

# On the Gumbel-Burr XII Distribution: Regression and Application

Raid Al-Aqtash<sup>1</sup>, Avishek Mallick<sup>1</sup>, G.G. Hamedani<sup>2</sup> & Mahmoud Aldeni<sup>3</sup>

<sup>1</sup> Department of Mathematics, Marshall University, West Virginia, USA

<sup>2</sup> Department of Mathematical and Statistical Sciences, Marquette University, Wisconsin, USA

<sup>3</sup> Department of Mathematics and Computer Science, Western Carolina University, North Carolina, USA

Correspondence: Raid Al-Aqtash, Department of Mathematics, Marshall University, West Virginia, USA

Received: September 7, 2021 Accepted: October 20, 2021 Online Published: October 26, 2021

doi:10.5539/ijsp.v10n6p31 URL: <https://doi.org/10.5539/ijsp.v10n6p31>

## Abstract

In this article, additional properties of the Gumbel-Burr XII distribution, denoted by (GBXII(L)), defined in (Osatohanmwen et al., 2017), are studied. We consider some useful characterizations for the GBXII(L) distribution and some of its properties. A simulation study is conducted to assess the performance of the MLEs and the usefulness of the GBXII(L) distribution is illustrated by means of three real data sets. The simulation study suggests that the maximum likelihood method can be used to estimate the distribution parameters, and the three examples show that the GBXII(L) is very flexible in fitting different shapes of data. A log-GBXII(L) regression model is proposed and a survival data is used in an application of the proposed regression model. The log-GBXII(L) regression model is adequate and can be used in comparison to other models.

**Keywords:** beta-generated, generalized class,  $T$ - $X(L)$  framework, regression, survival time, censored data

## 1. Introduction

Statistical distributions can be used to model many real life scenarios, such as reliability, actuarial science, survival analysis and lifetime (Tahir and Cordeiro, 2016). Seeking flexibility in modeling real life phenomena has been a strong reason to develop new statistical distributions. Many distributions have been developed by adding new shape, scale or location parameters.

Eugene et al. (2002) introduced the beta-generated class of distributions with cumulative distribution function (CDF)  $G(x) = \int_0^{F(x)} \frac{1}{B(\alpha, \beta)} t^{\alpha-1} (1-t)^{\beta-1} dt$ ,  $\alpha > 0, \beta > 0$ , where  $F(x)$  is the CDF of any continuous random variable  $X$ . Many distributions were developed utilizing this technique, such as the beta-normal (Eugene et al., 2002), beta-Weibull (Famoye et al., 2005), beta-generalized exponential (Barreto-Souza et al., 2010), beta-Cauchy (Alshawarbeh et al., 2012) and beta-Pareto (Akinsete et al., 2008).

Jones (2009) and Cordeiro and de Castro (2011) proposed the Kumaraswamy-generated distributions (KwG) with CDF  $F(x) = \int_0^{G(x)} \alpha \beta t^{\alpha-1} (1-t^\alpha)^{\beta-1} dt$ ,  $\alpha > 0, \beta > 0$ . Distributions developed using the KwG technique include the Kumaraswamy-Weibull (Cordeiro et al., 2010), the Kumaraswamy-generalized half-normal (Cordeiro et al., 2012) and the Kumaraswamy-geometric (Akinsete et al., 2014).

Alzaatreh et al. (2013) proposed a transformer-transformed technique to generate a class of distributions  $T$ - $X(W)$  by using a transformation  $W(F(x))$  of the CDF  $F(x)$  of a random variable  $X$ , that satisfies the conditions:

$$\begin{cases} (i) & W[F(x)] \in [a, b], \\ (ii) & W[F(x)] \text{ is differentiable and monotonically non-decreasing, and} \\ (iii) & W[F(x)] \rightarrow a \text{ as } x \rightarrow -\infty \text{ and } W[F(x)] \rightarrow b \text{ as } x \rightarrow \infty. \end{cases} \quad (1)$$

The class of  $T$ - $X(W)$  distributions is then defined by

$$G_{TX(W)}(x) = \int_a^{W(F_X(x))} r_T(t) dt, \quad (2)$$

where  $T$  is a random variable with probability density function (PDF)  $r_T(t)$  on  $(a, b)$ .

Aljarrah et al. (2014) used  $W(F(x)) = Q_Y(F(x))$ , the quantile of a random variable  $Y$ , as a transformation of a CDF  $F(x)$ . They defined the CDF of the  $T$ - $X\{Y\}$  family as

$$G_{TX\{Y\}}(x) = \int_a^{Q_Y(F_X(x))} r_T(t) dt. \quad (3)$$

Families of distributions developed using (2) and (3) include the Weibull- $G$  (Bourguignon, et al., 2014), the  $T$ -normal (Alzaatreh et al., 2014), the  $T$ -Weibull (Almheidat, et al., 2015) and the  $T$ - $X$ (Logit) (Al-Aqtash, 2013). For more on recent developments in distribution theory, we refer the interested reader to Lee et al. (2013) and Tahir and Cordeiro (2016).

Al-Aqtash (2013) and Al-Aqtash et al. (2015) utilized the logit of a CDF,  $W(F(x)) = Q_{logistic}(F(x)) = \ln\{F(x)/(1 - F(x))\}$  as a transformation in (2) and studied general properties such as symmetry, Shannon entropy and quantile function of the  $T$ - $X$ (Logit) family. Al-Aqtash et al. (2014) and Al-Aqtash et al. (2015) defined the Gumbel- $X$  (Logit) family and studied a member of the family, namely the Gumbel-Weibull distribution (GWD). The CDF for the Gumbel- $X$  (Logit) family is given by

$$G_{GX(logit)}(x) = \exp[-\beta\{F(x)/(1 - F(x))\}^{-1/\sigma}], \quad \beta > 0, \sigma > 0. \quad (4)$$

Osatohanmwun et al. (2017) used the CDF of the Burr XII distribution ( $F(x) = 1 - (1 + (x/\lambda)^c)^{-k}$ ) (Burr, 1942) in (4) to define the CDF of the five parameter GBXII(L) distribution

$$F_{GBXII(L)}(x) = \exp[-\beta\{(1 + (x/\lambda)^c)^k - 1\}^{-1/\sigma}], \quad x \geq 0; \beta, \sigma, c, k, \lambda > 0. \quad (5)$$

The PDF of the GBXII(L) distribution is given by

$$f_{GBXII(L)}(x) = \frac{\beta k c}{\sigma \lambda} (x/\lambda)^{c-1} (1 + (x/\lambda)^c)^{k-1} ((1 + (x/\lambda)^c)^k - 1)^{-1-1/\sigma} \exp[-\beta\{(1 + (x/\lambda)^c)^k - 1\}^{-1/\sigma}]. \quad (6)$$

Osatohanmwun et al. (2017) studied the shapes, hazard, Shannon entropy, moments and parameter estimation of the GBXII distribution. Figure 1 displays some of the different shapes of the GBXII density. In this article, we discuss additional properties of the GBXII distribution.

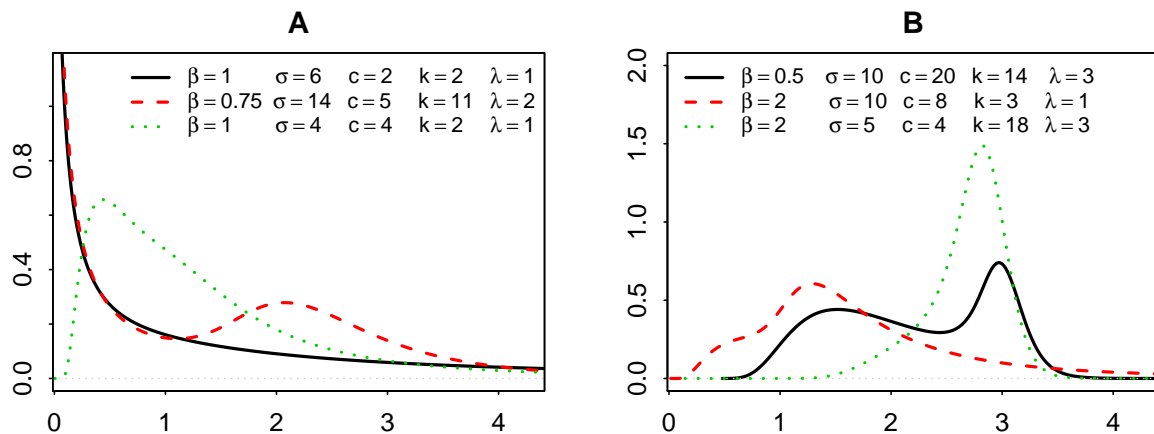


Figure 1. GBXII(L) density plot

The rest of the article is organized as follows: In section 2, we discuss some characterizations of the GBXII(L) model. In section 3, we present some additional properties of the GBXII(L) distribution. In section 4, the maximum likelihood method for parameter estimation is discussed and simulation results are presented to study and assess the performance of the maximum likelihood estimators. In section 5, three data sets are used to exhibit the flexibility of this distribution. In section 6, a log GBXII regression model is proposed and a survival time censored data set is used in an application of the proposed regression model. Finally, some concluding remarks are provided in section 7.

