

The Beta-Half-Cauchy Distribution

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

In this article, based on the half-Cauchy distribution, we propose the called beta-half-Cauchy distribution for modeling lifetime data. Various explicit expressions for its moments, generating and quantile functions, mean deviations, density function of the order statistics and their moments are provided. The parameters of the new model are estimated by maximum likelihood and the observed information matrix is derived. An application to lifetime real data shows that it can yield a better fit than three- and two-parameter Birnbaum–Saunders, gamma and Weibull models.

Keywords: Half-Cauchy distribution; Lifetime data; Maximum likelihood estimation; Mean deviation; Quantile function; Shape parameter.

1 Introduction

The statistics literature is filled with hundreds of continuous univariate distributions (see, for example, Johnson et al., 1994, 1995). Numerous classical distributions have been extensively used over the past decades for modeling data in several areas such as engineering, actuarial, environmental and medical sciences, biological studies, demography, economics, finance and insurance. However, in many applied areas like lifetime analysis, finance and insurance, there is a clear need for extended forms of these distributions, that is, new distributions which are more flexible to model real data in these areas since the data can present a high degree of skewness and kurtosis. So, we can give additional control over both skewness and kurtosis by adding new parameters, and hence the extended distributions become more flexible to model real data. Recent developments focus on new techniques for building meaningful distributions, including the generator approach pioneered by Eugene et al. (2002). In particular, these authors introduced the beta normal (BN) distribution, denoted by $BN(\mu, \sigma, a, b)$, where $\mu \in \mathbb{R}$, $\sigma > 0$ and a and b are positive shape parameters. These parameters control skewness through

the relative tail weights. The BN distribution is symmetric if $a = b$, it has negative skewness when $a < b$ and positive skewness when $a > b$. For $a = b > 1$, it has positive excess kurtosis and, for $a = b < 1$, it has negative excess kurtosis (Eugene et al., 2002). An application of this distribution to dose-response modeling is presented in Razzaghi (2009).

In this note, we use the generator approach suggested by Eugene et al. (2002) to define a new model called the beta-half-Cauchy (BHC) distribution, which extends the half-Cauchy (HC) model. In addition, we investigate some mathematical properties of the new model, discuss maximum likelihood estimation of its parameters and derive the observed information matrix. The proposed model is much more flexible than the HC distribution and can be used effectively for modeling lifetime data.

The HC distribution is derived from the Cauchy distribution by mirroring the curve on the origin so that only positive values can be observed. Its cumulative distribution function (cdf) is

$$G_\phi(t) = \frac{2}{\pi} \arctan\left(\frac{t}{\phi}\right), \quad t > 0, \quad (1)$$

where $\phi > 0$ is a scale parameter. The probability density function (pdf) corresponding to (1) is

$$g_\phi(t) = \frac{2}{\pi\phi} \left[1 + \left(\frac{t}{\phi}\right)^2\right]^{-1}, \quad t > 0. \quad (2)$$

For $k < 1$, the k th moment comes from (2) as $\mu'_k = \phi^k \sec\left(\frac{k\pi}{2}\right)$. As a heavy-tailed distribution, the HC distribution has been used as an alternative to model dispersal distances (Shaw, 1995), since the former predicts more frequent long-distance dispersal events than the latter. Additionally, Paradis et al. (2002) used the HC distribution to model ringing data on two species of tits (*Parus caeruleus* and *Parus major*) in Britain and Ireland.

The article is outlined as follows. In Section 2, we introduce the BHC distribution and plot the density and hazard rate functions. Explicit expressions for the density and cumulative functions, moments, moment generating function (mgf), a power series expansion for the quantile function, mean deviations, order statistics and Rényi entropy are derived in Section 3. In Section 4, we discuss maximum likelihood estimation and inference. An application in Section 5 shows the usefulness of the new distribution for lifetime data modeling. Finally, concluding remarks are addressed in Section 6.

2 The BHC distribution

Consider starting from an arbitrary baseline cumulative function $G(t)$, Eugene et al. (2002) demonstrated that any parametric family of distributions can be incorporated into larger families through an application of the probability integral transform. They defined the beta generalized (beta-G) cumulative distribution by

$$F(t) = I_{G(t)}(a, b) = \frac{1}{B(a, b)} \int_0^{G(t)} \omega^{a-1} (1 - \omega)^{b-1} d\omega, \quad (3)$$

where $a > 0$ and $b > 0$ are additional shape parameters whose role is to introduce skewness and to vary tail weight, $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function, $\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt$ is the gamma function, $I_y(a, b) = B_y(a, b)/B(a, b)$ is the incomplete beta function ratio and $B_y(a, b) = \int_0^y \omega^{a-1} (1-\omega)^{b-1} d\omega$ is the incomplete beta function. This mechanism for generating distributions from equation (3) is particularly attractive when $G(t)$ has a closed-form expression. One major benefit of the beta-G distribution is its ability of fitting skewed data that cannot be properly fitted by existing distributions.

The density function corresponding to (3) is

$$f(t) = \frac{g(t)}{B(a, b)} G(t)^{a-1} \{1 - G(t)\}^{b-1}, \quad (4)$$

where $g(t) = dG(t)/dt$ is the baseline density function. The density function $f(t)$ will be most tractable when both functions $G(t)$ and $g(t)$ have simple analytic expressions. Except for some special choices of these functions, $f(t)$ could be too complicated to deal with in full generality.

By using the probability integral transform (3), some beta-G distributions have been proposed in the last few years. In particular, Eugene et al. (2002), Nadarajah and Gupta (2004), Nadarajah and Kotz (2004), Nadarajah and Kotz (2006), Lee et al. (2007) and Akinsete et al. (2008) defined the BN, beta Fréchet, beta Gumbel, beta exponential, beta Weibull and beta Pareto distributions by taking $G(t)$ to be the cdf of the normal, Fréchet, Gumbel, exponential, Weibull and Pareto distributions, respectively. More recently, Barreto-Souza et al. (2010), Pescim et al. (2010), Silva et al. (2010), Paranaíba et al. (2011) and Cordeiro and Lemonte (2011a,b) defined the beta generalized exponential, beta generalized half-normal, beta modified Weibull, beta Burr XII, beta Birnbaum-Saunders and beta Laplace distributions, respectively.

In the same way, we can extend the HC distribution because it has closed-form cumulative function. By inserting (1) and (2) in (4), the BHC density function (for $t > 0$) with three positive parameters ϕ , a and b , say $\text{BHC}(\phi, a, b)$, follows as

$$f(t) = \frac{2^a}{\phi \pi^a B(a, b)} \left[1 + \left(\frac{t}{\phi} \right)^2 \right]^{-1} \left[\arctan \left(\frac{t}{\phi} \right) \right]^{a-1} \left\{ 1 - \frac{2}{\pi} \left[\arctan \left(\frac{t}{\phi} \right) \right] \right\}^{b-1}. \quad (5)$$

Evidently, the density function (5) does not involve any complicated function. Also, there is no functional relationship between the parameters and they vary freely in the parameter space. The density function (5) extends a few known distributions. The HC distribution arises as the basic exemplar when $a = b = 1$. The new model called the exponentiated half-Cauchy (EHC) distribution is obtained when $b = 1$. For a and b positive integers, the BHC density function reduces to the density function of the a th order statistic from the HC distribution in a sample of size $a + b - 1$. However, equation (5) can also alternatively be extended, when a and b are real non-integers, to define fractional HC order statistic distributions.

