

The Marshall-Olkin generalized gamma distribution

Gladys D.C. Barriga^a, Gauss M. Cordeiro^b, Dipak K. Dey^c, Vicente G. Cancho^d,
Francisco Louzada^d, Adriano K. Suzuki^{1,d}

^aFaculty of Engineering at Bauru, UNESP, Brazil;

^bDepartment of Statistics, Federal University of Pernambuco, Brazil;

^cDepartment of Statistics, University of Connecticut, USA;

^dDepartment of Applied Mathematics and Statistics, University of São Paulo, Brazil

Abstract

Attempts have been made to define new classes of distributions that provide more flexibility for modelling skewed data in practice. In this work we define a new extension of the generalized gamma distribution (Stacy, *The Annals of Mathematical Statistics*, **33**, 1187–1192, 1962) for Marshall-Olkin generalized gamma (MOGG) distribution, based on the generator pioneered by Marshall and Olkin (*Biometrika*, **84**, 641–652, 1997). This new lifetime model is very flexible including twenty one special models. The main advantage of the new family relies on the fact that practitioners will have a quite flexible distribution to fit real data from several fields, such as engineering, hydrology and survival analysis. Further, we also define a MOGG mixture model, a modification of the MOGG distribution for analyzing lifetime data in presence of cure fraction. This proposed model can be seen as a model of competing causes, where the parameter associated with the Marshall-Olkin distribution controls the activation mechanism of the latent risks (Cooner *et al.*, *Statistical Methods in Medical Research*, **15**, 307–324, 2006). The asymptotic properties of the maximum likelihood estimation approach of the parameters of the model are evaluated by means of simulation studies. The proposed distribution is fitted to two real data sets, one arising from measuring the strength of fibers and the other on melanoma data.

Keywords: cure fraction model, generalized gamma distribution, geometric distribution, maximum likelihood, lifetime data

1. Introduction

Standard lifetime distributions usually present very strong restrictions to produce bathtub curves, and thus appear to be inappropriate for data with this characteristic. The three-parameter generalized gamma (GG) (Stacy, 1962) distribution includes as special models the exponential, Weibull, gamma, and Rayleigh distributions, among others. It is suitable for modeling data with hazard rate function (hrf) of different forms (increasing, decreasing, bathtub and unimodal) and useful for estimating individual hazard functions and both relative hazards and relative times (Cox *et al.*, 2007). The GG distribution has been used in several research areas such as engineering, hydrology and survival analysis. Its probability density function (pdf) and cumulative distribution function (cdf) are given by (for $t > 0$)

$$f_{k,\beta,\tau}(t) = \frac{\tau}{\beta \Gamma(k)} \left(\frac{t}{\beta}\right)^{k\tau-1} \exp\left[-\left(\frac{t}{\beta}\right)^\tau\right] \quad (1.1)$$

¹ Corresponding author: Department of Applied Mathematics and Statistics, University of São Paulo, Avenida Trabalhador São-carlense, 400 - Centro CEP: 13566-59, São Carlos-SP, Brazil. E-mail: suzuki@icmc.usp.br

and

$$F_{k,\beta,\tau}(t) = \Gamma_G \left(\left[\frac{t}{\beta} \right]^\tau ; k \right), \quad (1.2)$$

respectively, where $\tau > 0$, $\beta > 0$, $k > 0$, $\Gamma_G(t; k) = \Gamma(k)^{-1} \int_0^t w^{k-1} e^{-w} dw$ is the incomplete gamma function ratio and $\Gamma(k) = \int_0^\infty w^{k-1} e^{-w} dw$ (for $k > 0$) is the gamma function. In the density function (1.1), β is a scale parameter and τ and k are shape parameters. The Weibull and gamma distributions are special models of (1.1) when $k = 1$ and $\tau = 1$, respectively. The GG distribution approaches the log-normal distribution when $\beta = 1$ and $k \rightarrow \infty$.

The GG distribution includes all four more common types of the hrf: monotonically increasing and decreasing, bathtub and unimodal (Cox *et al.*, 2007). This property is useful in reliability and survival analysis. This model has been used in several applied areas such as engineering, economics and survival analysis.

Now, we define an extended form of the density function (1.1) (for $t > 0$) given by

$$f(t) = \frac{|\tau|}{\beta \Gamma(k)} \left(\frac{t}{\beta} \right)^{k\tau-1} \exp \left[- \left(\frac{t}{\beta} \right)^\tau \right], \quad (1.3)$$

where τ is not zero and the other parameters are positive. The cdf corresponding to (1.3) becomes

$$F(t) = \Gamma_G \left(\left[\frac{t}{\beta} \right]^\tau ; k \right) \text{ for } \tau > 0 \quad \text{and} \quad F(t) = 1 - \Gamma_G \left(\left[\frac{t}{\beta} \right]^\tau ; k \right) \text{ for } \tau < 0. \quad (1.4)$$

In order to avoid convergence problems using the maximum likelihood method, Lawless (2002) proposed a re-parametrized density function with new parameters given by $\mu = \log(\beta) + \tau^{-1} \log(k)$, $\sigma = (\tau \sqrt{k})^{-1}$ and $\lambda = (\sqrt{k})^{-1}$ and adding the extra case $\lambda = 0$. So, we define the pdf

$$f(t) = \begin{cases} \frac{c(\lambda)}{\sigma t} \exp \left\{ \frac{(\log(t) - \mu)}{\lambda \sigma} - \frac{1}{\lambda^2} \exp \left[\lambda \frac{(\log(t) - \mu)}{\sigma} \right] \right\}, & \text{if } \lambda \neq 0, \\ \frac{1}{t \sqrt{2\pi\sigma}} \exp \left\{ -\frac{1}{2} \left[\frac{\log(t) - \mu}{\sigma} \right]^2 \right\}, & \text{if } \lambda = 0, \end{cases} \quad (1.5)$$

where $t > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $\lambda \in \mathbb{R}$ are the location, scale and shape parameters, respectively, and $c(\lambda) = |\lambda|/\Gamma(\lambda^{-2})$. The special case $\lambda = \sigma$ gives the two-parameter gamma distribution. The Weibull distribution arises when $\lambda = 1$, and the very special case $\lambda = \sigma = 1$ corresponds to the exponential distribution. The case $\lambda = 0$ is the log normal distribution and, for $\lambda = -1$, we obtain a reciprocal Weibull distribution. In addition, the half-normal distribution is obtained from (1.5) when $\sigma = \lambda = \sqrt{2}$.

The cdf (for $t > 0$) corresponding to (1.5) is given by

$$F(t) = \begin{cases} \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right] ; \lambda^{-2} \right\}, & \text{if } \lambda > 0, \\ \Phi \left[\frac{\log(t) - \mu}{\sigma} \right], & \text{if } \lambda = 0, \\ 1 - \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right] ; \lambda^{-2} \right\}, & \text{if } \lambda < 0, \end{cases} \quad (1.6)$$

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution.

Marshall and Olkin (1997) proposed a method of adding a parameter $\alpha > 0$ to define a class of distributions. If $\bar{F}(t)$ denotes a baseline survival function, they defined the Marshall and Olkin-F (MO-F) distribution by the survival function given by

$$\bar{G}(t) = \frac{\alpha \bar{F}(t)}{1 - \bar{\alpha} \bar{F}(t)} = \frac{\alpha \bar{F}(t)}{F(t) + \alpha \bar{F}(t)}, \quad -\infty < t < \infty, \alpha > 0, \tag{1.7}$$

where $\bar{\alpha} = 1 - \alpha$. The transformed distribution contains the baseline model as a special case when $\alpha = 1$. It has a stability property in the sense that the result of applying twice the transformation is also in the transformed model.

The MO-F density function, say $g(t)$, is given by

$$g(t) = \frac{\alpha f(t)}{[1 - \bar{\alpha} \bar{F}(t)]^2}, \quad -\infty < t < \infty, \tag{1.8}$$

where $f(t) = dF(t)/dt$ is the baseline density function.