## 2. Characterizations of GBXII(L) Distribution

This section deals with various characterizations of GBXII(L) distribution. These characterizations are presented in three directions:

- based on the ratio of two truncated moments,
- in terms of the reverse hazard function and
- based on the conditional expectation of certain function of the random variable.

It should be noted that characterization based on truncated moments can be employed even when the CDF does not have a closed form.

### 2.1 Characterization Based on Truncated Moments

Our first characterization employs a theorem due to Glänzel (1987), see Theorem A1 of Appendix A. The result, however, holds also when the interval  $H$  is not closed since the condition of Theorem A1 is on the interior of  $H$ . We like to mention that this kind of characterization based on a truncated moment is stable in the sense of weak convergence (Glänzel, 1990).

**Proposition 1.** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable and let

$q_2(x) = \exp \left[ \beta \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}^{-1/\sigma} \right]$  and  $q_1(x) = q_2(x) \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}^{-1/\sigma}$  for  $x > 0$ . The random variable  $X$  belongs to the family (6) if and only if the function  $\eta$  defined in Theorem A1 has the form

$$\eta(x) = 2 \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}^{1/\sigma}, \quad x > 0.$$

*Proof:* Assume  $X$  is a random variable with PDF (6), then

$$(1 - F(x)) E[q_1(X) | X \geq x] = \frac{\beta}{2} \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}^{-2/\sigma}$$

and

$$(1 - F(x)) E[q_2(X) | X \geq x] = \beta \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}^{-1/\sigma}.$$

Further,

$$\eta(x) q_1(x) - q_2(x) = q_2(x) > 0 \quad \text{for } x > 0.$$

Conversely, if  $\eta$  is given as above, then

$$s'(x) = \frac{\eta'(x) q_1(x)}{\eta(x) q_1(x) - q_2(x)} = \frac{2kc \left( \frac{x}{\lambda} \right)^{c-1} \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^{k-1}}{\lambda \sigma \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}}, \quad x > 0,$$

and hence

$$s(x) = \frac{2}{\sigma} \log \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}, \quad x > 0.$$

Now, according to Theorem A1, random variable  $X$  has density (6).

**Corollary 1.** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable and let  $q_2$  be as in Proposition 1. Then,  $X$  has PDF (6) if and only if there exist functions  $q_1$  and  $\eta$  defined in Theorem A1 satisfying the differential equation

$$\frac{\eta'(x) q_1(x)}{\eta(x) q_1(x) - q_2(x)} = \frac{2kc \left( \frac{x}{\lambda} \right)^{c-1} \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^{k-1}}{\lambda \sigma \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}}, \quad x > 0.$$

The general solution of the differential equation in Corollary 1 is

$$\eta(x) = \left\{ \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^k - 1 \right\}^{-1} \left[ - \int 2kc \left( \frac{x}{\lambda} \right)^{c-1} \left( 1 + \left( \frac{x}{\lambda} \right)^c \right)^{k-1} (q_1(x))^{-1} q_2(x) dx + D \right],$$

where  $D$  is a constant. Note that a set of functions satisfying the above differential equation is given in Proposition 1 with  $D = 0$ .

**Remark 1.** For  $k = 1$ ,  $q_2(x) = \exp \left[ \beta \left( \frac{x}{\lambda} \right)^{c/\sigma} \right]$  and  $q_1(x) = q_2(x) x^{-c/\sigma}$ , we have  $\eta(x) = 2x^{c/\sigma}$ ,  $x > 0$ ,  $s'(x) = \frac{c}{\sigma} x^{-1}$ ,  $x > 0$  and

$$\eta(x) = x \left[ - \int \frac{c}{\sigma} (q_1(x))^{-1} q_2(x) dx + D \right].$$

## 2.2 Characterization in Terms of the Reverse Hazard Function

The reverse hazard function,  $r_F$ , of a twice differentiable distribution function,  $F$ , is defined as

$$r_F(x) = \frac{f(x)}{F(x)}, \quad x \in \text{support of } F.$$

In this subsection we present characterization of GBXII(L) distribution in terms of the reverse hazard function.

**Proposition 2.** Let  $X : \Omega \rightarrow (0, \infty)$  be a continuous random variable. Then,  $X$  has PDF (6) if and only if its reverse hazard function  $r_F(x)$  satisfies the differential equation

$$r'_F(x) - (c-1)x^{-1}r_F(x) = \frac{\beta kc}{\sigma \lambda^c} x^{c-1} \frac{d}{dx} \left\{ \frac{\left(1 + \left(\frac{x}{\lambda}\right)^c\right)^{k-1}}{\left[\left(1 + \left(\frac{x}{\lambda}\right)^c\right)^k - 1\right]^{\frac{1}{\sigma}+1}} \right\}, \quad x > 0.$$

*Proof:* If  $X$  has PDF (6), then clearly the differential equation holds. Now, if the differential equation holds, then

$$\frac{d}{dx} \left\{ x^{-(c-1)} r_F(x) \right\} = \frac{\beta kc}{\sigma \lambda^c} \frac{d}{dx} \left\{ \frac{\left(1 + \left(\frac{x}{\lambda}\right)^c\right)^{k-1}}{\left[\left(1 + \left(\frac{x}{\lambda}\right)^c\right)^k - 1\right]^{\frac{1}{\sigma}+1}} \right\},$$

from which we arrive at the reverse hazard function of (6).

**Remark 2.** For  $k = 1$ , we have the following simple differential equation

$$r'_F(x) + \left(\frac{c+\sigma}{\sigma}\right)x^{-1}r_F(x) = 0.$$

## 2.3 Characterization Based on the Conditional Expectation of Certain Function of the Random Variable

In this subsection we employ a single function  $\psi$  of  $X$  and characterize the distribution of  $X$  in terms of the conditional expectation of  $\psi$ . The following proposition has already appeared in Hamedani(2013), so we will just state it here which can be used to characterize GBXII(L) distribution.

**Proposition 3.** Let  $X : \Omega \rightarrow (a, b)$  be a continuous random variable with CDF  $F$ . Let  $\psi(x)$  be a differentiable function on  $(a, b)$  with  $\lim_{x \rightarrow b^-} \psi(x) = 1$ . Then for  $\delta \neq 1$ ,

$$E[\psi(X) | X \leq x] = \delta \psi(x), \quad x \in (a, b),$$

implies

$$\psi(x) = (F(x))^{\frac{1}{\delta}-1}, \quad x \in (a, b).$$

**Remarks 3.**

- A. For  $(a, b) = (0, \infty)$ ,  $\psi(x) = \exp\left[-\left\{\left(1 + \left(\frac{x}{\lambda}\right)^c\right)^k - 1\right\}^{-1/\sigma}\right]$  and  $\delta = \frac{\beta}{1+\beta}$ , Proposition 3 provides a characterization of the GBXII(L) distribution.
- B. For  $k = 1$ ,  $\psi(x) = \exp\left[-x^{c/\sigma}\right]$  and  $\delta = \frac{\beta \lambda^{c/\sigma}}{1+\beta \lambda^{c/\sigma}}$ , Proposition 3 provides a characterization of this special submodel of the GBXII(L) model.

In the next section we present some structural properties of the GBXII(L) distribution.

## 3. Additional Properties of GBXII(L) Distribution

### 3.1 Mean Deviations

Two measures of dispersion for GBXII(L) are the mean absolute deviation from the mean and the mean absolute deviation from the median. Let  $X$  be a random variable from GBXII(L) with mean  $\mu$  and median  $M$ . The mean absolute deviation

from the mean is given by

$$\begin{aligned} E(|X - \mu|) &= \int_0^{\infty} |x - \mu| f_{GBXII(L)}(x) dx \\ &= 2 \int_0^{\mu} (\mu - x) f_{GBXII(L)}(x) dx + \int_0^{\infty} (x - \mu) f_{GBXII(L)}(x) dx \\ &= 2\mu F_{GBXII(L)}(\mu) - 2 \int_0^{\mu} x f_{GBXII(L)}(x) dx. \end{aligned} \quad (7)$$

The mean absolute deviation from the median is given by

$$\begin{aligned} E(|X - M|) &= \int_0^{\infty} |x - M| f_{GBXII(L)}(x) dx \\ &= 2 \int_0^M (M - x) f_{GBXII(L)}(x) dx + \int_0^{\infty} (x - M) f_{GBXII(L)}(x) dx \\ &= \mu - 2 \int_0^M x f_{GBXII(L)}(x) dx. \end{aligned} \quad (8)$$

The integral  $I_{1,(0,\tau)} = \int_0^{\tau} x f_{GBXII(L)}(x) dx$  in (7) and (8) can be calculated numerically.

### 3.2 Transformation

A special case of the GBXII(L) distribution, when  $c = 1$ , is the Gumbel-Lomax distribution, studied by Tahir et al. (2016). The connection between the GBXII(L) to the Gumbel, standard uniform, standard exponential and Weibull distributions were discussed in Osatohanmwun et al. (2017). The link between the GBXII(L) to the Weibull-Dagum distribution is given in the following lemma.

**Lemma 1.** If  $Y$  is a random variable following the Weibull-Dagum distribution (Tahir et al., 2016) with CDF  $F(y) = 1 - \exp((1 + \lambda_1 y^{-\delta_1})^{\beta_1} - 1)^{-b_1}$ ,  $y \geq 0$ ,  $b_1, \lambda_1, \delta_1, \beta_1 > 0$ , then the random variable  $X = 1/Y$  has a GBXII(L) distribution with parameters  $\beta = 1$ ,  $\sigma = 1/b_1$ ,  $c = \delta_1$ ,  $k = \beta_1$  and  $\lambda = \lambda_1^{-1/c}$ .

