The McDonald Inverted Beta Distribution: A Good Alternative to Lifetime Data

Gauss M. Cordeiro

Departamento de Estatística, Universidade Federal de Pernambuco, Recife/PE, Brazil

Artur J. Lemonte

Departamento de Estatística, Universidade de São Paulo, São Paulo/SP, Brazil

Abstract

We introduce a five-parameter continuous model, called the McDonald inverted beta distribution, to extend the two-parameter inverted beta distribution and provide new four- and three-parameter sub-models. We give a mathematical treatment of the new distribution including expansions for the density function, moments, generating and quantile functions, mean deviations, entropy and reliability. The model parameters are estimated by maximum likelihood and the observed information matrix is derived. An application of the new model to real data shows that it can give consistently a better fit than other important lifetime models.

Keywords: Inverted beta distribution; Maximum likelihood estimation; McDonald distribution; Moment; Moment generating function.

1 Introduction

The beta distribution with support in the standard unit interval $(0, 1)$ has been utilized extensively in statistical theory and practice for over one hundred years. It is very versatile and a variety of uncertainties can be usefully modeled by this distribution, since it can take an amazingly great variety of forms depending on the values of its parameters. On the other hand, the inverted beta (IB) distribution with support in $(0, \infty)$ can be used to model positive real data. It is also known as the beta prime distribution or beta distribution of the second kind. Its probability density function (pdf) with two positive parameters $\alpha > 0$ and $\beta > 0$ is given by

$$g_{\alpha,\beta}(x) = \frac{x^{\alpha-1}}{B(\alpha,\beta)\,(1+x)^{\alpha+\beta}}, \qquad x > 0, \tag{1}$$

where $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$ is the beta function and $\Gamma(\alpha) = \int_0^\infty w^{\alpha-1}\,\mathrm{e}^{-w}dw$ is the gamma function. The cumulative distribution function (cdf) corresponding to (1) is

$$G_{\alpha,\beta}(x) = I_{\frac{x}{1+x}}(\alpha, \beta), \qquad x > 0, \tag{2}$$

where $I_y(p,q) = B_y(p,q)/B(p,q)$ is the incomplete beta function ratio and $B_y(p,q) = \int_0^y \omega^{p-1}(1-\omega)^{q-1}d\omega$ is the incomplete beta function. The cdf (2) can be expressed in terms of the hypergeometric function as

$$G_{\alpha,\beta}(x) = \frac{x^\alpha}{\alpha\, B(\alpha,\beta)}\, {}_2F_1\left(\alpha, \alpha+\beta; \alpha+1; -x\right),$$

where

$${}_2F_1(p,q;r;y) = \frac{\Gamma(r)}{\Gamma(p)\,\Gamma(q)} \sum_{j=0}^{\infty} \frac{\Gamma(p+j)\,\Gamma(q+j)}{\Gamma(r+j)}\, \frac{y^j}{j!}.$$

The hypergeometric function can be computed, for example, using the MATHEMATICA software. For example, ${}_2F_1(p,q;r;y)$ is obtained from MATHEMATICA as `HypergeometricPFQ[{p,q},{r},y]`. For $s < \beta$, the sth moment about zero associated with (1) is

$$\mathrm{E}(X^s) = \frac{B(\alpha+s, \beta-s)}{B(\alpha,\beta)}.$$

Also, for $s \in \mathbb{N}$ and $s < \beta$, this equation simplifies to $\mathrm{E}(X^s) = \prod_{i=1}^{s}(\alpha+i-1)/(\beta-i)$. The mean and variance of X for $\beta > 1$ and $\beta > 2$ are given by

$$\mathrm{E}(X) = \frac{\alpha}{\beta-1} \qquad \text{and} \qquad \mathrm{var}(X) = \frac{\alpha(\alpha+\beta-1)}{(\beta-2)(\beta-1)^2},$$

respectively. If V has the beta distribution with positive parameters α and β, then $X = V/(1-V)$ has the IB distribution (1). It also arises from a linear transformation of the F distribution.

The IB distribution has been studied by several authors. McDonald and Richards (1987a) discussed various properties of this distribution and obtain the maximum likelihood estimates (MLEs) of the model parameters. The behavior of its hazard ratio function has been examined by McDonald and Richards (1987b). Bookstaber and McDonald (1987) showed that this distribution is quite useful in the empirical estimation of security returns and in facilitating the development of option pricing models (and other models) that depend on the specification and mathematical manipulation of distributions. Mixtures of two IB distributions have been considered by McDonald and Butler (1987) who have applied it in the analysis of unemployment duration. McDonald and Butler (1990) have used this distribution while discussing regression models for positive random variables. Other applications in modeling insurance loss processes have been illustrated by Cummins et al. (1990). McDonald and Bookstaber (1991) have developed an option pricing formula based on this distribution that includes the widely used Black Scholes formula based on the assumption of log-normally distributed returns.

The generalized beta distribution of first kind (or, beta type I) may be characterized by the density function (McDonald, 1984)

$$h(x) = \frac{c}{B(ac^{-1}, b)}\, x^{a-1}(1-x^c)^{b-1}, \quad 0 < x < 1, \tag{3}$$

where $a > 0$, $b > 0$ and $c > 0$ are shape parameters. Two important special models are the beta and Kumaraswamy (Kumaraswamy, 1980) distributions defined from (3) for $c = 1$ and $a = c$, respectively.

The statistics literature is filled with hundreds of continuous univariate distributions that have been extensively used over the past decades for modeling data in several fields such as environmental and medical sciences, engineering, demography, biological studies, actuarial, economics, finance and insurance. However, in many applied areas such as lifetime analysis, finance and insurance, there is a clear need for extended forms of these distributions. Recent developments focus on new techniques for building meaningful distributions, including the two-piece approach introduced by Hansen (1994) and the generator approach pioneered by Eugene et al. (2002) and Jones (2004). For any continuous baseline cdf $G(x)$ with parameter vector $\boldsymbol{\tau}$ and density function $g(x)$, the cumulative function $F(x)$ of the McDonald-G (denoted with the prefix "McG" for short) distribution is defined by

$$F(x) = I_{G(x)^c}(ac^{-1}, b) = \frac{1}{B\left(ac^{-1}, b\right)} \int_0^{G(x)^c} \omega^{\frac{a}{c}-1} \left(1-\omega\right)^{b-1} d\omega, \tag{4}$$

where $a > 0$, $b > 0$ and $c > 0$ are additional shape parameters to those in $\boldsymbol{\tau}$ to govern skewness and to provide greater flexibility of its tails. The density function corresponding to (4) can be reduced to

$$f(x) = \frac{c}{B(ac^{-1}, b)}\, g(x)\, G(x)^{a-1} \left[1 - G(x)^c\right]^{b-1}. \tag{5}$$

Clearly, the McDonald density (3) is a basic exemplar of (5) for $G(x) = x$, $x \in (0, 1)$.

The class of distributions (5) includes two important special sub-classes: the beta generalized (BG) and Kumaraswamy generalized (KwG) distributions when $c = 1$ (Eugene et al., 2002) and $a = c$ (Cordeiro and de Castro, 2011), respectively. It follows from (5) that the McG distribution with baseline cdf $G(x)$ is the BG distribution with baseline cdf $G(x)^c$. This simple transformation may facilitate the derivation of some of its structural properties. The BG and KwG distributions can be limited in one aspect. They introduce only two additional shape parameters, whereas three may be required to control both tail weights and the distribution of weight in the center. Hence, the McDonald distribution (5) is a more flexible model since it has one more shape parameter than the classical beta or Kumaraswamy generators that can give additional control over both skewness and kurtosis.

Clearly, for $G(x) = x$, we obtain as simple sub-models the classical beta and Kumaraswamy distributions for $c = 1$ and $a = c$, respectively. The Kumaraswamy distribution is commonly termed the "minimax" distribution. Jones (2009) advocates its tractability, especially in simulations because its quantile function takes a simple form, and its pedagogical appeal relative to the classical beta distribution.

Equation (5) will be most tractable when both functions $G(x)$ and $g(x)$ have simple analytic expressions. Its major benefit is the ability of fitting skewed data that cannot be properly fitted by existing distributions. Let $Q_G(u)$ be the quantile function of the G distribution. Application of $X = Q_G(V^{1/c})$ to a beta random variable V with positive parameters a/c and b generates X with cumulative function (4). The cumulative function (4) can also be expressed in terms of the hypergeometric function as

$$F(x) = \frac{c\, G(x)^a}{a\, B(ac^{-1}, b)}\, {}_2F_1\left(ac^{-1}, 1-b; ac^{-1}+1; G(x)^a\right).$$

Thus, for any parent $G(x)$, the properties of $F(x)$ could, in principle, be obtained from the well established properties of the hypergeometric function (see Gradshteyn and Ryzhik, 2007).

In this note, we study some mathematical properties of a new five-parameter distribution called the McDonald inverted beta (McIB) distribution, which is defined from (5) by taking $G(x)$ and $g(x)$

to be the cdf and pdf of the IB distribution, respectively. We adopt a different approach to much of the literature so far: rather than considering the classical beta generator (Eugene et al., 2002) or the Kumaraswamy generator (Cordeiro and de Castro, 2011) applied to a baseline distribution, we propose a more flexible McDonald generator applied to the IB distribution. We also discuss maximum likelihood estimation of its parameters.

The article is outlined as follows. In Section 2, we define the McIB distribution. Section 3 provides a useful expansion for its density function. In Section 4, we obtain a simple expansion for the moments. Section 5 provides an expansion for the moment generating function (mgf). Section 6 deals with non-standard measures for the skewness and kurtosis. Mean deviations, Bonferroni and Lorenz curves, Rénvy entropy and reliability are investigated in Sections 7, 8 and 9, respectively. Maximum likelihood estimation is discussed in Section 10. An empirical application is presented and discussed in Section 11. Finally, Section 12 offers some concluding remarks.

