6949	-0.0051	0.0511
	$\lambda$ (0.8)	0.7553	-0.0447	0.0588	0.8323	0.0323	0.0453
	$\beta$ (2.0)	2.0410	0.0410	0.0543	2.0377	0.0377	0.0324
N = 200	$p$ (0.1)	0.1169	0.0169	0.0163	0.1194	0.0194	0.0261
	$\alpha$ (0.7)	0.6920	-0.0080	0.0513	0.6970	-0.0030	0.0399
	$\lambda$ (0.8)	0.7953	-0.0047	0.0419	0.8234	0.0234	0.0340
	$\beta$ (2.0)	2.0562	0.0562	0.0412	2.0339	0.0339	0.0198
N = 300	$p$ (0.1)	0.1097	0.0097	0.0129	0.1207	0.0207	0.0256
	$\alpha$ (0.7)	0.6952	-0.0048	0.0395	0.6874	-0.0126	0.0353
	$\lambda$ (0.8)	0.7988	-0.0012	0.0319	0.8105	0.0105	0.0295
	$\beta$ (2.0)	2.0376	0.0376	0.0293	2.0285	0.0285	0.0140
N = 500	$p$ (0.1)	0.1080	0.0080	0.0097	0.1218	0.0218	0.0250
	$\alpha$ (0.7)	0.7017	0.0017	0.0238	0.7027	0.0027	0.0205
	$\lambda$ (0.8)	0.8076	0.0076	0.0191	0.8211	0.0211	0.0179
	$\beta$ (2.0)	2.0328	0.0328	0.0205	2.0228	0.0228	0.0098

MSEs = means squared errors; OMSG = odd Birnbaum-Saunders geometric.

marginal likelihood (LPML) defined by  $LPML = \sum_{i=1}^n \log(\widehat{CPO}_i)$ . The higher the LPML value, the better the fit of the model.

Other criteria, such as the deviance information criterion (DIC) proposed by Spiegelhalter *et al.* (2002), the expected Akaike information criterion (EAIC)-Brooks (2002), and the expected Bayesian (or Schwarz) information criterion (EBIC)-Carlin and Louis (2001) can also be used. These other criteria are based on the posterior mean of deviance, which can be approximated by  $\bar{d} = \sum_{q=1}^Q d(\theta_q)/Q$ , where  $d(\theta) = -2 \sum_{i=1}^n \log [f(t_i|\theta)]$ . The DIC criterion can be estimated using the MCMC output by  $\widehat{DIC} = \bar{d} + \hat{\rho}_d = 2\bar{d} - \hat{d}$ , where  $\rho_D$  is the effective number of parameters defined as  $E\{d(\theta)\} - d\{E(\theta)\}$ , where  $d\{E(\theta)\}$  is the deviance evaluated at the posterior mean. Similarly, EAIC and EBIC criteria can be estimated by means of  $\widehat{EAIC} = \bar{d} + 2\#(\theta)$  and  $\widehat{EBIC} = \bar{d} + \#(\theta) \log(n)$ , where  $\#(\theta)$  is the number of model parameters. Comparing alternative models, the preferred model is the one with the smallest criteria values.

### 3.3. Simulation

We simulate the OMSG distribution (for  $\alpha = 0.7$ ,  $\lambda = 0.8$ ,  $\beta = 2$ ,  $p = 0.1$ ) from equation (A.3) by using a random variable  $U$  with a uniform distribution in  $(0, 1)$ . We simulate  $n = 100, 200, 300$ , and  $500$  variates and, for each replication, we compute the MLEs and the posterior mean of the parameter  $\alpha$ ,  $\lambda$ ,  $\beta$ , and  $p$ . We repeat this process 1,000 times for classical inference and 200 times for the Bayesian procedure. We determine the average estimates (mean), biases, and means squared errors (MSEs). The results are given in Table 1 provides the results.

The results of the Monte Carlo study in Table 1 indicate that the MSEs of the mean of  $\alpha$ ,  $\lambda$ ,  $\beta$  and  $p$  decay toward zero as the sample size increases, as expected under standard asymptotic theory. As  $n$  increases, the mean of the parameters tend to be closer to the true parameter values. These results support a asymptotic normal distribution that provides an adequate approximation to the finite sample distribution of the MLEs. The normal approximation can often be improved using bias adjustments for these estimators. Approximations to the their biases in simple models may be obtained analytically.



Bias correction typically does a very good job correct the MLEs; however, it may either increase the MSEs. Whether bias correction is useful in practice depends on the shape of the bias function and the variance of the MLE. Several cumulants of log likelihood derivatives need to be obtained in order to improve the accuracy of the estimators using analytical bias reduction; however, these derivatives are notoriously cumbersome for the proposed model. Also, the results under a Bayesian approach (Table 1) are similar to those obtained via MLEs.

#### 4. The odd Birnbaum-Saunders model for survival data with long-term survivors

For censored survival times, the presence of an immune proportion of individuals who are not subject to death, failure or relapse may be indicated by a relatively high number of individuals with large censored survival times. Here, the OBS model is modified in two different ways for the possible presence of long-term survivors in the data.

##### 4.1. The odd Birnbaum-Saunders mixture model

In population-based cancer studies, cure is said to occur when mortality in the group of cancer patients returns to the same level as that expected in the general population. The cure fraction is a useful measure that is of interest to patients analyzing trends in cancer patient survival. Models for survival analysis typically assume that every subject in the study population is susceptible to the event under study and will eventually experience such event if follow-up is sufficiently long. However, there are situations when a fraction of individuals are not expected to experience the event of interest such as when those individuals are cured or not susceptible. Cure rate models have been used to model time-to-event data for various types of cancers, including breast cancer, non-Hodgkin's lymphoma, leukemia, prostate cancer, and melanoma. The most popular type of cure rate models are often the mixture models (MMs) introduced by Boag (1949), Berkson and Gage (1952), and Farewell (1982). Additionally, MMs allow both the cure fraction and the survival function of uncured patients (latency distribution) to depend on covariates. Longini and Halloran (1996) and Price and Manatunga (2001) have further introduced frailty to MMs for individual survival data. Further, Peng and Dear (2000) investigated a nonparametric MM for cure estimation, Sy and Taylor (2000) considered estimation in a proportional hazards cure model and Yu and Peng (2008) extended MMs to bivariate survival data by modeling marginal distributions. Fachini *et al.* (2014) recently proposed a scale model for bivariate survival times based on the copula that model the dependence of bivariate survival data with cure fraction, Hashimoto *et al.* (2015) introduced the new long-term survival model with interval-censored data, Ortega *et al.* (2015) proposed the power series beta Weibull regression model to predict breast carcinoma, Lanjoni *et al.* (2016) conducted extended Burr XII regression models and Ortega *et al.* (2017) proposed regression models generated by gamma random variables with long-term survivors.