Survival models with a surviving fraction (also known as cure rate models or long-term survival models) have generated significant interest in the survival analysis literature. Models that accommodate a cured fraction have widely developed. A very popular type of cure rate model is the mixture distribution introduced by Boag (1949) and Berkson and Gage (1952). Basic references on cure rate distributions are the books by Maller and Zhou (1996) and Ibrahim *et al.* (2001).

This paper introduces a new four-parameter model named the Marshall-Olkin generalized gamma (MOGG) distribution by inserting the cdf (1.6) in equation (1.7). This new lifetime model is very flexible and includes twenty one special models. The main advantage of the new family is because practitioners have a very flexible distribution to fit real data from several fields. The MOGG distribution is also modified to model the possibility that long-term survivors are presented in the data. In the proposed model, parameter α associated with the Marshall-Olkin distribution controls the activation mechanism of the latent risks (Cooner *et al.*, 2006).

The rest of the paper proceeds as follows. Sections 2–3 formulates the MOGG model and MOGG mixture model. Inference based on maximum likelihood for both models is addressed in Section 4. Two simulation studies are presented in Section 5 to investigate some finite sample properties. In Section 6, our methodology is illustrated on a real data set. Finally, Section 7 presents some concluding remarks.

2. The Marshall-Olkin generalized gamma distribution

The MOGG survival function is given by

$$\bar{G}(t) = \begin{cases} \frac{\alpha \left(1 - \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right]; \lambda^{-2} \right\}\right)}{1 - \bar{\alpha} \left(1 - \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right]; \lambda^{-2} \right\}\right)}, & \text{if } \lambda > 0, \\ \frac{\alpha \Phi \left[-\frac{\log(t) - \mu}{\sigma} \right]}{1 - \bar{\alpha} \Phi \left[-\frac{\log(t) - \mu}{\sigma} \right]}, & \text{if } \lambda = 0, \\ \frac{\alpha \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right]; \lambda^{-2} \right\}}{1 - \bar{\alpha} \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right]; \lambda^{-2} \right\}}, & \text{if } \lambda < 0. \end{cases} \tag{2.1}$$

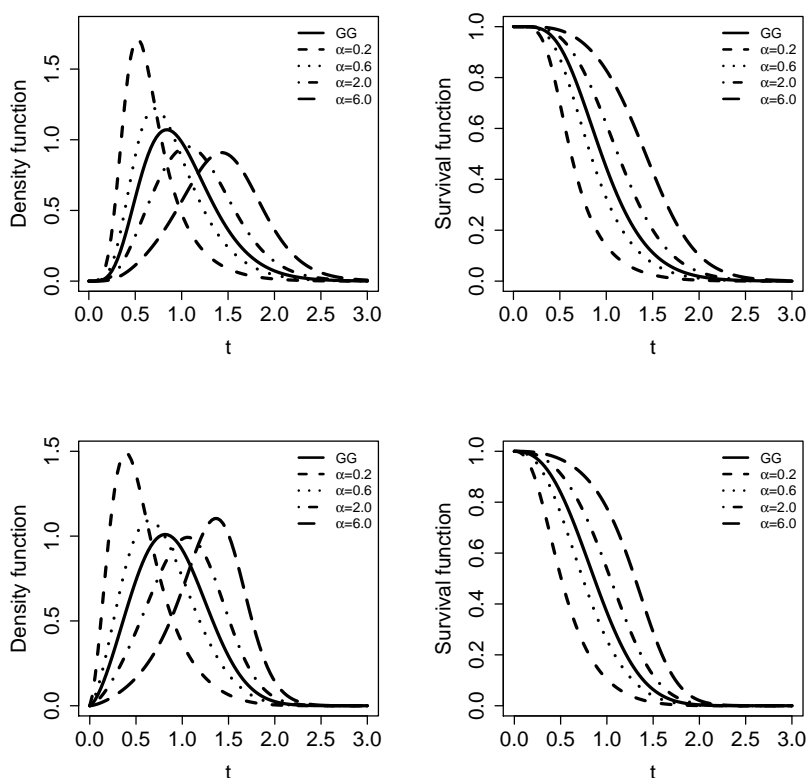


Figure 1: Density and survival functions of the Marshall-Olkin generalized gamma distribution and GG distribution with parameters $\mu = 0$, $\sigma = 1$, and $\lambda = 0.4$ (upper panel), $\lambda = 1$ (left panel). GG = generalized gamma.

The corresponding MOGG pdf becomes

$$g(t) = \begin{cases} \frac{c(\lambda)\alpha \exp\left\{\frac{1}{\lambda}\left[\frac{\log(t)-\mu}{\sigma}\right] - \frac{1}{\lambda^2} \exp\left\{\lambda\left[\frac{\log(t)-\mu}{\sigma}\right]\right\}\right\}}{t\sigma\left[1 - \bar{\alpha}\left(1 - \Gamma_G\left\{\lambda^{-2} \exp\left[\lambda\left(\frac{\log(t)-\mu}{\sigma}\right)\right]; \lambda^{-2}\right\}\right)\right]^2}, & \text{if } \lambda > 0, \\ \frac{\alpha \exp\left\{-\frac{1}{2}\left[\frac{\log(t)-\mu}{\sigma}\right]^2\right\}}{t\sqrt{2\pi}\sigma\left[1 - \bar{\alpha}\Phi\left[-\frac{\log(t)-\mu}{\sigma}\right]\right]^2}, & \text{if } \lambda = 0, \\ \frac{c(\lambda)\alpha \exp\left\{\frac{1}{\lambda}\left[\frac{\log(t)-\mu}{\sigma}\right] - \frac{1}{\lambda^2} \exp\left\{\lambda\left[\frac{\log(t)-\mu}{\sigma}\right]\right\}\right\}}{t\sigma\left[1 - \bar{\alpha}\Gamma_G\left\{\lambda^{-2} \exp\left[\lambda\left(\frac{\log(t)-\mu}{\sigma}\right)\right]; \lambda^{-2}\right\}\right]^2}, & \text{if } \lambda < 0. \end{cases} \quad (2.2)$$

Henceforth, we denote by T a random variable having the MOGG($\alpha, \mu, \sigma, \lambda$) density function (2.2). In Figure 1 plot the MOGG density functions for some fixed values of α and λ . These plots indicate that the new distribution is very flexible and that the values of λ and α have a substantial effect on skewness and kurtosis. The MOGG model includes several distributions listed as special models in Table 1. For example, the MO-exponential and MO-Weibull (Marshall and Olkin, 1997; Ghitany, 2005) distributions are obtained when $\lambda = \sigma = 1$ and $\lambda = 1$, respectively.

Table 1: Some special models of the MOGG distribution.

	Case	μ	σ	λ	Distribution	Reference
$\alpha = 1$	1	μ	σ	λ	GG	Cox <i>et al.</i> (2007)
	2	μ	σ	1	Weibull	
	3	μ	λ	λ	Gamma	
	4	μ	1/2	1	Rayleigh	
	5	μ	σ	0	Log-normal	
	6	μ	σ	-1	Inverse Weibull	
	7	μ	1	1	exponential	
$0 < \alpha < 1$	8	μ	σ	λ	Geometric GG	Ortega <i>et al.</i> (2011)
	9	μ	σ	1	Geometric Weibull	
	10	μ	λ	λ	Geometric Gamma	Barreto-Souza <i>et al.</i> (2011)
	11	μ	1/2	1	Geometric Rayleigh	
	12	μ	σ	0	Geometric Log-normal	
	13	μ	σ	-1	Geometric Inverse Weibull	
	14	μ	1	1	Geometric exponential	Adamidis and Loukas (1998)
$\alpha > 1$	8	μ	σ	λ	Complementary Geometric GG	Louzada <i>et al.</i> (2011)
	9	μ	σ	1	Complementary Geometric Weibull	
	10	μ	λ	λ	Complementary Geometric Gamma	
	11	μ	1/2	1	Complementary Geometric Rayleigh	
	12	μ	σ	0	Complementary Geometric Log-normal	
	13	μ	σ	-1	Complementary Geometric Inverse Weibull	
	14	μ	1	1	Complementary Geometric exponential	
$\alpha > 0$	15	μ	σ	λ	MOGG	Marshall and Olkin (1997)
	16	μ	σ	1	MO Weibull	
	17	μ	λ	λ	MO Gamma	
	18	μ	1/2	1	MO Rayleigh	
	19	μ	σ	0	MO Log-normal	
	20	μ	σ	-1	MO Inverse Weibull	
	21	μ	1	1	MO exponential	

MO = Marshall-Olkin; GG = generalized gamma.