2 The McIB Distribution

The McIB density function can be obtained from (5) as

$$f(x) = \frac{c\,x^{\alpha-1}}{B(\alpha,\beta)\,B(ac^{-1},b)\,(1+x)^{\alpha+\beta}}\, I_{\frac{x}{1+x}}(\alpha,\beta)^{a-1}\Big[1-I_{\frac{x}{1+x}}(\alpha,\beta)^{c}\Big]^{b-1}, \qquad x>0. \tag{6}$$

The cdf corresponding to (6) is given by $F(x) = I_{I_{\frac{x}{1+x}}(\alpha,\beta)^c}(ac^{-1},b)$, the survival function is $S(x) = 1 - I_{I_{\frac{x}{1+x}}(\alpha,\beta)^c}(ac^{-1},b)$ and the associated hazard rate function takes the form

$$r(x) = \frac{c\,x^{\alpha-1}}{B(\alpha,\beta)\,B(a,b)\,(1+x)^{\alpha+\beta}}\,\frac{I_{\frac{x}{1+x}}(\alpha,\beta)^{a-1}\Big[1-I_{\frac{x}{1+x}}(\alpha,\beta)^{c}\Big]^{b-1}}{\Big[1-I_{I_{\frac{x}{1+x}}(\alpha,\beta)^c}(ac^{-1},b)\Big]}. \tag{7}$$

The study of the new distribution is important since it includes as special sub-models some distributions not previously considered in the literature. In fact, the IB distribution (with parameters α and β) is clearly a basic exemplar for $a=b=c=1$. The beta IB (BIB) and Kumaraswamy IB (KwIB) distributions are new models when $c=1$ and $a=c$, respectively. For $b=c=1$, it leads to a new distribution refereed to as the exponentiated IB (EIB) distribution. The Lehmann type-II IB (LeIB) distribution arises with $a=c=1$. The McIB distribution can also be applied in engineering as the IB distribution and can be used to model reliability and survival problems. The proposed distribution allows for greater flexibility of its tails and can be widely applied in many areas.

Figures 1 and 2 illustrate some of the possible shapes of the density function (6) and hazard rate function (7), respectively, for selected parameter values. The density function and hazard rate function can take various forms depending on the parameter values.

Let $Q_{\alpha,\beta}(u)$ be the quantile function of the beta distribution with parameters α and β. The quantile function of the McIB(α,β,a,b,c) distribution, say $x=Q(u)$, can be easily obtained as

$$x = Q(u) = \frac{Q_{\alpha,\beta}\big(Q_{a/c,b}(u)^{1/c}\big)}{1-Q_{\alpha,\beta}\big(Q_{a/c,b}(u)^{1/c}\big)}. \tag{8}$$

This scheme is useful because of the existence of fast generators for beta random variables in most statistical packages.

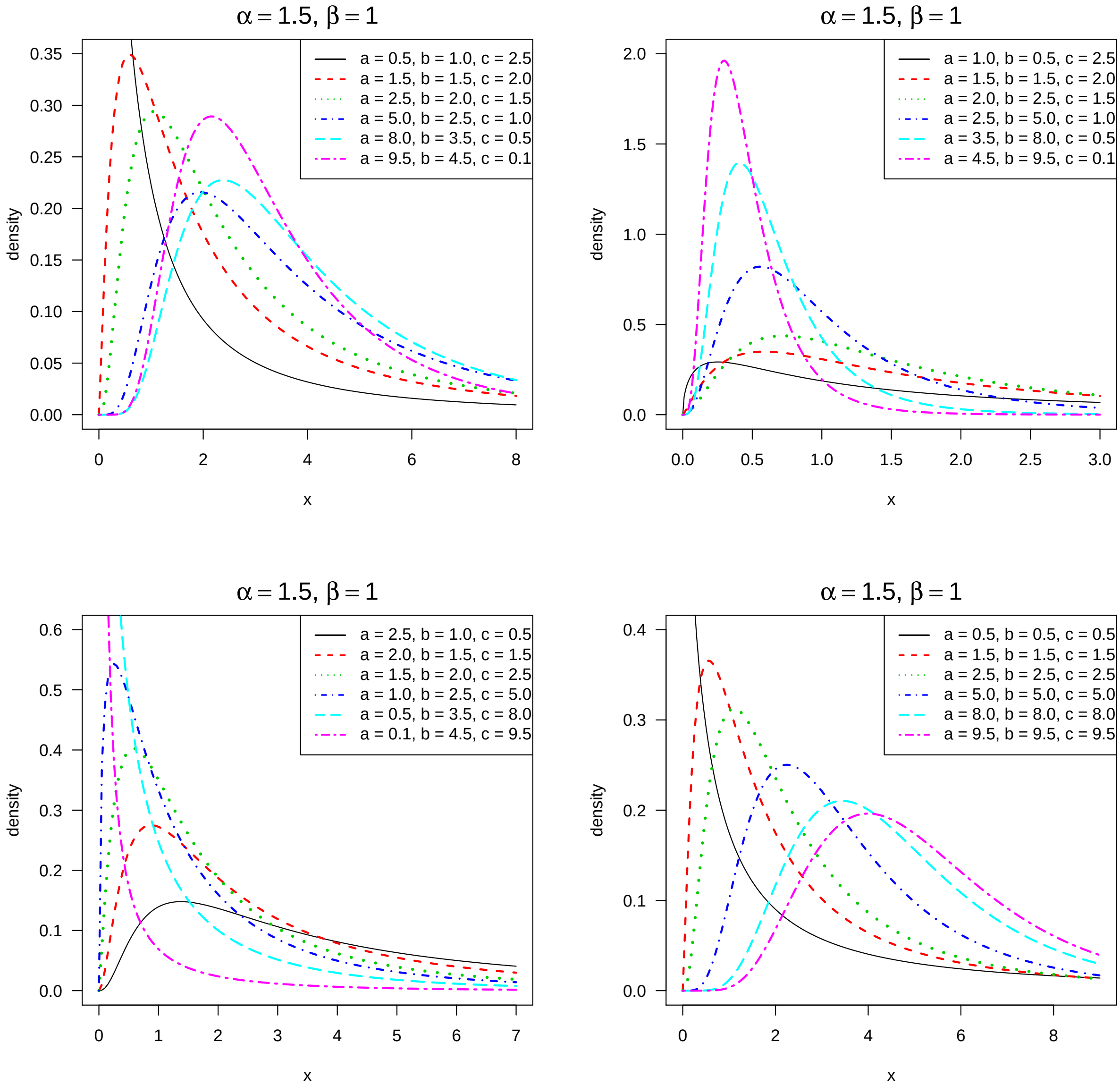

Figure 1: Plots of the density function (6) for some parameter values.

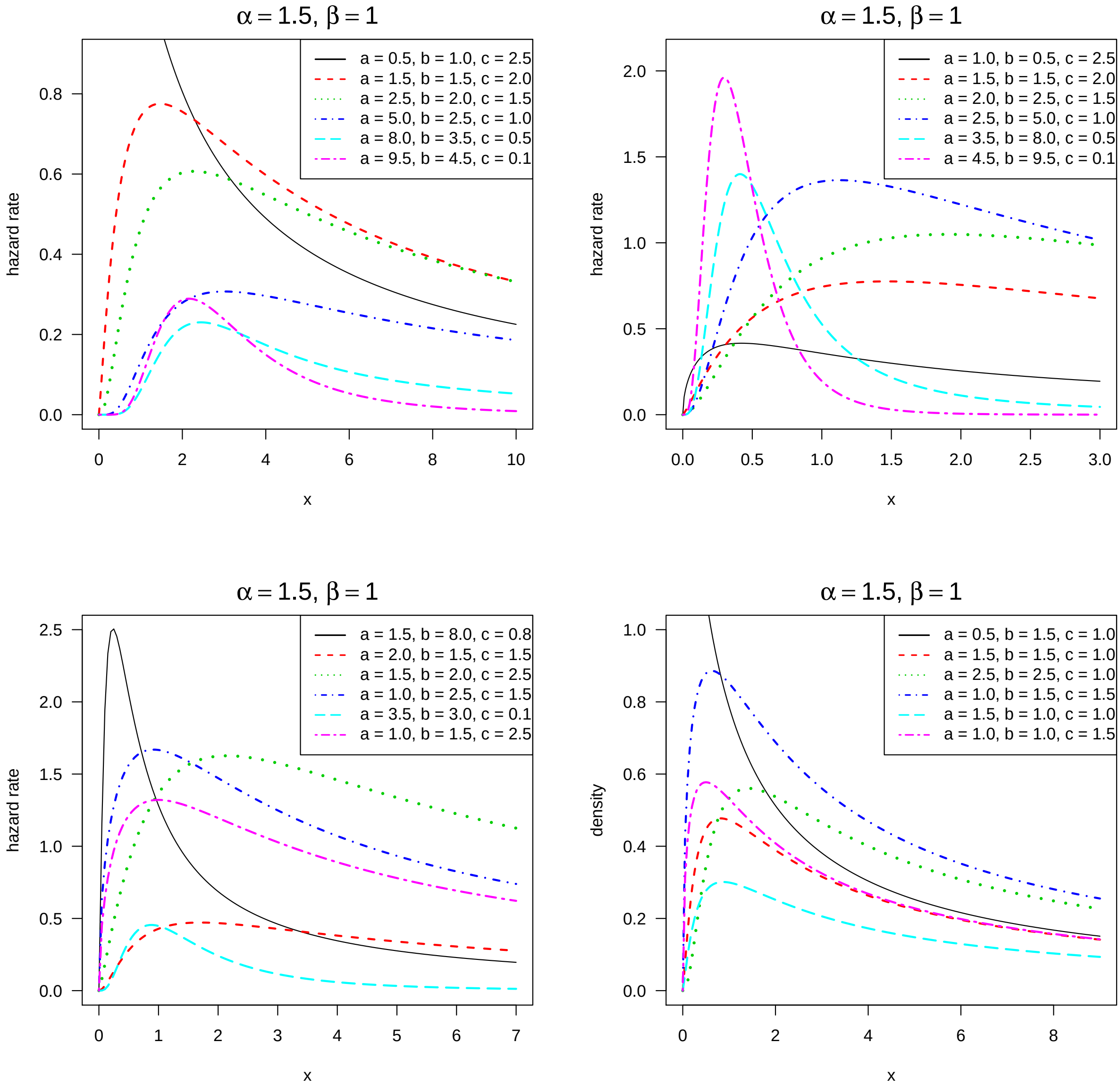

Figure 2: Plots of the hazard rate function (7) for some parameter values.

3 Density Function Expansion

We start this section by stating some useful expansions for the McG density function and, for brevity of notation, we shall drop the explicit reference to the parameter vector $\boldsymbol{\tau}$ in $G(x)$. A useful expansion for (5) can be derived as a linear combination of exponentiated-G distributions. For an arbitrary baseline G and $a > 0$, a random variable X having cdf and pdf given by

$$H_a(x) = G(x)^a \qquad \text{and} \qquad h_a(x) = a\, g(x)\, G(x)^{a-1},$$

respectively, is denoted by $X \sim Exp^a(G)$. The transformation $Exp^a(G)$ is called the exponentiated-G distribution but it is also refereed to as the Lehmann type-I distribution with parameter a. The properties of exponentiated distributions have been studied by many authors in recent years, see Mudholkar et al. (1995) and Mudholkar and Hutson (1996) for exponentiated Weibull distribution, Gupta et al. (1998) for exponentiated Pareto distribution, Gupta and Kundu (2001) for exponentiated exponential distribution, Nadarajah and Gupta (2007) for exponentiated gamma distribution and, more recently, Lemonte and Cordeiro (2011) for exponentiated generalized inverse Gaussian distribution.

