The Log-Beta Generalized Half-Normal Regression Model

Rodrigo R. Pescim[a], Edwin M. M. Ortega[b], Gauss M. Cordeiro[c],
Clarice G. B. Demtrio[d] and G. G. Hamedani[e]

[a]*Departamento de Cincias Exatas*,
Universidade de So Paulo,
Av. Pdua Dias 11, 13418-900 – Piracicaba, SP, Brazil
(edwin@usp.br)

[b]*Departamento de Cincias Exatas*,
Universidade de So Paulo,
Av. Pdua Dias 11, 13418-900 – Piracicaba, SP, Brazil
(rrpescim@usp.br)

[c]*Departamento de Estatstica*,
Universidade Federal de Pernambuco,
Cidade Universitria, 50740-540 – Recife, PE, Brazil
(gausscordeiro@uol.com.br);

[d]*Departamento de Cincias Exatas*,
Universidade de So Paulo,
Av. Pdua Dias 11, 13418-900 – Piracicaba, SP, Brazil
(clarice.demetrio@usp.br)

[e]*Department of Mathematics, Statistics and Computer Science*,
Marquette University,
Katharine Reed Cudahy Hall, Room 314, Milwaukee, Estados Unidos
(g.hamedani@mu.edu)

We introduce a log-linear regression model based on the beta generalized half-normal distribution (Pescim et al., 2010). We formulate and develop a log-linear model using a new distribution so-called the log-beta generalized half normal distribution. We derive expansions for the cumulative distribution and density functions which do not depend on complicated functions. We obtain formal expressions for the moments and moment generating function. We characterize the proposed distribution using a simple relationship between two truncated moments. An advantage of the new distribution is that it includes as special sub-models classical distributions reported in the lifetime literature. We also show that the new regression model can be applied to censored data since it represents a parametric family of models that includes as special cases several widely-known regression models. It therefore can be used more effectively in the analysis of survival data. We investigate the maximum likelihood estimates of the model parameters by considering censored data. We demonstrate that our extended regression model is very useful to the analysis of real data and may give more realistic fits than other special regression models.

Keywords: Beta generalized half normal; Censored data; Regression model; Survival function.

1. Introduction

The fatigue is a structural damage which occurs when a material is exposed to stress and tension fluctuations. Statistical models allow to study the random variation of the failure time associated to

Rodrigo R. Pescim, Edwin M. M. Ortega, Gauss M. Cordeiro, Clarice G. B. Demtrio and G. G. Hamedani

materials exposed to fatigue as a result of different cyclical patterns and strengths. Some popular models used to describe the lifetime process under fatigue are the half normal (HN) and Birnbaum-Saunders (BS) distributions. When modeling monotone hazard rates, the HN and BS distributions may be an initial choice because of its positively skewed density shapes. However, they do not provide a reasonable parametric fit for modeling phenomenon with non-monotone failure rates such as the bathtub shaped and unimodal failure rates, which are common in reliability and biological studies. Such bathtub hazard curves have nearly flat middle portions and the corresponding densities have a positive anti-mode. A highly flexible lifetime model, which admits different degrees of kurtosis and asymmetry, is the beta generalized half-normal (BGHN) distribution proposed by Pescim et al. (2010).

In many medical problems, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure, weight and many others. Parametric models to estimate univariate survival functions and for censored data regression problems are widely used. Different forms of regression models have been proposed in survival analysis. Among them, the location-scale regression model (Lawless, 2003) is distinguished since it is frequently used in clinical trials. Recently, the location-scale regression model has been used in several research areas such as engineering, hydrology and survival analysis. Lawless (2003) discussed the generalized log-gamma regression models with censored data, Barros et al. (2008) proposed a new class of lifetime regression models for which the errors follow the generalized BS distribution, Carrasco et al. (2008) introduced a regression model considering the modified Weibull distribution, Silva et al. (2008) studied a location-scale regression model using the Burr XII distribution and Silva et al. (2009) worked a location-scale regression model suitable for fitting censored survival times with bathtub-shaped hazard rates. Other applications, we have for example, Ortega et al. (2009) proposed a modified generalized log-gamma regression model to allow the possibility that long-term survivors may be presented in the data and Hashimoto et al. (2010) introduced the log-exponentiated Weibull regression model for interval-censored data. In this article, we propose a log-location regression model with censored observations, based on the BGHN distribution (Pescim et al., 2010), referred to as the log-beta generalized half-normal (LBGHN) regression model, which is a feasible alternative for modeling the four existing types of failure rate functions. Also, some properties of the proposed estimators useful for developing asymptotic inference are presented.

The article is organized as follows. In Section 2, we define the LBGHN distribution, present some special cases and provide expansions for its distribution and density functions. Section 3 gives general expansions for the moments and the moment generating function (mgf). In Section 4, we characterize the BGHNG distribution. In Section 5, we propose a LBGHN regression model, estimate the parameters by the method of maximum likelihood and derive the observed information matrix. In Section 6, a real data set is analyzed which shows the flexibility, practical relevance and applicability of our regression model. Section 7 ends with some concluding remarks.

2. The log-beta generalized half-normal distribution

Most classes of generalized beta distributions have been proposed in reliability literature to provide better fits to certain data sets. The BGHN distribution with four parameters $\alpha > 0$, $\theta > 0$, $a > 0$ and $b > 0$, introduced and studied by Pescim et al. (2010), extends the generalized half-normal (GHN) distribution (Cooray and Ananda, 2008) and provides good fits to various types of data. Its density

function for $t > 0$ is given by

$$
f(t) = \frac{\sqrt{\frac{2}{\pi}}\left(\frac{\alpha}{t}\right)\left(\frac{t}{\theta}\right)^{\alpha}\exp\left[-\frac{1}{2}\left(\frac{t}{\theta}\right)^{2\alpha}\right]}{B(a,b)}2^{b-1}\left\{2\Phi\left[\left(\frac{t}{\theta}\right)^{\alpha}\right]-1\right\}^{a-1} \times
$$
$$
\left\{1-\Phi\left[\left(\frac{t}{\theta}\right)^{\alpha}\right]\right\}^{b-1}, \tag{2.1}
$$

where $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function. Note that

$$
\Phi(x) = \frac{1}{2}\left[1+\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)\right] \quad \text{and} \quad \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)dt.
$$

This class of generalized distributions has been receiving considerable attention over the last years, in particular after the works of Eugene et al. (2002) and Jones (2004).

The BGHN distribution contains as special sub-models some well-known distributions. It simplifies to the GHN distribution when $a = b = 1$. If $\alpha = 1$, it reduces to the beta half normal (BHN) distribution. If $b = 1$, it leads to the exponentiated generalized half normal (EGHN) distribution. Further, if $a = b = 1$, in addition to $\alpha = 1$, becomes the HN distribution.

