Journal of Biostatistics and Epidemiology

J Biostat Epidemiol. 2023;9(1):46-58

Original Article

The Geometric Generalized Birnbaum–Saunders model with long-Term Survivors

Ahmad Reza Baghestani[1], Farid Zayeri[1], Mojtaba Meshkat[2*]

[1]Department of Biostatistics, School of Paramedicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.
[2]Department of Community Medicine, Faculty of Medicine, Mashhad Medical Sciences, Islamic Azad University, Mashhad, Iran.

ARTICLE INFO

Received 28.10.2022
Revised 23.12.2022
Accepted 03.02.2023
Published 15.03.2023

Key words:
Cure fraction models;
Generalized birnbaum-saunders distribution;
Geometric distribution;
Lifetime data;
Fatigue Life Distribution.

ABSTRACT

Introduction: A cure rate survival model was developed based on the assumption that the number of competing reasons for the event of interest has the Geometric distribution and the time allocated to the event of interest follows the Generalized Birnbaum-Saunders distribution.
Methods: The Geometric Generalized Birnbaum–Saunders distribution distribution was defined and two useful representations were represented for its density function which contributes to the creation of some mathematical properties. Furthermore, the parameters of the model with cure rate were estimated by using the maximum likelihood method.
Results: Several simulations were performed and a real data set was analyzed from the medical area for different sample sizes and censoring percentages.In the melanoma data set and regarding the AIC and SBC selection criteria, the Geometric Generalized Birnbaum–Saunders distribution model was preferred and was selected as the appropriate model in the present study.
Conclusion: Geometric Generalized Birnbaum–Saunders distribution is a highly flexible lifetime model which allows for different degrees of Kurtosis and asymmetry.by considering the advantages of the Geometric Generalized Birnbaum–Saunders distribution model, the model can be implemented as an appropriate alternative to explain or predict the survival time for long-term individuals.

Introduction

The model derived by Birnbaum and Saunders[1] is regarded as an important lifetime model beginning from a material fatigue problem. The Birnbaum-Saunders (BS) distribution is approximately related to the normal distribution, which is applied in different fields of study.[2] This kind of distribution which is positively skewed, and unimodal with non-negative support includes two parameters.[1, 3] In addition, Generalized Birnbaum–Saunders (GB-S) distribution is a highly flexible lifetime model which allows for different degrees of

*.Corresponding Author: m.meshkat@sbmu.ac.ir

Kurtosis and asymmetry while possessing unimodal and bimodal distribution proposed by Díaz-García, Leiva and Owen.[4-8] In addition, GB-S distributions are able to produce models whose parameter estimates are often non-sensitive to outliers and robust to atypical data.[8-11] The GB-S distribution is emphasized due to its physical and theoretical arguments, its relationship with the normal distribution and other attractive properties.

In the case of t>0, the survival function, cumulative distribution function and probability density function of the GB-S distribution are given by:[6, 7]

$$S_{GB-S} = \Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda} - \left(\frac{\beta}{t}\right)^{\lambda}\right]\right), \qquad (1)$$

$$F_{GB-S} = 1 - S_{GB-S} = \Phi\left(\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda} - \left(\frac{\beta}{t}\right)^{\lambda}\right]\right), \qquad (2)$$

$$f_{GB-S} = \frac{\lambda}{\alpha t\sqrt{2\pi}}\left(\left(\frac{t}{\beta}\right)^{\lambda} - \left(\frac{\beta}{t}\right)^{\lambda}\right) e^{\left\{-\frac{1}{2\alpha^2}\left[\left(\frac{t}{\beta}\right)^{2\lambda} + \left(\frac{\beta}{t}\right)^{2\lambda} - 2\right]\right\}}, \qquad (3)$$

Where $\phi(.)$ represents the cumulative distribution function for standard normal distribution, and $\alpha > 0$, $\beta > 0$ and $\lambda > 0$ indicate shape, scale and location parameters, respectively. The mean and variance of the GB-S distribution are expressed as follows:

$$E(T) = \beta\sum_{k=0}^{\infty}\sum_{s=0}^{\infty}\binom{\frac{1}{\lambda}}{k}\binom{\frac{1}{2\lambda}-\frac{k}{2}}{s}(I_1 + I_2),$$

$$Var(T) = \beta^2\sum_{k=0}^{\infty}\sum_{s=0}^{\infty}\binom{\frac{2}{\lambda}}{k}\binom{\frac{1}{\lambda}-\frac{k}{2}}{s}(I_1 + I_2) - \left[E(T)\right]^2$$

where

$$I_1 = 2^{\frac{-2s-k-1}{2}}\left[\Gamma\left(\frac{2s+k+1}{2}\right) - \Gamma\left(\frac{2s+k+1}{2}, \frac{2}{\alpha^2}\right)\right]\left[(-1)^{2s+k}+1\right],$$

$$I_2 = \frac{\alpha^{\left(\frac{2}{\lambda}-2s\right)}}{2^{\frac{\frac{2}{\lambda}-2s+1}{2}}}\left[(-1)^{\frac{2}{\lambda}-2s}+1\right]\Gamma\left(\frac{\frac{2}{\lambda}-2s+1}{2}, \frac{2}{\alpha^2}\right).$$

The models used for survival data or long-term survival models play an important role in survival analysis. Cure rate models include the situations where there are sampling units which are not sensitive to the occurrence of the event of interest. The proportion of such units is called "cured fractions." These models are very popular due to the considerable development in treatment therapies resulting in increasing cure rates. In medical studies, the event of interest may be related to the death of a patient due to different competing risks or a tumor recurrence because of metastasis-component tumor cells left active after an initial treatment.

