A Four-Parameter Generalized Logistic Distribution

Alice Lemos Morais\textsuperscript{a} and Gauss Moutinho Cordeiro\textsuperscript{b}

\textsuperscript{a}Departamento de Estatística, Universidade Federal de Sergipe, Jardim Rosa Else, 49100-000 – São Cristóvão, SE, Brazil (alice.lm@hotmail.com)

\textsuperscript{b}Departamento de Estatística, Universidade Federal Rural de Pernambuco, Rua Dom Manuel de Medeiros s/n, 50171-900 – Recife, PE, Brazil (gauss@deinfo.ufrpe.br)

Abstract

For the first time, we introduce the beta generalized logistic distribution which is obtained by compounding the beta and generalized logistic distributions. The shape of the new distribution is quite flexible, specially the skewness and the tail weights, due to the two extra shape parameters. We obtain general expansions for the moments and quantile functions. The estimation of the parameters is investigated by maximum likelihood. Some related distributions are studied in detail. An application to a real data set is given to show the flexibility and potentiality of the new distribution.

Keywords: Beta distribution; Generalized logistic distribution; Maximum likelihood; Order Statistic.

1 Introduction

Because of their flexibility, much attention has been given to the study of generalized distributions in recent times. Prentice (1976) [10] 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 (pdf) of the GLIV distribution, say GLIV(p,q), is given by

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

The cumulative distribution function (cdf) \(G_{p,q}(x)\) corresponding to (1) is

\[
G_{p,q}(x) = I_{\frac{x}{1+e^{-x}}}(p, q), \quad x \in \mathbb{R}, \ p > 0, \ q > 0, \tag{2}
\]
where $I_x(a, b) = B_x(a, b)/B(a, b)$ is the incomplete beta function ratio, $B(a, b)$ is the beta function and $B_x(a, b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$ is the incomplete beta function. The moment generating function (mgf) corresponding to (1) is $M_{p, q}(t) = \Gamma(p+t)\Gamma(q-t)/[\Gamma(p)\Gamma(q)]$ defined for $-p < t < q$.

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 (similar to that of the normal distribution) makes it simpler and also profitable on suitable occasions to replace the normal by the logistic distribution with negligible errors in the respective theories. 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 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 article, we mainly consider a generalization of the GLIV distribution by introducing two extra shape parameters which leads to a new distribution so-called the beta generalized logistic (BGL) distribution. The role of the two additional parameters is to introduce skewness and to vary tail weights. They provide greater flexibility in the form of the distribution and consequently in modeling observed data. It may be mentioned that although several skewed distribution functions exist on the positive real axis, but 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 rest of the paper 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 as linear combinations of GLIV densities are derived in Section 4. Expansions for the quantile function and moment generating function (mgf) are presented in Section 5. Section 6 is devoted to mean deviations about the mean and the median and to the Bonferroni and Lorenz curves. Expansions for the order statistics and their moments are given in Section 7. Maximum likelihood estimation and the observed information matrix are determined 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 cdf of a parent random variable. The method to generalize distributions discussed here consists to define a new cdf $F(x)$ from $G(x)$ by

$$F(x) = I_{G(x)}(a, b) = \frac{1}{B(a, b)} \int_0^{G(x)} w^{a-1} (1 - w)^{b-1} dw, \quad a > 0, \ b > 0,$$  \hfill (3)

It is easy to see that $F(x)$ coincides with $G(x)$ when $a = b = 1$.

The BGL distribution is obtained by setting the GLIV cdf (2) in (3). The BGL cdf is then

$$F(x) = I_{\frac{1}{1+e^{-x}}(p,q)}(a, b) = \frac{1}{B(a, b)} \int_0^{\frac{1}{1+e^{-x}}(p,q)} w^{a-1} (1 - w)^{b-1} dw,$$  \hfill (4)

where $x \in \mathbb{R}$ and all parameters $a$, $b$, $p$ and $q$ are real positive.

