

Parametric and Semiparametric Estimations of Bivariate Truncated Type I Generalized Logistic Models driven from Copulas

Lamya A Baharith¹

King Abdulaziz University, Department of Statistics, Faculty of Science, Jeddah, Saudi Arabia

Correspondence: Lamya A. Baharith,, Department of Statistics, Faculty of Science, King Abdulaziz University, P.O.Box 42805, Jeddah 21551, Saudi Arabia. E-mail: lbaharith@kau.edu.sa

Received: September 11, 2017 Accepted: October 8, 2017 Online Published: November 16, 2017

doi:10.5539/ijsp.v7n1p72

URL: <https://doi.org/10.5539/ijsp.v7n1p72>

Abstract

Truncated type I generalized logistic distribution has been used in a variety of applications. In this article, a new bivariate truncated type I generalized logistic (BTTGL) distributional models driven from three different copula functions are introduced. A study of some properties is illustrated. Parametric and semiparametric methods are used to estimate the parameters of the BTTGL models. Maximum likelihood and inference function for margin estimates of the BTTGL parameters are compared with semiparametric estimates using real data set. Further, a comparison between BTTGL, bivariate generalized exponential and bivariate exponentiated Weibull models is conducted using Akaike information criterion and the maximized log-likelihood. Extensive Monte Carlo simulation study is carried out for different values of the parameters and different sample sizes to compare the performance of parametric and semiparametric estimators based on relative mean square error.

Keywords: copula, inference function for margin, maximum likelihood method, maximum pseudo-likelihood method, method-of-moments, truncated type I generalized logistic distribution

1. Introduction

Truncated logistic distribution has been used effectively in different lifetime applications. It was first introduced by Kjelsberg (1962) and then Balakrishnan (1985) studied the half-logistic distribution and its use as lifetime model. AL-Angary (1997) introduced the truncated type I generalized logistic (TTGL) distribution. In addition, several researchers such as Al-Hussaini et al. (2006), Atea (2001), Al-Hussaini and Ateya (2003, 2005), Ateya and Ahmed (2013), and Rao (2015) have studied the properties and inferences of TTGL distribution and suggested many applications in different fields.

The TTGL distribution with two parameters has the following cumulative distribution function (Cdf)

$$F(t) = \frac{1}{2^\alpha - 1} \left[\frac{2^\alpha}{1 + \exp(-t/\sigma)} - 1 \right], \quad t \geq 0. \quad (1)$$

The probability density function (Pdf) is

$$f(t) = \frac{2^\alpha \alpha}{(2^\alpha - 1)\sigma} \frac{\exp(-t/\sigma)}{[1 + \exp(-t/\sigma)]^{\alpha+1}}, \quad t \geq 0, \quad (2)$$

where α is the shape parameter and σ is the scale parameter (Rao, 2015).

There are many lifetime applications where we need to consider bivariate lifetime distributions. Lately, copula has become a popular method to construct bivariate and multivariate distributions from given marginals due to its flexibility and practical use in a variety of fields, for example, biostatistics, financial, and actuarial fields (Nelsen, 2007; Trivedi & Zimmer, 2005). The main advantage of copula is it allows the marginal to be modeled and analyzed separately from the dependence structure. Also, the mathematical simplicity of copula is another advantage. Several authors considered different copula functions to construct bivariate and multivariate distributions. These include multivariate Gompertz-Type distribution (Adham & Walker, 2001), bivariate half-logistic-type distribution (Adham et al. 2009), bivariate Birnbaum-Saunders distribution derived from Gaussian copula (Kundu, Balakrishnan, & Jamalizadeh, 2010), bivariate

Sinh-normal distribution (Kundu, 2014), Absolute continuous bivariate generalized exponential distribution derived from the clayton copula (Kundu & Gupta, 2011), multivariate generalized exponential distribution (Al-Hussaini & Ateya, 2006; Ateya & Al-Alazwari, 2013), and bivariate Burr Type X (Elaal, Mahmoud, EL-Gohary, & Baharith, 2016).

The main aim of this article is to establish bivariate truncated type I generalized logistic distributional (BTTGL) models driven from three different copula functions. These are the Gaussian, Frank, and Clayton which are widely used in the literature to fit bivariate lifetime data since they represent different dependence structures between variables (Nelsen, 2007). The marginals of these distributional models are univariate truncated type I generalized logistic. The flexibility of the BTTGL distribution arises from the presence of the five parameters and the different shapes of its joint Pdf that can be used quite effectively to analyze bivariate lifetime data.

The rest of this article is organized as follows. Section 2 introduces BTTGL distributional models driven from different copula functions. Parametric and semiparametric methods are used to estimate the parameters of the proposed distribution in Section 3. In Section 4, a goodness of fit test is presented to check the appropriateness of different copula functions. A real data set is analyzed to illustrate the performance of the proposed bivariate distributional models in Section 5. In Section 6, a simulation study is carried out to investigate and compare the performance of the different estimators. Finally, concluding remarks are presented in Section 7.

2. Bivariate Truncated Type I Generalized Logistic Distributions Driven from Copulas

The copula approach is derived from Sklar's theorem (Sklar, 1959), who stated that any multivariate distribution can be disintegrated to a copula and its continuous marginal. For the bivariate case, copulas are used to link two marginal distributions with joint distribution such that for every bivariate distribution function $F(t_1, t_2)$ with continuous marginal $F(t_1), F(t_2)$, there exist a unique copula function $C: [0,1]^2 \rightarrow [0,1]$ as follows

$$F(t_1, t_2) = C\{F_1(t_1), F_2(t_2)\}, \quad (3)$$

where C is a distribution function with uniform(0,1) margins. Therefore, the density function of the bivariate distribution can be written as

$$f(t_1, t_2) = f_1(t_1)f_2(t_2)c(F_1(t_1), F_2(t_2)), \quad (4)$$

where $f_1(t_1), f_2(t_2)$ are the density functions corresponding to $F_1(t_1), F_2(t_2)$ and $c(F_1(t_1), F_2(t_2)) = \frac{\partial^2 C}{\partial u \partial v}$ is the copula density, see (Nelsen, 1999).

