



A bounded distribution derived from the shifted Gompertz law

P. Jodrá*

Dpto. de Métodos Estadísticos, EINA, Universidad de Zaragoza, 50018 Zaragoza, Spain

ARTICLE INFO

Article history:

Received 8 May 2018

Accepted 9 August 2018

Available online xxxxx

2000 AMS Classification:

60E05

62P10

33B30

Keywords:

Bounded distribution

Shifted Gompertz

Beta

Kumaraswamy

Power function

ABSTRACT

A two-parameter probability distribution with bounded support is derived from the shifted Gompertz distribution. It is shown that this model corresponds to the distribution of the minimum of a random number with shifted Poisson distribution of independent random variables having a common power function distribution. Some statistical properties are written in closed form, such as the moments and the quantile function. To this end, the incomplete gamma function and the Lambert W function play a central role. The shape of the failure rate function and the mean residual life are studied. Analytical expressions are also provided for the moments of the order statistics and the limit behavior of the extreme order statistics is established. Moreover, the members of the new family of distributions can be ordered in terms of the hazard rate order. The parameter estimation is carried out by the methods of maximum likelihood, least squares, weighted least squares and quantile least squares. The performance of these methods is assessed by means of a Monte Carlo simulation study. Two real data sets are used to illustrate the usefulness of the proposed distribution.

© 2018 The Author. Production and hosting by Elsevier B.V. on behalf of King Saud University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Over the last decades, there has been considerable interest in the development of new parametric probability distributions, which can be used in a wide range of practical applications. Almost all new proposals are distributions with unbounded support. However, there are many real situations in which the data take values in a bounded interval, such as percentages and proportions. The most popular two-parameter distributions with bounded support are the classical beta and Kumaraswamy laws (for the latter see Kumaraswamy, 1980, Jones, 2009). Other less known models are the transformed Leipnik distribution (Jorgensen, 1997, pp. 196–197), the Log–Lindley distribution (Gómez-Déniz et al., 2014; Jodrá and Jiménez-Gamero, 2016) and the standard two-sided power distribution (Van Dorp and Kotz, 2002). In the mathematical literature, there are proposals with more parameters such as the three-parameter generalized beta distribution (McDonald, 1984) and the reflected generalized Topp–Leone power series

distribution (Condino and Domma, 2017), the four-parameter Kumaraswamy Weibull distribution (Cordeiro et al., 2010), the exponentiated Kumaraswamy-power function distribution (Bursa and Ozel, 2017) and the two-sided generalized Kumaraswamy distribution (Korkmaz and Genç, 2017) and the five-parameter Kumaraswamy generalized gamma distribution (Pascoa et al., 2011), among others, but in these cases their handling is more complex due to the greater number of parameters. The usefulness of distributions with bounded support has been highlighted in Condino and Domma (2017), Cordeiro et al. (2010), Kotz and van Dorp (2004), Pascoa et al. (2011) and the references therein.

In this paper, our proposal is to provide a new two-parameter distribution for modelling data taking values in a bounded domain. Specifically, a new random variable defined on the interval $(0,1)$ is derived from the shifted Gompertz (SG) distribution. The SG law was introduced by Bemmaor (1994) as a model of adoption timing of a new product in a marketplace, showing its connection with the Bass diffusion model widely used in marketing research (see Bemmaor and Lee, 2002). The properties of the SG distribution were thoroughly studied in Jiménez and Jodrá (2009) and the parameter estimation in Jiménez (2014), Jukić and Marković (2017).

To start with, let Y be a random variable having a SG distribution with parameters $\alpha > 0$ and $\beta > 0$, whose cumulative distribution function (cdf) is given by

$$F_Y(y; \alpha, \beta) = (1 - e^{-\beta y}) \exp(-\alpha e^{-\beta y}), \quad y > 0. \quad (1)$$

* Address: María de Luna 3, 50018 Zaragoza, Spain.

E-mail address: pjodra@unizar.es

Peer review under responsibility of King Saud University.



Production and hosting by Elsevier

Now, we define the random variable $X = \exp(-Y)$, which has the following probability density function (pdf)

$$f(x; \alpha, \beta) = \beta [1 + \alpha(1 - x^\beta)] x^{\beta-1} \exp(-\alpha x^\beta), \quad 0 < x < 1, \quad (2)$$

where $\alpha > 0$ and $\beta > 0$ are two shape parameters. Fig. 1 represents (2) for several values of the parameters. The cdf of X is given by

$$F(x; \alpha, \beta) = 1 - (1 - x^\beta) \exp(-\alpha x^\beta), \quad 0 < x < 1. \quad (3)$$

Hereafter, the random variable with pdf (2) will be referred to as the Log-shifted Gompertz (LSG) distribution and will be denoted by $LSG(\alpha, \beta)$ assuming $\alpha > 0$ and $\beta > 0$.

It should be noted that the standard power function distribution with parameter $\beta > 0$ is obtained by setting $\alpha = 0$ in (2); in particular, the case $\alpha = 0$ and $\beta = 1$ corresponds to the uniform distribution. Since the power function distribution is a well-studied model, for the sake of simplicity the special case $\alpha = 0$ has been intentionally omitted in the current study.

Note also that the LSG law defined by (2) in the unit interval can be extended to any bounded domain in a straightforward manner, since a linear transformation $(b - a)X + a$ moves a random variable X defined on $(0, 1)$ to any other bounded support (a, b) , with $a < b$. Accordingly, there is no need to study such an extension further.

The remainder of this paper is organized as follows. In Section 2, some statistical properties of the LSG distribution are studied. First, a stochastic representation is given. Next, the moments are expressed in closed form in terms of the incomplete gamma function and the quantile function is written in terms of the Lambert W function. The shape of the failure rate function and the mean residual life are also studied. Analytical expressions are provided for the moments of the order statistics and the limit behavior of the extreme order statistics is established. It is also shown that the members of the new family of distributions can be ordered in terms of the hazard rate order. In Section 3, the parameter estimation is carried out by maximum likelihood, least squares, weighted least squares and quantile least squares. A Monte Carlo simulation study assesses the performance of these methods. Finally, in Section 4, two real data sets illustrate that the LSG model may provide a better fit than other two-parameter distributions with bounded support.

2. Statistical properties

2.1. Stochastic representation

The LSG distribution has been defined in (2) via an exponential transformation of the SG model. It is interesting to note that the LSG law can be obtained as the distribution of the minimum of a shifted Poisson random number of independent random variables having a common power function distribution. To be more precise, let N be a random variable having a Poisson distribution with mean $\alpha > 0$, so its probability mass function is given by

$$P(N = n) = e^{-\alpha} \frac{\alpha^n}{n!}, \quad n = 0, 1, 2, \dots$$

Let M be a shifted Poisson random variable defined by $M = N + 1$. Let $T_i, i = 1, 2, \dots$ be independent identically distributed random variables having T_i a standard power function distribution with parameter $\beta > 0$, that is, its cdf is $F_{T_i}(t; \beta) = t^\beta, 0 < t < 1$. Assume also that M is independent of T_i , for $i = 1, 2, \dots$

Proposition 2.1. *The random variable $T = \min\{T_1, T_2, \dots, T_M\}$ has a $LSG(\alpha, \beta)$ distribution.*

Proof. For any $m = 1, 2, \dots$ and $\beta > 0$, the conditional cdf of the random variable $T|M = m$ is given by $F_{T|M=m}(t; \beta) = 1 - (1 - t^\beta)^m, 0 < t < 1$. Therefore,

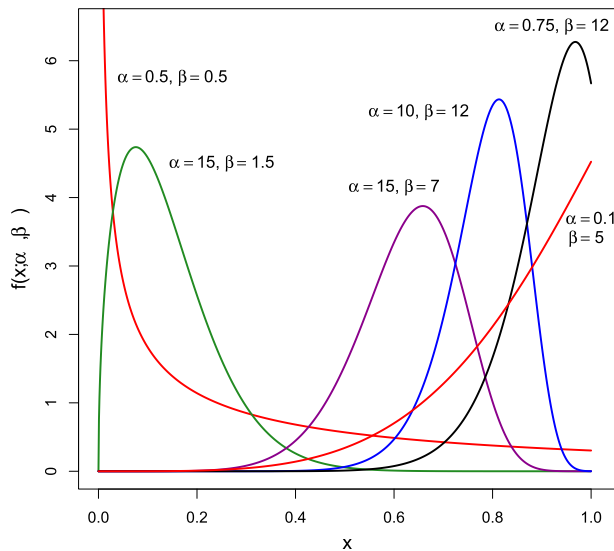


Fig. 1. $f(x; \alpha, \beta)$ for different values of α and β .

$$P(T \leq t, M = m) = [1 - (1 - t^\beta)^m] e^{-\alpha} \frac{\alpha^{m-1}}{(m-1)!}.$$

Hence, for any $\alpha > 0$ and $\beta > 0$ the marginal cdf of T is given by

$$F_T(t; \alpha, \beta) = \sum_{m=1}^{\infty} P(T \leq t, M = m) = 1 - (1 - t^\beta) \exp(-\alpha t^\beta), \quad 0 < t < 1,$$

which implies the result. \square

As contexts of potential application of the LSG distribution, note that in medical and veterinary studies a patient/animal may die due to different unknown causes such as infections. In this context, M denotes the number of causes and T_i denotes the lifetime in a bounded period associated with each cause i . In the described scenario, frequently only the minimum lifetime T among all causes will be observed. A similar scenario can be found in reliability, where a complex device may fail due to different unknown causes.

2.2. Shape and mode

The next result describes the different shapes of pdf (2) depending on the values of the parameters. The proof is omitted because it consists of routine calculations.

Proposition 2.2. *Let X be a random variable having a $LSG(\alpha, \beta)$ distribution.*

- (i) For any $0 < \beta \leq 1$ and $\alpha > 0, f(x; \alpha, \beta)$ is a decreasing function in x .
- (ii) For any $\beta > 1$ and $0 < \alpha \leq (\beta - 1)/(2\beta), f(x; \alpha, \beta)$ is an increasing function in x .
- (iii) For any $\beta > 1$ and $\alpha > (\beta - 1)/(2\beta), f(x; \alpha, \beta)$ has a mode given by

$$x = \left(\frac{\beta(\alpha + 3) - 1 - \sqrt{(\alpha\beta + 1)^2 + \beta(2\alpha\beta + 5\beta - 2)}}{2\alpha\beta} \right)^{1/\beta}.$$

Moreover, for any $\alpha > 0$

$$\lim_{x \rightarrow 0^+} f(x; \alpha, \beta) = \begin{cases} \infty & \text{if } \beta \in (0, 1), \\ 1 + \alpha & \text{if } \beta = 1, \\ 0 & \text{if } \beta > 1, \end{cases}$$

and $f(1; \alpha, \beta) = \beta e^{-\alpha}$ for any $\alpha > 0$ and $\beta > 0$.

From Proposition 2.2, a clear difference between the LSG law and the beta and Kumaraswamy distributions is the limit behavior at $x = 1$, since that limit can be equal to 0, 1 or ∞ in the beta and Kumaraswamy models.