We can note that the GLIV cdf (2) is also presented in the form (3), whereas $(1+e^{-x})^{-1}$ is the cdf of the standard logistic distribution. From (3) 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 written as

$$f(x) = \frac{1}{B(p, q)B(a, b)(1+e^{-x})^{p+q}} \left[ I_{\frac{1}{1+e^{-x}}(p, q)} \right]^{a-1} \left[ I_{\frac{1}{1+e^{-x}}(q, p)} \right]^{b-1},$$  \hfill (5)

A random variable $X$ having density function (5) is denoted by BGL($a, b, p, q$). Figure 1 plots the shapes of the BGL distribution for selected values of $a$ and $b$ by fixing $p = 0.2$ and $q = 1.5$. The BGL distribution is symmetric for $p = q$ and $a = b$. From (4) we have that the BGL($m, n, 1, 1$) and BGL($1, 1, m, n$) distributions are identical.

![Figure 1: The BGL density function for some values of $a$ and $b$ with $p = 0.2$ and $q = 1.5.$](image-url)
3 Related Distributions

Evidently, the density function (5) does not involve any complicated function but generalizes a few interesting distributions. We now present the distribution of some functions of a BGL random variable.

3.1 Type I Beta Generalized Logistic distribution

The type I generalized logistic (GLI) distribution is a special case of the GLIV distribution when $q = 1$. This property extends to the BGL distribution, i.e. the type I beta generalized logistic (BGLI) distribution is a special sub-model of the BGL distribution when $q = 1$. Then, the BGLI density function is given by

$$f(x) = \frac{pe^{-x}}{B(a,b)} \left[ (1 + e^{-ax})^p - 1 \right]^{b-1} \frac{1}{(1 + e^{-x})^{a+pb}}, \quad x \in \mathbb{R}, \ a, b, p > 0.$$ 

![Figure 2: The BGLI density function for some values of $a$ and $b$ with $p = 0.2$.](image)

A random variable $X$ having a BGLI distribution with parameters $a$, $b$ and $p$ is denoted by $X \sim BGLI(a,b,p)$. We can verify that if the parameter $p = 1$, the BGLI($a,b,p$) distribution coincides with the GLIV($a,b$) distribution. In fact, when $p = 1$, the baseline distribution of the BGLI distribution reduces to the standard logistic which is the baseline distribution of the GLI distribution. Figure 2 plots some special shapes of the BGLI distribution for selected values of $a$ and $b$ with $p = 0.2$.

We generate a random variable $X$ having the BGLI($a,b,p$) distribution from a random variate $V$ following a beta distribution with parameters $a$ and $b$ by $X = - \log(V^{-1/p} - 1)$. Other properties of this distribution will be omitted here because they can be obtained from the BGLIV distribution by setting $q = 1$. 


3.2 Type II Beta Generalized Logistic distribution

The type II beta generalized logistic (BGLII) distribution is also a special case of the BGL distribution when \( p = 1 \). This distribution can also be obtained through the transformation \( X = -Y \) and \( Y \) following the BGLI distribution. The BGLII density function is

\[
    f(x) = \frac{q e^{-bx}}{B(a, b)(1 + e^{-x})^{q b + 1}} \left[ 1 - \frac{e^{-qx}}{(1 + e^{-x})^{q b + 1}} \right], \quad x \in \mathbb{R}, \ a, b, q > 0.
\]

Figure 3: The BGLII density function for some values of \( a \) and \( b \) with \( p = 0.2 \).

A random variable \( X \) following the BGLII distribution with parameters \( a, b \) and \( q \) will be denoted by \( X \sim \text{BGLII}(a, b, q) \). Figure 3 plots some special shapes of the BGLII distribution for \( q = 0.5 \) and some values of \( a \) and \( b \).