There are several copula functions that can be used to construct BTTGL distribution with TTGL marginals given by (1). In this article, we will use the Gaussian, Frank, and Clayton functions to construct BTTGL distributional models.

The Gaussian copula with copula parameter ρ , which measures the degree of association and dependence is denoted by ρ for this copula to indicate the relation with normal distribution, takes the following form

$$C(u, v) = \varphi_\rho(\varphi^{-1}(u), \varphi^{-1}(v)),$$

where $u = F_1(t_1), v = F_2(t_2)$ are the marginals Cdf of TTGL distribution given by (1) for the random variables T_1 and T_2 , respectively, φ_ρ denotes the distribution function of a bivariate standard normal random variable, and φ^{-1} represents the inverse of standard normal (Mardia, 1970).

The joint Pdf of the BTTGL distribution with TTGL marginals driven from Gaussian copula become

$$f(t_1, t_2) = \prod_{j=1}^2 f_j(t_j) \left\{ \frac{1}{\sqrt{1-\rho^2}} \exp \left[\frac{-\rho}{2(1-\rho^2)} \{ \rho(z_1^2 + z_2^2) - 2z_1z_2 \} \right] \right\},$$

where $f_j(t_j)$ is the Pdf of TTGL distribution given by

$$f_j(t_j) = \frac{2^{\alpha_j} \alpha_j}{(2^{\alpha_j} - 1) \sigma_j} \frac{\exp(-t_j/\sigma_j)}{[1 + \exp(-t_j/\sigma_j)]^{\alpha_j+1}}, \quad j = 1, 2$$

Here, $\alpha_1, \alpha_2, \sigma_1, \sigma_2$ represent the shape and scale parameters, respectively, $\rho \in [-1, 1]$ is the copula parameter, and $z_1 = \varphi^{-1}(u), z_2 = \varphi^{-1}(v)$.

The Frank and Clayton copula functions are members of the Archimedean family which have been used widely in many disciplines. The Frank copula have symmetric dependence patterns where we have a strong association between T_1 and T_2 for large ρ in absolute value. The Frank copula takes the following form

$$C(u, v) = -\frac{1}{\theta} \ln \left[1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{(\exp(-\theta) - 1)} \right], \quad \theta \in (-\infty, 0) \cup (0, \infty).$$

Note that when $\theta=0$, implies independence between the two variables. The joint Pdf of T_1 and T_2 driven from Frank copula is given by

$$f(t_1, t_2) = \prod_{j=1}^2 f_j(t_j) \left\{ -\frac{1}{\theta} \ln \left[1 + \frac{\theta (\exp(\theta) - 1) (\exp(\theta(1+u+v)))}{(\exp(\theta) - \exp(\theta(1+u)) - \exp(-\theta(1+v)) + \exp(-\theta(u+v)))^2} \right] \right\},$$

where $u = F_1(t_1)$, $v = F_2(t_2)$ are the Cdf of TTGL distribution given by (1).

The Clayton copula have asymmetric dependence patterns with higher dependence between T_1 and T_2 at left tails. The corresponding joint Pdf of T_1 and T_2 are respectively, given by

$$C(u, v) = [u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, \quad \theta \in (0, \infty),$$

$$f(t_1, t_2) = \prod_{j=1}^2 f_j(t_j) (1 + \theta)(uv)^{-1-\theta} [u^{-\theta} + v^{-\theta} - 1]^{-2-1/\theta}.$$

The main advantage of the Frank and Clayton copulas is that their copula and density functions have simple closed form expressions. For details, see (Clayton, 1978; Frank, 1979; Hutchinson & Lai, 1990; Nelsen, 2007).

2.1 Dependency Measures

According to Schweizer and Wolff ("On nonparametric measures of dependence for random variables," n.d.), copula capture all the dependence between the variables T_1 and T_2 . This can be measured using Spearman's rho and Kendall's tau, which are respectively given by

$$\rho(\theta) = 12 \int_0^1 \int_0^1 uv dC(u, v) - 3 \quad (5)$$

$$\tau(\theta) = 4 \left(\int_0^1 \int_0^1 C(u, v) dC(u, v) \right) - 1 \quad (6)$$

Table 1 reports Kendall's tau and spearman's rho range for Gaussian, Frank, and Clayton copulas.

Table 1. Kendall's tau and spearman's rho range

Copula	Kendall's tau Range	Spearman's rho range
Gaussian	[-1,1]	[-1,1]
Frank	[-1,1]	[-1,1]
Clayton	[0,1]	

3. Parameter Estimation

In this section, parametric and semiparametric estimation are used to estimate the unknown parameters of the proposed distribution.

3.1 Parametric Methods of Estimation

There are two methods for fitting copula model. The first one is to estimate the parameters of the marginal distributions and the copula at once. The second method is the two-stage method where we estimate the parameters of the marginal distributions first by maximizing the marginal log-likelihood and the second stage estimate the copula parameter by using the resulting parameter estimates from the first stage and maximizing the log-likelihood for the copula.

Maximum likelihood estimation

Maximum likelihood (ML) estimation is usually used to estimate all the parameters simultaneously. Let $(T_{1i}, T_{2i})'$, $i=1, \dots, n$, be a bivariate random sample from $BTTGL(\alpha_1, \alpha_2, \sigma_1, \sigma_2, \theta)$ with density function given in (4), then the log-likelihood function of the joint distribution $L(\alpha_1, \alpha_2, \sigma_1, \sigma_2, \theta)$ can be expressed as

$$\begin{aligned}\ell_{ML} = \ln L &= \sum_{i=1}^n \ln f_1(t_{1i}) + \sum_{i=1}^n \ln f_2(t_{2i}) + \sum_{i=1}^n \ln c(F_1(t_{1i}), F_2(t_{2i})) \\ &= \prod_{j=1}^2 [n \ln(k_j) - \sum_{i=1}^n (t_{ji}/\sigma_j) - (\alpha_j + 1) \sum_{i=1}^n \ln(1 + \exp(-t_{ji}/\sigma_j))] + \ln c(F_1(t_{1i}), F_2(t_{2i})),\end{aligned}$$

where $k_j = \frac{2^{\alpha_j} \alpha_j}{(2^{\alpha_j} - 1) \sigma_j}$, $j=1,2$.