The mixture cure model (MCM) is regarded as the most popular type of cure rate model introduced by Boag (1949) in order to examine those cases in which a proportion of cured patients was available among the treatment receivers for mouth cancer.[12] In addition, Berkson and Gage[13] developed the model three years later and accordingly it was extensively investigated by some other authors.[13-16] Based on MCM distribution, a specific number of the patients (*p0*) are cured in such a way that they cannot present the event of interest

during a long period of time while they can be observed to be immune to the cause of death under study or cured. No literature, to the best of our knowledge, is available for observing a mixture GB-S (MGB-S) distribution with survial function as follows:

$$S_{MGB-S}(t) = p_0 + (1 - p_0) S_{GB-S} \quad t > 0, \tag{4}$$

where $S_{GB-S}(t)$ is defined by (1) and p_0 represents the cured fraction.

The Non-Mixture Cure Model (NMCM) is considered as another kind of cure rate model, which was first proposed by Yakovlev et al.[17] and was further discussed by Chen et al.,[18] Yin et al.,[19] Cooner et al.,[20] Rodriguesa et al.,[21] Castro et al.[22] and Borges et al.[23] NMCM model was motivated by the biological mechanism and developed based on the assumption that the number of cancer cells remaining active after cancer treatment is consistent with the Poisson distribution.[17, 18] There is a vast amount of literature on distributions which accommodate other different latent competing causes.

Cancho et al. introduced the Geometric Birnbaum–Saunders regression model with cure rate by utilizing a cutaneous melanoma data, which could provide better fitting, compared to the MB-S distribution.[24] Hashimoto et al. demonstrated the Poisson Birnbaum–Saunders model with long-term survivors in breast cancer data[25] and Cordeiro et al. presented Negative binomial Birnbaum–Saunders model with long-term survivors in melanoma data.[26] Recently, Meshkat et al. introduced the Poisson generalized Birnbaum–Saunders regression model in melanoma data, which could provide better fitting, compared to B-S distributions[27] and Taketomi et al. reviewed some of this models.[28]

Considering all the above-mentioned studies, the present study aimed to propose the geometric generalized Birnbaum–Saunders cure rate model (GGB-S), as a new distribution family, which considered a latent competing causes scenario with cure fraction. In Cancer study there are many risk factors that the exact cause of the individual death or tumor recurrence, no information is available. Among all of the risks, the minimum lifetime value is observed and a fraction of the population is not susceptible to the event of interest.

This paper is organized as follows: in the second section, the GGB-S model is formulated by defining the density, cumulative distribution and hazard rate functions of the Geometric Generalized Birnbaum–Saunders (GGB-S) distribution and some of its properties are also studied. In the third section, the maximum-likelihood estimation procedure and parameter inference are elaborated. The fourth section evaluates the performance of the parameter estimates using Monte Carlo simulation. The fifth section gives an application to a real data set on melanoma cancer. Finally, some conclusions are drawn.

Methods

Model formulation

Let M define the unobservable number of causes (risk factors) of the event under study for an individual in the population. We assume that M follows a Geometric distribution with parameter θ and mass probability

$$P(M=m)=(1-\theta)\,\theta^m \qquad m=0,1,\ldots . \tag{5}$$

The time for the j^{th} causes to produce the

event of interest is defined by Z_j , $j=1,...,M$. We assume that conditional on M, the Z_{js} are regarded as i.i.d. random variables having the Generalized Birnbaum–Saunders cumulative function given by equation (2). In addition, we assume that Z_1, Z_2, … are independent of M. The observable time to event of interest is defined by the random variable T = min (Z_1, ..., Z_M), and $T=\infty$ if $M=0$ with $P(T=\infty \mid M=0)=1$. Under this setup and based on the equation (5), the survival function for the population is as follows:

$$S_{POP}(t)=P(T>t)=P(M=0)+P(Z_1>t,\ldots,Z_M>t|M\geq 1)$$
$$=(1-\theta)+\sum_{m=1}^{\infty}[S_{GB-S}]^m P(M=m)=$$
$$(1-\theta)+\sum_{m=1}^{\infty}\left\{\left[\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)\right]^m(1-\theta)\theta^m\right\}$$
$$=\frac{1-\theta}{1-\theta\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)}. \quad (6)$$

The cure function is defined by $p_0=S_{POP}(\infty)=1-\theta$ and the corresponding density function is given below:

$$f_{POP}(t)=\theta(1-\theta)f_{GB-S}(t)exp\left\{1-\theta\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)\right\}^{-2},$$

where f_{GB-S} (t) represents the probability density function of the distribution given in (3). Furthermore, the population hazard function is as follows:

$$h_{POP}(t)=\frac{\theta f_{GB-S}(t)}{1-\theta\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)}. \quad (8)$$

The function $h_{POP}(t)$ is multiplicative in θ and $f_{GB-S}(t)$, including the proportional hazard structure when the covariates are modeled through θ. The functions $f_{POP}(t)$ in (7) and $h_{POP}(t)$ in (8) are regarded as improper functions since $S_{POP}(t)$ is not a proper survival function.

