

Submitted to South African Journal of Statistics  
Mar 2007

## **The Beta-Hyperbolic Secant (BHS) Distribution**

**Matthias J. Fischer<sup>1</sup>, David Vaughan<sup>2</sup>**

<sup>1</sup> Department of Statistics, University of Erlangen-Nürnberg, Nürnberg, Germany

<sup>2</sup> Department of Mathematics, Wilfrid Laurier University, Waterloo, Ontario, Canada

Received: date / Revised version: date

**Abstract** The shape of a probability distribution is often summarized by the distribution's skewness and kurtosis. Starting from a symmetric "parent" density  $f$  on the real line, we can modify its shape (i.e. introduce skewness and in-/decrease kurtosis) if  $f$  is appropriately weighted. In particular, every density  $w$  on the interval  $(0, 1)$  is a specific weighting function. Within this work, we follow up a proposal of Jones (2004) and choose the Beta distribution as underlying weighting function  $w$ . "Parent" distributions like the Student-t, the logistic and the normal distribution have already been investigated in the literature. Based on the assumption that  $f$  is the density of a hyperbolic secant distribution, we introduce the Beta-hyperbolic secant (BHS) distribution. In contrast to the Beta-normal distribution and the to Beta-Student-t distribution, BHS densities are always unimodal and all moments exist. In contrast to the Beta-logistic distribution, the BHS distribution is more flexible regarding the range of skewness and leptokurtosis combinations. Moreover, we propose a generalization which nests both the Beta-logistic and the BHS distribution. Finally, the goodness-of-fit between all above-mentioned distributions is compared for glass fibre data and aluminium returns.

## 1 Introduction

Several techniques can be applied to symmetric distributions in order to generate asymmetric ones with possibly lighter or heavier tails. In terms of density functions — provided their existence — most of these methods can be represented by

$$g(x; \theta) = f(x)w(F(x); \theta), \quad (1)$$

where  $g$  denotes the transformed density,  $f$  and  $F$  the (symmetric) pdf and cdf, respectively, of the original ("parent") distribution and  $w$  is an appropriate weighting function on the interval  $(0, 1)$  with parameter vector  $\theta$  (see, for instance, Ferreira and Steel, 2004). Choosing  $w(u; \lambda) = 2F(\lambda F^{-1}(u))$ , the skewing mechanism of Azzalini (1985, 1986) is recovered. Similarly, using

$$w(u; \lambda) = \frac{2}{\lambda + \frac{1}{\lambda}} \frac{f(\lambda^{\text{sign}(0.5-u)} F^{-1}(u))}{f(F^{-1}(u))} \quad (2)$$

corresponds to applying different parameters of scale to the positive and the negative part of a symmetric density (see, for example, Fernández, Osiewalski and Steel, 1995 and Theodossiou, 1998).

In particular, every probability density on  $(0, 1)$  which is not uniform can be used either to introduce skewness and/or to modify the kurtosis of

the parent distribution. A very attractive choice (due to its flexibility) is the density of a Beta distribution, i.e.

$$w(x; \beta_1, \beta_2) = \frac{1}{B(\beta_1, \beta_2)} x^{\beta_1-1} (1-x)^{\beta_2-1}, \quad \beta_1, \beta_2 > 0, \quad (3)$$

where  $B(a, b) = \int_0^1 t^{a-1} (1-t)^{b-1} dt$  denotes the Beta function (cf. Jones, 2004). Examples where (3) has been used in the literature are the following:

- Aroian (1941), Prentice (1975): Beta-logistic distribution (which is also termed as exponential generalized beta of the second kind or EGB2 distribution, or log  $F$  distribution),
- Eugene et al. (2002): Beta-normal (BN) distribution,
- Jones and Faddy (2003): Beta-Student-t distribution.

Within this work we introduce the BHS (Beta-hyperbolic secant) distribution as a weighted hyperbolic secant distribution with weights from (3). The hyperbolic secant distribution itself dates back to Perks (1932). It is symmetric, more leptokurtic than the normal, even more than the logistic distribution but still with existing moments. Both the cumulative distribution function and the inverse cumulative distribution function are given in closed form. Despite its interesting properties, the hyperbolic secant distribution has not received sufficient attention in the literature so far.

Whereas both Beta-normal and Beta-Student-t distribution do not guarantee unimodality — except for a special parameterization given in Ferreira and Steel (2004) — the BHS distribution does. In contrast to the Beta-Student-t distribution, all moments of the BHS distribution exist. Although the Beta-logistic and the BHS distribution are very similar, the BHS distribution will be seen to be more flexible regarding skew and leptokurtic data. In order to discriminate between both distribution models, a generalized Beta-GSH model — based on Vaughan's (2002) generalized secant hyperbolic (GSH) distribution — is proposed that includes both candidate distributions as special case.

The paper is structured as follows: The BHS distribution and some fundamental properties are introduced in section 2. Section 3 is devoted to the parameter estimation of the BHS distribution. A generalization of both the Beta-logistic distribution and the BHS distribution is proposed in section 4. In section 5, the BHS distribution is compared with its competitors derived from alternative parent distributions.