Expanding the binomial term in (5) yields the McG density function as a linear combination of exponentiated-G densities, namely

$$f(x) = \sum_{i=0}^{\infty} w_i\, h_{a(i+c)}(x), \tag{9}$$

where $h_{a\,(i+c)}(x)$ denotes the density function of the $Exp^{a(i+c)}(G)$ distribution and

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

We can derive some of the McG properties from the linear combination (9) and those corresponding properties of exponentiated-G distributions.

An expansion for (6) can be derived using the concept of exponentiated inverted beta (EIB) distributions. We define a random variable X having the EIB distribution with parameters α, β and $a > 0$, say $X \sim \text{EIB}(\alpha, \beta, a)$, if its cdf and pdf are given by

$$H_a(x) = I_{\frac{x}{1+x}}(\alpha, \beta)^a \quad \text{and} \quad h_a(x) = \frac{a\, x^{\alpha-1}}{B(\alpha, \beta)(1+x)^{\alpha+\beta}}\, I_{\frac{x}{1+x}}(\alpha, \beta)^{a-1}.$$

The McIB density function is then a linear combination of $\text{EIB}(\alpha, \beta, a(i+c))$ density functions.

We can expand $I_{\frac{x}{1+x}}(\alpha, \beta)^{a-1}$ as

$$I_{\frac{x}{1+x}}(\alpha, \beta)^{a-1} = \sum_{r=0}^{\infty} s_r(a-1)\, I_{\frac{x}{1+x}}(\alpha, \beta)^r, \tag{10}$$

where

$$s_r(a-1) = \sum_{j=r}^{\infty} (-1)^{r+j} \binom{a-1}{j} \binom{j}{r}. \tag{11}$$

Thus, from equations (1), (9) and (10), we can write

$$f(x)=\sum_{r=0}^{\infty}\frac{e_r\,x^{\alpha-1}}{(1+x)^{\alpha+\beta}}\,I_{\frac{x}{1+x}}(\alpha,\beta)^r, \tag{12}$$

where

$$e_r=a\,B(\alpha,\beta)^{-1}\sum_{i=0}^{\infty}(c+i)\,w_i\,s_r(a(i+c)-1). \tag{13}$$

The incomplete beta function expansion for β real non-integer

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

where $(f)_k=\Gamma(f+k)/\Gamma(f)$ is the ascending factorial, can be expressed as

$$I_{\frac{x}{1+x}}(\alpha,\beta)=\sum_{m=0}^{\infty}\frac{d_m\,x^{\alpha+m}}{(1+x)^{\alpha+m}},$$

where $d_m=(1-\beta)_m/[(\alpha+m)\,m!\,B(\alpha,\beta)]$. Further, we use an equation in Section 0.314 of Gradshteyn and Ryzhik (2007) for a power series raised to a positive integer r given by

$$\left(\sum_{m=0}^{\infty}d_m\,z^m\right)^r=\sum_{m=0}^{\infty}p_{r,m}\,z^m, \tag{14}$$

where the coefficients $p_{r,m}$ (for $m=1,2,\ldots$) can be obtained from the recurrence equation

$$p_{r,m}=(m\,d_0)^{-1}\sum_{k=1}^{m}(r\,k-i+k)\,d_k\,p_{r,m-k}, \tag{15}$$

and $p_{r,0}=d_0^r$. The coefficient $p_{r,m}$ can be determined from $p_{r,0},\ldots,p_{r,m-1}$ and then from $d_0,\ldots,d_i$. Clearly, $p_{r,m}$ can be written explicitly in terms of the quantities d_m, although it is not necessary for programming numerically our expansions in any algebraic or numerical software. From equations (12) and (14), we can write

$$f(x)=\sum_{r,m=0}^{\infty}t_{r,m}\,g_{\alpha^\star,\beta}(x). \tag{16}$$

Here, $\alpha^\star=\alpha^\star(r,m)=(r+1)\,\alpha+m$, $g_{\alpha^\star,\beta}(x)$ denotes the IB$(\alpha^\star,\beta)$ density function given by (1) and the coefficients $t_{r,m}$ are calculated from (13) and (15) as

$$t_{r,m}=\frac{a\,p_{r,m}\,B((r+1)\,\alpha+m,\beta)}{B(\alpha,\beta)}\sum_{i=0}^{\infty}(c+i)\,w_i\,s_r(a(i+c)-1).$$

Equation (16) reveals that the McIB density function is a double linear combination of IB density functions. So, some mathematical properties of the McIB distribution immediately follow from those of the IB properties.

4 Moments

From now on, let $X \sim \text{McIB}(\alpha, \beta, a, b, c)$. We derive a simple representation for the sth moment $\mu_s' = \text{E}(X^s)$. For $s < \beta$, we can write from (16)

$$\mu_s' = \sum_{r,m=0}^{\infty} t_{r,m} \, \frac{B((r+1)\alpha + m + s, \beta - s)}{B((r+1)\alpha + m, \beta)}. \tag{17}$$

The moments of the BIB and KwIB distributions are obtained from (17) when $c = 1$ and $a = c$, respectively. Further, the central moments (μ_s) and cumulants (κ_s) of X can be expressed from (17) as

$$\mu_s = \sum_{k=0}^{p} \binom{s}{k} (-1)^k \, \mu_1'^s \, \mu_{s-k}' \qquad \text{and} \qquad \kappa_s = \mu_s' - \sum_{k=1}^{s-1} \binom{s-1}{k-1} \kappa_k \, \mu_{s-k}',$$

respectively, where $\kappa_1 = \mu_1'$. Thus, $\kappa_2 = \mu_2' - \mu_1'^2$, $\kappa_3 = \mu_3' - 3\mu_2'\mu_1' + 2\mu_1'^3$, etc. The pth descending factorial moment of X is

$$\mu_{(p)}' = \text{E}[X^{(p)}] = \text{E}[X(X-1) \times \cdots \times (X - p + 1)] = \sum_{m=0}^{p} s(p, m) \, \mu_m',$$

where $s(r, m) = (m!)^{-1}[d^m m^{(r)}/dx^m]_{x=0}$ is the Stirling number of the first kind. Other kinds of moments related to the L-moments (Hosking, 1990) may also be obtained in closed form, but we consider only these moments for reasons of space.

5 Generating function

Here, we provide three representations for the mgf of X, say $M(t) = \text{E}\{\exp(t\,X)\}$. First, we require the Meijer G-function defined by

$$G_{p,q}^{m,n}\left(x \,\middle|\, \begin{array}{c} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{array}\right) = \frac{1}{2\pi \text{i}} \int_L \frac{H_1(m, n, a_j, b_j, t)}{H_2(n, m, p, q, a_j, b_j, t)} \, x^{-t} dt,$$

with

$$H_1(m, n, a_j, b_j, t) = \prod_{j=1}^{m} \Gamma(b_j + t) \prod_{j=1}^{n} \Gamma(1 - a_j - t),$$

$$H_2(n, m, p, q, a_j, b_j, t) = \prod_{j=n+1}^{p} \Gamma\left(a_j + t\right) \prod_{j=m+1}^{q} \Gamma\left(1 - b_j - t\right),$$

where $\text{i} = \sqrt{-1}$ is the complex unit and L denotes an integration path (see, Gradshteyn and Ryzhik, 2007, § 9.3). The Meijer G-function contains many integrals with elementary and special functions. Some of these integrals are included in Prudnikov et al. (1986).

For $\alpha > 0$ and $t > 0$, we have the following result (Prudnikov et al., 1990)

$$\int_0^{\infty} \exp(-t\,x) \, x^{\alpha - 1} \, (1 + x)^{\nu} dx = \Gamma(-\nu) \, t^{\alpha} \, G_{2,1}^{1,2}\left(t^{-1} \,\middle|\, \begin{array}{l} (1-\alpha), (\nu+1) \\ 0 \end{array}\right).$$

Hence, for $t > 0$, $M(-t) = \mathrm{E}\{\exp(-t\,X)\}$ can be expressed from the previous integral and (16) as

$$M(-t) = \sum_{r,m=0}^{\infty} A_{r,m}\, t^{(r+1)\alpha+m}\, G_{2,1}^{1,2}\left(t^{-1} \,\middle|\, \begin{array}{l} (1-(r+1)\alpha-m), (1-(r+1)\alpha-m-\beta) \\ 0 \end{array} \right), \tag{18}$$

where

$$A_{r,m} = t_{r,m}\, \frac{\Gamma((r+1)\alpha+m+\beta)^2}{\Gamma((r+1)\alpha+m)\,\Gamma(\beta)}.$$

A second representation for the mgf $M_{\alpha,\beta}(t)$ of the IB distribution follows from (1) by a simple transformation $u = x/(1+x)$. We obtain

$$M_{\alpha,\beta}(t) = \frac{1}{B(\alpha,\beta)} \int_0^1 \exp\{t\,u/(1-u)\}\, u^{\alpha-1}(1-u)^{\beta-1} du.$$

By expanding the binomial term and setting $v = 1-u$, we have

$$M_{\alpha,\beta}(t) = \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\infty} (-1)^j \binom{\alpha-1}{j} \int_0^1 \exp\{t\,(1-v)/v\}\, v^{\beta+j-1} dv.$$

We can use MAPLE to calculate the above integral for $t < 0$ as

$$\begin{aligned} M_{\alpha,\beta}(t) = \frac{-\mathrm{e}^{-t}}{B(\alpha,\beta)} \sum_{j=0}^{\infty} (-1)^j \binom{\alpha-1}{j} (-t)^{\beta+j} \\ \times \left[\frac{\pi \csc(\pi(\beta+j))}{\Gamma(\beta+j+1)} + \Gamma(-\beta-j) - \Gamma(-\beta-j,-t) \right], \end{aligned}$$

where $\Gamma(a,x) = \int_x^\infty w^{a-1}\,\mathrm{e}^{-w} dw$ is the complementary incomplete gamma function. So, the mgf of X can be expressed from (16) as

$$M(t) = \sum_{r,m=0}^{\infty} t_{r,m}\, M_{(r+1)\alpha+m,\beta}(t).$$

It can be further reduced (for $t < 0$) to

$$M(t) = -\mathrm{e}^{-t} \sum_{j=0}^{\infty} (-1)^j\, (-t)^{\beta+j}\, h_j \left[\frac{\pi \csc(\pi(\beta+j))}{\Gamma(\beta+j+1)} + \Gamma(-\beta-j) - \Gamma(-\beta-j,-t) \right], \tag{19}$$

where $h_j = \sum_{r,m=0}^{\infty} \frac{t_{r,m}}{B((r+1)\alpha+m,\beta)} \binom{(r+1)\alpha+m-1}{j}$.