The distribution in (6) is written as a mixture distribution:[13]

$$S_{POP}(t)=(1-)+\theta\left\{\frac{(1-\theta)\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)}{1-\theta\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)}\right\} \quad (9)$$

Then, based on the Equation (4) and (9), the survival function for the non-cured population from the Geometric Generalized Birnbaum Saunders (GGB-S) survival function is given below:

$$S_{GGB-S}(t)=P(T>t|M\geq 1)=\frac{(1-\theta)\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)}{1-\theta\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)}. \quad (10)$$

It is worth noting that $S_{GGB-S}(0)=1$ and $S_{GGB-S}(\infty)=0$, in order to be considered as a proper survival function. The density function of the GGB-S distribution is given by:

$$f_{GGB-S}(t)=\frac{(1-\theta)\lambda t^{-(\lambda+1)}\left(t^{2\lambda}+\beta^{2\lambda}\right)exp\left(-\frac{1}{2\alpha^2}\left[\left(\frac{t}{\beta}\right)^{2\lambda}+\left(\frac{\beta}{t}\right)^{2\lambda}-2\right]\right)}{\sqrt{2\pi}\alpha\beta^{\lambda}\left\{1-\theta\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)\right\}^2} \quad (11)$$

Based on Equation (11), the parameter β controls the scale of the density distribution while the parameters α, λ and θ are responsible for controlling its shape. As θ approaches

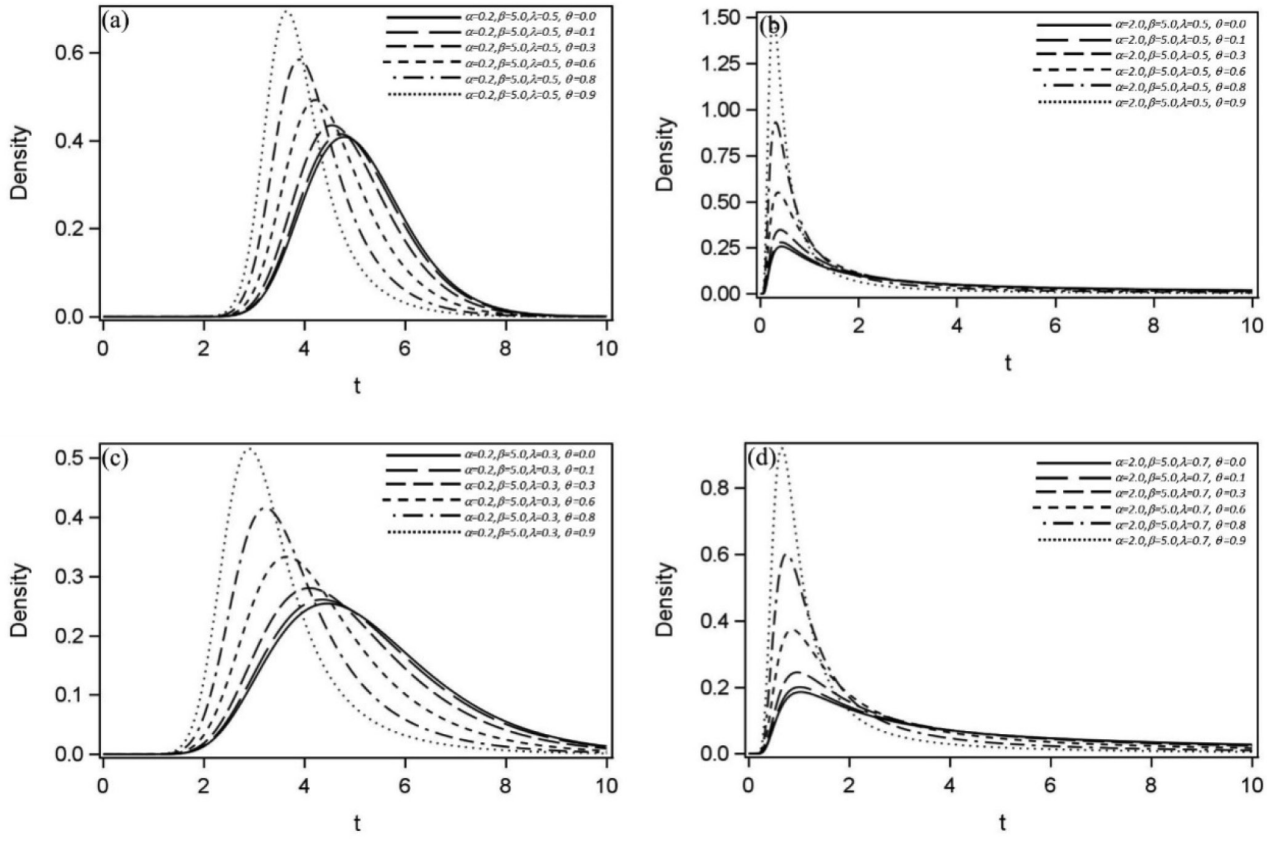

Figure 1. Probability density function of the GGB-S distribution. The parameters were fixed θ=0.0 0.1 0.3 0.6 0.8 0.9 and α=0.2, β=5.0, λ=0.5 (a), α=2.0, β=5.0, λ= 0.5 (b), α= 0.2, β=5.0, λ=3.0 (c), α=2.0, β=5.0, λ= 0.7 (d).

zero, the $G_{GB\text{-}S}$ distribution converges to a $G_{B\text{-}S}$ distribution. Figure 1 illustrates the $G_{GB\text{-}S}$ density function plotted for the selected values of the parameters. The $G_{GB\text{-}S}$ distribution is used to model data in survival analysis.

Based on the equations (10) and (11), it is easy to verify that the hazard rate function of the non-cured population is reduced to:

$$h_{GGB\text{-}S}(t)=\frac{\lambda t^{-(\lambda+1)}\left(t^{2\lambda}+\beta^{2\lambda}\right)exp\left(-\frac{1}{2\alpha^2}\left[\left(\frac{t}{\beta}\right)^{2\lambda}+\left(\frac{\beta}{t}\right)^{2\lambda}-2\right]\right)}{\sqrt{2\pi}\alpha\beta^{\lambda}\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)\left\{1-\theta\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)\right\}} \quad (11)$$

Numerical examples allow us to determine that the hazard rate function (12) is either increasing or unimodal (Figure 2).

