

ON A NEW SKEW GENERALIZED LOGISTIC DISTRIBUTION

Pushpa Narayan Rathie

Department of Statistics and Applied Mathematics,
Federal University of Ceara, Fortaleza, CE, Brazil.
E-mail: pushpanrathie@yahoo.com

Paulo Henrique Dourado da Silva and Gabriela Guimarães Olinto

Department of Statistics,
University of Brasilia, Brasilia, Brazil.

Dedicated to Prof. M.A. Pathan on his 75th birth anniversary

Abstract: In this paper, we propose a new skew generalized logistic distribution, which is an extension of the logistic distribution, the generalized logistic distribution, proposed by Rathie and Swamee [11] and the Kumaraswamy Logistic distribution, proposed by Santana et al. [13]. We study various mathematical properties of the new distribution, including explicit formulas for the probability density function, cumulative distribution function, survival, hazard and quantile functions. We also derive formulas for the moments, mean and median deviations, the Lorenz and Bonferroni curves. We also study some statistical properties, such as distributions of the minimum and the maximum order statistics. This distribution is used in two applications with real data: the first data is bimodal, and the second unimodal.

Keywords: Generalized logistic distribution, hazard rate function, Kumaraswamy distribution, maximum likelihood estimation, survival function.

Mathematics Subject Classification: 60E05, 62B15, 33C60, 60E10.

1. Introduction

The cumulative distribution function given by Kumaraswamy [6] is

$$G(x) = 1 - (1 - x^\alpha)^\beta, \quad 0 < x < 1, \quad (1)$$

where $\alpha, \beta > 0$. The probability density corresponding to (1) is given by

$$g(x) = \alpha\beta x^{\alpha-1}(1 - x^\alpha)^{\beta-1}, \quad 0 < x < 1, \quad (2)$$

which, as the beta distribution, may be unimodal, increasing, decreasing or constant, depending on the parameter values. Recently, Cordeiro and de Castro [3] defined the Kum-G distribution

$$G(t) = 1 - [1 - F^\alpha(t)]^\beta, \quad t, \alpha, \beta > 0, \quad (3)$$

where F denotes the cumulative distribution of a random variable. The parameters α and β control the asymmetry and the kurtosis of the distribution. The probability density function of (3) is given by

$$g(t) = \alpha\beta f(t)F(t)^{\alpha-1}[1 - F(t)^\alpha]^{\beta-1}, \quad (4)$$

where f is probability density function correspondent to F . According to Cordeiro et al. ([2], [4]), this generalization contains unimodal distributions and risk functions with bathtub forms.

Recently various distributions, following this construction, have been introduced just by modifying F in (3), for example Pascoa et al. [9] introduced and studied the Kumuraswamy generalized gamma distribution with applications in survival analysis. Santana et al. [13] introduced the Kumaraswamy log-logistic and logistic distributions. Cordeiro et al. [4] introduced and studied the Kumuraswamy generalized half-normal distribution. Yang [16] studied some statistical properties of the Kumaraswamy generalized inverse Weibull distribution. Cordeiro et al. [2] introduced and studied the Kumaraswamy Weibull distribution with application in failure data .

In this paper, we will use F as generalized logistic distribution and introduce the Kumaraswamy Rathie-Swamee generalized logistic distribution (KRS-GL). This distribution is a generalization of the Kumaraswamy logistic distribution, introduced by Santana et al. [13]. The KRS-GL has a bimodal nature for the density, which can be an important property in practical situations. We studied some mathematical properties such as, moments, quantile function, order statistics among other properties. To illustrate the versatility of the new distribution, we applied it to two sets of data. The results generated here, making $b = 0$, can be viewed as alternative results given by Santana et al. [13].

The paper is organized as follows. In Section 2, we introduce the KRS-GL, showing its cumulative distribution function and its probability density function, the survival and the hazard functions. In Section 3, we will give the mathematical expressions for the moments in terms of the H function (Mathai et al., [8]). In Section 4, we will give the mathematical expressions for the probability weight moment (PWM). Section 5 deals with the mathematical expressions for the Lorenz and Bonferroni curves.

In Section 6, we study the mathematical properties of the distributions of the minimum and the maximum respectively. In Section 7, we study some inferential properties of the distribution parameters beyond its log-likelihood function, which are of great importance in practical situations. The KRS-GL distribution is applied to two real data sets in Section 8. Finally, in Section 9, we point out some final considerations of the results obtained in this article.

Some important information about the results used in this article are detailed below.

1.1 Generalized logistic distribution

The generalized logistic distribution, RS-GL, introduced by Rathie and Swamee [11], has the following probability density function and cumulative distribution function

$$f(x) = \frac{[a + b(1 + p)|x|^p] \exp(-x(a + b|x|^p))}{(1 + \exp(-x(a + b|x|^p)))^2}, \quad (5)$$

and

$$F(x) = [1 + \exp(-x(a + b|x|^p))]^{-1}, \quad (6)$$

respectively. Here $-\infty < x < \infty$, $a, b, p \geq 0$ (Throughout this paper, it is assumed that all are not zeros simultaneously) and $F(x)$ is an invertible function (Rathie and Swamee [11]). For $a = 1.59413$, $b = 0.07443$ and $p = 1.939$, this distribution approximates very well to the standard normal distribution with a maximum error of 4×10^{-4} at $x = 0$ for the density function and 7.757×10^{-5} at $x = 2.81$ for the cumulative distribution function. Rathie et al. [12], show that, for some values of the parameters a , b and p , the distribution (5) closely approximates the Student- t distribution, and when the degrees of freedom of Student- t distribution increase, the distribution fits better. For example, for $a = 1.539966$, $b = 0.00848$ and $p = 0.687$, the mean square error between (5) and Student- t distribution with 7 degrees of freedom is 0.0002651130. Now, for $a = 1.589134$, $b = 0.0645$ and $p = 1.915$, the mean square error between (5) and Student- t with 60 degrees of freedom is 0.00008347, this is natural, since when the degrees of freedom of a Student- t distribution increases then the distribution is close to the standard normal distribution.