The survival function is given by

$$
S(t) = 1 - I_{2\Phi\left[\left(\frac{t}{\theta}\right)^{\alpha}\right]-1}(a,b) = 1 - \frac{1}{B(a,b)}\int_0^{2\Phi\left[\left(\frac{t}{\theta}\right)^{\alpha}\right]-1}\omega^{a-1}(1-\omega)^{b-1}d\omega, \tag{2.2}
$$

where $I_y(a,b) = B_y(a,b)/B(a,b)$ is the incomplete beta function ratio and

$$
B_y(a,b) = \int_0^y w^{a-1}(1-w)^{b-1}dw
$$

is the incomplete beta function. The hazard rate function corresponding to (2.1) becomes

$$
h(t) = \frac{\sqrt{\frac{2}{\pi}}\left(\frac{\alpha}{t}\right)\left(\frac{t}{\theta}\right)^{\alpha}e^{-\frac{1}{2}\left(\frac{t}{\theta}\right)^{2\alpha}}2^{b-1}\left\{2\Phi\left[\left(\frac{t}{\theta}\right)^{\alpha}\right]-1\right\}^{a-1}\left\{1-\Phi\left[\left(\frac{t}{\theta}\right)^{\alpha}\right]\right\}^{b-1}}{B(a,b)\left\{1-I_{2\Phi\left[\left(\frac{t}{\theta}\right)^{\alpha}\right]-1}(a,b)\right\}}. \tag{2.3}
$$

If T is a random variable with density function (2.1), we write $T \sim \text{BGHN}(\alpha,\theta,a,b)$. The BGHN distribution is easily simulated as follows: if V has a beta distribution with parameters a and b, then the solution of the nonlinear equation $\left(\frac{T}{\theta}\right)^{\alpha} = \Phi^{-1}\left(\frac{V+1}{2}\right)$ has the $\text{BGHN}(\alpha,\theta,a,b)$ distribution. For other properties of the BGHN distribution, see, for example, Pescim et al. (2010).

Henceforth, T is a random variable following the BGHN density function (2.1) and Y is defined by $Y = \log(T)$. It is easy to verify that the density function of Y obtained by replacing $\mu = \log(\theta)$ and $\sigma = \sqrt{2}/2\alpha$ reduces to

$$
f(y) = \frac{\exp\left\{-\frac{1}{2}\exp\left[\left(\frac{y-\mu}{\sigma}\right)\sqrt{2}\right]+\left(\frac{y-\mu}{\sigma}\right)\frac{\sqrt{2}}{2}\right\}2^{b-1}}{\sigma\sqrt{\pi}B(a,b)}\left\{2\Phi\left[\exp\left\{\left(\frac{y-\mu}{\sigma}\right)\frac{\sqrt{2}}{2}\right\}\right]-1\right\}^{a-1}
$$
$$
\times\left\{1-\Phi\left[\exp\left\{\left(\frac{y-\mu}{\sigma}\right)\frac{\sqrt{2}}{2}\right\}\right]\right\}^{b-1}, \quad -\infty < y < \infty, \tag{2.4}
$$

where $-\infty < \mu < \infty$, $\sigma > 0$, $a > 0$ and $b > 0$. We refer to equation (2.4) as the new LBGHN distribution, say $Y \sim \text{LBGHN}(\mu, \sigma, a, b)$, where $-\infty < \mu < \infty$ is the location parameter, $\sigma > 0$ is the scale parameter and $a > 0$ and $b > 0$ are shape parameters. Thus, if $T \sim BGHN(\alpha, \theta, a, b)$ then $Y = \log(T) \sim LBGHN(\mu, \sigma, a, b)$. The LBGHN distribution contains well-known distributions as special sub-models. It simplifies to the log-generalized half-normal (LGHN) distribution when $a = b = 1$. If $\sigma = \sqrt{2}/2$, it reduces to the log-beta half normal (LBHN) distribution. If $b = 1$, it leads to the log-exponentiated generalized half normal (LEGHN) distribution. Further, if $a = b = 1$, in addition to $\sigma = \sqrt{2}/2$, it reduces to the log-half-normal (LHN) distribution. Figures 1 and 2 plots this density function for selected values of the parameters μ, σ, a and b showing that the LBGHN distribution could be very flexible for modeling its kurtosis.
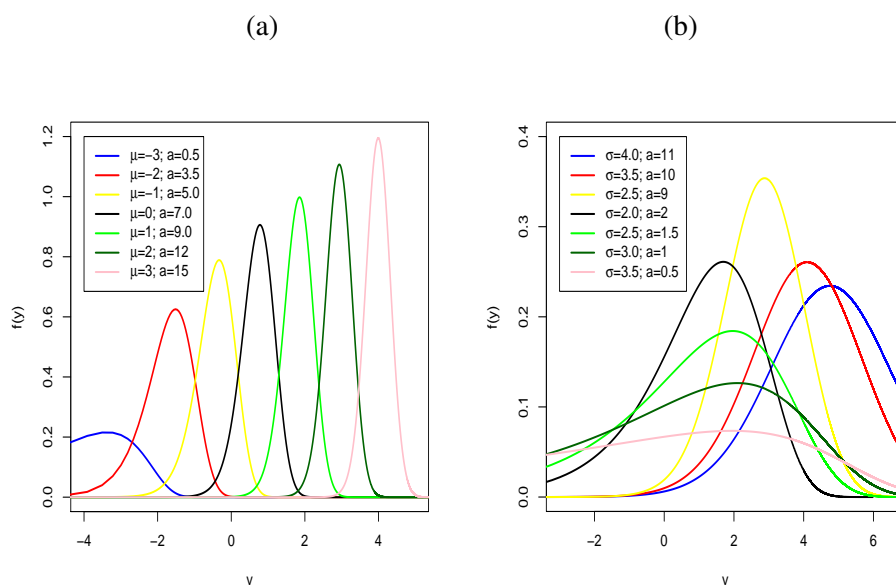
(a)   (b)

Fig. 1. The LBGHN density curves: (a) For some values of μ increasing and a increasing with $\sigma = 1$, and $b = 1$. (b) For some values of σ increasing and a decreasing with $\mu = 0$, and $b = 0.5$.

The corresponding survival function is

$$S(y) = 1 - I_{2\Phi\left[\exp\left\{\left(\frac{y-\mu}{\sigma}\right)\frac{\sqrt{2}}{2}\right\}\right]-1}(a,b). \tag{2.5}$$

The random variable $Z = (Y - \mu)/\sigma$ has density function given by

$$f(z) = \frac{\exp\left\{\frac{1}{2}\left[-\exp\left(z\sqrt{2}\right) + z\sqrt{2}\right]\right\}}{2^{-(b-1)}\sqrt{\pi}B(a,b)}\left\{2\Phi\left[\exp\left(\frac{z\sqrt{2}}{2}\right)\right]-1\right\}^{a-1} \times$$

$$\left\{1 - \Phi\left[\exp\left(\frac{z\sqrt{2}}{2}\right)\right]\right\}^{b-1}. \tag{2.6}$$

We provide two simple formulae for the cumulative distribution function (cdf), probability density function (pdf) and survival function of the LBGHN distribution depending if the parameter $b > 0$ is real non-integer or integer.
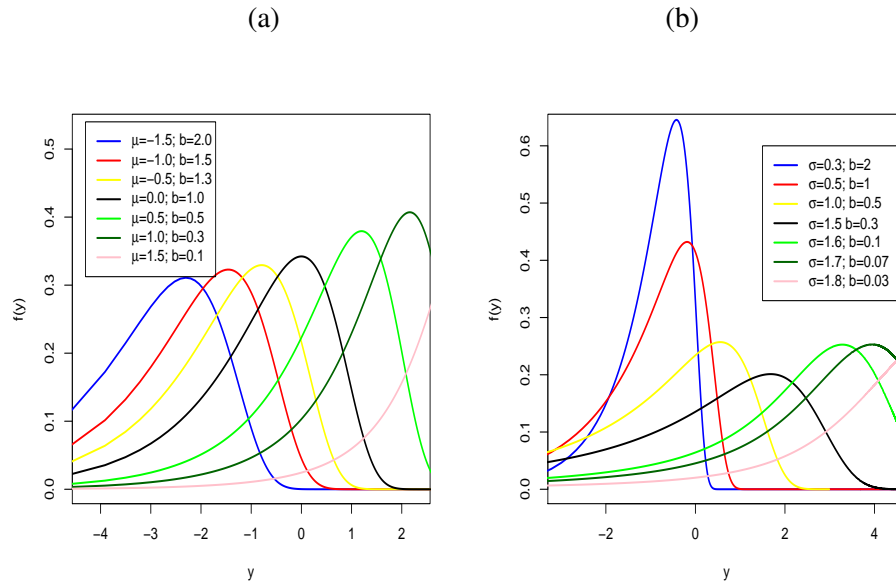
(a)   (b)

Fig. 2. The LBGHN density curves: (a) For some values of μ increasing and b decreasing with $a = 1$, and $\sigma = 1$. (b) For some values of σ increasing and b decreasing with $\mu = 0$, and $a = 0.5$.