The first derivatives of ℓ_{ML} are given by

$$\frac{\partial \ell_{ML}(\alpha_1, \alpha_2, \sigma_1, \sigma_2, \theta)}{\partial \sigma_j} = 0 \quad \frac{\partial \ell_{ML}(\alpha_1, \alpha_2, \sigma_1, \sigma_2, \theta)}{\partial \alpha_j} = 0 \quad \frac{\partial \ell_{ML}(\alpha_1, \alpha_2, \sigma_1, \sigma_2, \theta)}{\partial \theta} = 0$$

The solution of the above system of nonlinear equations gives the ML estimates of the unknown parameters for the proposed BTTGL distribution using nonlinear optimization algorithms such as a quasi-Newton algorithm.

Confidence intervals (CI) can be obtained by using the large sample approximation in which the ML estimates $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\sigma}_1, \hat{\sigma}_2, \hat{\rho})$ are approximately multivariate normal with mean $(\alpha_1, \alpha_2, \sigma_1, \sigma_2, \theta)$ and covariance matrix I^{-1} , where I^{-1} is the inverse of the information matrix, thus

$$I^{-1} = \begin{pmatrix} -\frac{\partial^2 \ell}{\partial \alpha_1^2} & -\frac{\partial^2 \ell}{\partial \alpha_1 \partial \alpha_2} & -\frac{\partial^2 \ell}{\partial \alpha_1 \partial \sigma_1} & -\frac{\partial^2 \ell}{\partial \alpha_1 \partial \sigma_2} & -\frac{\partial^2 \ell}{\partial \alpha_1 \partial \rho} \\ -\frac{\partial^2 \ell}{\partial \alpha_2 \partial \alpha_1} & -\frac{\partial^2 \ell}{\partial \alpha_2^2} & -\frac{\partial^2 \ell}{\partial \alpha_2 \partial \sigma_1} & -\frac{\partial^2 \ell}{\partial \alpha_2 \partial \sigma_2} & -\frac{\partial^2 \ell}{\partial \alpha_2 \partial \rho} \\ -\frac{\partial^2 \ell}{\partial \sigma_1 \partial \alpha_1} & -\frac{\partial^2 \ell}{\partial \sigma_1 \partial \alpha_2} & -\frac{\partial^2 \ell}{\partial \sigma_1^2} & -\frac{\partial^2 \ell}{\partial \sigma_1 \partial \sigma_2} & -\frac{\partial^2 \ell}{\partial \sigma_1 \partial \rho} \\ -\frac{\partial^2 \ell}{\partial \sigma_2 \partial \alpha_1} & -\frac{\partial^2 \ell}{\partial \sigma_2 \partial \alpha_2} & -\frac{\partial^2 \ell}{\partial \sigma_2 \partial \sigma_1} & -\frac{\partial^2 \ell}{\partial \sigma_2^2} & -\frac{\partial^2 \ell}{\partial \sigma_2 \partial \rho} \\ -\frac{\partial^2 \ell}{\partial \rho \partial \alpha_1} & -\frac{\partial^2 \ell}{\partial \rho \partial \alpha_2} & -\frac{\partial^2 \ell}{\partial \rho \partial \sigma_1} & -\frac{\partial^2 \ell}{\partial \rho \partial \sigma_2} & -\frac{\partial^2 \ell}{\partial \rho^2} \end{pmatrix} = \begin{pmatrix} \text{var}(\hat{\alpha}_1) & \text{cov}(\hat{\alpha}_1, \hat{\alpha}_2) & \text{cov}(\hat{\alpha}_1, \hat{\sigma}_1) & \text{cov}(\hat{\alpha}_1, \hat{\sigma}_2) & \text{cov}(\hat{\alpha}_1, \hat{\rho}) \\ \text{cov}(\hat{\alpha}_2, \hat{\alpha}_1) & \text{var}(\hat{\alpha}_2) & \text{cov}(\hat{\alpha}_2, \hat{\sigma}_1) & \text{cov}(\hat{\alpha}_2, \hat{\sigma}_2) & \text{cov}(\hat{\alpha}_2, \hat{\rho}) \\ \text{cov}(\hat{\sigma}_1, \hat{\alpha}_1) & \text{cov}(\hat{\sigma}_1, \hat{\alpha}_2) & \text{var}(\hat{\sigma}_1) & \text{cov}(\hat{\sigma}_1, \hat{\sigma}_2) & \text{cov}(\hat{\sigma}_1, \hat{\rho}) \\ \text{cov}(\hat{\sigma}_2, \hat{\alpha}_1) & \text{cov}(\hat{\sigma}_2, \hat{\alpha}_2) & \text{cov}(\hat{\sigma}_2, \hat{\sigma}_1) & \text{var}(\hat{\sigma}_2) & \text{cov}(\hat{\sigma}_2, \hat{\rho}) \\ \text{cov}(\hat{\rho}, \hat{\alpha}_1) & \text{cov}(\hat{\rho}, \hat{\alpha}_2) & \text{cov}(\hat{\rho}, \hat{\sigma}_1) & \text{cov}(\hat{\rho}, \hat{\sigma}_2) & \text{var}(\hat{\rho}) \end{pmatrix}$$