Finally, a third representation for $M(t)$ can be obtained using the WhittakerM ("WM" for short) function defined by

$$\mathrm{WM}(p,q,y) = \mathrm{e}^{-y/2}\, y^{q+1/2}\, {}_1F_1(q-p+1/2, 1+2q; y),$$

where

$${}_1F_1(p,q;y) = \frac{\Gamma(q)}{\Gamma(p)} \sum_{j=0}^{\infty} \frac{\Gamma(p+j)}{\Gamma(q+j)} \frac{y^j}{j!}$$

is the confluent hypergeometric function. For any real t, direct integration using MAPLE gives

$$
\begin{aligned}
M(t) = -\Gamma(a+b)^{-1}\,\mathrm{e}^{-t/2}\Big[&-(b-1)^{-1}\,(a+b-t)\,t^{b/2-1}\,\Gamma(a)\,\Gamma(b)\,\mathrm{WM}(a+b/2,(1-b)/2,t)\\
&+(b-1)^{-1}\,(a+1)\,t^{b/2-1}\,\Gamma(a)\,\Gamma(b)\,\mathrm{WM}(a+b/2+1,(1-b)/2,t)\\
&-(-1)^{b}\,(b+1)^{-1}\,(t-a)\,t^{b/2-1}\,\Gamma(a+b)\,\Gamma(-b)\,\mathrm{WM}(a+b/2,(b+1)/2,t)\\
&-(-1)^{b}\,(b+1)^{-1}\,(1+a+b)\,t^{b/2-1}\,\Gamma(a+b)\,\Gamma(-b)\,\mathrm{WM}(a+b/2+1,(b+1)/2,t)\Big].
\end{aligned}
\tag{20}
$$

Equations (18)-(20) are the main results of this section.

6 Quantile Measures

The McIB quantile function, say $Q(u) = F^{-1}(u)$, can be determined from the beta quantile function as given in (8). The effects of the shape parameters a, b and c on the skewness and kurtosis can be considered based on quantile measures. The shortcomings of the classical kurtosis measure are well-known. The Bowley skewness (Kenney and Keeping, 1962) is one of the earliest skewness measures defined by the average of the quartiles minus the median, divided by half the interquartile range, namely

$$B = \frac{Q(3/4) + Q(1/4) - 2Q(1/2)}{Q(3/4) - Q(1/4)}.$$

Since only the middle two quartiles are considered and the outer two quartiles are ignored, this adds robustness to the measure. The Moors kurtosis (Moors, 1998) is based on octiles

$$M = \frac{Q(3/8) - Q(1/8) + Q(7/8) - Q(5/8)}{Q(6/8) - Q(2/8)}.$$

Clearly, $M > 0$ and there is good concordance with the classical kurtosis measures for some distributions. For the normal distribution, $B = M = 0$. These measures are less sensitive to outliers and they exist even for distributions without moments. Because M is based on the octiles, it is not sensitive to variations of the values in the tails or to variations of the values around the median. The basic justification of M as an alternative measure of kurtosis is the following: keeping $Q(6/8) - Q(2/8)$ fixed, M clearly decreases as $Q(3/8) - Q(1/8)$ and $Q(7/8) - Q(5/8)$ decrease. If $Q(3/8) - Q(1/8) \to 0$ and $Q(7/8) - Q(5/8) \to 0$, then $M \to 0$ and half of the total probability mass is concentrated in the neighborhoods of the octiles $Q(2/8)$ and $Q(6/8)$.

In Figures 3, 4 and 5, we plot the measures B and M for some parameter values. These plots indicate that both measures B and M depend on all shape parameters. Figure 5 shows clearly that they can be very sensitive to the extra third parameter c even in the case when $a = b$.

7 Mean Deviations

The deviations from the mean and from the median can be used as a measure of spread in a population. We can derive the mean deviations about the mean and about the median from the relations $\delta_1(X) =$

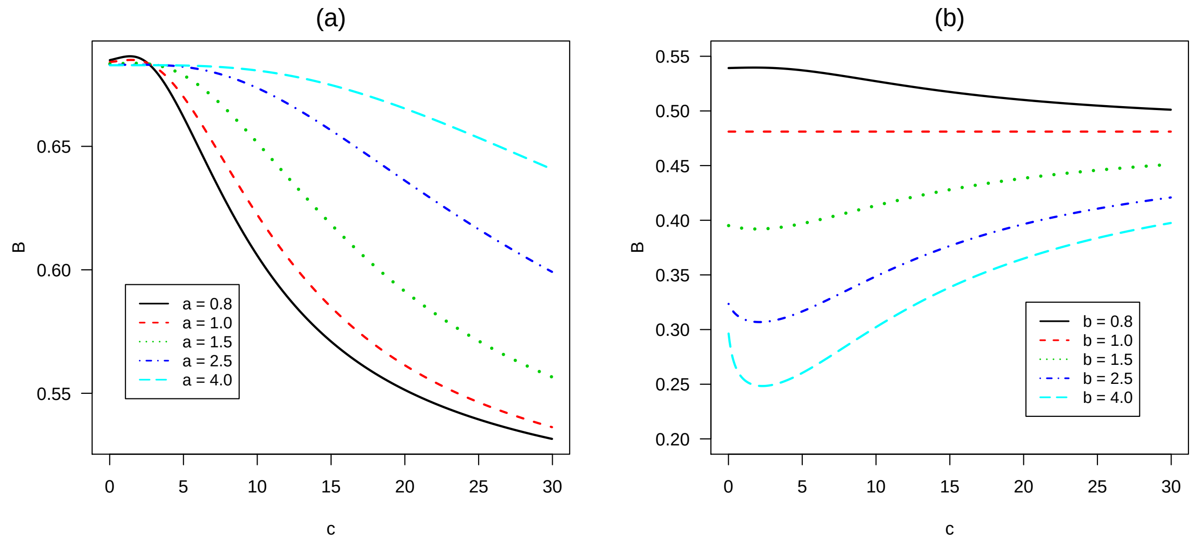

Figure 3: Plots of the measure B for some parameter values. (a) For values $\alpha = 1.5$, $\beta = 1.0$ and $b = 0.5$. (b) For values $\alpha = 1.5$, $\beta = 1.0$ and $a = 1.5$.

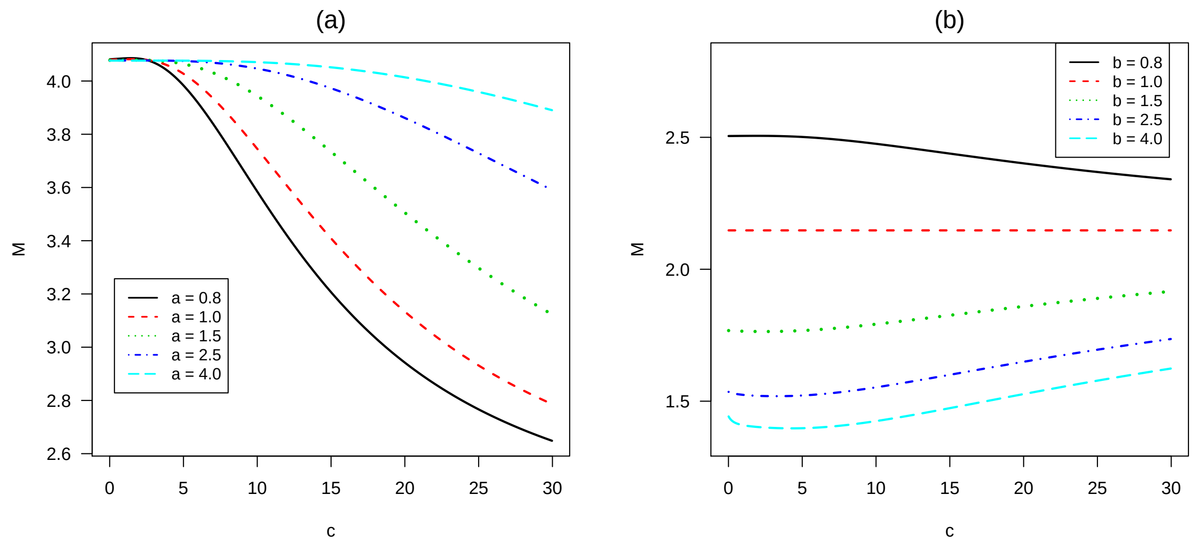

Figure 4: Plots of the measure M for some parameter values. (a) For values $\alpha = 1.5$, $\beta = 1.0$ and $b = 0.5$. (b) For values $\alpha = 1.5$, $\beta = 1.0$ and $a = 1.5$.

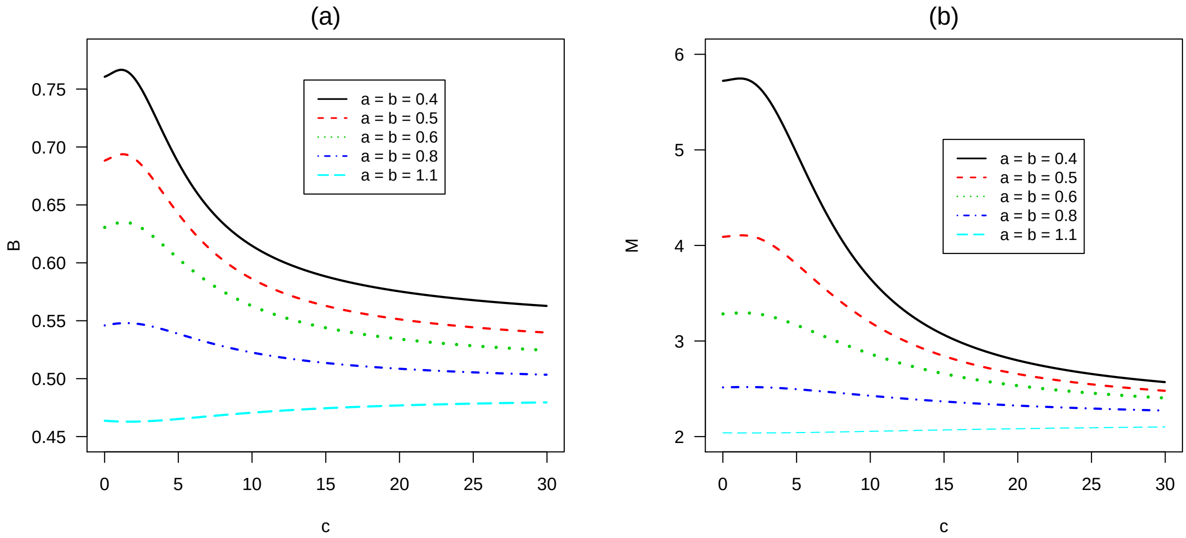

Figure 5: Plots of the measures B (a) and M (b) for some parameter values with $\alpha = 1.5$ and $\beta = 1.0$.