The r_{st} moment[29] of the GGB-S distribution is as follows:

$$\mu_r^{'}=E\left(T^r\right)=\int_0^{\infty}rt^{r-1}\frac{(1-\theta)\Phi\left(v_t\left(\alpha,\beta,\lambda\right)\right)}{1-\theta\Phi\left(v_t\left(\alpha,\beta,\lambda\right)\right)}dt, \quad (13)$$

Where,

$$v_t\left(\alpha,\beta,\lambda\right)=-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right],$$

and $\Phi(.)$ represents the standard normal cumulative distribution function.

In equation (13) for θ=0, the moments are corresponding to the respective moments

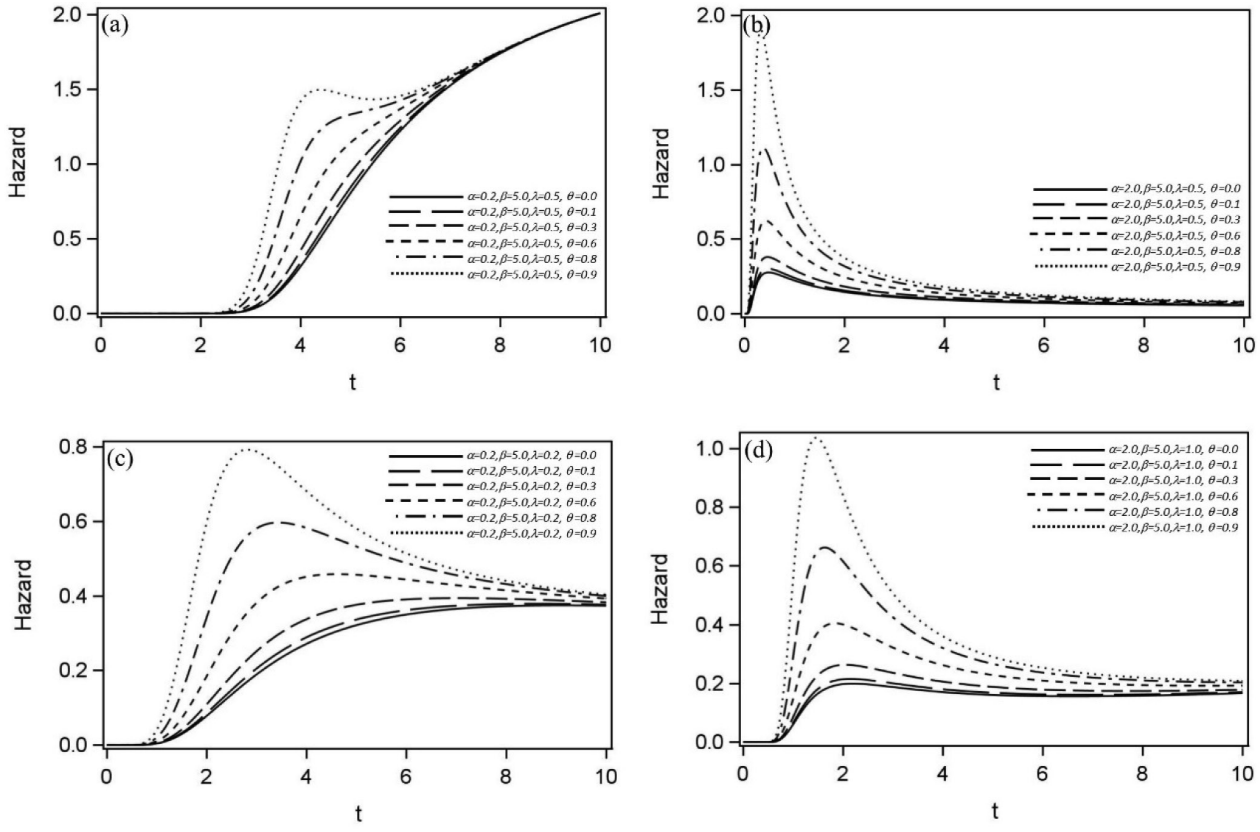

Figure 2. Hazard function of the GGB-S distribution .The parameters were fixed θ=0.0 0.1 0.3 0.6 0.8 0.9 and α=0.2, β=5.0, λ=0.5 (a), α=2.0, β=5.0, λ= 0.5 (b), α=0.2, β=5.0, λ=0.2 (c), α=2, β=5.0, λ=1.0 (d).

of the GB-S distribution. The skewness and kurtosis of the distribution is calculated from the ordinary moments given in equation (13) using well-known relationships where μ'_r are numerically obtainable from some common statistical software such as SAS (Guad function). Figure 3 illustrates the graphical representations of skewness and kurtosis indicating that both are regarded as increasing functions of α, θ and λ.

The lifetimes of all the sampling units may not be observable, due to censoring which are present in the data. Now, let us consider the case that the minimum lifetime of all the competing causes is not completely observed and is subject to right censoring. Let C_i represents the censoring time. Then, we observe δ_i =I(Ti $\leq C_i$) and $t_i = \min\{T_i, C_i\}$, where $\delta_i = 1$ if Ti is the observed time to the event defined before and $\delta_i = 0$ if it is correctly censored for $i = 1, \ldots, n$. Considering $\gamma = (\alpha, \beta, \lambda)$T as the parameter vector of the distribution function of the time-to-event, the observed likelihood function under non-informative right censoring based on n pairs of times and censoring indicators (t_1, δ_1), . . ., (t_n, δ_n) is as follows:

$$L(\theta,\tilde{\boldsymbol{a}}) \propto \prod_{i=1}^{n}\left[f_{POP}(t_i;\theta,\tilde{\boldsymbol{a}})\right]^{\delta_i}\left[S_{POP}(t_i;\theta,\tilde{\boldsymbol{a}})\right]^{1-\delta_i} \quad (14)$$