Theorem 1:   If $Y \sim LBGHN(\mu, \sigma, a, b)$, then we have the following approximations:

1.1   For $a > 0$ and $b > 0$ real non-integers,

$$F(y) = \frac{1}{B(a,b)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j) j! (a+j)} \left\{ 2\Phi\left[\exp\left\{ \left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right\} \right] - 1 \right\}^{a+j}. \quad (2.7)$$

1.2   For $a > 0$ and $b > 0$ real non-integers,

$$f(y) = \frac{1}{\sigma\sqrt{\pi}} \exp\left\{ -\frac{1}{2}\exp\left[\left(\frac{y-\mu}{\sigma}\right)\sqrt{2} \right] + \left(\frac{y-\mu}{\sigma}\right)\frac{\sqrt{2}}{2} \right\}$$
$$\sum_{j,k=0}^{\infty} \sum_{r=0}^{k} w_{j,k,r}(a,b) \left\{ \mathrm{erf}\left[\frac{1}{\sqrt{2}} \exp\left\{ \left(\frac{y-\mu}{\sigma}\right)\frac{\sqrt{2}}{2} \right\} \right] \right\}^r, \quad (2.8)$$

where $w_{j,k,r}(a,b) = \frac{(-1)^{j+k+r}\Gamma(a+b)\Gamma(a+j)\binom{k}{r}}{\Gamma(a)\Gamma(b-j)\Gamma(a+j-k)j!k!}$.

1.3   The survival function can be written as

$$S(y) = 1 - \sum_{j,k=0}^{\infty} v_{jk}(a,b) S^*(y)^k,$$

where $v_{j,k}(a,b) = \frac{(-1)^{j+k}\Gamma(a+b)\Gamma(a+j+1)}{\Gamma(a)\Gamma(b-j)j!(a+j)\Gamma(a+j+1-k)k!}$ and $S^*(y)$ is the survival function of the LGHN distribution.

Proof 1.1:

First, if $|z| < 1$ and $b > 0$ is real non-integer, it follows that

$$(1-z)^{b-1} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j) j!} z^j. \tag{2.9}$$

Using the representation (2.9), the LBGHN cumulative function for $b > 0$ real non-integer can be expanded from (2.5) as

$$F(y) = \frac{1}{B(a,b)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(b)}{\Gamma(b-j) j!} \int_0^{2\Phi\left\{ \exp\left[\left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right] \right\} - 1} \omega^{a+j-1} d\omega.$$

Solving the last integral, we have

$$F(y) = \frac{1}{\Gamma(a)} \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(a+b)}{\Gamma(b-j) j! (a+j)} \left\{ 2\Phi\left[\exp\left\{ \left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right\} \right] - 1 \right\}^{a+j}. \blacksquare$$

Proof 1.2:

Differentiating $F(y)$ given in Theorem 1.1, we obtain

$$f(y) = \frac{1}{\sigma \sqrt{\pi} \Gamma(a)} \exp\left\{ -\frac{1}{2} \exp\left[\left(\frac{y-\mu}{\sigma}\right) \sqrt{2} \right] + \left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right\}$$
$$\sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(a+b)}{\Gamma(b-j) j!} \left\{ 2\Phi\left[\exp\left\{ \left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right\} \right] - 1 \right\}^{a+j-1}. \tag{2.10}$$

Using the representation (2.9) in (2.10) for a¿0 real non-integer, it can be expanded as

$$f(y) = \frac{1}{\sigma \sqrt{\pi} \Gamma(a)} \exp\left\{ -\frac{1}{2} \exp\left[\left(\frac{y-\mu}{\sigma}\right) \sqrt{2} \right] + \left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right\}$$
$$\sum_{j,k=0}^{\infty} \frac{(-1)^{j+k} \Gamma(a+b) \Gamma(a+j)}{\Gamma(b-j) \Gamma(a+j-k) j! k!} \left\{ 1 - \left\{ 2\Phi\left[\exp\left\{ \left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right\} \right] - 1 \right\} \right\}^k. \tag{2.11}$$

Further, using the binomial expansion and the erf(.) function in (2.11), we obtain after some algebra

$$f(y) = \frac{1}{\sigma \sqrt{\pi}} \exp\left\{ -\frac{1}{2} \exp\left[\left(\frac{y-\mu}{\sigma}\right) \sqrt{2} \right] + \left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right\}$$
$$\sum_{j,k=0}^{\infty} \sum_{r=0}^{k} w_{j,k,r}(a,b) \left\{ \text{erf}\left[\frac{1}{\sqrt{2}} \exp\left\{ \left(\frac{y-\mu}{\sigma}\right) \frac{\sqrt{2}}{2} \right\} \right] \right\}^r,$$

where

$$w_{j,k,r}(a,b) = \frac{(-1)^{j+k+r} \Gamma(a+b) \Gamma(a+j) \binom{k}{r}}{\Gamma(a) \Gamma(b-j) \Gamma(a+j-k) j! k!}. \blacksquare$$

Proof 1.3:

Note that the Theorem 1.1 can be written in terms of the survival function $S^*(y)$ of the LGHN distribution. We have

$$F(y) = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(a+b)}{\Gamma(a)\Gamma(b-j)j!(a+j)} \left[1 - S^*(y)\right]^{a+j}. \tag{2.12}$$

Using the representation (2.9) in equation (2.12) for a¿0 real non-integer, it can be expanded as

$$F(y) = \sum_{j,k=0}^{\infty} \frac{(-1)^{j+k} \Gamma(a+b)\Gamma(a+j+1)}{\Gamma(a)\Gamma(b-j)j!(a+j)\Gamma(a+j+1-k)k!} S^*(y)^k.$$

Finally, we obtain

$$S(y) = 1 - \sum_{j,k=0}^{\infty} v_{j,k}(a,b)S^*(y)^k,$$

where

$$v_{j,k}(a,b) = \frac{(-1)^{j+k}\Gamma(a+b)\Gamma(a+j+1)}{\Gamma(a)\Gamma(b-j)j!(a+j)\Gamma(a+j+1-k)k!}. \blacksquare$$

3. Properties of the log-beta generalized half-normal distribution

We hardly need to emphasize the necessity and importance of the moments in any statistical analysis especially in applied work. Some of the most important features and characteristics of a distribution can be studied through moments (e.g., tendency, dispersion, skewness and kurtosis).

The moments of a random variable Y following the LBGHN density function (2.4) can be expressed parameterized in terms of μ, σ, a and b. We have the theorem.

