Brazilian Journal of Probability and Statistics
2013, Vol. 27, No. 2, 185–200
DOI: 10.1214/11-BJPS166

The beta generalized logistic distribution

Alice L. Morais, Gauss M. Cordeiro and Audrey H. M. A. Cysneiros

Universidade Federal de Pernambuco

Abstract. For the first time, a four-parameter beta generalized logistic distribution is obtained by compounding the beta and generalized logistic distributions. The new model extends some well-known distributions and its shape is quite flexible, specially the skewness and the tail weights, due to the extra shape parameters. We obtain general expansions for the moment generating and quantile functions. The estimation of the parameters is investigated by maximum likelihood. An application to a real data set is given to show the flexibility and potentiality of our distribution.

1 Introduction

Because of their flexibility, much attention has been given to the study of generalized distributions in recent times. Prentice (1976) proposed the type IV generalized logistic (GLIV) distribution as an extended distribution to modeling binary response data under the usual symmetric logistic distribution. The probability density function (p.d.f.) of the GLIV distribution, say GLIV(p, q), is given by

$$g_{p,q}(x) = \frac{1}{B(p,q)} \frac{e^{-qx}}{(1+e^{-x})^{p+q}}, \qquad x \in \mathbb{R}, p > 0, q > 0. \tag{1.1}$$

The cumulative distribution function (c.d.f.) $G_{p,q}(x)$ corresponding to (1.1) is

$$G_{p,q}(x) = I_{1/(1+e^{-x})}(p,q), \qquad x \in \mathbb{R}, p > 0, q > 0, \tag{1.2}$$

where $B(a,b)$ is the beta function, $B_x(a,b) = \int_0^x t^{a-1}(1-t)^{b-1}\,dt$ is the incomplete beta function and $I_x(a,b) = B_x(a,b)/B(a,b)$ is the incomplete beta function ratio. The associated moment generating function (m.g.f.) for $-p < t < q$ is given by $M_{p,q}(t) = \Gamma(p+t)\Gamma(q-t)[\Gamma(p)\Gamma(q)]^{-1}$.

The simplicity of the logistic distribution and its importance as a growth curve have made it one of the most important statistical models. The shape of the logistic distribution makes it simpler and also profitable on suitable occasions to replace the normal distribution. In order to improve the fit of the logistic model for bioassay and quantal response data, many generalized types of the logistic distribution have been proposed recently. These generalized distributions (indexed by one or

Key words and phrases. Beta distribution, generalized logistic distribution, maximum likelihood, order statistic.

Received May 2010; accepted July 2011.

more shape parameters) are developed to extend the scope of the logistic model to asymmetric probability curves and to improve the fit in the non-central probability regions.

In this note, we extend the GLIV distribution by introducing two extra shape parameters to define a new distribution refereed to as the beta generalized logistic (BGL) distribution. The role of the two additional parameters is to introduce skewness and to vary tail weights and provide greater flexibility in the shape of the generalized distribution and consequently in modeling observed data. It may be mentioned that although several skewed distribution functions exist on the positive real axis, not many skewed distributions are available on the whole real line, which are easy to use for data analysis purpose. The main idea is to introduce two shape parameters, so that the BGL distribution can be used to model skewed data, a feature which is very common in practice.

The BGL distribution is defined by the beta c.d.f. at the point $G_{p,q}(x)$, where $G_{p,q}(x)$ is the GLIV c.d.f. Many other generalized distributions obtained from the beta c.d.f. were introduced in the literature. First, Eugene, Lee and Famoye (2002) defined the beta normal distribution which has some advantages over the normal distribution. Nadarajah and Kotz (2004) defined the beta Gumbel distribution which has greater tail flexibility than the Gumbel distribution. Nadarajah and Gupta (2004) defined the beta Fréchet distribution and Barreto-Souza, Cordeiro and Simas (2011) presented some additional mathematical properties. Nadarajah and Kotz (2006) defined the beta exponential distribution, whose hazard function can be increasing and decreasing. The beta Weibull distribution was defined by Famoye, Lee and Olumolade (2005) and Lee, Famoye and Olumolade (2007) applied this distribution to censored data. Kong, Lee and Sepanski (2007) proposed the beta gamma distribution. Barreto-Souza, Santos and Cordeiro (2010) defined the beta generalized exponential distribution.

The rest of the article is organized as follows. In Section 2, we define the new distribution. In Section 3, we present some special sub-models and related distributions. General expansions for the BGL density function expressed as linear combinations of GLIV densities are derived in Section 4. Expansions for the quantile and generating functions are determined in Section 5. In Section 6, we obtain the mean deviations about the mean and the median and the Bonferroni and Lorenz curves. Expansions for the order statistics and their moments are derived in Section 7. Maximum likelihood estimation of the model parameters is discussed in Section 8. Section 9 provides an application to real data. Section 10 ends with some conclusions.

2 The new distribution

Let $G(x)$ be the c.d.f. of a random variable. A method to generalize distributions consists to define a new c.d.f. $F(x)$ from the baseline $G(x)$ by

$$\begin{aligned} F(x) &= I_{G(x)}(a,b) \\ &= \frac{1}{B(a,b)} \int_0^{G(x)} w^{a-1}(1-w)^{b-1}\,dw, \qquad a>0, b>0, \end{aligned} \tag{2.1}$$

Clearly, $F(x)$ coincides with $G(x)$ when $a=b=1$.