It is interesting to note that when the parameter a is near to zero, the distribution (5) has a bimodal shape, and for some values of a , b and p the shape can be trimodal. More detail about the generalized logistic distribution can be found in Rathie and Swamee [11], Swamee and Rathie [15] and Rathie [10].

1.2 Generalized Hypergeometric function

The H function is a generalization of the G function of Meijer and is defined by

(see Mathai et al. [8]; Springer, [14])

$$\begin{aligned} H_{p,q}^{m,n} \left[x \left| \begin{array}{c} (a_1, A_1), \dots, (a_n, A_n), (a_{n+1}, A_{n+1}), \dots, (a_p, A_p) \\ (b_1, B_1), \dots, (b_n, B_n), (b_{n+1}, B_{n+1}), \dots, (b_p, B_p) \end{array} \right. \right] = \\ = \frac{1}{2\pi i} \int_L \frac{\prod_{j=1}^m \Gamma(b_j - B_j s) \prod_{j=1}^n \Gamma(1 - a_j + A_j s)}{\prod_{j=m+1}^q \Gamma(1 - b_j + B_j s) \prod_{j=n+1}^p \Gamma(a_j - A_j s)} x^s ds. \end{aligned} \quad (7)$$

where

1. $i = \sqrt{-1}$;
2. $x (\neq 0)$ is a complex variable;
3. $x^s = \exp(s(\ln|x| + i \arg x))$;
4. An empty product is interpreted as one;
5. m, n, p and q are non-negative integers satisfying $0 \leq n \leq p, 0 \leq m \leq q$ (both n and m not equal to zero simultaneously);
6. A_j ($j = 1, \dots, p$) and B_j ($j = 1, \dots, q$) are assumed positive,
7. a_j ($j = 1, \dots, p$) and b_j ($j = 1, \dots, q$) are complex numbers such that none of the poles of $\Gamma(b_j + B_j s)$ ($j = 1, \dots, m$) coincide with the poles of $\Gamma(1 - a_j - A_j s)$ ($j = 1, \dots, n$) i.e. $A_k(b_h + v) \neq B_h(a_k - \lambda - 1)$ for $v, \lambda = 0, 1, \dots, h = 1, \dots, m$ and $k = 1, \dots, n$;
8. The contour L goes from $(-i\infty, +i\infty)$ such that the poles of $\Gamma(b_j + B_j s)$ ($j = 1, \dots, m$) are to the left of L and the poles of $\Gamma(1 - a_j - A_j s)$ ($j = 1, \dots, n$) are to the right of L .

The H function makes sense and is an analytic function of x in the following cases:

1. $\delta > 0, x \neq 0$ where

$$\delta = \sum_{j=1}^q B_j - \sum_{j=1}^p A_j. \quad (8)$$

2. $\delta = 0, 0 < |x| < D^{-1}$, where

$$D = \prod_{j=1}^p A_j^{A_j} / \prod_{j=1}^q B_j^{B_j}. \quad (9)$$

The value of the H function does not depend on the choice of L .

2. Kumaraswamy Rathie- Swamee generalized logistic distribution - KRS-GL

Let T be a random variable distributed as generalized logistic distribution (RS-GL), with parameters a, b, p . Thus, substituting (5) in (4), we obtain the KRS-GL distribution involving five parameters (a, b, p, α, β) with the probability density function as

$$g(t) = \alpha\beta \frac{[a + b(1+p)|t|^p] \exp(-t(a + b|t|^p))}{(1 + \exp(-t(a + b|t|^p)))^{1+\alpha}} \times \{1 - [1 + \exp(-t(a + b|t|^p))]\}^{\beta-1} \quad (10)$$

where $-\infty < x < \infty$, $a, b, p \geq 0$ and $\alpha, \beta > 0$. Now, replacing (6) in (3), we obtain the cumulative distribution function as

$$G(t) = 1 - [1 - [1 + \exp(-t(a + b|t|^p))]]^{-\alpha}]^{\beta} \quad (11)$$

The density KRS-GL is easily implemented and manipulated using any statistical software. Taking $b = 0$, we obtain the Kum-L distribution introduced by Santana et al. [13] as a special case. Taking $\alpha = 1$ and $\beta = 1$ we obtain the generalized logistic distribution as a special case.

We can write (11) in the following form

$$\frac{1}{[1 - (1 - G(t))^{1/\beta}]^{1/\alpha}} = 1 + \exp(-t(a + b|t|^p)) \quad (12)$$

or,

$$t = \frac{1}{a} \ln \left(\frac{G^*(t)}{1 - G^*(t)} \right) - \frac{b}{a} t |t|^p \quad (13)$$

where

$$G^*(t) = [1 - (1 - G(t))^{1/\beta}]^{1/\alpha} \quad (14)$$

Using the Lagrange expansion, we can write the quantile function of KRS-GL in the following form

$$t = \begin{cases} -\sum_{j=0}^{\infty} \frac{(-b/a)^j \Gamma(jp+j+1)}{j! \Gamma(jp+2)} \left[\frac{1}{a} \ln \left(\frac{1-G^*(t)}{G^*(t)} \right) \right]^{jp+1}, & G^*(t) \leq 0.5 \\ \sum_{j=0}^{\infty} \frac{(-b/a)^j \Gamma(jp+j+1)}{j! \Gamma(jp+2)} \left[\frac{1}{a} \ln \left(\frac{G^*(t)}{1-G^*(t)} \right) \right]^{jp+1}, & G^*(t) \geq 0.5. \end{cases} \quad (15)$$

The survival and hazard functions of (10) are given by

$$S(t) = [1 - ([1 + \exp(-t(a + b|t|^p))]^{-\alpha})]^\beta \quad t > 0, \quad (16)$$

and

$$\begin{aligned} h(t) = & \alpha\beta \frac{[a + b(1+p)|t|^p] \exp(-t(a + b|t|^p))}{(1 + \exp(-t(a + b|t|^p)))^{1+\alpha}} \times \\ & \times \{1 - ([1 + \exp(-t(a + b|t|^p))]^{-\alpha})\}^{-1} \end{aligned} \quad (17)$$

respectively.