Theorem 2: If $Y \sim LBGHN(\mu, \sigma, a, b)$, then the sth moment is given by

$$\mu_s' = \frac{1}{\sigma\sqrt{\pi}} \sum_{j,k=0}^{\infty} \sum_{r=0}^{k} \sum_{l=0}^{s} w_{j,k,r}(a,b) \binom{s}{l} \mu^{s-l} (\sqrt{2}\sigma)^{s+1}$$

$$\times \sum_{m_1=0}^{\infty} \cdots \sum_{m_r=0}^{\infty} \frac{(-1)^{m_1+\ldots+m_r}}{(2m_1+1)\ldots(2m_r+1)m_1!\ldots m_r!}$$

$$\times \sum_{q=0}^{s} \binom{s}{q} [\log(2)]^{s-q} 2^{-1/2-s} \Gamma^{(q)}\left(m_1+\ldots+m_r+\frac{r+1}{2}\right), \tag{3.1}$$

where $w_{j,k,r}(a,b)$ is just defined after (2.8) and

$$\Gamma^{(q)}\left(m_1+\ldots+m_r+\frac{r+1}{2}\right) = \frac{\partial^q \Gamma(m_1+\ldots+m_r+\frac{r+1}{2})}{\partial(m_1+\ldots+m_r+\frac{r+1}{2})^q}.$$

Proof:

The sth moment of the LBGHN distribution is $\mu_s' = \int_{-\infty}^{\infty} y^s f(y) dy$. Using the Theorem 1.2, we obtain

$$\mu'_s = \frac{1}{\sigma\sqrt{\pi}} \sum_{j,k=0}^{\infty} \sum_{r=0}^{k} w_{j,k,r}(a,b) \frac{2\sigma}{\sqrt{2}}$$
$$\times \int_0^\infty \left[\mu + \frac{2\sigma}{\sqrt{2}} \log(u)\right]^s \exp(-u^2/2) \left[\operatorname{erf}(u/\sqrt{2})\right]^r du. \tag{3.2}$$

Using the binomial expansion and the erf(.) function in equation (3.2), we obtain

$$\mu'_s = \frac{1}{\sigma\sqrt{\pi}} \sum_{j,k=0}^{\infty} \sum_{r=0}^{k} w_{j,k,r}(a,b) \sum_{l=0}^{s} \binom{s}{l} \mu^{s-l} \left(\frac{2\sigma}{\sqrt{2}}\right)^{s+1}$$
$$\times \sum_{m_1=0}^{\infty} \cdots \sum_{m_r=0}^{\infty} \frac{(-1)^{m_1+\ldots+m_r}}{2^{m_1+\ldots+m_r+r/2}(2m_1+1)\ldots(2m_r+1)m_1!\ldots m_r!}$$
$$\times \int_0^\infty u^{2(m_1+\ldots+m_r)+r} \exp(-u^2/2)[\log(u)]^s du. \tag{3.3}$$

Setting $x = u^2/2$ and after some manipulation, it reduces to

$$\mu'_s = \frac{1}{\sigma\sqrt{\pi}} \sum_{j,k=0}^{\infty} \sum_{r=0}^{k} w_{j,k,r}(a,b) \sum_{l=0}^{s} \binom{s}{l} \mu^{s-l} \left(\frac{2\sigma}{\sqrt{2}}\right)^{s+1}$$
$$\times \sum_{m_1=0}^{\infty} \cdots \sum_{m_r=0}^{\infty} \frac{(-1)^{m_1+\ldots+m_r}}{(2m_1+1)\ldots(2m_r+1)m_1!\ldots m_r!}$$
$$\times \sum_{q=0}^{s} \binom{s}{q} [\log(2)]^{s-q} 2^{-\frac{1}{2}-s} \int_0^\infty x^{m_1+\ldots+m_r+\frac{r+1}{2}-1} e^{-x} [\log(x)]^q dx. \tag{3.4}$$

Finally, we have

$$\mu'_s = \frac{1}{\sigma\sqrt{\pi}} \sum_{j,k=0}^{\infty} \sum_{r=0}^{k} \sum_{l=0}^{s} w_{j,k,r}(a,b) \binom{s}{l} \mu^{s-l} (\sqrt{2}\sigma)^{s+1}$$
$$\times \sum_{m_1=0}^{\infty} \cdots \sum_{m_r=0}^{\infty} \frac{(-1)^{m_1+\ldots+m_r}}{(2m_1+1)\ldots(2m_r+1)m_1!\ldots m_r!}$$
$$\times \sum_{q=0}^{s} \binom{s}{q} [\log(2)]^{s-q} 2^{-1/2-s} \Gamma^{(q)}\left(m_1+\ldots+m_r+\frac{r+1}{2}\right),$$

where $w_{j,k,r}(a,b)$ is just defined after equation (2.8) and

$$\Gamma^{(q)}\left(m_1+\ldots+m_r+\frac{r+1}{2}\right) = \frac{\partial^q \Gamma(m_1+\ldots+m_r+\frac{r+1}{2})}{\partial(m_1+\ldots+m_r+\frac{r+1}{2})^q}. \blacksquare$$

The skewness and kurtosis measures can now be calculated from the ordinary moments using well-known relationships. Plots of the skewness and kurtosis for some choices of the parameter b as function of a, and for some choices of the parameter a as function of b, for $\mu = 0.01$ and $\sigma = 90$, are shown in Figures 3 and 4, respectively. These figures immediately reveal that the skewness and kurtosis curves increase (decrease) with b (a) for fixed a (b).
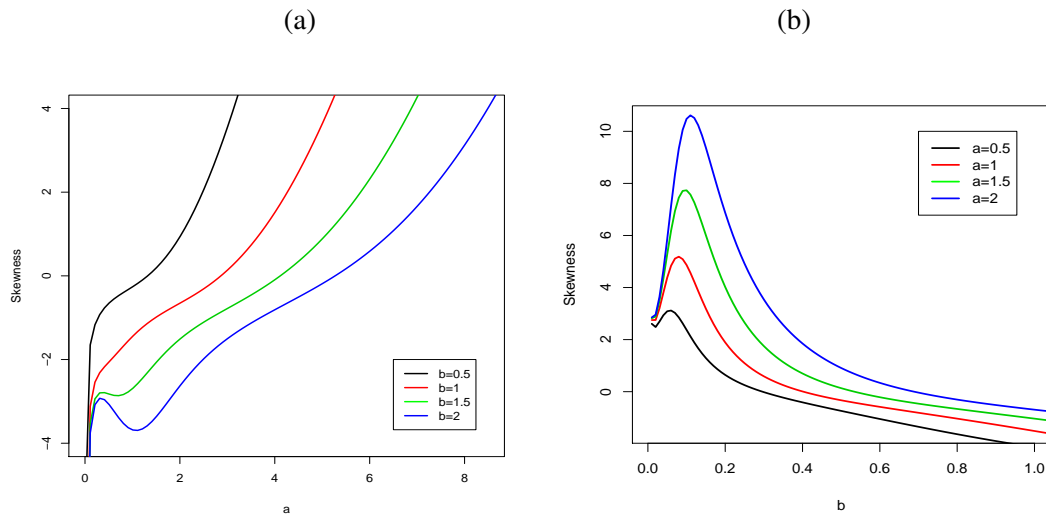
Fig. 3. Skewness of the LBGHN distribution. (a) Function of *a* for some values of *b*. (b) Function of *b* for some values of *a*.
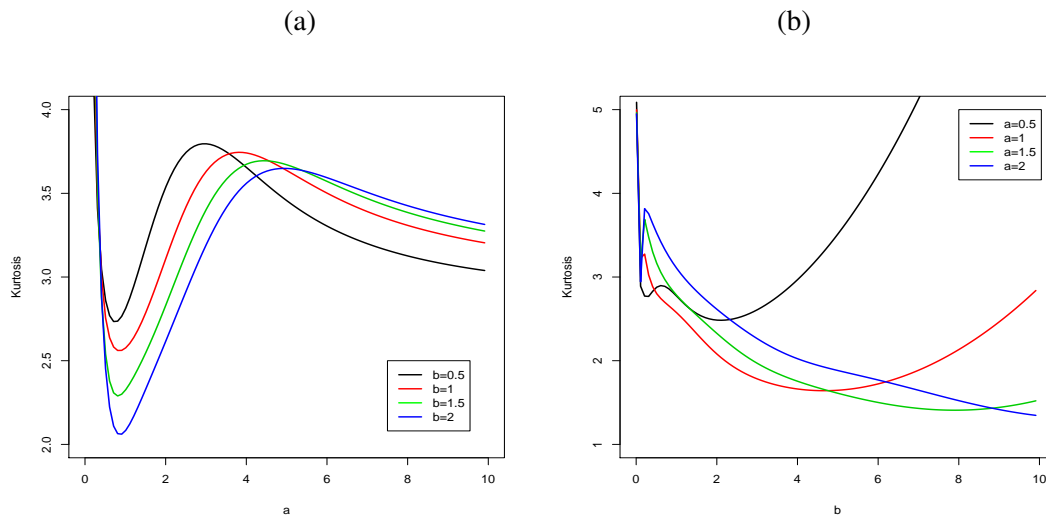

