The Beta Alpha Distribution

Gauss M. Cordeiro[a**], Fredy Castellares[b] and Lourdes C. Montenegro[b]

[a] Departamento de Estatística e Informática,
Universidade Federal Rural de Pernambuco,
Rua Dom Manoel de Medeiros s/n, 50171-900 – Recife, PE, Brazil
(gausscordeiro@uol.com.br)

[b] Universidade Federal de Minas Gerais
Av. Antonio Carlos 6627, 31270-901–Minas Gerais, MG, Brazil
(fredy@est.ufmg.br, lourdes@est.ufmg.br)

Abstract

For the first time, we introduce the so-called beta alpha distribution which generalizes the alpha distribution (Katsav (1968), Wager and Barash (1971)). Expansions for the cumulative distribution and density functions that do not involve complicated functions are derived. We obtain expressions for its moments and for the moments of order statistics. The estimation of parameters is approached by the method of maximum likelihood and the expected information matrix is derived. The usefulness of the beta alpha distribution is illustrated in an analysis of a real data set. The new model is quite flexible in analyzing positive data and it is an important alternative to the gamma, Weibull, generalized exponential, beta exponential and Birnbaum-Saunders distributions.

Keywords: Alpha distribution; Beta alpha distribution; Maximum likelihood estimation; Moment; Observed information matrix.

1 Introduction

The alpha distribution generally is used in tool wear problems (Katsav (1968), Wager and Barash (1971)). Sherif (1983) suggested its use in modeling lifetimes under accelerated test conditions and Salvia (1985) provided a characterization of the distribution and a number of accompanying properties. One of the main results of the alpha distribution is that the mean does not exist. However, this result does not prohibit its use as a model for accelerated life testing (as, for example, the Cauchy and certain Pareto models). A Cauchy model provides a symmetric distribution, whereas the Pareto distribution has a non-increasing probability density function (pdf) and is highly skewed to the right. The alpha model is also skewed to the right, but (unlike the Pareto) its mode is finite.

A random variable X has an alpha distribution with shape ($\alpha > 0$) and scale ($\beta > 0$) parameters, if its probability density function (pdf) is

$$g(x) = \frac{\beta}{\sqrt{2\pi}x^2\Phi(\alpha)} \exp\left\{-\frac{1}{2}\left(\alpha - \frac{\beta}{x}\right)^2\right\},\ x > 0. \tag{1}$$

**Corresponding author. Email: gausscordeiro@uol.com.br

Simple integration shows that the cumulative distribution function (cdf) of the alpha distribution is easily expressed in terms of the cdf of the standard normal distribution as

$$G(x) = \frac{\Phi\left(\alpha - \frac{\beta}{x}\right)}{\Phi\left(\alpha\right)}. \tag{2}$$

The hazard rate function corresponding to (1) is

$$h(x) = \frac{\beta}{\sqrt{2\pi}x^2\left[\Phi(\alpha) - \Phi(\alpha - \frac{\beta}{x})\right]} \exp\left\{-\frac{1}{2}\left(\alpha - \frac{\beta}{x}\right)^2\right\}. \tag{3}$$

Plots of the alpha density function for selected parameter values are given in Figure 1. The mode of the distribution is

$$M = \beta\frac{(\sqrt{\alpha^2+8} - \alpha)}{4}.$$

The mode M moves to the left (right) as α (β) increases. Figure 2 gives some of the possible shapes of the alpha cumulative function for selected parameter values. Figure 3 shows that the hazard rate function has an upside-down bathtub-shaped for different values of the parameters, which increases to a modal value and then decreases slowly.

In this paper, we introduce a new four-parameter distribution, so-called the beta alpha (BA) distribution, with the hope of wider applications for accelerated life testing in the area of engineering and other areas of research. This generalization includes the alpha distribution as special case. The new distribution due to its flexibility seems be an important model that can be used in a variety of problems in modeling of reliability.

It is interesting to know that there is a similarity between the density (1) and the inverse normal density function. The similarity is due to the fact that (1) is nothing but the density function of $X = 1/Y$ if Y has a normal $N(\mu, \sigma^2)$ distribution truncated to the left at zero (for $\alpha = \mu/\sigma$ and $\beta = 1/\sigma$).

The calculations in the paper involve some special functions, including the well-known error function defined by

$$\mathrm{erf(x)} = \frac{2}{\sqrt{\pi}}\int_0^{\mathrm{x}} \exp(-\mathrm{t}^2)\mathrm{dt},$$

the incomplete beta function ratio, i.e. the cdf of the beta distribution with parameters $a > 0$ and $b > 0$, given by

$$I_x(a, b) = \frac{1}{B(a, b)}\int_0^x t^{a-1}(1-t)^{b-1}dt,$$

the beta function defined by ($\Gamma(\cdot)$ is the gamma function)

$$B(a, b) = \int_0^1 w^{a-1}(1-w)^{b-1}dw = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

the well-known hypergeometric function (Gradshteyn and Ryzhik, 2000) defined by (for $\alpha_k > 0$, $\beta_k > 0$, $k = 1, 2, \cdots$)

$$_pF_q(\alpha_1, \cdots, \alpha_p; \beta_1, \cdots, \beta_q; x) = \sum_{k=0}^{\infty} \frac{(\alpha_1)_k \cdots (\alpha_p)_k}{(\beta_1)_k \cdots (\beta_q)_k}\frac{x^k}{k!},$$

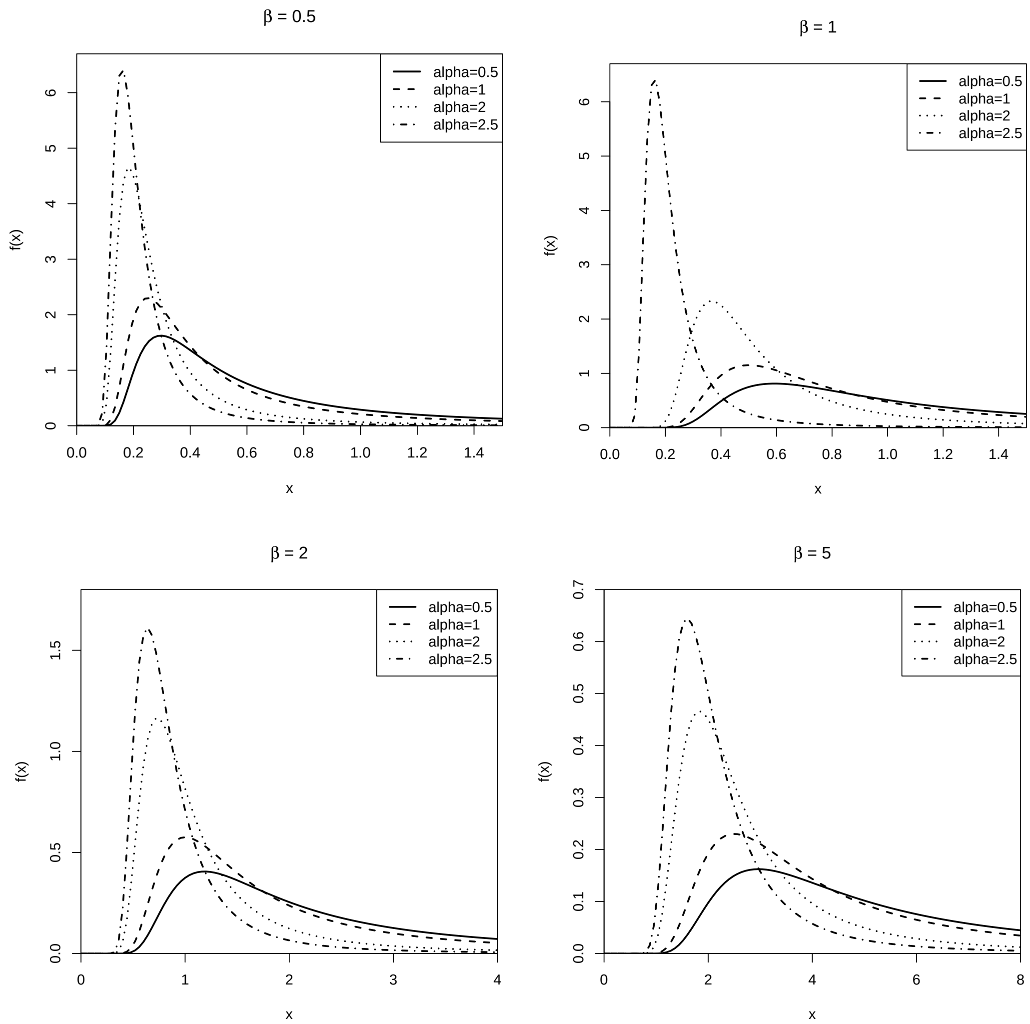

Figure 1: Plots of the alpha density (1) for some parameter values.