Then the asymptotic CI for the parameters $(\alpha_1, \alpha_2, \sigma_1, \sigma_2, \rho)$ using above approach are

$$\hat{\alpha}_1 \pm z_{\gamma/2} \sqrt{\text{var}(\hat{\alpha}_1)}, \quad \hat{\alpha}_2 \pm z_{\gamma/2} \sqrt{\text{var}(\hat{\alpha}_2)},$$

$$\hat{\sigma}_1 \pm z_{\gamma/2} \sqrt{\text{var}(\hat{\sigma}_1)}, \quad \hat{\sigma}_2 \pm z_{\gamma/2} \sqrt{\text{var}(\hat{\sigma}_2)},$$

$$\hat{\rho} \pm z_{\gamma/2} \sqrt{\text{var}(\hat{\rho})}$$

Two-stage estimation

The two-stage approach named as the method of inference functions for margins (IFM), where the first stage of this method will estimate the parameters of the marginal distributions by maximizing the marginal log-likelihood (Harry Joe, 1997; H Joe & Xu, 1996). The marginal log-likelihood for F_1 and F_2 are given by

$$\begin{aligned}\ell_{F_j}(\alpha_j, \sigma_j) &= n \ln(k_j) - \sum_{i=1}^n (t_{ij}/\sigma_j) - (\alpha_j + 1) \sum_{i=1}^n \ln[1 + \exp(-t_{ij}/\sigma_j)] \\ \frac{\partial \ell_{F_j}}{\partial \sigma_j} &= \frac{1}{\sigma_j} \left[\sum_{i=1}^n (t_{ij}/\sigma_j) - n - (\alpha_j + 1) \sum_{i=1}^n (t_{ij}/\sigma_j) [1 + \exp(-t_{ij}/\sigma_j)]^{-1} \right] = 0\end{aligned}$$

$$\frac{\partial \ell_{F_j}}{\partial \alpha_j} = n \left[\frac{1}{\alpha_j} - \ln 2 (2^{\alpha_j} - 1)^{-1} \right] - \sum_{i=1}^n \ln [1 + \exp(-t_{ij}/\sigma_j)] = 0$$

The solution of the above system of nonlinear equations gives the ML estimates of the marginal parameters $(\alpha_1, \alpha_2, \sigma_1, \sigma_2)$

Then, the second stage will estimate the copula parameter by maximizing the log-likelihood of the copula parameter after substituting the estimated marginal parameters in the first stage, that is

$$\ell_{IFM} = \sum_{i=1}^n \ln c(\hat{F}_1(t_{1i}), \hat{F}_2(t_{2i})),$$

Where $\hat{F}_1(t_1)$ and $\hat{F}_2(t_2)$ denote the ML estimates of the marginal parameters obtained from the first stage. This method has the advantage of reducing computational time and mathematical complexity (H Joe & Xu, 1996).

3.2 Semiparametric Methods of Estimation

Three semiparametric methods are conducted to estimate the copula parameter. These are the maximum Pseudo-Likelihood method, and the two Method-of-moments: inversion Kendall's and inversion of Spearman's rho.

Maximum Pseudo-Likelihood method

Genest et al. (1995) proposed the maximum pseudo-likelihood (MPL) method which estimate the copula parameter independently of the marginal fitting. The method uses the empirical distribution of the marginal where the data $T = (T_{1i}, T_{2i})'$ is transformed to pseudo-observations having uniform margins. That is, the maximum pseudo-likelihood estimator of θ is obtained by maximizing the following log-likelihood

$$\ell_{MPL}(\theta) = \sum_{i=1}^n \ln[c(\hat{U}_i, \hat{V}_i)],$$

\hat{U}_i, \hat{V}_i are the pseudo-observations from C calculated from data as follows

$$\hat{U}_i = \frac{R_{1i}}{n+1} = \frac{n}{n+1} \hat{F}_1(t_{1i}), \quad \hat{V}_i = \frac{R_{2i}}{n+1} = \frac{n}{n+1} \hat{F}_2(t_{2i}), \quad (7)$$

R_{1i}, R_{2i} are respectively, the ranks of t_{1i}, t_{2i} .

Method-of-Moments

This method estimate the copula parameter using the relation between the copula with Kendall's tau and Spearman's rho (Genest, 1987; Genest & Rivest, 1993). That is, this method estimate the copula parameter by matching the sample correlation Spearman's rho or Kendall's tau to the dependence measure given in (5) and (6) which is independent of the marginal distributions. Therefore, consistent estimators of θ that are called inversion of Kendall's (itau) and inversion of Spearman's rho (irho) are, respectively given by

$$\theta_{n,\rho} = \rho^{-1} \left(\frac{12}{n(n+1)(n-1)} \sum_{i=1}^n R_{1i} R_{2i} - 3 \frac{n+1}{n-1} \right),$$

$$\theta_{n,\tau} = \tau^{-1} \left(\frac{4}{n(n-1)} \sum_{i=1}^n 1[t_{1i} \leq t_{1\epsilon}] 1[t_{2i} \leq t_{2\epsilon}] - 1 \right),$$

where R_{1i}, R_{2i} are respectively, the ranks of t_{1i}, t_{2i} (Kojadinovic & Yan, 2010).

4. Goodness of Fit Tests for Copula

The idea of this test is to compare the empirical copula with the parametric estimator derived under the null hypothesis (Dobrić & Schmid, 2007; Fermanian, 2005). That is, test if the copula C is well-represented by a specific copula C_θ

$$H_0: C = C_\theta \quad Vs. \quad H_1: C \neq C_\theta$$

Two approaches are commonly used in the literature to test the goodness of fit of a copula; the parametric bootstrap (Genest & R émillard, 2008) or the fast multiplier approach (Genest, R émillard, & Beaudoin, 2009; Kojadinovic, Yan, & Holmes, 2011). The goodness of fit tests based on the empirical process is

$$C_n(u, v) = \sqrt{n} \{C_n(u, v) - C_{\theta_n}(u, v)\},$$

Where $C_n(u, v)$ is the empirical copula of the data of T_1 and T_2 and is given by

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^n 1(U_{i,n} \leq u, V_{i,n} \leq v), \quad u, v \in [0, 1],$$

Where U_i, V_i are given by (7), $C_n(u, v)$ is a consistent estimator and θ_n is an estimator of θ obtained using the pseudo observations. According to Genest et al. (2009), the test statistics is the Cramer-von Miss and is defined as

$$S_n = \sum_{i=1}^n \{C_n(U_i, V_i) - C_{\theta_n}(U_i, V_i)\}^2 \quad (8)$$

See for details (Genest & Rémillard, 2008; Genest et al., 2009; Kojadinovic et al., 2011).