The cdf and hazard rate function corresponding to (5) are

$$F(t) = I_{\frac{2}{\pi} \arctan(\frac{t}{\phi})}(a, b) \quad (6)$$

and

$$h(t) = \frac{2^a}{\phi \pi^a B(a, b)} \frac{\left[\arctan\left(\frac{t}{\phi}\right) \right]^{a-1} \left\{ 1 - \frac{2}{\pi} \left[\arctan\left(\frac{t}{\phi}\right) \right] \right\}^{b-1}}{\left[1 + \left(\frac{t}{\phi}\right)^2 \right] \left[1 - I_{\frac{2}{\pi} \arctan\left(\frac{t}{\phi}\right)}(a, b) \right]}, \quad (7)$$

respectively.

The BHC distribution can present several forms depending on the parameter values. In Figure 1, we illustrate some possible shapes of the density function (5) for selected parameter values. From Figure 1, we can see how changes in the parameters a and b modify the form of the density function. It is evident that the BHC distribution is much more flexible than the HC distribution. Plots of the hazard rate function (7) for some parameter values are shown in Figure 2. The new model is easily simulated as follows: if V is a beta random variable with parameters a and b , then $T = \phi \tan(\pi V/2)$ has the $\text{BHC}(\phi, a, b)$ distribution. This scheme is useful because of the existence of fast generators for beta random variables in statistical software.

3 Properties

In this section, we study some structural properties of the BHC distribution.

3.1 Expansion for the Density Function

The cdf $F(t)$ and pdf $f(t)$ of the beta-G distribution are usually straightforward to compute numerically from the baseline functions $G(t)$ and $g(t)$ from (3) and (4) using statistical software with numerical facilities. However, we provide expansions for these functions in terms of infinite (or finite if both a and b are integers) power series of $G(t)$ that can be useful when this function does not have a simple expression.

Expansions for the beta-G cumulative function are given by Cordeiro and Lemonte (2011a) and follow immediately from (3) (for $b > 0$ real non-integer) as

$$F(t) = \frac{1}{B(a, b)} \sum_{r=0}^{\infty} w_r G(t)^{a+r}, \quad (8)$$

where $w_r = (-1)^r (a+r)^{-1} \binom{b-1}{r}$. If b is an integer, the index r in (8) stops at $b-1$. If a is an integer, equation (8) gives the beta-G cumulative distribution as a power series of $G(t)$. Otherwise, if a is a real non-integer, we can expand $G(t)^a$ as

$$G(t)^a = \sum_{r=0}^{\infty} s_r(a) G(t)^r, \quad (9)$$

where $s_r(a) = \sum_{j=r}^{\infty} (-1)^{r+j} \binom{a}{j} \binom{j}{r}$, and then $F(t)$ can be expressed from (8) and (9) as

$$F(t) = \frac{1}{B(a, b)} \sum_{r=0}^{\infty} t_r G(t)^r, \quad (10)$$

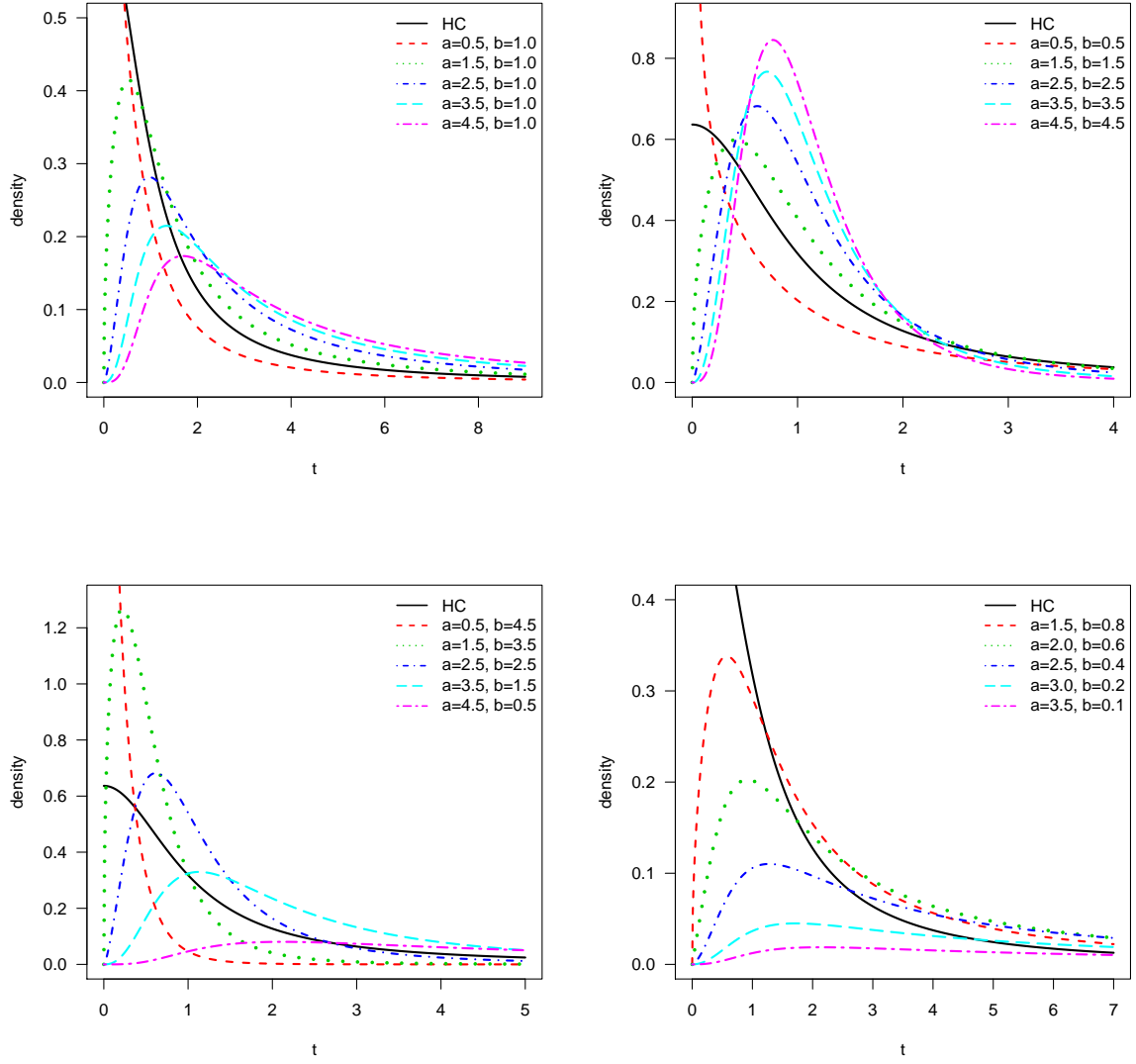


Figure 1: Plots of the density function (5) for some parameter values; $\phi = 1$.