The BGL distribution is obtained by setting the GLIV c.d.f. (1.2) in (2.1). The BGL c.d.f. becomes

$$\begin{aligned} F(x) &= I_{I_{1/(1+e^{-x})}(p,q)}(a,b) \\ &= \frac{1}{B(a,b)} \int_0^{I_{1/(1+e^{-x})}(p,q)} w^{a-1}(1-w)^{b-1}\,dw, \qquad x \in \mathbb{R}, \end{aligned} \tag{2.2}$$

where the parameters a, b, p and q are positive real numbers.

We can verify that the GLIV cumulative function (1.2) can also be written in form (2.1), whereas $(1+e^{-x})^{-1}$ is the standard logistic baseline distribution. From (2.1) and using the property of the incomplete beta function $B_x(a,b) = B(a,b) - B_{1-x}(b,a)$, the BGL density function can be expressed as

$$\begin{aligned} f(x) &= \frac{1}{B(p,q)B(a,b)} \frac{e^{-qx}}{(1+e^{-x})^{p+q}} \\ &\quad \times \left[I_{1/(1+e^{-x})}(p,q)\right]^{a-1} \left[I_{e^{-x}/(1+e^{-x})}(q,p)\right]^{b-1}. \end{aligned} \tag{2.3}$$

A random variable X having density function (2.3) is denoted by BGL(a,b,p,q). The BGL distribution is symmetric for $p=q$ and $a=b$. Using (2.2), we note that the BGL$(m,n,1,1)$ and BGL$(1,1,m,n)$ distributions are identical.

3 Related distributions

Evidently, the density function (2.3) does not involve any complicated function but generalizes a few interesting distributions. The GLIV distribution has three remarkable submodels. So, the BGL distribution generalizes the GLIV distribution and its submodels. They are formally the type I, II and III beta generalized distributions, which are denoted by BGLI, BGLII and BGLIII, respectively. We present in Table 1 the parameter restriction to obtain these submodels and their desnity functions. Let $Y \sim$ BGL(a,b,p,q). We also present in Table 1 the distribution of $X = e^{-Y}$ and $X = (p/q)e^{-Y}$, which returns the Beta Beta Prime (BBP) and Beta F-Snedecor (BFSn) distributions, respectively. Due to the transformation, the set $\mathbb{R}^+$ is the support of the last two distributions and then they may be used for model

Table 1 *Some distributions related to the BGL distribution*

Distribution	Condition	Density
BGLI	$q = 1$	$f(x) = \frac{pe^{-x}}{B(a,b)} \frac{[(1+e^{-ax})^p - 1]^{b-1}}{(1+e^{-x})^{a+pb}}$
BGLII	$p = 1$	$f(x) = \frac{qe^{-bqx}}{B(a,b)(1+e^{-x})^{qb+1}} [1 - \frac{e^{-qx}}{(1+e^{-x})^q}]$
BGLIII	$p = q$	$f(x) = \frac{B(p,p)^{1-a-b}}{B(a,b)} \frac{e^{-px}}{(1+e^{-x})^{2p}} [B_{1/(1+e^{-x})}(p, p)]^{a+b-2}$
BBP	$X = e^{-Y}$	$f(x) = \frac{B(p,q)^{1-a-b}}{B(a,b)} \frac{x^{q-1}}{(1+x)^{p+q}} [B_{x/(1+x)}(q, p)]^{(a-1)}$ $\times [B_{1/(1+x)}(p, q)]^{(b-1)}$
BFS	$X = (p/q)e^{-Y}$	$f(x) = \frac{B(a,b)^{-1}(q/p)^{q/2} x^{q/2-1}}{B(p/2,q/2)(1+(q/p)x)^{(p+q)/2}}$ $\times [I_{x/((p/q)+x)}(q/2, p/2)]^{a-1} [I_{1/(1+(q/p)x)}(p/2, q/2)]^{b-1}$

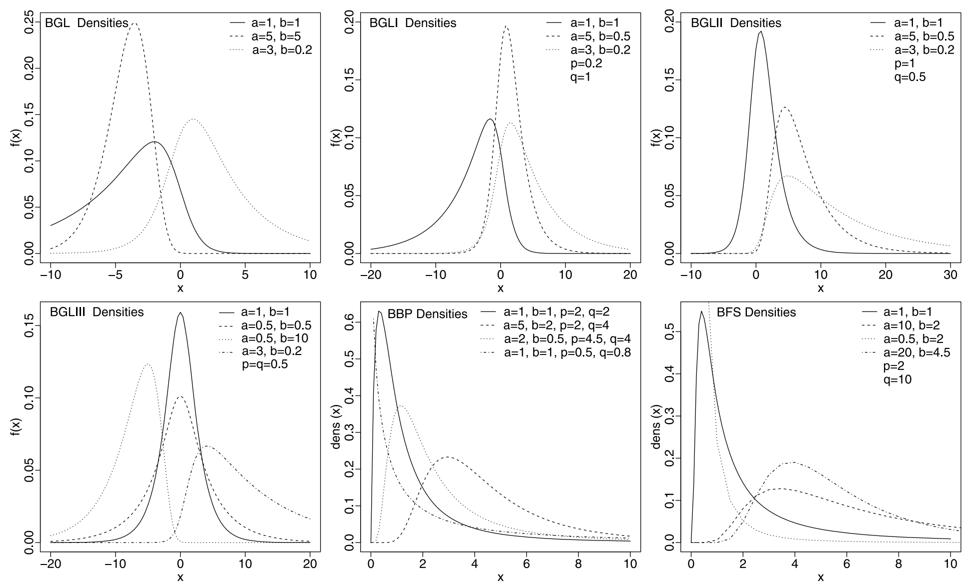

Figure 1 *Plots of the BGL, BGLI, BGLII, BGLIII, BBP and BFSn densities for selected parameter values.*