5. Illustrative Data Analysis

This section illustrates the applications and the performance of the proposed BTTGL distributional models by analyzing Football data set which describes the UEFA Champion's League for 2004-2005 and 2005-2006, see (Meintanis, 2007). The data represents the games in which the home team has at least one goal scored and one goal by any team. That is,

T_1 = time of the first goal scored by any team in minutes,

T_2 = time of the first goal scored by the home team.

Meintanis (2007) analyzed this data for the Marshall–Olkin bivariate exponential distribution while Kundu and Dey (2009) and Kundu and Gupta (2009) studied this data using the Marshall–Olkin bivariate Weibull and bivariate generalized exponential distributions, respectively. In this article, this data is also considered to analyze the BTTGL distribution. Some descriptive statistics of T_1 and T_2 are reported in Table 2. In addition, the scatter plot of T_1 and T_2 shows a positive correlation.

Table 2. Descriptive statistics of T_1 and T_2 .

Marginal	Min	Max	Mean	Median	Quartile	
					1st	3rd
T_1	2	82	40.89	41.00	25	54
T_2	2	85	32.86	28.00	41	48

There is no adequate goodness of fit test for general bivariate distribution as stated in (Kundu & Gupta, 2011). Therefore, the TTGL distribution was fitted first to the marginals separately, and the ML estimates of each marginal parameters are obtained to be $(\hat{\sigma}_1 = 16.98, \hat{\alpha}_1 = 6.78)$ and $(\hat{\sigma}_2 = 19.66, \hat{\alpha}_2 = 2.47)$, respectively. These estimates used as initial values for the model parameters of BTTGL distribution. Then Kolmogorov-Smirnov (K-S) goodness of fit test statistics is calculated to verify that the fits based on TTGL distribution are suitable for the data. The K-S test with associated p-value in bracket for T_1 are 0.0979 (0.8703) and for T_2 are 0.0952 (0.8904). Also, figure 1 shows the plots of the fitted and the empirical Cdf for both marginals based on ML estimates. These results indicate that TTGL distribution provides an appropriate fit for the data set and can be used to fit the marginals.

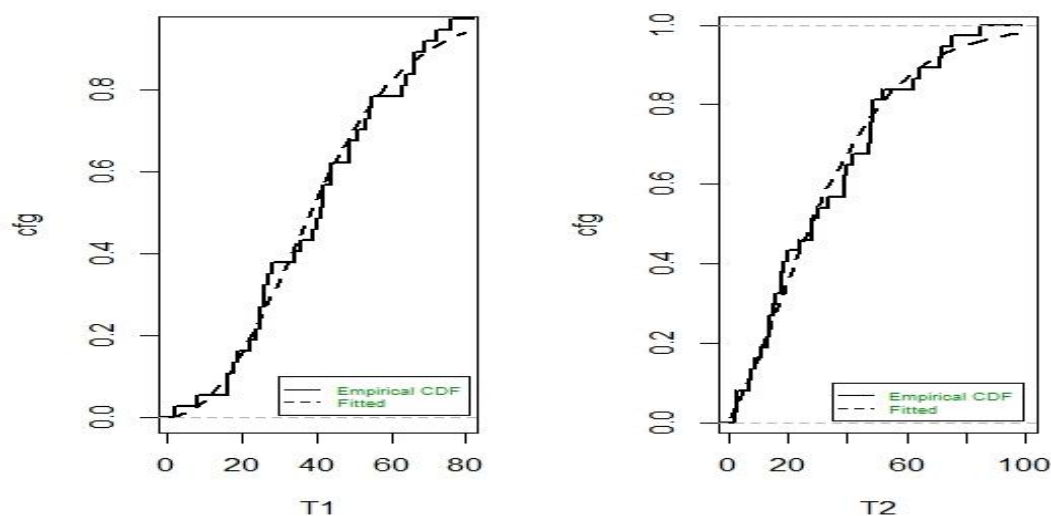


Figure 1. The plot for the fitted and empirical CDF of the TTGL distribution using ML estimates

Goodness of fit tests for copula functions

To check if the Gaussian, Frank, and Clayton copula functions are suitable for the data, the goodness of fit test statistics in Section 4 using equation (8) is calculated. The results in Table 3 show a non-significant p-values using parametric bootstrap which indicates that these copula functions provide suitable fit for the data.

Table 3. Cramer-von Miss goodness of fit test statistics with associated p-values.

Copula Function	statistic	p-value
Gaussian	0.0335	0.1703
Frank	0.0389	0.1414
Clayton	0.0561	0.0564

Table 4 display parameter estimates with associated 95% confidence intervals (CI) for the proposed BTTGL distributional models driven from different copula functions using parametric and semiparametric methods.

Table 4. ML, IFM, BML and moments estimates with associated 95% CI in bracket for BTTGL distributional models driven from selected copula functions.

Method	Copula Function	$\hat{\sigma}_1$	$\hat{\alpha}_1$	$\hat{\sigma}_2$	$\hat{\alpha}_2$	$\hat{\rho}$
ML	Gaussian	17.2558 (12.77,21.74)	6.52849 (2.96,10.10)	19.8838 (12.72,27.05)	2.4269 (0, 5.05)	0.4811 (0.22, 0.74)
	Frank	17.3091 (12.88,21.74)	6.4787 (2.93,10.03)	19.6912 (12.87,26.51)	2.6026 (0.02,5.19)	2.8241 (0.64,5.01)
	Clayton	17.5727 (12.98,22.17)	6.1637 (2.74, 9.59)	19.3822 (12.74,26.02)	2.7008 (0.12,5.28)	0.6380 (0.01,1.27)
IFM	Gaussian	16.98 (12.70,21.26)	6.78 (3.18,10.38)	19.66 (12.63,26.7)	2.47 (0,5.11)	0.4734 (0.24,0.71)
	Frank					2.7647 (0.70,4.83)
	Clayton					0.5838 (0.07, 1.10)
MPL	Gaussian					0.464 (0.26,0.66)
	Frank					2.948 (0.96, 4.94)
	Clayton					0.7198 (0.11, 1.33)
irho	Gaussian					0.423 (0.34, 5)
	Frank					2.668 (0,1.63)
	Clayton					0.778 (0, 1.63)
itau	Gaussian					0.469 (2.51, 3.58)
	Frank					3.044 (0, 1.88)
	Clayton					0.903 (0.12, 0.73)

Models Comparison

For comparison purposes, the univariate generalized exponential (GE) (Kundu & Gupta, 2009) and the exponentiated Weibull (EW) distributions are fitted to the marginals using this data set. Table 5 reports the K-S tests, associated p-value, and the maximized log-likelihood values (LL) for these distributions. It is clear that the p-values are not significant based on the TTGL, GE, and EW as marginals in which we can use these distributions to fit the data. Therefore, the BTTGL, the bivariate generalized exponential (BGE) and the bivariate exponentiated Weibull (BEW) distributions can be used for this data.