We can generate a random variable t having the MOGG distribution based on equation (2.1). Let $\Gamma_G^{-1}(u; \gamma)$ denote the quantile function (qf) of the gamma distribution with mean and variance equal to γ , i.e., $\Gamma_G(\Gamma_G^{-1}(u; \gamma); \gamma) = u$. For $\lambda > 0$, we have

$$\log(t) = \mu + \frac{\sigma}{\lambda} \log \left[\Gamma_G^{-1} \left(\frac{u\alpha}{1 - u\bar{\alpha}}; \lambda^{-2} \right) \right], \tag{2.3}$$

where $U \sim U(0, 1)$. Similarly, for $\lambda < 0$,

$$\log(t) = \mu + \frac{\sigma}{\lambda} \log \left[\Gamma_G^{-1} \left(\frac{(1 - u)\alpha}{1 - (1 - u)\bar{\alpha}}; \lambda^{-2} \right) \right]. \tag{2.4}$$

Therefore, the qf of T , say $t = Q(u)$, can be easily obtained from (2.3) and (2.4).

From equations (2.1) and (2.2), the MOGG hrf is given by

$$h_{\text{MOGG}}(t) = \frac{h(t)}{1 - (1 - \alpha)\bar{F}(t)}, \quad t > 0, \tag{2.5}$$

where $h(t)$ and $\bar{F}(t)$ are the hazard and survival functions of the GG distribution, respectively. Note that $h_{\text{MOGG}}(t)/h(t)$ is increasing in t for $0 < \alpha < 1$ and decreasing for $\alpha > 1$. Further, $h(t) \leq h_{\text{MOGG}}(t) \leq h(t)/\alpha$ for $0 < \alpha < 1$, $h(t)/\alpha \leq h_{\text{MOGG}}(t) \leq h(t)$ for $\alpha > 1$ and that $\lim_{t \rightarrow \infty} h_{\text{MOGG}}(t) = \lim_{y \rightarrow \infty} h(y)$. Hence, the limit behavior of the MOGG hrf is the same as that one of the GG hrf. Figure 2 displays the plots of the MOGG hrf for some parameter values.

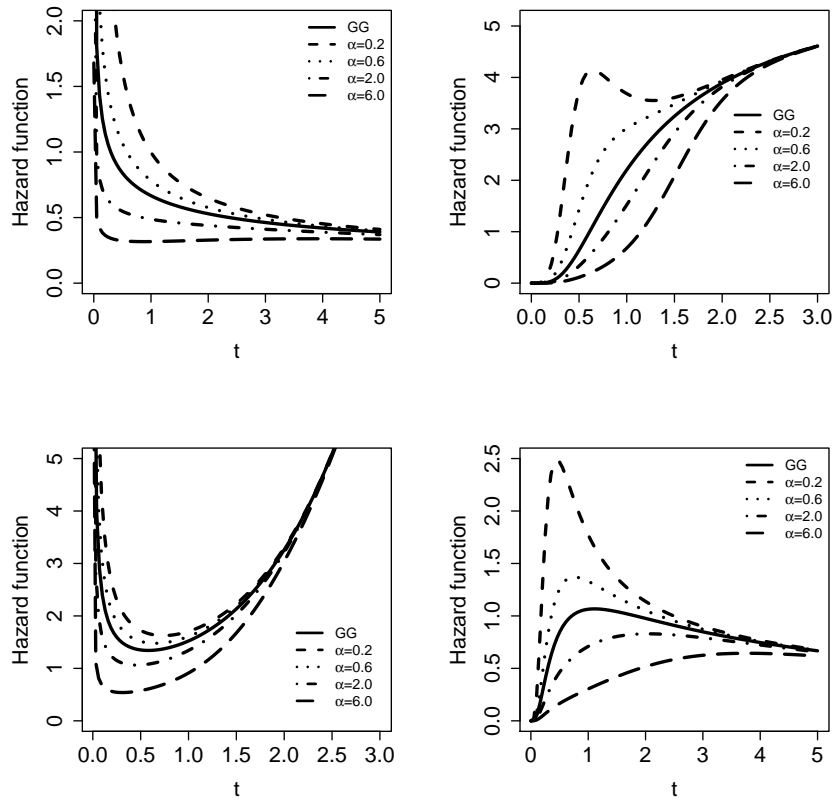


Figure 2: Marshall-Olkin generalized gamma hrf for some parameter values. hrf = hazard rate function; GG = generalized gamma.

3. The Marshall-Olkin generalized gamma mixture model

In survival and reliability studies, a part of the population may not be susceptible to the event of interest. Maller and Zhou (1996) indicate that it is adequate to consider a two components mixture model, in the sense that one component represents the failure or survival time of susceptible individuals to a certain event (in risk individuals; IR), while the other component represents the survival times of the non-susceptible individuals to the event (out of risk individuals; OR), allowing infinite survival times. An individual belongs to one group (or another) with certain probability. Then, the model formulation is described as follows. Let T be a random variable representing the time until the occurrence of an event of interest, and θ ($0 < \theta < 1$) be the probability of an individual belong to the OR group. Suppose a population for which there exists the possibility of cure. Then, the improper population survival function is given by (Maller and Zhou, 1996), $S_p(t) = \theta S_{OR}(t) + (1 - \theta)S_{IR}(t)$, where $S_{OR}(t)$ and $S_{IR}(t)$ are the survival functions of the OR and IR individuals, respectively. Following Maller and Zhou (1996), the OR individuals shall not present the event of interest, i.e., their failure times are infinite, so that $S_{OR}(t) = P(T > t|OR) = 1, \forall t > 0$. Then, we can rewrite $S_p(t)$ as

$$S_p(t) = \theta + (1 - \theta)S_{IR}(t). \quad (3.1)$$

All IR individuals will present the event of interest at the same time, i.e., $\lim_{t \rightarrow \infty} S_{IR}(t) = 0$. Consequently, we have $\lim_{t \rightarrow \infty} S_p(t) = \theta$, and therefore the survival function (not conditional) is improper and its limit corresponds to the OR individual proportion. The MOGG mixture is defined by selecting in (3.1) the MOGG survival function (2.1) ($S_{IR}(t)$), implying that

$$S_p(t) = \begin{cases} \theta + (1 - \theta) \frac{\alpha \left(1 - \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right]; \lambda^{-2} \right\}\right)}{1 - \bar{\alpha} \left(1 - \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right]; \lambda^{-2} \right\}\right)}, & \text{if } \lambda > 0, \\ \theta + (1 - \theta) \frac{\alpha \Phi \left[-\frac{\log(t) - \mu}{\sigma} \right]}{1 - \bar{\alpha} \Phi \left[-\frac{\log(t) - \mu}{\sigma} \right]}, & \text{if } \lambda = 0, \\ \theta + (1 - \theta) \frac{\alpha \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right]; \lambda^{-2} \right\}}{1 - \bar{\alpha} \Gamma_G \left\{ \lambda^{-2} \exp \left[\lambda \left(\frac{\log(t) - \mu}{\sigma} \right) \right]; \lambda^{-2} \right\}}, & \text{if } \lambda < 0. \end{cases} \quad (3.2)$$