lifetime data. Figure 1 gives plots of the density functions mentioned in this section. The bold lines represent the parent distributions and the dotted lines illustrates the flexibility inherited from the parameters added after the generalization.

4 Expansion for the density function

We provide an expansion for the BGL density function which will be helpful to obtain some mathematical properties for this distribution. Beyond the theoretical importance, these expansions can be used as an alternative way of numerical integration. For $b > 0$ real non-integer, the power series for $(1-w)^{b-1}$ in (2.1) yields

$$\int_0^x w^{a-1}(1-w)^{b-1}\,dw = \sum_{j=0}^{\infty} \frac{(-1)^j \binom{b-1}{j}}{(a+j)} x^{a+j}, \tag{4.1}$$

where the binomial coefficient is defined for any real. If b is an integer, the index j in (4.1) stops at $b-1$. From (1.2) and (4.1), we can express the BGL cumulative function as

$$F(x) = \sum_{r=0}^{\infty} w_r I_{1/(1+e^{-x})}(p,q)^{a+r}, \tag{4.2}$$

where the coefficients are

$$w_r = w_r(a,b) = \frac{(-1)^r \binom{b-1}{r}}{B(a,b)(a+r)}.$$

For a integer, (4.2) provides the BGL c.d.f. as an infinite power series expansion of GLIV c.d.f.'s. Otherwise, if a is real non-integer, we can expand $I_{1/(1+e^{-x})}(p,q)^{a+r}$ to obtain the BGL c.d.f. as an infinite power series of GLIV c.d.f.'s. We have

$$I_{1/(1+e^{-x})}(p,q)^{a+r} = \sum_{j=0}^{\infty} \binom{a+r}{j} (-1)^j \{1 - I_{1/(1+e^{-x})}(p,q)\}^j$$

and then

$$I_{1/(1+e^{-x})}(p,q)^{a+r} = \sum_{j=0}^{\infty}\sum_{k=0}^{j} (-1)^{j+k} \binom{a+r}{j}\binom{j}{k} I_{1/(1+e^{-x})}(p,q)^k.$$

We substitute $\sum_{j=0}^{\infty}\sum_{k=0}^{j}$ for $\sum_{k=0}^{\infty}\sum_{j=k}^{\infty}$ to obtain

$$I_{1/(1+e^{-x})}(p,q)^{a+r} = \sum_{k=0}^{\infty} s_k(a+r) I_{1/(1+e^{-x})}(p,q)^k,$$

where

$$s_k(\alpha) = \sum_{j=k}^{\infty} (-1)^{k+j} \binom{\alpha}{j}\binom{j}{k}. \tag{4.3}$$

Hence, from (4.2), we can write

$$F(x) = \sum_{r=0}^{\infty} t_r I_{1/(1+e^{-x})}(p,q)^r, \tag{4.4}$$

where

$$t_r = t_r(a,b) = \sum_{l=0}^{\infty} w_l s_r(a+l),$$

and $s_r(a+j)$ is calculated from (4.3). The functions $t_r(a,b)$ and $s_r(a+l)$ are easily computed in algebraic statistical software.

Expansions for the BGL density function can be obtained by simple differentiation of (4.2) (for $a > 0$ integer)

$$f(x) = g_{p,q}(x) \sum_{r=0}^{\infty} (a+r) w_r I_{1/(1+e^{-x})}(p,q)^{a+r-1} \tag{4.5}$$

and of (4.4) (for $a > 0$ real non-integer)

$$f(x) = g_{p,q}(x) \sum_{r=0}^{\infty} (r+1) t_{r+1} I_{1/(1+e^{-x})}(p,q)^{r}. \tag{4.6}$$

For both equations (4.5) and (4.6), we require a power series expansion for $I_{1/(1+e^{-x})}(p,q)^r$. We can use the incomplete beta function expansion for $q > 0$ real non-integer

$$I_x(p,q) = \frac{x^p}{B(p,q)} \sum_{m=0}^{\infty} \frac{(1-q)_m x^m}{(p+m)m!},$$

where $(f)_k = \Gamma(f+k)/\Gamma(f)$. First, we obtain an expansion for $I_{1/(1+e^{-x})}(p,q)^r$ from

$$\begin{aligned} I_{1/(1+e^{-x})}(p,q)^r &= \frac{1}{B(p,q)^r (1+e^{-x})^{pr}} \left(\sum_{m=0}^{\infty} d_m y^m \right)^r \\ &= \frac{1}{B(p,q)^r (1+e^{-x})^{pr}} \sum_{m=0}^{\infty} c_{r,m} y^m, \end{aligned} \tag{4.7}$$