## 2 Definition and Properties

### 2.1 Definition of the Beta-Hyperbolic Secant Distribution

The probability density function of a standardized (i.e. zero mean and unit variance) hyperbolic secant distribution is given by

$$f(x) = \frac{1}{\pi \cosh(x)} = \frac{2}{\pi(e^x + e^{-x})}, \quad x \in \mathbb{R}. \quad (4)$$

It is symmetric and the corresponding cumulative distribution function is

$$F(x) = \frac{2 \arctan(e^x)}{\pi}. \quad (5)$$

The inverse cumulative distribution function is  $F^{-1}(u) = \log(\tan(\frac{\pi u}{2}))$ . Combining (1), (3), (4) and (5), the density of the Beta-hyperbolic secant (BHS) distribution is defined by

$$g(x; \beta_1, \beta_2) = \frac{B(\beta_1, \beta_2)^{-1}}{\pi \cosh(x)} \frac{\left[\frac{2}{\pi} \arctan(\exp(x))\right]^{\beta_1-1}}{\left[1 - \frac{2}{\pi} \arctan(\exp(x))\right]^{1-\beta_2}}, \quad (6)$$

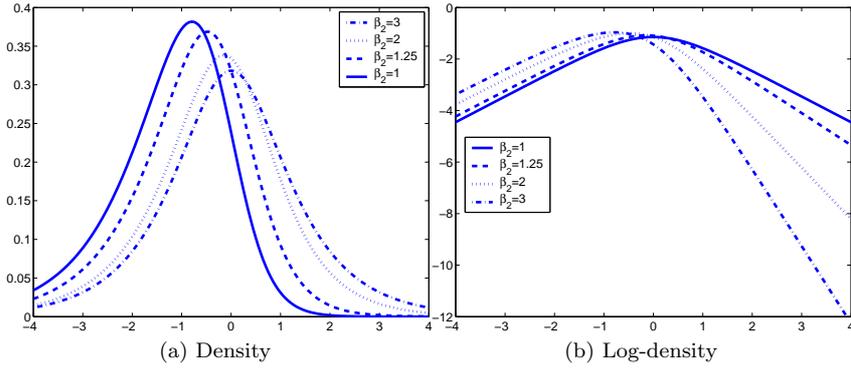
where  $\beta_1 > 0$  and  $\beta_2 > 0$  determine the shape of the density. The corresponding cumulative distribution function is

$$G(x; \beta_1, \beta_2) = \frac{B_{F^{-1}(x)}(\beta_1, \beta_2)}{B(\beta_1, \beta_2)} \quad \text{with} \quad B_u(p, q) = \int_0^u t^{p-1}(1-t)^{q-1} dt.$$

Introducing a location parameter  $\mu \in \mathbb{R}$  and a scale parameter  $\sigma > 0$ , the BHS density from (6) generalizes to

$$g(x) = \frac{B(\beta_1, \beta_2)^{-1}}{\sigma \pi \cosh(\frac{x-\mu}{\sigma})} \left[\frac{2}{\pi} \arctan(e^{\frac{x-\mu}{\sigma}})\right]^{\beta_1-1} \left[1 - \frac{2}{\pi} \arctan(e^{\frac{x-\mu}{\sigma}})\right]^{\beta_2-1}.$$

Different densities and their corresponding log-densities with  $\mu = 0$ ,  $\sigma = 1$  and varying  $\beta_1, \beta_2$  are plotted in figure 1.



**Fig. 1.** Density and log-density for  $\beta_1 = 1$

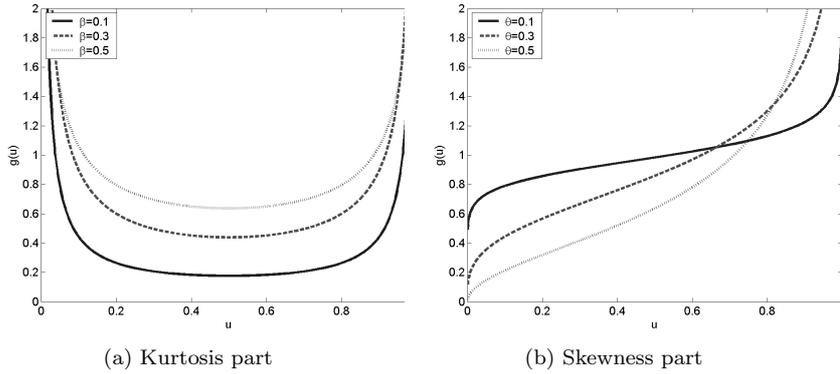
Define  $\theta \equiv \frac{\beta_1 - \beta_2}{2}$  and  $\beta \equiv \frac{\beta_1 + \beta_2}{2} > 0$ . Then  $\beta + \theta = \beta_1$  and  $\beta - \theta = \beta_2$ , and equation (3) can be rewritten as

$$\begin{aligned} w(x; \beta, \theta) &= \frac{1}{B(\beta + \theta, \beta - \theta)} x^{\beta + \theta - 1} (1 - x)^{\beta - \theta - 1} \\ &= C(\beta, \theta) \cdot \frac{x^{\beta - 1} (1 - x)^{\beta - 1}}{B(\beta, \beta)} \cdot \frac{\sin(\pi\theta) x^\theta (1 - x)^{-\theta}}{\pi\theta}, \end{aligned} \quad (7)$$

where  $C(\beta, \theta) = 1$  only if  $\beta = 1$ . Thus, the weighting density can be partitioned into two parts, where the first part essentially governs the amount of kurtosis and the second part the amount of skewness (see figure 2, where both parts are plotted separately). Consequently, a second parameterization of BHS density is given by

$$g(x) = \frac{1}{B(\beta + \theta, \beta - \theta)} \frac{\pi \cosh(x)}{\pi \cosh(x)} \left[ \frac{2 \arctan(e^x)}{\pi} \right]^{\beta + \theta - 1} \left[ 1 - \frac{2 \arctan(e^x)}{\pi} \right]^{\beta - \theta - 1},$$

where symmetry corresponds to  $\theta = 0$ .