Fig. 4. Kurtosis of the LBGHN distribution. (a) Function of *a* for some values of *b*. (b) Function of *b* for some values of *a*..

4. Characterization results

The problem of characterizing a distribution is an important problem which has recently attracted the attention of many researchers. Thus, various characterizations have been established in many different directions. In practice, an investigator will be vitally interested to know if their model fits the requirements of their proposed distribution. To this end, the investigator relies on the characterizations of the distribution which provide conditions under which the underlying distribution is indeed the proposed distribution.

Here, we present, without loss of generality, characterizations of the $LBGHN(0,1,a,b)$ distribution ,with pdf given by (6), in terms of: (i) a simple relationship between two truncated moments; (ii) truncated moment of certain functions of the nth order statistic.

Due to the format of cdf of $LBGHN(0,1,a,b)$, we believe characterizations in other directions may not be possible or if possible will be quite complicated.

4.1. *Characterizations of $LBGHN(0,1,a,b)$ based on the ratio of two truncated moments*

Our first set of characterizations will employ an interesting result due to Glänzel (1987) (Theorem 3 below).

Theorem 3. Let $(\Omega, \mathfrak{F}, \mathbf{P})$ be a given probability space and let $H = [a,b]$ be an interval for some $a < b$ $(a = -\infty, b = \infty$ might as well be allowed$)$. Let $X : \Omega \to H$ be a continuous random variable with the distribution function F and let g and h be two real functions defined on H such that

$$E[g(X) \mid X \geq x] = E[h(X) \mid X \geq x] \eta(x), \quad x \in H,$$

is defined with some real function η. Assume that g, $h \in C^1(H)$, $\eta \in C^2(H)$ and F is twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $h\eta = g$ has no real solution in the interior of H. Then, F is uniquely determined by the functions q, h and η, particularly

$$F(x) = C \int_a^x \left| \frac{\eta'(u)}{\eta(u)h(u) - g(u)} \right| \exp[-s(u)] \, du,$$

where the function s is a solution of the differential equation $s' = \frac{\eta' h}{\eta h - g}$ and C is a constant chosen to make $\int_H dF = 1$.

We like to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence, in particular, let us assume that there is a sequence $\{X_n\}$ of random variables with distribution functions $\{F_n\}$ such that the functions g_n, h_n and η_n $(n \in \mathbb{N})$ satisfy the conditions of Theorem 3 and let $g_n \to g$, $h_n \to h$ for some continuously differentiable real functions g and h. Let X be a random variable with distribution F. Under the condition that $g_n(X)$ and $h_n(X)$ are uniformly integrable and that the family is relatively compact, the sequence X_n converges to X in distribution if and only if η_n converges to η, where

$$\eta(x) = \frac{E[g(X) \mid X \geq x]}{E[h(X) \mid X \geq x]}.$$

This stability theorem makes sure that the convergence of distribution functions is reflected by corresponding convergence of the functions g, h and η, respectively. It guarantees, for instance, the "convergence" of characterization of the Wald distribution to that of the Lévy-Smirnov distribution if $\alpha \to \infty$.

A further consequence of the stability property of Theorem 3 is the application of this theorem to special tasks in statistical practice such as the estimation of the parameters of discrete distributions. For such purpose, the functions g, h and, specially, η should be as simple as possible. Since

the function triplet is not uniquely determined it is often possible to choose η as a linear function. Therefore, it is worth analyzing some special cases which helps to find new characterizations reflecting the relationship between individual continuous univariate distributions and appropriate in other areas of statistics.

Remark 1. (a) In Theorem 3, the interval H need not be closed. (b) The goal is to have the function η as simple as possible. For a more detailed discussion on the choice of η, we refer the reader to Glänzel and Hamedani (2001) and Hamedani (2006 and 2010).

Proposition 1. Let $X : \Omega \to \mathbb{R}$ be a continuous random variable and let

$$h(x) = \left\{ 2\Phi\left[\exp\left(x\sqrt{2}/2\right)\right] - 1 \right\}^{1-a} \quad \text{and} \quad g(x) = h(x)\left\{ 1 - \Phi\left[\exp\left(x\sqrt{2}/2\right)\right] \right\}$$

for $x \in \mathbb{R}$. The pdf of X is (6) if and only if the function η defined in Theorem 3 has the form

$$\eta(x) = \frac{b}{b+1}\left\{ 1 - \Phi\left[\exp\left(x\sqrt{2}/2\right)\right] \right\} , \quad x \in \mathbb{R}.$$

Proof:

Let X have pdf (6). Then,

$$[1 - F(x)] \, \mathbf{E}[h(X) \mid X \geq x] = \frac{1}{b2^{-b}B(a,b)}\left\{ 1 - \Phi\left[\exp\left(x\sqrt{2}/2\right)\right] \right\}^{b},$$

and

$$[1 - F(x)] \, \mathbf{E}[g(X) \mid X \geq x] = \frac{1}{(b+1)\,2^{-b}B(a,b)}\left\{ 1 - \Phi\left[\exp\left(x\sqrt{2}/2\right)\right] \right\}^{b+1},$$

where F is the cdf corresponding to the pdf f.

Finally,

$$\eta(x)h(x) - g(x) = -\frac{1}{b+1}\left\{ 1 - \Phi\left[\exp\left(x\sqrt{2}/2\right)\right] \right\}\left\{ 2\Phi\left[\exp\left(x\sqrt{2}/2\right)\right] - 1 \right\}^{1-a} < 0.$$

Conversely, if η is given as above, then

$$s'(x) = \frac{\eta'(x)\,h(x)}{\eta(x)\,h(x) - g(x)} = b\left\{ \frac{d}{dx}\Phi\left[\exp\left(x\sqrt{2}/2\right)\right] \right\}\left\{ 1 - \Phi\left[\exp\left(x\sqrt{2}/2\right)\right] \right\}^{-1},$$

and hence $s(x) = \log\left\{ 1 - \Phi\left[\exp\left(x\sqrt{2}/2\right)\right] \right\}^{-b}$, $x \in \mathbb{R}$. Now, in view of Theorem 3 (with C chosen appropriately), X has corresponding pdf (6).∎

Remark 2. Clearly, there are other triplets (h, g, η) satisfying the conditions of Proposition 1.