Table 5. Kolmogorov-Smirnov goodness of fit test statistics with associated p-values and maximized log-likelihood for the two marginal of TTGL, GE, and EW distributions.

Model	T_1			T_2		
	K-S	p-value	LL	K-S	p-value	LL
TTGL	0.0979	0.8703	-163.36	0.0952	0.8904	-163.60
GE	0.119	0.667	-165.82	0.121	0.656	-163.94
EW	0.1395	0.4679	-172.05	0.1426	0.4391	-167.21

Table 6. Maximum likelihood estimates for the copula parameter with the associated CI in bracket, AIC, and LL of the BTTGL, BGE, and BEW distributions.

Copula Function	Model	ρ	LL	AIC
Gaussian	BTTGL	0.4811 (0.22, 0.74)	-322.61	655.23
	BGE	0.4898 (0.24, 0.74)	-325.06	660.12
	BEW	0.5364 (0.29, 0.78)	-333.95	677.89
Frank	BTTGL	2.8241 (0.64, 5.01)	-323.37	656.75
	BGE	3.037 (0.70, 5.38)	-326.16	662.31
	BEW	3.7508 (0.96, 6.54)	-335.36	680.73
Clayton	BTTGL	0.6380 (0.01, 1.27)	-323.41	656.82
	BGE	0.4692 (-0.05, 0.99)	-326.06	662.12
	BEW	0.3671 (-0.05, 0.79)	-334.97	679.94

6. Simulation Studies

Here we will conduct a Monte Carlo simulation to illustrate the performance of the proposed BTTGL distribution driven from Gaussian copula function and compare between different methods of estimations. The Gaussian copula has the advantage that it is underline the multivariate normal distribution and incorporates dependency as the multivariate normal distribution does use only pairwise correlations among variables, which gives more flexibilities to the constructed

bivariate distribution. That is, the flexibility and analytical tractability of Gaussian copula suggest that it is a good choice to represent dependency.

To study the effect of the marginal parameters on the dependency measures, different sample sizes and different values of the parameters are used. In each case, parametric and semiparametric estimates of the parameters are obtained by computing the average estimates and their relative mean square errors (RMSE) over 1000 replications. Tables 7 and 8 present the results of the simulation.

Table 7 The average estimates and the corresponding RMSE (in brackets) of parametric and semiparametric estimates of BTTGL driven from Gaussian copula function for $\sigma_1=1.2$, $\sigma_2=1.1$, $\alpha_1=\alpha_2=2.8$ with different values of the copula parameter ρ .

n	method	ρ	$\hat{\sigma}_1$	$\hat{\alpha}_1$	$\hat{\sigma}_2$	$\hat{\alpha}_2$	$\hat{\rho}$
15	ML	0.5	1.1387(0.1159)	4.0674(5.1794)	1.0451(0.1139)	4.0721(2.8966)	0.4765(0.1808)
		0.7	1.1379(0.1146)	4.0879(3.1918)	1.0499(0.1096)	3.9821(2.4412)	0.6774(0.0777)
	IFM	0.5	1.1397(0.1168)	4.0611(5.3825)	1.0478(0.1148)	4.0508(2.9050)	0.4734(0.1919)
		0.7	1.1373(0.1187)	4.0976(3.2066)	1.0412(0.1121)	4.0030(2.4778)	0.6730(0.0865)
	MPL	0.5					0.4737(0.1986)
		0.7					0.6588(0.1195)
	Irho	0.5					0.4619(0.2514)
		0.7					0.6498(0.1430)
	itau	0.5					0.4757(0.2025)
		0.7					0.6683(0.1077)
50	ML	0.5	1.1905(0.0336)	3.0853(0.6057)	1.0982(0.0283)	3.0446(0.5890)	0.4932(0.0531)
		0.7	1.1984(0.0290)	2.9900(0.5241)	1.0989(0.0269)	3.0076(0.5339)	0.7016(0.0102)
	IFM	0.5	1.1906(0.0342)	3.0864(0.6177)	1.0978(0.0291)	3.0485(0.5961)	0.4920(0.0578)
		0.7	1.1980(0.0298)	2.9948(0.5401)	1.0996(0.0278)	3.0100(0.5624)	0.7002(0.0073)
	MPL	0.5					0.4877(0.0760)
		0.7					0.6930(0.0219)
	irho	0.5					0.4851(0.0904)
		0.7					0.6926(0.0243)
	itau	0.5					0.4904(0.0689)
		0.7					0.6996(0.0097)
80	ML	0.5	1.1928(0.0219)	2.9626(0.3501)	1.0980(0.0180)	2.9084(0.3407)	0.4985(0.0206)
		0.7	1.2016(0.0194)	2.9166(0.3523)	1.0966(0.0195)	2.9624(0.3546)	0.6974(0.0101)
	IFM	0.5	1.1927(0.0221)	2.9638(0.3536)	1.0985(0.0179)	2.9060(0.3486)	0.4979(0.023)
		0.7	1.1989(0.0196)	2.9399(0.3669)	1.0950(0.0210)	2.9745(0.3631)	0.6960(0.0131)
	MPL	0.5					0.4944(0.0317)
		0.7					0.6897(0.0262)
	irho	0.5					0.4941(0.0417)
		0.7					0.6905(0.0255)
	itau	0.5					0.4973(0.0289)
		0.7					0.6952(0.0156)