To formulate the odd Birnbaum-Saunders mixture (OBSM) model, we consider that the population under study is a mixture of susceptible (uncured) individuals, who may experience the event of interest, and non-susceptible (cured) individuals, who never will experience the event of interest (Maller and Zhou, 1996). This approach allows to estimate simultaneously whether the event of interest will occur, which is called incidence, and if it does occur then when it will occur, which is called latency. Let  $N_i$  (for  $i = 1, \dots, n$ ) be the indicator denoting that the  $i^{th}$  individual is susceptible ( $N_i = 1$ ) or non-susceptible ( $N_i = 0$ ). The MM is given by

$$S_{\text{pop}}(t_i) = \pi + (1 - \pi)S(t_i|N_i = 1), \quad (4.1)$$

where  $S_{\text{pop}}(t_i)$  is the unconditional survival function of  $T_i$  for the entire population,  $S(t_i|N_i = 1)$  is

the survival function for susceptible individuals and  $\pi = P(N_i = 0)$  is the probability of cure for an individual. The OBSM model is defined by assuming that the survival function for susceptible individuals in (4.1) is given by

$$S(t_i|N_i = 1) = \frac{[1 - \Phi(v_i)]^\alpha}{\Phi(v_i)^\alpha + [1 - \Phi(v_i)]^\alpha}. \quad (4.2)$$

Let  $\theta = (\pi, \alpha, \lambda, \beta)^\top$  be the vector of parameters of interest for the OBSM model. For  $\pi = 0$ , the OBSM model reduces to the OBS model.

Suppose we have data in the form  $t_i, i = 1, \dots, n$ , where  $t_i$  denotes the observed logarithm of time for the  $i^{\text{th}}$  subject, i.e.,  $t_i = \min\{T_i, C_i\}$ ,  $T_i$  is the lifetime for the  $i^{\text{th}}$  individual and  $C_i$  is the censoring time for the  $i^{\text{th}}$  individual. Under this assumption, the contribution of an individual that failed at  $t_i$  to the likelihood function is given by

$$\frac{(1 - \pi) \alpha \kappa(\lambda, \beta) t_i^{-\frac{3}{2}} (t_i + \beta) \exp\left\{-\frac{\tau(t_i/\beta)}{2\lambda^2}\right\} \{\Phi(v_i)[1 - \Phi(v_i)]^{\alpha-1}\}}{\{\Phi(v_i)^\alpha + [1 - \Phi(v_i)]^\alpha\}^2}$$

and the contribution of an individual that is at risk at  $t_i$  is

$$\pi + \frac{(1 - \pi)[1 - \Phi(v_i)]^\alpha}{\Phi(v_i)^\alpha + [1 - \Phi(v_i)]^\alpha}.$$

The full log-likelihood function for the parameter vector  $\theta = (\pi, \alpha, \lambda, \beta)^\top$  is given by

$$\begin{aligned} l(\theta) = & r \log[(1 - \pi) \alpha \kappa(\lambda, \beta)] - \frac{3}{2} \sum_{i \in F} \log(t_i) + \sum_{i \in F} \log(t_i + \beta) - \frac{1}{2\lambda^2} \sum_{i \in F} \tau\left(\frac{t_i}{\beta}\right) \\ & + (\alpha - 1) \sum_{i \in F} \log\{\Phi(v_i)[1 - \Phi(v_i)]\} - 2 \sum_{i \in F} \log\{\Phi^\alpha(v_i) + [1 - \Phi(v_i)]^\alpha\} \\ & + \sum_{i \in C} \log\left\{\pi + \frac{(1 - \pi)[1 - \Phi(v_i)]^\alpha}{\Phi(v_i)^\alpha + [1 - \Phi(v_i)]^\alpha}\right\}, \end{aligned} \quad (4.3)$$

where  $F$  and  $C$  denote the sets of individuals who failed and are censoring, respectively. The MLE  $\hat{\theta}$  of  $\theta$  can be obtained by maximization of (4.3). The asymptotic covariance matrix of  $\hat{\theta}$  can be estimated by the Hessian matrix evaluated at  $\hat{\theta}$ .

#### 4.2. The odd Birnbaum-Saunders geometric model with cure fraction

Models with a cure fraction play an important role in survival analysis. Recently, some of these models have been published in the literature. For example, Cooner *et al.* (2007) proposed a flexible cure rate under latent activation schemes, Ortega *et al.* (2009) presented the generalized log-gamma regression model with cure fraction and Cancho *et al.* (2011) defined a flexible cure rate survival model by considering that the number of competing causes of the event of interest follows the Conway-Maxwell distribution and the time for the event has the generalized gamma distribution. Recently, Hashimoto *et al.* (2014) proposed the Poisson BS model with long-term survivors and Suzuki *et al.* (2016) studied the Poisson inverse-Gaussian regression cure rate model. In the same context, we present the OBSG model with a cure fraction.

For an individual in the population, let  $N$  denote the unobservable number of causes of the event of interest for this individual. We assume that  $N$  has a geometric distribution. The time for the  $j^{\text{th}}$

cause to produce the event of interest is denoted by  $W_j, j = 1, \dots, N$ . We consider that, conditional on  $N$ , the  $W_j$ 's are iid random variables with the OBS cdf. Further, we assume that  $N$  is independent of  $W_1, W_2, \dots$ . The observable time to the event is defined by  $T = \min\{W_1, \dots, W_N\}$ , if  $N \geq 1$  and  $T = \infty$  if  $N = 0$ , with  $P(T = \infty | N = 0) = 1$ . Under this setup, the survival function for the population is given by

$$S_{\text{pop}}(t) = P(N = 0) + P(W_1 > t, \dots, W_N > t | N \geq 1) P(N \geq 1). \tag{4.4}$$

Tsodikov *et al.* (2003), among others, demonstrated that  $S_{\text{pop}}(t) = A_p[S(t)]$ , where  $A_p(\cdot)$  is the probability generating function (pgf) of the number of competing causes ( $N$ ). We consider that  $N$  has the geometric distribution with parameter  $0 \leq p < 1$ . The pgf of  $N$  is given by  $A_p(w) = \sum_{m=0}^{\infty} p_m w^m = (1-p)/(1-pw)$ ,  $0 \leq w \leq 1$ , and then the survival function for the population becomes

$$S_{\text{pop}}(t) = A_p[S(t)] = (1-p) \left\{ 1 - \frac{p[1 - \Phi(v)]^\alpha}{\Phi(v)^\alpha + [1 - \Phi(v)]^\alpha} \right\}^{-1}. \tag{4.5}$$

Equation for  $S_{\text{pop}}(t)$  is called *improper survival function* because it has two main properties:

$$\lim_{t \rightarrow 0} S_{\text{pop}}(t) = 1 \quad \text{and} \quad \lim_{t \rightarrow \infty} S_{\text{pop}}(t) = p_0,$$

where  $p_0$  denoted the cured fraction, in this model the cured fraction is given by  $p_0 = 1 - p$ . The pdf corresponding to (4.5) reduces to

$$\begin{aligned} f_{\text{pop}}(t) &= p(1-p)\alpha \kappa(\lambda, \beta) t^{-\frac{3}{2}}(t+\beta) \exp\left\{-\frac{\tau(t/\beta)}{2\lambda^2}\right\} \{\Phi(v)[1 - \Phi(v)]\}^{\alpha-1} \\ &\times \{\Phi(v)^\alpha + [1 - \Phi(v)]^\alpha\}^{-2} \left\{ 1 - \frac{p[1 - \Phi(v)]^\alpha}{\Phi(v)^\alpha + [1 - \Phi(v)]^\alpha} \right\}^{-2}. \end{aligned} \tag{4.6}$$

Equation (4.6) refers to the OMSG model with long-term survivors in a competitive risk structure. The associated population hrf is

$$\begin{aligned} h_{\text{pop}}(t) &= p\alpha \kappa(\lambda, \beta) t^{-\frac{3}{2}}(t+\beta) \exp\left\{-\frac{\tau(t/\beta)}{2\lambda^2}\right\} \{\Phi(v)[1 - \Phi(v)]\}^{\alpha-1} \\ &\times \{\Phi(v)^\alpha + [1 - \Phi(v)]^\alpha\}^{-2} \left\{ 1 - \frac{p[1 - \Phi(v)]^\alpha}{\Phi(v)^\alpha + [1 - \Phi(v)]^\alpha} \right\}^{-1}. \end{aligned}$$

We note that  $f_{\text{pop}}(t)$  and  $h_{\text{pop}}(t)$  are improper functions, since  $S_{\text{pop}}(t)$  is not a proper survival function.

