Sankhyā : The Indian Journal of Statistics
2017, Volume 79-B, Part 1, pp. 170-186

Characterization of Bivariate Generalized Logistic Family of Distributions Through Conditional Specification

Indranil Ghosh
University of North Carolina, Wilmington, USA
N. Balakrishnan
McMaster University, Hamilton, Canada

Abstract

The univariate logistic distribution and its properties and applications have been studied quite extensively in the literature. Some generalizations as well as multivariate extensions of it have also been proposed for greater flexibility in modeling univariate and multivariate data. In this paper, we construct three different types of generalized bivariate logistic type distributions through conditional specification, and discuss some of their properties. Finally, we use a data set to fit the proposed models for the purpose of illustration.

AMS (2000) subject classification. Primary: 62E, Secondary: 62F.
Keywords and phrases. Generalized logistic distributions, Generalized bivariate logistic distributions, Conditional specification, TP_2 property, Unimodality, Likelihood inference

1 Introduction

In the last two decades, considerable amount of work has been done on introducing various bivariate non-normal models and then discussing their properties, fit and applications; for elaborate details, one may refer to the books by Kotz et al. (2000) and Balakrishnan and Lai (2009). One such model that has been studied extensively in the literature is the bivariate logistic; see Balakrishnan (1992) for pertinent details. In this work, we focus on characterizing three different forms of generalized bivariate logistic distributions via conditional specification. The simplicity of the logistic distribution and its importance as a growth curve have attracted many researchers to study this distribution and various generalizations. Many of these generalizations were introduced in order to provide greater flexibility while modeling either skewed and/or heavy tailed data. Let us consider the following three forms of generalized logistic distributions described in detail in Balakrishnan (1992):

1. Type I generalized logistic distribution: The pdf is

$$f(x;b) = \frac{be^{-x}}{(1+e^{-x})^{b+1}}, \quad -\infty < x < \infty, \quad b > 0; \tag{1.1}$$

2. Type II generalized logistic distribution: The pdf is

$$f(x;b) = \frac{be^{-bx}}{(1+e^{-x})^{b+1}}, \quad -\infty < x < \infty, \quad b > 0; \tag{1.2}$$

3. Type III generalized logistic distribution: The pdf is

$$f(x;b) = \frac{e^{-bx}}{(1+e^{-x})^{2b}\, B(b,b)}, \quad -\infty < x < \infty, \quad b > 0. \tag{1.3}$$

It needs to be mentioned that Type I and Type II families provide skewed alternatives to the logistic with varying levels of skewness depending on the value of shape parameter b, while Type III provides symmetric alternatives to the logistic with light and heavier tails.

Specification of joint distributions by means of conditional densities has received considerable attention in the literature. A study of bivariate distributions can not be complete without the knowledge of univariate distributions, which would naturally form the marginal or conditional distributions. Although multivariate data sets with logistic like marginals have always been around, it was not until 1961 that a bivariate logistic model was proposed by Gumbel. Gumbel (1961) actually provided three candidate bivariate logistic models. Among them, the third bivariate logistic distribution function takes on the simple form $F(x,y) = (1+e^{-x}+e^{-y})^{-1}, \quad (x,y) \in \mathbb{R}^2$. It is evident that this distribution does have logistic marginals. However, the usefulness of this distribution is severely limited by the absence of any parameter concerning the association between x and y. The correlation coefficient is a constant value, 1/2. This provides a motivation for the present work to seek richer families of bivariate logistic type distributions with different properties and dependence structures.

The rest of the paper is organized as follows. In Section 2, we discuss briefly the idea of conditional specification of bivariate distributions. In

Section 3, we consider the construction of Type I generalized bivariate logistic distributions via conditional specification that both the conditionals (i.e., X given Y and Y given X) are in the same family of univariate Type I generalized logistic distributions and then discuss some of their properties. In Section 4, we consider the construction of bivariate Type II generalized logistic distributions via conditional specification that both the conditionals are of univariate Type II generalized logistic with appropriate parameters and discuss some of their properties. In Section 5, we consider the construction of Type III generalized logistic distributions via conditional specification that both the conditionals are of univariate Type III generalized logistic with appropriate parameters and discuss their properties and shape characteristics. In Section 6, we discuss the likelihood inference for these models, and in Section 7, we illustrate the fit of these models with a data set. Finally, in Section 8, we provide some concluding remarks.