Table 8 The average estimates and the corresponding RMSE (in brackets) of parametric and semiparametric estimates of

BTGL driven from Gaussian copula function for $\sigma_1 = 1.2, \sigma_2 = 1.1, \alpha_1 = \alpha_2 = 2.8$ with different values of the copula parameter ρ .

n	method	ρ	$\hat{\sigma}_1$	$\hat{\alpha}_1$	$\hat{\sigma}_2$	$\hat{\alpha}_2$	$\hat{\rho}$
15	ML	0.5	0.6760(0.0828)	3.5727(1.5449)	0.7351(0.0440)	4.6685(1.3828)	0.4833(0.1325)
		0.7	0.6657(0.1101)	3.8643(2.6098)	0.7211(0.0783)	5.1231(2.2994)	0.6849(0.0619)
	IFM	0.5	0.6758(0.0835)	3.5761(1.5532)	0.7340(0.0461)	4.6802(1.3722)	0.4807(0.1422)
		0.7	0.6666(0.1108)	3.8732(2.6932)	0.7210(0.0783)	5.1298(2.3755)	0.6810(0.0694)
	MPL	0.5					0.4709(0.1901)
		0.7					0.6654(0.1055)
	Irho	0.5					0.4611(0.2345)
		0.7					0.6570(0.1283)
	itau	0.5					0.4714(0.1988)
		0.7					0.6759(0.0912)
50	ML	0.5	0.7019(0.0219)	2.9762(0.5677)	0.7403(0.0157)	4.2152(0.4902)	0.4977(0.0334)
		0.7	0.7051(0.0279)	2.9250(0.5689)	0.7415(0.0164)	4.1469(0.4139)	0.7019(0.0111)
	IFM	0.5	0.7002(0.0171)	2.9909(0.5592)	0.7394(0.0160)	4.2233(0.4895)	0.496 (0.0394)
		0.7	0.7038(0.0251)	2.9419(0.5797)	0.7414(0.0162)	4.1459(0.4199)	0.7005(0.0083)
	MPL	0.5					0.4926(0.05401)
		0.7					0.6923(0.0234)
	irho	0.5					0.4878(0.0767)
		0.7					0.6942(0.0211)
	itau	0.5					0.4938(0.0526)
		0.7					0.7005(0.0099)
80	ML	0.5	0.7020(0.0157)	2.9497(0.3645)	0.7404(0.0097)	4.1553(0.2896)	0.4999(0.0143)
		0.7	0.7029(0.0176)	2.9139(0.3759)	0.7391(0.0101)	4.1324(0.2768)	0.6955(0.0140)
	IFM	0.5	0.7019(0.0154)	2.9540(0.3729)	0.7403(0.0097)	4.1583(0.2958)	0.4991(0.0174)
		0.7	0.7025(0.0171)	2.9202(0.3873)	0.7393(0.0099)	4.1302(0.2850)	0.6947(0.0157)
	MPL	0.5					0.4954(0.0327)
		0.7					0.6883(0.0290)
	irho	0.5					0.4939(0.0407)
		0.7					0.6905(0.0256)
	itau	0.5					0.4978(0.0252)
		0.7					0.6946(0.0172)

From Tables 7 and 8, we concluded the following:

1. In all methods of estimation, as the dependency parameter (copula parameter) value increases, its RMSE become smaller.
2. The ML method mostly performs better than IFM method based on RMSE for different sample sizes and different values of the parameters.
3. As the sample size increases, the RMSE decrease for all the parameters using parametric and semiparametric methods, as expected.
4. For semiparametric methods, the itau method provides better estimates for the dependency parameter (copula parameter) and smaller RMSE compared to MPL and irho methods.
5. Based on the RMSE of the copula parameter, the parametric methods perform better than the semiparametric methods. This is consistent with the results found by Genest et al. (1995).
6. The marginal parameters generally have small effect on estimating the copula parameter as seen in Tables 7 and 8 with different values of the parameters.

7. Concluding Remarks

In this article, bivariate truncated generalized logistic distributional models are introduced. The proposed bivariate distributional models derived from commonly used copula functions with truncated generalized logistic distribution as marginals. Different methods of estimation are used to estimate the unknown parameters. The proposed bivariate distributional models derived from Gaussian, Frank, and Clayton copula functions fitted to a real life data set and the results of the analysis showed that the BTTGL distribution provides more suitable fit than bivariate generalized exponential and the bivariate exponentiated Weibull distributions. Also, the results indicated that the BTTGL distribution driven from Gaussian copula provides a better fit for the data set compared to BTTGL distributional models that are driven from Frank and Clayton copulas. Therefore, particular attention was directed to BTTGL distribution driven from Gaussian copula, and a Monte Carlo simulation was performed to compare the performance of the ML, IFM, BML and moment estimators which shows that the parametric methods perform better than semiparametric methods based on RMSE.