3.3 Type III Beta Generalized Logistic

The type III beta generalized logistic (BGLIII) distribution is a special case of the BGL model when \( p = q \). The BGLIII density function is given by

\[
    f(x) = \frac{B(p, p)^{1-a-b}}{B(a, b)} \frac{e^{-px}}{(1 + e^{-x})^{2p}} \left[ B_{\frac{1}{1+e^{-x}}}(p, p) \right]^{a-1} \left[ B_{\frac{1}{1+e^{-x}}}(p, p) \right]^{b-1}, \quad x \in \mathbb{R}, \ a, b, p > 0.
\]

A random variable \( X \) following the BGLIII distribution with parameters \( a, b \) and \( p \) will be denoted by \( X \sim \text{BGLIII}(a, b, p) \). This distribution is symmetric when \( a = b \). Figure 4 plots some special shapes of the BGLIII distribution for selected values of \( a \) and \( b \) and \( p = 0.5 \).
3.4 Beta Beta Prime

Let $Y$ be a random variable following the BGL$(a,b,p,q)$ distribution. Thus, $X = e^{-Y}$ follows the beta beta prime (BBP) distribution with parameters $a$, $b$, $p$ and $q$. The BBP density function with support $(0, \infty)$ is given by

$$f(x) = \frac{B(p,q)^{1-a-b}}{B(a,b)} \frac{x^{p-1}}{(1+x)^{p+q}} \left[ B \frac{x}{1+x} (q,p) \right]^{(a-1)} \left[ B \frac{1}{x+1} (p,q) \right]^{(b-1)}, \ a, b, p, q, x > 0.$$
A random variable \( X \) following the BBP distribution with parameters \( a, b, p \) and \( q \) will be denoted by \( X \sim \text{BBP}(b, a, p, q). \) Figure 5 plots the BBP distribution for selected values of \( a, b, p \) and \( q. \) The associated hazard function is given by

\[
h(x) = \frac{B(p, q)^{1-a-b}}{B_{\frac{1}{1+x}}(q, p)} \frac{x^{q-1}}{(1 + x)^{p+q}} \left[ B_{\frac{1}{1+x}}(q, p) \right]^{(a-1)} \left[ B_{\frac{1}{1+x}}(p, q) \right]^{(b-1)}, \quad a, b, p, q, x > 0.
\]

### 3.5 Beta F

The density function of the F distribution is

\[
g(x) = \frac{(q/p)^{q/2}}{B(p/2, q/2)} x^{q/2-1} \frac{(1 + (q/p)x)^{(p+q)/2}}{(1 + (q/p)x)^{p/2}}, \quad p, q, x > 0
\]

and its cdf can be expressed as

\[
G(x) = \frac{1}{B(p/2, q/2)} B_{\frac{q/2}{1+(q/p)x}}(q/2, p/2).
\]

The beta F (BF) distribution is given by (for \( x > 0 \))

\[
f(x) = \frac{B(a, b)^{-1}}{B(p/2, q/2)} \frac{(q/p)^{q/2} x^{a-1}}{(1 + (q/p)x)^{p+q/2}} \left[ I_{\frac{q/2}{1+(q/p)x}}(q/2, p/2) \right]^{-a} \left[ I_{\frac{1}{1+(q/p)x}}(p/2, q/2) \right]^{-b},
\]

where \( a, b, p, q \) are positive real numbers. A random variable \( X \) following a BF distribution with these parameters will be denoted by \( X \sim \text{BF}(a, b, p, q). \) It is easy to verify that if \( Y \sim \text{BGL}(a, b, p, q), \) then \( F = \frac{1}{q} e^{-Y} \sim \text{BF}(a, b, 2p, 2q). \) Figure 6 plots the BF distribution for selected values of \( a \) and \( b \) and \( p = 2 \) and \( q = 10. \)