2.3. Moments

The moments of the LSG distribution can be expressed analytically in terms of the lower and upper incomplete gamma functions (see Paris, 2010, pp. 174), which are defined, respectively, by

$$\gamma(a, z) = \int_0^z t^{a-1} e^{-t} dt, \quad \text{Re}(a) > 0, \tag{4}$$

and

$$\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t} dt.$$

Proposition 2.3. Let X be a random variable having a LSG(α, β) distribution. For $k = 1, 2, \dots$, the k -th moment of X is given by

$$E[X^k] = e^{-\alpha} + \alpha^{-\left(\frac{k}{\beta}+1\right)} \left(\alpha - \frac{k}{\beta}\right) \gamma\left(\frac{k}{\beta} + 1, \alpha\right). \tag{5}$$

Proof. By making the change of variable $u = \alpha x^\beta$ and taking into account (4),

$$\begin{aligned} E[X^k] &= \int_0^1 x^k f(x; \alpha, \beta) dx = \alpha^{-\left(\frac{k}{\beta}+1\right)} \int_0^\alpha u^{k/\beta} e^{-u} (1 + \alpha - u) du \\ &= \alpha^{-\left(\frac{k}{\beta}+1\right)} \left[(1 + \alpha) \gamma\left(\frac{k}{\beta} + 1, \alpha\right) - \gamma\left(\frac{k}{\beta} + 2, \alpha\right) \right]. \end{aligned}$$

The result follows by applying to the above equation the recurrence formula $\gamma(a, z) = (a - 1)\gamma(a - 1, z) - z^{a-1}e^{-z}$ (Paris, 2010, pp. 178). \square

Computer algebra systems usually implement the upper incomplete gamma function instead of the lower one, so it is also useful to express (5) in terms of the former.

Corollary 2.4. Let X be a random variable having a LSG(α, β) distribution. For $k = 1, 2, \dots$, the k -th moment of X is given by

$$E[X^k] = e^{-\alpha} + \alpha^{-\left(\frac{k}{\beta}+1\right)} \left(\alpha - \frac{k}{\beta}\right) \left[\Gamma\left(\frac{k}{\beta} + 1\right) - \Gamma\left(\frac{k}{\beta} + 1, \alpha\right) \right]. \tag{6}$$

Proof. The result follows from Eq. (5) together with the fact that $\gamma(a, z) = \Gamma(a) - \Gamma(a, z), a \neq 0, -1, -2, \dots$ (Paris, 2010, pp. 174). \square

From Corollary 2.4, the usual statistical measures involving $E[X^k]$ can be computed efficiently. Figs. 2 and 3 illustrate, respectively, the behavior of the mean and the standard deviation (denoted by σ) for different values of the parameters.

Figs. 4 and 5 display, respectively, the behavior of the asymmetry coefficient, $E[(X - E[X])^3]/\sigma^3$, and the kurtosis coefficient, $E[(X - E[X])^4]/\sigma^4 - 3$, for different values of the parameters. As can be seen, the LSG distribution has a wide range of values for both coefficients depending on the values of α and β , which suggests that the proposed model is rich enough to model real data.

2.4. Quantile function

An outstanding property of the LSG distribution is that its cdf is invertible. Specifically, the quantile function F^{-1} can be expressed in closed form in terms of the principal branch of the Lambert W function.

For the sake of completeness, recall that the Lambert W function is a multivalued complex function defined as the solution of the equation $W(z)e^{W(z)} = z$, where z is a complex number. This function has two real branches if z is a real number such that $z \geq -1/e$. The real branch taking on values in $[-1, \infty)$ (resp. $(-\infty, -1]$) is called the principal (resp. negative) branch and is denoted by W_0 (resp. W_{-1}). A review of this special function can be found in the seminal paper by Corless et al. (1996).

Proposition 2.5. The quantile function of the LSG(α, β) distribution is given by

$$F^{-1}(u; \alpha, \beta) = \left(1 - \frac{1}{\alpha} W_0(\alpha e^{\alpha(1-u)}) \right)^{1/\beta}, \quad 0 < u < 1. \tag{7}$$

Proof. For any $0 < u < 1$, we have to solve with respect to x the equation $F(x; \alpha, \beta) = u$, which is equivalent to solve $F_Y(-\log x; \alpha, \beta) = 1 - u$, with F_Y given by Eq. (1). The solution of the latter equation is $x = \exp(-F_Y^{-1}(1 - u; \alpha, \beta))$. Now, by taking

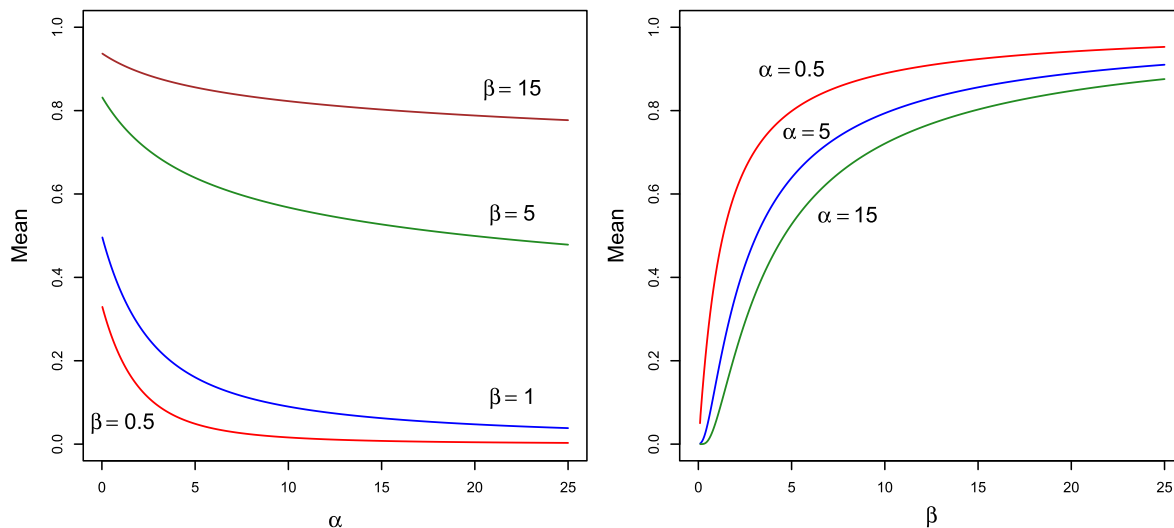


Fig. 2. Mean for several values of α and β .

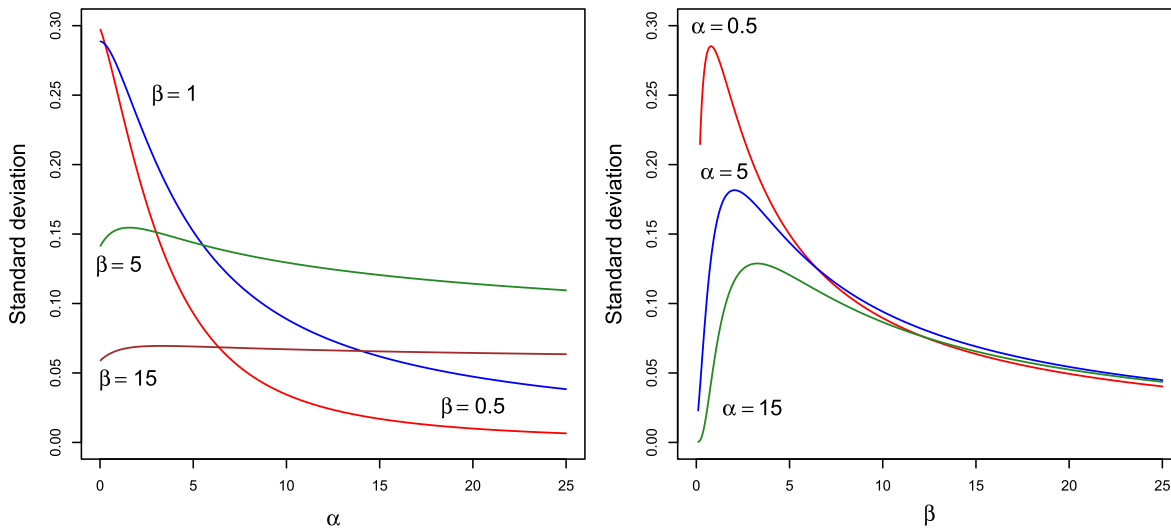


Fig. 3. Standard deviation for different values of α and β .

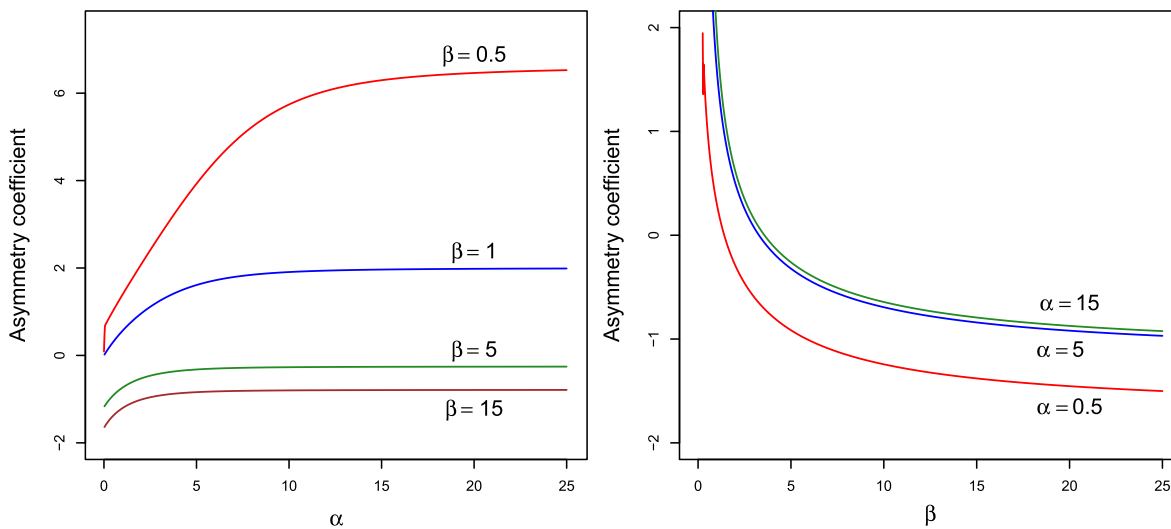


Fig. 4. Asymmetry coefficient for different values of α and β .

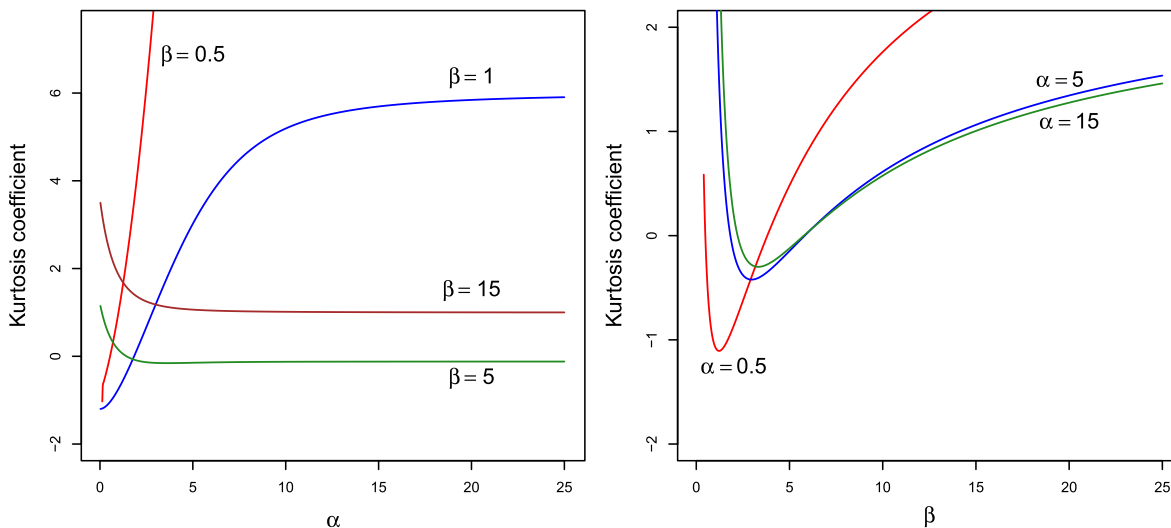


Fig. 5. Kurtosis coefficient for different values of α and β .

into account the closed-form expression for F_Y^{-1} provided in Jiménez and Jodrá (2009, Proposition 3.1), we get

$$F^{-1}(u; \alpha, \beta) = \exp\left(\frac{1}{\beta} \log\left(1 - \frac{1}{\alpha} W_0(\alpha e^\alpha(1-u))\right)\right),$$

which is equivalent to (7). □

For simulation purposes, Proposition 2.5 is very useful because pseudo-random data from the LSG model can be computer-generated from formula (7) in a straightforward manner by applying the inverse transform method (see, for example, Fishman (1996, Chapter 3)). More precisely, a pseudo-random datum from the LSG model is generated evaluating in formula (7) a pseudo-random datum from a uniform distribution defined on (0,1). In this regard, it should be remarked that the Lambert W function is available in computer algebra systems such as Maple (function LambertW), Mathematica (function ProductLog) and Matlab (function lambertw) and in programming languages such as R Development Core Team (2018) (functions lambertW0 and lambertWm1 for W_0 and W_{-1} , respectively, in the package lamW).