References

- Adham, S. A., AL-Dayian, G. R., El Beltagy, S. H., and Abd Elaal, M. K. (2009). Bivariate half- logistic-type distribution. *Academy of Business Journal, AL-Azhar University*, 2, 92–107.
- Adham, S. A., & Walker, S. G. (2001). A multivariate Gompertz-type distribution. *Journal of Applied Statistics*, 28(8), 1051–1065. <https://doi.org/10.1080/02664760120076706>
- AL-Angary, A. (1997). Truncated logistic distributions as lifetime models. King Abdulaziz University. MSc. thesis.
- Al-Hussaini, E. K., Al-Dayian, G. R., & Al-Angary, A. (2006). Bayesian prediction bounds under the truncated type I generalized logistic model. *J. Egyptian Math. Soc*, 14, 55–67.
- Al-Hussaini, E. K., & Ateya, S. F. (2003). Maximum likelihood estimations under a mixture of truncated type I generalized logistic components model. *J. Stat. Theory Appl*, 2(1), 47–60.
- Al-Hussaini, E. K., & Ateya, S. F. (2005). Bayes estimations under a mixture of truncated type I generalized logistic components model. *J. Stat. Theory Appl*, 4(2), 183–208.
- Al-Hussaini, E. K., & Ateya, S. F. (2006). A class of multivariate distributions and new copulas. *J. Egypt. Math. Soc*, 14(1), 45–54.
- Ateya, S. F. (2001). Mixtures Of Truncated Type I Generalized Logistic Components Model. King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia.
- Ateya, S. F., & Ahmad, A. E.-B. A. (2013). Inferences based on generalized order statistics under truncated type-I generalized logistic distribution. *Statistics*, 47(1), 141–155. <https://doi.org/10.1080/02331888.2011.586457>
- Ateya, S. F., & Al-Alazwari, A. A. (2013). A new multivariate generalized exponential distribution with application. *Far East Journal of Theoretical Statistics*, 45(1), 51.
- Balakrishnan, N. (1985). Order statistics from the half logistic distribution. *Journal of Statistical Computation and Simulation*, 20(4), 287–309. <https://doi.org/10.1080/00949658508810784>
- Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *In Biometrika*, 65, 141–151). Biometrika Trust. <https://doi.org/10.1093/biomet/65.1.141>
- Dobrić, J., & Schmid, F. (2007). A goodness of fit test for copulas based on Rosenblatt's transformation. *Computational*

- Statistics & Data Analysis*, 51(9), 4633–4642. <https://doi.org/10.1016/j.csda.2006.08.012>
- Elaal, M. K. A., Mahmoud, M. R., EL-Gohary, M. M., & Baharith, L. A. (2016). Univariate And Bivariate Burr Type X Distributions Based On Mixtures And Copula. *International Journal of Mathematics and StatisticsTM*, 17(1), 113–127.
- Fermanian, J.-D. (2005). Goodness-of-fit tests for copulas. *Journal of Multivariate Analysis*, 95(1), 119–152. <http://doi.org/10.1016/j.jmva.2004.07.004>
- Frank, M. J. (1979). On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. *Aequationes Mathematicae*, 19(1), 194–226. <https://doi.org/10.1007/BF02189866>
- Genest, C. (1987). Frank's family of bivariate distributions. *Biometrika*, 74(3), 549–555. <https://doi.org/10.1093/biomet/74.3.549>
- Genest, C., Ghoudi, K., & Rivest, L. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3), 543–552. <https://doi.org/10.1093/biomet/82.3.543>
- Genest, C., & R  millard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. In *Annales de l'IHP Probabilit  s et statistiques* (Vol. 44, pp. 1096–1127). Genest, C., R  millard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2), 199–213. <https://doi.org/10.1214/07-AIHP148>
- Genest, C., R  millard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2), 199–213. <https://doi.org/10.1016/j.insmatheco.2007.10.005>
- Genest, C., & Rivest, L. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, 88(423), 1034–1043. <https://doi.org/10.1080/01621459.1993.10476372>
- Hutchinson, T. P., & Lai, C. D. (1990). Continuous bivariate distributions, emphasising applications. Rumsby Scientific Publishing Adelaide.
- Joe, H. (1997). Multivariate models and multivariate dependence concepts. CRC Press. <https://doi.org/10.1201/b13150>
- Joe, H., & Xu, J. J. (1996). The Estimation Method of Inference Function for Margins for Multivariate Models. British Columbia.
- Kjelsberg, M. O. (1962). Estimation of the parameters of the logistic distribution under truncation and censoring. University of Minnesota, Minneapolis.
- Kojadinovic, I., & Yan, J. (2010). Comparison of three semiparametric methods for estimating dependence parameters in copula models. *Insurance: Mathematics and Economics*, 47(1), 52–63. <https://doi.org/10.1016/j.insmatheco.2010.03.008>
- Kojadinovic, I., Yan, J., & Holmes, M. (2011). Fast large-sample goodness-of-fit tests for copulas. *Statistica Sinica*, 21(2), 841. <https://doi.org/10.5705/ss.2011.037a>
- Kundu, D. (2014). Bivariate Sinh-Normal distribution and a related model. *Braz. J. Probab. Stat.*
- Kundu, D., Balakrishnan, N., & Jamalizadeh, A. (2010). Bivariate Birnbaum--Saunders distribution and associated inference. *Journal of Multivariate Analysis*, 101(1), 113–125. <https://doi.org/10.1016/j.jmva.2009.05.005>
- Kundu, D., & Dey, A. K. (2009). Estimating the parameters of the Marshall--Olkin bivariate Weibull distribution by EM algorithm. *Computational Statistics & Data Analysis*, 53(4), 956–965.
- Kundu, D., & Gupta, R. D. (2009). Bivariate generalized exponential distribution. *Journal of Multivariate Analysis*, 100(4), 581–593.
- Kundu, D., & Gupta, R. D. (2011). Absolute continuous bivariate generalized exponential distribution. *AStA Advances in Statistical Analysis*, 95(2), 169–185. <https://doi.org/10.1007/s10182-010-0151-0>
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530. <https://doi.org/10.1093/biomet/57.3.519>
- Meintanis, S. G. (2007). Test of fit for Marshall--Olkin distributions with applications. *Journal of Statistical Planning and Inference*, 137(12), 3954–3963. <https://doi.org/10.1016/j.jspi.2007.04.013>
- Nelsen, R. B. (1999). An Introduction to Copulas. Springer&Verla, New York. Inc. <https://doi.org/10.1007/978-1-4757-3076-0>
- Nelsen, R. B. (2007). An introduction to copulas. Springer Science & Business Media. On nonparametric measures of dependence for random variables. (n.d.).

- Rao, G. S. (2015). Estimation of stress-strength reliability from truncated type-I generalised logistic distribution. *International Journal of Mathematics in Operational Research*, 7(4), 372–382. <https://doi.org/10.1504/IJMOR.2015.070188>
- Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. Université Paris 8.
- Trivedi, P. K., & Zimmer, D. M. (2005). Copula modeling: an introduction for practitioners. *Foundations and Trends® in Econometrics*, 1(1), 1–111. <https://doi.org/10.1561/0800000005>

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/>).