where $y = (1+e^{-x})^{-1}$, $d_m = \frac{(1-q)_m}{(p+m)m!}$ and the coefficients $c_{r,m}$ (for $r = 1, 2, \ldots$) can be calculated from the expansion of a power series raised to a positive integer power (Gradshteyn and Ryzhik (2007)) given by the recurrence relation

$$c_{r,m} = (m d_0)^{-1} \sum_{j=1}^{i} (rj - m + j) d_j c_{r,m-j}, \tag{4.8}$$

and $c_{0,m} = d_0^r$. Hence, for $a > 0$ integer, we can write from (4.5) and (4.7)

$$f(x) = \sum_{r,m=0}^{\infty} \rho_{\text{int}}(r,m) g_{p(a+r)+m,q}(x), \tag{4.9}$$

where

$$\rho_{\mathrm{int}}(r,m) = \frac{(a+r)w_r c_{a+r-1,m} B(p(a+r)+m,q)}{B(p,q)^{a+r}}.$$

In a similar way, for a real non-integer, we obtain from (4.6) and (4.7)

$$f(x) = \sum_{r,m=0}^{\infty} \rho_{\mathrm{real}}(r,m) g_{p(r+1)+m,q}(x), \tag{4.10}$$

where

$$\rho_{\mathrm{real}}(r,m) = \frac{(r+1)t_{r+1} c_{r,m} B(p(r+1)+m,q)}{B(p,q)^{r+1}}.$$

Equations (4.9) and (4.10) are the main results of this section. They show that the BGL density function can be written as simple linear combinations of GLIV densities. Then, several mathematical properties of the BGL distribution can follow from those properties of the GLIV distributions. They (and other expansions in the article) can be evaluated in symbolic computation software such as Mathematica and Maple. These symbolic software have currently the ability to deal with analytic expressions of formidable size and complexity.

5 Quantile and generating functions

The quantile function $Q(u)$ of the BGL distribution follows from (2.1) as

$$Q(u) = -\log[\{Q_{p,q}(Q_{a,b}(u))\}^{-1} - 1], \tag{5.1}$$

where $Q_{p,q}(u)$ denotes the beta quantile function with parameters p and q. The following expansion for $Q_{p,q}(u)$ can be found in wolfram website[1]

$$\begin{aligned} Q_{p,q}(u) = w &+ \frac{q-1}{p+1}w^2 + \frac{(q-1)(p^2+3pq-p+5q-4)}{2(p+1)^2(p+2)}w^3 \\ &+ \frac{(q-1)[p^4+(6q-1)p^3+(q+2)(8q-5)p^2]}{3(p+1)^3(p+2)(p+3)}w^4 \\ &+ \frac{(q-1)[(33q^2-30q+4)p+q(31q-47)+18]}{3(p+1)^3(p+2)(p+3)}w^5 \\ &+ O(u^{6/p}), \end{aligned}$$

where $w = [pB(p,b)u]^{1/p}$ for $p > 0$.

[1] http://functions.wolfram.com/06.23.06.0004.01.

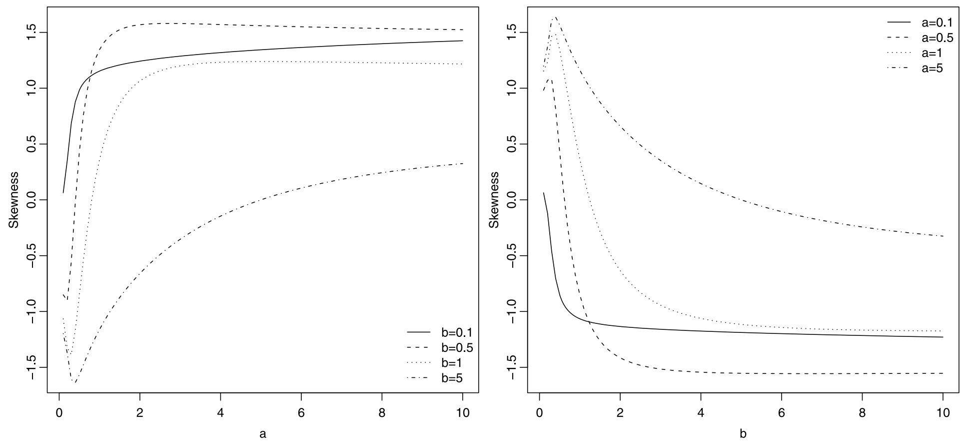

Figure 2 *Skewness of the BGL distribution as function of a and b for* $p = 0.5$ *and* $q = 0.4$.

The BGL generating function follows from (4.9) and (4.10). For $a > 0$ integer, (4.9) gives

$$M_X(t) = \frac{\Gamma(q-t)}{\Gamma(q)} \sum_{r,m=0}^{\infty} \rho_{\mathrm{int}}(r,m) \frac{\Gamma(p(a+r)+m+t)}{\Gamma(p(a+r)+m)},$$

whereas for $a > 0$ real non-integer, (4.10) yields

$$M_X(t) = \frac{\Gamma(q-t)}{\Gamma(q)} \sum_{r,m=0}^{\infty} \rho_{\mathrm{real}}(r,m) \frac{\Gamma(p(r+1)+m+t)}{\Gamma(p(r+1)+m)}.$$

The rth moment is determined by the rth derivative of $M_X(t)$ at $t = 0$. The skewness and kurtosis measures can now be calculated using well-known relationships. Plots of the skewness and kurtosis for some choices of the parameters a and b, fixing $p = 0.5$ and $q = 0.4$, are shown in Figures 2 and 3, respectively.