![Figure 6: The BF density function for some values of b and a and p = 2 and q = 10.](image-url)
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. For $b > 0$ real non-integer, the power series for $(1 - w)^{b-1}$ in (3) yields

$$
\int_0^x w^{a-1} (1 - w)^{b-1} \, dw = \sum_{j=0}^{\infty} \frac{(-1)^j \binom{b-1}{j}}{(a+j)} x^{a+j}, \tag{6}
$$

where the binomial coefficient is defined for any real. If $b$ is an integer, the index $j$ in (6) stops at $b - 1$. From (2) and (6), we can express the BGL cumulative function as

$$
F(x) = \sum_{r=0}^{\infty} w_r I_{\frac{1}{1+x^{-p}}} (p, q)^{a+r}, \tag{7}
$$

where the coefficients are

$$
w_j = w_j(a, b) = \frac{(-1)^j \binom{b-1}{j}}{B(a, b)(a+j)}. \tag{8}
$$

If $a$ is an integer, equation (7) provides the BGL cdf as an infinite power series expansion of GLIV cdf’s. Otherwise, if $a$ is real non-integer, we can expand $I_{\frac{1}{1+x^{-p}}} (p, q)^{a+r}$ to obtain the BGL cdf as an infinite power series of GLIV cdf’s. We have

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

and then

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

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

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

where

$$
s_k(\alpha) = \sum_{j=k}^{\infty} (-1)^{k+j} \binom{\alpha}{j} \binom{j}{k}. \tag{8}
$$

Hence, from (7), we obtain

$$
F(x) = \sum_{r=0}^{\infty} t_r I_{\frac{1}{1+x^{-p}}} (p, q)^r, \tag{9}
$$
where
\[ t_r = t_r(a, b) = \sum_{l=0}^{\infty} w_l s_r(a + l), \]
and \( s_r(a + j) \) comes from (8). 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 (7) (for \( a > 0 \) integer)
\[ f(x) = g_{p,q}(x) \sum_{r=0}^{\infty} (a + r) w_r I_{\frac{1}{1+e^{-x}}}(p, q)^{a+r-1} \quad (10) \]
and of (9) (for \( a > 0 \) real non-integer)
\[ f(x) = g_{p,q}(x) \sum_{r=0}^{\infty} (r + 1) t_{r+1} I_{\frac{1}{1+e^{-x}}}(p, q)^r. \quad (11) \]
In any case of equations (10) and (11), we need to obtain a power series expansion for \( I_{\frac{1}{1+e^{-x}}}(p, q)^r \). We can use the incomplete beta function expansion for \( q > 0 \) real non-integer
\[ I_x(p, q) = \frac{x^p}{B(p, q)} \sum_{m=0}^{\infty} \frac{(1-q)m^m}{(p+m)m!}, \]
where \((f)_k = \Gamma(f + k)/\Gamma(f)\). First, we need an expansion for \( I_{\frac{1}{1+e^{-x}}}(p, q)^r \). We have
\[ I_{\frac{1}{1+e^{-x}}}(p, q)^r = \frac{1}{B(p, q)^r(1 + e^{-x})^{pr}} \left( \sum_{m=0}^{\infty} d_m y^m \right)^r = \frac{1}{B(p, q)^r(1 + e^{-x})^{pr}} \sum_{m=0}^{\infty} c_{r,m} y^m, \quad (12) \]
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 nonnegative integer power (see Gradshteyn and Ryzhik, 2007 [4]) given by the recurrence equation
\[ c_{r,m} = (md_0)^{-1} \sum_{j=1}^{i} (rj - m + j)d_j c_{r,m-j}, \quad (13) \]
and \( c_{0,m} = d_0^m \). Hence, for \( a > 0 \) integer, we can write from (10) and (12)
\[ f(x) = \sum_{r,m=0}^{\infty} \rho_{int}(r, m) g_{p(a+r)+m,q}(x), \quad (14) \]
where
\[ \rho_{int}(r, m) = \frac{(a + r) w_r c_{a+r-1,m} B(p(a + r) + m, q)}{B(p, q)^{a+r}}. \]
In a similar way, for a real non-integer, we obtain from (11) and (12)

\[ f(x) = \sum_{r,m=0}^{\infty} \rho_{\text{real}}(r,m) g_{p(r+1)+m,q}(x), \]  