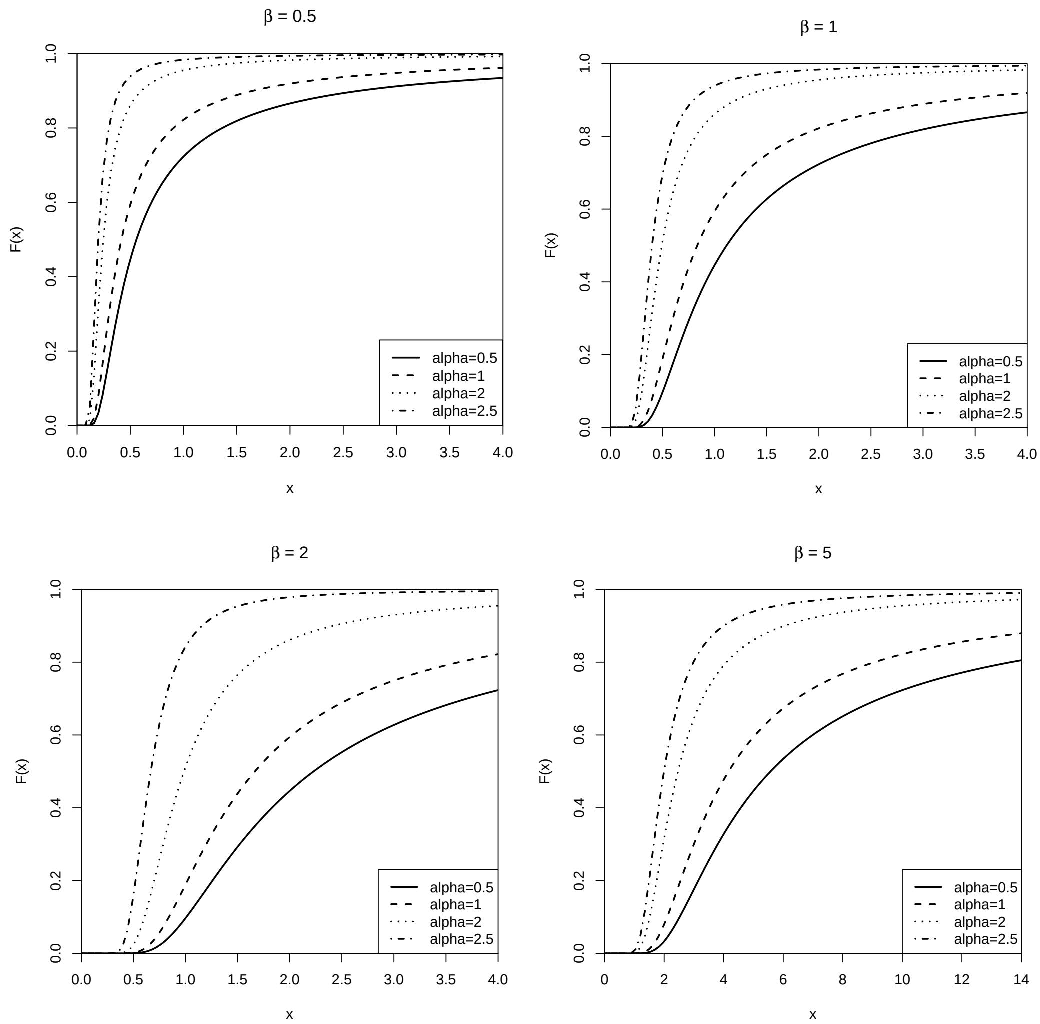

Figure 2: Plots of the alpha cdf (2) for some parameter values.

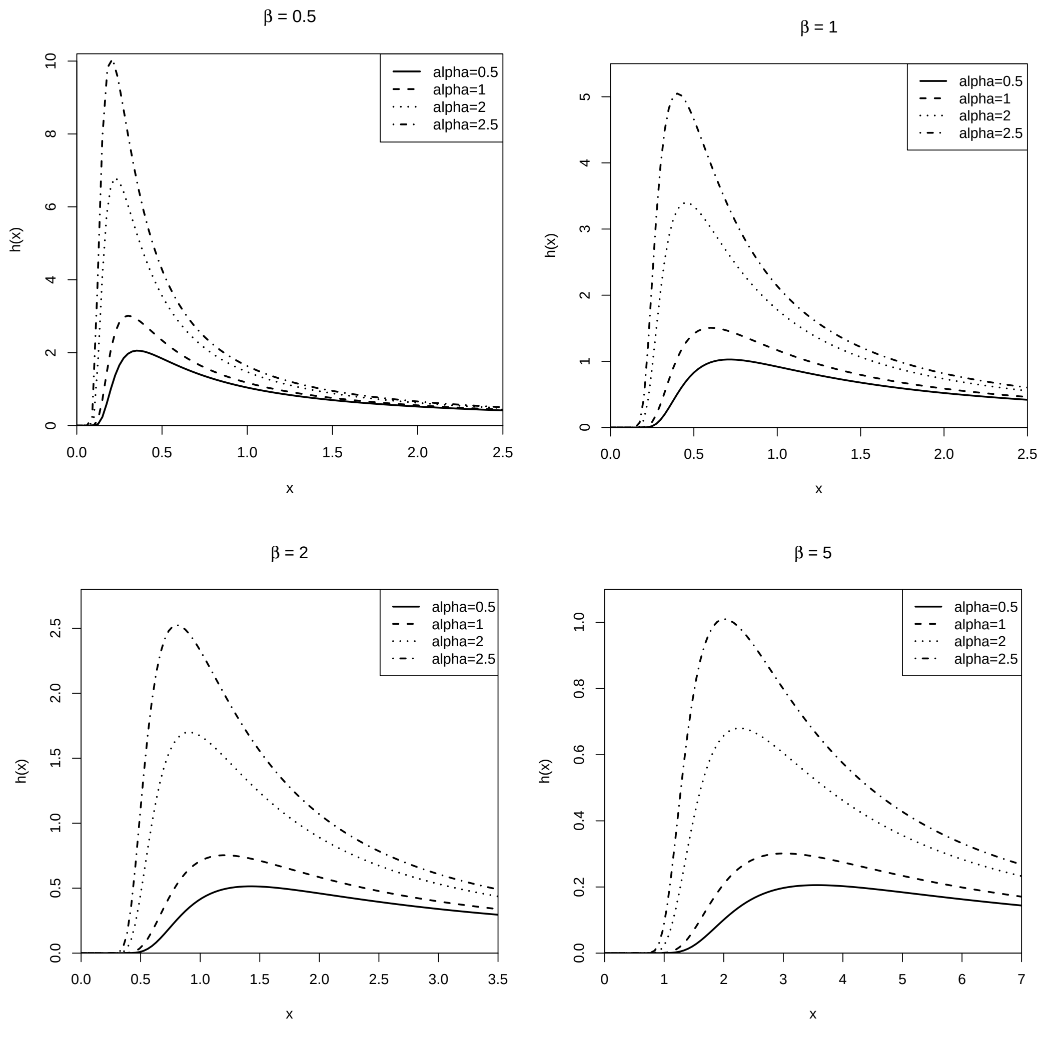

Figure 3: Plots of the alpha hazard rate (3) for some parameter values.

where $(\alpha)_i = \alpha(\alpha+1)\dots(\alpha+i-1)$ is the ascending factorial. Two important particular cases correspond to $p = 2$ and $q = 1$ giving ${}_2F_1(\alpha,\beta;\gamma;x)$ and $p = q = 1$ yielding the confluent hypergeometric function ${}_1F_1(\alpha;\beta;x)$.

Finally, we require the Laguerre function defined by

$$L_{p/2}^{1/2}\left(\frac{\alpha^2}{2}\right) = \frac{\Gamma(p/2+1/2+1)}{\Gamma(p/2)}\,{}_1F_1(p/2;1/2+1;\alpha^2/2).$$

The rest of the paper is organized as follows. In Section 2, we define the BA distribution. Probability weighted moments (PWMs) are expectations of certain functions of a random variable defined when the ordinary moments of the random variable exist. In Section 3, we derive the PWMs of the alpha distribution. Section 4 provides a general expansion for the moments of the BA distribution. Its moment generating function (mgf) is derived in Section 5. Section 6 is devoted to mean deviations. Section 7 provides expansions for the BA order statistics. In Section 8, we derive the moments of order statistics and expansions for the L-moments defined by Hosking (1990) as expectations of certain linear combinations of order statistics. In Section 9, we discuss maximum likelihood estimation and calculate the elements of the observed information matrix. Two applications to real data in Section 10 illustrate the importance of the BA distribution. Finally, concluding remarks are given in Section 11.

2 The New Model

The generalization of the alpha distribution is motivated by the work of Eugene et al. (2002). One major benefit of the class of beta generalized distributions is its ability of fitting skewed data that can not be properly fitted by existing distributions. Consider starting from a parent cumulative function $G(x)$, they defined a class of generalized beta distributions by

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

where $a > 0$ and $b > 0$ are two additional parameters whose role is to introduce skewness and to vary tail weight. The cdf $G(x)$ could be quite arbitrary and $F(x)$ is refereed to the beta G distribution. If V has a beta distribution with parameters a and b, application of $X = G^{-1}(V)$ yields X with cumulative distribution (4).