6 Mean deviations

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and median. If X has the BGL distribution with c.d.f. $F(x)$, we can derive the mean deviations about the mean $\nu = E(X)$ and about the median m from the relations

$$\delta_1 = \int_{-\infty}^{\infty} |x-\nu| f(x)\,dx \quad \text{and} \quad \delta_2 = \int_{-\infty}^{\infty} |x-m| f(x)\,dx,$$

respectively. The median m is given by

$$m = -\log[\{Q_{p,q}(Q_{a,b}(0.5))\}^{-1} - 1].$$

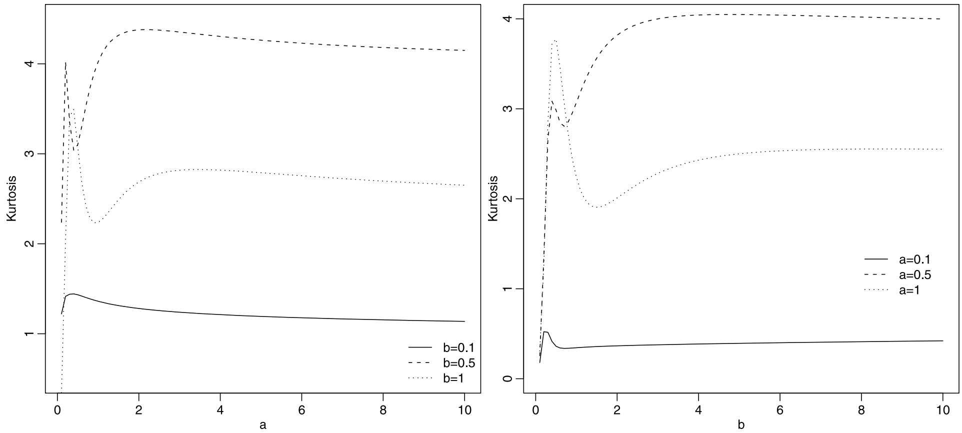

Figure 3 *Kurtosis of the BGL distribution as function of a and b for $p = 0.5$ and $q = 0.4$.*

These measures can be calculated from

$$\delta_1 = 2\nu F(\nu) - 2J(\nu) \quad \text{and} \quad \delta_2 = E(X) - 2J(m), \tag{6.1}$$

where

$$J(s) = \int_{-\infty}^{s} x f(x)\,dx = \frac{1}{B(a,b)} \int_{-\infty}^{s} x g(x) G(x)^{a-1} \{1 - G(x)\}^{b-1},$$

and $F(\nu)$ is easily calculated from (2.1). We derive a formula to obtain the integral $J(s)$. For $a > 0$ integer, we have from (4.2)

$$\begin{aligned} J(s) = \sum_{r,m,n=0}^{\infty} & \frac{\rho_{\text{int}}(r,m)}{B(p(a+r)+m,q)(n+1)} \\ & \times \bigl[B\bigl(p(a+r)+m+n+1,q\bigr)G_{p(a+r)+m+n+1,q}(s) \\ & \quad - B\bigl(p(a+r)+m,q+n+1\bigr)G_{p(a+r)+m,q+n+1}(s)\bigr], \end{aligned}$$

whereas for $a > 0$ real non-integer, we obtain from (4.4)

$$\begin{aligned} J(s) = \sum_{r,m,n=0}^{\infty} & \frac{\rho_{\text{int}}(r,m)}{B(p(r+1)+m,q)(n+1)} \\ & \times \bigl[B\bigl(p(r+1)+m+n+1,q\bigr)G_{p(r+1)+m+n+1,q}(s) \\ & \quad - B\bigl(p(r+1)+m,q+n+1\bigr)G_{p(r+1)+m,q+n+1}(s)\bigr]. \end{aligned}$$

The above formulae can be used to determine Bonferroni and Lorenz curves which have applications in economics, reliability, demography, insurance and medicine. They are defined by

$$B(\pi) = \frac{J(q)}{\pi \nu} \quad \text{and} \quad L(\pi) = \frac{J(q)}{\nu}, \tag{6.2}$$

respectively, where $\nu = E(X)$ and $q = Q(\pi)$ is calculated by (5.1) for given probability π.