(15)

where

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

Equations (14) and (15) are the main results of this section. They show that the BGL density function can be expressed 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 Moment Generating Functions

The quantile function \( \gamma \) \((x(\gamma))\) of the BGL distribution follows from (3) as

\[ x(\gamma) = -\log \left( (Q_{p,q}(Q_{a,b}(\gamma)))^{-1} - 1 \right), \]  

(16)

where \( Q_{p,q}(u) \) denotes the quantile function of the beta distribution with parameters \( p \) and \( q \).

The following expansion for the inverse of the beta incomplete function \( Q_{p,q}(u) \) can be found in wolfram website\(^1\)

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

where \( w = [puB(p,b)]^{1/p} \) for \( p > 0 \).

The mgf of the BGL distribution can be determined from equations (14) and (15). For \( a > 0 \) integer, (14) gives

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

and for \( a > 0 \) non-integer, (15) yields

\(^1\)http://functions.wolfram.com/06.23.06.0004.01
\[ M_X(t) = \frac{\Gamma(q-t)}{\Gamma(q)} \sum_{r,m=0}^{\infty} \rho_{\text{real}}(r,m) \frac{\Gamma(p(r+1)+m+t)}{\Gamma(p(r+1)+m)}. \]

The \( r \)th moment is obtained by the \( r \)th 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 7 and 8, respectively.

![Skewness vs a and b](image)

Figure 7: Skewness as function of \( a \) and \( b \).

### 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 cdf \( F(x) \), 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 the solution of the non-linear equation

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

Defining the integral

\[
J(s) = \int_{-\infty}^{s} x f(x)dx = \frac{1}{B(a,b)} \int_{-\infty}^{s} x g(x) G(x)^{a-1} \{1 - G(x)\}^{b-1},
\]
these measures can be calculated from
\[
\delta_1 = 2\nu F(\nu) - 2J(\nu) \quad \text{and} \quad \delta_2 = E(X) - 2J(m),
\]
where \( F(\nu) \) is easily obtained from (3). We now derive a formula to obtain the integral \( J(s) \). For \( a > 0 \) integer, we have from (7)
\[
J(s) = \sum_{r,m,n=0}^{\infty} \frac{\rho_{\text{int}}(r,m)}{B(p(a+r) + m,q)(n+1)} \left[ B(p(a+r) + m+n+1,q)G_{p(a+r)+m+n+1,q}(s) - B(p(a+r) + m,q + n + 1)G_{p(a+r)+m,q+n+1}(s) \right].
\]
For \( a > 0 \) non-integer, we obtain from (9)
\[
J(s) = \sum_{r,m,n=0}^{\infty} \frac{\rho_{\text{int}}(r,m)}{B(p(r+1) + m,q)(n+1)} \left[ B(p(r+1) + m+n+1,q)G_{p(r+1)+m+n+1,q}(s) - B(p(r+1) + m,q + n + 1)G_{p(r+1)+m,q+n+1}(s) \right].
\]

The above formulae can be used to determine Bonferroni and Lorenz curves which have applications not only in economics to study income and poverty, but also in other fields like reliability, demography, insurance and medicine. They are defined by
\[
B(p) = \frac{J(q)}{p\nu}, \quad \text{and} \quad L(p) = \frac{J(q)}{\nu},
\]
respectively, where \( \nu = E(X) \) and \( q = F^{-1}(p) \) is calculated by (16) for given \( p \).
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 now derive an explicit expression for the density of the \( r \)th 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

\[
f_{i:n}(x) = \frac{1}{B(i, n-i+1)} f(x) F(x)^{i-1} \{1 - F(x)\}^{n-i},
\]

for \( i = 1, \ldots, n \). Then,

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

(19)