**Fig. 2.** Decomposition of the weighting density

In order to ensure the existence of the Beta function in the last equation, both  $\beta + \theta$  and  $\beta - \theta$  have to be positive. Hence, it is required that  $|\theta| < \beta$ , i.e. highly leptokurtic data (that means small  $\beta$ ) induce higher restrictions on  $\theta$ . It also becomes obvious from the above parameterization that  $\beta_1$  and  $\beta_2$  commonly determine skewness and kurtosis (measured by the third and fourth standardized moment within this work).

## 2.2 Properties of the BHS distribution

**Lemma 1 (Asymmetry and kurtosis).** *The BHS distribution with parameters  $\mu, \sigma, \beta_1, \beta_2$  is symmetric about  $\mu$  for  $\beta \equiv \beta_1 = \beta_2$ . Moreover, it is skewed to the right for  $\beta_1 > \beta_2$  and skewed to the left for  $\beta_1 < \beta_2$ . Assume that  $\beta_1 = \beta_2 \equiv \beta$ . Then, kurtosis increases if  $\beta$  decreases and vice versa.*

**Lemma 2 (Tail behavior).** *The BHS distribution has exponentially decaying tails. In particular, the log-density is asymptotically linear with slope determined by  $\beta_1$  and  $\beta_2$ , respectively.*

*Proof.* Assume  $\mu = 0, \sigma = 1$  and focus on the right tail of the BHS distribution. From

$$\lim_{x \rightarrow \infty} \left( \frac{1}{\cosh(x)} - 2 \exp(-x) \right) = 0, \quad \lim_{x \rightarrow \infty} \left[ \frac{2}{\pi} \arctan(e^x) \right]^{\beta_1 - 1} = 1,$$

and

$$\left[ 1 - \frac{2}{\pi} \arctan(e^x) \right]^{\beta_2 - 1} \sim C \exp((1 - \beta_2)x)$$

we conclude that for large  $x$

$$g(x; \beta_1, \beta_2) \sim C \exp(-x) \exp((1 - \beta_2)x) = C \exp(-\beta_2 x), \quad C = \frac{(2/\pi)^{\beta_2}}{B(\beta_1, \beta_2)}.$$

In particular,  $\beta_2 < 1$  corresponds to distributions with heavier than plain exponential tails,  $\beta_2 > 1$  distributions with lighter than plain exponential tails. The same argument is true for the left tail.

Additionally, the score function for the BHS distribution is derived which plays an important role in the theory of rank test (see, e.g. Kravchuk, 2005, for  $\beta_1 = \beta_2 = 1$ )

**Lemma 3 (Score function).** *With  $\zeta(x) \equiv \arctan(e^x)$  the score function of a BHS variable is given by*

$$\begin{aligned} \psi(x; \beta_1, \beta_2) &= -\frac{g'(x; \beta_1, \beta_2)}{g(x; \beta_1, \beta_2)} \\ &= \frac{\tanh(x) \zeta(x) (e^{2x} + 1) (2\zeta(x) - \pi) + e^x \beta_1 (\pi - 2\zeta(x))}{(1 + e^{2x}) \zeta(x) (2\zeta(x) - \pi)} \\ &\quad - \frac{e^x \pi - 2e^x \zeta(x) (2 - \beta_2)}{(1 + e^{2x}) \zeta(x) (2\zeta(x) - \pi)}. \end{aligned}$$

Setting  $\beta_1 = \beta_2 = 1$ , the last equation reduces to  $\psi(x; 1, 1) = \tanh(x)$ .

Finally, it can be shown (see Appendix A for a detailed proof) that BHS densities are unimodal for all  $\beta_1, \beta_2 > 0$ . This is not valid for the Beta-normal and the Beta-Student-t distribution, in general.

**Lemma 4 (Unimodality).** *The BHS distribution is unimodal for  $\beta_1, \beta_2 > 0$ .*

### 2.3 Special and limiting cases

First of all, for  $\beta_1 = \beta_2 = 1$  the hyperbolic secant distribution is recovered. Setting  $\beta_2 = 1$  or  $\beta_1 = 1$ , skew hyperbolic secant distributions can be obtained. A generalized symmetric family of hyperbolic secant distributions is achieved for  $\beta_1 = \beta_2 = \beta$ , where  $\beta$  governs the amount of kurtosis. Like the Beta-logistic distribution and the Beta-normal distribution, the BHS distribution converges to the normal distribution for  $\beta_1, \beta_2 \rightarrow \infty$ .

## 2.4 Moments of the BHS distribution

Obviously, the exponential tail behaviour of the BHS distribution guarantees the existence of all moments. In particular, the  $m^{\text{th}}$  non-central moment of a BHS density is given by

$$E(X^m) = \frac{1}{B(\beta_1, \beta_2)} \int_0^1 \ln^m(\tan(\frac{\pi}{2}u)) u^{\beta_1-1} (1-u)^{\beta_2-1} du.$$

From Gradshteyn and Ryzhik (1994), formula 1.518.3 and 9.616 we can write

$$\tan(\frac{\pi}{2}u) = \ln(\frac{\pi}{2}u) + \sum_{k=1}^{\infty} \frac{(2^{2k-1} - 1)\zeta(2k)}{k2^{2k-1}} u^{2k} = \ln(\frac{\pi}{2}u) + u^2 \sum_{k=0}^{\infty} a_k u^{2k}$$

with the usual Riemann zeta function

$$\zeta(2k) = \sum_{l=1}^{\infty} \frac{1}{l^{2k}} \text{ and } a_k = \frac{(2^{2k+1} - 1)\zeta(2k+2)}{(k+1)2^{2k+1}}. \quad (8)$$

Using the notation

$$\frac{\partial^v}{\partial p^v} B(p, q) \equiv B^{v,0}(p, q), \quad B^{0,0}(p, q) = B(p, q),$$

the following lemma can be derived.