We can express (4) in terms of the hypergeometric function, since the properties of this function are well established in the literature. We have

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

Some generalized beta distributions were discussed in recent literature. Eugene et al. (2002) defined the beta normal (BN) distribution by taking $G(x)$ to be the cdf of the normal distribution and derived some of its first moments. Nadarajah and Kotz (2004) introduced the beta Gumbel (BGu) distribution by taking $G(x)$ to be the cdf of the Gumbel distribution, provided expressions for the moments, and discussed the asymptotic distribution of the extreme order statistics and maximum likelihood estimation. Nadarajah and Gupta (2004) defined the beta Fréchet (BF) distribution by taking $G(x)$ to be the Fréchet distribution, derived the analytical shapes of the density and hazard rate functions and calculated the asymptotic distribution of the extreme order statistics. Nadarajah and

Kotz (2005) proposed the beta exponential (BE) distribution and obtained the moment generating function, the first four cumulants, the asymptotic distribution of the extreme order statistics and estimated its parameters by the method of maximum likelihood.

The density function corresponding to (4) can be expressed as

$$f(x) = \frac{g(x)}{B(a,b)} G(x)^{a-1}\{1 - G(x)\}^{b-1}, \tag{5}$$

where $g(x) = dG(x)/dx$ is the density of the parent distribution. The density $f(x)$ will be most tractable when both functions $G(x)$ and $g(x)$ have simple analytic expressions. Except for some special choices of these functions, the density $f(x)$ will be difficult to deal with in generality.

The BA density function with four parameters α, β, a and b, from now on denoted by $BA(\alpha, \beta, a, b)$, is given by

$$f(x) = \frac{\beta \exp\left\{-\frac{1}{2}\left(\alpha - \frac{\beta}{x}\right)^2\right\}}{\sqrt{2\pi}x^2 B(a,b)\Phi(\alpha)^{a+b-1}} \Phi\left(\alpha - \frac{\beta}{x}\right)^{a-1} \left\{\Phi(\alpha) - \Phi\left(\alpha - \frac{\beta}{x}\right)\right\}^{b-1}, \quad x > 0. \tag{6}$$

Evidently, the density function (6) does not involve any complicated function but generalizes a few known distributions. The alpha distribution arises as the particular case for $a = b = 1$. If $b = 1$, it leads to a new distribution termed here the exponentiated alpha (EA) distribution. The BA distribution is easily simulated as follows: if V has a beta distribution with parameters a and b, then $X = \beta\{\alpha - \Phi^{-1}(\Phi(\alpha)V)\}^{-1}$ has the $BA(\alpha, \beta, a, b)$ distribution.

The cdf and hazard rate function corresponding to (6) are given by

$$F(x) = I_{\left[\frac{\Phi\left(\alpha - \frac{\beta}{x}\right)}{\Phi(\alpha)}\right]}(a,b) \tag{7}$$

and

$$h(x) = \frac{\beta \exp\left\{-\frac{1}{2}\left(\alpha - \frac{\beta}{x}\right)^2\right\} \Phi\left(\alpha - \frac{\beta}{x}\right)^{a-1} \left\{\Phi(\alpha) - \Phi\left(\alpha - \frac{\beta}{x}\right)\right\}^{b-1}}{\sqrt{2\pi}x^2 B(a,b)\Phi(\alpha)^{a+b-1}\left[1 - I_{\left[\frac{\Phi\left(\alpha - \frac{\beta}{x}\right)}{\Phi(\alpha)}\right]}(a,b)\right]}, \tag{8}$$

respectively.

It is clear that the BA distribution is much more flexible than the alpha distribution. Plots of the density (6), cumulative distribution (7) and hazard rate function (8) for selected parameter values are displayed in Figures 4, 5 and 6, respectively.

3 Probability Weighted Moments

First proposed by Greenwood et al. (1979), PWMs are expectations of certain functions of a random variable whose mean exists. A general theory for PWMs covers the summarization and description of theoretical probability distributions and observed data samples, nonparametric estimation of the underlying distribution of an observed sample, estimation of parameters, quantiles of probability distributions and hypothesis tests. The PWM method can generally be used for estimating parameters of a distribution whose inverse form cannot be expressed explicitly. For several distributions, such as normal, log-normal and Pearson type three distributions, the expressions relating PWMs to the parameters of

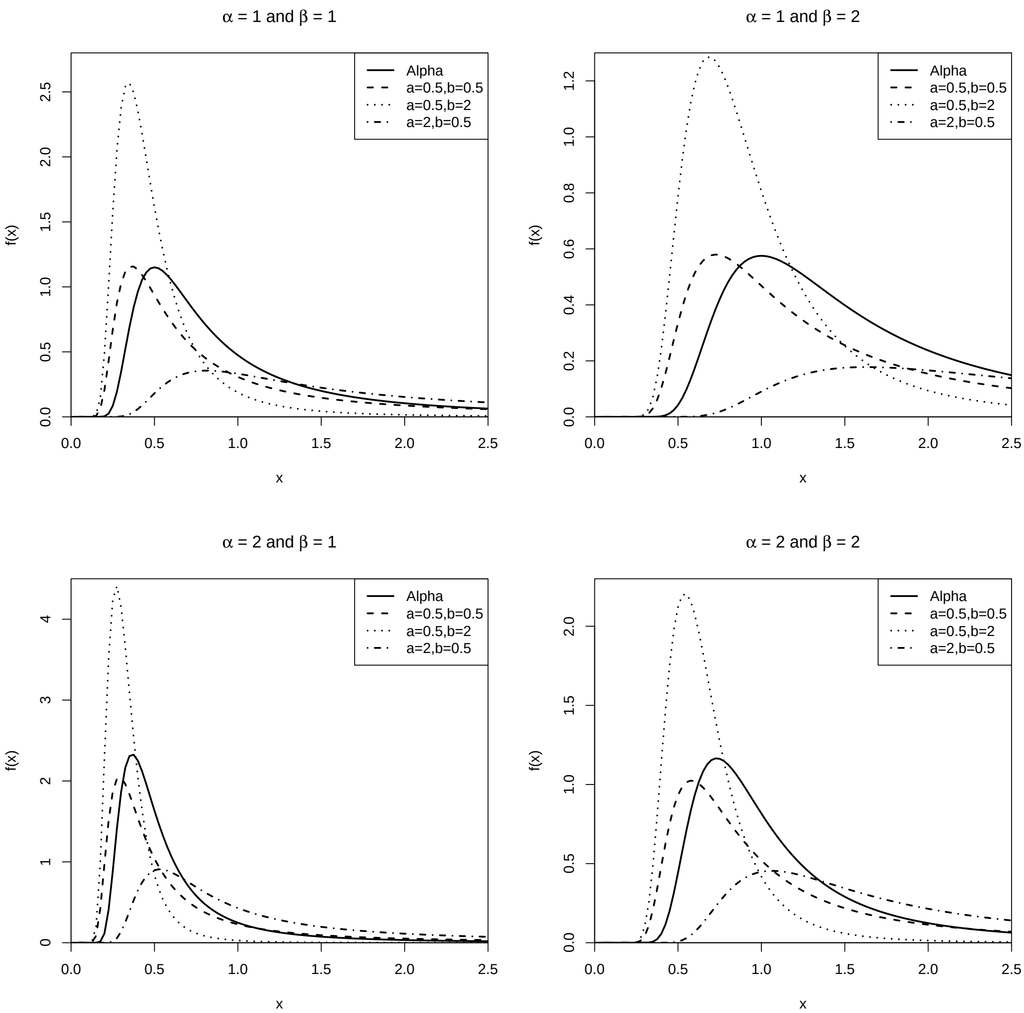

Figure 4: Plots of the BA density (6) for some parameter values.