The MOGG distribution in the MOGG mixture model can be interpreted as follows. Suppose that the event of interest in the IR group may be caused by an unknown competing cause leading to latent competing risk scenarios. Let M denote the unobservable number of causes of the event of interest for the IR group. Suppose that M follows a geometric distribution with mean $1/(1 - \alpha)$ ($0 < \alpha < 1$). The time for the j^{th} cause to produce the event of interest is denoted by Z_j (for $j = 1, \dots, M$). We assume that, conditional on M , the Z_j 's are independent and identically distributed random variables having the GG distribution given by (1.7). Further, we consider that Z_1, Z_2, \dots are independent of M . The observable time to the event of interest is defined by the random variable $T = \min\{Z_1, \dots, Z_M\}$. Under this setup, the survival function for an IR individual has the MOGG distribution (2.1). If $\alpha > 1$ and M has a geometric distribution with mean $1/(1 - \alpha^{-1})$ and $T = \max(T_1, \dots, T_M)$, then T has a survival function given by (2.1). Moreover, the proposed model in (3.2) with $\alpha = \theta$ yields a cure rate survival model with an activation mechanism (Cooner *et al.*, 2006; Cooner *et al.*, 2007). When the event of interest happens due to any one of the possible causes it gives the first activation scheme. The last activation is obtained when the event of interest only takes place after all M causes have been occurred. Finally, the model (2.1) with $\alpha = 1$ gives the GG mixture model, which is the survival cure rate model with random activation mechanism, where the distribution of activation of each cause is a discrete uniform distribution. Thus, the parameter α controls the activation mechanism of the risks in the proposed model.

The MOGG mixture is flexible, because the MOGG distribution is a wider family that contains most commonly used distributions, such as the exponential, Weibull, log normal and gamma models (Table 1).

4. Inference

4.1. Inference for the Marshall-Olkin generalized gamma model

Let t_1, \dots, t_n be a random sample of size n from the MOGG distribution with unknown parameter vector $\boldsymbol{\vartheta} = (\alpha, \mu, \sigma, \lambda)$. We estimate these parameters by the method of maximum likelihood. Setting $z_i = \sigma^{-1}[\log(t_i) - \mu]$, the log-likelihood function for $\boldsymbol{\vartheta}$ is given by

$$\ell(\boldsymbol{\vartheta}) \propto \sum_{i=1}^n \ell_i(\boldsymbol{\vartheta}), \quad (4.1)$$

where

$$\ell_i(\boldsymbol{\theta}) = \begin{cases} \log[c(\lambda)] + \log(\alpha) + \frac{z_i}{\lambda} + \frac{e^{\lambda z_i}}{\lambda^2} - \log(\sigma t_i) - 2 \log \left\{ 1 - \bar{\alpha} \left[1 - \Gamma_G(\lambda^{-2} e^{\lambda z_i}; \lambda^{-2}) \right] \right\}, & \text{if } \lambda > 0, \\ \log(\alpha) - 0.5z_i^2 - \log(\sigma t_i) - 2 \log \{ 1 - \bar{\alpha} \Phi(-z_i) \}, & \text{if } \lambda = 0, \\ \log[c(\lambda)] + \log(\alpha) + \frac{z_i}{\lambda} + \frac{e^{\lambda z_i}}{\lambda^2} - \log(\sigma t_i) - 2 \log \left\{ 1 - \bar{\alpha} \Gamma_G(\lambda^{-2} e^{\lambda z_i}; \lambda^{-2}) \right\}, & \text{if } \lambda < 0. \end{cases}$$

The maximum likelihood estimate (MLE) $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by maximizing the log-likelihood function (4.1). Numerical maximization of the log-likelihood function $\ell(\boldsymbol{\theta})$ is accomplished by using the R software (R Development Core Team, 2013). The computational program is available from the authors upon request. Under general regularity conditions (Maller and Zhou, 1996), we can approximate the distribution of $\hat{\boldsymbol{\theta}}$ by the multivariate normal distribution with mean vector $\boldsymbol{\theta}$ and covariance matrix $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}) = \{-\partial^2 \ell(\boldsymbol{\theta}; \mathbf{t}, \boldsymbol{\delta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top\}^{-1}$, which can be evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. The required second derivatives can be computed numerically.

We can easily check the adequacy of the fitted GG model by testing the null hypothesis $H_0 : \alpha = 1$. The log-likelihood ratio (LR) statistic for testing H_0 is given by $\Lambda = 2[\ell(\hat{\alpha}, \hat{\mu}, \hat{\sigma}, \hat{\lambda}) - \ell(1, \tilde{\mu}, \tilde{\sigma}, \tilde{\lambda})]$, where $\hat{\alpha}$, $\hat{\mu}$, $\hat{\sigma}$, and $\hat{\lambda}$ are the unrestricted estimates and $\tilde{\mu}$, $\tilde{\sigma}$, and $\tilde{\lambda}$ are the restricted estimates under H_0 . The limiting null distribution of this statistic is chi-square with one degree of freedom.

4.2. Inference for the Marshall-Olkin generalized gamma mixture model

Let us consider the situation when the failure time T in Section 3 is not completely observed and is subject to right censoring. Let C_i denote the censoring time. In a sample of size n , we then observe $y_i = \min\{T_i, C_i\}$ and $\delta_i = \mathbf{I}(T_i \leq C_i)$, where $\delta_i = 1$ if T_i is a failure time and $\delta_i = 0$ if it is right censored, for $i = 1, \dots, n$.

Let $\mathbf{x}_i = (x_{i1}, \dots, x_{ip_1})^\top$ and $\mathbf{w}_i = (w_{i1}, \dots, w_{ip_2})^\top$ denote the vectors of covariates for the i^{th} individual. Further, we relate θ_i (the cure fraction) to covariates \mathbf{x}_i by the logistic link and μ_i to covariates \mathbf{w}_i by the identity link, respectively, i.e.,

$$\log\left(\frac{\theta_i}{1 - \theta_i}\right) = \mathbf{x}_i^\top \boldsymbol{\beta}_1 \quad \text{and} \quad \mu_i = \mathbf{w}_i^\top \boldsymbol{\beta}_2, \quad (4.2)$$

where $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ denote the corresponding parameter vectors. The mixture model is not identifiable when the cure fraction is a constant θ , but is identifiable when it is modeled by a logistic regression with non-constant covariates (Li *et al.*, 2001).

We can write the likelihood function for $\boldsymbol{\theta} = (\sigma, \lambda, \boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)^\top$ from (4.2) under non-informative censoring as

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^n f_p(y_i; \boldsymbol{\theta})^{\delta_i} S_p(y_i; \boldsymbol{\theta})^{1-\delta_i}, \quad (4.3)$$

where $S_p(y; \boldsymbol{\theta})$ is the improper survival function in (3.2) and $f_p(y; \boldsymbol{\theta}) = \partial S_p(y; \boldsymbol{\theta}) / \partial y$ is the corresponding improper pdf.