**Lemma 5 (Moments of the BHS distribution).** *Assume that  $m > 0$ .*

$$E(X^m) = \frac{1}{B(\beta_1, \beta_2)} \left[ \sum_{j=0}^m \binom{m}{j} \ln^{m-j}(\frac{\pi}{2}) B^{j,0}(\beta_1, \beta_2) + \sum_{k=0}^{\infty} \sum_{j=1}^m \binom{m}{j} a_k^{(j)} \sum_{i=0}^{m-j} \binom{m-j}{i} \ln^{m-j-i}(\frac{\pi}{2}) B^{i,0}(2k+2j+\beta_1, \beta_2) \right],$$

where

$$a_0^{(j)} = a_0^j, \quad a_k^{(j)} = \frac{1}{ka_0} \sum_{i=1}^k (ij - k + i) a_i a_{k-i}^{(j)}, \quad k \geq 1.$$

In particular, the mean of the BHS distribution is given by

$$E(X) = \ln(\frac{\pi}{2}) + \psi(\beta_1) - \psi(\beta_1 + \beta_2) + \sum_{k=0}^{\infty} a_k \frac{B(2k+2+\beta_1, \beta_2)}{B(\beta_1, \beta_2)}. \quad (9)$$

with  $a_k$  from (8). Note that  $\psi$  denotes the digamma function in the last equation. In contrast to (9), the corresponding formula for the Beta-logistic distribution is given by

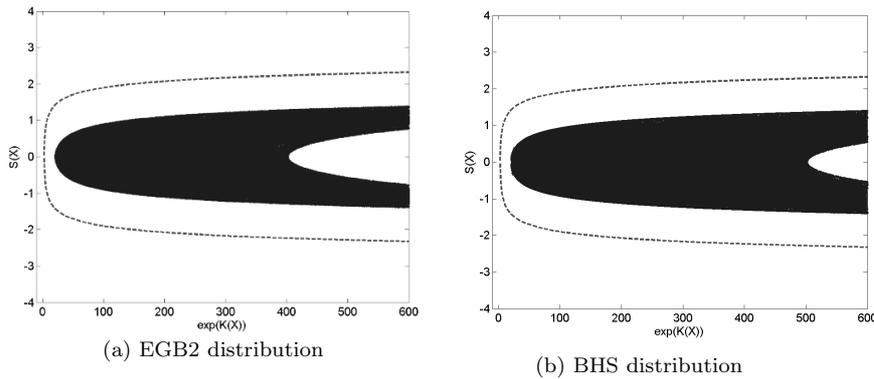
$$E(X) = \psi(\beta_1) - \psi(\beta_2).$$

From the first four moments we can deduce the skewness and kurtosis coefficients  $M_3$  and  $M_4$  (i.e the third and fourth standardized moments) for different parameter combinations of the BHS distribution.

## 2.5 Moment ratio diagrams

Moment ratio diagrams have been introduced for Pearson-type distributions by Elderton and Johnson (1969) in order to provide a useful visual assessment of skewness and kurtosis. The classical moment ratio plot consists of all possible pairs  $(M_3, M_4)$  that can be obtained through different combinations of the shape parameters of the underlying distributions. In general, the relation  $M_3^2 < M_4 - 1$  for  $M_4 > 0$  holds, i.e. for a given level of kurtosis only a finite range of skewness may be spanned.

Due to the bi-modality of the Beta-normal distribution and the non-existence of some moments for the Beta-Student- $t$  distribution we only compare the BHS distribution with the Beta-logistic (EGB2) distribution in figure 3, below.



**Fig. 3.** Moment ratio diagrams

The possible combinations of skewness and kurtosis (for a given distribution) are indicated by the black area which was generated using a

large number of random numbers from the domain of the shape parameters  $(\beta_1, \beta_2)$ . The dashed line (encompassing the black area) corresponds to the boundary mentioned above. Note that we plotted the *exponentiated* kurtosis against the skewness in order to highlight the differences between EGB2 distribution and BHS distribution. It then becomes visible that the achievable area of the BHS distribution includes that of the EGB2 distribution.

### 3 Parameter estimation using maximum likelihood

Suppose that  $X_1, \dots, X_n$  are an *iid* random sample from a BHS distribution. Introducing a scale parameter  $\sigma > 0$  and a location parameter  $\mu \in \mathbb{R}$ , the log-likelihood function is given by

$$\begin{aligned} \ell(\theta) = n \log \left( \frac{(2/\pi)^{\beta_1 + \beta_2 - 2}}{B(\beta_1, \beta_2) \pi \sigma} \right) + \sum_{i=1}^n \{ & (\beta_1 - 1) \log(\arctan(\exp(x_i^*))) \\ & + (\beta_2 - 1) \log \left( \frac{\pi}{2} - \arctan(\exp(x_i^*)) \right) - \log(\cosh(x_i^*)) \}. \end{aligned}$$

where  $x_i^* = (x_i - \mu)/\sigma$  and  $\theta = (\mu, \sigma, \beta_1, \beta_2)$ . Taking the partial derivative of the log-likelihood with respect to the parameters  $\mu, \sigma, \beta_1, \beta_2$  we obtain