*Proof:* The results follow directly from the transformation technique.

### 3.3 Quantile Function

Osatohanmwun et al. (2017) derived the formulas for the quantile function and the median for GBXII(L), respectively, as

$$\begin{aligned} Q_{GBXII(L)}(p) &= \lambda \{ ((-\ln p)/\beta)^{-\sigma} + 1 \}^{1/k} - 1 \}^{1/c}, \quad 0 < p < 1. \\ M &= Q_{GBXII(L)}(0.5) = \lambda \{ ((0.69314718/\beta)^{-\sigma} + 1 \}^{1/k} - 1 \}^{1/c}. \end{aligned} \quad (9)$$

**Remarks 4.** It follows directly from (9) that the quantile function for GBXII(L) is

- an increasing function of  $\lambda$ , when all other parameters are held fixed,
- a decreasing function of  $k$ , when all other parameters are held fixed,
- an increasing function of  $\beta$ , when all other parameters are held fixed,
- decreasing, increasing, or constant function of  $\sigma$  when all other parameters are held fixed, if  $p < e^{-\beta}$ ,  $p > e^{-\beta}$ , or  $p = e^{-\beta}$  respectively.
- decreasing, increasing, or constant function of  $c$  when all other parameters are held fixed, if  $p > \exp(-\beta(2^k - 1)^{-1/\sigma})$ ,  $p < \exp(-\beta(2^k - 1)^{-1/\sigma})$ , or  $p = \exp(-\beta(2^k - 1)^{-1/\sigma})$  respectively.

## 4. Parameter Estimation and Simulation

According to Nadarajah and Okorie (2018), the log likelihood function for the GBXII(L) distribution reported in Osatohanmwun et al. (2017) is incorrect and hence might affect the accuracy of the simulation and estimation results. Nadarajah and Okorie (2018) provided the correct log likelihood function for the GBXII(L) which can be written as

$$\begin{aligned} \ell = \ell(\Theta) &= n(\ln \beta + \ln k + \ln c - \ln \sigma - \ln \lambda) + (c - 1) \sum_{i=1}^n \ln \left( \frac{x_i}{\lambda} \right) + (k - 1) \sum_{i=1}^n \ln \left( 1 + \left( \frac{x_i}{\lambda} \right)^c \right) \\ &\quad - \left( 1 + \frac{1}{\sigma} \right) \sum_{i=1}^n \ln \left\{ \left( 1 + \left( \frac{x_i}{\lambda} \right)^c \right)^k - 1 \right\} - \beta \sum_{i=1}^n \left\{ \left( 1 + \left( \frac{x_i}{\lambda} \right)^c \right)^k - 1 \right\}^{-1/\sigma}, \end{aligned}$$

where  $x_1, \dots, x_n$  are the observed values of GBXII(L) with parameters  $\beta, \sigma, c, k$  and  $\lambda$ .

The correct log likelihood function for the GBXII(L), provided by Nadarajah and Okorie (2018), is used in estimation of the unknown parameters of the GBXII(L) distribution, using the method of maximum likelihood.

The components of the score vector,  $\mathbf{U}(\Theta) = \frac{\partial \ell}{\partial \Theta} = \left( \frac{\partial \ell}{\partial \beta}, \frac{\partial \ell}{\partial \sigma}, \frac{\partial \ell}{\partial c}, \frac{\partial \ell}{\partial k}, \frac{\partial \ell}{\partial \lambda} \right)^T = (U_\beta, U_\sigma, U_c, U_k, U_\lambda)^T$ , are

$$U_\beta = \frac{n}{\beta} - \sum_{i=1}^n \left\{ \left( 1 + \left( \frac{x_i}{\lambda} \right)^c \right)^k - 1 \right\}^{-1/\sigma},$$

$$U_\sigma = -\frac{n}{\sigma} + \frac{1}{\sigma^2} \sum_{i=1}^n \ln \left\{ \left( 1 + \left( \frac{x_i}{\lambda} \right)^c \right)^k - 1 \right\} - \frac{\beta}{\sigma^2} \sum_{i=1}^n \left\{ \left( 1 + \left( \frac{x_i}{\lambda} \right)^c \right)^k - 1 \right\}^{-1/\sigma} \ln \left\{ \left( 1 + \left( \frac{x_i}{\lambda} \right)^c \right)^k - 1 \right\},$$

$$U_c = \frac{n}{c} + \sum_{i=1}^n \ln \left( \frac{x_i}{\lambda} \right) + (k-1) \sum_{i=1}^n \frac{(x_i/\lambda)^c \ln(x_i/\lambda)}{1 + (x_i/\lambda)^c} - k \left( 1 + \frac{1}{\sigma} \right) \sum_{i=1}^n \frac{(1 + (x_i/\lambda)^c)^{k-1} (x_i/\lambda)^c \ln(x_i/\lambda)}{(1 + (x_i/\lambda)^c)^k - 1} \\ + \frac{\beta k}{\sigma} \sum_{i=1}^n \left\{ (1 + (x_i/\lambda)^c)^k - 1 \right\}^{-1-1/\sigma} (1 + (x_i/\lambda)^c)^{k-1} (x_i/\lambda)^c \ln(x_i/\lambda),$$

$$U_k = \frac{n}{k} + \sum_{i=1}^n \ln \left( 1 + \left( \frac{x_i}{\lambda} \right)^c \right) - \left( 1 + \frac{1}{\sigma} \right) \sum_{i=1}^n \frac{(1 + (x_i/\lambda)^c)^k \ln(1 + (x_i/\lambda)^c)}{(1 + (x_i/\lambda)^c)^k - 1} \\ + \frac{\beta}{\sigma} \sum_{i=1}^n \left\{ (1 + (x_i/\lambda)^c)^k - 1 \right\}^{-1-1/\sigma} (1 + (x_i/\lambda)^c)^k \ln(1 + (x_i/\lambda)^c),$$

$$U_\lambda = -\frac{nc}{\lambda} - \frac{c(k-1)}{\lambda} \sum_{i=1}^n \frac{(x_i/\lambda)^c}{1 + (x_i/\lambda)^c} - \frac{kc}{\lambda} \left( 1 + \frac{1}{\sigma} \right) \sum_{i=1}^n \frac{(1 + (x_i/\lambda)^c)^{k-1} (x_i/\lambda)^{c-1}}{(1 + (x_i/\lambda)^c)^k - 1} \\ + \frac{\beta kc}{\sigma \lambda} \sum_{i=1}^n \left\{ (1 + (x_i/\lambda)^c)^k - 1 \right\}^{-1-1/\sigma} (1 + (x_i/\lambda)^c)^{k-1} (x_i/\lambda)^{c-1}.$$

Solving  $U_\beta = U_\sigma = U_c = U_k = U_\lambda = 0$  simultaneously yields the MLE  $\widehat{\Theta} = (\widehat{\beta}, \widehat{\sigma}, \widehat{c}, \widehat{k}, \widehat{\lambda})$ .

The initial values are calculated using the criteria reported in Osatohanmwen et al. (2017) as follows: Since  $c, k$  and  $\lambda$  are coming from Burr XII distribution (BXII), and  $\beta$  and  $\sigma$  are coming from Gumbel distribution (GD), we can assume that the random sample  $x_1, \dots, x_n$  is following BXII, and use the BXII MLEs  $\widetilde{c}, \widetilde{k}$ , and  $\widetilde{\lambda}$  as initial values  $c_0, k_0$  and  $\lambda_0$ . The next step is to transform the sample from GBXII(L) to GD using the transformation  $y_i = \ln((1 + (x_i/\widetilde{\lambda})^{\widetilde{c}})^{\widetilde{k}} - 1)$ . Then the initial values for the last two parameters are  $\beta_0 = e^{\nu_0/\sigma_0}$  and  $\sigma_0 = s_y \sqrt{6}/\pi$ , where  $\nu_0 = \bar{y} - \gamma\sigma_0$  and  $\sigma_0$  are the GD moment estimates,  $\gamma$  is the Euler's constant,  $\bar{y}$  and  $s_y$  are the mean and standard deviation of  $y_1, \dots, y_n$  respectively (Johnson et al., 1995, p. 12).

#### 4.1 Simulation Study

A simulation study is conducted to assess the performance of the maximum likelihood estimators based on the bias and the standard deviation. For this purpose we fix the value of scale parameter  $\lambda$  as 1.0 and  $k$  as 2.0 and use a total of 10 parameter combinations and different sample sizes:

- $\beta = 0.9, 1.6, 2.4, 3.0, 3.6$
- $\sigma = 0.8, 1.2$
- $c = 2.0$
- $n = 200, 400, 600, 1000$ .