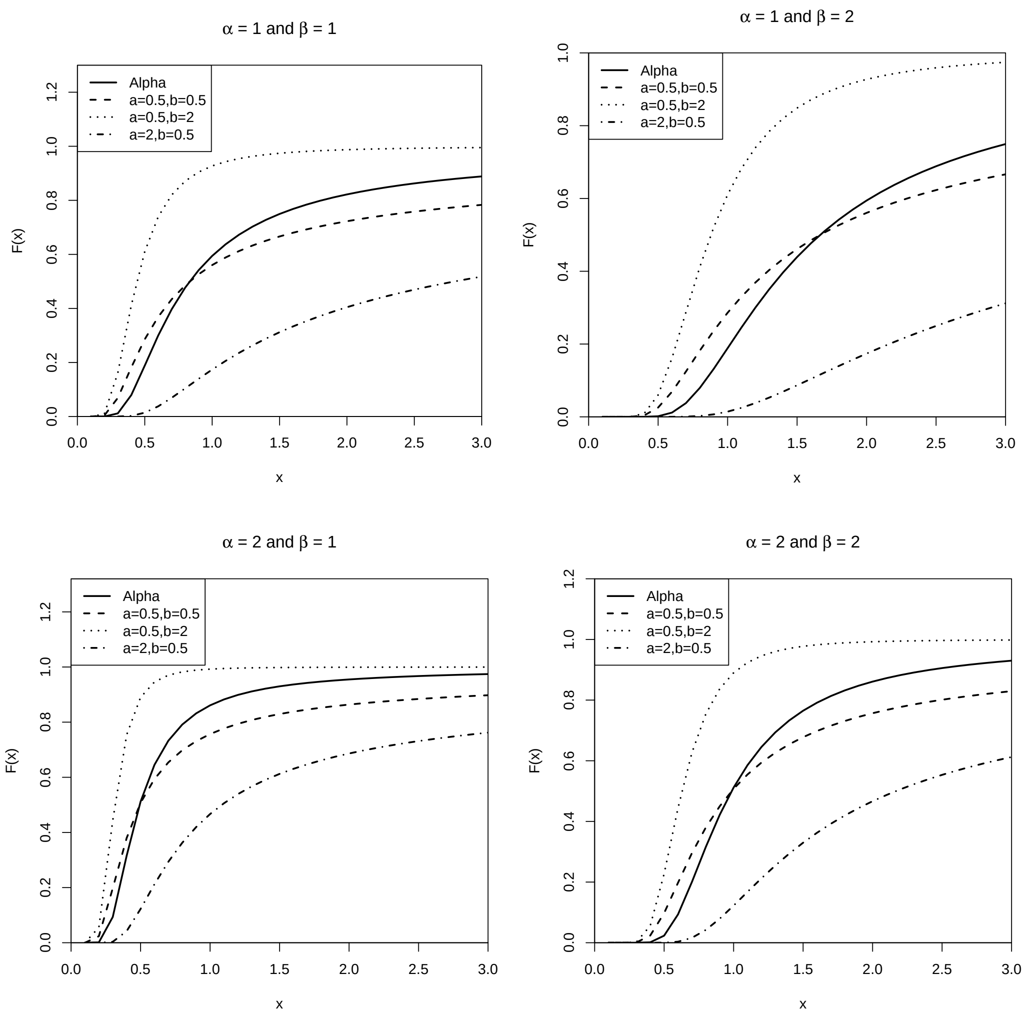

Figure 5: Plots of the BA cumulative function (7) for some parameter values.

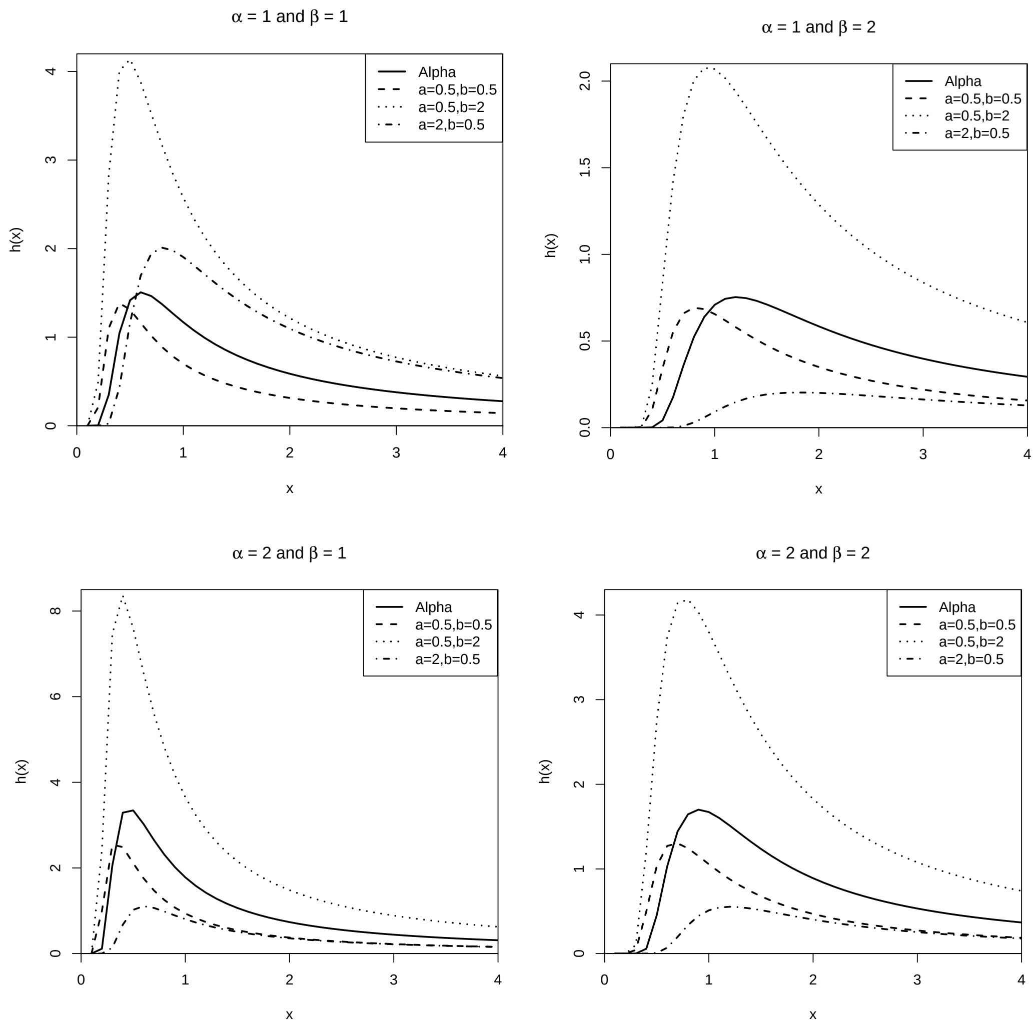

Figure 6: Plots of the BA hazard rate function (8) for some parameter values.

the model have the same forms. Such expressions may be readily employed in practice for estimating the parameters. We calculate the PWMs of the alpha distribution since they are required to obtain the ordinary moments of the BA distribution.

The PWMs of the alpha distribution are formally defined by

$$\tau_{s,r} = \int_0^\infty x^s G(x)^r g(x) dx.$$

Equations (1) and (2) lead to

$$\tau_{s,r} = c_1 \int_0^\infty x^{s-2} \left\{ \Phi\left(\alpha - \frac{\beta}{x}\right) \right\}^r \exp\left\{ -\frac{1}{2}\left(\alpha - \frac{\beta}{x}\right)^2 \right\} dx, \tag{9}$$

where $c_1 = \dfrac{\beta}{\sqrt{2\pi}\Phi(\alpha)^{r+1}}$. First, we obtain

$$\left\{ \Phi\left(\alpha - \frac{\beta}{x}\right) \right\}^r = \frac{1}{2^r}\left\{ 1 + \mathrm{erf}\left(\frac{\alpha - \beta \mathrm{x}^{-1}}{\sqrt{2}}\right) \right\}^{\mathrm{r}}$$

and the binomial expansion implies

$$\left\{ \Phi\left(\alpha - \frac{\beta}{x}\right) \right\}^r = \frac{1}{2^r} \sum_{j=0}^{r} \binom{r}{j} \mathrm{erf}\left(\frac{\alpha - \beta \mathrm{x}^{-1}}{\sqrt{2}}\right)^{\mathrm{j}}.$$

From the series expansion for the error function erf(.)

$$\mathrm{erf(x)} = \frac{2}{\sqrt{\pi}} \sum_{\mathrm{k}=0}^{\infty} \frac{(-1)^{\mathrm{k}} \mathrm{x}^{2\mathrm{k}+1}}{(2\mathrm{k}+1)\mathrm{k}!},$$

the last equation becomes

$$\left\{ \Phi\left(\alpha - \frac{\beta}{x}\right) \right\}^r = \frac{1}{2^r} \sum_{j=0}^{r} \binom{r}{j} \left\{ \sum_{k=0}^{\infty} a_k (\alpha - \beta x^{-1})^{2k+1} \right\}^j,$$

where the coefficients a_k are given by $a_k = \dfrac{(-1)^k 2^{(1-2k)/2}}{\sqrt{\pi}(2k+1)k!}$. Hence,

$$\left\{ \Phi\left(\alpha - \frac{\beta}{x}\right) \right\}^r = \frac{1}{2^r} \sum_{j=0}^{r} \binom{r}{j} \sum_{k_1,\ldots,k_j=0}^{\infty} A(k_1, ..., k_j) (\alpha - \beta x^{-1})^{2s_j + j},$$

where $A(k_1, \ldots, k_j) = a_{k_1} \ldots a_{k_j}$ and $s_j = k_1 + \cdots + k_j$. Using the binomial expansion, we have

$$\left\{ \Phi\left(\alpha - \frac{\beta}{x}\right) \right\}^r = \frac{1}{2^r} \sum_{j=0}^{r} \binom{r}{j} \sum_{k_1,\ldots,k_j=0}^{\infty} \sum_{m=0}^{2s_j+j} A(k_1, ..., k_j) \binom{2s_j + j}{m} \alpha^{2s_j+j-m} (-\beta x^{-1})^m.$$