2.5. Failure rate and mean residual life

The failure (or hazard) rate and the mean residual life are essential in reliability and lifetime data analysis. The failure rate gives a description of the random variable in an infinitesimal interval of time whereas the mean residual life describes it in the whole remaining interval of time.

The failure rate function of the $LSG(\alpha, \beta)$ distribution is given by

$$h(x; \alpha, \beta) = \frac{f(x; \alpha, \beta)}{1 - F(x; \alpha, \beta)} = \frac{\beta[1 + \alpha(1 - x^\beta)]x^{\beta-1}}{1 - x^\beta}, \quad 0 < x < 1. \tag{8}$$

Fig. 6 displays $h(x; \alpha, \beta)$ for different values of the parameters. If this function increases then the item under consideration is usually degrading in some probabilistic sense, as the conditional probability of failure in the infinitesimal interval of time increases with time. If the failure rate decreases the reasoning is analogous. For the LSG law we establish the following.

Proposition 2.6. (i) If $\beta \geq 1$ then the LSG law has an increasing failure rate (IFR). (ii) If $\beta \in (0, 1)$ then there exists $x_0 = x_0(\alpha, \beta) \in (0, 1)$ so that $h(x; \alpha, \beta)$ is (strictly) decreasing when $x < x_0$ and (strictly) increasing when $x > x_0$.

Proof. The first derivative of h can be written as below

$$h'(x; \alpha, \beta) = \frac{\partial}{\partial x} h(x; \alpha, \beta) = \frac{\beta x^{\beta-2}}{(1-x^\beta)^2} [(\beta-1)(1+\alpha(1-x^\beta)^2) + x^\beta].$$

Part (i) follows because $h'(x; \alpha, \beta) \geq 0$ for any $\beta \geq 1, \alpha > 0$ and $x \in (0, 1)$. To prove part (ii), let $c(x; \alpha, \beta) = (\beta-1)[\alpha(1-x^\beta)^2 + 1] + x^\beta$. For any $\alpha > 0, \beta \in (0, 1)$ and $x \in (0, 1)$, note that $c(x)$ is an increasing function in x since

$$c'(x; \alpha, \beta) = \frac{\partial}{\partial x} c(x; \alpha, \beta) = \beta x^{\beta-1}[1 + 2\alpha(1-\beta)(1-x^\beta)] > 0.$$

Moreover, $\lim_{x \rightarrow 0} c(x; \alpha, \beta) = (\alpha+1)(\beta-1) < 0, \lim_{x \rightarrow 1} c(x; \alpha, \beta) = \beta \in (0, 1)$ and $\beta x^{\beta-2}/(1-x^\beta)^2 > 0$. As a consequence, there exists $x_0 = x_0(\alpha, \beta) \in (0, 1)$ such as $h'(x; \alpha, \beta) < 0$ for $x < x_0$ and $h'(x; \alpha, \beta) > 0$ for $x > x_0$, which implies the result. □

The different shapes of the failure rate function depending on the values of the parameters suggest a great flexibility of the LSG distribution to model real data.

On the other hand, the mean residual life function is defined by $r(t; \alpha, \beta) = E[X - t | X > t], 0 < t < 1$. It is known that if a random variable has an IFR then the mean residual life is decreasing. Therefore, as a consequence of Proposition 2.6, for $\beta \geq 1$ the LSG distribution has decreasing mean residual life. In the following result, the mean residual life is written in closed form.

Proposition 2.7. For any $0 < t < 1$, the mean residual life of the $LSG(\alpha, \beta)$ distribution is given by

$$r(t; \alpha, \beta) = \frac{\exp(\alpha t^\beta)}{\beta \alpha^{1/\beta} (1-t^\beta)} \times \left\{ \left(1 - \frac{1}{\alpha \beta}\right) \left[\Gamma\left(\frac{1}{\beta}, \alpha t^\beta\right) - \Gamma\left(\frac{1}{\beta}, \alpha\right) \right] + \alpha^{1/\beta-1} [e^{-\alpha} - t \exp(-\alpha t^\beta)] \right\}.$$

Proof. Let $S(x; \alpha, \beta) = 1 - F(x; \alpha, \beta)$. For any $0 < t < 1$, we have

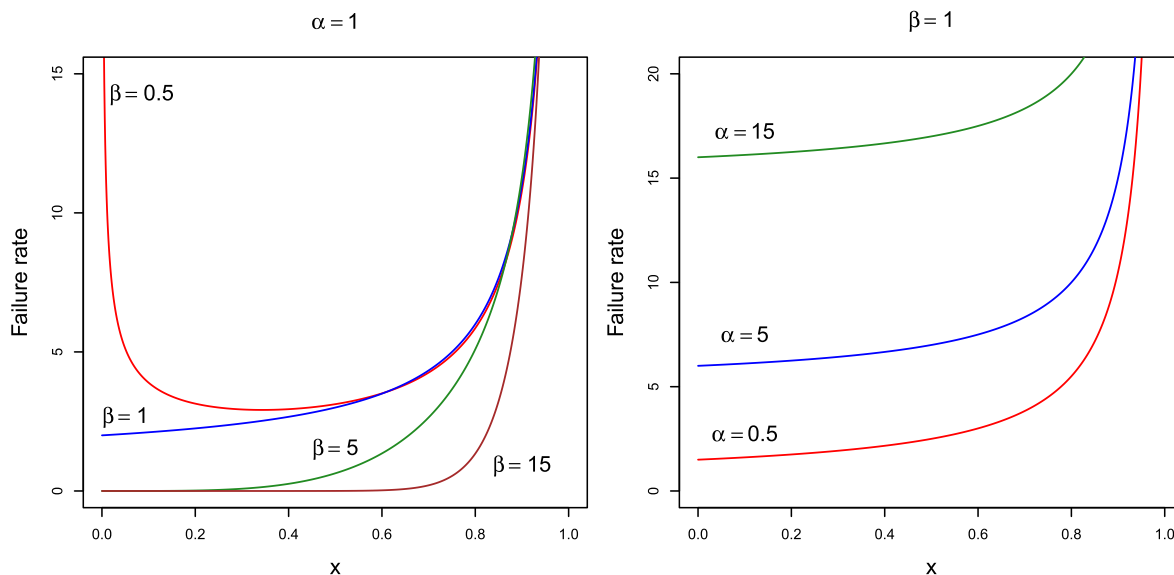


Fig. 6. Failure rate functions for $\alpha = 1$ (left) and $\beta = 1$ (right).

$$r(t; \alpha, \beta) = E[X - t | X > t] = \frac{1}{S(t; \alpha, \beta)} \int_t^1 S(x; \alpha, \beta) dx$$

$$= \frac{\exp(\alpha t^\beta)}{\beta \alpha^{1/\beta} (1 - t^\beta)} \left[\int_{\alpha t^\beta}^\alpha u^{1/\beta - 1} e^{-u} du - \frac{1}{\alpha} \int_{\alpha t^\beta}^\alpha u^{1/\beta} e^{-u} du \right]$$

$$= \frac{\exp(\alpha t^\beta)}{\beta \alpha^{1/\beta} (1 - t^\beta)} \left[\gamma\left(\frac{1}{\beta}, \alpha\right) - \gamma\left(\frac{1}{\beta}, \alpha t^\beta\right) - \frac{1}{\alpha} \gamma\left(\frac{1}{\beta} + 1, \alpha\right) + \frac{1}{\alpha} \gamma\left(\frac{1}{\beta} + 1, \alpha t^\beta\right) \right],$$

where we have made the change of variable $u = \exp(\alpha x^\beta)$ in the third equality and Eq. (4) was considered in the last one. Finally, the desired result is obtained by taking into account that $\gamma(a, z) = \Gamma(a) - \Gamma(a, z)$, $a \neq 0, -1, -2, \dots$, and $\Gamma(a + 1, z) = a\Gamma(a, z) + z^a e^{-z}$ (Paris, 2010, pp. 178). □

2.6. Order statistics

This section provides analytical expressions to compute the moments of the order statistics and also studies the limit behavior of the extreme order statistics.

Let X_1, \dots, X_n be n independent random variables having LSG (α, β) distribution. Let $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ be the order statistics obtained by arranging X_i , $i = 1, \dots, n$, in non-decreasing order of magnitude. In particular, the minimum $X_{1:n}$ and maximum $X_{n:n}$ are called the extreme order statistics. For any $n = 1, 2, \dots, r = 1, \dots, n$ and $k = 1, 2, \dots$, it is known that the k th moment of $X_{r:n}$ can be computed using the following formula (see, for example, Balakrishnan and Rao, 1998, pp. 7)

$$E[X_{r:n}^k] = r \binom{n}{r} \int_0^\infty x^k (F(x; \alpha, \beta))^{r-1} (1 - F(x; \alpha, \beta))^{n-r} dF(x; \alpha, \beta). \quad (9)$$

The next result gives an analytical expression for the moments of the smallest order statistic from the LSG law.