$$\begin{aligned} 0 = \frac{\partial \ell}{\partial \mu} &= \frac{1}{\sigma} \sum_{i=1}^n \left( \tanh(x_i^*) + \frac{(1 - \beta_1) \exp(x_i^*)}{(1 + e^{2x_i^*}) \arctan(e^{x_i^*})} \right. \\ &\quad \left. + \frac{(\beta_2 - 1) \exp(x_i^*)}{(1 + e^{2x_i^*}) (\pi/2 - \arctan(e^{x_i^*}))} \right), \\ 0 = \frac{\partial \ell}{\partial \sigma} &= - \left( \frac{(2/\pi)^{\beta_1 + \beta_2 - 2}}{B(\beta_1, \beta_2) \pi \sigma} \right)^n \frac{n}{\sigma} \\ &\quad + \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) \left( \tanh(x_i^*) + \frac{(1 - \beta_1) \exp(x_i^*)}{(1 + e^{2x_i^*}) \arctan(e^{x_i^*})} \right. \\ &\quad \left. + \frac{(\beta_2 - 1) \exp(x_i^*)}{(1 + e^{2x_i^*}) (\pi/2 - \arctan(e^{x_i^*}))} \right) \\ 0 = \frac{\partial \ell}{\partial \beta_1} &= \frac{n}{4^n} \left( \frac{(2/\pi)^{\beta_1 + \beta_2} \pi}{B(\beta_1, \beta_2) \sigma} \right)^n \left( \log \left( \frac{2}{\pi} \right) - \frac{B^{(1,0)}(\beta_1, \beta_2)}{B(\beta_1, \beta_2)} \right) \\ &\quad + \sum_{i=1}^n \log(\arctan(\exp(x_i^*))), \end{aligned}$$

$$0 = \frac{\partial \ell}{\partial \beta_2} = \frac{n}{4^n} \left( \frac{\left(\frac{2}{\pi}\right)^{\beta_1 + \beta_2} \pi}{B(\beta_1, \beta_2) \sigma} \right)^n \left( \log \left( \frac{2}{\pi} \right) - \frac{B^{(0,1)}(\beta_1, \beta_2)}{B(\beta_1, \beta_2)} \right) + \sum_{i=1}^n \log(\pi/2 - \arctan(\exp(x_i^*))).$$

We solve the equations above iteratively to obtain  $\hat{\beta}_1, \hat{\beta}_2, \hat{\mu}, \hat{\sigma}$ .

#### 4 Generalizations: EGB2 versus BHS distribution

In order to discriminate between Beta-logistic (*EGB2*) and BHS distribution we can plug a parent distribution into (3) which includes both logistic distribution and hyperbolic secant distribution. A promising choice is the GSH distribution of Vaughan (2002) with kurtosis parameter  $t$  and density

$$f_{GSH}(x; t) = c_1(t) \cdot \frac{\exp(x)}{\exp(2x) + 2a(t)\exp(x) + 1}, \quad x \in \mathbb{R} \quad (10)$$

$$\text{with } \begin{cases} a(t) = \cos(t), & c_1(t) = \frac{\sin(t)}{t} \text{ for } -\pi < t \leq 0, \\ a(t) = \cosh(t), & c_1(t) = \frac{\sinh(t)}{t} \text{ for } t > 0 \end{cases}.$$

The GSH distribution includes the logistic distribution ( $t = 0$ ) and the hyperbolic secant distribution ( $t = -\pi/2$ ) as special cases and has cumulative distribution function given by

$$F_{GSH}(x; t) = \begin{cases} 1 + \frac{1}{t} \operatorname{arccot} \left( -\frac{\exp(x) + \cos(t)}{\sin(t)} \right) & \text{for } t \in (-\pi, 0), \\ \frac{\exp(\pi x / \sqrt{3})}{1 + \exp(\pi x / \sqrt{3})} & \text{for } t = 0, \\ 1 - \frac{1}{t} \operatorname{arccoth} \left( \frac{\exp(x) + \cosh(t)}{\sinh(t)} \right) & \text{for } t > 0. \end{cases}$$

Thus, we can apply a simple likelihood ratio test to the hypothesis

$$H_0 : t = 0 \text{ (EGB2)} \quad \text{against} \quad H_0 : t = -\pi/2 \text{ (BHS)}.$$

## 5 Applications

### 5.1 Strength of glass fibre

Our first example corresponds to that of Jones and Faddy (2003) who analyzed the strengths of glass fibre. This data set is 'sample 1' of Smith and Naylor (1987) and deals with the breaking strength of  $n = 63$  glass fibres of length 1.5 cm, originally obtained by workers at the UK National Physical Laboratory. Due to the apparent skewness in the data set (see figure 4(a) for a histogram of the data), Jones and Faddy (2003) fitted a Beta-Student- $t$  distribution – using a reparameterized version – to the data, estimating the unknown parameters by means of maximum likelihood.