For a beta generalized model defined for the parent pdf \( g(x) \) and cdf \( G(x) \), \( f_{i:n}(x) \) can be written

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

(20)

Expanding \( F(x)^{i+j-1} \) similarly to (12) we have

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

(21)

where the coefficients \( c_{r,m} \) are given in (13) with \( d_m = w_m \). Replacing (21) in (20), we have

\[
f_{i:n}(x) = \frac{1}{B(a, b)B(i, n-i+1)} \sum_{k=0}^{\infty} \sum_{j=0}^{n-i} (-1)^j c_{i+j-1,k} \binom{n-i}{j} f_g(a, b, p, q, x).
\]

(22)

where \( f_g(x) \) is the density of the BGL\((m, b, p, q)\). For \( a > 0 \) integer, we can write from (14) and (22)

\[
f_{i:n}(x) = \sum_{r,m,k=0}^{\infty} \sum_{j=0}^{n-i} \tau_{int}(r, m, k, j) g_p(a, b, p, q, x),
\]

(23)

where

\[
\tau_{int}(r, m, k, j) = \binom{n-i}{j} \left( -1 \right)^j w_r c_{i+j-1,k} e_{a(i+j)+k+r-1,m} B(p(a, b, p, q, x)+k+r, m, q) \left( a(i+j)+k+r \right)^{-1} B(p, q, a(i+j)+k+r, m, q) \left( a(i+j)+k+r \right)^{-1} B(a, b, p, q, x).
\]

and \( e_{r,m} = c_{r,m} \) with \( d_m = \frac{(1-a)^m}{(p+m)^m} \). For \( a > 0 \) non-integer,

\[
f_{i:n}(x) = \sum_{r,m,k=0}^{\infty} \sum_{j=0}^{n-i} \tau_{real}(r, m, k, j) g_p(a, b, p, q, x),
\]

(24)
where
\[ \tau_{\text{real}}(r, m, k, j) = \left( \frac{n - i}{j} \right) \frac{(-1)^j t_{r+1}(a(i+j) + k, b)c_{i+j-1, k}e_{r,v}B(p(r+1) + m, q)}{(r+1)^{-1}B(p, q)^{r+1}B(a, b)B(i, n-i+1)}. \]

The \( r \)th moment of the \( i \)th order statistic \( X_{i,n} \) can be obtained directly from (23) and (24). For \( a > 0 \) integer, (23) gives
\[ E(X^r_{i,n}) = \sum_{r,m,k=0}^{\infty} \sum_{j=0}^{n-i} \tau_{\text{int}}(r, m, k, j) E(Y^r_{p(a(i+j)+k+r)+m,q}), \]

and for \( a > 0 \) non-integer, (24) yields
\[ E(X^r_{i,n}) = \sum_{r,m,k=0}^{\infty} \sum_{j=0}^{n-i} \tau_{\text{real}}(r, m, k, j) E(Y^r_{p(r+1)+m,q}), \]

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

8 Estimation

Let \( x_1, \ldots, x_n \) be an independent random sample from the BGL distribution. The total log-likelihood is given by
\[ \ell = \ell(a, b, p, q; x) = -n \log B(a, b) + (1 - a - b)n \log B(p, q) - (p + q) \sum_{i=1}^{n} \log(1 + e^{-x_i}) + (a - 1) \sum_{i=1}^{n} \log B \left( \frac{1}{1+e^{-x_i}} \right) (p, q) + (b - 1) \sum_{i=1}^{n} \log B \left( \frac{e^{-x_i}}{1+e^{-x_i}} \right) (q, p) - q \sum_{i=1}^{n} x_i. \]