Table 1. Bias and standard deviation of MLEs

$n$	Parameter Values			Bias			Standard deviation		
	$\beta$	$\sigma$	$c$	$\hat{\beta}$	$\hat{\sigma}$	$\hat{c}$	$\hat{\beta}$	$\hat{\sigma}$	$\hat{c}$
200	0.9	0.8	2.0	0.0405	-0.1099	-0.2869	0.2584	0.1965	0.5151
		1.2	2.0	-0.0522	-0.0831	-0.1445	0.2413	0.3504	0.6509
	1.6	0.8	2.0	0.1048	-0.1327	-0.3389	0.4419	0.1984	0.5287
		1.2	2.0	-0.0708	-0.1124	-0.1768	0.4254	0.3679	0.6061
	2.4	0.8	2.0	0.1866	-0.1539	-0.3727	0.7029	0.2155	0.5439
		1.2	2.0	-0.1125	-0.0852	-0.1181	0.6022	0.3292	0.5119
	3.0	0.8	2.0	0.2090	-0.1371	-0.3220	0.8689	0.2007	0.4975
		1.2	2.0	-0.1481	-0.0753	-0.0997	0.7513	0.3192	0.4094
	3.6	0.8	2.0	0.2908	-0.1410	-0.3272	0.9849	0.2000	0.4989
		1.2	2.0	-0.1868	-0.0615	-0.0725	0.8934	0.3093	0.4723
400	0.9	0.8	2.0	-0.0055	-0.0678	-0.1824	0.2692	0.1952	0.5327
		1.2	2.0	-0.0491	-0.0514	-0.0794	0.2203	0.3218	0.5503
	1.6	0.8	2.0	0.0240	-0.0836	-0.2115	0.4559	0.1986	0.5196
		1.2	2.0	-0.0787	-0.0539	-0.0794	0.3683	0.3079	0.5197
	2.4	0.8	2.0	-0.0180	-0.0669	-0.1641	0.6924	0.1949	0.5006
		1.2	2.0	-0.1328	-0.0450	-0.0579	0.5631	0.3156	0.5154
	3.0	0.8	2.0	0.0210	-0.0784	-0.1896	0.8723	0.2009	0.5086
		1.2	2.0	-0.2187	-0.0315	-0.0342	0.7729	0.3279	0.5264
	3.6	0.8	2.0	0.0848	-0.0886	-0.2067	0.9795	0.2044	0.5119
		1.2	2.0	-0.2238	-0.0363	-0.0393	0.8898	0.3134	0.4933
600	0.9	0.8	2.0	0.0096	-0.0701	-0.1838	0.2370	0.1834	0.4824
		1.2	2.0	-0.0056	-0.0912	-0.1564	0.1708	0.2995	0.5166
	1.6	0.8	2.0	0.0312	-0.0779	-0.1995	0.4125	0.1871	0.4857
		1.2	2.0	-0.0978	-0.0439	-0.0647	0.3992	0.3188	0.5299
	2.4	0.8	2.0	0.0692	-0.0786	-0.1945	0.5941	0.1809	0.4542
		1.2	2.0	-0.1537	-0.0417	-0.0572	0.6028	0.3258	0.5237
	3.0	0.8	2.0	0.1197	-0.0883	-0.2148	0.7443	0.1843	0.4581
		1.2	2.0	-0.1559	-0.0456	-0.0605	0.7134	0.3109	0.4860
	3.6	0.8	2.0	0.1697	-0.0942	-0.2253	0.8988	0.1941	0.4729
		1.2	2.0	-0.1421	-0.0577	-0.0784	0.8394	0.3103	0.4806
1000	0.9	0.8	2.0	-0.0111	-0.0571	-0.1545	0.2536	0.1975	0.5211
		1.2	2.0	-0.0474	-0.0348	-0.0594	0.1995	0.3003	0.5249
	1.6	0.8	2.0	0.0322	-0.0814	-0.2140	0.4513	0.1957	0.5093
		1.2	2.0	-0.0928	-0.0118	-0.0147	0.3381	0.2625	0.4435
	2.4	0.8	2.0	0.0672	-0.0747	-0.1911	0.6325	0.1747	0.4406
		1.2	2.0	-0.1895	-0.0100	-0.0233	0.5365	0.2649	0.4324
	3.0	0.8	2.0	0.0679	-0.0706	-0.1763	0.7933	0.1741	0.4325
		1.2	2.0	-0.2058	-0.0068	-0.0055	0.7023	0.2791	0.4466
	3.6	0.8	2.0	0.0977	-0.0719	-0.1738	0.9267	0.1709	0.4202
		1.2	2.0	-0.1815	-0.0211	-0.0237	0.8076	0.2709	0.4289

For each parameter combination and each sample size, a random sample is simulated from the GBXII(L) distribution. The SAS NLPTR subroutine is used to estimate the parameters by maximizing the log-likelihood function. This process is repeated 100 times and the bias and the standard deviation are presented in Table 1. These results show that the maximum likelihood estimation method performs reasonably well. It is observed that the shape parameters  $\sigma$  and  $c$  are negatively biased and  $\beta$  is positively (or negatively) biased depending on  $\sigma$  less (or greater) than 1. In general, bias and the standard deviation of the parameters are reasonable and decrease as the sample size increases. The simulation study suggests that the maximum likelihood method can be used efficiently to estimate the parameters of the GBXII(L) distribution.

## 5. Application

In this section, three examples are used for the purpose of illustration. The MLEs are computed and the fit is compared to other well known distributions based on the p-value of the Kolmogorov-Smirnov (K-S) test statistic, the Akaike infor-

mation criterion (AIC), and the log-likelihood value. These data sets were selected because they come from completely diverse disciplines, have different sample sizes, have non-negative values, beside that they differ structurally, i.e., in terms of skewness and kurtosis.

### 5.1 Actuarial Science Data

This Actuarial Science data (skewness = 0.07, kurtosis = 0.08) is taken from Balakrishnan et. al. (2009) and later on was also used for model fitting by Tahir et. al. (2016). In order to be able to do long and short term financial estimation, such as the assessment of the reserve required to pay the 'minimum pensions', it is important for the Mexican Institute of Social Security (IMSS) to study the distributional behavior of the mortality of retired people on disability. This data set corresponding to lifetime (in years) of retired women with temporary disabilities, which are incorporated in the Mexican insurance public system and who died during 2004, are:

22, 24, 25(2), 27, 28, 29(4), 30, 31(6), 32(7), 33(3), 34(6), 35(4), 36(11), 37(5), 38(3), 39(6), 40(14), 41(12), 42(6), 43(5), 44(7), 45(10), 46(6), 47(5), 48(11), 49(8), 50(8), 51(8), 52(14), 53(10), 54(13), 55(11), 56(10), 57(15), 58(11), 59(9), 60(7), 61(2), 62, 63, 64(4), 65(2), 66(3), 71, 74, 75, 79, 86.

Tahir et. al. (2016) have compared the performance of four different distributions namely Dagum (D), Beta-Dagum (BD), McDonald-Dagum (McD) and Weibull-Dagum (WD) in modeling the above data set based on the following goodness-of-fit criteria: Kolmogorov-Smirnov test(K-S), Akaike information criterion (AIC) and Bayesian information criterion (BIC) and concluded that WD gives the overall best fit, followed by McD. We fit the Weibull-Dagum distribution, the beta-normal (BN) distribution, the Gumbel-Weibull (GW) distribution, and the GBXII(L) distribution to the data. The maximum likelihood estimates along with their standard errors and different goodness-of-fit statistics or model selection criteria are reported in Table 2. It is evident from this table that GBXII(L) has the smallest lowest AIC, BIC and K-S test statistic value and consequently the highest p-value. Figure 2 presents the histogram of the Actuarial Science data along with the estimated densities and also the empirical and estimated CDFs of the four models. Thus it can be concluded that the GBXII(L) model outperforms the other distributions under consideration.

Table 2. MLEs (standard errors in parentheses) and model selection criteria for the Actuarial Science data

Distribution	WD*	BN	GW	GBXII(L)
Parameter estimates (standard error)	$\hat{b} = 3.4237$ (2.0727)	$\hat{a} = 0.4095$ (0.1240)	$\hat{\beta} = 2.4296$ (0.8430)	$\hat{\beta} = 0.7250$ (0.1071)
	$\hat{\lambda} = 4.5590$ (9.5120)	$\hat{b} = 0.2673$ (0.0274)	$\hat{\sigma} = 1.9914$ (0.2775)	$\hat{\sigma} = 8.5001$ (2.9947)
	$\hat{\delta} = 1.0948$ (0.6243)	$\hat{\mu} = 44.4852$ (2.7850)	$\hat{a} = 3.6281$ (0.5796)	$\hat{c} = 25.1386$ (9.3949)
	$\hat{\beta} = 11.7383$ (13.7411)	$\hat{\sigma} = 5.2866$ (0.3680)	$\hat{\lambda} = 36.9150$ (4.2805)	$\hat{k} = 4.7764$ (0.8112)
				$\hat{\lambda} = 52.3397$ (1.0324)
Log Likelihood	-1050.59	-1051.02	-1050.86	-1041.01
AIC	2109.18	2110.04	2109.72	2092.02
BIC	2123.72	2124.58	2124.26	2110.19
K-S	0.0687	0.0729	0.0700	0.0416
(p-value)	(0.1427)	(0.1017)	(0.1288)	(0.7189)

\* parameter estimates are from Tahir et al. (2016).

### 5.2 Skin Folds Data

The second example considered here is taken from Cook and Weisberg (1994). The data was recorded at the Australian Institute of Sport (AIS) and contains 13 variables on 100 female Australian athletes. The sum of skin folds (SSF) for the 100 female Australian athletes given below is used to illustrate the application of the GBXII(L) distribution. The data is right skewed (skewness = 0.79, kurtosis = 0.73) and shows a bimodal trend.