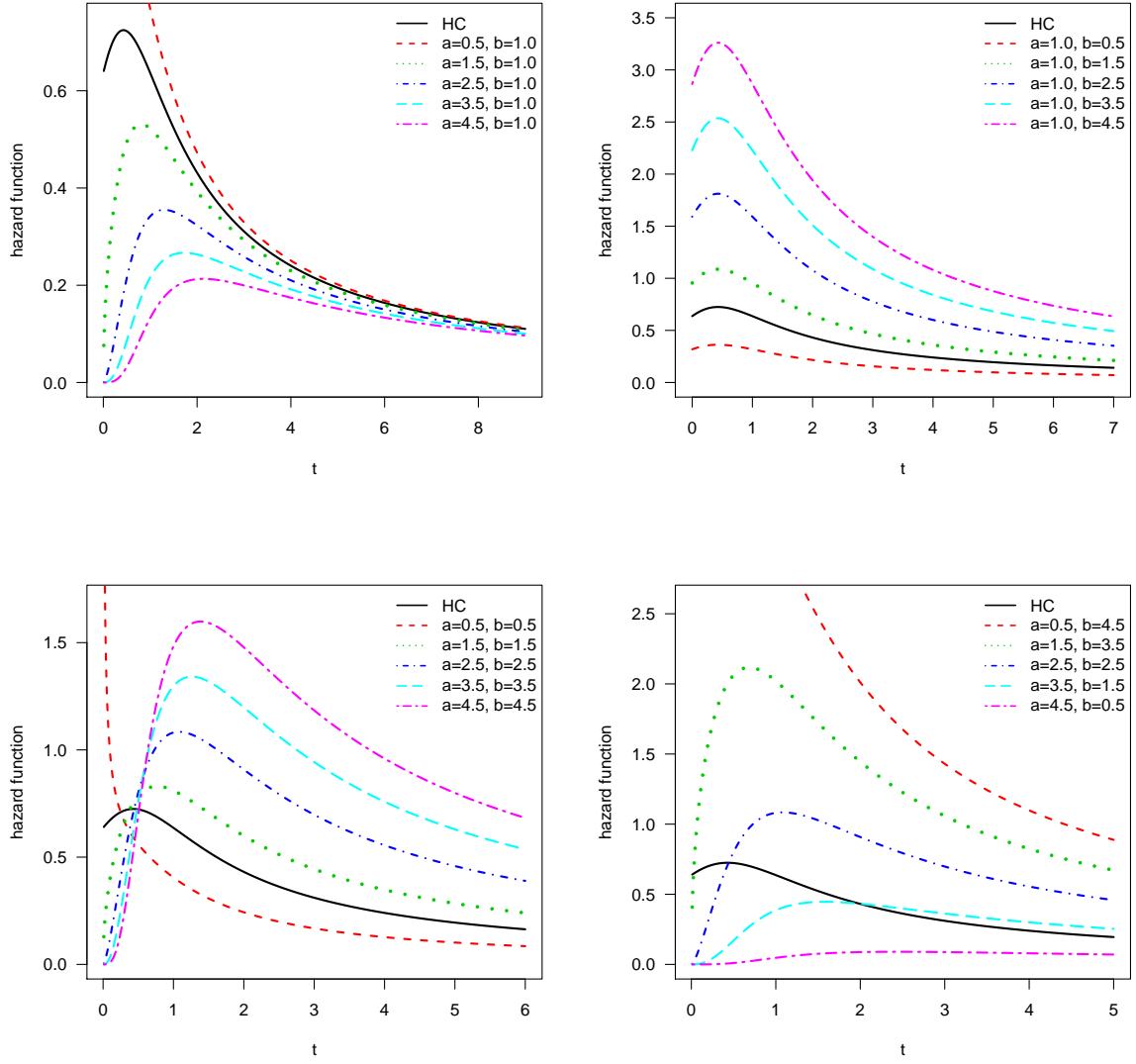


Figure 2: Plots of the hazard rate function (7) for some parameter values; $\phi = 1$.

where $t_r = \sum_{m=0}^{\infty} w_m s_r(a+m)$. By simple differentiation, it is immediate from (8) and (10) that

$$f(t) = \frac{g(t)}{B(a,b)} \sum_{r=0}^{\infty} (a+r) w_r G(t)^{a+r-1} \quad (11)$$

and

$$f(t) = \frac{g(t)}{B(a,b)} \sum_{r=0}^{\infty} (r+1) t_{r+1} G(t)^r, \quad (12)$$

which hold if a is an integer and a is a real non-integer, respectively. Using the expansion

$$\arctan(x) = \sum_{i=0}^{\infty} a_i \frac{x^{2i+1}}{(1+x^2)^{i+1}},$$

where $a_i = 2^{2i} (i!)^2 / [(2i+1)!]$, $G_{\phi}(t)$ can be expanded as

$$G_{\phi}(t) = \left(\frac{t}{\phi^2 + t^2} \right) \sum_{i=0}^{\infty} b_i \left(\frac{t^2}{\phi^2 + t^2} \right)^i,$$

where $b_i = (2\phi a_i)/\pi$.

By application of an equation from Gradshteyn and Ryzhik (2007) for a power series raised to a positive integer j , we obtain

$$G_{\phi}(t)^j = \left(\frac{t}{\phi^2 + t^2} \right)^j \sum_{i=0}^{\infty} c_{j,i} \left(\frac{t^2}{\phi^2 + t^2} \right)^i \quad (13)$$

where the coefficients $c_{j,i}$ (for $i = 1, 2, \dots$) can be determined from the recursive equation ($c_{j,0} = b_0^j$)

$$c_{j,i} = (i b_0)^{-1} \sum_{m=1}^i [(j+1)m - i] b_m c_{j,i-m}. \quad (14)$$

The coefficient $c_{j,i}$ follows recursively from $c_{j,0}, \dots, c_{j,i-1}$ and then from b_0, \dots, b_i . Here, $c_{j,i}$ can be written explicitly in terms of the quantities b_m , although it is not necessary for programming numerically our expansions in any algebraic or numerical software. Now, we can rewrite (11) and (12) as

$$f(t) = \sum_{i,r=0}^{\infty} A_{i,r} \frac{t^{a+r+2i-1}}{(\phi^2 + t^2)^{a+r+i}} \quad \text{and} \quad f(t) = \sum_{i,r=0}^{\infty} B_{i,r} \frac{t^{r+2i}}{(\phi^2 + t^2)^{r+i+1}}, \quad (15)$$

where

$$A_{i,r} = \frac{2\phi(a+r) w_r c_{a+r-1,i}}{\pi B(a,b)} \quad \text{and} \quad B_{i,r} = \frac{2\phi(r+1) t_{r+1} c_{r,i}}{\pi B(a,b)}.$$

Equations (15) are the main results of this section.