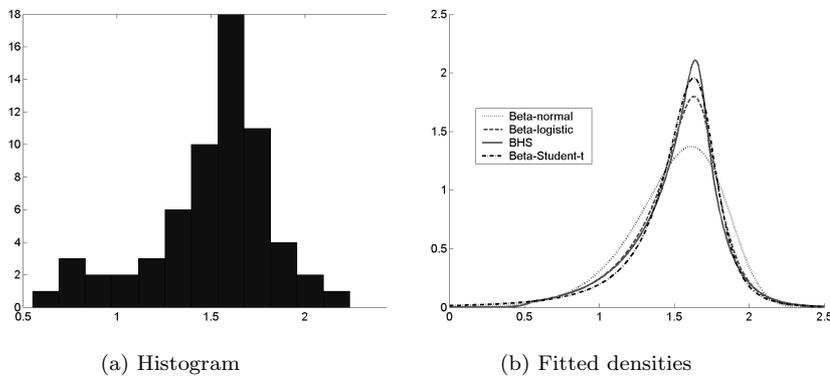


Fig. 4. Strength of glass fibre

Additionally, we fitted a Beta-normal, a Beta-logistic (EGB2), a Beta-hyperbolic secant (BHS) and a Beta-GSH distribution to the data. The estimation results are summarized in table 1, below. Graphs of the fitted densities are provided by figure 4(b).

Regarding the log-likelihood value  $\mathcal{L}$ , the Beta-normal distribution seems to fit worse. Both Beta-logistic and Beta-hyperbolic secant distribution outperform the Beta-Student- $t$  distribution, in particular, if we account for the number of parameters  $k$  and focus on the criterion of Akaike, i.e.  $AIC = -2\mathcal{L} + 2k$ . Moreover, the log-likelihood value of the BHS distribution is higher than that of the EGB2 distribution. Finally, the Beta-GSH distribution provides evidence in favor of the BHS distribution against the EGB2 distribution.

**Table 1.** Estimation results for the glass fibre data set

Distribution	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\nu}/\hat{t}$	$\mathcal{L}$	AIC
Normal	1.51 [0.0409]	0.322 [0.0287]	-	-	-	-17.91	39.82
Beta-Normal	2.60 [0.2005]	0.475 [0.1558]	0.5946 [0.37]	23.66 [4.7340]	-	-14.06	36.11
Beta-Logistic	1.67 [0.0460]	0.041 [0.0393]	0.1450 [0.14]	0.31 [0.3085]	-	-10.49	28.99
BHS	1.65 [0.0400]	0.043 [0.0662]	0.1451 [0.23]	0.28 [0.4638]	-	-10.02	28.03
Beta-Student- $t$	1.70 [0.0695]	0.621 [0.2099]	49.345 [43.4]	56.83 [46.484]	0.12 [0.0688]	-11.41	32.82
Beta-Student- $t_2$	1.70 [0.0763]	0.226 [0.0958]	1.1073 [0.60]	2.08 [1.1947]	-	-11.93	33.86
Beta-GSH	1.65 [0.0383]	0.071 [0.0861]	0.2270 [0.26]	0.43 [0.4761]	-2.00 [0.5939]	-9.90	29.80

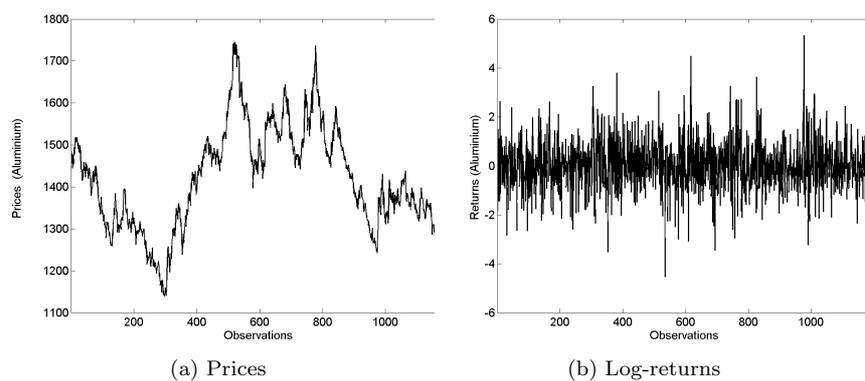
Concerning the estimation results of the Student- $t$ , the parameters  $\beta_1, \beta_2, \nu$  seem to be poorly identified. We therefore fix the number of degrees at 2 as in Jones and Faddy (2003). Note that the 6th column of table 1 contains the estimated shape parameter beyond  $\hat{\beta}_1$  and  $\hat{\beta}_2$ , i.e. the estimated degrees of freedom  $\hat{\nu}$  for a Beta-Student- $t$  distribution and the estimated  $t$  of the Beta-GSH distribution, respectively.

## 5.2 Returns aluminium

Secondly, we focus on the series of the daily aluminium prices (in US-Dollar/Tonne) from January 1999 to September 2002 ( $N = 1195$  observations) which can be obtained from the LME (London Metal Exchange).<sup>1</sup> The series of prices and corresponding log-returns (i.e. difference of consecutive log-prices) are displayed in figure 5.

The (sample) mean of the log-returns is  $-0.0139$  with a (sample) standard deviation of 1.0560. Moreover, there seems to be a certain amount of skewness in the data set (the skewness coefficient – measured by the third standardized moments – is given by 0.2398), whereas the kurtosis coefficient – in terms of the fourth standardized moments – is 4.4250, reflecting the leptokurtosis of the data. The results of a maximum likelihood estimation are summarized in table 2, below.

<sup>1</sup> Download under <http://www.lme.co.uk>.