Inserting the last equation into (9) and interchanging terms, we obtain

$$\tau_{s,r} = \frac{c_1}{2^r} \sum_{j=0}^{r} \binom{r}{j} \sum_{k_1,\ldots,k_j=0}^{\infty} \sum_{m=0}^{2s_j+j} A(k_1, ..., k_j) \binom{2s_j + j}{m} \alpha^{2s_j+j-m} (-\beta)^m I(s - m - 2, \alpha, \beta). \tag{10}$$

Here, $I(p,\alpha,\beta)$ is the integral easily obtained by

$$
\begin{aligned}
I(p,\alpha,\beta) &= \int_0^\infty x^p \exp\left\{-\frac{1}{2}\left(\alpha-\frac{\beta}{x}\right)^2\right\} dx \\
&= \frac{1}{\sqrt{\pi}}\left[2^{(-3/2-p/2)}\beta^{(p+1)}\exp^{-\alpha^2/2}\left[-\frac{1}{2}\frac{\sqrt{2}\pi^2\alpha\beta L_{p/2}^{1/2}\left(\frac{\alpha^2}{2}\right)}{\sqrt{\beta^2}\sin\left(\frac{p}{2}\pi\right)\Gamma\left(\frac{3}{2}+\frac{p}{2}\right)}\right.\right. \\
&\left.\left. -\frac{1}{2}\frac{\pi^2(\alpha^2+1)L_{p/2}^{1/2}\left(\frac{\alpha^2}{2}\right)}{\cos\left(\frac{p}{2}\pi\right)\Gamma\left(2+\frac{p}{2}\right)}+\frac{1}{2}\frac{\pi^2\alpha^2 L_{p/2}^{3/2}\left(\frac{\alpha^2}{2}\right)}{\cos\left(\frac{1}{2}\pi\right)\Gamma\left(2+\frac{p}{2}\right)}\right]\right],
\end{aligned}
$$

where $L_{p/2}^{1/2}\left(\frac{\alpha^2}{2}\right)$ and ${}_1F_1(p/2;1/2+1;\alpha^2/2)$ are the Laguerre and confluent hypergeometric functions (see Section 1). Equation (10) for the PWM of the alpha distribution is the main result of this section.

4 Moments

The cdf $F(x)$ and pdf $f(x)$ of the beta G distribution are usually straightforward to compute numerically from the baseline functions $G(x)$ and $g(x)$ from equations (4) and (5) using statistical software with numerical facilities. Here, we provide expansions for these functions in terms of infinite (or finite) weighted sums of powers of $G(x)$ which will prove useful in our case where $G(x)$ does not have a simple expression. In subsequent sections, we use these expansions to obtain formal expressions for the moments of the BA distribution and for the density of the order statistics and their moments.

For $b>0$ real non-integer, the series representation for $(1-w)^{b-1}$ yields

$$
\int_0^x w^{a-1}(1-w)^{b-1}dw = \sum_{j=0}^{\infty}\frac{(-1)^j\binom{b-1}{j}}{(a+j)}x^{a+j}, \tag{11}
$$

where the binomial coefficient is defined for any real. If b is an integer, the index j in the sum (11) stops at $b-1$. We define the constants

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

Combining equations (4) and (11), the cumulative distribution of any beta G can be written as

$$
F(x) = \frac{1}{B(a,b)}\sum_{r=0}^{\infty} w_r(a,b)G(x)^{a+r}. \tag{12}
$$

If a is an integer, equation (12) gives the cdf of the beta G distribution as an infinite sum of powers of $G(x)$. Otherwise, if a is real non-integer, we can expand $G(x)^{a+j}$ using equation (27) in the Appendix, and then the cumulative function $F(x)$ can be expressed as an infinite power series expansion of the baseline $G(x)$

$$
F(x) = \frac{1}{B(a,b)}\sum_{r=0}^{\infty} t_r(a,b)G(x)^r, \tag{13}
$$

where

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

and the quantities $s_r(a+j)$ are easily determined from equation (28) in the Appendix.

Expansions for the density of the beta G distribution are immediately obtained by simple differentiation of equations (12) and (13) for $a > 0$ integer and $a > 0$ real non-integer, respectively. We have

$$f(x) = \frac{g(x)}{B(a,b)} \sum_{r=0}^{\infty} (a+r)\, w_r(a,b) G(x)^{a+r-1} \tag{14}$$

and

$$f(x) = \frac{g(x)}{B(a,b)} \sum_{r=0}^{\infty} (r+1)\, t_{r+1}(a,b) G(x)^{r}. \tag{15}$$

Expansions (14) and (15) are the main results of this section. The sth moment of the beta G distribution can then be written as an infinite sum of convenient PWMs of the parent distribution G. These expansions are readily computed numerically using standard statistical software. They (and other expansions in the paper) can also 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. In numerical applications, a large natural number N can be used in the sums instead of infinity.

For a integer, equation (14) yields

$$E(X^s) = \sum_{r=0}^{\infty} \frac{(a+r) w_r(a,b)}{B(a,b)} \int_0^{\infty} x^s G(x)^{a+r-1} g(x) dx$$

and then

$$E(X^s) = \sum_{r=0}^{\infty} \frac{(a+r) w_r(a,b)}{B(a,b)} \tau_{s,a+r-1}.$$

For a real non-integer, equation (15) implies

$$E(X^s) = \sum_{r=0}^{\infty} \frac{(r+1) t_{r+1}(a,b)}{B(a,b)} \int_0^{\infty} x^s G(x)^r g(x) dx$$

and then

$$E(X^s) = \sum_{r=0}^{\infty} \frac{(r+1) t_{r+1}(a,b)}{B(a,b)} \tau_{s,r}.$$

From these expansions, we can obtain the moments of the BA distribution as infinite sums of certain PWMs of the alpha distribution.

5 Moment generating function

For $a > 0$ integer, the moment generating function (mgf) of the BA distribution can be determined from equation (14) as

$$M_X(t) = \sum_{r=0}^{\infty} \frac{(a+r)w_r(a,b)}{B(a,b)} \int_0^{\infty} \exp(tx)\, G(x)^{a+r-1} g(x)dx.$$

We have

$$\int_0^{\infty} \exp(tx)\, G(x)^{a+r-1} g(x)dx = \sum_{l=0}^{\infty} \frac{t^l}{l!} \int_0^{\infty} x^l G(x)^{a+r-1} g(x)dx$$

and then

$$M_X(t) = \sum_{r=0}^{\infty}\sum_{l=0}^{\infty} \frac{(r+1)t_{r+1}(a,b)t^l}{B(a,b)l!} \tau_{l,a+r-1}.$$

For $a > 0$ real non-integer, the mgf is obtained from (15). It reduces to

$$M_X(t) = \sum_{r=0}^{\infty} \frac{(r+1)t_{r+1}(a,b)}{B(a,b)} \int_0^{\infty} \exp(tx)\, G(x)^r g(x)dx$$

and then

$$\int_0^{\infty} \exp(tx)\, G(x)^r g(x)dx = \sum_{l=0}^{\infty} \frac{t^l}{l!} \int_0^{\infty} x^l G(x)^r g(x)dx.$$

Hence,

$$M_X(t) = \sum_{r=0}^{\infty}\sum_{l=0}^{\infty} \frac{(r+1)t_{r+1}(a,b)t^l}{B(a,b)l!} \tau_{l,r}.$$

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 BA distribution with cdf $F(x)$, we can derive the mean deviations about the mean $\mu = E(X)$ and about the median m from the relations

$$\delta_1 = \int_0^{\infty} |x-\mu|\, f(x)dx \text{ and } \delta_2 = \int_0^{\infty} |x-m|\, f(x)dx.$$

respectively. The median is the solution of the non-linear equation

$$I_{\left[\frac{\Phi\left(\alpha-\frac{\beta}{m}\right)}{\Phi(\alpha)}\right]}(a,b) = 1/2.$$

Defining the integral $I(s) = \int_0^s xf(x)dx$, these measures can be calculated from

$$\delta_1 = 2\mu F(\mu) - 2I(\mu) \text{ and } \delta_2 = E(X) - 2I(m), \tag{16}$$

where $F(\mu)$ is easily obtained from equation (7). We now derive formulas to obtain the integral $I(s)$. Setting

$$\rho(s,r;\alpha,\beta)=\int_0^s xg(x)G(x)^r dx,$$

we can obtain from equation (14) for $a>0$ integer

$$I(\mu)=\sum_{r=0}^{\infty}\frac{(a+r)\,w_r(a,b)}{B(a,b)}\rho(\mu,a+r-1;\alpha,\beta)$$

and from equation (15) for $a>0$ real non-integer

$$I(\mu)=\sum_{r=0}^{\infty}\frac{(r+1)\,t_{r+1}(a,b)}{B(a,b)}\rho(\mu,r;\alpha,\beta).$$