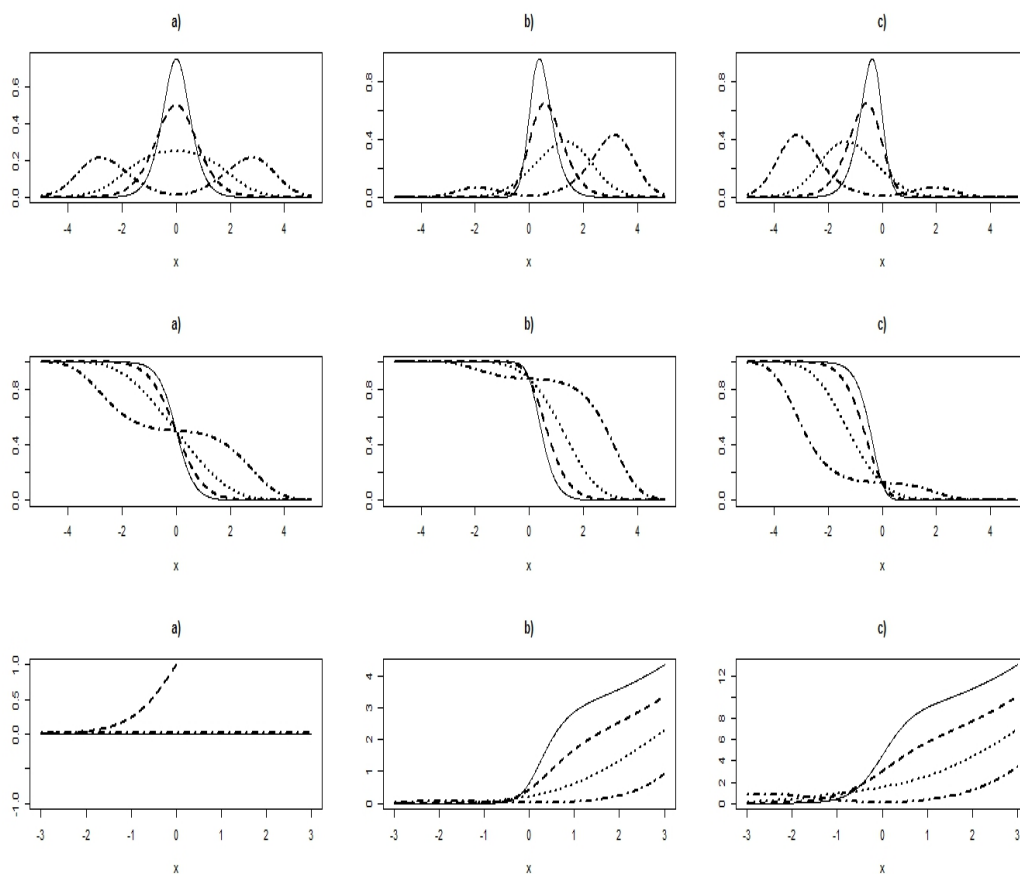


Figure 1 KRS-GL shape for density function, first line. Survival function, second line. Hazard function, third line.

Plots for probability density, survival and hazard functions for the KRS-GL distribution are illustrated in Figure 1. We can see the flexibility of the parameters α and β . For values of parameter a near to zero, the KRS-GL distribution has a bimodal shape, which can be very important in practical applications. It is interesting to note that Kum-L distribution, introduced by Santana et al. [13], is a particular case of KRS-GL.

Adding a location parameter $\mu \in \Re$ and a scale parameter $\sigma > 0$, for practical purposes, we can generalize (10) and (11) in the following form

$$g(t) = \frac{\alpha\beta}{\sigma} \frac{[a + b(1+p) \left(\frac{t-\mu}{\sigma}\right)^p] \exp\left(-\left(\frac{t-\mu}{\sigma}\right) \left(a + b \left|\frac{t-\mu}{\sigma}\right|^p\right)\right)}{\left(1 + \exp\left(-\left(\frac{t-\mu}{\sigma}\right) \left(a + b \left|\frac{t-\mu}{\sigma}\right|^p\right)\right)\right)^{1+\alpha}} \times \\ \times \left\{1 - \left(\left[1 + \exp\left(-\left(\frac{t-\mu}{\sigma}\right) \left(a + b \left|\frac{t-\mu}{\sigma}\right|^p\right)\right)\right]^{-\alpha}\right)\right\}^{\beta-1} \quad (18)$$

and

$$G(t) = 1 - \left[1 - \left[1 + \exp\left(-\left(\frac{t-\mu}{\sigma}\right) \left(a + b \left|\frac{t-\mu}{\sigma}\right|^p\right)\right)\right]^{-\alpha}\right]^{\beta} \quad (19)$$

In fact, (18) and (19) are transformed into identifiable expressions on defining new parameters A and B by $A = \frac{a}{\sigma}$ and $B = \frac{b}{\sigma^{p+1}}$.

3. Moments

If $T \sim \text{KRS-GL}(a, b, p, \alpha, \beta)$ with $\beta > 1$ real. Then, the n -th moments of T , when $a > 0$ and $b > 0$, are given by

$$E(T^n) = a\alpha\beta \sum_{j=0}^{\infty} (-1)^j \binom{\beta-1}{j} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(j+1)+1}{k} (-1)^n I_n(\alpha(j+1)+k) + \\ + a\alpha\beta \sum_{j=0}^{\infty} (-1)^j \binom{\beta-1}{j} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(j+1)+1}{k} I_n(1+k) + \\ + b(1+p)\alpha\beta \sum_{j=0}^{\infty} (-1)^j \binom{\beta-1}{j} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(j+1)+1}{k} (-1)^n I_{n+p}(\alpha(j+1)+k) + \\ + b(1+p)\alpha\beta \sum_{j=0}^{\infty} (-1)^j \binom{\beta-1}{j} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(j+1)+1}{k} I_{n+p}(1+k). \quad (20)$$

where

$$I_{\gamma}(\theta) = (a\theta)^{-\gamma-1} H_{1,1}^{1,1} \left[\frac{b\theta^{-p}}{a^{p+1}} \middle| \begin{matrix} (-\gamma, p+1) \\ (0, 1) \end{matrix} \right]. \quad (21)$$