Proposition 2.8. For $k = 1, 2, \dots$, the k -th moment of the smallest order statistic from the LSG (α, β) distribution is given by

$$E[X_{1:n}^k] = \frac{1}{\alpha(n\alpha)^{k/\beta}} \sum_{j=0}^{n-1} \binom{n-1}{j} \frac{(-1)^j}{(n\alpha)^j} \left[(1 + \alpha) - \frac{1}{n} \left(\frac{k}{\beta} + j + 1 \right) \right] \times \gamma\left(\frac{k}{\beta} + j + 1, n\alpha\right).$$

Proof. The result is derived from formula (9) setting $n = 1$, by applying the change of variable $u = \alpha x^\beta$, the binomial theorem, the recurrence formula $\gamma(a, z) = (a - 1)\gamma(a - 1, z) - z^{a-1} e^{-z}$ and also that $\sum_{j=0}^{n-1} (-1)^j \binom{n-1}{j} = 0$ (the details are omitted here). □

The expression of $E[X_{1:n}^k]$ in Proposition 2.8 can be used to compute $E[X_{r:n}^k]$, for $r = 2, \dots, n$ and $k = 1, 2, \dots$, avoiding the use of Eq. (9). With this aim, it can be used the recurrence formula (see Balakrishnan and Rao, 1998, pp. 156)

$$E[X_{r:n}^k] = \sum_{j=n-r+1}^n (-1)^{j-(n-r+1)} \binom{n}{j} \binom{j-1}{n-r} E[X_{1:j}^k], \quad r = 2, \dots, n.$$

Next, the limit behaviour of the extreme order statistic is established. First, recall that the cdfs of $X_{n:n}$ and $X_{1:n}$ are given, respectively, by $F_{X_{n:n}}(x; \alpha, \beta) = F^n(x; \alpha, \beta)$ and $F_{X_{1:n}}(x) = 1 - (1 - F(x; \alpha, \beta))^n$, with $F(x; \alpha, \beta)$ defined by (3). When n increases to ∞ , it is well-known that their limit distributions degenerate being needed to apply linear transformations to avoid degeneration. If there exist a non-degenerate cdf H and normalizing constants a_n and $b_n > 0$ such that

$$\lim_{n \rightarrow \infty} F^n(a_n + b_n x; \alpha, \beta) = H(x), \quad (10)$$

then F belongs to the domain of maximal attraction of H . Similarly, if there exist a non-degenerate cdf L and constants c_n and $d_n > 0$ such that

$$\lim_{n \rightarrow \infty} 1 - (1 - F(c_n + d_n x; \alpha, \beta))^n = L(x), \quad (11)$$

then F belongs to the domain of minimal attraction of L . Moreover, H and L follow either a Fréchet, Gumbel or Weibull distribution (for further details see, for example, Arnold et al., 1992, Chapter 8).

Proposition 2.9. The LSG (α, β) distribution belongs to:

- (i) The domain of maximal attraction of the Weibull distribution, with norming constants $a_n = 1$ and $b_n = 1 - F^{-1}(1 - n^{-1}; \alpha, \beta)$, with F^{-1} given by (7).
- (ii) The domain of minimal attraction of the Weibull distribution, with norming constants $c_n = F^{-1}(0; \alpha, \beta)$ and $d_n = F^{-1}(n^{-1}; \alpha, \beta) - F^{-1}(0; \alpha, \beta)$, with F^{-1} given by (7).

Proof. Part (i) is a consequence of Arnold et al. (1992, Theorems 8.3.2 and 8.3.4) by taking into account that

$$\lim_{t \rightarrow 0^+} \frac{1 - F(F^{-1}(1; \alpha, \beta) - tx; \alpha, \beta)}{1 - F(F^{-1}(1; \alpha, \beta) - t; \alpha, \beta)} = x,$$

where $F^{-1}(1; \alpha, \beta) = 1$ from (7) (details of the calculations are omitted here). Similarly, part (ii) is a consequence of Arnold et al. (1992, Theorems 8.3.6) by taking into account that

$$\lim_{t \rightarrow 0^+} \frac{F(F^{-1}(0; \alpha, \beta) + tx; \alpha, \beta)}{F(F^{-1}(0; \alpha, \beta) + t; \alpha, \beta)} = x^\beta.$$

The proof is complete. □

2.7. Stochastic orderings

To conclude Section 2, it is shown that the members of the LSG family can be ordered in terms of the hazard rate order, which is stronger than the usual stochastic order and the mean residual life order. For the sake of completeness these definitions are given below (see Shaked and Shanthikumar, 2007 for more details).

Definition 2.10. Let X_1 and X_2 be two random variables with cdfs F_1 and F_2 , respectively. Let h_1 and h_2 be the hazard rate functions of X_1 and X_2 , respectively. Let r_1 and r_2 be the mean residual life functions of X_1 and X_2 , respectively.

- (i) X_1 is said to be smaller than X_2 in the hazard rate order, denoted by $X_1 \leq_{HR} X_2$, if $h_1(x) \geq h_2(x)$ for all x .
- (ii) X_1 is said to be stochastically smaller than X_2 , denoted by $X_1 \leq_{ST} X_2$, if $F_1(x) \geq F_2(x)$ for all x .
- (iii) X_1 is said to be smaller than X_2 in the mean residual life order, denoted by $X_1 \leq_{MR} X_2$, if $r_1(x) \leq r_2(x)$ for all x .

The hazard rate order has applications in reliability theory and life insurance for comparing remaining lifetimes. Given two random variables X_1 and X_2 representing times until death, $X_1 \leq_{HR} X_2$ means that items with remaining lifetime X_2 will tend to live longer than those with remaining lifetime X_1 . The LSG family is ordered according to the value of the parameters in the following way.

Proposition 2.11. Let X_1 and X_2 be two random variables having LSG(α_1, β) and LSG(α_2, β) distributions, respectively. If $\alpha_1 \geq \alpha_2$ then $X_1 \leq_{HR} X_2$.

Proof. From Eq. (8), $h(x; \alpha_1, \beta) - h(x; \alpha_2, \beta) = (\alpha_1 - \alpha_2)\beta x^{\beta-1} (1 - x^\beta)/(1 - x^\beta)$. The result follows since clearly $h(x; \alpha_1, \beta) - h(x; \alpha_2, \beta) \geq 0$ for any $\alpha_1 \geq \alpha_2$. \square

Corollary 2.12. Let X_1 and X_2 be two random variables having LSG(α_1, β) and LSG(α_2, β) distributions, respectively. If $\alpha_1 \geq \alpha_2$ then the following relations hold:

- (i) $X_1 \leq_{ST} X_2$,
- (ii) $X_1 \leq_{MR} X_2$,
- (iii) $E(X_1^k) \leq E(X_2^k)$ for all $k > 0$.

Proof. Part (i) follows from the fact that $X_1 \leq_{HR} X_2 \Rightarrow X_1 \leq_{ST} X_2$ (see Shaked and Shanthikumar, 2007, pp. 18). Part (ii) follows from the property $X_1 \leq_{HR} X_2 \Rightarrow X_1 \leq_{MR} X_2$ (see Shaked and Shanthikumar, 2007, pp. 83). Part (iii) is a consequence of part (i) and the fact that $X_1 \leq_{ST} X_2$ if and only if $E[g(X_1)] \leq E[g(X_2)]$ for all non-decreasing functions g (see Shaked and Shanthikumar, 2007, pp. 4). \square

In particular, Corollary 2.12 shows that, for fixed β , the mean of the LSG family decreases as α increases.

3. Parameter estimation

The estimation of parameters of the LSG distribution is considered in this section. Initially, the maximum likelihood (ML) method is applied and the performance of this method is assessed via a Monte Carlo simulation study. Asymptotic as well as bootstrap confidence intervals for the model parameters are also discussed. Next, methods based on least squares are applied and their performance evaluated via a simulation study.

3.1. Maximum likelihood method

Let X_1, \dots, X_n be a random sample of size n from a LSG(α, β) law with both parameters unknown and denote by x_1, \dots, x_n the observed values. From the likelihood function, $L(\alpha, \beta) = \prod_{i=1}^n f(x_i; \alpha, \beta)$, the log-likelihood function is given by

$$\log L(\alpha, \beta) = n \log \beta - \alpha \sum_{i=1}^n x_i^\beta + (\beta - 1) \sum_{i=1}^n \log x_i + \sum_{i=1}^n \log(1 + \alpha(1 - x_i^\beta)). \tag{12}$$

The ML estimates of α and β are the values $\hat{\alpha}$ and $\hat{\beta}$ that maximize $\log L(\alpha, \beta)$. The system of partial derivatives of $\log L(\alpha, \beta)$ with respect to each parameter set equal to zero is given by:

$$\begin{aligned} \frac{\partial}{\partial \alpha} \log L(\alpha, \beta) &= - \sum_{i=1}^n x_i^\beta + \sum_{i=1}^n \frac{1 - x_i^\beta}{1 + \alpha(1 - x_i^\beta)} = 0, \\ \frac{\partial}{\partial \beta} \log L(\alpha, \beta) &= \frac{n}{\beta} + \sum_{i=1}^n \log x_i - \alpha \sum_{i=1}^n x_i^\beta \log x_i - \sum_{i=1}^n \frac{\alpha x_i^\beta \log x_i}{1 + \alpha(1 - x_i^\beta)} = 0. \end{aligned} \tag{13}$$

3.1.1. Practical considerations and simulation results

It is clear that the system of Eqs. (13) does not have an explicit solution. Consequently, to obtain the ML estimates it is preferable to solve the following optimization problem

$$\begin{aligned} \max \quad & \log L(\alpha, \beta) \\ \text{s.t.} \quad & \alpha > 0 \\ & \beta > 0. \end{aligned} \tag{14}$$

Problem (14) has linear inequality constraints so it can be solved using, for example, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm. The BFGS algorithm requires the gradient function of $\log L(\alpha, \beta)$, which is obtained from Eq. (13), together with a starting point in the interior of the feasible region. To determine a suitable starting point, we propose to solve with respect to α and β the system of equations

$$F^{-1}(0.25; \alpha, \beta) = q_1, \quad F^{-1}(0.75; \alpha, \beta) = q_3, \tag{15}$$

where q_1 and q_3 denote the first and third sample quartiles, respectively, and F^{-1} is given by Eq. (7).

A Monte Carlo simulation study was carried out to evaluate the performance of the ML method. Let N be the number of random samples generated and n the size of each random sample. The following quantities were computed for the simulated estimates $\hat{\alpha}_j, j = 1, \dots, N$:

- (i) The mean: $\bar{\alpha} = (1/N) \sum_{j=1}^N \hat{\alpha}_j$.
- (ii) The bias: $\text{Bias}(\hat{\alpha}) = \bar{\alpha} - \alpha$.
- (iii) The mean-square error: $\text{MSE}(\hat{\alpha}) = (1/N) \sum_{j=1}^N (\hat{\alpha}_j - \alpha)^2$.

The quantities $\bar{\beta}, \text{Bias}(\hat{\beta})$ and $\text{MSE}(\hat{\beta})$ are analogously defined and were also computed.

We generated $N = 10,000$ random samples of sizes $n = 50, 100, 200, 500, 1000$ for different values of α and β . Pseudo-random data from the LSG distribution were computer-generated by using formula (7). Problem (14) was solved using the BFGS algorithm available in the function `constrOptim` in the R language (R Development Core Team, 2018). Table 1 shows some simulation results, where the mean, bias and MSE of the simulated estimates are reported together with the corresponding asymptotic variances denoted by $\text{Var}(\hat{\alpha})$ and $\text{Var}(\hat{\beta})$. The asymptotic variances were obtained from the diagonal elements of the inverse of the expected Fisher information matrix. In this respect, the expected Fisher information matrix is defined as $I_n(\alpha, \beta) = -E[H(\alpha, \beta)]$, where $H(\alpha, \beta)$ is the Hessian matrix of $\log L(\alpha, \beta)$ (see Lehmann and Casella, 1998, for further details). Specifically,

$$I_n(\alpha, \beta) = - \begin{bmatrix} I_{\alpha\alpha} & I_{\alpha\beta} \\ I_{\beta\alpha} & I_{\beta\beta} \end{bmatrix} = - \begin{bmatrix} E \left[\frac{\partial^2}{\partial \alpha^2} \log L(\alpha, \beta) \right] & E \left[\frac{\partial^2}{\partial \alpha \partial \beta} \log L(\alpha, \beta) \right] \\ E \left[\frac{\partial^2}{\partial \beta \partial \alpha} \log L(\alpha, \beta) \right] & E \left[\frac{\partial^2}{\partial \beta^2} \log L(\alpha, \beta) \right] \end{bmatrix},$$

where

$$\begin{aligned} I_{\alpha\alpha} &= -n \int_0^1 \frac{(1 - x^\beta)^2}{[1 + \alpha(1 - x^\beta)]^2} f(x; \alpha, \beta) dx, \\ I_{\alpha\beta} = I_{\beta\alpha} &= -n \int_0^1 x^\beta \log x f(x; \alpha, \beta) dx - n \int_0^1 \frac{x^\beta \log x}{1 + \alpha(1 - x^\beta)} f(x; \alpha, \beta) dx \\ &\quad + n \int_0^1 \frac{\alpha x^\beta (1 - x^\beta) \log x}{[1 + \alpha(1 - x^\beta)]^2} f(x; \alpha, \beta) dx, \\ I_{\beta\beta} &= -\frac{n}{\beta^2} - n \int_0^1 \alpha x^\beta (\log x)^2 f(x; \alpha, \beta) dx - n \int_0^1 \frac{\alpha x^\beta (\log x)^2}{1 + \alpha(1 - x^\beta)} f(x; \alpha, \beta) dx \\ &\quad - n \int_0^1 \frac{\alpha^2 x^{2\beta} (\log x)^2}{[1 + \alpha(1 - x^\beta)]^2} f(x; \alpha, \beta) dx. \end{aligned}$$

As far as we have been able to determine, the integral expressions defining the elements of $I_n(\alpha, \beta)$ cannot be written in closed form

Table 1
ML estimates.