2 Conditional Specification of a Bivariate Distribution

There are different conditional specifications through which one can identify or (classify) a family of bivariate distributions. If we are given both families of conditional densities, of X given Y and Y given X, then the information is more than enough to characterize the joint density of (X, Y). We focuses on cases in which the conditional densities are only assumed to be known to belong to specified parametric families. The models thus derived are called conditionally specified models.

Before proceeding further, we mention the following theorem which plays a key role in the rest of our discussion. It is originally due to Aczel (see also Arnold et al. (1999)).

Theorem 1. *All solutions of the equation*

$$\sum_{i=1}^{r} f_i(x)\phi_i(y) = \sum_{j=1}^{s} g_j(y)\Psi_j(x), x \in S(X), y \in S(Y),$$

where ϕ_i (for $i = 1, 2, .., r$) and Ψ_j (for $j = 1, 2, .., s$) are given systems of mutually linearly independent functions, are of the form

$$\underline{f}(x) = C\underline{\Psi}(x)$$

and

$$\underline{g}(y) = D\underline{\phi}(y).$$

where $D = C'$

3 Bivariate Type I Generalized Logistic Distribution

For simplicity, we assume that the location parameters are zero and scale parameters are same; otherwise, an explicit solution for the resulting functional equation is not achievable. Suppose we want for each fixed x,

$$f(y|X=x) = \frac{\delta_1(x)e^{-y/\sigma}}{\left(1+e^{-y/\sigma}\right)^{\delta_1(x)+1}}, \quad y \in \mathbb{R}$$

and also for each fixed y,

$$f(x|Y=y) = \frac{\delta_2(y)e^{-x/\sigma}}{\left(1+e^{-x/\sigma}\right)^{\delta_2(y)+1}}, \quad x \in \mathbb{R},$$

where $\delta_1(x)$ is a function depending on x and $\delta_2(y)$ is a function depending on y and that both are unknown. We are now interested in identifying the class of all bivariate distributions for which both conditionals are Type I generalized logistic of the forms mentioned above. Now, let $g(x)$ and $h(y)$ be the marginal densities of X and Y, respectively. Then, we can write the joint density of (X, Y) as a product of marginals and conditionals, that is,

$$g(x)f(y|X=x) = h(y)f(x|Y=y), \tag{3.1}$$

or equivalently

$$f(x,y) = g(x)\delta_1(x)\frac{e^{-y/\sigma}}{\left(1+e^{-y/\sigma}\right)^{\delta_1(x)+1}} = h(y)\delta_2(y)\frac{e^{-x/\sigma}}{\left(1+e^{-x/\sigma}\right)^{\delta_2(y)+1}}. \tag{3.2}$$

Define

$$\theta_1(x) = g(x)\delta_1(x)$$

and

$$\theta_2(y) = h(y)\delta_2(y)$$

Then, we can write (3.2) equivalently as

$$\begin{aligned} &\exp\left(\log\theta_1(x) - y/\sigma - (\delta_1(x)+1)\log\left(1+e^{-y/\sigma}\right)\right) \\ &= \exp\left(\log\theta_2(y) - x/\sigma - (\delta_2(y)+1)\log\left(1+e^{-x/\sigma}\right)\right). \end{aligned} \tag{3.3}$$

Now, we rewrite (3.3) as

$$\begin{aligned} &\exp\left(\log\left(\theta_1(x)(1+e^{-x/\sigma})(e^{(x/\sigma)})\right) - \delta_1(x)\log\left(1+e^{-y/\sigma}\right)\right) \\ &= \exp\left(\log\left(\theta_2(y)(1+e^{-y/\sigma})(e^{(y/\sigma)})\right) - \delta_2(y)\log\left(1+e^{-x/\sigma}\right)\right). \end{aligned} \tag{3.4}$$