**Fig. 5.** Prices and log-returns of aluminium 05/01/98 to 30/09/02

**Table 2.** Unconditional fit to the aluminium returns

Distribution	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\nu}/\hat{t}$	$\mathcal{L}$	AIC
Normal	-0.014 [0.031]	1.056 [0.022]	-	-	-	-1758.8	3521.5
Beta-Normal	-1.218 [0.697]	1.728 [0.586]	3.980 [3.058]	1.520 [0.792]	-	-1753.0	3514.0
Beta-Logistic	-0.248 [0.087]	0.497 [0.091]	0.932 [0.243]	0.719 [0.182]	-	-1733.6	3475.2
BHS	-0.282 [0.099]	0.921 [0.156]	1.738 [0.445]	1.355 [0.336]	-	-1733.5	3474.9
Beta-Student-t	-0.331 [0.167]	3.874 [3.310]	49.65 [128.1]	47.013 [126.384]	0.26 [0.473]	-1734.0	3478.0
Beta-Student- $t_2$	-0.294 [0.285]	1.643 [0.563]	3.804 [3.013]	3.146 [2.171]	-	-1734.3	3478.7
Beta-GHS	-0.286 [0.089]	0.966 [0.895]	1.819 [1.517]	1.419 [1.285]	-1.63 [1.026]	-1733.5	3476.9

Though this data set is totally different to the glass fibre data, the results are nearly identical (concerning the order of the log-likelihood values). Again, the Beta-GSH distribution favors the BHS distribution against the Beta-Logistic distribution with  $\hat{t} = -1.63$ , both of which outperform Beta-normal and Beta-Student-t. Again, the shape parameters of the Beta-Student-t seem to be unidentified.

## 6 Summary

A new class of probability densities (the so-called BHS-distribution family) is introduced which arises as special case from the general family explored by Jones (2004) if the hyperbolic secant distribution is chosen as "parent distribution". It exhibits similar behavior and properties like the log- $F$  or EGB2 distribution. In particular, the range of possible skewness and kurtosis combinations of the BHS distribution includes that of the EGB2 distribution. Moreover, a generalized distribution model is introduced which includes both EGB2 and BHS distribution. Application to glass fibre data and aluminium returns provides empirical evidence in favor of the BHS distribution.

## 7 Appendix: Proof of uni-modality

In the Jones and Faddy formulation, the density function for a family of Skew Generalized Secant Hyperbolic Distributions is given by

$$g(x) = \frac{1}{B(\beta_1, \beta_2)} f(x) [F(x)]^a [1 - F(x)]^b$$

where  $f(x) = \frac{1}{\pi \cosh(x)}$  so that  $F(x) = 2/\pi \arctan(\exp(x))$ , and we assume  $a, b > -1$ . We want to show this density is unimodal for all choices of  $a$  and  $b$ . Since the functions are all continuous and continuously differentiable, the only critical points for the function  $g$  satisfy  $g'(x) = 0$ . Thus we want to prove that this has exactly one root, and that this yields a relative maximum. Since  $\lim_{x \rightarrow \pm\infty} g(x) = 0$ , then if there is one critical point, it must yield the absolute maximum, so we need to prove there is exactly one root to the derivative equation. After simplification, this can be seen to be equivalent to proving

$$-\sin h(x) + \frac{a}{2 \tan^{-1}(\exp(x))} - \frac{b}{\pi(1 - \frac{2}{\pi} \tan^{-1}(\exp(x)))} = 0$$

has exactly one root. Note that if we set  $u = \tan^{-1}(\exp(x))$ , the last statement is equivalent to showing

$$-(\tan(u) - \cot(u))u(\pi/2 - u) = -\frac{\pi a}{2} + (a + b)u$$

has exactly one root in  $(0, \pi/2)$ . Define

$$h(u) = -(\tan(u) - \cot(u))u(\pi/2 - u)$$

on  $(0, \pi/2)$ . Note that  $h(u + \pi/4)$  is odd on  $(-\pi/4, \pi/4)$ . Also,  $h(\pi/4) = 0$  and we set  $h(0) = \lim_{u \rightarrow 0^+} h(u) = \pi/2$ , and  $h(\pi/2) = \lim_{u \rightarrow (\pi/2)^-} h(u) = -\pi/2$ . Note that

$$h'(u) = -\frac{u(\pi/2 - u)}{\sin^2(u) \cos^2(u)} - (\tan(u) - \cot(u))(\pi/2 - 2u)$$

and

$$h''(u) = \frac{4}{\sin^3(2u)} [4 \cos(2u)u(\pi/2 - u) - 2 \sin(2u)(\pi/2 - 2u) - \cos(2u) \sin^2(2u)].$$

We want to prove that  $h$  is concave down on  $(0, \pi/4)$  and concave up on  $(\pi/4, \pi/2)$ . The second fact will follow from the first, and the symmetry property of  $h$  noted earlier. Thus, we want to prove that  $h''(u) < 0$  on  $(0, \pi/4)$ . By using trigonometric identities, we can show this is equivalent to proving the function  $k(v) = v(\pi - v) \cos v - 2(\pi/2 - v) \sin v - \cos v \sin^2 v < 0$

on  $(0, \pi/2)$ . Now  $k(0) = 0$ ,  $k(\pi/2) = 0$ . Note that  $k'(v) = \sin v[v^2 - \pi v + 3 \sin^2(v)]$ .