7 Expansions for the order statistics

Moments of order statistics play an important role in quality control testing and reliability, where a practitioner needs to predict the failure of future items based on the times of a few early failures. These predictors are often based on moments of order statistics. We derive an explicit expression for the density of the ith order statistic $X_{i:n}$, say $f_{i:n}(x)$, in a random sample of size n from the BGL distribution. It is well-known that

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1} \tag{7.1}$$

for $i = 1, \ldots, n$. For a beta generalized model defined from the parent functions $g(x)$ and $G(x)$, $f_{i:n}(x)$ can be written as

$$f_{i:n}(x) = \frac{g(x) G(x)^{a-1} \{1 - G(x)\}^{b-1}}{B(a,b) B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1}. \tag{7.2}$$

Expanding $F(x)^{i+j-1}$ in a similar way to (4.7), we have

$$F(x)^{i+j-1} = \sum_{k=0}^{\infty} c_{i+j-1,k} G(x)^{a(i+j-1)+k}, \tag{7.3}$$

where the coefficients $c_{r,m}$ come from (4.8) by taking $d_m = w_m$. Replacing (7.3) in (7.2) yields

$$f_{i:n}(x) = \frac{1}{B(a,b) B(i, n-i+1)} \times \sum_{k=0}^{\infty} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} c_{i+j-1,k} f_{a(i+j)+k}(x), \tag{7.4}$$

where $f_m(x)$ is the density of the BGL(m, b, p, q) distribution. For $a > 0$ integer, we can write from (4.9) and (7.4)

$$f_{i:n}(x) = \sum_{r,m,k=0}^{\infty} \sum_{j=0}^{n-i} \tau_{\text{int}}(r, m, k, j) g_{p(a(i+j)+k+r)+m,q}(x), \tag{7.5}$$

where

$$
\begin{aligned}
\tau_{\mathrm{int}}&(r,m,k,j)\\
&=\binom{n-i}{j}\\
&\quad\times\frac{(-1)^j w_r c_{i+j-1,k} e_{a(i+j)+k+r-1,m} B(p(a(i+j)+k+r)+m,q)}{(a(i+j)+k+r)^{-1} B(p,q)^{a(i+j)+k+r} B(a,b) B(i,n-i+1)}
\end{aligned}
$$

and $e_{r,m}=c_{r,m}$ comes from (4.8) by taking $d_m=\frac{(1-q)_m}{(p+m)m!}$. For $a>0$ real non-integer, we have

$$
f_{i:n}(x)=\sum_{r,m,k=0}^{\infty}\sum_{j=0}^{n-i}\tau_{\mathrm{real}}(r,m,k,j)g_{p(r+1)+m,q}(x), \tag{7.6}
$$

where

$$
\begin{aligned}
\tau_{\mathrm{real}}&(r,m,k,j)\\
&=\binom{n-i}{j}\frac{(-1)^j t_{r+1}(a(i+j)+k,b) c_{i+j-1,k} e_{r,m} B(p(r+1)+m,q)}{(r+1)^{-1} B(p,q)^{r+1} B(a,b) B(i,n-i+1)}.
\end{aligned}
$$

The rth moment of $X_{i:n}$ follows from equations (7.5) and (7.6). For $a>0$ integer, (7.5) gives

$$
E\big(X_{i:n}^r\big)=\sum_{r,m,k=0}^{\infty}\sum_{j=0}^{n-i}\tau_{\mathrm{int}}(r,m,k,j)E\big(Y_{p(a(i+j)+k+r)+m,q}^r\big),
$$

and for $a>0$ real non-integer, (7.6) yields

$$
E\big(X_{i:n}^r\big)=\sum_{r,m,k=0}^{\infty}\sum_{j=0}^{n-i}\tau_{\mathrm{real}}(r,m,k,j)E\big(Y_{p(r+1)+m,q}^r\big),
$$

where $Y_{p,q}\sim \mathrm{GLIV}(p,q)$.

8 Estimation

Let $x_1,\ldots,x_n$ be an independent random sample from the BGL distribution. The total log-likelihood is given by

$$
\begin{aligned}
\ell&=\ell(a,b,p,q;x)\\
&=-n\log B(a,b)+n(1-a-b)\log B(p,q)\\
&\quad-(p+q)\sum_{i=1}^{n}\log\big(1+e^{-x_i}\big)+(a-1)\sum_{i=1}^{n}\log B_{1/(1+e^{-x_i})}(p,q)\\
&\quad+(b-1)\sum_{i=1}^{n}\log B_{e^{-x_i}/(1+e^{-x_i})}(q,p)-q\sum_{i=1}^{n}x_i.
\end{aligned}
$$

The components of the score function are ($\psi(\cdot)$ is the digamma function)

$$\frac{\partial \ell}{\partial a} = n\psi(a+b) - n\psi(a) - n\log B(p,q) + \sum_{i=1}^{n} \log B_{1/(1+e^{-x_i})}(p,q),$$

$$\frac{\partial \ell}{\partial b} = n\psi(a+b) - n\psi(b) - n\log B(p,q) + \sum_{i=1}^{n} \log B_{e^{-x_i}/(1+e^{-x_i})}(q,p),$$

$$\begin{aligned}\frac{\partial \ell}{\partial p} &= (1-a-b)n\{\psi(p) - \psi(p+q)\} - \sum_{i=1}^{n} \log(1+e^{-x_i}) \\ &\quad + (a-1)\frac{\partial}{\partial p}\sum_{i=1}^{n} \log B_{1/(1+e^{-x_i})}(p,q) \\ &\quad + (b-1)\frac{\partial}{\partial p}\sum_{i=1}^{n} \log B_{e^{-x_i}/(1+e^{-x_i})}(q,p),\end{aligned}$$

$$\begin{aligned}\frac{\partial \ell}{\partial q} &= (1-a-b)n\{\psi(q) - \psi(p+q)\} - \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \log(1+e^{-x_i}) \\ &\quad + (a-1)\frac{\partial}{\partial q}\sum_{i=1}^{n} \log B_{1/(1+e^{-x_i})}(p,q) \\ &\quad + (b-1)\frac{\partial}{\partial q}\sum_{i=1}^{n} \log B_{e^{-x_i}/(1+e^{-x_i})}(q,p).\end{aligned}$$