When $a = 0$ and $b > 0$ we have

$$\begin{aligned} E(T^n) &= b(1+p)\alpha\beta \sum_{j=0}^{\infty} \binom{\beta-1}{j} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(j+1)+1}{k} (-1)^n I_{n+p}(\alpha(j+1)+k) + \\ &+ b(1+p)\alpha\beta \sum_{j=0}^{\infty} \binom{\beta-1}{j} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(j+1)+1}{k} I_{n+p}(1+k), \end{aligned} \quad (22)$$

where

$$I_{\gamma}(\theta) = \frac{\Gamma\left(\frac{\gamma}{p+1}\right)}{(p+1)[b\theta]^{\frac{\gamma}{p+1}}}. \quad (23)$$

Finally, when $a > 0$ and $b = 0$, we get

$$\begin{aligned} E(T^n) &= a\alpha\beta \sum_{j=0}^{\infty} \binom{\beta-1}{j} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(j+1)+1}{k} (-1)^n I_n(\alpha(j+1)+k) + \\ &+ a\alpha\beta \sum_{j=0}^{\infty} \binom{\beta-1}{j} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(j+1)+1}{k} I_n(1+k), \end{aligned} \quad (24)$$

where

$$I_{\gamma}(\theta) = \frac{\Gamma(\gamma)}{(a\theta)^{\gamma}}. \quad (25)$$

The results given above are valid for integer β by replacing $\sum_{j=0}^{\infty}$ by $\sum_{j=0}^{\beta-1}$. Alternative expressions for the moments, when $b = 0$, are given in Santana et al. [13].

4. Probability Weighted Moments

By definition, the (s, r) -th Probability Weighted Moment (PWM) of a random variable T following a distribution $f(t)$ is given by

$$\delta_{s,r} = E[T^s F(T)^r] = \int_{-\infty}^{\infty} t^s F(t)^r f(t) dt. \quad (26)$$

For the case where integer $\beta > 1$, assuming that $f(t)$ is given by a KRS-GL, then the PWM is given by

$$\delta_{s,r} = \alpha\beta \sum_{j=0}^r (-1)^j \binom{r}{j} \left[\sum_{i=0}^{\beta(j+1)-1} (-1)^i \binom{\beta(j+1)-1}{i} \tau_i(s) \right], \quad (27)$$

where,

$$\begin{aligned} \tau_i(s) = & a \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(i+1)+1}{k} [((-1)^s I_s(\alpha(i+1)+k) + I_s(1+k))] + \\ & + b(1+p) \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(i+1)+1}{k} [((-1)^s I_{s+p}(\alpha(i+1)+k) + I_{s+p}(1+k))]. \end{aligned} \quad (28)$$

Here $I_\gamma(\theta)$ is defined by the expression (21). For the case where $a = 0$ and $b > 0$, we have

$$\tau_i(s) = b(1+p) \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(i+1)+1}{k} \{(-1)^s I_{s+p}(\alpha(i+1)+k) + I_{s+p}(1+k)\}. \quad (29)$$

For the case where $b = 0$ and $a > 0$, we have

$$\tau_i(s) = a \sum_{k=0}^{\infty} (-1)^k \binom{\alpha(i+1)+1}{k} \{(-1)^s I_s(\alpha(i+1)+k) + I_s(1+k)\}. \quad (30)$$

For the case where $\beta > 1$ real, assuming that $f(t)$ is given by a KRS-GL, then the PWM is given by

$$\delta_{s,r} = \alpha\beta \sum_{j=0}^r (-1)^j \binom{r}{j} \left[\sum_{i=0}^{\infty} (-1)^i \binom{\beta(j+1)-1}{i} \tau_i(s) \right], \quad (31)$$

where τ_i defined by (28). The PWM will be important for calculating the k -th moments for the distribution of the minimum and the maximum, which will be illustrated in Section 6.

5. Other measures

In this section, we calculate the mean deviation and the Lorenz and Bonferroni curves. If T has KRS-GL distribution, then we can derive the average deviation around the mean and the median through the following relations

$$\delta_1 = 2\mu'_1 F(\mu'_1) - 2T(\mu'_1) \quad (32)$$

and

$$\delta_2 = \mu'_1 - 2T(m) \quad (33)$$

respectively. For $\beta > 1$ integer and $a, b > 0$, one gets

$$\begin{aligned}
 T(z) &= \int_0^z t f(t) dt = \alpha \beta \int_0^z t g(t) G^{\alpha-1}(t) [1 - G^\alpha(t)]^{\beta-1} dt \\
 &= \alpha \beta \sum_{k=0}^{\beta-1} \binom{\beta-1}{k} (-1)^k \int_0^z t g(t) G^{\alpha(k+1)-1}(t) dt \\
 &= \alpha \beta \sum_{k=0}^{\beta-1} \binom{\beta-1}{k} (-1)^k \sum_{j=0}^{\infty} \binom{2-\alpha(k+1)}{j} (-1)^j J_1(1+j) \\
 &\quad + b(1+p) \alpha \beta \sum_{k=0}^{\beta-1} \binom{\beta-1}{k} (-1)^k \sum_{j=0}^{\infty} \binom{2-\alpha(k+1)}{j} (-1)^j J_{p+1}(1+j) \quad (34)
 \end{aligned}$$

The median is given by

$$m = \sum_{j=0}^{\infty} \frac{(-b/a)^j \Gamma(jp + j + 1)}{j! \Gamma(jp + 2)} \left[\frac{1}{a} \ln \left(\frac{M}{1-M} \right) \right]^{jp+1}, \quad (35)$$

where $M = \left[1 - (0.5)^{1/\beta} \right]^{1/\alpha}$.