From the likelihood function in (4.3), the maximum likelihood estimation of the parameter $\boldsymbol{\theta}$ can be conducted. Numerical maximization of the log-likelihood function $\ell(\boldsymbol{\theta}) = \log(L(\boldsymbol{\theta}))$ is performed using the R software (R Development Core Team, 2013). Under general regularity conditions (Maller

Table 2: Averages of maximum likelihood estimates, SD, RMSE, CP of the parameters of the Marshall-Olkin generalized gamma model

n		$\alpha = 0.2$				$\alpha = 2.0$			
		$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\alpha}$
50	Mean	0.932	0.502	2.207	0.256	1.065	0.451	2.266	1.982
	SD	0.242	0.143	0.660	0.112	0.215	0.139	0.568	1.244
	RMSE	0.251	0.143	0.691	0.125	0.224	0.147	0.627	1.244
	CP	0.913	0.918	0.921	0.920	0.934	0.931	0.925	0.957
100	Mean	0.988	0.477	2.297	0.227	1.052	0.462	2.158	1.971
	SD	0.233	0.144	0.731	0.100	0.175	0.109	0.360	1.076
	RMSE	0.233	0.146	0.789	0.104	0.183	0.116	0.393	1.076
	CP	0.932	0.945	0.934	0.945	0.952	0.943	0.948	0.933
200	Mean	1.031	0.462	2.299	0.202	1.033	0.477	2.096	1.987
	SD	0.185	0.119	0.661	0.073	0.134	0.083	0.227	0.851
	RMSE	0.187	0.125	0.725	0.073	0.137	0.086	0.246	0.851
	CP	0.943	0.950	0.954	0.948	0.954	0.948	0.952	0.953
400	Mean	1.028	0.471	2.163	0.195	1.032	0.478	2.063	1.915
	SD	0.136	0.086	0.396	0.050	0.093	0.057	0.149	0.548
	RMSE	0.139	0.091	0.428	0.050	0.098	0.061	0.161	0.554
	CP	0.948	0.946	0.951	0.946	0.951	0.945	0.952	0.947

SD = standard deviation; RMSE = square root of mean square error; CP = coverage probability.

and Zhou, 1996), the MLE $\hat{\boldsymbol{\vartheta}}$ has an approximate multivariate normal distribution with mean vector $\boldsymbol{\vartheta}$ and covariance matrix $\boldsymbol{\Sigma}(\hat{\boldsymbol{\vartheta}})$, which can be estimated by $\hat{\boldsymbol{\Sigma}}(\hat{\boldsymbol{\vartheta}}) = \{-\partial^2 \ell(\boldsymbol{\vartheta}; \mathbf{t}, \boldsymbol{\delta}) / (\partial \boldsymbol{\vartheta} \partial \boldsymbol{\vartheta}^T)\}^{-1}$, evaluated at $\boldsymbol{\vartheta} = \hat{\boldsymbol{\vartheta}}$. The second derivatives of this matrix can be computed numerically.

Hypothesis tests can also be conducted. Let $\boldsymbol{\vartheta}_1$ and $\boldsymbol{\vartheta}_2$ be proper disjoint subsets of $\boldsymbol{\vartheta}$. We aim to test $H_0 : \boldsymbol{\vartheta}_1 = \boldsymbol{\vartheta}_{01}$ against $H_1 : \boldsymbol{\vartheta}_1 \neq \boldsymbol{\vartheta}_{01}$ ($\boldsymbol{\vartheta}_2$ unspecified). Let $\hat{\boldsymbol{\vartheta}}_0$ maximize $L(\boldsymbol{\vartheta})$ constrained to H_0 and define the LR statistic as $\Lambda = 2[\ell(\hat{\boldsymbol{\vartheta}}) - \ell(\hat{\boldsymbol{\vartheta}}_0)]$, where $\ell(\cdot)$ is the log-likelihood. Under H_0 and general regularity conditions, Λ converges in distribution to the chi-square distribution with $\dim(\boldsymbol{\vartheta}_1)$ degrees of freedom.

Alternatively, non-nested models can be compared using the Akaike information criterion (AIC) given by $AIC = -2\ell(\hat{\boldsymbol{\vartheta}}) + 2\#(\boldsymbol{\vartheta})$ and the Schwartz-Bayesian criterion (SBC) defined by $SBC = -2\ell(\hat{\boldsymbol{\vartheta}}) + \#(\boldsymbol{\vartheta}) \log(n)$, where $\#(\boldsymbol{\vartheta})$ is the number of model parameters. The model with the smallest value of any of these criteria (among all models considered) is commonly taken as the preferred model for describing a given dataset.

5. Simulation study

Here, we evaluate the performance of the MLEs of the parameters of the MOGG model and MOGG mixture model by means of two simulation studies.

5.1. Simulated Marshall-Olkin generalized gamma mixture model

From equation (2.3), we generate 1,000 samples of size $n = 50, 100, 200,$ and 400 from the MOGG model with parameters $\mu = 1.0, \sigma = 0.5, \lambda = 2.0,$ and $\alpha = 0.2$ and 2.0 . For each configuration, we compute the average of the MLEs of the model parameters, their standard deviations (SDs), the square root of the mean squared errors (RMSEs) and the coverage probabilities (CPs) of the 95% intervals of the MLEs.

Table 2 reports the simulation results. We note that the averages of the MLEs of the parameters of the MOGG model are close to the true values. As expected, the SDs and RMSEs decrease as

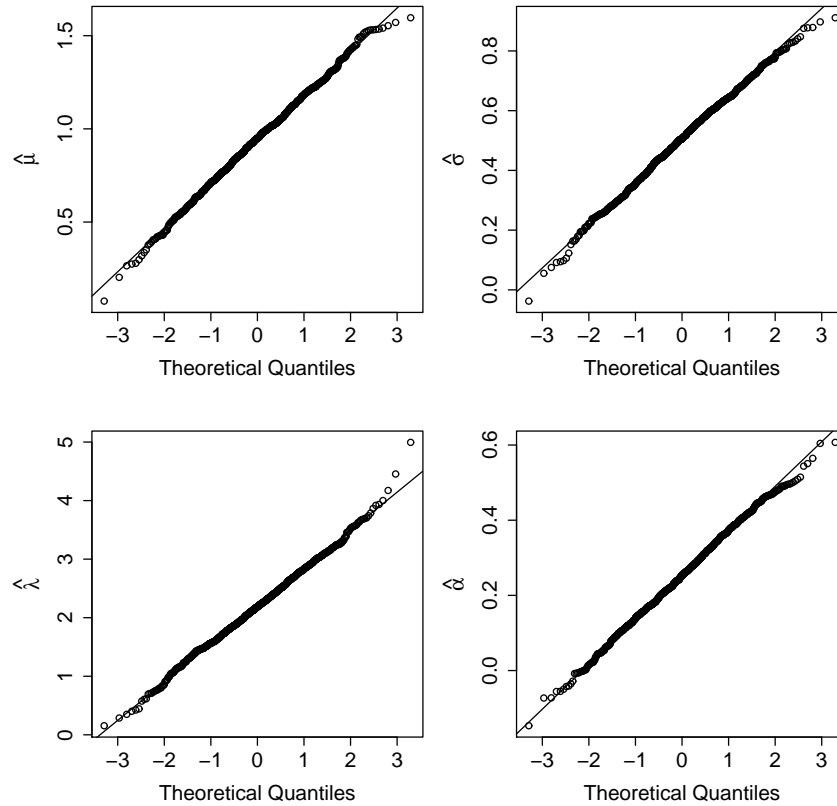


Figure 3: QQ-normal plots of the maximum likelihood estimates of the parameters ($\mu = 1.0$, $\sigma = 0.5$, $\lambda = 2.0$, and $\alpha = 0.2$) for the Marshall-Olkin generalized gamma model with sample size $n = 50$.

the sample size increases. Table 2 also shows that the CP becomes closer to the nominal value as the sample size increases. Further, we plot the empirical distributions of the MLEs $\hat{\mu}$, $\hat{\sigma}$, $\hat{\lambda}$, and $\hat{\alpha}$ for the sample size 50 (Figure 3). These plots reveal that normal distribution provides a reasonable approximation for the distributions of these estimates.

5.2. Simulated Marshall-Olkin generalized gamma mixture model

In this study, we consider the MOGG mixture model given in (3.2) with parameters μ_i , $\sigma = 0.5$, $\lambda = 2$, $\alpha = 0.2, 2$, and θ_i for $i = 1, \dots, n$. In the simulation study, we have two covariates, say x_i and w_i , such that x_i is generated from a Bernoulli(0.5) distribution and w_i is generated from the $N(0, 1)$ distribution. Thus, under the logit link, $\log(\theta_i/(1 - \theta_i)) = \beta_{10} + \beta_{11}x_i$ and $\mu_i = \beta_{20} + \beta_{21}w_i$, where $\beta_{10} = -0.5$, $\beta_{11} = 0.7$, $\beta_{20} = 1$, and $\beta_{21} = 0.5$. The censoring times are sampled from the Uniform(0, τ), where τ is set in order to control the proportion of censored observations on average to be approximately 60%.