We consider a situation where the time to the event of interest is not completely observed but subjected to right censoring. Let  $C_i$  denote the censoring time. We observe  $t_i = \min\{T_i, C_i\}$ , for  $i = 1, \dots, n$ . Let  $\theta = (p, \alpha, \lambda, \beta)^T$  be the parameter vector, from  $n$  pairs of times and censoring indicators  $t_1, \dots, t_n$ . The full log-likelihood function under non-informative censoring can be expressed as

$$\begin{aligned} l(\theta) &= r \log[p(1-p)\alpha \kappa(\lambda, \beta)] + (n-r) \log(1-p) - \frac{3}{2} \sum_{i \in F} \log(t_i) + \sum_{i \in F} \log(t_i + \beta) \\ &- \frac{1}{2\lambda^2} \sum_{i \in F} \tau\left(\frac{t_i}{\beta}\right) + (\alpha-1) \sum_{i \in F} \log\{\Phi(v_i)[1 - \Phi(v_i)]\} - 2 \sum_{i \in F} \log\{\Phi^\alpha(v_i) + [1 - \Phi(v_i)]^\alpha\} \\ &- 2 \sum_{i \in F} \log\left\{ 1 - \frac{p[1 - \Phi(v_i)]^\alpha}{\Phi(v_i)^\alpha + [1 - \Phi(v_i)]^\alpha} \right\} - \sum_{i \in C} \log\left\{ 1 - \frac{p[1 - \Phi(v_i)]^\alpha}{\Phi(v_i)^\alpha + [1 - \Phi(v_i)]^\alpha} \right\}, \end{aligned} \tag{4.7}$$

Table 2: Mean, biases, and MSEs of the OBSGcr distribution for  $\alpha = 1.5$ ,  $\lambda = 0.5$ ,  $\beta = 2.0$ , and  $p = 0.25$  and  $n = 100, 200, 300$ , and  $500$

Sample size	Parameter (true value)	Classical estimation			Bayesian estimation		
		Mean	Bias	MSE	Mean	Bias	MSE
$N = 100$	$p$ (0.25)	0.2595	0.0095	0.0026	0.2197	-0.0303	0.0043
	$\alpha$ (1.5)	1.3903	-0.1097	0.0845	1.5038	0.0038	0.0816
	$\lambda$ (0.5)	0.4737	-0.0263	0.0156	0.6379	0.1379	0.0382
	$\beta$ (2.0)	2.0158	0.0158	0.0454	2.3942	0.3942	0.3605
$N = 200$	$p$ (0.25)	0.2544	0.0044	0.0015	0.2379	-0.0121	0.0016
	$\alpha$ (1.5)	1.3790	-0.1210	0.0948	1.5196	0.0196	0.0763
	$\lambda$ (0.5)	0.4737	-0.0263	0.0128	0.5573	0.0573	0.0160
	$\beta$ (2.0)	1.9872	-0.0128	0.0194	2.1135	0.1135	0.0688
$N = 300$	$p$ (0.25)	0.2533	0.0033	0.0009	0.2443	-0.0057	0.0012
	$\alpha$ (1.5)	1.4113	-0.0887	0.0872	1.5860	0.0860	0.0878
	$\lambda$ (0.5)	0.4854	-0.0146	0.0104	0.5547	0.0547	0.0119
	$\beta$ (2.0)	1.9966	-0.0034	0.0132	2.0483	0.0483	0.0191
$N = 500$	$p$ (0.25)	0.2476	-0.0024	0.0005	0.2463	-0.0037	0.0005
	$\alpha$ (1.5)	1.4343	-0.0657	0.0767	1.5209	0.0209	0.0824
	$\lambda$ (0.5)	0.4844	-0.0156	0.0067	0.5252	0.0252	0.0087
	$\beta$ (2.0)	2.0176	0.0176	0.0058	2.0316	0.0316	0.0058

MSEs = means squared errors; OBSGcr = odd Birnbaum-Saunders geometric cure.

where  $r$  is the number of failures and  $F$  and  $C$  denote the uncensored and censored sets of observations, respectively.

The MLE  $\hat{\theta}$  of the parameter vector  $\theta$  can be determined by maximizing (4.7). We use the NLMIXED procedure in SAS to compute  $\hat{\theta}$ . The inference method for  $\theta = (p, \alpha, \lambda, \beta)^T$  can be based on the normal approximation

$$(\hat{p}, \hat{\alpha}, \hat{\lambda}, \hat{\beta})^T \sim N_4 \left\{ (p, \alpha, \lambda, \beta)^T, J^{-1}(\hat{\theta}) \right\}, \quad (4.8)$$

where  $J^{-1}(\theta) = \{-\partial^2 l(\theta) / \partial \theta \theta^T\}$  is the  $4 \times 4$  observed information matrix.

The Bayesian procedure was used analogously to Section 3.2. We simulate the OBSG cure fraction distribution (for  $\alpha = 1.5$ ,  $\lambda = 1.5$ ,  $\beta = 0.5$ ,  $p = 0.25$ ) and the OBSM distribution (for  $\alpha = 2.0$ ,  $\lambda = 0.5$ ,  $\beta = 1.5$ ,  $p = 0.2$ ). Censoring times were then sampled from the uniform distribution in the  $(0, \tau)$  interval, where  $\tau$  controlled the censoring proportion of the uncured population. Here, the proportions of censored observations were approximately 42% and 38%, respectively.

Analogously to the previous simulation study, 1,000 data sets for classical inference and 200 data sets for Bayesian procedure were generated each one with  $n = 100, 200, 300$ , and  $500$  variates. For each replication, we compute the MLEs and the posterior mean of the parameter  $\alpha$ ,  $\lambda$ ,  $\beta$ , and  $p$ . We determine the average estimates (mean), biases, and MSEs.

Tables 2 and 3 show the results for the OBSGcr model and OBSM model, respectively. Similar results to the previously simulation study can also be observed.

## 5. Applications

In this section, we provide two applications to real data to demonstrate the flexibility of the OBS, OBSM, and OBSG models. The computations are performed using the NLMixed subroutine in the SAS package.