Corollary 1. Let $X : \Omega \to \mathbb{R}$ be a continuous random variable and let

$$h(x) = \left\{ 2\Phi\left[\exp\left(x\sqrt{2}/2\right)\right] - 1 \right\}^{1-a}$$

for $x \in \mathbb{R}$. The pdf of X is (6) if and only if there exist functions g and η defined in Theorem 3 satisfying the differential equation

$$\frac{\eta'(x)\left\{2\Phi\left[\exp\left(x\sqrt{2}/2\right)\right]-1\right\}^{1-a}}{\eta(x)\left\{2\Phi\left[\exp\left(x\sqrt{2}/2\right)\right]-1\right\}^{1-a}-g(x)}$$

$$=\frac{b}{2\sqrt{\pi}}\exp\left\{\frac{1}{2}\left[-\exp\left(x\sqrt{2}\right)+x\sqrt{2}\right]\right\}\left\{1-\Phi\left[\exp\left(x\sqrt{2}/2\right)\right]\right\}^{-1},\quad x\in\mathbb{R}.$$

Remark 3. The general solution of the differential equation given in Corollary 1 is

$$\eta(x)=\left\{1-\Phi\left[\exp\left(x\sqrt{2}/2\right)\right]\right\}^{-1}\times$$
$$\left[\begin{array}{c}-\int g(x)\,\frac{b}{(b+1)\sqrt{\pi}}\exp\left\{\frac{1}{2}\left[-\exp\left(x\sqrt{2}\right)+x\sqrt{2}\right]\right\}\times\\\left\{2\Phi\left[\exp\left(x\sqrt{2}/2\right)\right]-1\right\}^{a-1}dx+D\end{array}\right],$$

for $x\in\mathbb{R}$, where D is a constant. One set of appropriate functions is given in Proposition 1 with $D=0$.

4.2. *Characterization of* $LBGHN(0,1,a,b)$ *based on truncated moment of certain functions of the nth order statistic*

Let $X_{1:n}\leq X_{2:n}\leq ...\leq X_{n:n}$ be n order statistics from a continuous cdf F. We state here a characterization result based on certain functions of the nth order statistic. Our characterization of $LBGHN(0,1,a,b)$ here will be a consequence of the following proposition, which is similar to the one appeared in our previous work (Hamedani 2010).

Proposition 2. Let $X:\Omega\to(0,\infty)$ be a continuous random variable with cdf F. Let $\psi(x)$ and $q(x)$ be two differentiable functions on $(0,\infty)$ such that $\lim_{x\to0}\psi(x)\left[F(x)\right]^n=0$ and $\int_0^\infty\frac{q'(t)}{\left[\psi(t)-q(t)\right]}dt=\infty$.
Then

$$E\left[\psi(X_{n:n})\,|\,X_{n:n}<t\right]=q(t),\quad t>0,\tag{4.1}$$

implies

$$F(x)=\exp\left\{-\int_x^\infty\frac{q'(t)}{n\left[\psi(t)-q(t)\right]}dt\right\},\quad x\geq0.\tag{4.2}$$

Remarks 4. (c) Taking, e.g., $\psi(x)=\left(I_{2\Phi\left[\left(\frac{x}{\theta}\right)^\alpha\right]-1}(a,b)\right)^n$ and $q(x)=\frac{1}{2}\psi(x)$ in Proposition 2, (4.2) will reduce to the cdf F corresponding to the pdf (6). (d) For $b=1$, we may take $\psi(x)=\left(2\Phi\left[\left(\frac{x}{\theta}\right)^\alpha\right]-1\right)^n$ and $q(x)=\frac{1}{2}\psi(x)$.

5. The log-beta generalized half-normal regression model

In many practical applications, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure, weight and many others. Parametric models to estimate univariate survival functions and for censored data regression problems are widely used. A parametric model

that provides a good fit to lifetime data tends to yield more precise estimates of the quantities of interest. Based on the LBGHN density, we propose a linear location-scale regression model or log-linear regression model linking the response variable y_i and the explanatory variable vector $\mathbf{x}_i^T = (x_{i1}, \ldots, x_{ip})$ as follows

$$y_i = \mathbf{x}_i^T \beta + \sigma z_i, \ i = 1, \ldots, n, \tag{5.1}$$

where the random error z_i has density function (2.6), $\beta = (\beta_1, \ldots, \beta_p)^T$, $\sigma > 0$, $a > 0$ and $b > 0$ are unknown parameters. The parameter $\mu_i = \mathbf{x}_i^T \beta$ is the location of y_i. The location parameter vector $\mu = (\mu_1, \ldots, \mu_n)^T$ is represented by a linear model $\mu = \mathbf{X}\beta$, where $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$ is a known model matrix. The LBGHN model (5.1) opens new possibilities for fitting many different types of data. This model is referred to as the LBGHN regression model for censored data. It is an extension of an accelerated failure time model using the BGHN distribution for censored data.

It contains as special sub-models the following well-known regression models:

- **Log-beta half normal (LBHN) regression model (new)**
  For $\sigma = \sqrt{2}/2$, the survival function is

$$S(y|\mathbf{x}) = 1 - I_{2\Phi\left[\exp\left(y - \mathbf{x}^T\beta\right)\right] - 1}(a, b).$$

  Note that the LBHN regression model is a new model. If $a = b = 1$ in addition to $\sigma = \sqrt{2}/2$, it reduces to the new log-half normal (LHN) regression model.
- **Log-exponentiated generalized half-normal (LEGHN) regression model (new)**
  For $b = 1$, the survival function is

$$S(y|\mathbf{x}) = 1 - \frac{1}{a}\left\{2\Phi\left[\exp\left\{\left(\frac{y - \mathbf{x}^T\beta}{\sigma}\right)\frac{\sqrt{2}}{2}\right\}\right] - 1\right\}^a,$$

  which is the new LEGHN regression model.
- **Log-generalized half-normal (LGHN) distribution (new)**
  For $a = b = 1$, the survival function becomes

$$S(y|\mathbf{x}) = 2 - 2\Phi\left\{\exp\left[\left(\frac{y - \mathbf{x}^T\beta}{\sigma}\right)\frac{\sqrt{2}}{2}\right]\right\},$$

  which is the new LGHN regression model.

Consider a sample $(y_1, \mathbf{x}_1), \ldots, (y_n, \mathbf{x}_n)$ of n independent observations, where each random response is defined by $y_i = \min\{\log(t_i), \log(c_i)\}$. We assume non-informative censoring such that the observed lifetimes and censoring times are independent. Let F and C be the sets of individuals for which y_i is the log-lifetime or log-censoring, respectively. Conventional likelihood estimation techniques can be applied here. The log-likelihood function for the vector of parameters $\gamma = (a, b, \sigma, \beta^T)^T$ from model (5.1) has the form $l(\gamma) = \sum_{i \in F} l_i(\gamma) + \sum_{i \in C} l_i^{(c)}(\gamma)$, where $l_i(\gamma) = \log[f(y_i|\mathbf{x}_i)]$, $l_i^{(c)}(\gamma) = \log[S(y_i|\mathbf{x}_i)]$, $f(y_i|\mathbf{x}_i)$ is the density (2.4) and $S(y_i|\mathbf{x}_i)$ is the survival function (2.5) of Y_i.