3.2 Moments

Here and henceforth, let $T \sim \text{BHC}(\phi, a, b)$. Then, for a an integer and a a real non-integer, the moments of T can be expressed from (15) as

$$E(T^s) = \sum_{i,r=0}^{\infty} A_{i,r} \int_0^{\infty} \frac{t^{s+a+r+2i-1}}{(\phi^2 + t^2)^{a+r+i}} dt \quad \text{and} \quad E(T^s) = \sum_{i,r=0}^{\infty} B_{i,r} \int_0^{\infty} \frac{t^{s+r+2i}}{(\phi^2 + t^2)^{r+i+1}} dt,$$

respectively. For $0 < \alpha < 2\rho$, these integrals can be calculated from Prudnikov et al. (1986, Volume 1, page 309, integral 7) as

$$\int_0^{\infty} \frac{x^{\alpha-1}}{(c^2 + x^2)^{\rho}} dx = c^{\alpha-2\rho} B(\alpha, 2\rho - \alpha) {}_2F_1\left(\frac{\alpha}{2}, \rho - \frac{\alpha}{2}; \rho + \frac{1}{2}; 1\right),$$

where

$${}_2F_1(p, q; c; z) = \sum_{i=0}^{\infty} \frac{(p)_i (q)_i z^i}{(c)_i i!}$$

is the hypergeometric function and $(p)_i = p(p+1) \cdots (p+i-1)$ is the ascending factorial (with the convention that $(p)_0 = 1$). The function ${}_2F_1\left(\frac{\alpha}{2}, \rho - \frac{\alpha}{2}; \rho + \frac{1}{2}; 1\right)$ is absolutely convergent since $c - p - q = 1/2 > 0$.

Hence, for a a positive integer and $s < a$, we can express the moments of T as

$$E(T^s) = \sum_{i,r=0}^{\infty} P_{i,r}(s) {}_2F_1\left(\frac{s+a+r+2i}{2}, \frac{r+a-s}{2} + 1; a+r+i + \frac{1}{2}; 1\right), \quad (16)$$

where $P_{i,r}(s) = \phi^{s-r-a} B(s+a+r+2i, r+a-s) A_{i,r}$. The moments of the HC distribution for $s < 1$ can be computed from (16) with $a = b = 1$.

On the other hand, for a a positive real non-integer and $s < 1$, we can obtain

$$E(T^s) = \sum_{i,r=0}^{\infty} Q_{i,r}(s) {}_2F_1\left(\frac{s+1+r+2i}{2}, \frac{r+1-s}{2} + 1; r+i + \frac{3}{2}; 1\right), \quad (17)$$

where $Q_{i,r}(s) = \phi^{s-r-1} B(s+r+1+2i, r+1-s) B_{i,r}$. The moments functions (16) and (17) show that the method of moments will not work for this distribution.

3.3 Generating Function

The mgf $M(-v) = E\{\exp(-vT)\}$ of T can be derived from the following result due to Prudnikov et al. (1986, Volume 1, page 309, integral 7)

$$K_{m,n}(v; \phi) = \int_0^{\infty} \frac{x^m \exp(-vx)}{(\phi^2 + x^2)^n} dx = \frac{(-1)^{m+n-1}}{2^{n-1} (n-1)!} \frac{\partial^m}{\partial v^m} \left(\frac{\partial}{v \partial v} \right)^{n-1} H(v; \phi), \quad (18)$$

which holds for any v , where

$$H(v; \phi) = \phi^{-1} [\sin(\phi v) \text{ci}(\phi v) - \cos(\phi v) \text{si}(\phi v)],$$

and $\text{ci}(\phi v) = -\int_{\phi v}^{\infty} t^{-1} \cos(t) dt$ and $\text{si}(\phi v) = -\int_{\phi v}^{\infty} t^{-1} \sin(t) dt$ are the cosine integral and sine integral, respectively.

For a an integer and a a real non-integer, the BHC generating function can be determined, from equations (15) and (18), as linear combinations of $K_{\cdot,\cdot}(v; \phi)$ functions

$$M(-v) = \sum_{i,r=0}^{\infty} A_{i,r} K_{a+r+2i-1, a+r+i}(v; \phi) \quad (19)$$

and

$$M(-v) = \sum_{i,r=0}^{\infty} B_{i,r} K_{r+2i, r+i+1}(v; \phi), \quad (20)$$

respectively. Equations (19) and (20) are the main results of this section.

3.4 Quantile Expansion

The BHC quantile function $t = Q(u)$ is straightforward to be computed from the beta quantile function $Q_B(u)$, which is available in most statistical packages, by

$$t = Q(u) = \phi \tan \left(\frac{\pi Q_B(u)}{2} \right).$$

Power series methods are at the heart of many aspects of applied mathematics and statistics. Here, we provide a power series expansion for $Q(u)$ that can be useful to derive some mathematical measures of the new distribution. Further, we propose alternative expressions for the BHC moments based upon this expansion.

First, an expansion for the beta quantile function, say $Q_B(u)$, can be found in Wolfram website¹ as $Q_B(u) = \sum_{i=0}^{\infty} g'_i u^{i/a}$, where $g'_0 = 0$ and $g'_i = q_i [a B(a, b)]^{i/a}$ (for $i \geq 1$) and the quantities q_i 's (for $i \geq 2$) can be derived from the cubic recursive equation

$$q_i = \frac{1}{[i^2 + (a-2)i + (1-a)]} \left\{ (1 - \delta_{i,2}) \sum_{r=2}^{i-1} q_r q_{i+1-r} [r(1-a)(i-r) - r(r-1)] \right. \\ \left. + \sum_{r=1}^{i-1} \sum_{s=1}^{i-r} q_r q_s q_{i+1-r-s} [r(r-a) + s(a+b-2)(i+1-r-s)] \right\},$$

where $\delta_{i,2} = 1$ if $i = 2$ and $\delta_{i,2} = 0$ if $i \neq 2$. For example, $q_0 = 0$, $q_1 = 1$, $q_2 = (b-1)/(a+1)$, $q_3 = [(b-1)(a^2 + 3ab - a + 5b - 4)]/[2(a+1)^2(a+2)]$, and so on. We can expand $Q(u)$ (since $E_0 = 0$) as

$$Q(u) = \phi \sum_{k=1}^{\infty} E_k Q_B(u)^k,$$

where $E_{2k} = 0$, $E_{2k-1} = (2^{2k} - 1) \pi^{2k-1} [2(2k)!]^{-1} B_{2k}$ (for $k = 1, 2, \dots$) and B_{2k} are the Bernoulli numbers. We have $B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42$, $B_8 = -1/30, \dots$. The beta quantile

¹<http://functions.wolfram.com/06.23.06.0004.01>.