$\mathrm{E}(|X - \mu_1'|)$ and $\delta_2(X) = \mathrm{E}(|X - m|)$, respectively, where the mean $\mu_1' = \mathrm{E}(X)$ comes from (17) and the median m can be obtained from (8) as $m = Q(1/2)$. These measures can be expressed as

$$\delta_1(X) = 2\mu_1'\, F(\mu_1') - 2J(\mu_1') \qquad \text{and} \qquad \delta_2(X) = \mu_1' - 2J(m),$$

where $J(q) = \int_0^q x\, f(x) dt$. In what follows, we obtain an expression for the integral $J(q)$. We can write from (16)

$$\int_0^q x\, f(x)dx = \sum_{r,m=0}^{\infty} t_{r,m}\, \frac{B(\alpha^\star + 1, \beta - 1)}{B(\alpha^\star, \beta)} \int_0^q g_{\alpha^\star+1,\beta-1}(x)dx.$$

But

$$\int_0^q g_{\alpha^\star+1,\beta-1}(x)dx = I_{\frac{q}{1+q}}(\alpha^\star + 1, \beta - 1),$$

and then

$$J(q) = \sum_{r,m=0}^{\infty} t_{r,m}\, \frac{B(\alpha^\star + 1, \beta - 1)}{B(\alpha^\star, \beta)}\, I_{\frac{q}{1+q}}(\alpha^\star + 1, \beta - 1).$$

The result

$$I_{\frac{q}{1+q}}(\alpha^\star + 1, \beta - 1) = \frac{q^{\alpha^\star+1}}{(\alpha^\star + 1)\, B(\alpha^\star + 1, \beta - 1)}\, {}_2F_1\left(\alpha^\star + 1, \alpha^\star + \beta; \alpha^\star + 2; -q\right),$$

allows us to write $J(q)$ as

$$J(q) = \sum_{r,m=0}^{\infty} \frac{t_{r,m}\, q^{\alpha^\star+1}}{(\alpha^\star + 1)\, B(\alpha^\star, \beta)}\, {}_2F_1\left(\alpha^\star + 1, \alpha^\star + \beta; \alpha^\star + 2; -q\right).$$

The mean deviations can be applied to obtain the Lorenz and Bonferroni curves which are important in some fields (economics, reliability, demography, insurance and medicine). They are defined (for a given probability π), by $L(\pi) = J(q)/\mu'_1$ and $B(\pi) = J(q)/(\pi\,\mu'_1)$, respectively, where $q = Q(\pi)$ is determined from (8). In economics, if $\pi = F(q)$ is the proportion of units whose income is lower than or equal to q, $L(\pi)$ gives the proportion of total income volume accumulated by the set of units with an income lower than or equal to q. The Lorenz curve is increasing and convex and given the mean income, the density function of X can be obtained from the curvature of $L(\pi)$. In a similar manner, the Bonferroni curve $B(\pi)$ gives the ratio between the mean income of this group and the mean income of the population. In summary, $L(\pi)$ yields fractions of the total income, while the values of $B(\pi)$ refer to relative income levels.

8 Entropy

The entropy of a random variable X with density function $f(x)$ is a measure of variation of the uncertainty. The Rényi entropy is defined by

$$I_R(\delta) = (1-\delta)^{-1}\log\left\{\int_{-\infty}^{\infty} f(x)^{\delta} dx\right\},$$

where $\delta > 0$ and $\delta \neq 1$. Entropy has been used in various situations in science and engineering. For further details, the reader is referred to Song (2001).

If $X \sim$McIB(α, β, a, b, c), we can show by using the binomial expansion and the results derived in Section 3 that

$$\frac{I_{\frac{x}{1+x}}(\alpha,\beta)^{\delta(a-1)}}{\left[1 - I_{\frac{x}{1+x}}(\alpha,\beta)^{c}\right]^{\delta(1-b)}} = \sum_{r,j=0}^{\infty} (-1)^j \binom{\delta(b-1)}{j} s_r(\delta(a-1)+cj) \sum_{m=0}^{\infty} \frac{\nu_{r,m}\, x^{\alpha r+m}}{(1+x)^{\alpha r+m}},$$

where $\nu_{r,m}$ can be obtained from the recurrence equation $\nu_{r,m} = (m\,d_0)^{-1}\sum_{k=1}^{m}(rk-i+k)\,d_k\,\nu_{r,m-k}$, $\nu_{r,0} = d_0^r$, $d_k = (1-\beta)_k/[(\alpha+k)\,k!\,B(\alpha,\beta)]$ and $s_r(\cdot)$ is defined in (11). Hence, after some algebra, we can write

$$f(x)^{\delta} = \frac{c^{\delta}\,B(\alpha,\beta)^{-\delta}}{B(ac^{-1},b)^{\delta}} \sum_{r,j,m=0}^{\infty} (-1)^j \binom{\delta(b-1)}{j} s_r(\delta(a-1)+cj)\,\nu_{r,m}\,B(\alpha^*,\beta^*)\,g_{\alpha^*,\beta^*}(x),$$

where $\alpha^* = \delta(\alpha-1)+\alpha r+m+1$, $\beta^* = \delta(\beta+1)-1$ and $g_{\alpha^*,\beta^*}(x)$ denotes the IB(α^*,β^*) density function given by (1). Then, we have

$$I_R(\delta) = (1-\delta)^{-1}\log\left\{\frac{c^{\delta}\,B(\alpha,\beta)^{-\delta}}{B(ac^{-1},b)^{\delta}} \sum_{r,m=0}^{\infty} f_r\,\nu_{r,m}\,B(\alpha^*,\beta^*)\right\},$$

where

$$f_r = \sum_{j=0}^{\infty}(-1)^j \binom{\delta(b-1)}{j} s_r(\delta(a-1)+cj).$$

9 Reliability

In the context of reliability, the stress-strength model describes the life of a component which has a random strength X_1 that is subjected to a random stress X_2. The component fails at the instant that the stress applied to it exceeds the strength, and the component will function satisfactorily whenever $X_1 > X_2$. Hence, $R = \Pr(X_2 < X_1)$ is a measure of component reliability which has many applications in engineering. Here, we derive the reliability R when X_1 and X_2 have independent McIB$(\alpha, \beta, a_1, b_1, c_1)$ and McIB$(\alpha, \beta, a_2, b_2, c_2)$ distributions, respectively, with the same baseline parameters α and β.

The pdf of X_1 and the cdf of X_2 can be written from (16) as

$$f_1(x) = \sum_{r,m=0}^{\infty} t_{r,m}^{(1)}\, g_{\alpha_1^\star,\beta}(x) \qquad \text{and} \qquad F_2(x) = \sum_{k,l=0}^{\infty} t_{k,l}^{(2)}\, G_{\alpha_2^\star,\beta}(x),$$

where $\alpha_1^\star = (r+1)\alpha + m$, $\alpha_2^\star = (k+1)\alpha + l$, $g_{\alpha_1^\star,\beta}(x)$ denotes the IB$(\alpha_1^\star, \beta)$ density function given by (1) and $G_{\alpha_2^\star,\beta}(x)$ denotes the IB$(\alpha_2^\star, \beta)$ cumulative function given by (2). Here,

$$t_{r,m}^{(1)} = \frac{a_1\, p_{r,m}\, B((r+1)\,\alpha + m, \beta)}{B(\alpha,\beta)} \sum_{i=0}^{\infty} (c_1 + i)\, w_i^{(1)}\, s_r(a_1(i + c_1) - 1),$$

$$t_{k,l}^{(2)} = \frac{a_2\, p_{k,l}\, B((k+1)\,\alpha + l, \beta)}{B(\alpha,\beta)} \sum_{i=0}^{\infty} (c_2 + i)\, w_i^{(2)}\, s_k(a_2(i + c_2) - 1),$$

$$w_i^{(1)} = \frac{(-1)^i \binom{b_1}{i}}{(a_1 + i)\, B(a_1, b_1 + 1)}, \qquad w_i^{(2)} = \frac{(-1)^i \binom{b_2}{i}}{(a_2 + i)\, B(a_2, b_2 + 1)},$$

where $s_r(\cdot)$ and $p_{r,m}$ are given by (11) and (15), respectively. We have

$$R = \int_0^\infty f_1(x)\, F_2(x)\, dx$$

and then

$$R = \sum_{r,m,k,l=0}^{\infty} t_{r,m}^{(1)}\, t_{k,l}^{(2)} \int_0^\infty g_{\alpha_1^\star,\beta}(x)\, G_{\alpha_2^\star,\beta}(x) dx.$$

After some algebra, we obtain

$$\int_0^\infty g_{\alpha_1^\star,\beta}(x)\, G_{\alpha_2^\star,\beta}(x) dx = \sum_{n=0}^{\infty} \frac{d_n^\star\, B(\alpha_1^\star + \alpha_2^\star + n, \beta)}{B(\alpha_1^\star, \beta)},$$

where $d_n^\star = (1-\beta)_n / [(\alpha_2^\star + n)\, n!\, B(\alpha_2^\star, \beta)]$. Finally, R reduces to the form

$$R = \sum_{r,m,k,l=0}^{\infty} t_{r,m}^{(1)}\, t_{k,l}^{(2)} \sum_{n=0}^{\infty} \frac{d_n^\star\, B(\alpha_1^\star + \alpha_2^\star + n, \beta)}{B(\alpha_1^\star, \beta)}.$$