The total log-likelihood function for γ reduces to

$$
\begin{aligned}
l(\gamma) = {} & r\log\left[\frac{2^{b-1}}{\sigma\sqrt{\pi}B(a,b)}\right] - \frac{1}{2}\sum_{i\in F}\exp\left(z_i\sqrt{2}\right) + \frac{\sqrt{2}}{2}\sum_{i\in F}z_i \\
& + (a-1)\sum_{i\in F}\log\left\{2\Phi\left[\exp\left(\frac{z_i\sqrt{2}}{2}\right)\right] - 1\right\} \\
& + (b-1)\sum_{i\in F}\log\left\{1 - \Phi\left[\exp\left(\frac{z_i\sqrt{2}}{2}\right)\right]\right\} + \sum_{i\in C}\log\left[1 - I_{G(y_i|\mathbf{x}_i)}(a,b)\right],
\end{aligned}
\tag{5.2}
$$

where $G(y_i|\mathbf{x}_i) = 2\Phi\left[\exp\left(z_i\sqrt{2}/2\right)\right] - 1$, $z_i = (y_i - \mathbf{x}_i^T\beta)/\sigma$ and r is the number of uncensored observations (failures). The maximum likelihood estimate (MLE) $\widehat{\gamma}$ of the vector of unknown parameters can be calculated by maximizing the log-likelihood (5.2). We use the subroutine NLMixed in SAS to calculate the estimate $\widehat{\gamma}$. Initial values for β and σ are taken from the fit of the LGHN regression model with $a = b = 1$. The fit of the LBGHN model produces the estimated survival function for y_i ($\hat{z}_i = (y_i - \mathbf{x}_i^T\hat{\beta})/\hat{\sigma}$) given by

$$
S(y_i;\hat{a},\hat{b},\hat{\sigma},\widehat{\beta}^T) = 1 - I_{\widehat{G}(y_i|\mathbf{x}_i)}(\hat{a},\hat{b}),
\tag{5.3}
$$

where

$$
\widehat{G}(y_i|\mathbf{x}_i) = 2\Phi\left\{\exp\left[\left(\frac{y_i - \mathbf{x}_i^T\widehat{\beta}}{\hat{\sigma}}\right)\frac{\sqrt{2}}{2}\right]\right\} - 1.
$$

Similarly, using the invariance property of the MLEs, we have

$$
S(t_i;\hat{a},\hat{b},\hat{\alpha},\widehat{\theta}) = 1 - I_{2\Phi\left[\left(\frac{t_i}{\hat{\theta}}\right)^{\hat{\alpha}}\right]-1}(\hat{a},\hat{b}),
\tag{5.4}
$$

where $\hat{\alpha} = \frac{\sqrt{2}}{2\hat{\sigma}}$ and $\hat{\theta} = \exp(-\mathbf{x}^T\hat{\beta})$.

Under conditions that are fulfilled for the parameter vector γ in the interior of the parameter space but not on the boundary, the asymptotic distribution of $\sqrt{n}(\widehat{\gamma} - \gamma)$ is multivariate normal $N_{p+3}(0, K(\gamma)^{-1})$, where $K(\gamma)$ is the information matrix. The asymptotic covariance matrix $K(\gamma)^{-1}$ of $\widehat{\gamma}$ can be approximated by the inverse of the $(p+3) \times (p+3)$ observed information matrix $-\ddot{\mathbf{L}}(\gamma)$. The elements of the observed information matrix $-\ddot{\mathbf{L}}(\gamma)$, are are calculated numerically. The approximate multivariate normal distribution $N_{p+3}(0, -\ddot{\mathbf{L}}(\gamma)^{-1})$ for $\widehat{\gamma}$ can be used in the classical way to construct approximate confidence regions for some parameters in γ.

We can use the likelihood ratio (LR) statistic for comparing some special sub-models with the LBGHN model. We consider the partition $\gamma = (\gamma_1^T, \gamma_2^T)^T$, where γ_1 is a subset of parameters of interest and γ_2 is a subset of the remaining parameters. The LR statistic for testing the null hypothesis $H_0 : \gamma_1 = \gamma_1^{(0)}$ versus the alternative hypothesis $H_1 : \gamma_1 \neq \gamma_1^{(0)}$ is given by $w = 2\{\ell(\widehat{\gamma}) - \ell(\widetilde{\gamma})\}$, where $\widetilde{\gamma}$ and $\widehat{\gamma}$ are the estimates under the null and alternative hypotheses, respectively. The statistic w is asymptotically (as $n \to \infty$) distributed as χ_k^2, where k is the dimension of the subset of parameters γ_1 of interest.

6. Application- Voltage data

In this section, we illustrate the usefulness of the BGHN and LBGHN distributions with one application. Lawless (2003) reports an experiment in which specimens of solid epoxy electrical-insulation

were studied in an accelerated voltage life test. The sample size is $n = 60$, the percentage of censored observations was 10% and are considered three levels of voltage $52.5, 55.0$ and 57.5. The variables involved in the study are: t_i - failure times for epoxy insulation specimens (in min); $cens_i$ - censoring indicator (0=censoring, 1=lifetime observed); x_{i1} - voltage (kV).

We started the analysis of data considering only failure (t_i) and censoring ($cens_i$) data. An appropriate model for fitting such data could be the BGHN distribution. Table 1 gives the MLEs (and the corresponding standard errors in parentheses) of the model parameters and the values of the following statistics for some models: AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and CAIC (Consistent Akaike Information Criterion). The computations were done using the subroutine NLMixed in SAS. These results indicate that the BGHN model has the lowest AIC, BIC and CAIC values among those values of the fitted models, and therefore it could be chosen as the best model. Note that α_1 and γ are the parameters of the Weibull distribution.

Table 1. Estimates of the model parameters for the voltage data, the corresponding SEs (given in parentheses) and the statistics AIC, CAIC and BIC.

Model	α	θ	a	b	AIC	CAIC	BIC
BGHN	0.099	1020.98	59.443	30.219	839.1	839.8	847.5
	(0.0097)	(289.73)	(1.406)	(0.077)			
GHN	0.622	1030.0	1	1	861.3	861.5	865.5
	(0.066)	(151.07)	(-)	(-)			
HN	1	1776.36	1	1	870.0	870.1	872.1
	(-)	(175.04)	(-)	(-)			
	α_1	γ					
Weibull	1132.0	0.947			873.3	873.5	877.5
	(167.83)	(0.094)					

Formal tests for other sub-models of the BGHN distribution are conducted using LR statistics as described before. Applying these statistics to the voltage data, the results are given in Table 2. Clearly, we reject the null hypotheses for the three LR tests in favor of the BGHN distribution.

Table 2. LR tests.

Voltage	Hypotheses	Statistic w	p-value
BGHN vs GHN	$H_0 : a = b = 1$ vs $H_1 : H_0$ is false	26.20	< 0.0001
BGHN vs HN	$H_0 : a = b = \alpha = 1$ vs $H_1 : H_0$ is false	36.90	<0.0001

In order to assess if the model is appropriate, Figure 5a plots the empirical survival function and the estimated survival function of the BGHN distribution which shows a good fit to these data.

Now, we present results by fitting the model

$$y_i = \beta_0 + \beta_1 x_{i1} + \sigma z_i,$$

where the random variable Y_i follows the LBGHN distribution (2.4) for $i = 1, \ldots, 60$. The MLEs of the model parameters are calculated using the procedure NLMixed in SAS. Iterative maximization

of the logarithm of the likelihood function (5.2) starts with initial values for β and σ which are taken from the fit of the LGHN regression model with $a = b = 1$.