Where, $S_{POP}(t_i)$ and $f_{POP}(t_i)$ were defined in equations (6) and (7), respectively.[30] In most of the medical studies, the lifetimes are usually

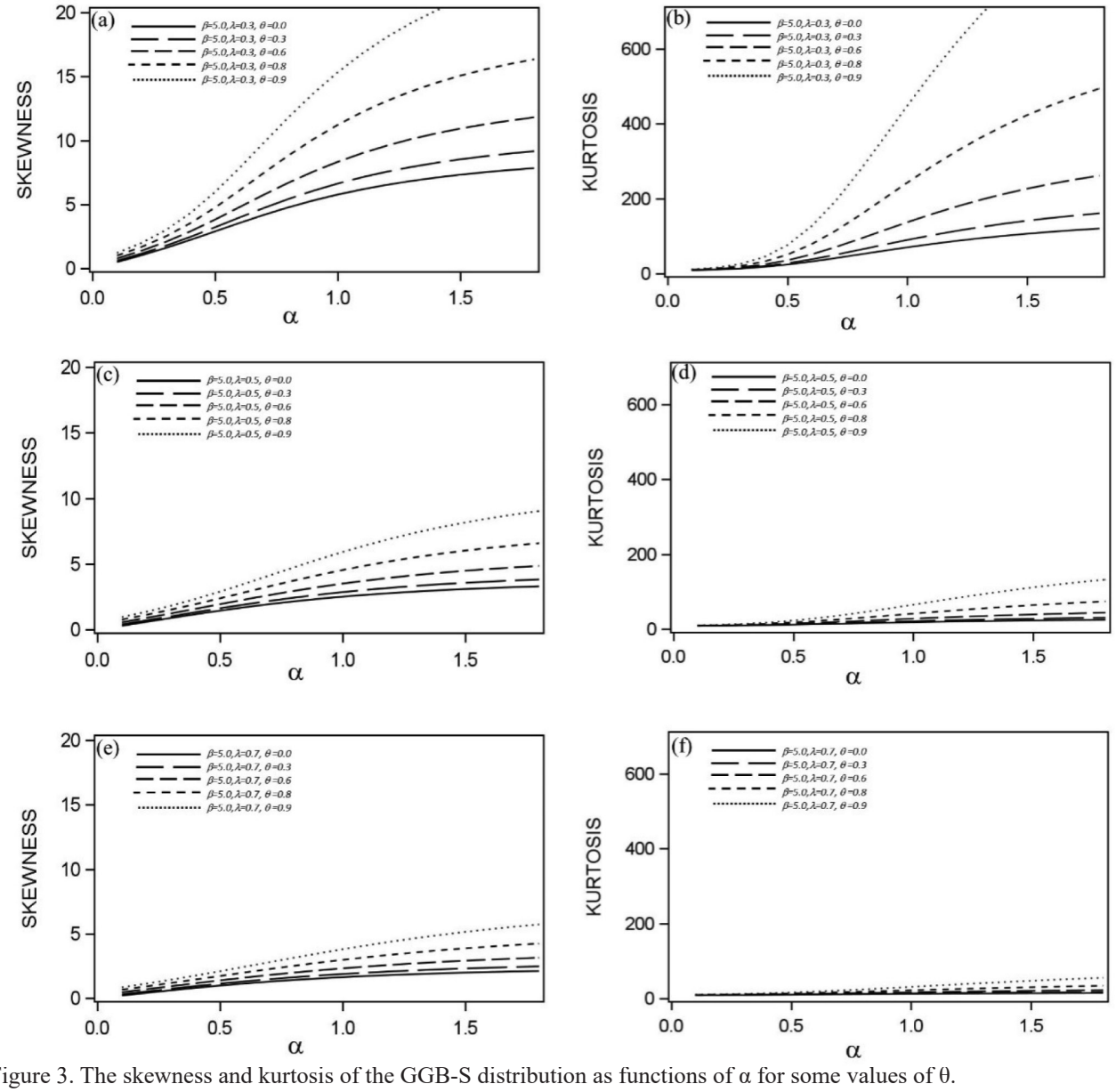

Figure 3. The skewness and kurtosis of the GGB-S distribution as functions of α for some values of θ.

influenced by explanatory variables such as the sex, age, blood pressure and the like. The parameter θ in equation (6) is now linked to a vector X_i of explanatory variables by assuming a logistic link, $\log(p_{0i}/(1-p_{0i}) = X_i^T \eta \; i = 1,\ldots n$ as follows:

$$p_{0i} = 1 - \theta = \frac{\exp\left(X_i^T \eta\right)}{1 + \exp\left(X_i^T \eta\right)}, \qquad (15)$$

where $(\eta_1, \ldots \eta_p)$ denotes the vector of regression coefficients. However, we consider $\omega = (\gamma, \eta)$ as the vector of model parameters. By substituting equations (6) and (7) into equation (14), and the use of equation (15), the

log-likelihood function is given below:

$$l(\omega)=\sum_{i=1}^{n}\delta_i\left\{\begin{array}{l}log\left[\left(1-\frac{\exp(X_i^T\eta)}{1+\exp(X_i^T\eta)}\right)\left(\frac{\exp(X_i^T\eta)}{1+\exp(X_i^T\eta)}\right)f_{GB-S}(t)\right]+\\ \left\{1-\left(1-\frac{exp(X_i^T\eta)}{1+exp(X_i^T\eta)}\right)\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)\right\}^{-2}\end{array}\right\}$$
$$+\sum_{i=1}^{n}(1-\delta_i)\left\{\left(\frac{exp(X_i^T\eta)}{1+exp(X_i^T\eta)}\right)\left\{1-\left(\frac{exp(X_i^T\eta)}{1+exp(X_i^T\eta)}\right)\Phi\left(-\frac{1}{\alpha}\left[\left(\frac{t}{\beta}\right)^{\lambda}-\left(\frac{\beta}{t}\right)^{\lambda}\right]\right)\right\}^{-1}\right\}. \quad (16)$$