The Lorenz and Bonferroni curves are widely applied in economics for the study of income and poverty, as well as other areas such as reliability, demography, medicine, among others. Using $T(z)$ we can generate the Lorenz and Bonferroni curves by the following relationships

$$B(\pi) = \frac{T(q)}{\pi \mu'_1} \quad (36)$$

and

$$L(\pi) = \frac{T(q)}{\mu'_1} \quad (37)$$

respectively, where the quantile function $q = Q(\pi)$ is determined by (15), for $Q(\pi) = t$ and $\pi = G^*(t)$.

6. Order Statistics

Let $X_1, X_2, \dots, X_n$ be a random sample of size n obtained from a cumulative distribution function $F(\cdot)$. Then $Y_1 \leq Y_2 \leq \dots \leq Y_n$, where Y_i 's are the X_i 's arranged in ascending order of magnitude, are defined as the order statistics corresponding to the random sample $X_1, X_2, \dots, X_n$. Thus, the cumulative distribution of Y_k , with $k = 1, 2, \dots, n$ is given by

$$F_{Y_k}(t) = \sum_{k=1}^n \binom{n}{k} [F(t)]^k [1 - F(t)]^{n-k}, \quad (38)$$

and the corresponding probability density function is given by

$$f_{Y_k}(t) = \frac{n!}{(k-1)!(n-k)!} [F(t)]^k [1 - F(t)]^{n-k} f(t), \quad (39)$$

Then, the distribution of the minimum and the maximum are obtained as special cases of (38) and (39) by taking $k = 1$ and $k = n$ respectively. For the minimum, the cumulative distribution function is given by

$$F_{Y_1}(t) = 1 - [1 - F(t)]^n \quad (40)$$

and the probability density function by

$$f_{Y_1}(t) = n f(t) [1 - F(t)]^{n-1} \quad (41)$$

For the maximum, the cumulative distribution function and the probability density function are given by

$$F_{Y_n}(t) = [F(t)]^n \quad (42)$$

and

$$f_{Y_n}(t) = n f(t) [F(t)]^{n-1} \quad (43)$$

respectively.

In the next two subsections, we will detail the mathematical properties of the distributions of the minimum and the maximum, respectively. We will also give expressions for the quantile function and the k -th moments.

6.1 Distribution of the minimum

Suppose that $Y_1 \leq Y_2 \leq \dots \leq Y_n$ are the order statistics for a random sample of size n obtained from a KRS-GL. Then, the cumulative distribution function and the probability density function for the minimum Y_1 are given by

$$F_{Y_1}(t) = 1 - \left[1 - (1 - G^\alpha(t))^\beta \right] \quad (44)$$

and

$$f_{Y_1}(t) = n\alpha\beta g(t) G^{\alpha-1}(t) (1 - G^\alpha(t))^{\beta-1} \left[1 - (1 - G^\alpha(t))^\beta \right]^{n-1} \quad (45)$$

respectively. $G(t)$ and $g(t)$ are given by the expressions (10) and (11) respectively. We can write (45) in the following form

$$t = \frac{1}{a} \ln \left(\frac{F_{Y_1}^{**}(t)}{1 - F_{Y_1}^{**}(t)} \right) - \frac{b}{a} t |t|^p \quad (46)$$

where

$$F_{Y_1}^{**}(t) = \left[\left(1 - (1 - F_{Y_1}(t))^{1/n} \right)^{1/\beta} \right]^{1/\alpha} \quad (47)$$

Using the Lagrange expansion, we obtain the quantile function for (46) as

$$t = \begin{cases} -\sum_{j=0}^{\infty} \frac{(-b/a)^j \Gamma(jp+j+1)}{j! \Gamma(jp+2)} \left[\frac{1}{a} \ln \left(\frac{1-F_{Y_1}^{**}(t)}{F_{Y_1}^{**}(t)} \right) \right]^{jp+1}, & F_{Y_1}^{**}(t) \leq 0.5 \\ \sum_{j=0}^{\infty} \frac{(-b/a)^j \Gamma(jp+j+1)}{j! \Gamma(jp+2)} \left[\frac{1}{a} \ln \left(\frac{F_{Y_1}^{**}(t)}{1-F_{Y_1}^{**}(t)} \right) \right]^{jp+1}, & F_{Y_1}^{**}(t) \geq 0.5. \end{cases} \quad (48)$$

The k -th moment for the distribution of the minimum is given by

$$E(T^k) = \int_{-\infty}^{\infty} t^k f_{Y_1}(t) dt = n \sum_{j=0}^n (-1)^j \binom{n}{j} \delta_{k,jn}, \quad (49)$$

where the expressions for $\delta_{k,n}$ are given in Section 4.

6.2 Distribution of the maximum

Again, suppose that $Y_1 \leq Y_2 \leq \dots \leq Y_n$ are the order statistics for a random sample of size n obtained from a KRS-GL. Then, the cumulative distribution function and the probability density function for the maximum Y_n are given by

$$F_{Y_n}(t) = \left[1 - (1 - G^\alpha(t))^\beta \right]^n \quad (50)$$

and

$$f_{Y_n}(t) = n\alpha\beta g(t) G^{\alpha-1}(t) (1 - G^\alpha(t))^{\beta-1} \left[1 - (1 - G^\alpha(t))^\beta \right]^{n-1} \quad (51)$$

respectively. $G(t)$ and $g(t)$ are given by the expressions (10) and (11) respectively. We can write (51) in the following form

$$t = \frac{1}{a} \ln \left(\frac{F_{Y_n}^{**}(t)}{1 - F_{Y_n}^{**}(t)} \right) - \frac{b}{a} t |t|^p \quad (52)$$

where

$$F_{Y_n}^{**}(t) = \left[\left(1 - (1 - F_{Y_n}^{1/n}(t)) \right)^{1/\beta} \right]^{1/\alpha} \quad (53)$$