We consider sample sizes of $n = 100, 300$, and 600 . For each of these schemes, we perform 1,000 simulations to calculate the average of the MLEs, the mean squared errors (MSE) of the MLEs and coverage probabilities of 95% confidence intervals for the parameters in model (3.2). Table 3 given the simulation results. We note that the averages of MLEs are close to the true values, the MSEs

Table 3: Averages of maximum likelihood estimates, SD, RMSE, CP of the parameters of the Marshall-Olkin generalized gamma model

α	n		$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}_{1_0}$	$\hat{\beta}_{1_1}$	$\hat{\beta}_{2_0}$	$\hat{\beta}_{2_1}$
0.2	100	Mean	0.572	1.868	0.271	-0.515	0.708	0.871	0.541
		SD	0.138	0.428	0.132	0.222	0.310	0.257	0.223
		RMSE	0.156	0.448	0.150	0.222	0.310	0.287	0.245
		CP	0.967	0.987	0.978	0.959	0.948	0.960	0.939
	300	Mean	0.548	1.933	0.255	-0.512	0.710	0.899	0.518
		SD	0.124	0.378	0.108	0.172	0.231	0.234	0.219
		RMSE	0.133	0.383	0.121	0.172	0.231	0.254	0.222
		CP	0.956	0.958	0.962	0.946	0.937	0.958	0.949
	600	Mean	0.526	1.982	0.232	-0.494	0.688	0.947	0.482
		SD	0.115	0.358	0.089	0.113	0.150	0.204	0.195
		RMSE	0.117	0.358	0.095	0.113	0.150	0.210	0.200
		CP	0.954	0.948	0.952	0.946	0.937	0.959	0.949
2.0	100	Mean	0.502	2.083	2.273	-0.507	0.706	1.017	0.511
		SD	0.125	0.389	1.289	0.251	0.348	0.205	0.233
		RMSE	0.125	0.397	1.317	0.251	0.348	0.206	0.235
		CP	0.977	0.977	0.968	0.950	0.952	0.958	0.943
	300	Mean	0.500	2.074	2.235	-0.514	0.700	1.010	0.501
		SD	0.114	0.316	1.169	0.183	0.240	0.188	0.215
		RMSE	0.114	0.325	1.192	0.183	0.240	0.188	0.216
		CP	0.960	0.948	0.960	0.953	0.948	0.951	0.954
	600	Mean	0.490	2.069	2.084	-0.501	0.699	1.023	0.452
		SD	0.092	0.248	0.908	0.132	0.172	0.151	0.115
		RMSE	0.093	0.257	0.912	0.132	0.172	0.152	0.117
		CP	0.949	0.962	0.981	0.946	0.948	0.956	0.949

SD = standard deviation; RMSE = square root of mean square error; CP = coverage probability.

Table 4: The AIC and SBC statistics for the fitted distributions

Distributions	Criterion		
	$-2 \max \ell()$	AIC	SBC
G	47.90	51.90	56.20
GG	29.17	35.17	41.60
MOGG	24.06	32.07	40.63

AIC = Akaike information criterion; SBC = Schwartz-Bayesian criterion; G = gamma; GG = generalized gamma; MOGG = Marshall-Olkin generalized gamma

decrease as sample size increases and the empirical coverage probabilities are closer to the nominal coverage level when sample size increases.

6. Applications

6.1. Strength of fibers

The data set is obtained from Smith and Naylor (1987) and describe the strengths of 1.5 cm glass fibers, measured at the National Physical Laboratory, England. This data set is of size $n = 63$ whose lowest value, first quartile, mean, median, highest value, and SD are equal to 0.550, 1.375, 1.507, 1.590, 2.240, and 0.3241, respectively.

The gamma (G), GG, and MOGG distributions are fitted to these data. For comparing the fitted models, we compute the AIC and SBC statistics. Table 4 lists the values of these criteria. According to both criteria, the MOGG and GG distributions are the best models. We also emphasize the gain provided by the MOGG distribution in relation to beta generalized gamma distribution (Cordeiro *et*

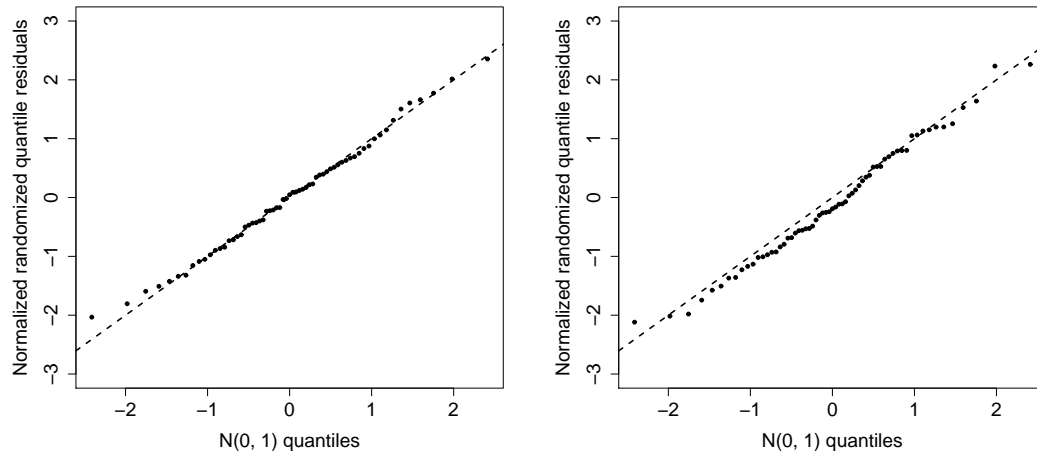


Figure 4: QQ plot of the normalized quantile residuals with an identity line for the distributions MOGG (left panel) and GG (right panel). MOGG = Marshall-Olkin generalized gamma; GG = generalized gamma.

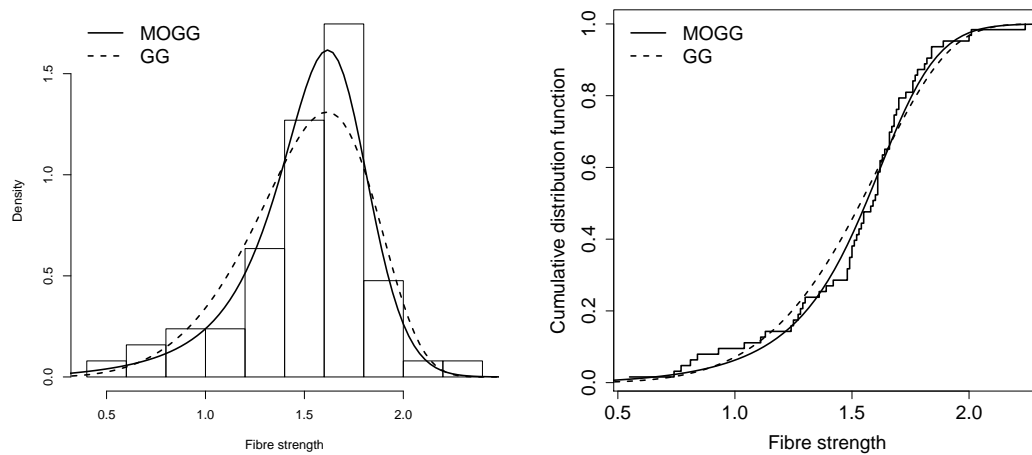


Figure 5: Histogram of strength and fitted density functions (left panel) and empirical cumulative function of strength and fitted cumulative functions (right panel). MOGG = Marshall-Olkin generalized gamma; GG = generalized gamma.

al., 2013) (Table 2).

The LR statistics for testing the hypotheses $H_0 : G$ versus $H_1 : MOGG$ and $H_0 : GG$ versus $H_1 : MOGG$ are $\Lambda = 22.84$ (2 d.f., p -value < 0.0001) and $\Lambda = 5.29$ (1 d.f., p -value = 0.021), respectively. Therefore, we reject the null hypotheses in both cases in favor of the MOGG distribution at the 5% level of significance. Figure 5 displays the plots of the MOGG and GG fitted densities to these data. They indicate that the MOGG distribution provides a better fit than the GG model.