The obtained equation cannot be analytically solved. Thus, statistical software should be used to solve them numerically. The model parameters are derived by the numerical maximization of the log-likelihood function obtained from equation (16) by using the SAS university software. The computational program is available from the authors upon request. We can use iterative techniques such as a quasi-Newton method to calculate the estimate $\hat{\omega}$.[31]

Under suitable regularity conditions, the inference procedures for $\omega = (\gamma^T, \eta^T)^T$ are based on the multivariate normal approximation[32]

$$\left(\hat{\gamma}^T,\hat{\eta}^T\right)^T \sim MVN_{p+3}\left(\left(\gamma^T,^T\right)^T,\Sigma(\omega)\right),$$

Where, covariance matrix $\sum(\hat{\omega})$ is estimated at $\omega=\hat{\omega}$. In other words, the information matrix is as follows:

$$\Sigma(\omega)=\left[-\frac{\vartheta^2 l(\omega)}{\vartheta\omega\,\vartheta\omega^T}\right]^{-1}. \quad (19)$$

Different models may be compared by penalizing over-fitting using the Akaike information criterion (*AIC*) and the Schwartz Bayesian criterion (*SBC*). Among all of the competing models, the model with the smallest value for any of these criteria is taken as the preferred model for describing the given data set.

Results

Simulation study

In the present study, the model was proposed based on the GGB-S distribution for the event times (T) by considering the parameters $\alpha = 2$, $\beta=10$ and $\lambda= (0.3, 0.5)$. Single binary covariate x was considered with the values drawn from a Bernoulli distribution with a parameter of 0.5. For each individual *i,* i=1,…,n, the number of causes (N_i) of the event of interest for the i^{th} individual was generated from the Geometric distribution with parameter

$$(1-\theta_i)=p_{0i}=\exp(\eta_0+\eta_1 x_i)/$$
$$(1+\exp(\eta_0+\eta_1 x_i)).$$

$\eta_0=0$ and $\eta_1 = -1$ were considered so that the cured fraction for the two levels of x_1 involves

$p_0^{(0)}=0.62$ and $p_0^{(1)}=0.38$, respectively.

The censoring times were selected from the uniform distribution on the interval (0, τ), where τ was adjusted in order to control the proportion of censored observations. The proportion of censored observations equaled to on average 50%.

The sample sizes of 50, 100, 500 and 1000 were selected in this study. For each setup, we calculated the mean of maximum likelihood estimates of the cured fraction ($p_0^{(0)}$ *and* $p_0^{(1)}$, standard deviations (SD) and square root of mean square errors (SRMSE) of the maximum

Table 1. Means for maximum likelihood estimates, standard deviation (SD) and square root of mean square error (RMSE) of cured fractions for the simulated data of GGB-S cure rate model.

		$p_0^{(0)}$			$p_0^{(1)}$		
		MEAN	SD	SRMSE	MEAN	SD	SRMSE
η= 0.3	50	0.643	0.095	0.098	0.402	0.091	0.094
	100	0.639	0.065	0.068	0.402	0.063	0.067
	500	0.620	0.029	0.029	0.401	0.029	0.036
	1000	0.620	0.021	0.021	0.400	0.020	0.028
η= 0.5	50	0.615	0.096	0.096	0.359	0.091	0.094
	100	0.617	0.069	0.069	0.362	0.062	0.064
	500	0.620	0.032	0.032	0.370	0.029	0.031
	1000	0.620	0.022	0.022	0.374	0.021	0.022

likelihood estimates for 1000 conducted simulations. Table 1 indicates the results for the simulated data of GGB-S cure models. Based on the results, an increase in the sample size resulted in approaching the average of maximum likelihood estimates to the true values of cured fraction and decreasing the SDs and SRMSEs.

Application

An example employing the modeling approach was presented in Section 2. The data set includes 205 patients observed after the operation for removing malignant melanoma during a follow-up period of 15 years. These data are available in the "timereg" package in R (Scheike, 2009). The observed time (T) ranges from 10 to 5565 days (during 0.0274-15.25 years, with Mean±SD of 5.9±3.1 years) and refers to the time until the patient's death or the censoring time. The patients who died from other causes and those who were still alive at the end of the study were censored observations (72%). Ulceration status (x_1) (absent, n=115; present, n=90) and tumor thickness (x_2) (M=2.92±2.96 mm) were regarded as covariates.

Figure 4 illustrates the Kaplan-Meier estimate of the surviving function. The flat line above 0.6 indicates that the models ignoring the possibility of cure are not regarded as appropriate choices for analyzing the data set.

In the next step, the mixture GB-S was fitted with p_{0i}. Table 2 represents the maximum likelihood estimates of the coefficients as well as *AIC* and *SBC* selection criteria In addition, the maximum likelihood estimates of the models coefficients based on the results of Cancho,[24] and *AIC* and *SBC* selection criteria based on the method of Hashimoto,[25] Cordeiroa[26] and Meshkat[27] were calculated in this data set. Regarding both criteria, the GGB-S model was preferred and was selected as the appropriate model in the present study.

As shown in Figure 5, the QQ plot of the normalized randomized quantile residuals from Dunn and Smyth[33] and Rigby & Stasinopoulos[34] suggests that the GGB-S model has yielded a good fit.

In order to compare the nested models, used for comparing the GGB-S and GB-S models, the maximum values of log-likelihoods are computed to obtain the likelihood ratio (LR) statistics for testing H_0:$\lambda=0.5$ versus H_0:$\lambda\neq0.5$. Based on the results, the null hypothesis is rejected (p=0.021).