Table 3. MLEs of the parameters from the LBGHN regression model fitted to the voltage data set, the corresponding SEs (given in parentheses), p-value in [.] and the statistics AIC, CAIC and BIC.

Model	σ	β_0	β_1	a	b	AIC	CAIC	BIC
LBGHN	5.306	10.632	-0.201	102.14	1.564	167.1	168.2	177.5
	(0.666)	(3.304)	(0.056)	(3.989)	(0.672)			
		[0.0021]	[0.0007]					
LGHN	0.778	23.637	-0.301	1	1	178.8	179.2	185.1
	(0.089)	(2.928)	(0.053)	(-)	(-)			
		[< 0.001]	[< 0.001]					
LHN	$\sqrt{2}/2$	24.0607	-0.3079	1	1	177.6	177.8	181.7
	(-)	(2.6276)	(0.0478)	(-)	(-)			
		[< 0.001]	[< 0.001]					
Log-Weibull	0.8454	22.032	-0.275			173.4	173.8	179.7
	(0.090)	(3.046)	(0.055)					
		[< 0.001]	[< 0.001]					

We note from the fitted LBGHN regression model that x_1 is significant at 1% and that there is a significant difference between the voltages 52.5, 55.0 and 57.5 for the survival times.

Formal tests for other sub-models of the LBGHN distribution are conducted using LR statistics as described before. Applying these statistics to the voltage data, the results are given in Table 4. Clearly, we reject the null hypotheses for the three LR tests in favor of the LBGHN distribution.

Table 4. LR tests.

Voltage	Hypotheses	Statistic w	p-value
LBGHN vs LGHN	$H_0 : a = b = 1$ vs $H_1 : H_0$ is false	15.70	0.0004
LBGHN vs LHN	$H_0 : a = b = 1$ and $\sigma = \sqrt{2}/2$ vs $H_1 : H_0$ is false	16.10	<0.0011

In order to assess if the model is appropriate, Figure 5b plots the empirical survival function and the estimated survival function given by (5.4) from the fitted LBGHN regression model. We conclude that the LBGHN regression model provides a good fit to these data.

7. Concluding Remarks

We introduce the so-called log-beta generalized half-normal (LBGHN) distribution whose hazard rate function accommodates four types of shape forms, namely increasing, decreasing, bathtub and unimodal. We derive expansions for the moments and moment generating function. Some important properties are addressed. Based on this new distribution, we propose a LBGHN regression model which is very suitable for modeling censored and uncensored lifetime data. The new regression model allows testing the goodness of fit of some known regression models as special sub-models. Hence, the proposed regression model serves as a good alternative for lifetime data analysis. Further,

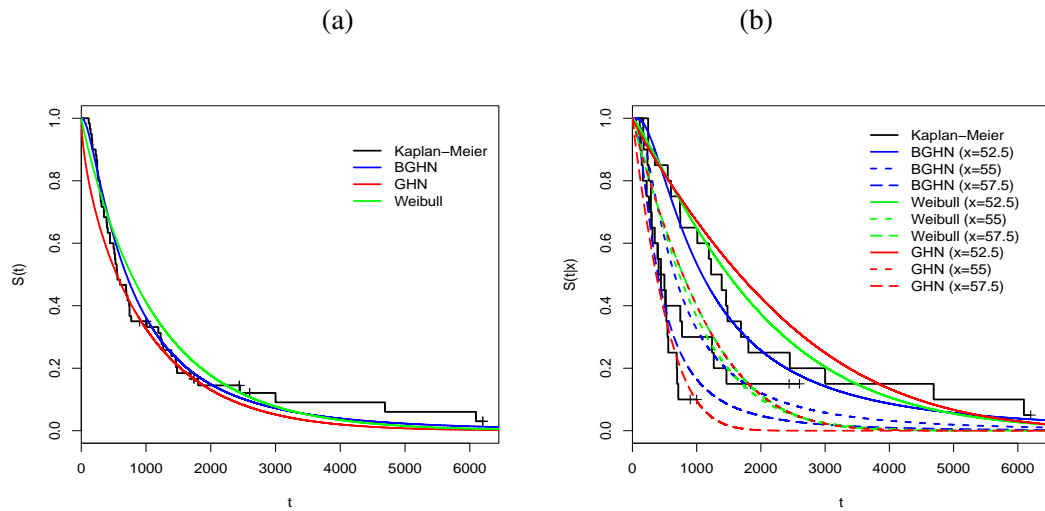(a)   (b)

Fig. 5. (a) Estimated survival functions and the empirical survival for voltage data. (b) Estimated survival functions for the BGHN distribution and some of its sub-models and the empirical survival for stress level data.

the new regression model is much more flexible than the generalized half normal, exponentiated generalized half normal and half normal sub-models. We use the procedure NLMixed (SAS) to obtain the maximum likelihood estimates and perform asymptotic tests for the parameters based on the asymptotic distribution of these estimates. We demonstrate in one application to real data that the LBGHN model can produce better fit than its sub-models.

Acknowledgment

References

[1] Barros, M., Paula, G. A., Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis,* **14**, 316-332.

[2] Carrasco, J. M. F., Ortega, E. M. M., Paula, G. A. (2008). Log-modified Weibull regression models with censored data: sensitivity and residual analysis. *Computational Statistics and Data Analysis,* **52**, 4021-4029.

[3] Cooray, K., Ananda, M.M.A. (2008). A generalization of the half-normal distribution with applications to lifetime data. *Communication in Statistics - Theory and Methods,* **37**, 1323-1337.

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

[5] Glänzel, W. (1987) A characterization theorem based on truncated moments and its application to some distribution families. *Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986), Vol. B, Reidel, Dordrecht*, 75-84.

[6] Glänzel, W., Hamedani, G.G. (2001). Characterizations of univariate continuous distributions. *Studia Scientiarum Mathematicarum Hungarica,* **37**, 83-118.

[7] Hamedani, G.G. (2006). Characterizations of univariate continuous distributions III. Studia Scientiarum Mathematicarum Hungarica, **43** , 361-385.

[8] Hamedani, G.G. (2010). Characterizations of continuous univariate distributions based on the truncated moments of functions of order statistics. *Studia Scientiarum Mathematicarum Hungarica*, **47**, 462-484.

[9] Hashimoto, E. M., Ortega, E. M. M., Cancho, V. G., Cordeiro, G. M. (2010). The log-exponentiated Weibull regression model for interval-censored data. *Computational Statistics and Data Analysis,* **54**, 1017-1035.

[10] Jones, M. C. (2004). Family of distribution of order statistics. *Test,* **13**, 1-43.

[11] Lawless, J. F.(2003). *Statistical Models and Methods for Lifetime Data*. Wiley: New York.

[12] Ortega, E. M. M., Cancho, V. G., Paula G. A. (2009). Generalized log-gamma regression models with cure fraction. *Lifetime Data Analysis,* **15**, 79-106.

[13] Pescim, R. R., Demtrio, 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.

[14] Silva, G. O., Ortega, E. M. M., Garibay, V. C., Barreto, M. L. (2008). Log-Burr XII regression models with censored Data. *Computational Statistics and Data Analysis,* **52**, 3820-3842.

[15] Silva, G. O., Ortega, E. M. M., Cordeiro, G. M. (2009). A log-extended Weibull regression model. *Computational Statistics and Data Analysis,* **12**, 4482-4489.