function can be rewritten as $Q_B(u) = u^{1/a} \sum_{i=0}^{\infty} g_i u^{i/a}$ because $g'_0 = 0$, where $g_i = g'_{i+1} = q_{i+1} [aB(a, b)]^{(i+1)/a}$ for $i = 0, 1, \dots$. So, $g_0 = [aB(a, b)]^{1/a}$, $g_1 = [(b-1)/(a+1)] [aB(a, b)]^{2/a}$, and so on. Now, we obtain

$$Q(u) = \phi \sum_{k=1}^{\infty} E_k \left(u^{1/a} \sum_{i=0}^{\infty} g_i u^{i/a} \right)^k$$

and then

$$Q(u) = \phi \sum_{k=1, i=0}^{\infty} E_k h_{k,i} u^{(k+i)/a},$$

where the constants $h_{k,i}$ can be evaluated recursively using (14) from the quantities g_i by $h_{k,0} = g_0^k$ and $h_{k,i} = (i g_0)^{-1} \sum_{m=1}^i [(k+1)m - i] g_m h_{k,i-m}$, for $i = 1, 2, \dots$. Further,

$$Q(u) = \phi \sum_{p=1}^{\infty} N_p u^{p/a}, \quad (21)$$

where $N_p = \sum_{k=1}^p E_k h_{k,p-k}$ for $p = 1, 2, \dots$. The power series (21) for the BHC quantile can be used to obtain some mathematical properties of this distribution. For example, the s th moment of T (for a a real non-integer) can be expressed as

$$E(T^s) = \int_0^{\infty} x^s f(x) dx = \int_0^1 Q(u)^s du.$$

This integral in $(0, 1)$ yields an alternative formula for (17) as

$$E(T^s) = \phi^s \int_0^1 \left(\sum_{p=0}^{\infty} M_p u^{(p+1)/a} \right)^s du = \phi^s \sum_{p=0}^{\infty} L_{s,p} \int_0^1 u^{(p+s)/a} du = a \phi^s \sum_{p=0}^{\infty} \frac{L_{s,p}}{(p+s+a)},$$

where $M_p = N_{p+1} = \sum_{k=1}^{p+1} E_k h_{k,p+1-k}$ and $L_{s,p}$ can be computed from (14) by ($L_{s,0} = M_0^s$)

$$L_{s,p} = (p M_0)^{-1} \sum_{m=1}^p [(s+1)m - p] M_m L_{s,p-m}.$$

3.5 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. We can derive the BHC mean deviations about the mean $\mu = E(T)$ and about the median M ($M = Q(1/2)$) from the relations

$$\delta_1 = 2\mu F(\mu) - 2H(\mu) \quad \text{and} \quad \delta_2 = E(T) - 2H(M),$$

respectively, where μ can be computed from (16) with $s = 1$ for $a > 1$, $F(\mu)$ and $F(M)$ are calculated from (6) and $H(s) = \int_0^s t f(t) dt$. After some algebra from (21), $H(s)$ takes the form

$$H(s) = \phi \int_0^{F(s)} \left(\sum_{p=1}^{\infty} N_p u^{p/a} \right) du = a \phi \sum_{p=1}^{\infty} \frac{N_p F(s)^{p/a+1}}{(a+p)}. \quad (22)$$

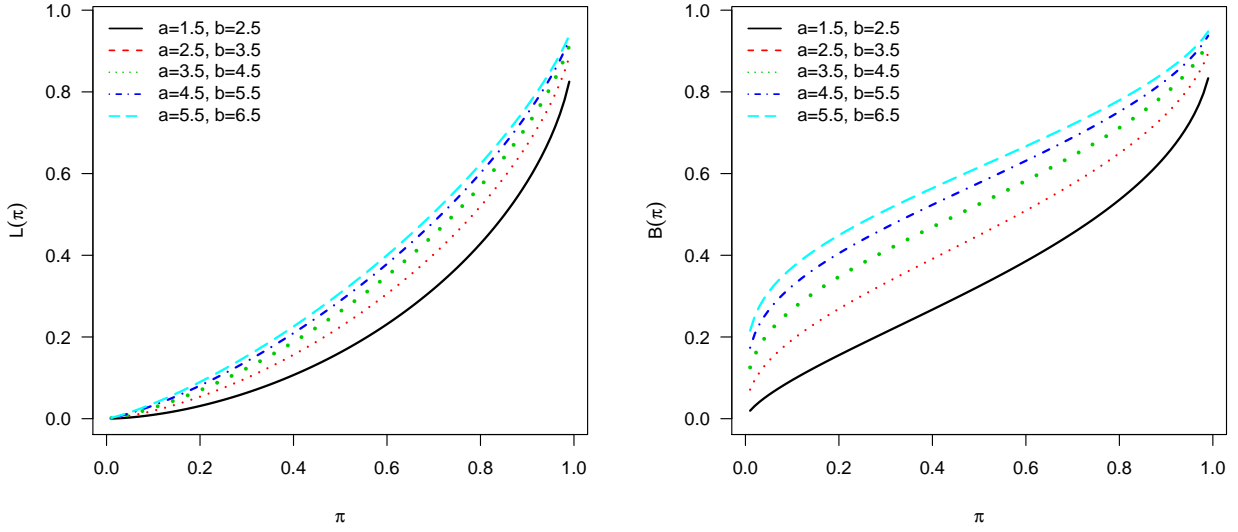


Figure 3: Plots of $L(\pi)$ and $B(\pi)$ with $\phi = 1$ and $\mu = 1$.

An application of the mean deviations is to the Lorenz and Bonferroni curves that are important in fields like economics, reliability, demography, insurance and medicine. They are defined for a given probability π by $L(\pi) = H(q)/\mu$ and $B(\pi) = H(q)/(\pi\mu)$, respectively, where $q = Q(\pi)$ comes from (21). 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 T 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. The curves $L(\pi)$ and $B(\pi)$ for the BHC distribution as functions of π are readily calculated from (22). They are plotted for selected parameter values in Figure 3.

3.6 Order statistics and Moments

Order statistics make their appearance in many areas of statistical theory and practice. The density function $f_{i:n}(t)$ of the i th order statistic, say $T_{i:n}$, for $i = 1, 2, \dots, n$, from data values T_1, \dots, T_n having the beta-G distribution can be obtained from (4) as

$$f_{i:n}(t) = \frac{g(t) G(t)^{a-1} \{1 - G(t)\}^{b-1}}{B(a, b) B(i, n - i + 1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(t)^{i+j-1}. \quad (23)$$

From equations (10), (13) and (14), we can write

$$F(t)^{i+j-1} = \frac{1}{B(a, b)^{i+j-1}} \sum_{r=0}^{\infty} d_{i+j-1, r} G(t)^r,$$

where $d_{i+j-1, r} = (r t_0)^{-1} \sum_{\ell=1}^r [(i+j)\ell - r] t_{\ell} d_{i+j-1, r-\ell}$ and $d_{i+j-1, 0} = t_0^{i+j-1}$.

Inserting this equation in (23), $f_{i:n}(t)$ can be further reduced to

$$f_{i:n}(t) = g(t) \sum_{k=0}^{\infty} M_{i:n}(k) G(t)^k,$$

where

$$M_{i:n}(k) = \sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j}}{B(a, b)^{i+j} B(i, n-i+1)} \sum_{r, m=0}^{\infty} (-1)^m \binom{b-1}{m} d_{i+j-1, r} s_k(a+r+m-1).$$

If b is an integer, the index m in the above quantity stops at $b-1$.