Using the Lagrange expansion, we obtain the quantile function for (52) as

$$t = \begin{cases} -\sum_{j=0}^{\infty} \frac{(-b/a)^j \Gamma(jp+j+1)}{j! \Gamma(jp+2)} \left[\frac{1}{a} \ln \left(\frac{1-F_{Y_n}^{**}(t)}{F_{Y_n}^{**}(t)} \right) \right]^{jp+1}, & F_{Y_n}^{**}(t) \leq 0.5 \\ \sum_{j=0}^{\infty} \frac{(-b/a)^j \Gamma(jp+j+1)}{j! \Gamma(jp+2)} \left[\frac{1}{a} \ln \left(\frac{F_{Y_n}^{**}(t)}{1-F_{Y_n}^{**}(t)} \right) \right]^{jp+1}, & F_{Y_n}^{**}(t) \geq 0.5. \end{cases} \quad (54)$$

The k -th moment for the distribution of the maximum is given by

$$E(T^k) = \int_{-\infty}^{\infty} t^k f_{Y_n}(t) dt = n\delta_{k,n}, \quad (55)$$

where the expressions for $\delta_{k,n}$ are given in Section 4.

7. Likelihood function

In this section the maximum likelihood estimation is used for the parameter estimation of the model (10). Suppose that T_i are random variables distributed as (10) with parameter vector $\boldsymbol{\theta} = (a, b, p, \alpha, \beta)$. The log-likelihood function is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \sum_{i=1}^n \ln(a + b(1+p)|t_i|^p) - \sum_{i=1}^n t_i(a + b|t_i|^p) + n\ln(\alpha\beta) \\ &- (\alpha + 1) \sum_{i=1}^n \{1 + \exp(-t_i(a + b|t_i|^p))\} \\ &+ (\beta + 1) \sum_{i=1}^n \ln(1 - (1 + \exp(-t_i(a + b|t_i|^p)))^\alpha) \end{aligned} \quad (56)$$

Due to the complexity of the log-likelihood function, numerical methods for calculating the gradient and the Hessian matrix can be used to obtain the score vector $U(\boldsymbol{\theta})$ and Fisher information matrix. Thus, the analytical expressions for the vector score and the Hessian matrix will be omitted.

As we use the maximum likelihood method to estimate the parameters of KRS-GL model, hypothesis tests of the type $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ versus $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ can be tested via likelihood ratio test. Such testing aims to check the adjustment of KRS-GL model and compare with some of the sub-models mentioned in Section 2. For example, suppose we are interested in testing whether the data are distributed as KRS-GL or if the data are distributed as Kum-L proposed by Santana et al (2012). One way to verify this, is by conducting the test $H_0 : b = 0$ versus $H_1 : b \neq 0$, for which the statistics LR is given by $w = 2 \left[\ell(\hat{a}, \hat{b}, \hat{p}, \hat{\alpha}, \hat{\beta}) - \ell(\tilde{a}, 0, \tilde{p}, \tilde{\alpha}, \tilde{\beta}) \right]$, where $\hat{a}, \hat{b}, \hat{p}, \hat{\alpha}$ and $\hat{\beta}$ are the ML estimators under H_1 (unrestricted model) and $\tilde{a}, \tilde{p}, \tilde{\alpha}$ and $\tilde{\beta}$ are the parameters estimated under H_0 .

8. Applications

In the previous sections we have given some theoretical aspects of the KRS-GL distribution. We have seen that this distribution can be used to model data which are unimodal, bimodal, symmetrical or asymmetrical. In this section we will give

two applications of the distribution. The first is related to waiting time between eruptions of Old Faithful Geyser, such data is frequently modelled by mixtures of distributions. Here we show that the distribution proposed in this paper can be used as an alternative to mixtures of distributions. The second application is related to the survival time of pigs, injected with different doses of tubercle bacilli, studied by Kundu, Kannan, and Balakrishnan [7] and originally reported by Bjerkedal [1].

8.1 Waiting time between eruptions of Old Faithful Geyser

This application shows the versatility of the new KRS-GL distribution. Using the data available in the free statistical software R (<http://www.r-project.org/>) we show how the bimodal distribution can take shape. The waiting time between eruptions of Old Faithful geyser in Yellowstone National Park, Wyoming, USA was used. The data stored in *faithful\$waiting* has 272 observations given in minutes. The new distribution has proven to be a good option to fit the data, because it allows a bimodal shape. Thus, after standardizing the distribution, we obtain the estimated parameters using the maximum likelihood method as $A = 0.04501322$, $B = 0.0004560879$, $p = 1.9732434$, $\alpha = 0.7664678$, $\beta = 0.4810185$ and $\mu = 65.5794003$. In the Figure 2, we can see the distribution adjustment to its empirical distribution and histogram.

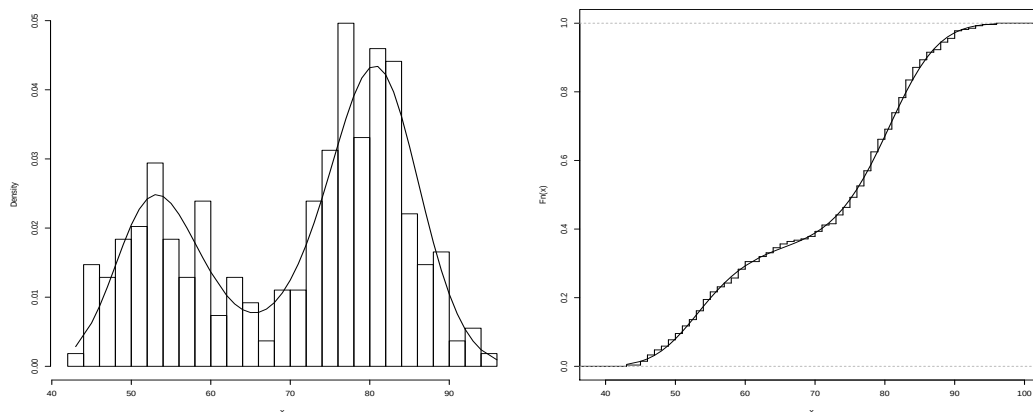


Figure 2 Distribution KRS-GL adjusted for waiting time between eruptions of Old Faithful geyser, with empirical distribution (right) and histogram (left).