10 Estimation and Inference

Let $\boldsymbol{x} = (x_1, \ldots, x_n)^\top$ be a random sample of size n from the McIB distribution with unknown parameter vector $\boldsymbol{\theta} = (\alpha, \beta, a, b, c)^\top$. We consider estimation by the method of maximum likelihood. However, some of the other estimators like the percentile estimators, estimators based on order statistics, weighted least squares and estimators based on L-moments can also be explored. The total log-likelihood function for $\boldsymbol{\theta}$ is

$$\ell(\boldsymbol{\theta}) = n\log(c) - n\log(B(\alpha,\beta)) - n\log(B(ac^{-1},b)) + (\alpha-1)\sum_{i=1}^{n}\log(x_i)$$
$$- (\alpha+\beta)\sum_{i=1}^{n}\log(1+x_i) + (a-1)\sum_{i=1}^{n}\log(\dot{z}_i) + (b-1)\sum_{i=1}^{n}\log(1-\dot{z}_i^{c}),$$

where $\dot{z}_i = I_{\frac{x_i}{1+x_i}}(\alpha,\beta)$ for $i=1,\ldots,n$. The components of the score vector $\boldsymbol{U_\theta} = (U_\alpha, U_\beta, U_a, U_b, U_c)^\top$ are obtained by taking the partial derivatives of the log-likelihood function with respect to the five parameters. After some algebra, we obtain

$$U_\alpha = n(\psi(\alpha+\beta)-\psi(\alpha)) + \sum_{i=1}^{n}\log(x_i) - \sum_{i=1}^{n}\log(1+x_i)$$
$$+ (a-1)\sum_{i=1}^{n}\frac{\dot{w}_i + (\psi(\alpha+\beta)-\psi(\alpha))\dot{z}_i}{\dot{z}_i}$$
$$- c(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}[\dot{w}_i + (\psi(\alpha+\beta)-\psi(\alpha))\dot{z}_i]}{1-\dot{z}_i^{c}},$$

$$U_\beta = n(\psi(\alpha+\beta)-\psi(\beta)) - \sum_{i=1}^{n}\log(1+x_i)$$
$$+ (a-1)\sum_{i=1}^{n}\frac{\dot{y}_i + (\psi(\alpha+\beta)-\psi(\beta))\dot{z}_i}{\dot{z}_i}$$
$$- c(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}[\dot{y}_i + (\psi(\alpha+\beta)-\psi(\beta))\dot{z}_i]}{1-\dot{z}_i^{c}},$$

$$U_a = \frac{n\psi(a/c+b)}{c} - \frac{n\psi(a/c)}{c} + \sum_{i=1}^{n}\log(\dot{z}_i),$$

$$U_b = n(\psi(a/c+b)-\psi(b)) + \sum_{i=1}^{n}\log(1-\dot{z}_i^{c}),$$

$$U_c = \frac{n}{c} + \frac{na}{c^2}\big(\psi(a/c)-\psi(a/c+b)\big) - (b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c}\log(\dot{z}_i)}{1-\dot{z}_i^{c}}.$$

Here, $\psi(\cdot)$ is the digamma function, $\dot{w}_i = \dot{I}^{(0)}_{\frac{x_i}{1+x_i}}(\alpha,\beta)$ and $\dot{y}_i = \dot{I}^{(1)}_{\frac{x_i}{1+x_i}}(\alpha,\beta)$, for $i=1,\ldots,n$, and

$$\dot{I}^{(k)}_{\frac{x_i}{1+x_i}}(\alpha,\beta) = \frac{1}{B(\alpha,\beta)}\int_0^{\frac{x_i}{1+x_i}}[\log(w)]^{1-k}[\log(1-w)]^{k}w^{\alpha-1}(1-w)^{\beta-1}dw.$$

The maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}} = (\widehat{\alpha},\widehat{\beta},\widehat{a},\widehat{b},\widehat{c})^\top$ of $\boldsymbol{\theta} = (\alpha,\beta,a,b,c)^\top$ is obtained by setting $U_\alpha = U_\beta = U_a = U_b = 0 = U_c = 0$ and solving these equations numerically using iterative techniques such as a Newton–Raphson type algorithm. The Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (see, for example, Nocedal and Wright, 1999; Press et al., 2007) with analytical derivatives has been used for maximizing the log-likelihood function $\ell(\boldsymbol{\theta})$. After fitting the model, the survival function can be readily estimated (for $i = 1, \ldots, n$) by

$$\widehat{S}(x_i) = 1 - I_{I_{\frac{x_i}{1+x_i}}(\widehat{\alpha},\widehat{\beta})^{\widehat{c}}}(\widehat{a}/\widehat{c}, \widehat{b}).$$

Approximate confidence intervals and hypothesis tests on the parameters α, β, a, b and c can be constructed using the normal approximation for the MLE of $\boldsymbol{\theta}$. Under conditions that are fulfilled for the parameters in the interior of the parameter space, we have $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \overset{A}{\sim} \mathcal{N}_5(\mathbf{0}, \boldsymbol{K}_{\boldsymbol{\theta}}^{-1})$, where $\overset{A}{\sim}$ means approximately distributed and $\boldsymbol{K}_{\boldsymbol{\theta}}$ is the unit expected information matrix. We have the asymptotic result $\boldsymbol{K}_{\boldsymbol{\theta}} = \lim_{n\to\infty} n^{-1}\boldsymbol{J}_n(\boldsymbol{\theta})$, where $\boldsymbol{J}_n(\boldsymbol{\theta})$ is the observed information matrix. The average matrix evaluated at $\widehat{\boldsymbol{\theta}}$, say $n^{-1}\boldsymbol{J}_n(\widehat{\boldsymbol{\theta}})$, can estimate $\boldsymbol{K}_{\boldsymbol{\theta}}$. The observed information matrix $\boldsymbol{J}_n(\boldsymbol{\theta}) = -\partial^2\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top$ is given in the Appendix. The multivariate normal $\mathcal{N}_5(\mathbf{0}, \boldsymbol{J}_n(\widehat{\boldsymbol{\theta}})^{-1})$ distribution can be used to construct approximate confidence intervals and confidence regions for the parameters. In fact, asymptotic $100(1-\eta)\%$ confidence intervals for α, β, a, b and c are given, respectively, by $\widehat{\alpha} \pm z_{\eta/2} \times [\widehat{\mathrm{var}}(\widehat{\alpha})]^{1/2}$, $\widehat{\beta} \pm z_{\eta/2} \times [\widehat{\mathrm{var}}(\widehat{\beta})]^{1/2}$, $\widehat{a} \pm z_{\eta/2} \times [\widehat{\mathrm{var}}(\widehat{a})]^{1/2}$, $\widehat{b} \pm z_{\eta/2} \times [\widehat{\mathrm{var}}(\widehat{b})]^{1/2}$ and $\widehat{c} \pm z_{\eta/2} \times [\widehat{\mathrm{var}}(\widehat{c})]^{1/2}$, where $\mathrm{var}(\cdot)$ is the diagonal element of $\boldsymbol{J}_n(\widehat{\boldsymbol{\theta}})^{-1}$ corresponding to each parameter, and $z_{\eta/2}$ is the quantile $(1-\eta/2)$ of the standard normal distribution.

We can compute the maximum values of the unrestricted and restricted log-likelihood functions to obtain the likelihood ratio (LR) statistics for testing some sub-models of the McIB distribution. For example, we can use the LR statistic to check if the fit using the McIB distribution is statistically "superior" to a fit using the BIB distribution for a given data set. We consider the partition $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$ of the parameter vector of the McIB distribution, where $\boldsymbol{\theta}_1$ is a subset of parameters of interest and $\boldsymbol{\theta}_2$ is a subset of the remaining parameters. The LR statistic for testing the null hypothesis $\mathcal{H}_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ against the alternative hypothesis $\mathcal{H}_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$ is given by $w = 2\{\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})\}$, where $\widehat{\boldsymbol{\theta}}$ and $\widetilde{\boldsymbol{\theta}}$ are the MLEs under the alternative and null hypotheses, respectively, and $\boldsymbol{\theta}_1^{(0)}$ is a specified parameter vector. The statistic w is asymptotically $(n \to \infty)$ distributed as χ^2_k, where k is the dimension of the subset $\boldsymbol{\theta}_1$ of interest. Then, we can compare the McIB model against the BIB model by testing $\mathcal{H}_0 : c = 1$ versus $\mathcal{H}_1 : c \neq 1$ and the LR statistic becomes $w = 2\{\ell(\widehat{\alpha}, \widehat{\beta}, \widehat{a}, \widehat{b}, \widehat{c}) - \ell(\widetilde{\alpha}, \widetilde{\beta}, \widetilde{a}, \widetilde{b}, 1)\}$, where $\widehat{\alpha}, \widehat{\beta}, \widehat{a}, \widehat{b}$ and $\widehat{c}$ are the MLEs under $\mathcal{H}_1$ and $\widetilde{\alpha}, \widetilde{\beta}, \widetilde{a}$ and $\widetilde{b}$ are the MLEs under $\mathcal{H}_0$.

11 Application

We provide an application of the McIB distribution and their sub-models: BIB, KwIB, EIB, LeIB and IB distributions. We compare the results of the fits of these models. We shall consider the real data set corresponding to daily ozone concentrations in New York during May-September, 1973. They were provided by the New York State Department of Conservation and are reported in Nadarajah (2008). All the computations were done using the Ox matrix programming language (Doornik, 2006) which is freely distributed for academic purposes and available at http://www.doornik.com.

Table 1 lists the MLEs (and the corresponding standard errors in parentheses) of the model parameters and the following statistics: AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and HQIC (Hannan–Quinn Information Criterion). These results show that the BIB distribution has the lowest AIC, BIC and HQIC values among all fitted models, and so it could be chosen as the best model. Additionally, it is evident that the IB distribution presents the worst fit to the current data and that the proposed models outperform this distribution. In order to assess if the model is appropriate, the Kaplan–Meier (K-M) estimate and the estimated survival functions of the fitted McIB, BIB, KwIB, EIB, LeIB and IB distributions are shown in Figure 6. From these plots, we can conclude that the McIB and BIB models yield the best fits and hence can be adequate for these data. Again, the IB model presents the worst fit to the data.

Table 1: MLEs (standard errors in parentheses) and the measures AIC, BIC and HQIC.