33.8, 36.8, 38.2, 41.1, 41.6, 42.3, 43.5, 43.5, 46.1, 46.2, 46.3, 47.5, 47.6, 48.4, 49.0, 49.9, 50.0, 52.5, 52.6, 54.6, 54.6, 55.6, 56.8, 57.9, 58.9, 59.4, 61.9, 62.6, 62.9, 65.1, 67.0, 68.3, 68.9, 69.9, 70.0, 71.3, 71.6, 73.9, 74.7, 74.9, 75.1, 75.2, 76.2, 76.8, 77.0, 80.1, 80.3, 80.3, 80.3, 80.6, 83.0, 87.2, 88.2, 89.0, 90.2, 90.4, 91.0, 91.2, 95.4, 96.8, 97.2, 97.9, 98.0, 98.1, 98.3, 98.5, 99.8, 99.9, 101.1, 102.8, 102.8, 103.6, 103.6, 104.6, 106.9, 109.0, 109.1, 109.5, 109.6, 110.2, 110.7, 111.1, 113.5, 114.0, 115.9, 117.8, 122.1, 123.6, 125.9, 126.4, 126.4, 131.9, 136.3, 143.5, 148.9, 156.6, 156.6, 171.1, 181.7, 200.8



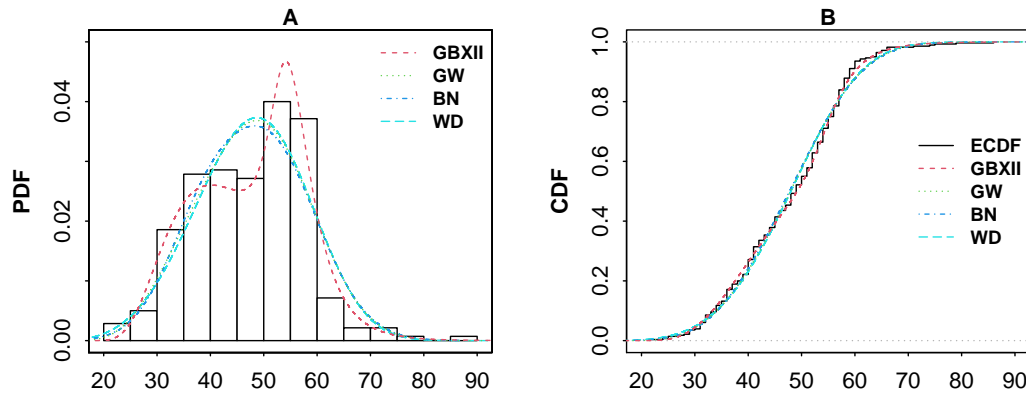


Figure 2. (A) Histogram and the estimated PDFs and (B) The empirical and the estimated CDFs for the Actuarial Science data

Four distributions are used to fit the data, namely, the exponentiated-Weibull (EW) distribution (Mudholkar et al., 1995), the beta-normal (BN) distribution, the Gumbel-Weibull (GW) distribution and the GBXII(L) distribution. The MLEs and the goodness of fit statistics are included in Table 3. Figure 3 contains the histogram and the fitted densities, in addition to the empirical and the fitted CDFs. The GBXII(L) has the best fit when compared to the other distributions. The GW and the BN can be unimodal or bimodal, however in this case they are unable to capture the bimodal trend of the data.

Table 3. MLEs (standard errors in parentheses) and model selection criteria for the skin folds data

Distribution	EW*	BN*	GW*	GBXII(L)
Parameter estimates (standard error)	$\hat{\alpha} = 1.2245$ (0.4378)	$\hat{\alpha} = 9.7706$ (0.4008)	$\hat{\beta} = 1.0299$ (1.0094)	$\hat{\beta} = 0.5779$ (0.1028)
	$\hat{\theta} = 6.5038$ (6.3413)	$\hat{b} = 0.1967$ (0.0223)	$\hat{\sigma} = 1.3303$ (0.3218)	$\hat{\sigma} = 14.3367$ (1.7127)
	$\hat{\sigma} = 41.4980$ (24.4832)	$\hat{\mu} = 9.1517$ (4.3612)	$\hat{\alpha} = 2.2896$ (0.9184)	$\hat{c} = 30.1079$ (1.3824)
		$\hat{\sigma} = 25.4309$ (0.9049)	$\hat{\lambda} = 81.7841$ (27.2786)	$\hat{k} = 2.7477$ (0.5391)
				$\hat{\lambda} = 94.6603$ (3.4293)
Log Likelihood	-487.17	-487.06	-485.89	-483.09
AIC	980.34	982.12	979.78	976.18
BIC	988.16	992.54	990.20	989.21
K-S	0.0808	0.0711	0.0705	0.0422
(p-value)	(0.5307)	(0.6925)	(0.7022)	(0.9941)

\*parameter estimates are from Al-Aqtash et al. (2014)

### 5.3 Glass Fibers Data

The strengths of 1.5 cm glass fibers data presented below was recorded at the National Physical Laboratories, England. This data is reported in Smith and Naylor (1987) and is left skewed with skewness = -0.95 and kurtosis = 1.10. The data was used by Barreto-Souza et al. (2010) in an application of the beta-generalized exponential (BGE) distribution, and was also used by Al-Aqtash et al. (2015) as an application of the Gumbel-Weibull (GW) distribution.

0.55, 0.74, 0.77, 0.81, 0.84, 0.93, 1.04, 1.11, 1.13, 1.24, 1.25, 1.27, 1.28, 1.29, 1.30, 1.36, 1.39, 1.42, 1.48(2), 1.49(2), 1.50(2), 1.51, 1.52, 1.53, 1.54, 1.55(2), 1.58, 1.59, 1.60, 1.61(4), 1.62(2), 1.63, 1.64, 1.66(3), 1.67, 1.68(2), 1.69, 1.70(2), 1.73, 1.76(2), 1.77, 1.78, 1.81, 1.82, 1.84(2), 1.89, 2.00, 2.01, 2.24.

Again four distributions are used to fit the data, namely, the beta-generalized exponential distribution, the beta-Birnbaum-

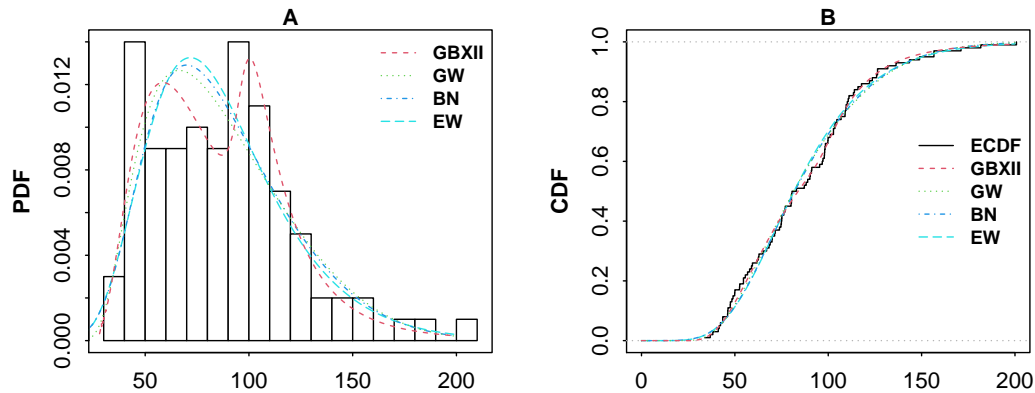


Figure 3. (A) Histogram and the estimated PDFs and (B) The empirical and the estimated CDFs for the skin folds data

Saunders (BBS) distribution (Cordeiro and Lemonte, 2011), the Gumbel-Weibull distribution, and the GBXII(L) distribution. The MLEs and goodness of fit statistics are presented in Table 4. Figure 4 contains the histogram of the data and the fitted PDFs as well as the empirical and fitted CDFs.

Table 4. MLEs (standard errors in parentheses) and model selection criteria for the glass fibers data

Distribution	BGE*	BBS**	GW**	GBXII(L)
Parameter estimates (standard error)	$\hat{\alpha} = 0.4125$	$\hat{\alpha} = 0.2938$ (0.1133)	$\hat{\beta} = 2.5598$ (0.6502)	$\hat{\beta} = 1.4292$ (0.2573)
	$\hat{b} = 93.4655$	$\hat{b} = 1080.71$ (24.4015)	$\hat{\sigma} = 5.5826$ (1.7978)	$\hat{\sigma} = 16.9395$ (4.9064)
	$\hat{\alpha} = 22.6124$	$\hat{\alpha} = 0.3995$ (0.1316)	$\hat{\alpha} = 5.2124$ (0.8858)	$\hat{c} = 24.2216$ (5.8116)
	$\hat{\lambda} = 0.9227$	$\hat{\beta} = 6.1727$ (1.9987)	$\hat{\lambda} = 1.0675$ (0.1271)	$\hat{k} = 10.0843$ (2.9373) $\hat{\lambda} = 1.5232$ (0.0421)
Log Likelihood	-15.60	-16.72	-11.52	-8.39
AIC	39.20	41.44	31.04	26.78
BIC	47.77	50.01	39.61	37.50
K-S (p-value)	0.1673 (0.0588)	0.1830 (0.0294)	0.0979 (0.5819)	0.075 (0.9989)

\* parameter estimates are from Barreto-Souza et al. (2010)

\*\* parameter estimates are from Al-Aqtash et al. (2015)

By comparing the goodness of fit statistics among the four distributions, we observe that the GBXII(L) outperforms the other distributions.