The maximum likelihood estimates (MLEs) of the parameters can be obtained by solving the system of nonlinear equations $\nabla\ell = 0$. Let $\theta = (a, b, p, q)^T$ be the parameter vector of the BGL distribution. The total observed information matrix, say $K(\theta)$, has elements given by ($\psi'(\cdot)$ is the trigamma function)

$$\frac{\partial^2 \ell}{\partial a^2} = n\psi'(a+b) - n\psi'(a), \qquad \frac{\partial^2 \ell}{\partial a\,\partial b} = n\psi'(a+b),$$

$$\frac{\partial^2 \ell}{\partial a\,\partial p} = n\psi(p+q) - n\psi(p) + \frac{\partial}{\partial p}\sum_{i=1}^{n} \log B_{1/(1+e^{-x_i})}(p,q),$$

$$\frac{\partial^2 \ell}{\partial a\,\partial q} = n\psi(p+q) - n\psi(q) + \frac{\partial}{\partial q}\sum_{i=1}^{n} \log B_{1/(1+e^{-x_i})}(p,q),$$

$$\frac{\partial^2 \ell}{\partial b^2} = n\psi'(a+b) - n\psi'(b),$$

$$\frac{\partial^2 \ell}{\partial b\,\partial p} = n\psi(p+q) - n\psi(p) + \frac{\partial}{\partial p}\sum_{i=1}^{n} \log B_{e^{-x_i}/(1+e^{-x_i})}(q,p),$$

$$\frac{\partial^2 \ell}{\partial b\,\partial q} = n\psi(p+q) - n\psi(q) + \frac{\partial}{\partial q}\sum_{i=1}^{n} \log B_{e^{-x_i}/(1+e^{-x_i})}(q, p),$$

$$\begin{aligned}\frac{\partial^2 \ell}{\partial p^2} &= n(1-a-b)\{\psi'(p) - \psi'(p+q)\}\\ &\quad + (a-1)\frac{\partial^2}{\partial p^2}\sum_{i=1}^{n} \log B_{1/(1+e^{-x_i})}(p, q)\\ &\quad + (b-1)\frac{\partial^2}{\partial p^2}\sum_{i=1}^{n} \log B_{e^{-x_i}/(1+e^{-x_i})}(q, p),\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \ell}{\partial p\,\partial q} &= -n(1-a-b)\psi'(p+q) + (a-1)\frac{\partial^2}{\partial p\,\partial q}\sum_{i=1}^{n} \log B_{1/(1+e^{-x_i})}(p, q),\\ &\quad + (b-1)\frac{\partial^2}{\partial p\,\partial q}\sum_{i=1}^{n} \log B_{e^{-x_i}/(1+e^{-x_i})}(q, p),\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 \ell}{\partial q^2} &= n(1-a-b)\{\psi'(q) - \psi'(p+q)\}\\ &\quad + (a-1)\frac{\partial^2}{\partial q^2}\sum_{i=1}^{n} \log B_{1/(1+e^{-x_i})}(p, q)\\ &\quad + (b-1)\frac{\partial^2}{\partial q^2}\sum_{i=1}^{n} \log B_{e^{-x_i}/(1+e^{-x_i})}(q, p).\end{aligned}$$

For the BGL distribution, it seems complicated to obtain the expected value of $K(\theta)$. Since the observed and expected information matrix converge to the same matrix, we believe it is reasonable to give the observed matrix. Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, we can consider the multivariate normal approximation $N_4(0, K(\theta)^{-1})$ for $\sqrt{n}(\hat{\theta} - \theta)$, where $K(\theta)^{-1}$ is the unit observed information matrix. The approximate multivariate normal $N_4(0, n^{-1}K(\hat{\theta})^{-1})$ distribution of $\hat{\theta}$ can be used to construct confidence intervals and confidence regions for the parameters and for the hazard and survival functions.

The likelihood ratio (LR) statistic is useful for testing the goodness of fit of the BGL distribution and for comparing this distribution with some of its special submodels. If we consider the partition $\theta = (\theta_1^T, \theta_2^T)^T$, tests of hypotheses of the type $H_0 : \theta_1 = \theta_1^{(0)}$ versus $H_A : \theta_1 \neq \theta_1^{(0)}$ can be performed using LR statistics. The LR statistic for testing the null hypothesis H_0 is $w = 2\{\ell(\hat{\theta}) - \ell(\tilde{\theta})\}$, where $\hat{\theta}$ and $\tilde{\theta}$ are the MLEs of θ under H_A and H_0, respectively. Under the null hypothesis, $w \xrightarrow{d} \chi_q^2$, where q is the dimension of the vector θ_1 of interest. The LR test rejects H_0 if $w >$

ξ_γ, where ξ_γ denotes the upper $100\gamma\%$ point of the χ^2_q distribution. For example, we can check if the fit using the BGL distribution is statistically "superior" to a fit using the GLIV distribution for a given data set by testing $H_0: a = b = 1$ versus $H_A: H_0$ is not true.