Combining (2) and (10) and defining

$$J(s,-m-1;\alpha,\beta)=\int_0^s x^{-m-1}\exp\left\{\frac{1}{2}\left(\alpha-\frac{\beta}{x}\right)^2\right\}dx,$$

we can write

$$\begin{aligned}\rho(s,r;\alpha,\beta) &= \frac{\beta}{\sqrt{2\pi}2^r\Phi^{r+1}(\alpha)}\sum_{j=0}^{r}\binom{r}{j}\sum_{k_1,\ldots,k_j=0}^{\infty}\sum_{m=0}^{2s_j+j}A(k_1,...,k_j)\\ &\quad\binom{2s_j+j}{m}\alpha^{2s_j+j-m}(-\beta)^m J(s,-m-1;\alpha,\beta),\end{aligned}$$

where s_j and $A(k_1,\ldots,k_j)$ were defined before. Setting $t=\alpha-\beta/x$, the last integral reduces to

$$J(s,-m-1;\alpha,\beta)=\int_{-\infty}^{\alpha-\beta/s}\frac{1}{\beta^m}(\alpha-t)^{m-1}\exp\left(-\frac{t^2}{2}\right)dt$$

and using the binomial expansion and interchanging terms, it becomes

$$J(s,-m-1;\alpha,\beta)=\sum_{l=0}^{\infty}\frac{(-1)^l\alpha^{m-1-l}}{\beta^m}\binom{m-1}{l}\int_{-\infty}^{\alpha-\beta/s}t^l\exp\left(-\frac{t^2}{2}\right)dt.$$

We now define

$$G(l)=\int_0^{\infty}x^l e^{-x^2/2}dx=2^{(l-1)/2}\,\Gamma(l+1/2).$$

In order to evaluate the integral in $J(s,-m-1;\alpha,\beta)$, it is necessary to consider two cases. If $\alpha-\beta/s<0$, we have

$$\int_{-\infty}^{\alpha-\beta/s}t^l\exp(-\frac{t^2}{2})dt=(-1)^lG(l)+(-1)^{l+1}\int_0^{-(\alpha-\beta/s)}t^l\exp(-\frac{t^2}{2})dt.$$

If $\alpha-\beta/s>0$, we have

$$\int_{-\infty}^{\alpha-\beta/s}t^l\exp(-\frac{t^2}{2})dt=(-1)^lG(l)+\int_0^{\alpha-\beta/s}t^l\exp(-\frac{t^2}{2})dt.$$

Further, the integrals of the type $\int_0^q x^l e^{-x^2/2}dx$ can be easily determined as Whittaker and Watson (1990).

$$\int_0^q x^l \exp\left(-\frac{x^2}{2}\right)dx = \frac{2^{l/4+1/4}q^{l/2+1/2}e^{-q^2/4}}{(l/2+1/2)(l+3)}M_{l/4+1/4,l/4+3/4}(q^2/4) + \frac{2^{l/4+1/4}q^{l/2-3/2}e^{-q^2/4}}{l/2+1/2}M_{l/4+5/4,l/4+3/4}(q^2/4),$$

where $M_{k,m}(x)$ is the Whittaker function. This function can be expressed in terms of the confluent hypergeometric function ${}_1F_1$ (see Section 1) as $M_{k,m}(x) = e^{-x/2}x^{m+1/2}{}_1F_1(\frac{1}{2}+m-k; 1+2m; x)$. Hence, we have all quantities to calculate $J(s, -m-1; \alpha, \beta)$, $\rho(s, r; \alpha, \beta)$, $I(\mu)$ and then the mean deviations (16).

7 Order Statistics

Order statistics make their appearance in many areas of statistical theory and practice. The density $f_{i:n}(x)$ of the ith order statistic for $i = 1, \ldots, n$ from data values $X_1, \ldots, X_n$ following the beta G distribution is

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

and then

$$f_{i:n}(x) = \frac{1}{B(i, n-i+1)} f(x) \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1}. \tag{17}$$

Combining (5) and (17), the density of the ith order statistic becomes

$$f_{i:n}(x) = \frac{g(x)G(x)^{a-1}\{1-G(x)\}^{b-1}}{B(a,b)B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1}. \tag{18}$$

By application of an equation in Section 0.314 of Gradshteyn and Ryzhik (2000) for power series raised to powers, we have for any j positive integer

$$\left(\sum_{i=0}^{\infty} a_i u^i\right)^j = \sum_{i=0}^{\infty} c_{j,i} u^i, \tag{19}$$

where the coefficients $c_{j,i}$ for $i = 1, 2, \ldots$ can be easily obtained from the recurrence equation

$$c_{j,i} = (ia_0)^{-1} \sum_{m=1}^{i} (jm - i + m) a_m c_{j,i-m}, \tag{20}$$

with $c_{j,0} = a_0^j$. The coefficient $c_{j,i}$ comes from $c_{j,0}, \ldots, c_{j,i-1}$ and therefore are obtained from $a_0, \ldots, a_i$. The coefficients $c_{j,i}$ can be given explicitly in terms of the quantities a_i, although it is not necessary for programming numerically our expansions in any algebraic or numerical software. We now use equations (19) and (20).

For $a > 0$ integer, we have

$$F(x)^{i+j-1} = \left(\frac{G(x)^a}{B(a,b)}\right)^{i+j-1} \left(\sum_{r=0}^{\infty} w_r G(x)^r\right)^{i+j-1}.$$

Substituting $u = G(x)$ and using equations (19) and (20) yields

$$f_{i:n}(x) = \sum_{j=0}^{n-i}(-1)^j \frac{g(x)u^{a-1}(1-u)^{b-1}}{B(a,b)^{i+j}B(i,n-i+1)} \binom{n-i}{j} u^{a(i+j-1)} \sum_{r=0}^{\infty} c_{i+j-1,r}u^r.$$

For $b > 0$ real non-integer, we have

$$f_{i:n}(x) = \sum_{j=0}^{n-i}\sum_{r=0}^{\infty}(-1)^j c_{i+j-1,r} \binom{n-i}{j} \frac{g(x)(1-u)^{b-1}u^{r+a(i+j)-1}}{B(a,b)^{i+j}B(i,n-i+1)},$$

and then the power series for $(1-u)^{b-1}$ gives

$$f_{i:n}(x) = \sum_{j=0}^{n-i}\sum_{r,l=0}^{\infty}(-1)^{j+l} c_{i+j-1,r} \binom{n-i}{j} \frac{g(x)G(x)^{r+l+a(i+j)-1}}{B(a,b)^{i+j}B(i,n-i+1)}, \tag{21}$$

where

$$c_{i+j-1,r} = (rw_0)^{-1}\sum_{m=1}^{r}\{(i+j)m-r\}w_m c_{i+j-1,r-m}. \tag{22}$$

If b is a integer, the index l in the sum (21) stops at $b-1$.

For $a > 0$ real non-integer, we have

$$F(x)^{i+j-1} = \left(\frac{1}{B(a,b)}\right)^{i+j-1} \left(\sum_{r=0}^{\infty} t_r G(x)^r\right)^{i+j-1}.$$

In the same way, using equations (19) and (20), it follows

$$f_{i:n}(x) = \sum_{j=0}^{n-i}(-1)^j \frac{g(x)u^{a-1}(1-u)^{b-1}}{B(a,b)^{i+j}B(i,n-i+1)} \binom{n-i}{j} \sum_{r=0}^{\infty} c_{i+j-1,r}u^r$$

For $b > 0$ real non-integer, we have

$$f_{i:n}(x) = \sum_{j=0}^{n-i}\sum_{r=0}^{\infty}(-1)^j d_{i+j-1,r} \binom{n-i}{j} \frac{g(x)(1-u)^{b-1}u^{r+a-1}}{B(a,b)^{i+j}B(i,n-i+1)},$$

and the power series for $(1-u)^{b-1}$ yields

$$f_{i:n}(x) = \sum_{j=0}^{n-i}\sum_{r,l=0}^{\infty}(-1)^{j+l}\ d_{i+j-1,r} \binom{n-i}{j} \frac{g(x)G(x)^{r+l+a-1}}{B(a,b)^{i+j}B(i,n-i+1)}, \tag{23}$$

where

$$d_{i+j-1,r} = (rt_0)^{-1}\sum_{m=1}^{r}\{(i+j)m-r\}t_m d_{i+j-1,r-m}. \tag{24}$$

If b is a integer, the index l in the sum (23) stops at $b-1$.

Equations (21) (for $a > 0$ integer) and (23) (for $a > 0$ real non-integer) are the main results of this section.