## 6. The Log-GBXII Regression Model With Application to Survival Data

Suppose that the survival time  $X$  follows the GBXII distribution in (5) with the parameters  $\beta, \sigma, c, k, \lambda > 0$ , then the survival function for the GBXII distribution is given by

$$S_{GBXII}(x) = 1 - \exp[-\beta\{(1 + (x/\lambda)^c)^k - 1\}^{-1/\sigma}], \quad 0 < x < \infty; \beta, \sigma, c, k, \lambda > 0.$$

If we take the log transform of  $X$ , and redefine the parameters as  $c = 1/\tau$  and  $\lambda = e^\mu$ , then,  $Y = \log(X)$ , can be written as a log linear model,  $Y = \mu + \tau W$ , where the random variable  $W = (Y - \mu)/\tau$  is the standardized log-GBXII (SLGBXII) distribution with PDF,

$$\pi_{SLGBXII}(w) = \frac{\beta k}{\sigma} e^w (1 + e^w)^{k-1} \left[ (1 + e^w)^k - 1 \right]^{-1-1/\sigma} \exp \left\{ -\beta \left[ (1 + e^w)^k - 1 \right]^{-1/\sigma} \right\} \quad (10)$$

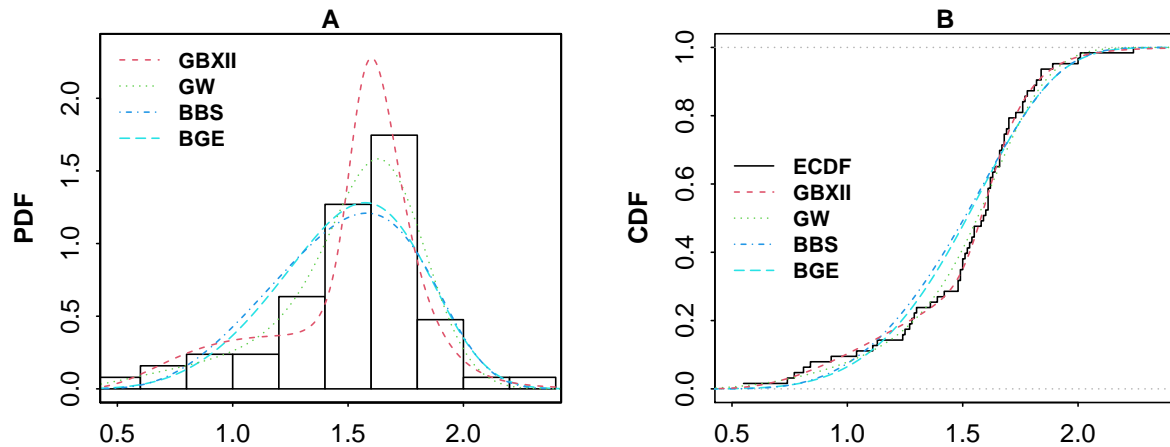


Figure 4. (A) Histogram and the estimated PDFs and (B) The empirical and the estimated CDFs for the glass fibers data

Thus, the underlying PDF and survival function, respectively, for  $Y$ , are respectively given by

$$g_{LGBXII}(y) = \frac{\beta k}{\sigma \tau} e^{((y-\mu)/\tau)} \left(1 + e^{((y-\mu)/\tau)}\right)^{k-1} \left[\left(1 + e^{((y-\mu)/\tau)}\right)^k - 1\right]^{-1-1/\sigma} \exp\left\{-\beta \left[\left(1 + e^{((y-\mu)/\tau)}\right)^k - 1\right]^{-1/\sigma}\right\} \quad (11)$$

and

$$S_{LGBXII}(y) = 1 - \exp\left\{-\beta \left[\left(1 + e^{((y-\mu)/\tau)}\right)^k - 1\right]^{-1/\sigma}\right\},$$

where  $y \in \mathbb{R}$ ,  $\mu \in \mathbb{R}$ ,  $\tau, \beta, \sigma, k > 0$ . Plots of the LGBXII density function in (11) for some parameter values are given in Figure 5

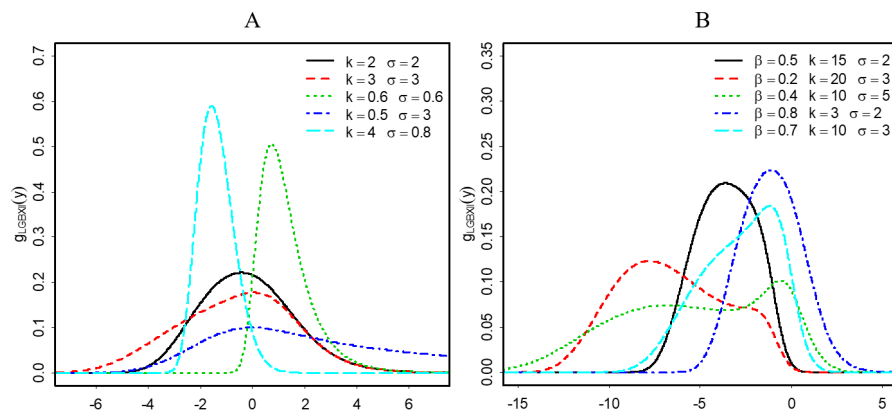


Figure 5. Plots of LGBXII density function for some parameter values: (A)  $\beta = 1, \mu = 0$  and  $\tau = 1$  and (B)  $\beta < 1, \mu = 0$  and  $\tau = 1$

In the analysis of most survival data, the relationship between the covariates and the survival time  $X$  is of interest. This relationship might be written as a linear relationship between the log of the survival time  $X$  and the covariate values as follows:

Consider the survival time  $X_i$  of the  $i^{th}$  individual in the sample, for  $i = 1, \dots, n$ , and suppose that we also have a set of  $p$  covariates such that  $Z_i = (1, z_{i1}, \dots, z_{ip})^T$ , where the 1 is for the intercept term. The log-linear (or location-scale) regression model which links the dependent variable  $Y_i = \log(X_i)$  and the set of covariates is given by

$$Y_i = \log(X_i) = \gamma^T Z_i + \tau W_i, \quad (12)$$

where  $\gamma^T = (\gamma_0, \gamma_1, \dots, \gamma_p)$  are the unknown regression coefficients of the values of  $p$  covariates,  $\tau$  is an unknown scale parameter, and  $W_i$  is the error variable.

In order to incorporate covariates into the LGBXII model, we use the log-linear model (12) for the survival time  $X_i$ , where  $W_i$  has the SLGBXII distribution (10) such that  $\mu_i = \gamma^T Z_i$  is the location parameter of  $Y_i$  and  $\tau, \beta, \sigma, k > 0$  are unknown parameters. With the model in (12), the survival function for  $Y$  is expressed as

$$S_{LGBXII}(y|Z) = 1 - \exp \left\{ -\beta \left[ \left( 1 + \exp \left( \frac{y - \gamma^T Z}{\tau} \right) \right)^k - 1 \right]^{-1/\sigma} \right\}.$$

Now, consider a sample of  $n$  independent observations, and let the random variables  $X_i$  and  $C_i$  denote the lifetime and censoring time for the  $i^{th}$  individual, whereas the response  $Y_i$  represents a log-lifetime or a log-censoring time for the  $i^{th}$  individual, such that  $Y_i = \min(\log(X_i), \log(C_i))$  for  $i = 1, 2, \dots, n$ . If all the observations are uncensored, then based on the LGBXII distribution in (11), the log likelihood for the model parameters  $\theta = (\beta, k, \sigma, \tau, \gamma^T)^T$  can be written as

$$\begin{aligned} l(\theta) &= \sum_{i=1}^n \log(g_{LGBXII}(y_i)) \\ &= n \log(\beta k) - n \log(\sigma \tau) + \sum_{i=1}^n w_i + (k-1) \sum_{i=1}^n \log(1 + e^{w_i}) \\ &\quad - (1 + (1/\sigma)) \sum_{i=1}^n \log[(1 + e^{w_i})^k - 1] - \beta \sum_{i=1}^n [(1 + e^{w_i})^k - 1]^{-1/\sigma}, \end{aligned} \quad (13)$$

where  $w_i = (y_i - \gamma^T z_i)/\tau$ .

Assume some of the observations are right censored and let  $C$  and  $F$  be the sets of censored and uncensored observations, respectively. In addition, if we assume non-informative censoring such that the censoring times are independent of the observed survival times, then the log-likelihood function for the model parameters  $\Theta = (\beta, k, \sigma, \tau, \gamma^T)^T$  is given by

$$\begin{aligned} l(\theta) &= \sum_{i \in F} \log(g_{LGBXII}(y_i)) + \sum_{i \in C} \log(S_{LGBXII}(y_i)) \\ &= m \log(\beta k) - m \log(\sigma \tau) + \sum_{i \in F} w_i + (k-1) \sum_{i \in F} \log(1 + e^{w_i}) \\ &\quad - (1 + (1/\sigma)) \sum_{i \in F} \log[(1 + e^{w_i})^k - 1] - \beta \sum_{i \in F} [(1 + e^{w_i})^k - 1]^{-1/\sigma} \\ &\quad + \sum_{i \in C} \log \left( 1 - \exp \left\{ -\beta [(1 + e^{w_i})^k - 1]^{-1/\sigma} \right\} \right), \end{aligned} \quad (14)$$

where  $m$  is the number of uncensored observations.