We used the nonparametric Kolmogorov-Smirnov test to check the fit obtained. With mean square error between the new distribution and empirical distribution

equal to 0.0004092406 and p-value of 0.7344, we can conclude that there is no evidence to reject the hypothesis that the data follow this distribution.

A comparison of the proposed distribution with some of its sub-models using LR statistics is shown in Table 1. In the Figure 3, we can see the survival KRS-GL

Hypotheses	Statistic w	p-value
$H_0 : b = 0$ vs $H_1 : b \neq 0$	1086.294	0
$H_0 : \alpha = \beta = 0$ vs $H_1 : \alpha \neq 0$ or $\beta \neq 0$	1043.59	5.938254×10^{-06}

Table 1: LR test for Old Faithful Geyser data

distribution adjustment to its empirical distribution.

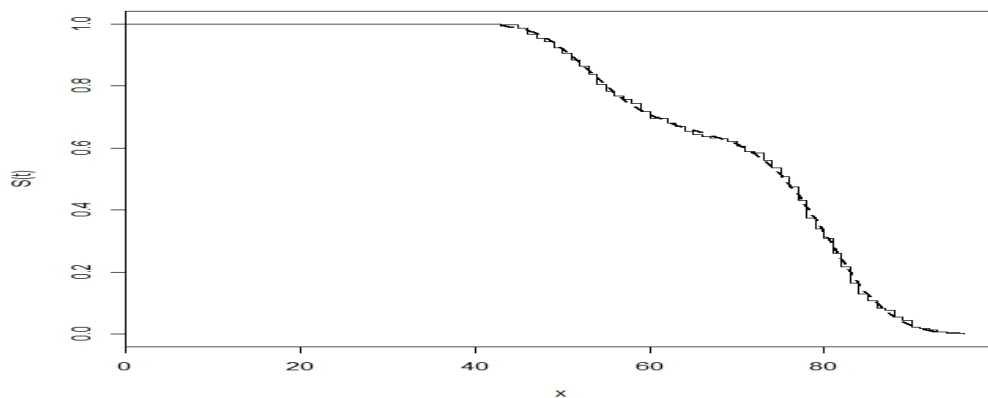


Figure 3 Kaplan-Meier nonparametric estimator versus survival KRS-GL adjusted for waiting time between eruptions of Old Faithful geyser, with empirical distribution.

8.2 Survival time in guinea pigs

This another application shows how the new KRS-GL distribution can be used to adjust an asymmetric shape data. The data were analyzed by Kundu, Kannan, and Balakrishnan [7] and originally reported by Bjerkedal [1] and correspond to the survival times of guinea pigs injected with different doses of tubercle bacilli.

It is known that guinea pigs have high susceptibility to human tuberculosis and that is why they were used in this study. Here, we are primarily concerned with the animals in the same cage and were under the same regime. The number of arrangements is the common logarithm of the number of units of bacillus in 0.5 ml solution, i.e. the system 6.6 corresponds to 4×10^6 bacilli per 0.5 ml

($\log(04 \times 106) = 6.6$) there were 72 observations in the regime corresponding to 6.6, which are listed in the Table 2. The new distribution has proven to be a good

2.00	15.00	22.00	24.00	24.00	32.00
32.00	33.00	34.00	38.00	38.00	43.00
44.00	48.00	52.00	53.00	54.00	54.00
55.00	56.00	57.00	58.00	58.00	59.00
60.00	60.00	60.00	60.00	61.00	62.00
63.00	65.00	65.00	67.00	68.00	70.00
70.00	72.00	73.00	75.00	76.00	76.00
81.00	83.00	84.00	85.00	87.00	91.00
95.00	96.00	98.00	99.00	109.00	110.00
121.00	127.00	129.00	131.00	143.00	146.00
146.00	175.00	175.00	211.00	233.00	258.00
258.00	263.00	297.00	341.00	341.00	376.00

Table 2: Guinea pigs data

option to fit the data, because it allows fitting well on the asymmetry of the data. Thus, after standardizing the distribution, we obtain the estimated parameters using the maximum likelihood method as $A = 0.06757096$, $B = 1.044953 \times 10^{-25}$, $p = 9.003746$, $\alpha = 5.097374$, $\beta = 2.160622 \times 10^{-1}$ and $\mu = 9.749253$. In the Figure 4, we can see the distribution adjustment KRS-GL to the empirical distribution and histogram.

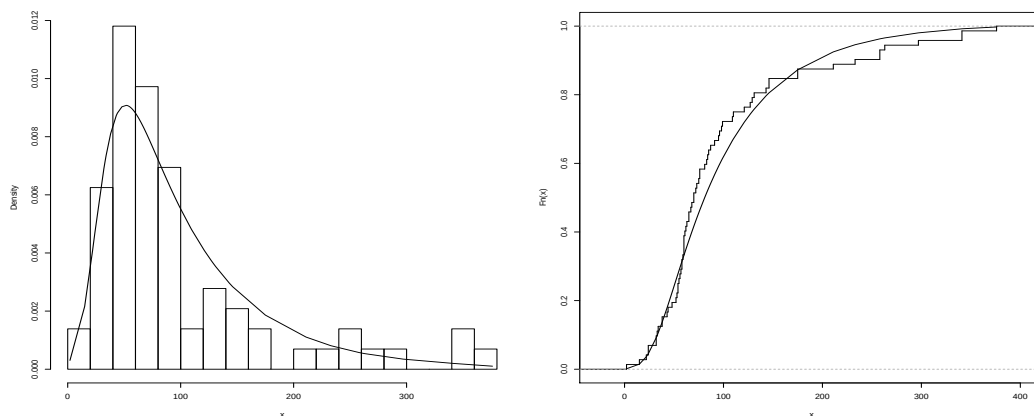


Figure 4 Distribution KRS-GL adjusted for survival times of guinea pigs injected with different doses of tubercle bacilli with empirical distribution (right) and histogram (left).