	$\alpha = 0.5$				$\beta = 0.5$				$\alpha = 0.5$				$\beta = 5.0$			
	$\hat{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Var($\hat{\alpha}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Var($\hat{\beta}$)	$\hat{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Var($\hat{\alpha}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Var($\hat{\beta}$)
$n = 50$	0.635	0.1356	0.2647	0.2595	0.524	0.0242	0.0109	0.0109	0.637	0.1372	0.2704	0.2595	5.245	0.2455	1.0987	1.0984
$n = 100$	0.558	0.0584	0.1306	0.1297	0.510	0.0107	0.0054	0.0054	0.556	0.0569	0.1317	0.1297	5.103	0.1032	0.5498	0.5492
$n = 200$	0.524	0.0241	0.0654	0.0648	0.504	0.0043	0.0027	0.0027	0.529	0.0296	0.0671	0.0648	5.045	0.0455	0.2755	0.2746
$n = 500$	0.509	0.0091	0.0263	0.0259	0.501	0.0018	0.0010	0.0010	0.508	0.0087	0.0262	0.0259	5.013	0.0133	0.1119	0.1098
$n = 1000$	0.503	0.0030	0.0129	0.0129	0.500	0.0005	0.0005	0.0005	0.504	0.0042	0.0130	0.0129	5.008	0.0081	0.0559	0.0549

	$\alpha = 5.0$				$\beta = 0.5$				$\alpha = 10.0$				$\beta = 10.0$			
	$\hat{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Var($\hat{\alpha}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Var($\hat{\beta}$)	$\hat{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	Var($\hat{\alpha}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	Var($\hat{\beta}$)
$n = 50$	5.505	0.5058	3.3367	2.0388	0.513	0.0136	0.0042	0.0036	11.415	1.4156	17.9270	8.9093	10.310	0.3105	1.5630	1.2876
$n = 100$	5.240	0.2403	1.3001	1.0194	0.506	0.0069	0.0020	0.0018	10.619	0.6197	6.0600	4.4546	10.145	0.1454	0.7034	0.6438
$n = 200$	5.105	0.1056	0.5638	0.5097	0.502	0.0026	0.0009	0.0009	10.293	0.2934	2.6612	2.2273	10.065	0.0655	0.3399	0.3219
$n = 500$	5.049	0.0492	0.2140	0.2038	0.501	0.0012	0.0003	0.0003	10.132	0.1327	0.9432	0.8909	10.029	0.0290	0.1289	0.1287
$n = 1000$	5.029	0.0291	0.1061	0.1019	0.500	0.0008	0.0001	0.0001	10.052	0.0523	0.4590	0.4454	10.011	0.0117	0.0643	0.0643

and, accordingly, for the particular values of α and β in Table 1 the elements $I_{\alpha\alpha}$, $I_{\alpha\beta}$ and $I_{\beta\beta}$ were calculated by numerical integration. The diagonal elements of $I_n^{-1}(\alpha, \beta)$ provided the variances in Table 1. Note that, alternatively, the asymptotic variances can also be approximated from the diagonal elements of the inverse of the observed Fisher information matrix. The observed information matrix will be provided in the next subsection.

Looking at Table 1, it can be noted that the ML method tended to slightly overestimate the value of both parameters. In spite of this fact, the ML method provided acceptable estimates of the parameters. As expected, the bias and MSE decrease as n increases and the values of MSE and variance are, in general, close.

3.1.2. Interval estimation

Based on the asymptotic normal approximation for $(\hat{\alpha}, \hat{\beta})$, interval estimation (similarly hypothesis tests) can be easily performed from the observed information matrix $J_n(\alpha, \beta)$. This matrix is given by

$$J_n(\alpha, \beta) = - \begin{bmatrix} J_{\alpha\alpha} & J_{\alpha\beta} \\ J_{\beta\alpha} & J_{\beta\beta} \end{bmatrix} = - \begin{bmatrix} \frac{\partial^2}{\partial \alpha^2} \log L(\alpha, \beta) & \frac{\partial^2}{\partial \alpha \partial \beta} \log L(\alpha, \beta) \\ \frac{\partial^2}{\partial \beta \partial \alpha} \log L(\alpha, \beta) & \frac{\partial^2}{\partial \beta^2} \log L(\alpha, \beta) \end{bmatrix},$$

where

$$J_{\alpha\alpha} = - \sum_{i=1}^n \frac{(1 - x_i^\beta)^2}{[1 + \alpha(1 - x_i^\beta)]^2},$$

$$J_{\alpha\beta} = J_{\beta\alpha} = - \sum_{i=1}^n x_i^\beta \log x_i - \sum_{i=1}^n \frac{x_i^\beta \log x_i}{1 + \alpha(1 - x_i^\beta)} + \sum_{i=1}^n \frac{\alpha x_i^\beta (1 - x_i^\beta) \log x_i}{[1 + \alpha(1 - x_i^\beta)]^2},$$

$$J_{\beta\beta} = - \frac{n}{\beta^2} - \sum_{i=1}^n \alpha x_i^\beta (\log x_i)^2 - \sum_{i=1}^n \frac{\alpha x_i^\beta (\log x_i)^2}{1 + \alpha(1 - x_i^\beta)} - \sum_{i=1}^n \frac{\alpha^2 x_i^{2\beta} (\log x_i)^2}{[1 + \alpha(1 - x_i^\beta)]^2}.$$

The observed covariance matrix is the inverse of $J_n(\alpha, \beta)$. $J_n^{-1}(\alpha, \beta)$, and the diagonal elements of $J_n^{-1}(\hat{\alpha}, \hat{\beta})$ are the variances of $\hat{\alpha}$ and $\hat{\beta}$, which we denote by $\widehat{\text{var}}(\hat{\alpha})$ and $\widehat{\text{var}}(\hat{\beta})$, respectively. Then, the asymptotic $(1 - \delta)100\%$ confidence intervals for α and β are $\hat{\alpha} \pm z_{\delta/2} \sqrt{\widehat{\text{var}}(\hat{\alpha})}$ and $\hat{\beta} \pm z_{\delta/2} \sqrt{\widehat{\text{var}}(\hat{\beta})}$, respectively, where $z_{\delta/2}$ stands for the upper percentile $\delta/2$ of the standard normal distribution.

It is well-known that the confidence intervals based on the asymptotic normal distribution of the ML estimates may not be suitable for samples of small size. To determine in practice for what values of n is observed that normal behavior, we generated 10,000 samples of different sizes n for different values of the

parameters. From our simulation experiments, we observed that the approximate normality of the ML estimates requires sample sizes n greater than 1000. This fact was checked by means of histograms, Q-Q plots and normality tests such as Kolmogorov–Smirnov and Anderson–Darling (the details are omitted here).

Accordingly, for samples of size 1000 or smaller we recommend constructing the confidence intervals using other techniques, such as parametric bootstrap. Table 2 shows the coverage probability (CP) and average width (AW) of 95% percentile bootstrap confidence intervals for each parameter. These simulation results were obtained by generating 10,000 samples of different sizes n and considering 1000 bootstrap samples in each case. As can be seen, as the sample size n increases the coverage probabilities are close to the nominal level of 95% and the average widths decrease.

3.2. Estimation methods based on least squares

This Section describes the estimation of parameters using the methods of least squares (LS), weighted least squares (WLS) and quantile least squares (QLS). Next, a Monte Carlo simulation study evaluates the performance of these methods.

Before going further, some notation is introduced. Let $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ be the order statistics of a random sample of size n from a LSG(α, β) distribution with both parameters unknown. Denote by $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$ the ordered observed data. It is well-known that the empirical distribution function is considered an estimator of $F(x_{(i)}; \alpha, \beta)$, which is defined as below

$$F_n(x_{(i)}; d) = \frac{i - d}{n - 2d + 1}, \quad i = 1, \dots, n, \tag{16}$$

for some real number $d, 0 \leq d < 1$. For a justification of choosing (16) together with suitable values of d (the most common values are $d = 0, 0.3, 0.375, 0.5$) we refer the reader to Barnett (1975), D’Agostino and Stephens (1986, Chapter 2).

3.2.1. Least squares method

The approach proposed by Bain (1974) was applied to obtain the LSG estimates of the parameters, which are the values $\hat{\alpha}$ and $\hat{\beta}$ that minimize the following least squares function

$$LS_d(\alpha, \beta) = \sum_{i=1}^n [\log(1 - F_n(x_{(i)}; d)) - \log(1 - F(x_{(i)}; \alpha, \beta))]^2, \tag{17}$$

where F and F_n are given by (3) and (16), respectively. The BFGS algorithm can be used to minimize (17), subject to the constraints $\alpha > 0$ and $\beta > 0$, and taking into account that the system of partial derivatives of $LS_d(\alpha, \beta)$ is the following:

Table 2
Percentile bootstrap 95% confidence intervals. CP and AW for α and β .

α	β	n	CP(α)	AW(α)	CP(β)	AW(β)
0.5	0.5	100	0.9539	1.3154	0.9464	0.2782
		200	0.9547	0.9559	0.9447	0.2001
		500	0.9469	0.6306	0.9467	0.1296
		1000	0.9517	0.4471	0.9508	0.0918
0.5	5.0	100	0.9540	1.3154	0.9464	2.7828
		200	0.9575	0.9576	0.9502	2.0012
		500	0.9477	0.6310	0.9446	1.2963
		1000	0.9505	0.4470	0.9510	0.9183
5.0	0.5	100	0.9335	4.5627	0.9433	0.1730
		200	0.9437	2.9945	0.9432	0.1200
		500	0.9464	1.8086	0.9450	0.0750
		1000	0.9482	1.2635	0.9527	0.0528
10.0	10.0	100	0.9326	10.1106	0.9396	3.2698
		200	0.9368	6.4646	0.9401	2.2634
		500	0.9431	3.8292	0.9465	1.4116
		1000	0.9483	2.653	0.9517	0.9938

$$\frac{\partial}{\partial \alpha} \text{LS}_d(\alpha, \beta) = 2 \sum_{i=1}^n [\log(1 - F_n(x_{(i)}; d)) - \log(1 - F(x_{(i)}; \alpha, \beta))] x_{(i)}^\beta,$$

$$\frac{\partial}{\partial \beta} \text{LS}_d(\alpha, \beta) = 2 \sum_{i=1}^n \left\{ [\log(1 - F_n(x_{(i)}; d)) - \log(1 - F(x_{(i)}; \alpha, \beta))] \right.$$

$$\left. \times x_{(i)}^\beta \left(\frac{\log x_i}{1 - x_{(i)}^\beta} + \alpha \log x_{(i)} \right) \right\}.$$

3.2.2. Weighted least squares method

The parameters of the LSG distribution were also estimated by the WLS method. Following Bickel and Doksum (2001, pp. 316–317), a weight $w_{i,d} = (1 - F_n(x_{(i)}; d))^2, i = 1, \dots, n$, was considered in (17) and then the weighted least squares function is given by

$$\text{WLS}_d(\alpha, \beta) = \sum_{i=1}^n w_{i,d} [\log(1 - F_n(x_{(i)}; d)) - \log(1 - F(x_{(i)}; \alpha, \beta))]^2.$$

The WLS estimates can be obtained following similar steps as in the LS method.