Table 3: The mean, biases, and MSE of the OMSG mixture distribution for  $\alpha = 2.0, \lambda = 0.5, \beta = 1.5$ , and  $p = 0.2$  and  $n = 100, 200, 300$ , and  $500$

Sample size	Parameter (true value)	Classical estimation			Bayesian estimation		
		Mean	Bias	MSE	Mean	Bias	MSE
N = 100	$p$ (0.2)	0.1960	-0.0040	0.0019	0.1838	-0.0162	0.0027
	$\alpha$ (2.0)	1.8903	-0.1097	0.0837	2.2731	0.2731	0.1187
	$\lambda$ (0.5)	0.4820	-0.018 0	0.0074	0.6573	0.1573	0.0291
	$\beta$ (1.5)	1.4378	-0.0622	0.0063	1.4437	-0.0563	0.0061
N = 200	$p$ (0.2)	0.2034	0.0034	0.0012	0.1943	-0.0057	0.0011
	$\alpha$ (2.0)	1.9346	-0.0654	0.0946	2.1971	0.1971	0.0844
	$\lambda$ (0.5)	0.4891	-0.0109	0.0066	0.5926	0.0926	0.0122
	$\beta$ (1.5)	1.4331	-0.0669	0.0055	1.4352	-0.0648	0.0053
N = 300	$p$ (0.2)	0.2047	0.0047	0.0008	0.1987	-0.0013	0.0008
	$\alpha$ (2.0)	1.9164	-0.0836	0.0921	2.1151	0.1151	0.0820
	$\lambda$ (0.5)	0.4914	-0.0086	0.0062	0.5588	0.0588	0.0082
	$\beta$ (1.5)	1.4342	-0.0658	0.0051	1.4357	-0.0643	0.0048
N = 500	$p$ (0.2)	0.2024	0.0024	0.0004	0.1992	-0.0008	0.0004
	$\alpha$ (2.0)	1.8891	-0.1109	0.0912	2.074	0.0740	0.0838
	$\lambda$ (0.5)	0.4815	-0.0185	0.0050	0.5425	0.0425	0.0067
	$\beta$ (1.5)	1.4371	-0.0629	0.0045	1.4370	-0.0630	0.0044

MSE = means squared error; OMSG = odd Birnbaum-Saunders geometric.

Table 4: MLEs of the model parameters for the ethylene data, the corresponding standard errors, and the AIC, CAIC, and BIC statistics

Model	Estimates					AIC	CAIC	BIC
KwN( $\mu, \sigma, a, b$ )	-32.77 (2.55)	29.40 (0.81)	13.47 (1.42)	0.45 (0.03)		5775.1	5775.2	5792.9
BN( $\mu, \sigma, a, b$ )	-56.17 (2.16)	32.24 (0.96)	50.93 (2.57)	0.41 (0.02)		5709.9	5710.0	5727.7
McN( $\mu, \sigma, a, b, c$ )	-186.04 (7.92)	47.99 (1.77)	10021.00 (8.85)	0.46 (0.03)	4.63 (0.63)	5638.3	5638.4	5660.5
KwOLLN( $\mu, \sigma, a, b, \alpha$ )	6.53 (2.77)	113.18 (14.60)	2.26 (0.33)	0.27 (0.01)	11.29 (2.31)	5547.0	5547.1	5569.3
BS( $\lambda, \beta$ )	0.71 (0.02)	27.47 (0.73)				5424.0	5424.0	5432.9
OBS( $\alpha, \lambda, \beta$ )	0.21 (0.01)	0.28 (0.01)	27.81 (0.87)			5425.8	5425.9	5439.2
<b>OMSG(<math>\alpha, \lambda, \beta, p</math>)</b>	<b>0.249</b> <b>(0.013)</b>	<b>0.269</b> <b>(0.008)</b>	<b>30.641</b> <b>(0.920)</b>	<b>0.309</b> <b>(0.072)</b>		<b>5254.8</b>	<b>5254.8</b>	<b>5272.5</b>

MLEs = maximum likelihood estimators; AIC = Akaike information criterion; CAIC = consistent Akaike information criterion; BIC = Bayesian information criterion; KwN = Kumaraswamy normal; BN = beta normal; McN = McDonald normal; KwOLLN = Kumaraswamy odd log-logistic-normal; BS = Birnbaum-Saunders; OBS = odd Birnbaum-Saunders; OMSG = odd Birnbaum-Saunders geometric.

### 5.1. Application 1: Ethylene data

These data are taken from a study by the University of São Paulo, ESALQ (Laboratory of Physiology and Post-colheita Biochemistry), which evaluated the effects of mechanical damage on banana fruits (genus *Musa* spp.); see Saavedra del Aguila *et al.* (2010) for more details. The major problem affecting bananas during and after harvest is the susceptibility of the mature fruit to physical damage caused by transport and marketing. Ethylene is a plant hormone important in post-harvest fruit. A high ethylene production can generate a fast senescence of fruit. We use  $n = 630$  points on ethylene data and assume a parent normal distribution.

Table 5: Likelihood ratio tests

Ethylene	Hypotheses	Statistic $w$	$p$ -value
OBSG vs. OBS	$H_0 : p = 0$ vs. $H_1 : H_0$ is false	173.0	<0.0001
OBSG vs. BS	$H_0 : \alpha = 1$ and $p = 0$ vs. $H_1 : H_0$ is false	173.2	<0.0001

OBSG = odd Birnbaum-Saunders geometric; OBS = odd Birnbaum-Saunders; BS = Birnbaum-Saunders.

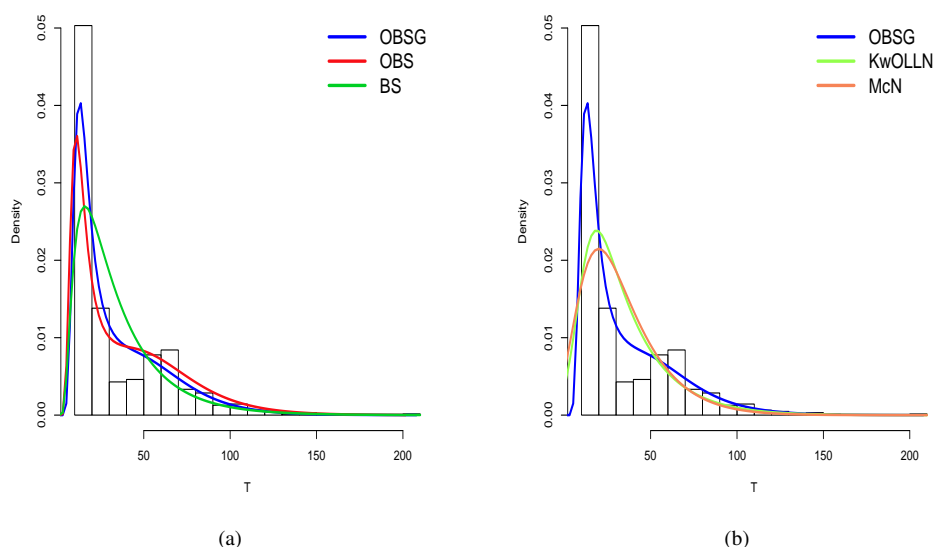


Figure 4: Estimated densities for the: (a) OBSG, OBS, and BS models; (b) OBSG, KwOLLN, and McN models fitted to the ethylene data. OBSG = odd Birnbaum-Saunders geometric; OBS = odd Birnbaum-Saunders; BS = Birnbaum-Saunders; KwOLLN = Kumaraswamy odd log-logistic-normal; McN = McDonald normal.

Recently, Alizadeh *et al.* (2015) proved that the Kumaraswamy odd log-logistic-normal (KwOLLN) distribution yields a better fit to these data than the McDonald normal (McN) (Cordeiro *et al.*, 2012), beta normal (BN) (Alexander *et al.*, 2012), and Kumaraswamy normal (KwN) (Cordeiro *et al.*, 2010). We now compare the results of the fits of the KwOLLN, McN, BN, and KwN models with the OBSG model as well as compare two sub-models of the OBSG model, namely the OBS and BS models.