The score function has the following components
\[ \frac{\partial \ell}{\partial a} = n \psi(a + b) - n \psi(a) - n \log B(p, q) + \sum_{i=1}^{n} \log B \left( \frac{1}{1+e^{-x_i}} \right) (p, q), \]
\[ \frac{\partial \ell}{\partial b} = n \psi(a + b) - n \psi(b) - n \log B(p, q) + \sum_{i=1}^{n} \log B \left( \frac{e^{-x_i}}{1+e^{-x_i}} \right) (q, p), \]
\[ \frac{\partial \ell}{\partial p} = (1 - a - b)n \{ \psi(p) - \psi(p + q) \} - \sum_{i=1}^{n} \log(1 + e^{-x_i}) + (a - 1) \sum_{i=1}^{n} \log B \left( \frac{1}{1+e^{-x_i}} \right) (p, q) + (b - 1) \sum_{i=1}^{n} \log B \left( \frac{e^{-x_i}}{1+e^{-x_i}} \right) (q, p), \]
\[ \frac{\partial \ell}{\partial q} = (1 - a - b)n \{ \psi(q) - \psi(p + q) \} - \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \log(1 + e^{-x_i}) + (a - 1) \sum_{i=1}^{n} \log B \left( \frac{1}{1+e^{-x_i}} \right) (p, q) + (b - 1) \sum_{i=1}^{n} \log B \left( \frac{e^{-x_i}}{1+e^{-x_i}} \right) (q, p), \]
where $\psi(\cdot)$ is the digamma function.

Solving the system of nonlinear equations $\nabla \ell = 0$, the maximum likelihood estimates (MLEs) of the parameters can be obtained. The observed information matrix, say $J(\theta)$, used in the estimation equations and for making interval inference has elements given by

\[
\frac{\partial^2 \ell}{\partial a^2} = n\psi'(a + b) - n\psi'(a), \\
\frac{\partial^2 \ell}{\partial a \partial b} = n\psi(a + b), \\
\frac{\partial^2 \ell}{\partial a \partial p} = n\psi(p + q) - n\psi(p) + \frac{\partial}{\partial p} \sum_{i=1}^{n} \log B \frac{1}{1 + e^{-x_i}}(p, q), \\
\frac{\partial^2 \ell}{\partial a \partial q} = n\psi(p + q) - n\psi(q) + \frac{\partial}{\partial q} \sum_{i=1}^{n} \log B \frac{1}{1 + e^{-x_i}}(p, q), \\
\frac{\partial^2 \ell}{\partial b^2} = n\psi'(a + b) - n\psi'(b), \\
\frac{\partial^2 \ell}{\partial b \partial p} = n\psi(p + q) - n\psi(p) + \frac{\partial}{\partial p} \sum_{i=1}^{n} \log B \frac{1}{1 + e^{-x_i}}(q, p), \\
\frac{\partial^2 \ell}{\partial b \partial q} = n\psi(p + q) - n\psi(q) + \frac{\partial}{\partial q} \sum_{i=1}^{n} \log B \frac{1}{1 + e^{-x_i}}(q, p), \\
\frac{\partial^2 \ell}{\partial p^2} = (1 - a - b)n\{\psi'(p) - \psi'(p + q)\} + (a - 1) \frac{\partial^2}{\partial p^2} \sum_{i=1}^{n} \log B \frac{1}{1 + e^{-x_i}}(p, q), \\
\frac{\partial^2 \ell}{\partial p \partial q} = -(1 - a - b)n\psi'(p + q) + (a - 1) \frac{\partial^2}{\partial p \partial q} \sum_{i=1}^{n} \log B \frac{1}{1 + e^{-x_i}}(p, q), \\
\frac{\partial^2 \ell}{\partial q^2} = (1 - a - b)n\{\psi'(q) - \psi'(p + q)\} + (a - 1) \frac{\partial^2}{\partial q^2} \sum_{i=1}^{n} \log B \frac{1}{1 + e^{-x_i}}(p, q),
\]

where $\psi'(\cdot)$ is the trigamma function.