The MLE  $\hat{\theta}$  can be obtained by maximizing the log-likelihood function in (13) or (14). It is common to use numerical non-linear optimization methods for that purpose. The NLMIXED procedure in SAS is used in this article to obtain the values in  $\hat{\theta}$ .

In the remaining part of this section, we apply the LGBXII regression model to fit the data from a two-arms (different treatments) clinical trial, which was previously analyzed by Efron (1988), from a study comparing treatment with radiotherapy alone (Arm A) to radiotherapy plus chemotherapy (Arm B), for head and neck cancer. The response for each patient is the survival time in days. Nine patients out of 51 in Arm A and 14 patients out of 45 in Arm B were lost to follow-up (censored).

Now, let  $Y_i$  be the log survival time for the  $i^{th}$  patient and  $z_{i1}$  be a binary covariate: two-arms, which is coded as (0 for Arm A, or 1 for Arm B). We fit this data using the LGBXII regression model. The log linear model is defined as  $Y_i = \gamma_0 + \gamma_1 z_{i1} + \tau W_i$ , for  $i = 1, 2, \dots, 96$ , where the random variable  $W_i$  follows the SLGBXII distribution with PDF (10).

We also compare the fit of the LGBXII regression model with the fits of other competitive lifetime models, namely: log-Gumbel-Weibull (LGW) (Al-Aqtash et al., 2015), log-beta-Weibull (LBW) (Ortega et al., 2011), and log-Pareto-Weibull generalized lambda (LPWGL) (Aldeni et al., 2017). The LGW, LBW, and LPWGL survival functions (for  $y \in \mathbb{R}$ ,  $-\infty < \mu = \gamma_0 + \gamma_1 z_{i1} < \infty$ ,  $\tau > 0$  and the remaining parameters are all positive) are given below:

LGW survival function

$$S(y; \beta, \sigma, \tau, \mu) = 1 - \exp \left\{ -\beta (\exp(\exp((y - \mu)/\tau)) - 1)^{-1/\sigma} \right\}.$$

LBW survival function

$$S(y; a, b, \tau, \mu) = 1 - I_{1 - \exp[-\exp((y - \mu)/\tau)]}(a, b).$$

LPWGL survival function

$$S(y; \beta, b, s, \tau, \mu) = \left( \frac{\beta}{\beta - 1 + \exp(b \exp((y - \mu)/\tau))} \right)^s$$

Table 5. Parameter estimates and fit statistics for the two-arms data (standard errors in parentheses) and [*p*-values in brackets]

Model	LPWGL	LGW	LBW	LGBXII
Parameter estimates (standard error) [ <i>p</i> -value]	$\hat{\tau} = 1.2044$ (0.1104)	$\hat{\tau} = 5.4525$ (3.5904)	$\hat{\tau} = 4.2531$ (1.2971)	$\hat{\tau} = 0.0734$ (0.0134)
	$\hat{\gamma}_0 = 6.0071$ (0.1232)	$\hat{\gamma}_0 = 2.4528$ (7.4229)	$\hat{\gamma}_0 = -2.0623$ (1.2296)	$\hat{\gamma}_0 = 4.0116$ (0.1698)
	[<0.0001]	[0.7418]	[0.0967]	[<0.0001]
	$\hat{\gamma}_1 = 0.7552$ (0.2789)	$\hat{\gamma}_1 = 0.5820$ (0.2692)	$\hat{\gamma}_1 = 0.5573$ (0.2660)	$\hat{\gamma}_1 = 0.3186$ (0.1821)
	[0.0080]	[0.0331]	[0.0388]	[0.0833]
	$\hat{b} = 0.0410$ (0.0051)	$\hat{\sigma} = 0.4626$ (0.3481)	$\hat{b} = 0.4912$ (0.3976)	$\hat{\sigma} = 39.6338$ (22.7883)
	$\hat{s} = 279.1100$ (48.2961)	$\hat{\beta} = 19.8368$ (85.6856)	$\hat{a} = 81.8537$ (8.3787)	$\hat{k} = 2.7447$ (1.3543)
	$\hat{\beta} = 12.2901$ (2.1498)			$\hat{\beta} = 2.9077$ (0.4447)
AIC	316.2	293.9	293.3	287.2
BIC	331.6	306.7	306.1	302.5

Among the fitted models, the LGBXII model has the lowest AIC and BIC values as indicated in Table 5, so the LGBXII model provides the best fit to the data, followed successively by the LBW and LGW models. In the fitted LGBXII regression model, we see that the covariate two-arms is not significant at the 5% level. In other words, Arm A clinical trial is not significantly different from Arm B clinical trial for the survival times. The plots of the empirical survival (Kaplan-Meier) function and the estimated survival functions of the LGBXII, LBW, and LGW models are depicted in Figure 6. These plots suggest that the LGBXII model is appropriate to fit this data. The two shape parameters,  $k$  and  $\sigma$  provide more flexibility to this distribution, so it can be considered a very competitive model to other lifetime models.

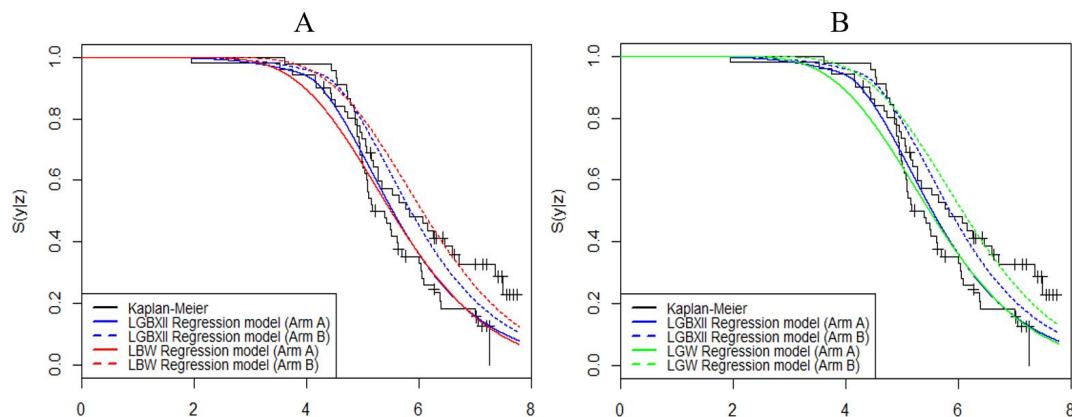


Figure 6. Kaplan-Meier and the estimated survival functions for the two-arms data (A) LGBXII model versus LBW model. (B) LGBXII model versus LGW model

## 7. Concluding Remarks

In this article, we discuss a member of the  $T$ - $X$ (Logit) family of distributions, namely the GBXII(L) distribution. Some useful characterizations and additional properties of the GBXII(L) are discussed. A simulation study is conducted to

assess the estimation of the parameters using the maximum likelihood method. From the results of the simulation study, the MLEs perform reasonably well and improve as the sample size increases. To illustrate the usefulness of GBXII(L) model in application, three real data sets arising from diverse disciplines and having different characteristics (different sample sizes and different shapes) are used. The fit is compared with other existing models based on the K-S, AIC and BIC statistics. The GBXII(L) performed very well in fitting different right skewed or left skewed data that can be unimodal or bimodal. A LGBXII regression model is proposed to explain and predict a response variable. A two-arm (clinical trial) survival time data is used in an application of the proposed LBGXII regression model and the fit is compared to three other comparable models. The LBGXII regression model performed very well based on the AIC and the BIC statistics. Thus our results exhibit the fact that GBXII(L) probability model and the LGBXII regression model may be used as alternatives to some other well-known models. Future research includes exploring efficient parameter estimation, extension to the multivariate cases and associated inference.