8 Moments of order statistics

The sth ordinary moment of the ith order statistic, say $X_{i:n}$, for $a > 0$ integer, follows from equation (21)

$$E(X_{i:n}^s) = \sum_{j=0}^{n-i} \sum_{r,l=0}^{\infty} \frac{(-1)^{j+l} c_{i+j-1,r} \binom{n-i}{j}}{B(a,b)^{i+j} B(i, n-i+1)} \tau_{s,r+a(i+j)+l-1}, \tag{25}$$

where the coefficient $c_{i+j-1,r}$ is defined in equation (22). If b is an integer, the index l in the above sum stops at $b-1$. For $a > 0$ real non-integer, equation (23) gives

$$E(X_{i:n}^s) = \sum_{j=0}^{n-i} \sum_{r,l=0}^{\infty} \frac{(-1)^{j+l} d_{i+j-1,r} \binom{n-i}{j}}{B(a,b)^{i+j} B(i, n-i+1)} \tau_{s,r+a+l-1}, \tag{26}$$

where $d_{i+j-1,r}$ is defined in equation (24). If b is an integer, the index l in the above sum stops at $b-1$.

Expansions (25) and (26) are the main results of this section. The L-moments are analogous to the ordinary moments but can be estimated by linear combinations of order statistics. They are linear functions of expected order statistics defined by Hosking (1990)

$$\lambda_{r+1} = r(r+1)^{-1} \sum_{k=0}^{r} \frac{(-1)^k}{k} E(X_{r+1-k:r+1}),\ r = 0, 1, \ldots$$

The first four L-moments are $\lambda_1 = E(X_{1:1})$, $\lambda_2 = \frac{1}{2}E(X_{2:2} - X_{1:2})$, $\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3})$ and $\lambda_4 = \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4})$. The L-moments have the advantage that they exist whenever the mean of the distribution exists, even though some higher moments may not exist, and are relatively robust to the effects of outliers. Setting $s = 1$ in equations (25) and (26), the L-moments follow easily from the means of the order statistics for $a > 0$ integer and $a > 0$ real non-integer, respectively.

9 Estimation and Inference

Consider that X follows the BA distribution and let $\theta = (\alpha, \beta, a, b)^T$ be the parameter vector. The log-likelihood $\ell = \ell(\alpha, \beta, a, b)$ for a single observation x of X is

$$\begin{aligned} \ell &= \log(\beta) - 2\log(x) - \log\{B(a,b)\} + \log\left(\frac{\phi(t)}{\Phi(\alpha)}\right) \\ &+ (a-1)\log\left(\frac{\Phi(t)}{\Phi(\alpha)}\right) + (b-1)\log\left(1 - \frac{\Phi(t)}{\Phi(\alpha)}\right), \quad x > 0, \end{aligned}$$

where $t = (\alpha - \beta/x)$. The unit score vector $U = (\frac{\partial \ell}{\partial \alpha}, \frac{\partial \ell}{\partial \beta}, \frac{\partial \ell}{\partial a}, \frac{\partial \ell}{\partial b})^T$ has components

$$\begin{aligned} \frac{\partial \ell}{\partial \alpha} &= -t - \frac{\phi(\alpha)}{\Phi(\alpha)} + (a-1)\left\{\frac{\phi(t)}{\Phi(t)} - \frac{\phi(\alpha)}{\Phi(\alpha)}\right\} \\ &- (b-1)\left\{\frac{\phi(t) - \phi(\alpha)}{\Phi(\alpha) - \Phi(t)} - \frac{\phi(\alpha)}{\Phi(\alpha)}\right\}, \\ \frac{\partial \ell}{\partial \beta} &= \frac{1}{\beta} + \frac{t}{x} - (a-1)\left\{\frac{\phi(t)}{x\Phi(t)}\right\} + (b-1)\left[\frac{\phi(t)}{y\{\Phi(\alpha) - \Phi(t)\}}\right], \\ \frac{\partial \ell}{\partial a} &= \log\left(\frac{\Phi(t)}{\Phi(\alpha)}\right) + \psi(a+b) - \psi(a), \\ \frac{\partial \ell}{\partial b} &= \log\left(1 - \frac{\Phi(t)}{\Phi(\alpha)}\right) + \psi(a+b) - \psi(b). \end{aligned}$$

The expected value of the score vector vanishes and then

$$E\left\{\frac{\phi(t)}{\Phi(t)}\right\}=\frac{\phi(\alpha)}{\Phi(\alpha)},\quad E\left\{\frac{\phi(\alpha)-\phi(t)}{\Phi(\alpha)-\Phi(t)}\right\}=\frac{\phi(\alpha)}{\Phi(\alpha)},$$
$$E\left\{\frac{\phi(t)}{x\Phi(t)}\right\}=0 \text{ and } E\left[\frac{\phi(t)}{x\{\Phi(\alpha)-\Phi(t)\}}\right]=0.$$

For a random sample $x=(x_1,\ldots,x_n)$ of size n from X, the total log-likelihood is $\ell_n=\ell_n(\alpha,\beta,a,b)=\sum_{i=1}^n \ell^{(i)}$, where $\ell^{(i)}$ is the log-likelihood for the ith observation ($i=1,\ldots,n$). The total score function is $U_n=\sum_{i=1}^n U^{(i)}$, where $U^{(i)}$ has the form given before for $i=1,\ldots,n$. The MLE $\hat\theta$ of θ is obtained numerically from the nonlinear equations $U_n=0$. For interval estimation and hypothesis testing on the parameters in θ we require the 4×4 unit information matrix

$$K=K(\theta)=\begin{bmatrix}\kappa_{\alpha,\alpha} & \kappa_{\alpha,\beta} & \kappa_{\alpha,a} & \kappa_{\alpha,b}\\ \kappa_{\beta,\alpha} & \kappa_{\beta,\beta} & \kappa_{\beta,a} & \kappa_{\beta,b}\\ \kappa_{a,\alpha} & \kappa_{a,\beta} & \kappa_{a,a} & \kappa_{a,b}\\ \kappa_{b,\alpha} & \kappa_{b,\beta} & \kappa_{b,a} & \kappa_{b,b}\end{bmatrix}.$$

We define the following expectations $m_{r,s}=E\left[\frac{1}{y^r}\left\{\frac{\phi(t)}{\Phi(t)}\right\}^s\right]$ for $r=0,\ldots,3$ and $s=1,2$ and $n_{r,s,t}=E\left[\frac{\phi(t)^s}{y^r\{\Phi(\alpha)-\Phi(t)\}^t}\right]$ for $r=0,\ldots,3$ and $s,t=1,2$ which can be obtained by numerical integration. The elements of the information matrix K are given by

$$\begin{aligned}
\kappa_{\alpha,\alpha} &= 1-\left[\frac{\alpha\phi(\alpha)}{\Phi(\alpha)}+\left\{\frac{\phi(\alpha)}{\Phi(\alpha)}\right\}^2\right]+(a+b-1)\left\{\frac{\phi(\alpha)}{\Phi(\alpha)}\right\}^2+(a-1)\,m_{0,1}^2\\
&+ (b-1)E\left\{\frac{\phi(\alpha)-\phi(t)}{\Phi(\alpha)-\Phi(t)}\right\}^2,\\
\kappa_{\alpha,\beta} &= -\left\{\alpha+\frac{\phi(\alpha)}{\Phi(\alpha)}\right\}\frac{1}{\beta}+(a-1)\beta\,m_{2,1}-(a-1)\,m_{1,2}-(b-1)\beta\,n_{2,1,2}\\
&+ (b-1)\phi(\alpha)\,n_{0,1,2}-(b-1)\,n_{0,2,2},\\
\kappa_{\alpha,a} &= 0,\quad \kappa_{\alpha,b}=0,\\
\kappa_{\beta,\beta} &= \frac{1}{\beta^2}+\frac{1}{\beta}\left[1-\alpha\left\{\alpha+\frac{\phi(\alpha)}{\Phi(\alpha)}\right\}\right]+(a-1)\,(\alpha\,m_{2,1}-\beta\,m_{3,1})-(b-1)\\
&\times (\alpha\,n_{2,1,1}+\beta\,n_{3,1,1}+n_{1,1,1}^2),\\
\kappa_{\beta,a} &= 0,\quad \kappa_{\beta,b}=0,\quad \kappa_{a,a}=\psi'(a)-\psi'(a+b),\quad \kappa_{b,b}=\psi'(b)-\psi'(a+b),\\
\kappa_{a,b} &= -\psi'(a+b).
\end{aligned}$$

Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of

$$\sqrt{n}(\hat\theta-\theta)\ \text{ is }\ N_4(0,K(\theta)^{-1}).$$

The asymptotic multivariate normal $N_4(0,K_n(\hat\theta)^{-1})$ 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 γ for each parameter θ_i is

$$ACI(\theta_i,100(1-\gamma)\%)=(\hat\theta_i-z_{\gamma/2}\sqrt{\kappa^{\theta_i,\theta_i}},\hat\theta_i+z_{\gamma/2}\sqrt{\kappa^{\theta_i,\theta_i}}),$$

where $\kappa^{\theta_i,\theta_i}$ is the ith diagonal element of $K_n(\theta)^{-1}$ for $i = 1, \ldots, 4$ 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 BA distribution and for comparing this distribution with some of its special sub-models. 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 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 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 $w > \xi_\gamma$, where ξ_γ denotes the upper $100\gamma\%$ point of the χ_q^2 distribution. For example, we can check if the fit using the BA distribution is statistically "superior" to a fit using the alpha distribution for a given data set by testing $H_0 : a = b = 1$ versus $H_A : H_0$ is not true.