	Estimates					Statistic		
Distribution	α	β	a	b	c	AIC	BIC	HQIC
McIB	248.4614	0.0901	0.1044	12573.5	1.8889	1071.70	1085.42	1077.27
	(37.620)	(0.0888)	(0.0177)	(31.460)	(0.4946)			
BIB	559.3609	0.0189	0.0487	12573.1		1069.78	1080.76	1074.23
	(28.318)	(0.0054)	(0.0050)	(5190.8)				
KwIB	3166.72	17.9265	0.0105	0.9149		1075.84	1086.82	1080.30
	(1.0014)	(1.0230)	(0.0015)	(0.1606)				
EIB	3167.95	16.6888	0.0109			1074.96	1083.20	1078.31
	(36.833)	(0.9992)	(0.0011)					
LeIB	25.3831	0.0052		448.962		1080.55	1088.79	1083.89
	(2.9312)	(0.0009)		(36.615)				
IB	41.6087	1.7771				1083.34	1088.83	1085.57
	(5.7768)	(0.2160)						

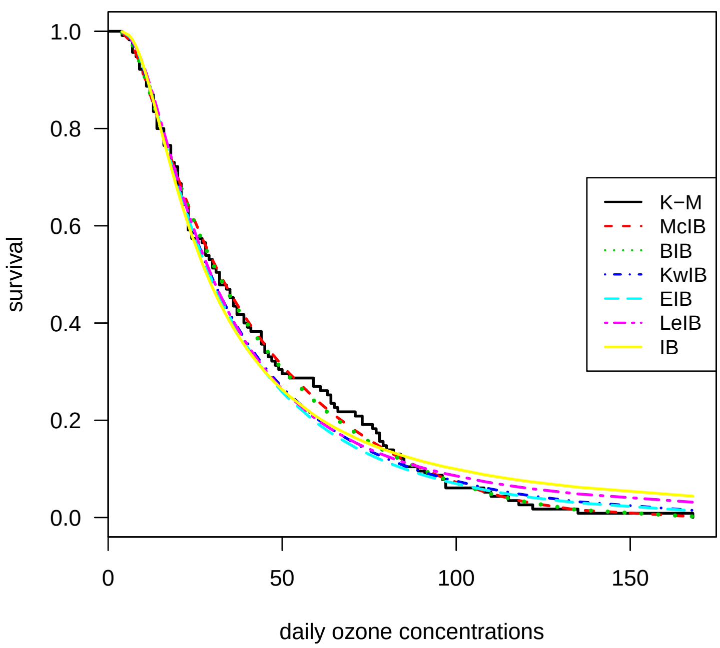

Figure 6: Empirical survival and estimated survival functions of the McIB, BIB, KwIB, EIB, LeIB and IB distributions.

Further, we compare these models using two other criteria. First, we consider LR statistics and then formal goodness-of-fit tests. The McIB model includes some sub-models (described in Section 2) thus allowing their evaluation relative to each other and to a more general model. As mentioned before, we can compute the maximum values of the unrestricted and restricted log-likelihoods to obtain LR statistics for testing some McIB sub-models. The values of the LR statistics are listed in Table 2. From the figures in this table, we conclude that there is no difference among the fitted McIB, BIB, KwIB and EIB models to the current data. In addition, these models provide a better representation for the data than the IB model based on the LR test at the 5% significance level.

Table 2: LR tests.

Model	w	p-value
McIB versus BIB	0.0757	0.7832
McIB versus KwIB	6.1425	0.0132
McIB versus EIB	7.2639	0.0265
McIB versus LeIB	12.8507	0.0016
McIB versus IB	17.6374	0.0005
BIB versus EIB	7.1882	0.0073
BIB versus LeIB	12.7750	0.0004
BIB versus IB	17.5617	0.0002
KwIB versus EIB	1.1214	0.2896
KwIB versus LeIB	6.7083	0.0096
KwIB versus IB	11.4949	0.0032
EIB versus IB	10.3735	0.0013
LeIB versus IB	4.7866	0.0287

Now, we apply formal goodness-of-fit tests in order to verify which distribution fits better to these data. We consider the Cramér–von Mises (W^*) and Anderson–Darling (A^*) statistics. The statistics W^* and A^* are described in details in Chen and Balakrishnan (1995). In general, the smaller the values of these statistics, the better the fit to the data. Let $H(x;\boldsymbol{\theta})$ be the cdf, where the form of H is known but $\boldsymbol{\theta}$ (a k-dimensional parameter vector, say) is unknown. To obtain the statistics W^* and A^*, we can proceed as follows: (i) Compute $v_i = H(x_i;\widehat{\boldsymbol{\theta}})$, where the x_i's are in ascending order, and then $y_i = \Phi^{-1}(v_i)$, where $\Phi(\cdot)$ is the standard normal cdf and $\Phi^{-1}(\cdot)$ its inverse; (ii) Compute $u_i = \Phi\{(y_i - \bar{y})/s_y\}$, where $\bar{y} = (1/n)\sum_{i=1}^{n} y_i$ and $s_y^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$; (iii) Calculate $W^2 = \sum_{i=1}^{n}\{u_i - (2i-1)/(2n)\}^2 + 1/(12n)$ and $A^2 = -n - (1/n)\sum_{i=1}^{n}\{(2i-1)\ln(u_i) + (2n+1-2i)\ln(1-u_i)\}$ and then $W^* = W^2\,(1 + 0.5/n)$ and $A^* = A^2\,(1 + 0.75/n + 2.25/n^2)$. The values of the statistics W^* and A^* for all models are listed in Table 3. Based on these statistics, we conclude that the BIB model fits the current data better than the other models. Additionally, all the proposed models outperform the IB model according to these statistics.

We also fit for the sake of comparison the Birnbaum–Saunders (BS), gamma and Weibull models to the data. The density functions of the BS, gamma and Weibull distributions are (for $x > 0$)

$$f(x) = \frac{1}{2\sqrt{2\pi}\alpha\beta}\left[\left(\frac{\beta}{x}\right)^{1/2} + \left(\frac{\beta}{x}\right)^{3/2}\right]\exp\left\{-\frac{1}{2\alpha^2}\left(\frac{x}{\beta} + \frac{\beta}{x} - 2\right)\right\},$$

Table 3: Goodness-of-fit tests.

Distribution	Statistic W^*	A^*
McIB	0.02820	0.17537
BIB	0.02669	0.17217
KwIB	0.05896	0.44217
EIB	0.06272	0.47069
LeIB	0.10039	0.74268
IB	0.13965	1.01826

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x) \qquad \text{and} \qquad f(x) = \alpha\, \beta\, x^{\alpha-1} \exp(-\beta x^\alpha),$$

respectively, with $\alpha > 0$ and $\beta > 0$. The MLEs (standard errors in parentheses) and the statistics W^* and A^* are listed in Table 4. Based on the statistics W^* and A^*, the BS model presents the

Table 4: MLEs (standard errors in parentheses) and the measures W^* and A^*.

	Estimates		Statistic	
Distribution	α	β	W^*	A^*
BS	0.8555	31.0439	0.04939	0.35286
	(0.0564)	(2.2542)		
Gamma	1.8102	0.0426	0.15794	0.92646
	(0.2203)	(0.0060)		
Weibull	1.3720	0.0051	0.20885	1.20586
	(0.0976)	(0.0021)		

best fit. On the other hand, according to these statistics, the McIB and BIB models outperform the BS model (compare the figures in Tables 3 and 4) and the new models outperform the gamma and Weibull models. So, the proposed distributions can yield a better fit than the BS, gamma and Weibull models and therefore may be an interesting alternative to these distributions for modeling positive real data sets.

In summary, the new McIB distribution (and their sub-models) produce better fits for the current data than the IB distribution. Additionally, among all the proposed models, the BIB distribution presents the best fit and should be chosen, since it yields the lowest AIC, BIC and HQIC values (see Table 1) and the lowest W^* and A^* values (see Table 3). These results illustrate the potentiality of the new distribution (and their sub-models) and the necessity of the additional shape parameters.

12 Concluding remarks

We propose a new five-parameter distribution, called the McDonald inverted beta (McIB) distribution, and study some of its general structural properties. This distribution has the support on the positive real line and it can be used to analyze lifetime data. We provide expansions for the density function, moments, generating function, mean deviations, entropy and reliability. The parameter estimation is approached by maximum likelihood and the observed information matrix is derived. The usefulness of the new model is illustrated in an application to real data using likelihood ratio statistics and goodness-of-fit tests. In a real application, we show that the proposed model is a very competitive model to the Birnbaum–Saunders, gamma and Weibull distributions. The formulae related with the new model are manageable and may turn into adequate tools comprising the arsenal of applied statisticians. The McIB model has the potential to attract wider applications in survival analysis.

Acknowledgments

We gratefully acknowledge grants from CNPq and FAPESP (Brazil).

Appendix

The observed information matrix for the parameter vector $\boldsymbol{\theta} = (\alpha, \beta, a, b, c)^\top$ is

$$\boldsymbol{J}_n(\boldsymbol{\theta}) = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} = -\begin{pmatrix} U_{\alpha\alpha} & U_{\alpha\beta} & U_{\alpha a} & U_{\alpha b} & U_{\alpha c} \\ \cdot & U_{\beta\beta} & U_{\beta a} & U_{\beta b} & U_{\beta c} \\ \cdot & \cdot & U_{aa} & U_{ab} & U_{ac} \\ \cdot & \cdot & \cdot & U_{bb} & U_{bc} \\ \cdot & \cdot & \cdot & \cdot & U_{cc} \end{pmatrix},$$

whose elements are, after extensive algebraic manipulations, given by

$$\begin{aligned} U_{\alpha\alpha} = & -n\psi'_\alpha + (a-1)\sum_{i=1}^{n} \frac{\ddot{w}_i - 2\psi_\alpha \dot{w}_i + \psi_\alpha^2 \dot{z}_i - \psi'_\alpha \dot{z}_i}{\dot{z}_i} - (a-1)\sum_{i=1}^{n} \frac{\dot{w}_i(\dot{w}_i - \psi_\alpha \dot{z}_i)}{\dot{z}_i^2} \\ & + (a-1)\psi_\alpha \sum_{i=1}^{n} \frac{\dot{w}_i - \psi_\alpha \dot{z}_i}{\dot{z}_i} - c^2(b-1)\sum_{i=1}^{n} \frac{\dot{z}_i^{c-2}(\dot{w}_i - \psi_\alpha \dot{z}_i)^2}{1-\dot{z}_i^c} \\ & - c(b-1)\sum_{i=1}^{n} \frac{\dot{z}_i^{c-1}(\ddot{w}_i - 2\psi_\alpha \dot{w}_i + \psi_\alpha^2 \dot{z}_i - \psi'_\alpha \dot{z}_i)}{1-\dot{z}_i^c} \\ & + c(b-1)\sum_{i=1}^{n} \frac{\dot{z}_i^{c-2}\dot{w}_i(\dot{w}_i - \psi_\alpha \dot{z}_i)}{1-\dot{z}_i^c} - c(b-1)\psi_\alpha \sum_{i=1}^{n} \frac{\dot{z}_i^{c-1}(\dot{w}_i - \psi_\alpha \dot{z}_i)}{1-\dot{z}_i^c} \\ & - c^2(b-1)\sum_{i=1}^{n} \frac{\dot{z}_i^{2(c-1)}(\dot{w}_i - \psi_\alpha \dot{z}_i)^2}{(1-\dot{z}_i^c)^2}, \end{aligned}$$