As it is evident from Table 2, the thickness and ulceration as covariates played a significant role in the reduction of the cured fraction. Comparing patients with and without ulceration, the odds ratio of the cured fraction was estimated as $e^{1.445}=4.24$. The estimate can be easily computed from Table 2, due to the parameterization of the cured fraction. Table 2 indicates the maximum likelihood estimates of the coefficients as well as AIC and SBC selection criteria. Based on both criteria, the GGB-S model was selected a as the best model.

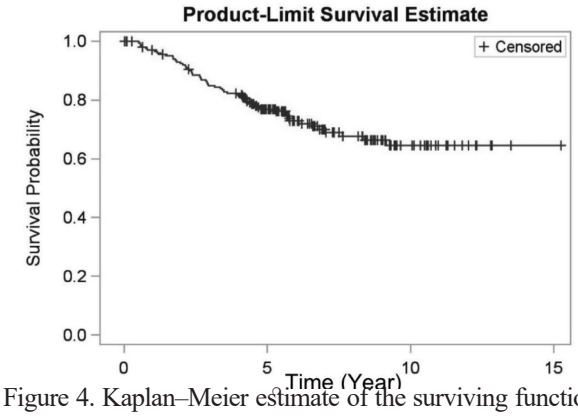

Figure 4. Kaplan–Meier estimate of the surviving function to the melanoma.

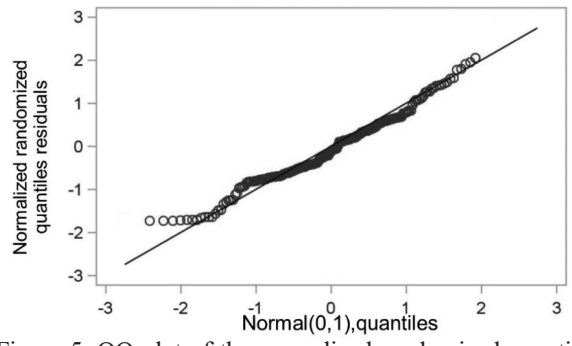

Figure 5. QQ plot of the normalized randomized quantile residuals with the identity line under the GGB-S model to melanoma.

Table 2. Maximum likelihood estimates for the parameters (SD) of the GGB-S model with the cure rate fitted to the melanoma data set.

Model	Estimate						Statistics	
	a	b	λ	$\eta_{intercept}$	$\eta_{thickness}$	$\eta_{ulceration}$	AIC	SBC
MB-S (Cancho G,2012)	1.312 (0.201)	6.442 (1.565)	-	2.421 (0.763)	-0.956 (0.556)	-1.331 (0.821)	430.9	447.5
GB-S (Cancho G,2012)	1.357 (0.560)	10.091 (3.437)	-	1.257 (0.586)	-0.168 (0.053)	-1.422 (0.353)	426.7	443.3
PB-S (Hashimoto EM,2014)	1.336 (0.162)	8.195 (1.999)	-	1.316 (0.313)	-0.110 (0.034)	-1.216 (0.301)	428.2	444.8
NBB-S* (Cordeiroa GM,2012)	1.352 (0.074)	10.064 (3.334)	-	1.261 (0.331)	-0.167 (0.067)	-1.415 (0.562)	428.6	448.6
PGB-S (Meshkat M,2018)	0.011 (0.071)	6.905 (3.119)	0.005 (0.033)	1.580 (0.444)	-0.112 (0.036)	-1.171 (0.306)	426.9	450.2
MGB-S	0.793 (0.113)	6.369 (1.812)	0.329 (0.178)	2.389 (1.055)	-0.917 (0.922)	-1.349 (1.021)	430.3	446.9
GGB-S	0.0010 (0.0002)	7.765 (2.408)	0.0005 (0.0003)	1.432 (0.419)	-0.179 (0.056)	-1.445 (0.338)	423.8	443.3

*The estimated value of the parameter in the negative binomial distribution is 1.034.

Discussion

We usually deal with the presence of cure fraction and covariates, especially in medical applications, in order to analyze the lifetime data. The Birnbaum-Saunders distribution has been extensively used for modeling the fatigue lifetimes in different fields such as medical sciences, biological studies, engineering and insurance. The generalized Birnbaum–Saunders (GB-S) distribution is regarded as a highly flexible lifetime model which allows for different values of kurtosis and asymmetry. In addition, it includes unimodality and bimodality introduced by Owen.[7]

In the present study, a GGB-S model was proposed for analyzing survival data with cure fraction, and was considered as a latent competing cause for the scenario with cure fraction. For this purpose, a four-parameter continuous model called "the GGB-S distribution" was introduced as the extension of the GB-S distribution. A mathematical relationship is available among the GB-S cure rate model the mixture cure rate model,[12, 13] and the GGB-S distribution. Hence, the proposed cure rate model involves the structure of the mixture cure rate model, and the GGB-S distribution represents the distribution related to the individuals at risk.

The quasi-Newton and Nelder-Mead methods (we can use option TECH in PROC NLP of SAS) were utilized to obtain the maximum likelihood estimates and accordingly asymptotic tests were performed for the parameters. The interpretation of the covariates is straightforward due to the parameterization of the cure fraction. By considering the advantages of the GGB-S model, the model can be implemented as an appropriate alternative to explain or predict the survival time for long-term individuals.

Conclusion

To compared to other models , Geometric Generalized Birnbaum–Saunders distribution is a highly flexible lifetime model but in cases where the sample size is not enough, it is better to use Bayesian models. Also, we suggest to use other different types of Birman Saunders distribution.

Acknowledgements

The authors thank the anonymous reviewers.

Funding

The author(s) disclosed not receipt financial support for the research, authorship, and/or publication of this article.