3.2.3. Quantile least squares method

Finally, the QLS method was applied. The QLS method minimizes the squares of the differences between the observed and the theoretical quantiles, so the QLS estimates are the values $\hat{\alpha}$ and $\hat{\beta}$ that minimize the quantile least squares function given below

$$\text{QLS}_d(\alpha, \beta) = \sum_{i=1}^n [x_{(i)} - F^{-1}(F_n(x_{(i)}; d); \alpha, \beta)]^2, \tag{18}$$

where F^{-1} is given by (7). The BFGS algorithm can be used to minimize (18), subject to the constraints $\alpha > 0$ and $\beta > 0$, with an initial point calculated as in Section 3.1.1 and taking into account that the system of partial derivatives of $\text{QLS}_d(\alpha, \beta)$ is given by:

$$\frac{\partial}{\partial \alpha} \text{QLS}_d(\alpha, \beta) = \frac{2}{\alpha \beta} \sum_{i=1}^n \left\{ \frac{W_0(\alpha e^\alpha (1 - F_n(x_{(i)}; d)))}{1 + W_0(\alpha e^\alpha (1 - F_n(x_{(i)}; d)))} \right.$$

$$\left. \times \left[x_{(i)} F^{-1}(F_n(x_{(i)}; d); \alpha, \beta) - \left(F^{-1}(F_n(x_{(i)}; d); \alpha, \beta) \right)^2 \right] \right\},$$

$$\frac{\partial}{\partial \beta} \text{QLS}_d(\alpha, \beta) = \frac{2}{\beta} \sum_{i=1}^n \left\{ F^{-1}(F_n(x_{(i)}; d); \alpha, \beta) \log \left(F^{-1}(F_n(x_{(i)}; d); \alpha, \beta) \right) \right.$$

$$\left. \times \left[x_{(i)} - F^{-1}(F_n(x_{(i)}; d); \alpha, \beta) \right] \right\},$$

where it has been considered that (see Corless et al., 1996)

$$\frac{\partial}{\partial x} W_0(x) = \frac{W_0(x)}{x(1 + W_0(x))}, \quad x \neq 0.$$

3.2.4. Simulation results

As in Section 3.1.1, a simulation study was carried out to evaluate the performance of the parameter estimates obtained by LS, WLS and QLS. To do this, we generated $N = 10,000$ random samples of sizes $n = 50, 100, 200, 500, 1000$, for different values of the parameters and took the values $d = 0, 0.3, 0.375, 0.5$. The best LS estimates were obtained for $d = 0$ and $d = 0.3$. These results are shown in Tables 3 and 4, respectively. As can be seen from both tables, the LS method tended to underestimate the value of the

Table 3
LS estimates (case $d = 0$).

	$\alpha = 0.5$			$\beta = 0.5$			$\alpha = 0.5$			$\beta = 5.0$		
	$\hat{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	$\hat{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)
$n = 50$	0.424	-0.0757	0.3552	0.509	0.0095	0.0232	0.413	-0.0863	0.3537	5.153	0.1536	6.7726
$n = 100$	0.403	-0.0965	0.2352	0.487	-0.0122	0.0138	0.402	-0.0972	0.2336	4.906	-0.0937	4.2687
$n = 200$	0.409	-0.0904	0.1605	0.481	-0.0180	0.0096	0.410	-0.0893	0.1592	4.810	-0.1895	0.9420
$n = 500$	0.432	-0.0677	0.0891	0.483	-0.0164	0.0056	0.427	-0.0725	0.0887	4.820	-0.1797	0.5573
$n = 1000$	0.454	-0.0452	0.0499	0.488	-0.0118	0.0032	0.453	-0.0463	0.0497	4.880	-0.1193	0.3191
	$\alpha = 5.0$			$\beta = 0.5$			$\alpha = 10.0$			$\beta = 10.0$		
	$\hat{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	$\hat{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)
$n = 50$	4.528	-0.4714	4.2637	0.465	-0.0347	0.0100	9.260	-0.7394	22.0944	9.424	-0.5751	3.1613
$n = 100$	4.533	-0.4670	2.0084	0.471	-0.0290	0.0055	9.236	-0.7633	9.4129	9.535	-0.4650	1.7096
$n = 200$	4.650	-0.3494	1.0350	0.478	-0.0211	0.0029	9.317	-0.6828	4.5549	9.627	-0.3727	0.9405
$n = 500$	4.789	-0.2108	0.4175	0.487	-0.0128	0.0011	9.539	-0.4603	1.8904	9.752	-0.2478	0.4033
$n = 1000$	4.859	-0.1404	0.2157	0.491	-0.0082	0.0006	9.663	-0.3364	0.9692	9.830	-0.1699	0.2065

Table 4
LS estimates (case $d = 0.3$).

	$\alpha = 0.5$			$\beta = 0.5$			$\alpha = 0.5$			$\beta = 5.0$		
	$\bar{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\bar{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	$\bar{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\bar{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)
$n = 50$	0.573	0.0738	0.4866	0.537	0.0370	0.0309	0.566	0.0667	0.4715	5.360	0.3606	2.9907
$n = 100$	0.507	0.0072	0.2775	0.512	0.0123	0.0170	0.525	0.0253	0.2863	5.151	0.1519	1.7450
$n = 200$	0.485	-0.0150	0.1740	0.500	0.0001	0.0107	0.485	-0.0150	0.1713	4.987	-0.0129	1.0666
$n = 500$	0.473	-0.0263	0.0888	0.494	-0.0060	0.0056	0.477	-0.0227	0.0888	4.947	-0.0522	0.5684
$n = 1000$	0.481	-0.0188	0.0487	0.495	-0.0048	0.0031	0.480	-0.0190	0.0499	4.952	-0.0477	0.3257

	$\alpha = 5.0$			$\beta = 0.5$			$\alpha = 10.0$			$\beta = 10.0$		
	$\bar{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\bar{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	$\bar{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\bar{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)
$n = 50$	4.997	-0.0024	4.9770	0.487	-0.0129	0.0097	10.494	0.4944	30.9899	9.863	-0.1370	3.1586
$n = 100$	4.894	-0.1053	2.1290	0.487	-0.0124	0.0050	9.907	-0.0928	10.0955	9.806	-0.1932	1.6139
$n = 200$	4.872	-0.1269	1.0204	0.489	-0.0103	0.0026	9.837	-0.1626	4.7350	9.848	-0.1513	0.8728
$n = 500$	4.914	-0.0853	0.3948	0.493	-0.0059	0.0010	9.817	-0.1826	1.8534	9.879	-0.1205	0.3687
$n = 1000$	4.939	-0.0604	0.2038	0.495	-0.0040	0.0005	9.864	-0.1350	0.9430	9.916	-0.0837	0.1917

parameters. On the other hand, the WLS results did not improve the ones obtained by the LS method, so we omit them for the sake of space. Finally, Table 5 reports the QLS estimates for $d = 0$ and, in general, this method tended to underestimate the value of the parameters. Clearly, the bias and the mean-square error decrease as n increases.

3.3. Simulation findings

From the simulation results in Sections 3.1.1 and 3.2.4, it must be highlighted that, in general, the performance of ML was better than that of the methods based on least squares because the ML method produced estimates with smaller MSE. It is interesting to

Table 5
QLS estimates (case $d = 0$).

	$\alpha = 0.5$			$\beta = 0.5$			$\alpha = 0.5$			$\beta = 5.0$		
	$\bar{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\bar{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	$\bar{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\bar{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)
$n = 50$	0.539	0.0395	0.3633	0.513	0.0133	0.0224	0.426	-0.0732	0.2012	4.426	-0.1765	0.8711
$n = 100$	0.513	0.0131	0.2191	0.503	0.0037	0.0138	0.457	-0.0422	0.1318	4.855	-0.1445	0.5553
$n = 200$	0.490	-0.0097	0.1333	0.496	-0.0032	0.0086	0.454	-0.0451	0.0741	4.877	-0.1226	0.3003
$n = 500$	0.485	-0.0145	0.0627	0.495	-0.0042	0.0042	0.478	-0.0215	0.0321	4.942	-0.0572	0.1258
$n = 1000$	0.493	-0.0068	0.0330	0.497	-0.0026	0.0022	0.488	-0.0118	0.0158	4.971	-0.0284	0.0626

	$\alpha = 5.0$			$\beta = 0.5$			$\alpha = 10.0$			$\beta = 10.0$		
	$\bar{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\bar{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	$\bar{\alpha}$	Bias($\hat{\alpha}$)	MSE($\hat{\alpha}$)	$\bar{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)
$n = 50$	4.634	-0.3653	7.5943	0.450	-0.0495	0.0281	9.511	-0.4889	13.2308	9.650	-0.3496	1.7177
$n = 100$	4.644	-0.3551	3.3963	0.463	-0.0369	0.01512	9.594	-0.4058	6.1707	9.765	-0.2349	0.8499
$n = 200$	4.702	-0.2974	1.5329	0.473	-0.0269	0.0075	9.699	-0.3009	3.0247	9.849	-0.1504	0.4260
$n = 500$	4.825	-0.1747	0.6278	0.484	-0.0153	0.0030	9.844	-0.1560	1.1999	9.930	-0.0691	0.1695
$n = 1000$	4.897	-0.1025	0.3190	0.491	-0.0089	0.0015	9.904	-0.0955	0.5859	9.958	-0.0413	0.0827

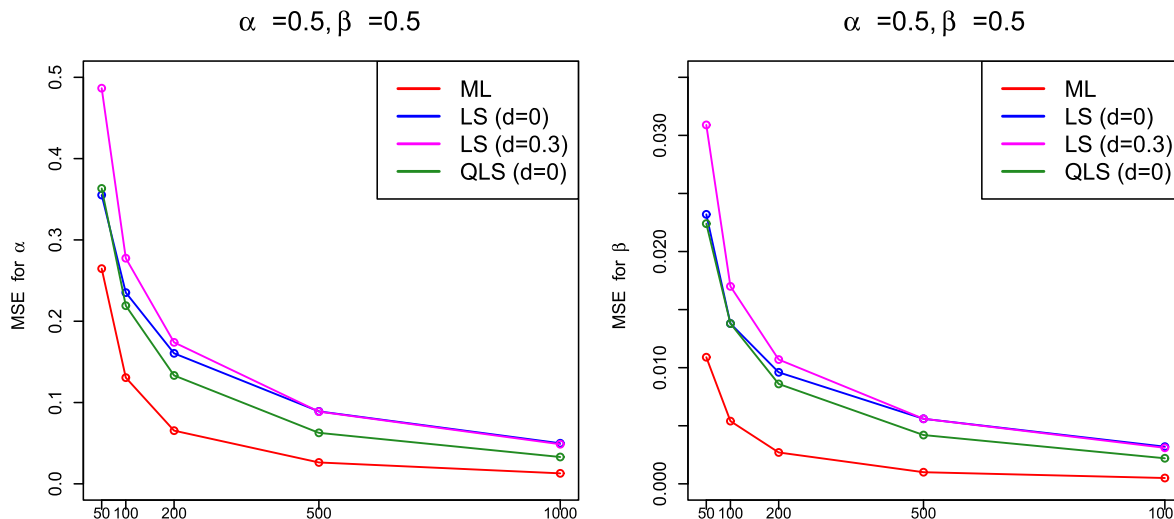


Fig. 7. MSE($\hat{\alpha}$) –left– and MSE($\hat{\beta}$) –right– for $\alpha = 0.5$ and $\beta = 0.5$.

note that the QLS method also provided good estimates if $\beta > 1$ but, as far as we have been able to determine, the QLS method did not improve the results obtained by ML. Figs. 7–10 display

graphically the behavior of the MSE for the estimation methods under comparison for the values of α, β and n in Table 1 and Tables 3–5.