$$
\begin{aligned}
U_{\alpha\beta} &= n\psi'(\alpha+\beta) + (a-1)\sum_{i=1}^{n}\frac{\dot{w}_i - \psi_\beta \dot{w}_i - \psi_\alpha \dot{y}_i + \psi_\alpha\psi_\beta \dot{z}_i + \psi'(\alpha+\beta)\dot{z}_i}{\dot{z}_i} \\
&\quad - (a-1)\sum_{i=1}^{n}\frac{\dot{y}_i(\dot{w}_i - \psi_\alpha \dot{z}_i)}{\dot{z}_i^2} + (a-1)\psi_\beta\sum_{i=1}^{n}\frac{\dot{w}_i - \psi_\alpha \dot{z}_i}{\dot{z}_i} \\
&\quad - c^2(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-2}(\dot{y}_i - \psi_\beta \dot{z}_i)(\dot{w}_i - \psi_\alpha \dot{z}_i)}{1-\dot{z}_i^c} \\
&\quad - c(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}(\dot{y}_i - \psi_\beta \dot{w}_i - \psi_\alpha \dot{y}_i + \psi_\alpha\psi_\beta \dot{z}_i + \psi'(\alpha+\beta)\dot{z}_i)}{1-\dot{z}_i^c} \\
&\quad + c(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-2}\dot{y}_i(\dot{w}_i - \psi_\alpha \dot{z}_i)}{1-\dot{z}_i^c} - c(b-1)\psi_\beta\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}(\dot{w}_i - \psi_\alpha \dot{z}_i)}{1-\dot{z}_i^c} \\
&\quad - c^2(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{2(c-1)}(\dot{y}_i - \psi_\beta \dot{z}_i)(\dot{w}_i - \psi_\alpha \dot{z}_i)}{(1-\dot{z}_i^c)^2},
\end{aligned}
$$

$$
U_{\alpha a} = \sum_{i=1}^{n}\frac{\dot{w}_i - \psi_\alpha \dot{z}_i}{\dot{z}_i}, \qquad U_{\alpha b} = -c\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}(\dot{w}_i - \psi_\alpha \dot{z}_i)}{1-\dot{z}_i^c},
$$

$$
U_{\alpha c} = -(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}(\dot{w}_i - \psi_\alpha \dot{z}_i)(1 + c\log(\dot{z}_i))}{1-\dot{z}_i^c},
$$

$$
\begin{aligned}
U_{\beta\beta} &= -n\psi'_\beta + (a-1)\sum_{i=1}^{n}\frac{\ddot{y}_i - 2\psi_\beta \dot{y}_i + \psi_\beta^2 \dot{z}_i - \psi'_\beta \dot{z}_i}{\dot{z}_i} - (a-1)\sum_{i=1}^{n}\frac{\dot{y}_i(\dot{y}_i - \psi_\beta \dot{z}_i)}{\dot{z}_i^2} \\
&\quad + (a-1)\psi_\beta\sum_{i=1}^{n}\frac{\dot{y}_i - \psi_\beta \dot{z}_i}{\dot{z}_i} - c^2(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-2}(\dot{y}_i - \psi_\beta \dot{z}_i)^2}{1-\dot{z}_i^c} \\
&\quad - c(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}(\ddot{y}_i - 2\psi_\beta \dot{y}_i + \psi_\beta^2 \dot{z}_i - \psi'_\beta \dot{z}_i)}{1-\dot{z}_i^c} \\
&\quad + c(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-2}\dot{y}_i(\dot{y}_i - \psi_\beta \dot{z}_i)}{1-\dot{z}_i^c} - c(b-1)\psi_\beta\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}(\dot{y}_i - \psi_\beta \dot{z}_i)}{1-\dot{z}_i^c} \\
&\quad - c^2(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{2(c-1)}(\dot{y}_i - \psi_\beta \dot{z}_i)^2}{(1-\dot{z}_i^c)^2},
\end{aligned}
$$

$$
U_{\beta a} = \sum_{i=1}^{n}\frac{\dot{y}_i - \psi_\beta \dot{z}_i}{\dot{z}_i}, \qquad U_{\beta b} = -c\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}(\dot{y}_i - \psi_\beta \dot{z}_i)}{1-\dot{z}_i^c},
$$

$$
U_{\alpha c} = -(b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^{c-1}(\dot{y}_i - \psi_\beta \dot{z}_i)(1 + c\log(\dot{z}_i))}{1-\dot{z}_i^c},
$$

$$
U_{aa} = -\frac{n}{c^2}\big(\psi'(a/c) - \psi'(b + a/c)\big), \qquad U_{ab} = \frac{n\,\psi'(a/c + b)}{c},
$$

$$
U_{ac} = \frac{n}{c^2}\big(\psi(a/c) - \psi(b + a/c)\big) + \frac{na}{c^3}\big(\psi'(a/c) - \psi'(b + a/c)\big),
$$

$$
U_{bb} = n\big(\psi'(b + a/c) - \psi'(b)\big), \quad U_{bc} = -\frac{na\psi'(b + a/c)}{c^2} - \sum_{i=1}^{n}\frac{\dot{z}_i^c \log(\dot{z}_i)}{1-\dot{z}_i^c},
$$

$$
\begin{aligned}
U_{cc} &= -\frac{n}{c^2} - \frac{2na\psi(a/c)}{c^3} - \frac{na^2\psi'(a/c)}{c^4} + \frac{2na\psi(b + a/c)}{c^3} \\
&\quad + \frac{na^2\psi'(b + a/c)}{c^4} - (b-1)\sum_{i=1}^{n}\frac{\dot{z}_i^c\,[\log(\dot{z}_i)]^2}{(1-\dot{z}_i^c)^2},
\end{aligned}
$$

where $\psi'(\cdot)$ is the trigamma function, $\psi_\alpha = \psi(\alpha) - \psi(\alpha+\beta)$, $\psi_\beta = \psi(\beta) - \psi(\alpha+\beta)$, $\psi'_\alpha = \psi'(\alpha) - \psi'(\alpha+\beta)$, $\psi'_\beta = \psi'(\beta) - \psi'(\alpha+\beta)$, $\ddot{w}_i = \ddot{I}^{(0)}_{\frac{x_i}{1+x_i}}(\alpha,\beta)$, $\ddot{y}_i = \ddot{I}^{(1)}_{\frac{x_i}{1+x_i}}(\alpha,\beta)$, for $i = 1,\ldots,n$, with

$$\ddot{I}^{(k)}_{\frac{x_i}{1+x_i}}(\alpha,\beta) = \frac{1}{B(\alpha,\beta)}\int_0^{\frac{x_i}{1+x_i}} [\log(w)]^{2(1-k)}[\log(1-w)]^{2k} w^{\alpha-1}(1-w)^{\beta-1}dw,$$

and $\dot{z}_i$, $\dot{w}_i$ and $\dot{y}_i$ were defined in Section 10.

References

Bookstaber, R.M., McDonald, J.B. (1987). A general distribution for describing security price returns. *The Journal of Business* 60, 401–424.

Chen, G., Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology* 27, 154–161.

Cordeiro, G.M., de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*. DOI:10.1080/00949650903530745.

Cummins, J.D., Dionne, G., McDonald, J.B., Pritchett, B.M. (1990). Applications of the GB2 family of distributions in modeling insurance loss processes. *Insurance: Mathematics and Economics* 9, 257–272.

Doornik, J.A. (2006). *An Object-Oriented Matrix Language – Ox 4*, 5th ed. Timberlake Consultants Press, London.

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

Gradshteyn, I.S., Ryzhik, I.M. (2007). *Table of Integrals, Series, and Products*. Academic Press, New York.

Gupta, R.C., Gupta, R.D., Gupta, P.L. (1998). Modeling failure time data by Lehmann alternatives. *Communications in Statistics – Theory and Methods* 27, 887–904.

Gupta, R.C., Kundu, D. (2001). Exponentiated exponential distribution: An alternative to gamma and Weibull distributions. *Biometrical Journal* 43, 117–130.

Hansen, B.E. (1994). Autoregressive conditional density estimation. *International Economic Review* 35, 705–730.

Hosking, J.R.M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society B* 52, 105–124.

Jones, M.C. (2004). Families of distributions arising from distributions of order statistics. *Test* 13, 1–43.

Jones, M.C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology* 6, 70–91.

Kenney, J.F., Keeping, E.S. (1962). *Mathematics of Statistics*, pp. 101-102, Part 1, 3rd ed. Princeton, NJ.

Kumaraswamy, P. (1980). Generalized probability density-function for double-bounded random-processes. *Journal of Hydrology* 46, 79–88.

Lemonte, A.J., Cordeiro, G.M. (2011). The exponentiated generalized inverse Gaussian distribution. *Statistics and Probability Letters* 81, 506–517.

McDonald, J.B. (1984). Some generalized functions for the size distribution of income. *Econometrica* 52, 647–663.

McDonald, J.B., Bookstaber, R.M. (1991). Option pricing for generalized distributions. *Communications in Statistics – Theory and Methods* 20, 4053–4068.

McDonald, J.B., Butler, R.J. (1987). Some generalized mixture distributions with an application to unemployment duration. *The Reriew of Economics and Statistics* 69, 232–240.

McDonald, J.B., Butler, R.J. (1990). Regression models for positive random variables. *Journal of Econometrics* 43, 227–251.

McDonald, J.B., Richards, D.O. (1987a). Model selection: Some generalized distributions. *Communications in Statistics – Theory and Methods* 16, 1049–1074.

McDonald, J.B., Richards, D.O. (1987b). Hazard rates and generalized beta distribution. *IEEE Transactions on Reliability* 36, 463–466.

Moors, J.J.A. (1998). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society D* 37, 25–32.

Mudholkar, G.S., Hutson, A.D. (1996). The exponentiated Weibull family: Some properties and a flood data application. *Communication in Statistics – Theory and Methods* 25, 3059–3083.

Mudholkar, G.S., Srivastava, D.K., Freimer, M. (1995). The exponentiated Weibull family. *Technometrics* 37, 436-445.

Nadarajah, S. (2008). A truncated inverted beta distribution with application to air pollution data. *Stochastic Environmental Research and Risk Assessment* 22, 285–289.

Nadarajah, S., Gupta, A.K. (2007). The exponentiated gamma distribution with application to drought data. *Calcutta Statistical Association Bulletin* 59, 29–54.

Nocedal, J., Wright, S.J. (1999). *Numerical Optimization*. Springer: New York.

Press, W.H., Teulosky, S.A., Vetterling, W.T., Flannery, B.P. (2007). *Numerical Recipes in C: The Art of Scientific Computing*, 3rd ed. Cambridge University Press.

Prudnikov, A.P., Brychkov, Y.A., Marichev, O.I. (1986). *Integrals and Series, Volume 1.* Gordon and Breach Science Publishers: Amsterdam.

Prudnikov, A.P., Brychkov, Y.A., Marichev, O.I. (1990). *Integrals and Series, Volume 3: More special functions.* Gordon and Breach Science Publishers: Amsterdam.

Song, K.S. (2001). Rényi information, loglikelihood and an intrinsic distribution measure. *Journal of Statistical Planning and Inference* 93, 51–69.