We used the nonparametric Kolmogorov-Smirnov test to check the fit obtained. With mean square error between the new distribution and empirical distribution equal to 0.003873119 and p-value of 0.6272, we can conclude that there is no evidence to reject the hypothesis that the data follow this distribution. A comparison of the proposed distribution with some of its sub-models using LR statistics is shown in Table 3. In the Figure 5, we can see the survival Kumaraswamy - Rathie

Hypotheses	Statistic w	p-value
$H_0 : b = 0$ vs $H_1 : b \neq 0$	390.6542	0.3274894
$H_0 : \alpha = \beta = 0$ vs $H_1 : \alpha \neq 0$ or $\beta \neq 0$	411.2976	8.050094×10^{-11}

Table 3: LR test for guinea pigs data

- Swamee adjustment to its empirical distribution.

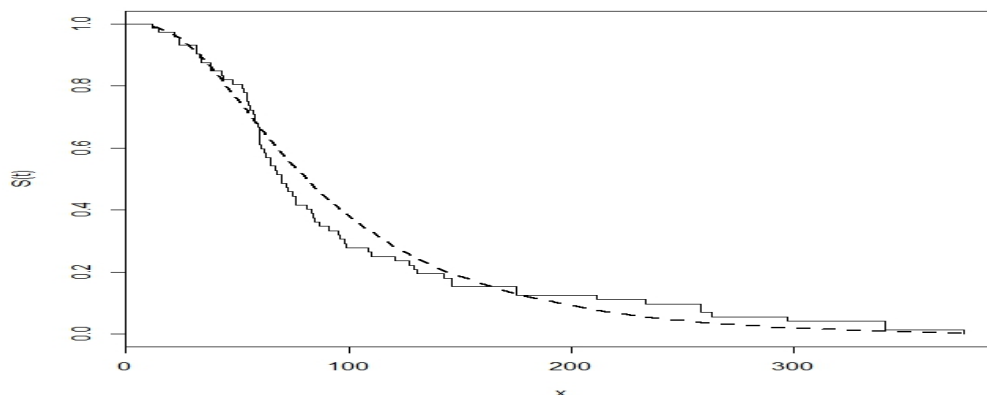


Figure 5 Kaplan-Meier nonparametric estimator versus survival KRS-GL adjusted with empirical distribution for survival times of guinea pigs injected with different doses of tubercle bacilli.

9. Conclusions

In this paper we proposed the KRS-GL distribution, which extends the logistic, generalized logistic (Rathie and Swamee, [11]) and Kumaraswamy logistic (Santana et al., [13]) distributions. We derived the mathematical formulas for the quantile function, based on the Lagrange expansion, the moments, Probability Weight Moments which are expressed in terms of the H function (Mathai et al., [8]). Also we derived other measures such as the mean and median deviations, and Bonferroni and Lorenz curves. The expression for the log-likelihood function was given

as well. In this article we used the likelihood ratio test to compare the proposed distribution with some of its sub-models.

Finally, two applications were given, one with bimodal data and the other with unimodal data, both asymmetrical, showing the versatility of the proposed distribution across its sub-models.

Acknowledgment

P. N. Rathie thanks CAPES, Brazil, for supporting his Senior National Visiting Professorship.

References

- [1] T. Bjerkedal, *Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli*, American Journal of Hygiene, 72 (1960), 130-148.
- [2] G. M. Cordeiro, E. M. M. Ortega, S. Nadarajah, *The Kumaraswamy Weibull distribution*, Journal of the Franklin Institute, 347 (2010), 1399-1429.
- [3] G. M. Cordeiro, M. de Castro, *A new family of generalized distributions*, Journal of Statistical Computation and Simulation, 81 (2011), 883-898.
- [4] G. M. Cordeiro, R.R. Pescim, E. M. M. Ortega, *The Kumaraswamy Generalized Half-Normal Distribution for Skewed Positive Data*. Journal of Data Science, 10 (2012), 195-224.
- [5] Y. L. Luke, *The Special Functions and Their Approximations*, Academic Press, New York, 1969.
- [6] P. Kumaraswamy, *Generalized probability density-function for double-bounded random-processes*, Journal of Hydrology, 46 (1980), 79-88.
- [7] D. Kundu, N. Kannan, N. Balakrishnan, *On the hazard function of Birnbaum-Saunders distribution and associated inference*, Computational Statistics and Data Analysis, 52 (2008), 2692-2702.
- [8] A. M. Mathai, R. K. Saxena, H. J. Haubold, *The H-Function: Theory and Applications*, Springer, New York, 2010.
- [9] M. A. R. Pascoa, E. M. M. Ortega, G. M. Cordeiro, *The Kumaraswamy generalized gamma distribution with application in survival analysis*, Statistical Methodology, 8 (2011), 411-433.

- [10] P. N. Rathie, *Normal distribution, Univariate*. International Encyclopedia of Statistical Sciences (2011), 1012-1013.
- [11] P. N. Rathie, P. K. Swamee, *On a new invertible generalized logistic distribution approximation to normal distribution*, Technical Research Report in Statistics, 07/2006 (2006), Department of Statistics, University of Brasilia, Brasilia, Brazil.
- [12] P. N. Rathie, P. K. Swamee, G. G. Matos, M. Coutinho, T. B. Carrijo, *H-function and statistical distributions*, Ganita, 59 (2008), 23-37.
- [13] T. V. F. Santana, E. M. M. Ortega, G. M. Cordeiro, G. O. Silva, *The Kumaraswamy-Log-Logistic distribution*, Journal of Statistical Theory and Applications, 11 (2012), 265-291.
- [14] M. D. Springer, *Algebra of Random Variables*, John Wiley, New York, 1979.
- [15] P. K. Swamee, P. N. Rathie, *Invertible alternatives to normal and lognormal distributions*, J. Hydrol. Eng. ASCE, 12 (2007), 218-221.
- [16] T. Yang, *Statistical Properties of Kumaraswamy Generalized Inverse Weibull Distribution*, Electronic Theses & Dissertations, Paper 678 (2012), <http://digitalcommons.georgiasouthern.edu/etd/678>.