Table 4 lists the MLEs (and the corresponding standard errors (SEs) in parentheses) of the model parameters and the values for the fitted models of the statistics: Akaike information criterion (AIC), Bayesian information criterion (BIC), and consistent Akaike information criterion (CAIC). The figures in this table indicate that the OBSG model could be selected as the best model because it has the lowest AIC, BIC, and CAIC values among the values for the fitted models. In addition, SEs of the estimates for all fitted models are also quite small.

Formal tests for the extra skewness parameters in the OBSG model can be based on LR statistics described in Section 3. Table 5 list the values of applying the LR statistics to the ethylene data. We then reject the null hypotheses of the two LR tests in favor of the OBSG distribution (Table 5). More information is provided by a visual comparison of the histogram of the data with the fitted density functions. Figure 4 displays the plots of the fitted OBSG, OBS, KwOLLN, and McN densities.

The empirical scaled total-time-on-test (TTT) transform can be used to identify the shape of the hazard function. The TTT plot for the ethylene data given in Figure 5(a) shows a unimodal and bathtub-shaped hrf and therefore indicates the appropriateness of the OBSG distribution as indicated

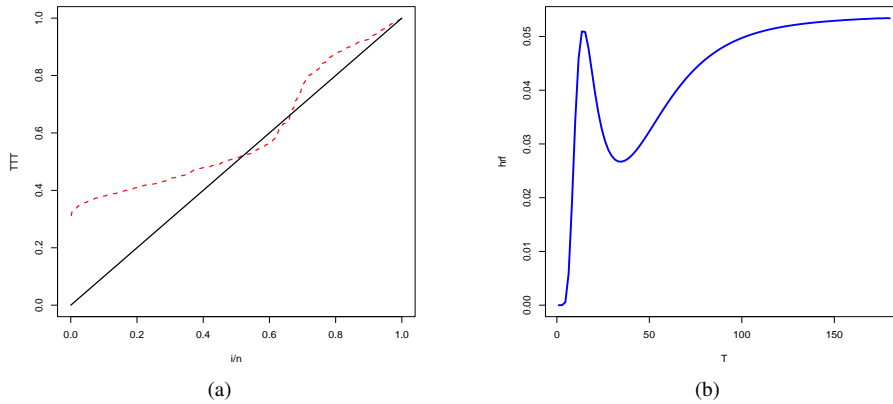


Figure 5: TTT plot (a) and estimated hazard function (b) for the OMSG distribution for the ethylene data. TTT = total-time-on-test; OMSG = odd Birnbaum-Saunders geometric.

Table 6: Posterior mean, SD, and 95% CI of the parameters for the ethylene data and the DIC, EAIC, EBIC, and LPML statistics

	Parameter	Mean	SD	95% CI		DIC	EAIC	EBIC	LPML
				2.50%	97.50%				
OBSG	$\alpha$	0.267	0.010	0.249	0.287	5220.931	5227.269	5245.052	-2610.026
	$\beta$	31.854	0.141	31.481	31.996				
	$\lambda$	0.266	0.003	0.262	0.274				
	$p$	0.305	0.012	0.276	0.320				
OBS	$\alpha$	0.248	0.002	0.244	0.250	5347.109	5353.03	5366.368	-2673.559
	$\beta$	27.943	0.057	27.786	27.998				
	$\lambda$	0.287	0.001	0.286	0.290				
BS	$\beta$	27.521	0.724	26.110	28.962	5423.964	5425.992	5434.883	-2711.748
	$\lambda$	0.716	0.020	0.677	0.757				

SD = standard deviation; CI = credible interval; DIC = deviance information criterion; EAIC = expected Akaike information Scriterion; EBIC = expected Bayesian information criterion; LPML = log pseudo marginal likelihood; OBSG = odd Birnbaum-Saunders geometric; OBS = odd Birnbaum-Saunders; BS = Birnbaum-Saunders.

by Figure 3. Further, Figure 5(b) reveals that the estimated hrf of the OMSG distribution can capture both behaviors (unimodal and bathtub-shaped) are presented in the ethylene data. The OMSG distribution provides a closer fit to the histogram of the current data than the other models.

Under a Bayesian approach we also fitted the OMSG, OBS, and BS models for the ethylene data as well as obtained the values DIC, EAIC, EBIC, and LPML criteria to compare these models. Table 6 shows the results that indicate estimates close to those obtained via MLEs; therefore, the OMSG model presented the best fit when using criteria found in Table 6.

### 5.2. Application 2: Melanoma data with cure fraction

Here, we discuss thee application of the OBSM and OMSG models with cure fraction to cancer recurrence with long-term survivors. These data are part of a study on cutaneous melanoma (a type of malignant cancer) for the evaluation of postoperative treatment performance with a high dose of a certain drug (interferon alfa-2b) in order to prevent recurrence. Patients from the years 1991 to 1995 were included in the study with follow-up conducted until 1998. The data were collected by Ibrahim *et al.* (2001). The survival time  $T$  is defined as the time until the patient’s death. The original sample

Table 7: MLEs of the model parameters for the melanoma data, the corresponding standard errors (given in parentheses), and the AIC, CAIC, and BIC statistics

Model	$\pi$	$\alpha$	$\lambda$	$\beta$	AIC	CAIC	BIC
OBSM	0.4672 (0.0384)	7.3671 (3.8938)	5.9259 (1.1123)	1.8335 (0.1705)	1057.2	1057.3	1073.3
BSM	0.4475 (0.0512)	1	0.9378 (0.0998)	1.8843 (0.2582)	1062.5	1062.6	1074.6
	$p$	$\alpha$	$\lambda$	$\beta$			
OBSG	0.5511 (0.0495)	504.560 (29.2503)	441.570 (33.4233)	2.9234 (0.6284)	1058.1	1058.2	1074.2
BSG	0.6201 (0.0877)	1	1.3107 (0.3799)	4.6027 (2.6620)	1069.5	1069.6	1081.6

MLEs = maximum likelihood estimators; AIC = Akaike information criterion; CAIC = consistent Akaike information criterion; BIC = Bayesian information criterion; OBSM = odd Birnbaum-Saunders mixture; BSM = Birnbaum-Saunders mixture; OBSG = odd Birnbaum-Saunders geometric; BSG = Birnbaum Saunders geometric.

Table 8: Posterior mean, SD, and 95% CI of the parameters for the melanoma data and the DIC, EAIC, EBIC, and LPML statistics

Parameter	Mean	SD	95% CI		DIC	EAIC	EBIC	LPML	
			2.50%	97.50%					
OBSM	$\pi$	0.442	0.057	0.310	0.526	1055.365	1060.597	1076.73	-527.929
	$\alpha$	7.063	2.076	2.597	9.886				
	$\lambda$	6.057	1.856	2.275	9.253				
	$\beta$	1.989	0.324	1.605	2.758				
BSM	$\pi$	0.450	0.032	0.402	0.520	1061.477	1065.042	1077.141	-531.093
	$\lambda$	0.954	0.081	0.808	1.123				
	$\beta$	1.920	0.206	1.567	2.369				
OBSG	$p$	0.569	0.032	0.510	0.633	1053.903	1060.063	1076.195	-527.008
	$\alpha$	491.035	12.154	470.983	510.809				
	$\lambda$	450.027	26.685	403.082	497.458				
	$\beta$	3.187	0.377	2.530	3.911				
BSG	$\pi$	0.615	0.023	0.569	0.662				
	$\lambda$	1.305	0.073	1.168	1.451	1067.451	1071.499	1083.598	-534.168
	$\beta$	4.506	0.288	4.020	4.978				

SD = standard deviation; CI = credible interval; DIC = deviance information criterion; EAIC = expected Akaike information criterion; EBIC = expected Bayesian information criterion; LPML = log pseudo marginal likelihood; OBSM = odd Birnbaum-Saunders mixture; BSM = Birnbaum-Saunders mixture; OBSG = odd Birnbaum-Saunders geometric; BSG = Birnbaum Saunders geometric.

size was  $n = 427$  patients, 10 of whom did not present a value for the explanatory variable tumor thickness. When such cases were removed, a sample of size  $n = 417$  patients was retained. A sample of size  $n = 417$  patients was retained after the removal of these cases. The percentage of censored observations was 56%. Tables 7 and 8 lists the MLEs and Bayesian estimates of the model parameters.