Using (13), we obtain

$$f_{i:n}(t) = g_{\phi}(t) \sum_{k, p=0}^{\infty} c_{k, p} M_{i:n}(k) \frac{t^{2p+k}}{(\phi^2 + t^2)^{p+k}}, \quad (24)$$

where $c_{k, p}$ is given by (14). By equation (24), we can derive some mathematical properties of $T_{i:n}$. For example, the s th moment of $T_{i:n}$ follows immediately as

$$\begin{aligned} E(T_{i:n}^s) &= \frac{2}{\pi} \sum_{k, p=0}^{\infty} \phi^{s-k+2} B(2p+k+s+1, k-s-1) c_{k, p} M_{i:n}(k) \\ &\quad \times {}_2F_1\left(\frac{2p+k+s+1}{2}, \frac{k-s-1}{2}; p+k+\frac{1}{2}; 1\right). \end{aligned}$$

L-moments are summary statistics for probability distributions and data samples (Hosking, 1990). They 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. The L-moments can be expressed as linear combinations of the ordered data values

$$\lambda_r = \sum_{j=0}^{r-1} (-1)^{r-1-j} \binom{r-1}{j} \binom{r-1+j}{j} \eta_j,$$

where $\eta_j = E\{TF(T)^j\} = (j+1)^{-1}E(T_{j+1:j+1})$. In particular, $\lambda_1 = \eta_0$, $\lambda_2 = 2\eta_1 - \eta_0$, $\lambda_3 = 6\eta_2 - 6\eta_1 + \eta_0$ and $\lambda_4 = 20\eta_3 - 30\eta_2 + 12\eta_1 - \eta_0$. The L-moments of the BHC distribution can be obtained from the results of this section.

3.7 Entropy

The entropy of a random variable T with density function $f(t)$ is a measure of variation of the uncertainty. Rényi entropy is defined by $I_R(\rho) = (1 - \rho)^{-1} \log\{\int f(t)^\rho dt\}$, where $\rho > 0$ and $\rho \neq 1$. If a random variable T has a BHC distribution, we have

$$f(t)^\rho = L(\rho) \left[1 + \left(\frac{t}{\phi} \right)^2 \right]^{-\rho} G_\phi(t)^{(a-1)\rho} \{1 - G_\phi(t)\}^{(b-1)\rho}, \quad (25)$$

where $L(\rho) = 2^\rho [\pi \phi B(a, b)]^{-\rho}$. By expanding the binomial term, we obtain

$$f(t)^\rho = L(\rho) \left[1 + \left(\frac{t}{\phi} \right)^2 \right]^{-\rho} \sum_{j=0}^{\infty} R_j G_\phi(t)^{(a-1)\rho+j},$$

where $R_j = (-1)^j \binom{(b-1)\rho}{j}$. By equation (9), we can write

$$f(t)^\rho = L(\rho) \left[1 + \left(\frac{t}{\phi} \right)^2 \right]^{-\rho} \sum_{r=0}^{\infty} N_r(\rho) \left[\arctan \left(\frac{t}{\phi} \right) \right]^r,$$

where

$$N_r(\rho) = \sum_{j=0}^{\infty} M_j s_r((a-1)\rho + j) \left(\frac{2}{\pi} \right)^r,$$

and $s_r((a-1)\rho + j)$ is defined after equation (9). We obtain

$$\left[\arctan \left(\frac{t}{\phi} \right) \right]^r = \phi^r \sum_{k=0}^{\infty} f_{r,k} \frac{t^{2k+r}}{(\phi^2 + t^2)^{k+r}},$$

where $f_{r,0} = a_0^r$, $f_{r,k} = (i a_0)^{-1} \sum_{m=1}^k [(r+1)m - k] a_m f_{r,k-m}$, and $a_k = 2^{2k} (k!)^2 / [(2k+1)!]$. Thus,

$$\int_0^\infty f(t)^\rho dt = L(\rho) \sum_{r,k=0}^{\infty} \phi^{2\rho+r} N_r(\rho) f_{r,k} \int_0^\infty \frac{t^{2k+r}}{(\phi^2 + t^2)^{k+r+\rho}} dt.$$

Finally, the Rényi entropy can be determined from

$$\int_0^\infty \frac{t^{2k+r}}{(\phi^2 + t^2)^{k+r+\rho}} dt = \frac{B(2k+r+1, r+2\rho-1)}{\phi^{r+2\rho-1}} {}_2F_1 \left(\frac{2k+r+1}{2}, \rho + \frac{r-1}{2}; k+r+\rho + \frac{1}{2}; 1 \right).$$

4 Estimation and Inference

The estimation of the model parameters is investigated by the method of maximum likelihood. Let $\mathbf{t} = (t_1, \dots, t_n)^\top$ be a random sample of size n from the BHC distribution with unknown parameter vector $\boldsymbol{\theta} = (\phi, a, b)^\top$. The total log-likelihood function for $\boldsymbol{\theta}$ can be written as

$$\begin{aligned} \ell(\boldsymbol{\theta}) = & n a \log(2/\pi) - n \log(\phi) - n \log\{B(a, b)\} - \sum_{i=1}^n \log(\dot{w}_i) \\ & + (a-1) \sum_{i=1}^n \log(\dot{z}_i) + (b-1) \sum_{i=1}^n \log(\dot{d}_i), \end{aligned}$$

where $\dot{v}_i = \dot{v}_i(\phi) = t_i/\phi$, $\dot{w}_i = \dot{w}_i(\phi) = 1 + \dot{v}_i^2$, $\dot{z}_i = \dot{z}_i(\phi) = \arctan(\dot{v}_i)$ and $\dot{d}_i = \dot{d}_i(\phi) = 1 - 2\dot{z}_i/\pi$, for $i = 1, \dots, n$. The maximization of the log-likelihood over three parameters looks easy in practice. The components of the score vector $\mathbf{U}_\theta = (U_\phi, U_a, U_b)^\top$ are

$$U_\phi = -\frac{n}{\phi} + \frac{2}{\phi^3} \sum_{i=1}^n \frac{t_i^2}{\dot{w}_i} - \frac{(a-1)}{\phi^2} \sum_{i=1}^n \frac{t_i}{\dot{w}_i \dot{z}_i} + \frac{2(b-1)}{\pi \phi^2} \sum_{i=1}^n \frac{t_i}{\dot{w}_i \dot{d}_i},$$

$$U_a = n \log(2/\pi) + n\{\psi(a+b) - \psi(a)\} + \sum_{i=1}^n \log(\dot{z}_i),$$

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

where $\psi(\cdot)$ is the digamma function. The maximum likelihood estimates (MLEs) $\hat{\boldsymbol{\theta}} = (\hat{\phi}, \hat{a}, \hat{b})^\top$ of $\boldsymbol{\theta} = (\phi, a, b)^\top$ are the simultaneous solutions of the equations $U_\phi = U_a = U_b = 0$. They can be solved numerically using iterative methods such as a Newton–Raphson type algorithm.