10 Application

In this section we compare the results of fitting the BA distribution, alpha and Birnbaum-Saunders (BS) to two real data sets.

10.1 Data set Glass fibres

The data set studied by Smith, Naylor (1987), which represent the strengths of 1.5 cm glass fibres, measured at the National Physical Laboratory, England. Unfortunately, the units of measurement are not given in the paper. The data set is: 0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2 ,0.74, 1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11, 1.28, 1.42, 1.5, 1.54, 1.6, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29, 1.48, 1.5, 1.55, 1.61, 1.62, 1.66, 1.7, 1.77, 1.84, 0.84, 1.24, 1.3, 1.48, 1.51, 1.55, 1.61, 1.63, 1.67, 1.7, 1.78, 1.89.

The MLEs and the maximized log-likelihood $\hat{l}_{BA}$ for the BA distribution are

$$\hat{\alpha} = 0.0100,\ \hat{\beta} = 6.5302,\ \hat{a} = 0.1542,\ \hat{b} = 900.2565 \text{ and } \hat{l}_{BA} = -22.0736,$$

whereas the MLEs and the maximized log-likelihood $\hat{l}_A$ for the alpha distribution are

$$\hat{\alpha} = 3.0217,\ \ \hat{\beta} = 4.2616 \ \text{ and } \ \hat{l}_A = -45.6037$$

and the MLEs and the maximized log-likelihood $\hat{l}_{BS}$ for the BS distribution are

$$\hat{\alpha} = 0.2621,\ \ \hat{\beta} = 1.4566 \ \text{ and } \ \hat{l}_{B-S} = -28.5305.$$

The LR statistic for testing the hypotheses H_0 : alpha $\times$ H_A : BA is $w = 2\{-22.0736 - (-45.6037)\} = 47.0602$ (p-value $= 6.0401 \times 10^{-11}$) and, therefore, we reject the alpha distribution in favor of the BA distribution at significance level of 5%.

We also compute the values of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) for the BA, alpha and BS models. We obtain the values $AIC = 52.1473$ and $BIC = 60.7198$ for the BA model, $AIC = 95.2073$ and $BIC = 99.4936$ for the alpha model and $AIC = 61.0606$ and $BIC = 65.3473$ for the BS model. These results indicate that the BA model has the lowest values for the AIC and BIC statistics among the fitted models, and therefore it could be chosen as the best model.

10.2 Data set Cattle of Nelore

The data set is obtained from Gusmão (2008). The commercial production of beef cattle in Brazil, which usually comes from cattle of Nelore, seeks to optimize the process trying to get a short time to reach cattle the specific weight in the period from birth to weaning or from weaning to slaughter.For the data with 155 bulls of the Nelore study time (in days) until the animals reach a weight of 160kg for the period from birth to weaning, we use only 69 of these 155 animals. The data set is: 138, 140, 141, 143, 145, 146, 146, 148, 148, 149, 149, 150, 151, 151, 151, 151, 152, 152, 153, 153, 153, 155, 156, 156, 157, 158, 158, 159, 159, 159, 159, 159, 159, 160, 161, 161, 161, 162, 163, 163, 163, 163, 164, 164, 164, 164, 165, 166, 166, 166, 166, 167, 167, 170, 170, 172, 172, 173, 174, 176, 179, 179, 180, 183, 184, 185, 187, 189, 197.

The MLEs and the maximized log-likelihood $\hat{l}_{BA}$ for the BA distribution are

$$\hat{\alpha} = 3.90479,\ \hat{\beta} = 331.20523,\ \hat{a} = 241.7684,\ \hat{b} = 8.5502 \,\text{and}\, \hat{l}_{BA} = -269.4547,$$

whereas for the alpha distribution are

$$\hat{\alpha} = 13.4166,\ \ \hat{\beta} = 2155.7369\ \text{ and }\ \hat{l}_A = -272.5163$$

and for the BS distribution are

$$\hat{\alpha} = 0.0100,\ \ \hat{\beta} = 170.6750\ \text{ and }\ \hat{l}_{BS} = -3208.401.$$

The LR statistic for testing the hypotheses $H_0 : alpha \times H_A : BA$ is 6.1232 (p-value= 0.0468) and, therefore, we reject the alpha distribution in favor of the BA distribution at significance level of 5%. The values $AIC = 542.9094$ and $BIC = 548.1396$ for the BA distribution, $AIC = 546.9094$ and $BIC = 553.5008$ for the alpha distribution and $AIC = 6420.802$ and $BIC = 6425.27$ for the BS distribution indicate once again that the BA model has the lowest AIC and BIC values and then it could be taken as the best model.

11 Conclusions

We introduce the four-parameter beta alpha (BA) distribution which generalizes the alpha distribution proposed by Salvia (1985). 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 new distribution is quite flexible in analyzing positive data in place of gamma, Weibull, Birnbaum-Saunders, generalized exponential and beta exponential distributions. It is useful to model asymmetric data to the right and uni-modal distributions. We provide a mathematical treatment of the distribution including expansions for the cumulative and density functions, moment generating function, ordinary moments, mean deviations, moments of order statistics and L-moments. The estimation of parameters is approached by the method of maximum likelihood and the expected information matrix is derived. Two applications of the BA distribution are given to show that this distribution could give better fit than other statistical models widely used in lifetime analysis.

Appendix

We derive an expansion for $G(x)^\rho$ which holds for $\rho > 0$ real non-integer. We can write

$$G(x)^\rho = [1 - \{1 - G(x)\}]^\rho = \sum_{j=0}^{\infty} \binom{\rho}{j} (-1)^j \{1 - G(x)\}^j$$

and then

$$G(x)^\rho = \sum_{j=0}^{\infty} \sum_{r=0}^{j} (-1)^{j+r} \binom{\rho}{j} \binom{j}{r} G(x)^r.$$

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

$$G(x)^\rho = \sum_{r=0}^{\infty} \sum_{j=r}^{\infty} (-1)^{j+r} \binom{\rho}{j} \binom{j}{r} G(x)^r.$$

and then

$$G(x)^\rho = \sum_{r=0}^{\infty} s_r(\rho) G(x)^r, \qquad (27)$$

where

$$s_r(\rho) = \sum_{j=r}^{\infty} (-1)^{r+j} \binom{\rho}{j} \binom{j}{r}. \qquad (28)$$

Equations (27) and (28) are used in Section 4.

References

Eugene, N., Lee, C., Famoye, F. (2002). Beta-normal distribution and its applications. Commun. Statist. - Theory and Methods 31:497-512.

Gradshteyn, I. S., Ryzhik, I. M. (2000). Table of integrals, series, and products. Academic Press, San Diego.

Greenwood, J., Landwehr, J., Matalas, N. (1979). Probability Weighted Moments: Definition and Relation to parameters of Several Distributions Expressable in Inverse Form. Water Resources Research vol 15 No 5, 1049-1054.

Gusmão, S. F. R. (2008). Uma abordagem bayesiana: distribution Weibull inversa generalizada. Dissertação de Mestrado. UFRPE. Recife.

Hosking, J. R. M. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. J. Royal Statist. Soc. B 52:105-124.

Katsev, P. G. (1968). Statistical methods in the study of cutting tools(Russian), Mashinostroyeniye, Moscow.

Nadarajah, S., Gupta, A. K. (2004). The beta Fréchet distribution. Far East Journal of Theoretical Statistics 14:15-24.

Nadarajah, S., Kotz, S. (2004). The beta Gumbel distribution. Math. Probab. Eng. 10:323-332.

Nadarajah, S. and Kotz, S. (2005). The beta exponential distribution. Reliability Engineering and System Safety 91:689-697.

Salvia, A. A. (1985). Reliability application of the Alpha distribution. IEEE Transactions on Reliability. 34:251-252.

Sherif, Y. S. (1983). Models for accelerated life-testing. Proc. Ann. Reliability and Maintainability Symp.

Smith, R. L. and Naylor, J. C. (1987). Comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. Applied Statistics. 36: 358-369.

Wager, J. G. and Barash, M. M. (1971). Study of the distribution of the life of HSS tools. ASME Transactions. 93:1044-1050.

Whittaker, E. T. and Watson, G. N. A. (1990). Course in Modern Analysis, 4th ed. Cambridge University Press, Cambridge.