Let $\theta = (a, b, p, q)^T$ be the parameter vector of the BGL distribution. Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, we can consider the following approximation

$$\sqrt{n}(\hat{\theta} - \theta) \sim N_4(0, J(\theta)^{-1}).$$
The approximate multivariate normal $N_d(0, J_n(\hat{\theta})^{-1})$ distribution of $\hat{\theta}$ can be used to construct approximate confidence intervals and confidence regions for the parameters and for the hazard and survival functions. An asymptotic confidence interval with significance level $\gamma$ for each parameter $\theta_r$ is

$$ACI(\theta_r, 100(1 - \gamma)\%) = (\hat{\theta}_r - z_{\gamma/2} \sqrt{k^{\theta_r, \theta_r}}, \hat{\theta}_r + z_{\gamma/2} \sqrt{k^{\theta_r, \theta_r}}),$$

where $k^{\theta_r, \theta_r}$ is the $r$th diagonal element of $K_n(\theta)^{-1}$ for $r = 1, \ldots, 4$ and $z_{\gamma/2}$ is the quantile $1 - \gamma/2$ of the standard normal distribution.

The likelihood ratio (LR) statistic is useful for testing goodness of fit of the BGL distribution and for comparing this distribution with some of its special sub-models. If we consider the partition $\theta = (\theta^T_1, \theta^T_2)^T$, 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 via LR tests. 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 $\theta$ under $H_A$ and $H_0$, respectively. Under the null hypothesis, $w \overset{d}{\to} \chi^2_q$, where $q$ is the dimension of the vector $\theta_1$ of interest. The LR test rejects $H_0$ if $w > \xi_{\gamma}$, where $\xi_{\gamma}$ denotes the upper $100\gamma\%$ point of the $\chi^2_q$ 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 [http://www.ibge.gov.br/series_estatisticas/exibedados.php?dnivel=BR&dsérie=PRECO101](http://www.ibge.gov.br/series_estatisticas/exibedados.php?dnivel=BR&dsérie=PRECO101).

We fit the BGL model and some of its special sub-models discussed in Section 3 to these data. The MLES of the model parameters, the maximized log-likelihoods ($\hat{\ell}$) and the p-values for the log-likelihood ratio tests are listed in Table 1.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{a}$</th>
<th>$\hat{b}$</th>
<th>$\hat{p}$</th>
<th>$\hat{q}$</th>
<th>$\ell$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BGL</td>
<td>179.92</td>
<td>0.39</td>
<td>0.92</td>
<td>6.96</td>
<td>-118.97</td>
<td>-</td>
</tr>
<tr>
<td>GLIV</td>
<td>-</td>
<td>-</td>
<td>9.43</td>
<td>5.18</td>
<td>-134.01</td>
<td>$2.9 \times 10^{-07}$</td>
</tr>
<tr>
<td>BGLI</td>
<td>9.34</td>
<td>5.21</td>
<td>1.02</td>
<td>-</td>
<td>-134.05</td>
<td>$3.9 \times 10^{-08}$</td>
</tr>
<tr>
<td>BGLII</td>
<td>19.03</td>
<td>1.03</td>
<td>-</td>
<td>3.21</td>
<td>-125.20</td>
<td>$4.1 \times 10^{-04}$</td>
</tr>
<tr>
<td>BGLIII</td>
<td>1.59</td>
<td>0.26</td>
<td>16.27</td>
<td>16.27</td>
<td>-125.74</td>
<td>$2.3 \times 10^{-04}$</td>
</tr>
</tbody>
</table>

Table 1: MLEs for the BGL, GLIV, BGLI, BGLII and BGLIII distributions.

Clearly, for the usual significance levels in all tests, we accept the BGL model. These results illustrate the potentiality of BGL distribution and the necessity to introduce 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 by (the well known technique) 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, Bonferroni and Lorenz curves, order statistics, their moments and L-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 logistic type models widely used data analysis.

References