The normal approximation of the estimate $\hat{\boldsymbol{\theta}}$ can be used for constructing approximate confidence intervals and for testing hypotheses on the parameters ϕ , a and b . Under standard regularity conditions, we have $\sqrt{n}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \overset{A}{\sim} \mathcal{N}_3(\mathbf{0}, \mathbf{K}_\theta^{-1})$, where $\overset{A}{\sim}$ means approximately distributed and \mathbf{K}_θ is the unit expected information matrix. The asymptotic result $\mathbf{K}_\theta = \lim_{n \rightarrow \infty} n^{-1} \mathbf{J}_n(\boldsymbol{\theta})$ holds, where $\mathbf{J}_n(\boldsymbol{\theta})$ is the observed information matrix. The average matrix evaluated at $\hat{\boldsymbol{\theta}}$, say $n^{-1} \mathbf{J}_n(\hat{\boldsymbol{\theta}})$, can estimate \mathbf{K}_θ . The elements of the observed information matrix $\mathbf{J}_n(\boldsymbol{\theta}) = -\partial^2 \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top = -\{U_{ij}\}$, for $i, j = \phi, a$ and b , are

$$U_{\phi\phi} = \frac{n}{\phi^2} - \frac{6}{\phi^4} \sum_{i=1}^n \frac{t_i^2}{\dot{w}_i} + \frac{4}{\phi^6} \sum_{i=1}^n \frac{t_i^4}{\dot{w}_i^2} + \frac{2(a-1)}{\phi^3} \sum_{i=1}^n \frac{t_i}{\dot{w}_i \dot{z}_i} \left[1 - \frac{t_i^2}{\phi^2 \dot{w}_i} - \frac{t_i}{2\phi \dot{w}_i \dot{z}_i} \right]$$

$$- \frac{4(b-1)}{\pi \phi^3} \sum_{i=1}^n \frac{t_i}{\dot{w}_i \dot{d}_i} \left[1 - \frac{t_i^2}{\phi^2 \dot{w}_i} + \frac{t_i}{\pi \phi \dot{w}_i \dot{d}_i} \right],$$

$$U_{\phi a} = -\frac{1}{\phi^2} \sum_{i=1}^n \frac{t_i}{\dot{w}_i \dot{z}_i}, \quad U_{\phi b} = \frac{2}{\pi \phi^2} \sum_{i=1}^n \frac{t_i}{\dot{w}_i \dot{d}_i},$$

$$U_{aa} = n\{\psi'(a+b) - \psi'(a)\}, \quad U_{ab} = n\psi'(a+b), \quad U_{bb} = n\{\psi'(a+b) - \psi'(b)\},$$

where $\psi'(\cdot)$ is the trigamma function. Thus, the multivariate normal $\mathcal{N}_3(\mathbf{0}, \mathbf{J}_n(\hat{\boldsymbol{\theta}})^{-1})$ distribution can be used to construct approximate confidence intervals $\hat{\phi} \pm z_{\eta/2} \times [\widehat{\text{var}}(\hat{\phi})]^{1/2}$, $\hat{a} \pm z_{\eta/2} \times [\widehat{\text{var}}(\hat{a})]^{1/2}$ and $\hat{b} \pm z_{\eta/2} \times [\widehat{\text{var}}(\hat{b})]^{1/2}$ for the parameters ϕ , a and b , respectively, where $\text{var}(\cdot)$ is the diagonal element of $\mathbf{J}_n(\hat{\boldsymbol{\theta}})^{-1}$ corresponding to each parameter and $z_{\eta/2}$ is the quantile $100(1 - \eta/2)\%$ of the standard normal distribution.

We can easily check if the fit using the BHC model is statistically “superior” to a fit using the HC model for a given data set by computing the likelihood ratio (LR) statistic $w = 2\{\ell(\hat{\phi}, \hat{a}, \hat{b}) - \ell(\tilde{\phi}, 1, 1)\}$, where $\hat{\phi}$, \hat{a} and \hat{b} are the unrestricted MLEs and $\tilde{\phi}$ is the restricted estimate. The statistic w is asymptotically distributed, under the null model, as χ_2^2 . Further, the LR test rejects the null hypothesis if $w > \xi_\eta$, where ξ_η denotes the upper $100\eta\%$ point of the χ_2^2 distribution.

5 Application

Here, we present an application of the BHC distribution to a real data set. We shall compare the fits of the BHC, EHC and HC distributions. We also consider for the sake of comparison the two-parameter Birnbaum–Saunders (BS), gamma and Weibull models, and the three-parameter BS and Weibull models. The BHC distribution may be an interesting alternative to these distributions for modeling positive real data sets. The cdf's of the exponentiated BS (ExpBS), exponentiated Weibull (ExpWeibull) and gamma models are (for $t > 0$)

$$F(t) = \Phi\left(\frac{1}{\alpha}\left[\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right]\right)^\gamma, \quad F(t) = (1 - e^{-\beta t^\alpha})^\gamma, \quad F(t) = \frac{\zeta(\alpha, \beta t)}{\Gamma(\alpha)},$$

respectively, where $\alpha > 0$, $\beta > 0$, $\gamma > 0$. Here, $\Phi(\cdot)$ is the cdf of the standard normal distribution and $\zeta(\cdot, \cdot)$ is the ordinary incomplete gamma function. If $\gamma = 1$ we have the two-parameter BS and Weibull models. All the computations were done using the Ox matrix programming language (Doornik, 2006) which is freely distributed for academic purposes at <http://www.doornik.com>. The maximization was performed by the BFGS method with analytical derivatives. For further details about this method the reader is referred to Nocedal and Wright (1999, §8.1) and Press et al. (2007, §10.7). We shall consider the data set originally due to Bjerkedal (1960), which has also been analyzed by Gupta et al. (1997). The data represent the survival times of guinea pigs injected with different doses of *tubercle bacilli*.

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 BHC distribution has the lowest AIC, BIC and HQIC values in relation to their sub-models, and so it could be chosen as the best model. The LR statistics for testing the hypotheses \mathcal{H}_0 : EHC against \mathcal{H}_1 : BHC and \mathcal{H}_0 : HC against \mathcal{H}_1 : BHC are 22.9462 and 40.7366, respectively, and all yield p -values < 0.001 . Thus, we can reject the null hypotheses in all cases in favor of the BHC distribution at any usual significance level, i.e. the BHC model is significantly better than the EHC and HC distributions. In order to assess if the model is appropriate, plots of the estimated density functions are given in Figure 4. They also indicate that the BHC model provides a better fit than the other models.