The values of AIC, CAIC, and BIC statistics (Table 7) and of the DIC, EAIC, EBIC, and LPML statistics (Table 8) are smaller for the OBSM and OBSG models when compared to the values of the Birnbaum-Saunders mixture (BSM) and BSG models thus indicating that the first two models could be chosen as the best models to fit these data.

The LR statistics given in Table 9 indicate that the OBSM and OBSG models are superior to BSM and BSG models in terms of model fitting.

In order to assess if the model is appropriate, Figures 6 and 7 display the empirical survival functions and the estimated OBS, OBSM, BSG, and OBSG survival functions, respectively, from which a significant fraction of survivors can be observed. It can be verified that the OBSM and OBSG



Table 9: Likelihood ratio tests

Melanoma	Hypotheses	Statistic $w$	$p$ -value
OBSM vs. BSM	$H_0 : \alpha = 1$ vs. $H_1 : H_0$ is false	7.3	0.0069
OBSG vs. BSG	$H_0 : \alpha = 1$ vs. $H_1 : H_0$ is false	13.0	0.0003

OBSM = odd Birnbaum-Saunders mixture; BSM = Birnbaum-Saunders mixture; OBSG = odd Birnbaum-Saunders geometric; BSG = Birnbaum Saunders geometric.

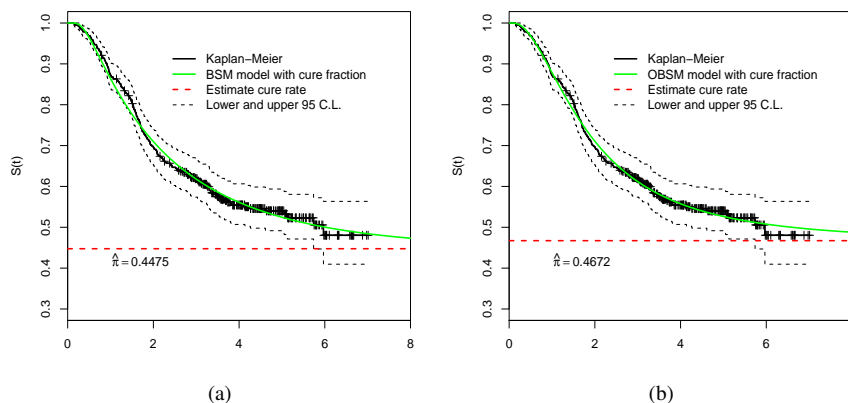


Figure 6: (a) Kaplan-Meier curves and BSM model and the estimated cure fraction for the cutaneous melanoma data and (b) Kaplan-Meier curves and estimated survival function for the OBSM model. BSM = Birnbaum-Saunders mixture; OBSM = odd Birnbaum-Saunders mixture; CL = confidence level.

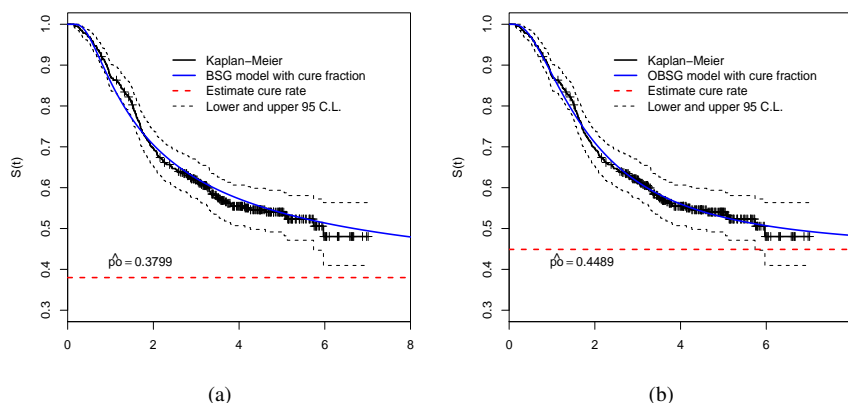


Figure 7: (a) Kaplan-Meier curves and BSG model and the estimated cure fraction for the cutaneous melanoma data and (b) Kaplan-Meier curves and estimated survival function for the OBSG model. BSG = Birnbaum Saunders geometric; OBSG = odd Birnbaum-Saunders geometric.

models with cure fraction provide good fits to these data.

### 6. Conclusions

Many lifetime distributions have been constructed with a view for applications in several areas, in particular, survival analysis, reliability engineering, demography, and actuarial studies, and hydrol-

ogy. Generalizing continuous univariate distributions by introducing additional shape parameters is an important way to provide very different shapes for the density and hrf of the generated model. We introduce and study a new lifetime model called the OMSG distribution. We derive some general mathematical properties based on a simple mixture representations for its density function. Cure rate models have been used to model time-to-event data for various types of cancers. Based on OMSG distribution, we define two new models with cure rate called the OMSM and OMSG models, where the unobservable number  $N$  of causes of the event of interest has a geometric distribution and the time for the  $j^{\text{th}}$  cause to produce this event is modeled by a random variable  $W_j$  such that, conditional on  $N$ , the  $W_j$ 's are considered iid random variables with the OBS distribution independent of  $N$ . Several simulation studies are performed for different parameter settings, sample sizes, and censoring percentages. Maximum likelihood and Bayesian approaches are described to estimate the model parameters. Two empirical applications to real data are presented in order to illustrate the flexibility of the proposed models.

### Appendix: Structural properties of the OMSG distribution

The necessity and importance of the quantile function, moments and generating function is always important in any statistical analysis such as applied work.

#### Quantile function

Let  $Q_{\text{OBS}}(u) = \Pi_{\alpha, \lambda, \beta}(t)^{-1}(u)$  be the qf of the OBS model (for  $0 < u < 1$ ). Inverting  $F(t) = u$  in (2.2), we obtain the qf of  $T$  as

$$t = F^{-1}(u) = Q_{\text{OBS}} \left[ \frac{u(1-p)}{1-pu} \right]. \quad (\text{A.1})$$

Hence, equation (A.1) reveals that the OMSG qf can be expressed in terms of the OBS qf. Using (2.2), we obtain

$$\Phi(v) = a_{u, \alpha, p} = \frac{[u(1-p)]^{\frac{1}{\alpha}}}{(1-u)^{\frac{1}{\alpha}} + [u(1-p)]^{\frac{1}{\alpha}}},$$

where  $v = \lambda^{-1} \rho(t/\beta)$  and  $\rho(u) = u^{1/2} - u^{-1/2}$ . After some algebra, we can write from (A.1)

$$t - \lambda(\beta t)^{\frac{1}{2}} Q_N(a_{u, \alpha, p}) - \beta = 0, \quad (\text{A.2})$$

where  $Q_N(\cdot)$  denotes the standard normal qf. Solving equation (A.2), the OMSG qf becomes

$$t = 0.5 \beta \lambda Q_N(a_{u, \alpha, p}) \left\{ \lambda Q_N(a_{u, \alpha, p}) + \sqrt{[\lambda Q_N(a_{u, \alpha, p})]^2 + 4} \right\} + \beta. \quad (\text{A.3})$$