Set  $z(v) = 3 \sin^2 v + v^2 - \pi v$ , and note that  $z(0) = 0$  and  $z(\pi/2) = 3 - \pi^2/4 > 0$ . We have  $z'(v) = 3 \sin(2v) + 2v - \pi$  and  $z''(v) = 6 \cos(2v) + 2$ . Clearly,  $z'' > 0$  if  $\cos(2v) > -1/3$  and  $z'' < 0$  if  $\cos(2v) < -1/3$ . In the interval  $(0, \pi/2)$  there is a unique value, say  $\alpha_0$  so that  $\cos(2\alpha_0) = -1/3$  and hence on  $(0, \alpha_0)$ ,  $z''(v) > 0$  and on  $(\alpha_0, \pi/2)$ ,  $z''(v) < 0$ . Because  $z'(0) = -\pi$  and  $z'(\pi/2) = 0$ , there is a unique value  $\alpha_1 \in (0, \alpha_0)$  for which  $z'(\alpha_1) = 0$ . We then have  $z'(v) < 0$  on  $(0, \alpha_1)$  and  $z'(v) > 0$  on  $(\alpha_1, \pi/2)$ . From the values  $z(0) = 0$  and  $z(\pi/2) > 0$ , and the properties of  $z'$ , there is a unique value  $\alpha_2 \in (0, \pi/2)$  for which  $z(\alpha_2) = 0$ .

The above shows that  $k'$  has exactly one root in  $(0, \pi/2)$ , call it  $\beta_0$ . It is clear that  $k'(v) < 0$  on  $(0, \beta_0)$ , and  $k'(v) > 0$  on  $(\beta_0, \pi/2)$ . This in turn implies  $k(v) < 0$  on  $(0, \pi/2)$ , since  $k(0) = 0 = k(\pi/2)$ .

The above argument establishes that  $h''(u) < 0$  on  $(0, \pi/4)$ , and therefore  $h$  is concave down on  $(0, \pi/4)$  and concave up on  $(\pi/4, \pi/2)$ . Set  $w(u) = -\frac{\pi a}{2} + (a + b)u$  on  $(0, \pi/2)$ . Since  $w(0) = -\frac{\pi a}{2} < h(0) = \pi/2$  and  $w(\pi/2) = \frac{\pi b}{2} > h(\pi/2) = \pi/2$ , then  $w$  and  $h$  intersect. If these curves intersect on  $(0, \pi/4)$ , they cannot intersect a second time on  $(0, \pi/4)$  (otherwise, since  $h$  is concave down, the line through the intersection points cannot intersect  $h$  a third time on this interval. This means the vertical axis intercept for the line is  $> \pi/2$ , and this is not possible, given the line must intersect the vertical axis at  $(0, -\frac{\pi a}{2})$ ). Further, the line  $w$  cannot intersect  $h$  on  $(\pi/4, \pi/2)$  in this case, since  $w$ 's slope is greater than  $-2$ , and  $-2$  is the slope of the line  $y = \pi/2 - 2u$  joining the points  $(0, \pi/2)$ ,  $(\pi/4, 0)$  and  $(\pi/2, -\pi/2)$  on  $h$ . Hence  $w$  and  $y$  have a unique intersection point, so that if  $w$  intersects  $h$  on  $(\pi/4, \pi/2)$ , this will force  $w$  and  $y$  to intersect again, a contradiction. A similar analysis shows that if  $w$  and  $h$  do intersect on  $(\pi/4, \pi/2)$ , they do so uniquely, and do not intersect on  $(0, \pi/4)$ .

Altogether, this means that  $g'$  has exactly one root in  $(-\infty, \infty)$ . It then also follows this yields a relative maximum (and hence absolute maximum) since  $g'$  is positive to the left of the root, and negative to the right.  $\square$

## References

- Aroian LA (1941) A Study of R. A. Fisher's  $z$  Distribution and the Related  $f$  Distribution. *Annals of Mathematical Statistics* **12**: 429-448
- Azzalini A (1985) A Class of Distributions which includes the Normal Ones. *Scandinavian Journal of Statistics* **12**: 171-178.
- Azzalini A (1986) Further Results on a Class of Distributions which includes the Normal Ones. *Statistica* **46**: 199-208.
- Elderton WP, Johnson NL (1969) *Systems of Frequency Curves*. Addison Wesley, Reading.
- Fernández C, Osiewalski J, Steel MFJ (1995) Modelling and Inference with  $\nu$ -spherical Distributions. *Journal of the American Statistical Society* **90**: 1331-1340.

- Eugene N, Lee C, Famoye F (2002) Beta-Normal Distribution and its Applications. *Communications in Statistics, Theory and Methods* **31**(4): 497-512.
- Ferreira JTAS, Steel, MFJ (2004) A Constructive Representation of Univariate Skewed Distributions. Working Paper, Department of Statistics, University of Warwick.
- Gradshteyn IS, Ryzhik IM (1994) *Tables of Integrals, Series, and Products*. Academic Press, New York.
- Jones MC (2004) Families of Distributions Arising from Distributions of Order Statistics. *Test* **13**(1): 1-43.
- Jones MC, Faddy, MJ (2004) A Skew Extension of the t-distribution, with Applications. *Journal of the Royal Statistical Society, Series B* **65**(1): 159-174.
- Kravchuk, OY (2005) Rank Test of Location Optimal for Hyperbolic Secant Distribution. *Communications in Statistics (Theory and Methods)*, **34**(7): 1617-1630.
- Perks, WF (1932) On Some Experiments in the Graduation of Mortality Statistics. *Institute of Actuaries Journals, Series B* **58**: 12-57.
- Prentice RL (1975) Discrimination among some Parametric Models. *Biometrika* **62**: 607-614.
- Smith RL, Naylor JC (1987) A Comparison of Maximum Likelihood and Bayesian Estimators for the Three-Parameter Weibull Distribution. *Applied Statistics* **36**(3): 358-369.
- Theodossiou P (1998) Financial data and the skewed generalized t distribution. *Mathematical Science* **44**(12): 1650-1660.
- Vaughan DC (2002) The Generalized Hyperbolic Secant Distribution and its Application. *Communications in Statistics, Theory and Methods* **31**(2): 219-238.