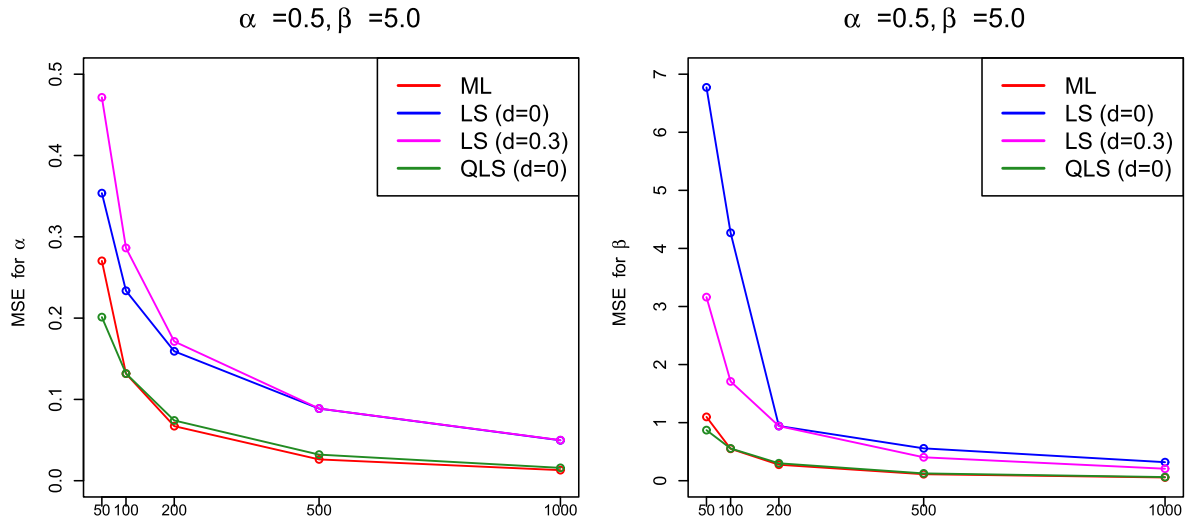


Fig. 8. $MSE(\hat{\alpha})$ –left– and $MSE(\hat{\beta})$ –right– for $\alpha = 0.5$ and $\beta = 5.0$.

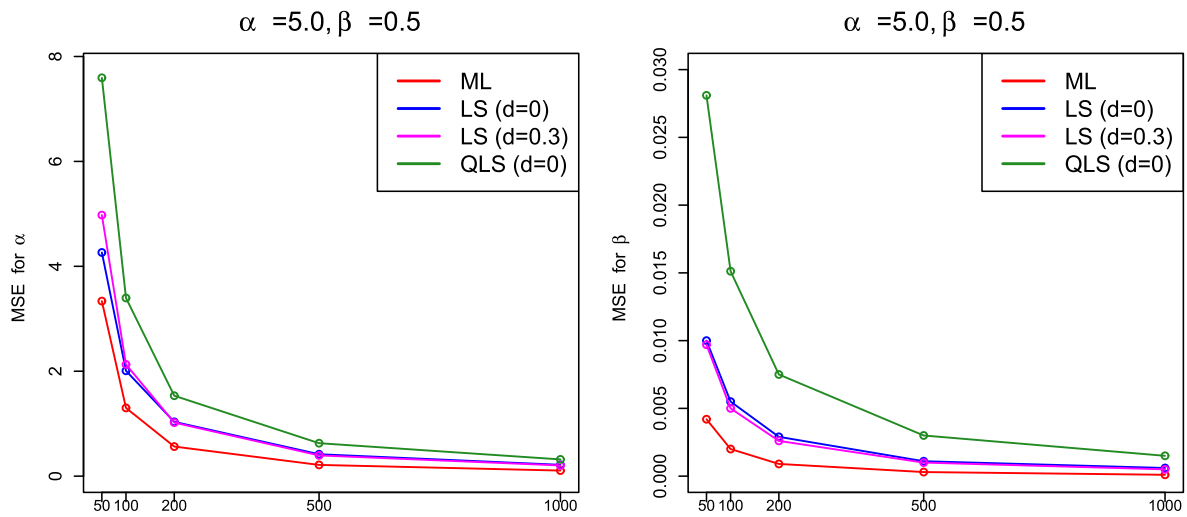


Fig. 9. $MSE(\hat{\alpha})$ –left– and $MSE(\hat{\beta})$ –right– for $\alpha = 5.0$ and $\beta = 0.5$.

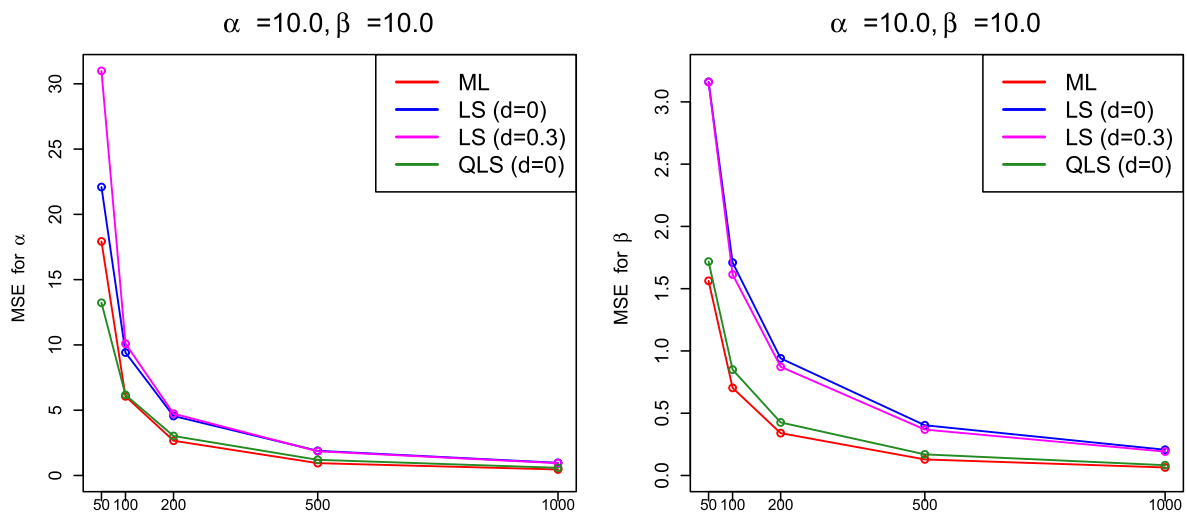


Fig. 10. $MSE(\hat{\alpha})$ –left– and $MSE(\hat{\beta})$ –right– for $\alpha = 10.0$ and $\beta = 10.0$.

Table 6
Antimicrobial resistance. Goodness-of-fit tests.

	W^2	U^2	A^2	D^+	D^-	D	V
p-value	0.7763	0.7561	0.8503	0.9693	0.4030	0.6499	0.8228

Table 7
Antimicrobial resistance. Models, ML estimates, AIC and BIC values.

Distributions with support (0, 1)	ML estimates	log L	AIC	BIC
LSG(α, β) $f(x; \alpha, \beta) = \beta x^{\beta-1} [1 + \alpha(1 - x^\beta)] e^{-\alpha x^\beta}$	$\hat{\alpha} = 2.5361, \hat{\beta} = 1.0693$	8.41	-12.83	-14.14
Beta(a, b) $f(x; a, b) = \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1}$	$\hat{a} = 0.8562, \hat{b} = 2.1635$	7.52	-11.05	-12.37
Kumaraswamy(a, b) $f(x; a, b) = abx^{a-1} (1-x^a)^{b-1}$	$\hat{a} = 0.8649, \hat{b} = 2.1116$	7.56	-11.13	-12.45
Log-Lindley(a, b) $f(x; a, b) = a[b + a(b-1) \log x] x^{a-1}$	$\hat{a} = 1.0918, \hat{b} = 0.0643$	7.87	-11.74	-13.06
Transformed Leipnik(μ, λ) $f(x; \mu, \lambda) = \frac{[x(1-x)]^{\frac{1}{2}}}{B(\frac{\mu+1}{2}, \frac{\lambda}{2})} \left[1 + \frac{(x-\mu)^2}{x(1-x)} \right]^{-\frac{\lambda}{2}}$	$\hat{\mu} = 0.0999, \hat{\lambda} = 2.5347$	7.80	-11.60	-12.92
Two-sided power(θ, v) $f(x; \theta, v) = \begin{cases} v(\frac{x}{\theta})^{v-1}, & 0 < x \leq \theta \\ v(\frac{1-x}{1-\theta})^{v-1}, & \theta \leq x < 1 \end{cases}$	$\hat{\theta} = 0.010, \hat{v} = 2.5168$	7.68	-11.37	-12.69

4. Data analysis

In this section, two real data sets are considered to illustrate that the LSG distribution may be an alternative to other two-parameter models with bounded support.

4.1. Data set 1: Antimicrobial resistance

The first data set is available in the report (European Centre for Disease Prevention and Control, 2013, pp. 178) from of the European Centre for Disease Prevention and Control, which is the main EU surveillance system for antimicrobial (antibiotic) resistance. The data represent the annual percentage of antimicrobial resistant isolates in Portugal in 2012. The $n = 24$ (sorted) values are the following: 1, 1, 3, 5, 8, 12, 14, 15, 15, 16, 19, 20, 20, 23, 26, 30, 32, 36, 39, 43, 54, 58, 59, 94.

The LSG law was fitted to the above data (divided by 100) and the ML estimates were $\hat{\alpha} = 2.5361$ and $\hat{\beta} = 1.0693$. The theoretical cumulative probabilities of the LSG model fit the empirical ones fairly well since the correlation coefficient between them is 0.9927. In order to evaluate the goodness of fit, the following tests were applied: Cramér von Mises statistic W^2 , Watson statistic U^2 , Anderson-Darling statistic A^2 , Kolmogorov-Smirnov statistics D^+, D^- and D , and Kuiper statistic V (see D’Agostino and Stephens, 1986, Chapter 4 for details). To get the corresponding p-values a parametric bootstrap was applied by generating 10,000 bootstrap samples (see Babu and Rao, 2004). The results are shown in Table 6.