We can write the standard normal qf in terms of the inverse error function, namely

$$Q_N(a_{u, \alpha, p}) = \sqrt{2} \operatorname{erf}^{-1}(2a_{u, \alpha, p} - 1).$$

An expansion of the inverse error function  $\operatorname{erf}^{-1}(x)$  can be obtained from Wolfram website (<http://mathworld.wolfram.com/InverseErf.html>) as

$$\operatorname{erf}^{-1}(x) = \sqrt{\pi} \left( \frac{1}{2}x + \frac{1}{24}\pi x^3 + \frac{7}{960}\pi^2 x^5 + \frac{127}{80640}\pi^3 x^7 + \dots \right).$$

Quantiles of interest can be determined from (A.1) or (A.3) by substituting appropriate values for  $u$ . In particular, the median of  $T$  follows when  $u = 1/2$ . By using (A.3), we have

$$t = \frac{\beta\lambda}{2} Q_N(a_{0.5,\alpha,p}) \left\{ \lambda Q_N(a_{0.5,\alpha,p}) + \sqrt{[\lambda Q_N(a_{0.5,\alpha,p})]^2 + 4} \right\} + \beta,$$

where

$$Q_N(a_{0.5,\alpha,p}) = \frac{(1-p)^{\frac{1}{\alpha}}}{1+(1-p)^{\frac{1}{\alpha}}}.$$

We can use (A.3) for simulating OBSG random variables by setting  $u$  as a uniform random variable in the unit interval  $(0, 1)$ .

**Mixture representation**

For any real  $\alpha > 0$ , we can expand  $\Phi(v)^\alpha$  as

$$\Phi(v)^\alpha = \sum_{k=0}^{\infty} v_k \Phi(v)^k, \tag{A.4}$$

where

$$v_k = v_k(\alpha) = \sum_{j=k}^{\infty} (-1)^{k+j} \binom{\alpha}{j} \binom{j}{k}.$$

Further, the generalized binomial expansion holds

$$[1 - \Phi(v)]^\alpha = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} \Phi(v)^k. \tag{A.5}$$

Inserting (A.4) and (A.5) in equation (2.2), we obtain

$$F(t) = \frac{\sum_{k=0}^{\infty} v_k \Phi(v)^k}{\sum_{k=0}^{\infty} w_k \Phi(v)^k},$$

where  $w_k = w_k(\alpha) = v_k + (-1)^k (1-p) \binom{\alpha}{k}$ .

The ratio of the two power series can be expressed as

$$F(t) = \sum_{k=0}^{\infty} d_k \Phi(v)^k, \tag{A.6}$$

where  $d_0 = v_0/w_0$  and the coefficients  $d_k$ 's (for  $k \geq 1$ ) are determined from the recurrence equation

$$d_k = w_0^{-1} \left( v_k - w_0^{-1} \sum_{r=1}^k w_r v_{k-r} \right).$$

By differentiating (A.6), we obtain

$$f(t) = \sum_{k=0}^{\infty} d_{k+1} h_{k+1}(t), \quad (\text{A.7})$$

where  $h_{k+1}(t) = (k+1)\Phi(\nu)^k g_{\lambda,\beta}(t)$  is the exponentiated Birnbaum-Saunders (exp-BS) model with power parameter  $k+1$ .

Cordeiro and Lemonte (2011) recently proposed the  $\beta$ BS distribution which has the exp-BS distribution as a special case when  $b=1$ . Therefore, some properties of the OMSG distribution can be obtained from the exp-BS distribution.

The formulae derived from equation (A.7) can be easily handled in most symbolic computation software platforms such as Maple, Mathematica, and Matlab, which have the ability to deal with analytic expressions of formidable size and complexity. In addition, we can change infinity by twenty or thirty for the majority of these formulae

### Moments

Some of the most important features and characteristics of a distribution can be studied through moments. Cordeiro and Lemonte (2011) determined the probability weighted moments (PWMs) of the BS distribution since they are required for the ordinary moments of the  $\beta$ BS distribution. Here, we also use the PWMs of the BS distribution to demonstrate the moments of the OMSG distribution. The PWMs (for  $p \geq 1, r \geq 0$ ) of the BS distribution are defined by  $\tau_{p,r} = E[T^p G_{\lambda,\beta}(T)^r]$ . We obtain

$$\tau_{p,r} = \frac{\beta^p}{2^r} \sum_{j=0}^r \binom{r}{j} \sum_{k_1, \dots, k_j=0}^{\infty} A(k_1, \dots, k_j) \sum_{m=0}^{2s_j+j} (-1)^m \binom{2s_j+j}{m} \beta^{-\frac{2s_j+j-2m}{2}} I\left(p + \frac{2s_j+j-2m}{2}, \lambda\right), \quad (\text{A.8})$$

where

$$I(p, \lambda) = \frac{K_{p+\frac{1}{2}}(\lambda^{-2}) + K_{p-\frac{1}{2}}(\lambda^{-2})}{2K_{\frac{1}{2}}(\lambda^{-2})}$$

and  $K_\nu(z)$  denotes the modified Bessel function of the third kind with  $\nu$  representing its order and  $z$  the argument. Its integral representation is  $K_\nu(z) = 0.5 \int_{-\infty}^{\infty} \exp\{-z \cosh(t) - \nu t\} dt$  (Watson, 1995).

From now on, let  $Y_{m+1}$  have the exp-BS distribution with power parameter  $m+1$ . A first formula for the  $s^{\text{th}}$  moment of  $T$  follows from (A.7) as

$$E(T^s) = \sum_{m=0}^{\infty} d_{m+1} E(Y_{m+1}^s). \quad (\text{A.9})$$

Based on the results by Cordeiro and Lemonte (2011), the moments of  $T$  are given by

$$E(T^s) = \sum_{m=0}^{\infty} (m+1) d_{m+1} \tau_{s,m},$$

where  $\tau_{s,m}$  comes from (A.8).

A second formula for  $E(T^s)$  follows from (A.7) in terms of the OBS qf as

$$E(T^s) = \sum_{k=0}^{\infty} (k+1) d_{k+1} \kappa(s, k),$$

where

$$\kappa(s, k) = \int_0^{\infty} t^s G_{\lambda, \beta}^k(t) g_{\lambda, \beta}(t) dx = \int_0^1 Q_{OBS}(u)^s u^k du.$$

### Generating function

The moment generating function of  $T$  can be determined from (A.7) using a result by Cordeiro and Lemonte (2011) that can be expressed as

$$M(s) = \sum_{m,k=0}^{\infty} \frac{(m+1) d_{m+1} \tau_{k,m}}{k!} s^k,$$

where  $\tau_{k,m}$  is given by (A.8).

### Acknowledgments

We are very grateful to a referee and an associate editor for helpful comments that improved the paper. We gratefully acknowledge financial support from CAPES and CNPq.