## References

- Akinsete, A., Famoye, F., & Lee, C. (2008). The beta-Pareto distribution. *Statistics*, 42(6), 547-563.
- Akinsete, A., Famoye, F., & Lee, C. (2014). The Kumaraswamy-geometric distribution. *Journal of Statistical Distributions and Applications*, 1(17), 1-21. <https://doi.org/10.1186/s40488-014-0017-1>
- Al-Aqtash, R. (2013). *On Generating New Families of Distributions Using the Logit Function*. Unpublished doctoral dissertation, Central Michigan University, Mt Pleasant, MI.
- Al-Aqtash, R., Famoye, F., & Lee, C. (2015). On generating a new family of distributions using the logit function, *Journal of Probability and Statistical Science*, 1, 135-152.
- Al-Aqtash, R., Lee, C., & Famoye, F. (2014). The Gumbel-Weibull distribution: Properties and applications, *Journal of Modern Applied Statistical Methods*, 13(1), 201-225. <https://doi.org/10.22237/jmasm/1414815000>
- Aldeni, M., Lee, C., & Famoye, F. (2017). Families of distributions arising from the quantile of generalized lambda distribution. *Journal of Statistical Distributions and Applications*, 4(25), 1-18. <https://doi.org/10.1186/s40488-017-0081-4>
- Aljarrah, M. A., Lee, C., & Famoye, F. (2014). On generating T-X family of distributions using quantile functions. *Journal of Statistical Distributions and Applications*, 1(2), 1-17. <https://doi.org/10.1186/2195-5832-1-2>
- Almheidat, M., Famoye, F., & Lee, C. (2015). Some Generalized Families of Weibull Distribution: Properties and Applications. *International Journal of Statistics and Probability*, 4(3), 18-35. <https://doi.org/10.5539/ijsp.v4n3p18>
- Alshawarbeh, E., Lee, C., & Famoye, F. (2012). The beta-Cauchy distribution. *Journal of Probability and Statistical Science*, 10(1), 41-57.
- Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions, *Metron*, 71(1), 63-79. <https://doi.org/10.1007/s40300-013-0007-y>
- Alzaatreh, A., Lee, C., & Famoye, F. (2014). T-normal family of distributions: a new approach to generalize the normal distribution. *Journal of Statistical Distributions and Applications*, 1(16), 1-18. <https://doi.org/10.1186/2195-5832-1-16>
- Balakrishnan, N., Leiva, V., Sanhueza, A., & Cabrera, E. (2009). Mixture inverse Gaussian distributions and its transformations, moments and applications. *Statistics*, 43, 91-104.
- Barreto-Souza, W., Santos, A. H. S., & Cordeiro, G. M. (2010). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*, 80(2), 159-172. <https://doi.org/10.1080/00949650802552402>
- Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull-G Family of Probability Distributions, *Journal of Data Science*, 12, 53-68.
- Burr, I. W. (1942). Cumulative frequency functions. *Annals of Mathematical Statistics*, 13, 215-232. <https://doi.org/10.1214/aoms/1177731607>
- Cook, R. D., & Weisberg, S. (1994). *An Introduction to Regression Graphics*. John Wiley and Sons, Inc., New York. <https://doi.org/10.1002/9780470316863>
- Cordeiro, G. M., & deCastro, M. (2011). A new family of generalized distributions, *Journal of Statistical Computation and Simulation*, 81(7), 883-898. <https://doi.org/10.1080/00949650903530745>
- Cordeiro, G. M., & Lemonte, A. J. (2011). The  $\beta$ -Birnbbaum-Saunders distribution: An improved distribution for fatigue life modeling. *Computational Statistics and Data Analysis*, 55(3), 1445-1461.

<https://doi.org/10.1016/j.csda.2010.10.007>

- Cordeiro, G. M., Ortega, E. M. M., & Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*, 347, 1399-1429. <https://doi.org/10.1016/j.jfranklin.2010.06.010>
- Cordeiro, G. M., Pescim, R. R., & Ortega, E. M. M. (2012). The Kumaraswamy generalized half-normal distribution for skewed positive data. *Journal of Data Science*, 10, 195-224.
- Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American statistical Association*, 83(402), 414-425. <https://doi.org/10.1080/01621459.1988.10478612>
- Eugene, N., Lee, C., & Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, 31(4), 497-512. <https://doi.org/10.1081/STA-120003130>
- Famoye, F., Lee, C., & Olumolade, O. (2005). The beta-Weibull distribution. *Journal of Statistical Theory and Applications*, 4(2), 121-136.
- Glänzel, W. (1986). A Characterization Theorem Based on Truncated Moments and its Application to Some Distribution Families, *Mathematical Statistics and Probability*, P. Bauer, F. Konecny, W. Wertz (Eds), Vol B, 75-84. [https://doi.org/10.1007/978-94-009-3965-3\\_8](https://doi.org/10.1007/978-94-009-3965-3_8)
- Glänzel, W. (1990). Some consequences of a characterization theorem based on truncated moments. *Statistics: A Journal of Theoretical and Applied Statistics*, 21(4), 613-618. <https://doi.org/10.1080/02331889008802273>
- Hamedani, G. G. (2013). *On certain generalized gamma convolution distributions II*. Technical Report No. 484, MSCS, Marquette University.
- Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous Univariate Distributions*. (2nd ed.) John Wiley and Sons, Inc., New York.
- Jones, M. C. (2009). Kumaraswamy's distribution: A beta-type distribution with tractability advantages. *Statistical Methodology*, 6, 70-81. <https://doi.org/10.1016/j.stamet.2008.04.001>
- Lee, C., Famoye, F., & Alzaatreh, A. (2013). Methods for generating families of univariate continuous distributions in the recent decades. *WIREs Computational Statistics*, 5, 219-238. <https://doi.org/10.1002/wics.1255>
- Mudholkar, G. S., Srivastava, D. K., & Freimer, M. (1995). The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics*, 37(4), 436-445. <https://doi.org/10.1080/00401706.1995.10484376>
- Nadarajah, S., & Okorie, I. (2018). A Note on a New Member from the T-X Family of Distributions: The Gumbel-Burr XII Distribution and its Properties. *Sankhya A: The Indian Journal of Statistics*. <https://doi.org/10.1007/s13171-018-0152-8>.
- Ortega, E. M., Cordeiro, G. M., & Hashimoto, E. M. (2011). A log-linear regression model for the Beta-Weibull distribution. *Communications in Statistics-Simulation and Computation*, 40(8), 1206-1235. <https://doi.org/10.1080/03610918.2011.568150>
- Osatohanmwun, P., Oyegun F. O., & Ogbonmwan, S. M. (2017). A New Member from the T-X Family of Distributions: The Gumbel-Burr XII Distribution and Its Properties. *Sankhya A: The Indian Journal of Statistics*. <https://doi.org/10.1007/s13171-017-0110-x>.
- Smith, R. L., & Naylor, J. C. (1987). A comparison of maximum likelihood and Bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics*, 36, 358-369.
- Tahir, M. H., & Cordeiro, G. M. (2016). Compounding of distributions: a survey and new generalized classes. *Journal of Statistical Distributions and Applications*, 3(13), 1-35. <https://doi.org/10.1186/s40488-016-0052-1>
- Tahir, M. H., Cordeiro, G. M., Mansoor, M., Zubair, M., & Alizadeh, M. (2016). The Weibull-Dagum distribution: Properties and applications. *Communications in Statistics - Theory and Methods*, 45(24), 7376-7398.
- Tahir, M. H., Hussain, M. A., Cordeiro, G. M., Hamedani, G. G., Mansoor, M., & Zubair, M. (2016). The Gumbel-Lomax Distribution: Properties and Applications. *Journal of Statistical Theory and Applications*, 15(1), 61-79.

## Appendix A

**Theorem A1.** Let  $(\Omega, \mathcal{F}, \mathbf{P})$  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  $X : \Omega \rightarrow H$  be a continuous random variable with the distribution function  $F$  and let  $q_1$

and  $q_2$  be two real functions defined on  $H$  such that

$$\mathbf{E}[q_2(X) | X \geq x] = \mathbf{E}[q_1(X) | X \geq x] \eta(x), \quad x \in H,$$

is defined with some real function  $\eta$ . Assume that  $q_1, q_2 \in C^1(H)$ ,  $\eta \in C^2(H)$  and  $F$  is twice continuously differentiable and strictly monotone function on the set  $H$ . Finally, assume that the equation  $\eta q_1 = q_2$  has no real solution in the interior of  $H$ . Then  $F$  is uniquely determined by the functions  $q_1, q_2$  and  $\eta$ , particularly

$$F(x) = \int_a^x C \left| \frac{\eta'(u)}{\eta(u) q_1(u) - q_2(u)} \right| \exp(-s(u)) du,$$

where the function  $s$  is a solution of the differential equation  $s' = \frac{\eta' q_1}{\eta q_1 - q_2}$  and  $C$  is the normalization constant, such that  $\int_H dF = 1$ .

We like to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence (Glänzel, 1990), in particular, let us assume that there is a sequence  $\{X_n\}$  of random variables with distribution functions  $\{F_n\}$  such that the functions  $q_{1n}, q_{2n}$  and  $\eta_n$  ( $n \in \mathbb{N}$ ) satisfy the conditions of Theorem A1 and let  $q_{1n} \rightarrow q_1, q_{2n} \rightarrow q_2$  for some continuously differentiable real functions  $q_1$  and  $q_2$ . Let, finally,  $X$  be a random variable with distribution  $F$ . Under the condition that  $q_{1n}(X)$  and  $q_{2n}(X)$  are uniformly integrable and the family  $\{F_n\}$  is relatively compact, the sequence  $X_n$  converges to  $X$  in distribution if and only if  $\eta_n$  converges to  $\eta$ , where

$$\eta(x) = \frac{\mathbf{E}[q_2(X) | X \geq x]}{\mathbf{E}[q_1(X) | X \geq x]}.$$

This stability theorem makes sure that the convergence of distribution functions is reflected by corresponding convergence of the functions  $q_1, q_2$  and  $\eta$  respectively. It guarantees, for instance, the ‘convergence’ of characterization of the Wald distribution to that of the Lévy-Smirnov distribution if  $\alpha \rightarrow \infty$ .

A further consequence of the stability property of Theorem A1 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  $q_1, q_2$  and, specially,  $\eta$  should be as simple as possible. Since the function triplet is not uniquely determined it is often possible to choose  $\eta$  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.

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).