The LSG fitting was compared to the ones provided by other two-parameter laws used to model data in the unit interval, specifically the distributions beta, Kumaraswamy, Log-Lindley (with the reparametrization given in Jodrá and Jiménez-Gamero, 2016), transformed Leipnik and standard two-sided power. The aforesaid models were compared through the Akaike information criterion (AIC) and the Bayesian information criterion (BIC), namely, $AIC = 2r - 2 \log L$ and $BIC = -2 \log L + r[\log n - \log(2\pi)]$, where r is the number of parameters, L denotes the maximized value of the likelihood function and n is the sample size. As it is well-known, the model with lower values of AIC and/or BIC is preferred. Table 7 reports the corresponding results and Fig. 11 shows the

Antimicrobial resistance: empirical cdf and fitted cdfs

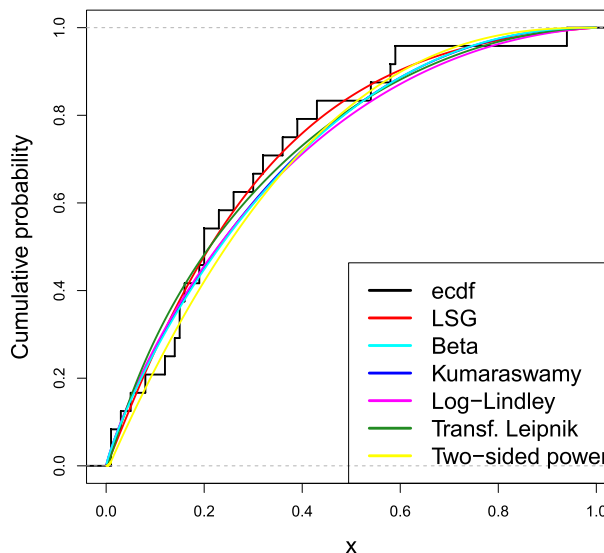


Fig. 11. Empirical cdf and fitted cdfs for the antimicrobial resistance data.

empirical cdf together with the fitted distributions. The LSG model provided a suitable fit with the smallest AIC and BIC values.

4.2. Data set 2: French speakers

The second data set corresponds to percentages of French speakers in 88 countries in 2014¹. The $n = 88$ (sorted) values are the following: 0.1, 0.1, 0.3, 0.4, 0.7, 0.7, 0.8, 0.8, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 6, 6, 6, 6, 7, 7, 8, 8, 9, 10, 10, 10, 11, 11, 11, 11, 12, 13, 13, 13, 13, 14, 15, 15, 16, 17, 20, 20, 21, 22, 23, 24, 25, 29, 29, 29, 29, 31, 31, 33, 34, 35, 38, 39, 40, 42, 47, 50, 53, 54, 58, 61, 66, 70, 72, 73, 79, 96, 97.

¹ The data set is available from the website: https://en.wikipedia.org/wiki/Geographical_distribution_of_French_speakers.

Table 8
French speakers. Goodness-of-fit tests.

	W^2	U^2	A^2	D^+	D^-	D	V
p-value	0.5574	0.5911	0.2744	0.1329	0.8670	0.2363	0.4294

Table 9
French speakers. Models, ML estimates, AIC and BIC values.

Distributions with support (0, 1)	ML estimates	log L	AIC	BIC
LSG(α, β)	$\hat{\alpha} = 1.9347, \hat{\beta} = 0.7052$	57.69	-111.38	-110.11
Beta(a, b)	$\hat{a} = 0.5381, \hat{b} = 1.8577$	54.07	-104.14	-102.86
Kumaraswamy(a, b)	$\hat{a} = 0.5765, \hat{b} = 1.7958$	54.89	-105.78	-104.50
Log-Lindley(a, b)	$\hat{a} = 0.7754, \hat{b} = 0.1099$	56.89	-109.78	-108.50
Transformed Leipnik(μ, λ)	$\hat{\mu} = 0.0232, \hat{\lambda} = 2.7316$	56.36	-108.72	-107.45
Two-sided power(θ, ν)	$\hat{\theta} = 0.0010, \hat{\nu} = 3.0952$	39.85	-75.71	-74.43

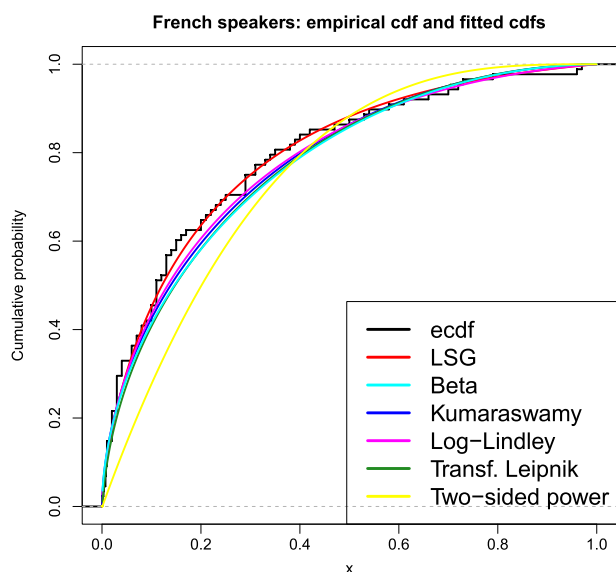


Fig. 12. Empirical cdf and fitted cdfs for the French speakers data.

The LSG distribution was fitted to the data set (divided by 100) and the ML estimates were $\hat{\alpha} = 1.9347$ and $\hat{\beta} = 0.7052$. The correlation coefficient between the theoretical cumulative probabilities of the LSG model and the empirical ones is 0.9967. As in Section 4.1, the same goodness-of-fit tests were applied and the p -values are shown in Table 8.

The LSG fitting was compared to the ones provided by the distributions beta, Kumaraswamy, Log-Lindley, transformed Leipnik and standard two-sided power. Table 9 shows the results and the LSG model yielded the smallest AIC and BIC values. Fig. 12 also shows the empirical cdf together with the fitted distributions. Overall, it can be concluded that the LSG model provided a good fit.

5. Conclusions

A two-parameter probability distribution defined on (0,1) is derived from the shifted Gompertz law. The proposed model also arises as the probability distribution of the minimum of a shifted Poisson random number of independent random variables having a common power function distribution. The new family of distributions has tractable properties. Analytical expressions are provided for the moments, quantile function and moments of the order statistics. The limit behavior of the extreme order statistics is also established. The parameters can be easily estimated by the method of maximum likelihood, which provided better results than other procedures based on least squares.

Acknowledgements

The author thanks the anonymous referees for their constructive comments, which led to a substantial improvement of the paper. Research in this paper has been partially funded by Diputación General de Aragón –Grupo E24-17R– and ERDF funds.

References

- Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., 1992. A First Course in Order Statistics. Wiley & Sons Inc, New York.
- Babu, G.J., Rao, C.R., 2004. Goodness-of-fit tests when parameters are estimated. *Sankhyā* 66 (1), 63–74.
- Bain, L.J., 1974. Analysis for the linear failure-rate life-testing distribution. *Technometrics* 16, 551–559.
- Balakrishnan, N., Rao, C.R. (Eds.), 1998. Order Statistics: Theory and Methods. Handbook of Statistics, volume 16. Elsevier, Amsterdam.
- Barnett, V., 1975. Probability plotting methods and order statistics. *J. Roy. Statist. Soc. Ser. C Appl. Statist.* 24, 95–108.
- Bemmaor, A.C., 1994. Modelling the diffusion of new durable goods: word-of-mouth effect versus consumer heterogeneity. In: Laurent, G., Lilien, G.L., Pras, B. (Eds.), *Research Traditions in Marketing*. Kluwer, Boston, pp. 201–229.
- Bemmaor, A.C., Lee, J., 2002. The impact of heterogeneity and ill-conditioning on diffusion model parameter estimates. *Market. Sci.* 21 (2), 209–220.
- Bickel, P.J., Doksum, K.A., 2001. *Mathematical Statistics, Basic Ideas and Selected Topics*, 2nd edition., vol. 1. Prentice Hall, New Jersey.
- Bursa, N., Ozel, G., 2017. The exponentiated Kumaraswamy-power function distribution. *Hacet. J. Math. Stat.* 46 (2), 277–292.
- Condino, F., Domma, F., A new distribution function with bounded support: the reflected generalized Topp-Leone power series distribution, *Metron* 75 (1) 51–68, 2017.
- Cordeiro, G.M., Ortega, E.M.M., Nadarajah, S., 2010. The Kumaraswamy Weibull distribution with application to failure data. *J. Franklin Inst.* 347 (8), 1399–1429.
- Corless, R.M., Gonnet, G.H., Hare, D.E.G., Jeffrey, D.J., Knuth, D.E., 1996. On the Lambert W function. *Adv. Comput. Math.* 5 (4), 329–359.
- D'Agostino, R.B., Stephens, M.A. (Eds.), 1986. *Goodness-of-Fit-Techniques*. Marcel Dekker, New York.
- European Centre for Disease Prevention and Control, 2013. Antimicrobial resistance surveillance in Europe 2012, Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net). ECDC, Stockholm. URL <https://ecdc.europa.eu>.
- Fishman, G.S., 1996. *Monte Carlo: Concepts, Algorithms and Applications*, Springer Series in Operations Research. Springer-Verlag, New York.
- Gómez-Déniz, E., Sordo, M.A., Calderín-Ojeda, E., 2014. The Log-Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insur. Math. Econ.* 54, 49–57.
- Jiménez, F., 2014. Estimation of parameters of the shifted Gompertz distribution using least squares, maximum likelihood and moments methods. *J. Comput. Appl. Math.* 255, 867–877.
- Jiménez, F., Jodrá, P., 2009. A note on the moments and computer generation of the shifted Gompertz distribution. *Commun. Stat.-Theory Methods* 38 (1–2), 75–89.
- Jodrá, P., Jiménez-Gamero, M.D., 2016. A note on the Log-Lindley distribution. *Insur. Math. Econ.* 71, 189–194.
- Jones, M.C., 2009. Kumaraswamy's distribution: a beta-type distribution with some tractability advantages. *Stat. Methodol.* 6 (1), 70–81.
- Jørgensen, B., 1997. *The Theory of Dispersion Models*. Chapman & Hall, London.
- Jukić, D., Marković, D., 2017. Nonlinear least squares estimation of the shifted Gompertz distribution. *Eur. J. Pure Appl. Math.* 10 (2), 157–166.
- Korkmaz, M.Ç., Genç, A.I., 2017. A new generalized two-sided class of distributions with an emphasis on two-sided generalized normal distribution. *Comm. Statist. Simul. Comput.* 46 (2), 1441–1460.

- Kotz, S., van Dorp, J.R., 2004. Beyond beta. Other continuous families of distributions with bounded support and applications. World Scientific Publishing Co., Pte. Ltd., Hackensack, NJ.
- Kumaraswamy, P., 1980. Generalized probability density function for double-bounded random processes. *J. Hydrol.* 46 (1–2), 79–88.
- Lehmann, E.L., Casella, G., 1998. *Theory of Point Estimation*. Springer Texts in Statistics, Springer-Verlag, New York.
- McDonald, J.B., 1984. Some generalized functions for the size distribution of income. *Econometrica* 52 (3), 647–663.
- Paris, R.B., 2010. Incomplete gamma and related functions. In: Olver, F.W.F., Lozier, D.W., Boisvert, R.F., Clark, C.W. (Eds.), *NIST Handbook of Mathematical Functions*, National Institute of Standards and Technology. Cambridge University Press, Cambridge, Washington, DC, pp. 173–192.
- Pascoa, M.A.R., Ortega, E.M.M., Cordeiro, G.M., 2011. The Kumaraswamy generalized gamma distribution with application in survival analysis. *Stat. Methodol.* 8, 411–433.
- R Development Core Team, 2018. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Shaked, M., Shanthikumar, J.G., 2007. *Stochastic Orders*. Springer-Verlag, New York.
- Van Dorp, J.R., Kotz, S., 2002. The standard two-sided power distribution and its properties: with applications in financial engineering. *Amer. Statist.* 56 (2), 90–99.