The QQ plot of the normalized randomized quantile residuals (Dunn and Smyth, 1996; Rigby and Stasinopoulos, 2005) in Figure 4 (left panel) suggests that the MOGG model is acceptable. Each point in Figure 4 corresponds to the median of five sets of ordered residuals. The values of the criteria in

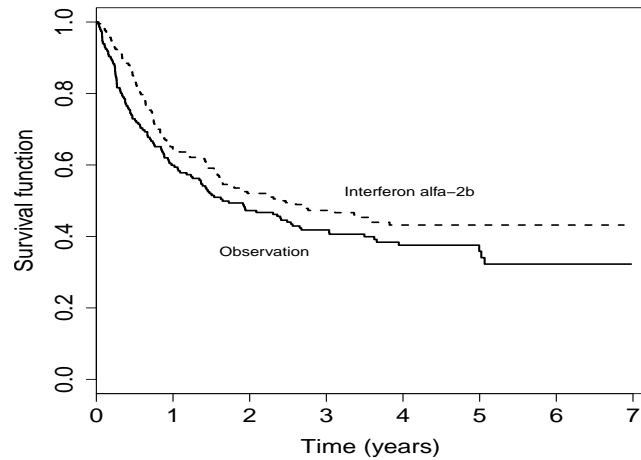


Figure 6: Kaplan-Meier estimate of the surviving function of high-dose interferon and observation groups.

Table 4, the LR statistics and the QQ plots in Figure 4, reveal that the MOGG model is the best model to these data. The parameter estimates (and 95% asymptotic confidence intervals) for the MOGG distribution are: $\hat{\alpha} = 17.801$ (1.062, 298.508), $\hat{\mu} = 0.114$ (-0.3901, 0.618), $\hat{\sigma} = 0.334$ (0.144, 0.773), and $\hat{\lambda} = 1.240$ (0.025, 2.45), so that the null hypothesis $H_0 : \lambda = 0$ is rejected at the significance level of 5%. As the confidence interval for alpha is huge, for mu includes zero and for lambda includes 1, so the related model is the Complementary Geometric Weibull (presented in case 9 on Table 1).

6.2. Melanoma data

In this section, we demonstrate an application of our models described in Section 3 to a well-known dataset on a Phase III cutaneous melanoma clinical trial conducted by the Eastern Cooperative Oncology Group (Kirkwood *et al.*, 2000). The incidence of melanoma ranks among the highest among solid tumor growths, with high mortality rates (between 60–75%) despite early detection and screening (Cooner *et al.*, 2007). The dataset here comes from an assay for the evaluation of postoperative treatment performance with a high dose of a certain drug (interferon alpha-2b) in order to prevent recurrence. Patients included in the study were from 1991 to 1995, and follow-up was conducted until 1998. The data were taken from Ibrahim *et al.* (2001) (labeled as E1690 data, available at <http://merlot.stat.uconn.edu/~mhchen/survbook/>). After deleting subjects with incomplete data and missing observation times, we have a subset of $n = 408$ patients with approximately 43% of censoring. We consider the relapse-free survival (RFS) time (in years) as the response variable.

The following information were collected from each patient: Observed time (in years, mean = 2.31, SD = 1.93); x_{1i} : treatment (0: observation, $n = 198$; 1: interferon alfa-2b, $n = 210$); x_{2i} : age (in years, mean = 48.1, SD = 13.1); x_{3i} : nodal number (1: $n = 110$; 2: $n = 131$; 3: $n = 86$; 4: $n = 81$), and x_{4i} : tumor thickness (in mm, mean = 3.98 and SD = 3.22), $i = 1, \dots, 408$. Kaplan-Meier curves stratified by treatment in Figure 6 level off between 0.25 and 0.42. This behavior indicates that models that ignore the possibility of cure will not be suitable for these data.

We fit the MOGG mixture model. Table 5 presents the MLEs, the standard errors and the p -values for the estimates of the model parameters. The estimate of the parameter (α) presents an evidence against the mixture GG model ($H_0 : \alpha = 1$). This estimate indicates that the event of interest happens

Table 5: MLEs of the parameters for the MOGG model with the covariate treatment

MOGG model	Parameter	Estimate	Standard error	p -value
Complete	α	-0.082	0.165	-
	σ	-1.501	0.692	-
	λ	-0.198	0.434	-
	$\beta_{1,\text{intercept}}$	-2.477	0.932	0.008
	$\beta_{1,\text{treatment}}$	-0.154	0.329	0.639
	$\beta_{1,\text{age}}$	-0.028	0.013	0.037
	$\beta_{1,\text{module}}$	-0.642	0.234	0.006
	$\beta_{1,\text{thickness}}$	-0.186	0.098	0.057
	$\beta_{2,\text{intercept}}$	-2.537	2.408	0.292
	$\beta_{2,\text{treatment}}$	-0.372	0.198	0.061
	$\beta_{2,\text{age}}$	-0.004	0.007	0.523
	$\beta_{2,\text{module}}$	-0.274	0.091	0.003
	$\beta_{2,\text{thickness}}$	-0.008	0.028	0.771
	Reduced	α	-0.035	0.063
σ		-2.240	1.286	-
λ		-0.158	0.606	-
$\beta_{1,\text{intercept}}$		-1.560	0.816	0.056
$\beta_{1,\text{module}}$		-0.789	0.311	0.011
$\beta_{1,\text{thickness}}$		-0.257	0.134	0.055
$\beta_{2,\text{intercept}}$		-4.329	2.857	0.130
$\beta_{2,\text{treatment}}$		-0.422	0.185	0.022
$\beta_{2,\text{module}}$		-0.283	0.088	0.001

MLEs = maximum likelihood estimates; MOGG = Marshall-Olkin generalized gamma.

due to any one of the possible causes (first activation scheme), since, $\hat{\alpha} \in (0, 1)$. Considering the LR statistic, we test the effect of some covariates in the model, i.e., $H_0 : \beta_{1,\text{age}} = \beta_{1,\text{treatment}} = \beta_{2,\text{thickness}} = \beta_{2,\text{age}} = 0$ versus H_1 : at least one of the β 's are different from zero, yielding $\Lambda = 5.054$ (p -value = 0.2812), and thus indicating that the effects of the covariates are not significant. Hence, Table 5 presents the MLEs, the standard errors and the p -values for the parameters of the MOGG model without those covariates (reduced model). We can observe that the covariate $\beta_{1,\text{thickness}}$ is significant at the level of 10% and the others are significant at the 5% level. Now, we consider the reduced model as our working model, and present further analysis results based on this model. The QQ plot of the normalized randomized quantile residuals (Dunn and Smyth, 1996; Rigby and Stasinopoulos, 2005) for the reduced model is presented in the left panel of Figure 7 suggesting that the MOGG mixture model produces an adequate fit.

The MLEs of the cure fraction (and standard errors) for patients with tumor thickness of 3.175 mm (median thickness) and stratified by nodal category from 1 to 4 are: 0.4886 (0.0817), 0.3027 (0.0751), 0.1647 (0.0801), and 0.0822 (0.0651), respectively. Standard errors are obtained after application of the delta method. The right panel of Figure 7 shows that the cure fraction decreases more rapidly for patients with a lower nodal category.

We conclude our application dealing with the MLE of the proportion of patients who survived beyond a certain fixed time, which is the practical interest to practitioners. For the sake of illustration, we choose five years. This proportion is estimated from $S_p(5)$. Table 6 gives the MLE of $S_p(5)$ stratified by nodal category (from 1 to 4) and treatment with median tumor thickness (3.175 mm). Figure 8 displays the plots of the surviving functions for patients stratified by nodal category and median tumor thickness (3.175 mm). We note that the survival probability diminishes rapidly with increasing nodal category; in addition, the survival probability is greater for patients treated with interferon alfa-2b.