Conflict of interest

The Authors declare that there is no conflict of interest.

References

1. Birnbaum ZW, Saunders SC. A new family of life distributions. Journal of Applied Probability. 1969;6(2):319-27.

2. Pescim RR, Cordeiro GM, Nararajah S, Demétrio CG, Ortega EM. The Kummer beta Birnbaum-Saunders: An alternative fatigue life distribution. Hacettepe Journal of Mathematics and Statistics. 2014;43:473-510.

3. Johnson NL, Kotz S, Balakrishnan N. Continuous univariate distributions, vol. 2 of wiley series in probability and mathematical statistics: applied probability and statistics. Wiley, New York; 1995.

4. Dıaz-Garcıa JA, Leiva V. A new family of life distributions based on Birnbaum-Saunders distribution. Technical report I-02-17 (PE/CIMAT), Mexico. www. cimat. mx/ biblioteca/RepTec; 2002.

5. Díaz-García JA, Leiva-Sánchez Vc. A new family of life distributions based on the elliptically contoured distributions. Journal of Statistical Planning and Inference. 2005;128(2):445-57.

6. Owen W. Another look at the Birnbaum-Saunders distribution. Available from: http:// www.stat.lanl.gov/MMR2004/Extended%20 Abstract/WOwnn.pdf; 2004.

7. Owen WJ. A new three-parameter extension to the Birnbaum-Saunders distribution. IEEE Transactions on Reliability. 2006;55(3):475-9.

8. Leiva V. Chapter 1 - Genesis of the Birnbaum–Saunders Distribution. The Birnbaum-Saunders Distribution: Academic Press; 2016. p. 1-15.

9. Elsayed E, editor Accelerated Life Testing Model for a Generalized Birnbaum-Saunders Distribution. QUALITA2013; 2013.

10. Barros M, Paula GA, Leiva V. A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. Lifetime Data Analysis. 2008;14(3):316-32.

11. Barros M, Paula GA, Leiva V. An R implementation for generalized Birnbaum–Saunders distributions. Computational Statistics & Data Analysis. 2009;53(4):1511-28.

12. Boag JW. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. Journal of the Royal Statistical Society Series B (Methodological). 1949;11(1):15-53.

13. Berkson J, Gage RP. Survival curve for cancer patients following treatment. Journal of the American Statistical Association. 1952;47(259):501-15.

14. Farewell VT. The use of mixture models for the analysis of survival data with long-term survivors. Biometrics. 1982:1041-6.

15. Kuk AY, Chen C-H. A mixture model combining logistic regression with proportional hazards regression. Biometrika. 1992;79(3):531-41.

16. Peng Y, Dear KB, Denham J. A generalized F mixture model for cure rate estimation. Statistics in Medicine. 1998;17(8):813-30.

17. Yakovlev AY, Tsodikov AD. Stochastic models of tumor latency and their biostatistical applications: World Scientific; 1996.

18. Chen M-H, Ibrahim JG, Sinha D. A new Bayesian model for survival data with a surviving fraction. Journal of the American

Statistical Association. 1999;94(447):909-19.

19. Yin G, Ibrahim JG. A general class of Bayesian survival models with zero and nonzero cure fractions. Biometrics. 2005;61(2):403-12.

20. Cooner F, Banerjee S, Carlin BP, Sinha D. Flexible cure rate modeling under latent activation schemes. Journal of the American Statistical Association. 2007;102(478):560-72.

21. Rodrigues J, de Castro M, Cancho VG, Balakrishnan N. COM–Poisson cure rate survival models and an application to a cutaneous melanoma data. Journal of Statistical Planning and Inference. 2009;139(10):3605-11.

22. Castro Md, Cancho VG, Rodrigues J. A Bayesian Long-term Survival Model Parametrized in the Cured Fraction. Biometrical Journal. 2009;51(3):443-55.

23. Borges P, Rodrigues J, Balakrishnan N. Correlated destructive generalized power series cure rate models and associated inference with an application to a cutaneous melanoma data. Computational Statistics & Data Analysis. 2012;56(6):1703-13.

24. Cancho VG, Louzada F, Barriga GD. The Geometric Birnbaum–Saunders regression model with cure rate. Journal of Statistical Planning and Inference. 2012;142(4):993-1000.

25. Hashimoto EM, Ortega EM, Cordeiro GM, Cancho VG. The Poisson Birnbaum–Saunders model with long-term survivors. Statistics. 2014;48(6):1394-413.

26. Cordeiro GM, Cancho VG, Ortega EM, Barriga GD. A model with long-term survivors: negative binomial Birnbaum-Saunders. Communications in Statistics-Theory and Methods. 2016;45(5):1370-87.

27. Meshkat M, Baghestani A, Zayeri F. The Poisson Generalized Birnbaum-Saunders Cure Model and Application in Breast Cancer Data. J Biom Biostat. 2018;9(1):389.

28. Taketomi N, Yamamoto K, Chesneau C, Emura T. Parametric distributions for survival and reliability analyses, a review and historical sketch. Mathematics. 2022;10(20):3907.

29. Casella G, Berger RL. Statistical inference: Duxbury Pacific Grove, CA; 2002.

30. Klein JP, Moeschberger ML. Survival analysis: techniques for censored and truncated data: Springer Science & Business Media; 2005.

31. Betts JT. Practical methods for optimal control and estimation using nonlinear programming: SIAM; 2010.

32. Coles S, Bawa J, Trenner L, Dorazio P. An introduction to statistical modeling of extreme values: Springer; 2001.

33. Dunn PK, Smyth GK. Randomized quantile residuals. Journal of Computational and Graphical Statistics. 1996;5(3):236-44.

34. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2005;54(3):507-54.