9 Application

The INPC is a national index of consumer prices of Brazil, produced by the IBGE since 1979. The period of collection extends from the day 01 to 30 of the reference month. The INPC measures the cost of living of households with heads employees. The search is done in the metropolitan regions of Rio de Janeiro, Porto Alegre, Belo Horizonte, Recife, São Paulo, Belém, Fortaleza, Salvador and Curitiba, in addition to Brasília and the city of Goiânia. This index can be found on seriesestatisticas.ibge.gov.br.

We fit the BGL model and some of its special sub-models discusssed in Section 3 to these data. The MLEs of the model parameters followed by their estimated standard errors, the maximized log-likelihoods ($\hat{\ell}$) and the p-values for the LR statistics are listed in Table 2.

Clearly, for the usual significance levels in all tests, we can accept the BGL model. In Figure 4 we show the histogram of the data and the fitted density functions. These results illustrate the potentiality of the BGL distribution and the necessity to adopt extra shape parameters.

10 Conclusions

In this article, we introduce the four-parameter beta generalized logistic (BGL) distribution that extends the type IV generalized logistic distribution. This is achieved following the idea of the cumulative distribution function of the class of beta generalized distributions proposed by Eugene et al. (2002). The BGL distribution is quite flexible in analyzing positive data in place of several other logistic distributions. We provide a mathematical treatment of the new distribution including expansions for the density function, moment generating function, mean deviations,

Table 2 *MLEs for the BGL, GLIV, BGLI, BGLII and BGLIII distributions*

Model	$\hat{a}$	$\hat{b}$	$\hat{p}$	$\hat{q}$	$\hat{\ell}$	p-value
BGL	179.92 (0.0570)	0.39 (0.0046)	0.92 (0.0100)	6.96 (0.0767)	−118.97	–
GLIV	–	–	9.43 (1.0582)	5.18 (0.5690)	−134.01	2.9×10^{-07}
BGLI	11.13 (10.8725)	5.13 (0.5677)	0.86 (0.7319)	–	−133.94	4.5×10^{-08}
BGLII	19.03 (6.0510)	1.03 (0.5795)	–	3.21 (1.0054)	−125.20	4.1×10^{-04}
BGLIII	1.59 (0.4909)	0.26 (0.1359)	16.27 (8.8041)	–	−125.74	2.3×10^{-04}

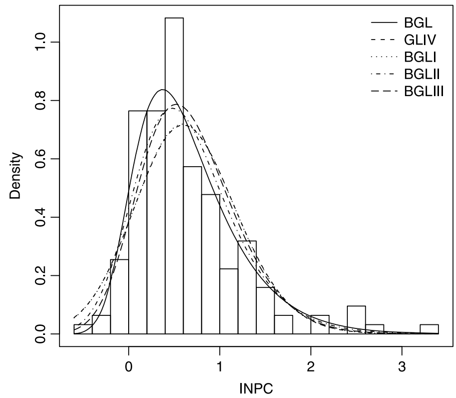

Figure 4 *Plots of the fitted density functions.*

Bonferroni and Lorenz curves, order statistics and their ordinary moments. The estimation of parameters is approached by the method of maximum likelihood and the observed information matrix is derived. One application of the BGL distribution shows that the new distribution could provide a better fit than other well-known logistic type models.

Acknowledgments

We thank the Editor and two anonymous referees for comments and suggestions. We also gratefully acknowledge financial support from CAPES and CNPq (Brazil).

References

Barreto-Souza, W., Cordeiro, G. M. and Simas, A. B. (2011). Some results for beta Fréchet distribution. *Communications in Statistics. Theory and Methods* **40**, 798–811. MR2762926

Barreto-Souza, W., Santos, A. and Cordeiro, G. M. (2010). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation* **80**, 159–172. MR2603623

Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics. Theory and Methods* **31**, 497–512. MR1902307

Famoye, F., Lee, C. and Olumolade, O. (2005). The beta–Weibull distribution. *Journal of Statistical Theory and Applications* **4**, 121–136. MR2210672

Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of Integrals, Series, and Products*, 7nd ed. Amsterdam: Elsevier. MR2360010

Kong, L., Lee, C. and Sepanski, J. H. (2007). On the properties of beta–gamma distribution. *Journal of Modern Applied Statistical Methods* **6**, 187–211.

Lee, C., Famoye, F. and Olumolade, O. (2007). Beta–Weibull distribution: Some properties and applications to censored data. *Journal of Modern Applied Statistical Methods* **6**, 173–186.

Nadarajah, S. and Gupta, A. K. (2004). The beta Fréchet distribution. *Far East Journal of Theoretical Statistics* **14**, 15–24. MR2108090

Nadarajah, S. and Kotz S. (2004). The beta Gumbel distribution. *Mathematical Problems in Engineering* **4**, 323–332. MR2109721

Nadarajah, S. and Kotz S. (2006). The beta exponential distribution. *Reliability Engineering and System Safety* **91**, 689–697.

Prentice, R. L. (1976). A generalization of the probit and logit models for dose response curves. *Biometrics* **32**, 761–768.

Departamento de Estatística
Universidade Federal de Pernambuco
Av. Luiz Freire s/n
50740-540 – Recife, PE
Brazil
E-mail: alice.lm@hotmail.com
gausscordeiro@uol.com.br
audrey@de.ufpe.br