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 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 $H(x; \theta)$ be the cdf, where the form of H is known but θ (a k -dimensional parameter vector, say) is unknown. To obtain the statistics W^* and A^* , we can proceed as follows: (i) Compute $v_i = H(x_i; \hat{\theta})$, 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 $\bar{y} = n^{-1} \sum_{i=1}^n y_i$ 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 $A^2 = -n - (1/n) \sum_{i=1}^n \{(2i-1) \log(u_i) + (2n+1-2i) \log(1-u_i)\}$ and

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

Distribution	Estimates			Statistic		
	ϕ	a	b	AIC	BIC	HQIC
BHC	56.6890 (23.1921)	3.7238 (1.1825)	2.7033 (0.6056)	785.58	792.41	788.30
EHC	20.9790 (11.6134)	4.1938 (2.3670)		806.53	811.08	808.34
HC	75.8253 (10.3629)			822.32	824.60	823.23

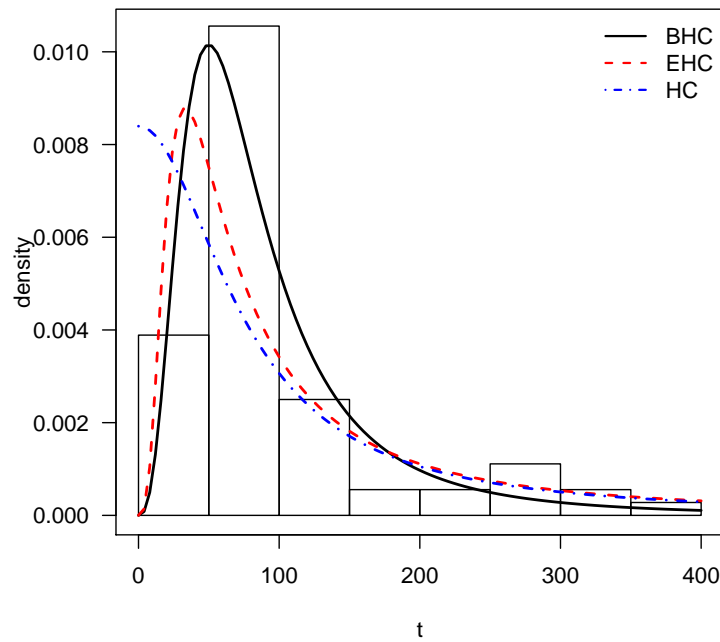


Figure 4: Estimated densities of the BHC, EHC and HC models.

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 the models are listed in Table 2, thus indicating that the BHC model should be chosen to fit the current data.

Table 2: Goodness-of-fit tests.

Distribution	Statistic	
	W^*	A^*
BHC	0.10682	0.60255
EHC	0.13318	0.79202
HC	0.13099	0.72207

The MLEs (standard errors in parentheses) of the model parameters of the ExpBS, ExpWeibull, BS, gamma and Weibull models and the statistics W^* and A^* are listed in Table 3. Based on these statistics, the ExpWeibull model yields a better fit than the ones of the other distributions. Overall, by comparing the figures in Tables 2 and 3, we conclude that the BHC model outperforms all the models considered in Table 3. So, the proposed distribution can yield a better fit than the classical three- and two-parameter BS, gamma and Weibull models and therefore may be an interesting alternative to these distributions for modeling positive real data sets. These results illustrate the potentiality of the new distribution and the necessity of additional shape parameters.

Table 3: MLEs (standard errors in parentheses) and the measures W^* and A^* .

Distribution	Estimates			Statistic	
	α	β	γ	W^*	A^*
ExpBS	0.5845 (0.9407)	131.3672 (377.8605)	0.3984 (2.2449)	0.18182	0.98014
ExpWeibull	0.4611 (0.1709)	0.4744 (0.5344)	22.4424 (30.1960)	0.14017	0.76577
BS	0.7600 (0.0633)	77.5348 (6.4508)		0.18824	1.01205
Gamma	2.0815 (0.3305)	0.0209 (0.0037)		0.33952	1.85891
Weibull	1.3932 (0.1184)	0.0014 (0.0009)		0.43476	2.39383

6 Concluding remarks

We introduce a new lifetime model, called the beta half-Cauchy (BHC) distribution, that extends the Half-Cauchy (HC) distribution, and study some of its general structural properties. We provide a mathematical treatment of the new distribution including expansions for the density function, moments, generating function, order statistics, quantile function, Rényi entropy, mean deviations and Lorentz and Bonferroni curves. The model parameters are estimated by maximum likelihood. Our formulas related with the BHC model are manageable and with the use of modern computer resources with analytic and numerical capabilities, may turn into adequate tools comprising the arsenal of applied statisticians. The usefulness of the proposed model is illustrated in an application to real data using likelihood ratio statistics and formal goodness-of-fit tests. The new model provides consistently better fit than other models available in the literature. We hope that the proposed model may attract wider applications in survival analysis for modeling positive real data sets.

Acknowledgments

We gratefully acknowledge grants from CNPq and FAPESP (Brazil). The authors thank an anonymous referee for some comments which improved the original version of the paper.

References

- Akinsete, A., Famoye, F., Lee, C. (2008). The beta-Pareto distribution. *Statistics* **42**, 547–563.
- Barreto-Souza, W., Santos, A., Cordeiro, G.M. (2010). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation* **80**, 159–172.
- Bjerkedal, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *American Journal of Hygiene* **72**, 130–148.
- Chen, G., Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology* **27**, 154–161.
- Cordeiro, G.M., Lemonte, A.J. (2011a). The β -Birnbaum–Saunders distribution: An improved distribution for fatigue life modeling. *Computational Statistics and Data Analysis* **55**, 1445–1461.
- Cordeiro, G.M., Lemonte, A.J. (2011b). The beta Laplace distribution. *Statistics and Probability Letters* **81**, 973–982.
- 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., Kannan, N., RayChoudhuri, A. (1997). Analysis of log-normal survival data. *Mathematical Biosciences* **139**, 103–115.
- 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.
- Johnson, N., Kotz, S., Balakrishnan, N. (1994). *Continuous Univariate Distributions – Volume 1*, 2nd ed. Wiley: New York.
- Johnson, N., Kotz, S., Balakrishnan, N. (1995). *Continuous Univariate Distributions – Volume 2*, 2nd ed. Wiley: New York.
- Lee, C., Famoye, F., Olumolade, O. (2007). Beta-Weibull distribution: Some properties and applications to censored data. *Journal of Modern Applied Statistical Methods* **6**, 173–186.
- 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. *Mathematical Problems in Engineering* **10**, 323–332.
- Nadarajah, S., Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering and System Safety* **91**, 689–697.
- Nocedal, J., Wright, S.J. (1999). *Numerical Optimization*. Springer: New York.
- Paradis, E., Baillie, S.R, Sutherland, W.J. (2002). Modeling large-scale dispersal distances. *Ecological Modelling* **151**, 279–292.
- Paranaíba, P.F., Ortega, E.M.M., Cordeiro, G.M., Pescim, R.R. (2011). The beta Burr XII distribution with application to lifetime data. *Computational Statistics and Data Analysis* **55**, 1118–1136.
- Pescim, R.R., Demétrio, C.G.B, Cordeiro, G.M., Ortega, E.M.M., Urbano, M.R. (2010). The beta generalized half-normal distribution. *Computational Statistics and Data Analysis* **54**, 945–957.
- 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 (Volumes 1, 2, 3 and 4)*. Amsterdam: Gordon and Breach Science Publishers.
- Razzaghi, M. (2009). Beta-normal distribution in dose-response modeling and risk assessment for quantitative responses. *Environmental and Ecological Science* **16**, 21–36.

- Shaw, M.W. (1995). Simulation of population expansion and spatial pattern when individual dispersal distributions do not decline exponentially with distance. *Proceedings of the Royal Society B: Biological Sciences* **259**, 243–248.
- Silva, G.O., Ortega, E.M.M., Cordeiro, G.M. (2010). The beta modified Weibull distribution. *Lifetime Data Analysis* **16**, 409–430.