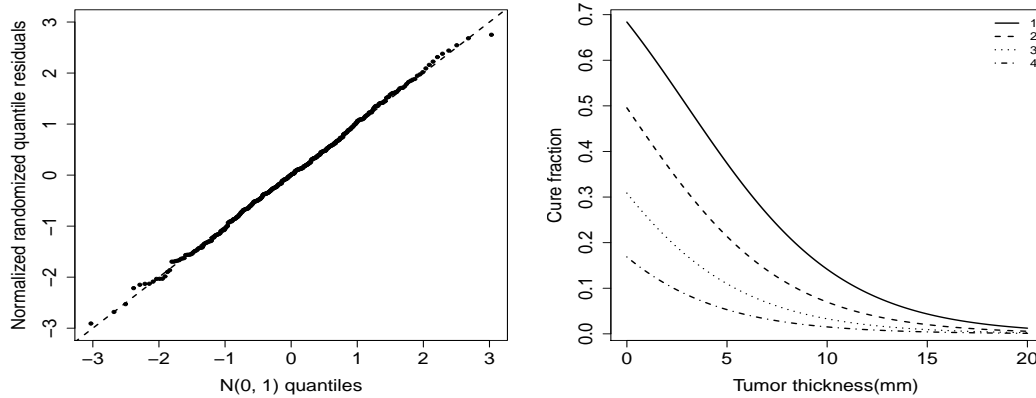


Figure 7: Left Panel: QQ plot of the normalized randomized quantile residuals, with the identity line where each point corresponds to the median of 5 sets of ordered residuals; Right panel: Cure fraction stratified by nodule category and tumor thickness, for the MOGG mixture model. MOGG = Marshall-Olkin generalized gamma.

Table 6: Survivor probability of patients after five years for various nodal categories stratified by treatment

Treatment	Nodal category	MLE	Standard error	95% confidence interval	
				LL	RL
Observation	1	0.592	0.108	0.381	0.802
	2	0.443	0.088	0.271	0.615
	3	0.333	0.108	0.122	0.543
	4	0.267	0.130	0.012	0.522
Interferon alfa-2b	1	0.627	0.136	0.361	0.893
	2	0.492	0.106	0.283	0.701
	3	0.391	0.157	0.084	0.699
	4	0.331	0.103	0.129	0.533

MLE = maximum likelihood estimates; LL = lower limit; UL = upper limit.

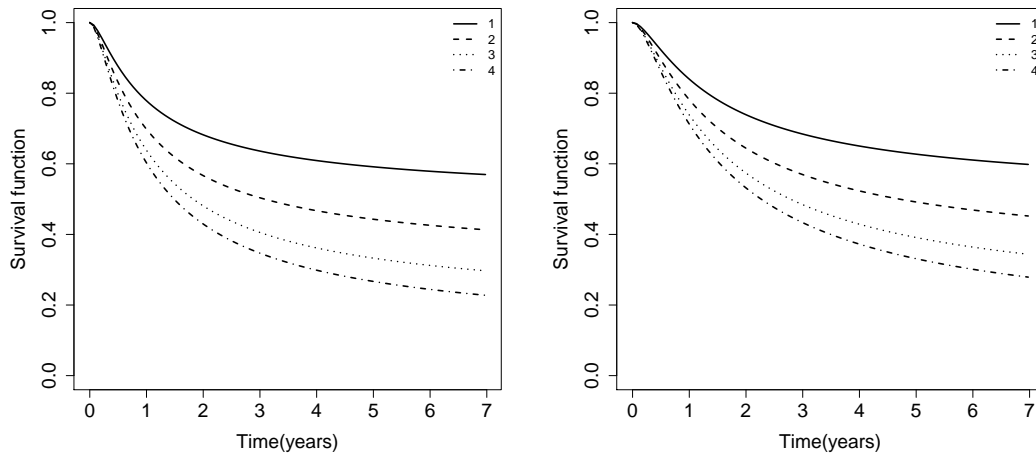


Figure 8: Surviving function of patients stratified by ulceration status stratified by nodal category (from 1 to 4) with median tumor thickness (3.175 mm) and treatment (left panel: observation; right panel:interferon alfa-2b).

7. Conclusion

In this paper, we define a new lifetime model named the *Marshall-Olkin generalized gamma* (MOGG) distribution as an extension of the generalized gamma distribution (Stacy, 1962). The proposed model includes twenty one special models. Some structural properties of the proposed distribution are provided such as moments, quantile and generating functions. We base the inference on maximum likelihood estimation. We also define a MOGG mixture model for the analysis of lifetime data with cure fraction. Two simulation studies are presented to investigate some finite sample properties of the maximum likelihood estimates. The MOGG mixture model can be seen as a model of competing causes, where activation mechanism of the causes is controlled by a parameter of the model. We apply the new models to two real data sets to illustrate their potentiality.

References

- Adamidis K and Loukas S (1998). A lifetime distribution with decreasing failure rate, *Statistics & Probability Letters*, **39**, 35–42.
- Barreto-Souza W, de Morais AL and Cordeiro GM (2011). The Weibull-Geometric distribution, *Journal of Statistical Computation and Simulation*, **81**, 645–657.
- Berkson J and Gage RP (1952). Survival curve for cancer patients following treatment, *Journal of the American Statistical Association*, **47**, 501–515.
- 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.
- 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.
- Cooner F, Banerjee S, and McBean AM (2006). Modelling geographically referenced survival data with a cure fraction, *Statistical Methods in Medical Research*, **15**, 307–324.
- Cordeiro GM, Castellares F, Montenegro LC, and de Castro M (2013). The beta generalized gamma distribution, *Statistics*, **47**, 888–900.
- Cox C, Chu H, Schneider MF, and Muñoz A (2007). Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution, *Statistics in Medicine*, **26**, 4352–4374.
- Dunn PK and Smyth GK (1996). Randomized quantile residuals, *Journal of Computational and Graphical Statistics*, **5**, 236–244.
- Ghitany ME (2005). Marshall-Olkin extended Pareto distribution and its application, *International Journal of Applied Mathematics*, **18**, 17–32.
- Ibrahim JG, Chen MH, and Sinha D (2001). *Bayesian Survival Analysis*, Springer, New York.
- Kirkwood JM, Ibrahim JG, Sondak VK, *et al.* (2000). High- and low-dose interferon alfa-2b in high-risk melanoma: first analysis of intergroup trial E1690/S9111/C9190, *Journal of Clinical Oncology*, **18**, 2444–2458.
- Lawless JF (2002). *Statistical Models and Methods for Lifetime Data* (2nd ed), Wiley, New York.
- Li CS, Taylor JMG, and Sy JP (2001). Identifiability of cure models, *Statistics & Probability Letters*, **54**, 389–395.
- Louzada F, Roman M, and Cancho VG (2011). The complementary exponential geometric distribution: model, properties, and a comparison with its counterpart, *Computational Statistics & Data Analysis*, **55**, 2516–2524.
- Maller RA and Zhou, X (1996). *Survival Analysis with Long-Term Survivors*, Wiley, New York.
- Marshall AW and Olkin I (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families, *Biometrika*, **84**, 641–652.

- Ortega EMM, Cordeiro GM, and de Pascoa MAR (2011). The generalized Gamma Geometric distribution, *Journal of Statistical Theory and Applications*, **3**, 433–454.
- R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
- Rigby RA and Stasinopoulos DM (2005). Generalized additive models for location, scale and shape (with discussion), *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **54**, 507–554.
- Smith RL and Naylor JC (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **36**, 358–369.
- Stacy EW (1962). A generalization of the gamma distribution, *The Annals of Mathematical Statistics*, **33**, 1187–1192.
- Tojeiro C, Louzada F, Roman M, and Borges P (2012). The complementary Weibull geometric distribution, *Journal of Statistical Computation and Simulation*, **84**, 1345–1362.

Received November 24, 2017; Revised March 27, 2018; Accepted April 30, 2018