### References

- Alexander C, Cordeiro GM, Ortega EMM, and Sarabia JM (2012). Generalized beta-generated distributions, *Computational Statistics and Data Analysis*, **56**, 1880–1897.
- Alizadeh M, Emadi M, Doostparast M, Cordeiro GM, Ortega EMM, and Pescim RR (2015). A new family of distributions: the Kumaraswamy odd log-logistic, properties and applications, *Hacetatepe Journal of Mathematics and Statistics*, **44**, 1491–1512.
- Berkson J and Gage RP (1952). Survival curve for cancer patients following treatment, *Journal of the American Statistical Association*, **47**, 505–515.
- Boag JW (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy, *Journal of the Royal Statistical Society Series B (Methodological)*, **11**, 15–53.
- Brooks SP (2002). Discussion on the paper by Spiegelhalter, Best, Carlin, and van der Linde, *Journal of the Royal Statistical Society Series B (Methodological)*, **64**, 616–618.
- Cancho VG, Ortega EMM, Barriga GDC, and Hashimoto EM (2011). The Conway-Maxwell-Poisson-generalized gamma regression model with long-term survivors, *Journal of Statistical Computation and Simulation*, **81**, 1461–1481.
- Carlin BP and Louis TA (2001). *Bayes and Empirical Bayes Methods for Data Analysis* (2nd ed), Chapman and Hall, Boca Raton.
- Castillo NO, Gómez HW, and Bolfarine H (2011). Epsilon Birnbaum-Saunders distribution family: properties and inference, *Statistical Papers*, **52**, 871–883.
- Chib S and Greenberg E (1995). Understanding the metropolis-Hastings algorithm, *The American Statistician*, **49**, 327–335.
- Cooner F, Banerjee S, Carlin BP, and Sinha D (2007). Flexible cure rate modeling under latent activation schemes, *Journal of the American Statistical Association*, **102**, 560–572.
- Cordeiro GM, Cintra RJ, Rego LC, and Ortega EMM (2012). The McDonald normal distribution, *Pakistan Journal of Statistics and Operation Research*, **8**, 301–329.
- Cordeiro GM and Lemonte AJ (2011). The  $\beta$ -Birnbaum-Saunders distribution: an improved distribution for fatigue life modeling, *Computational Statistics and Data Analysis*, **55**, 1445–1461.

- Cordeiro GM and Lemonte AJ (2014). The exponentiated generalized Birnbaum-Saunders distribution, *Applied Mathematics and Computation*, **247**, 762–779.
- Cordeiro GM, Lemonte AJ, and Ortega EMM (2013). An extended fatigue life distribution, *Statistics*, **47**, 626–653.
- Cordeiro GM, Ortega EMM, and Nadarajah S (2010). The Kumaraswamy Weibull distribution with application to failure data, *Journal of the Franklin Institute*, **347**, 1399–1429.
- Cowles MK and Carlin BP (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review, *Journal of the American Statistical Association*, **91**, 883–904.
- Fachini JB, Ortega EMM, and Cordeiro GM (2014). A bivariate regression model with cure fraction, *Journal of Statistical Computation and Simulation*, **84**, 1580–1595.
- Farewell VT (1982). The use of mixture models for the analysis of survival data with long-term survivors, *Biometrics*, 1041–1046.
- Gelfand AE, Dey DK, and Chang H (1992). Model determination using predictive distributions with implementation via sampling based methods (with discussion). In JM Bernardo, JO Berger, AP Dawid, and AFM Smith (Eds), *Bayesian Statistics 4* (pp. 7–167), Oxford University Press, New York.
- Hashimoto EM, Ortega EMM, Cordeiro GM, and Cancho VG (2014). The Poisson Birnbaum-Saunders model with long-term survivors, *Statistics*, **48**, 1394–1413.
- Hashimoto EM, Ortega EMM, Cordeiro GM, and Cancho VG (2015). A new long-term survival model with interval-censored data, *Sankhya B*, **77**, 207–239.
- Ibrahim JG, Chen MH, and Sinha D (2001). *Bayesian Survival Analysis*, Springer-Verlag, New York.
- Kundu D, Kannan N, and Balakrishnan N (2008). On the hazard function of Birnbaum-Saunders distribution and associated inference, *Computational Statistics and Data Analysis*, **52**, 2692–2702.
- Lanjoni BR, Ortega EMM, and Cordeiro GM (2016). Extended Burr XII regression models: theory and applications, *Journal of Agricultural Biological and Environmental Statistics*, **21**, 203–224.
- Lemonte AJ (2013). A new extension of the Birnbaum-Saunders distribution, *Brazilian Journal of Probability and Statistics*, **27**, 133–149.
- Longini IM and Halloran ME (1996). A frailty mixture model for estimating vaccine efficacy, *Journal of the Royal Statistical Society Series C Applied Statistics*, **45**, 165–173.
- Maller RA and Zhou X (1996). *Survival Analysis with Long-Term Survivors*, Wiley, New York.
- Ortega EMM, Cancho VG, and Paula GA (2009). Generalized log-gamma regression models with cure fraction, *Lifetime Data Analysis*, **15**, 79–106.
- Ortega EMM, Cordeiro GM, Campelo AK, Kattan MW, and Cancho VG (2015). A power series beta Weibull regression model for predicting breast carcinoma, *Statistics in Medicine*, **34**, 1366–1388.
- Ortega EMM, Cordeiro GM, Hashimoto EM, and Suzuki AK (2017). Regression models generated by gamma random variables with long-term survivors, *Communications for Statistical Applications and Methods*, **24**, 43–65.
- Ortega EMM, Cordeiro GM, and Lemonte AJ (2012). A log-linear regression model for the  $\beta$ -Birnbaum-Saunders distribution with censored data, *Computational Statistics Data Analysis*, **56**, 698–718.
- Peng Y and Dear KB (2000). A nonparametric mixture model for cure rate estimation, *Biometrics*, **56**, 237–243.
- Price DL and Manatunga AK (2001). Modelling survival data with a cured fraction using frailty models, *Statistics in Medicine*, **20**, 1515–1527.
- Saavedra del Aguila J, Heiffig-del Aguila LS, Sasaki FF, Tsumanuma GM, das Graças Ongarelli M,

- Spoto MHF, Jacomino AP, Ortega EMM, and Kluge RA (2010). Postharvest modifications of mechanically injured bananas, *Revista Iberoamericana de Tecnologia Postcosecha*, **10**, 73–85.
- Spiegelhalter DJ, Best NG, Carlin BP, and van der Linde A (2002). Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, **64**, 583–639.
- Spiegelhalter DJ, Thomas A, Best N, and Lunn D (2007). OpenBUGS user manual version 3.0.2, Retrieved July 2, 2017, from <http://mathstat.helsinki.fi/openbugs>
- Suzuki AK, Cancho VG, and Louzada F (2016). The Poisson-Inverse-Gaussian regression model with cure rate: a Bayesian approach and its case influence diagnostics, *Statistical Papers*, **57**, 133–159.
- Sy JP and Taylor JM (2000). Estimation in a Cox proportional hazards cure model, *Biometrics*, **56**, 227–236.
- Tsodikov AD, Ibrahim JG, and Yakovlev AY (2003). Estimating cure rates from survival data: an alternative to two-component mixture models, *Journal of the American Statistical Association*, **98**, 1063–1078.
- Watson GN (1995). *A Treatise on the Theory of Bessel Functions* (2nd ed), Cambridge University Press, Cambridge.
- Yu B and Peng Y (2008). Mixture cure models for multivariate survival data, *Computational Statistics and Data Analysis*, **52**, 1524–1532.

Received May 10, 2017; Revised July 6, 2017; Accepted July 7, 2017