If we let

$$\begin{aligned} f_1(x) &= \log\left(\theta_1(x)(1+e^{-\frac{x}{\sigma}})(e^{(x/\sigma)})\right), \\ f_2(x) &= \delta_1(x), \\ g_1(y) &= \log\left(\theta_2(y)(1+e^{-\frac{y}{\sigma}})(e^{(y/\sigma)})\right), \\ g_2(y) &= \delta_2(y), \\ \phi_1(y) &= 1, \\ \phi_2(y) &= -\log\left(1+e^{-y/\sigma}\right), \\ \Psi_1(x) &= 1, \\ \Psi_2(x) &= -\log\left(1+e^{-x/\sigma}\right), \end{aligned} \tag{3.5}$$

then, (3.4) can be rewritten as

$$\sum_{i=1}^{2} f_i(x)\phi_i(y) = \sum_{i=1}^{2} g_i(x)\Psi_i(x).$$

This implies according to Aczel's theorem that a general solution to the above equation, is given by

$$f_1(x) = a - b\log\left(1+e^{-\frac{x}{\sigma}}\right), \tag{3.6}$$

$$f_2(x) = c - d\log\left(1+e^{-\frac{x}{\sigma}}\right), \tag{3.7}$$

$$g_1(y) = a - c\log\left(1+e^{-\frac{y}{\sigma}}\right), \tag{3.8}$$

$$g_2(y) = b - d\log\left(1+e^{-\frac{y}{\sigma}}\right), \tag{3.9}$$

where a, b, c, d are unknown parameters. Since, $f_2(x) = c - d\log\left(1+e^{-\frac{x}{\sigma}}\right) = \delta_1(x) > 0$, for all $x \in \mathbb{R}$, and $-\log\left(1+e^{-\frac{x}{\sigma}}\right) < 0$, for all $x \in \mathbb{R}$, it follows that c should be positive ($c > 0$) and $d \leq 0$. Similarly, from $g_2(y) = \delta_2(y) > 0$, we obtain that $b > 0$ and $d \leq 0$.

The constant a is a normalizing constant and it can be evaluated from the condition $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = 1$, or from the condition that the marginal densities when integrated over $(-\infty, \infty)$ is 1. To get the exact expression of the joint density $f(x,y)$, we need to simplify $\theta_1(x)$ and/ or $\theta_2(y)$. We have

$$f_1(x) = a - b\log\left(1+e^{-\frac{x}{\sigma}}\right) = \log\left(\theta_1(x)(1+e^{-x/\sigma})(e^{(x/\sigma)})\right).$$

From this, we get

$$\theta_1(x) = \frac{e^a e^{-\frac{x}{\sigma}}}{\left(1+e^{-\frac{x}{\sigma}}\right)^{b+1}}.$$

Using the expression of $\theta_1(x)$ in (3.2), we obtain the joint density of (X, Y) as

$$f_{X,Y}(x,y) = \frac{e^a e^{-x/\sigma} e^{-y/\sigma} e^{d\log(1+e^{-x/\sigma})\log(1+e^{-y/\sigma})}}{\left[\left(1+e^{-x/\sigma}\right)^{b+1}\left(1+e^{-y/\sigma}\right)^{c+1}\right]} \quad (x,y) \in \mathbb{R}^2. \quad (3.10)$$

The marginals densities can be obtained as follows. We have $\theta_1(x) = g(x)\delta_1(x)$. Next, using explicit expressions for $\theta_1(x)$ and $\delta_1(x)$ as obtained earlier, we get, the marginal density of X

$$g(x) = \frac{e^a e^{-x/\sigma}}{\left[c - d\log\left(1+e^{-\frac{x}{\sigma}}\right)\right]\left(1+e^{-x/\sigma}\right)^{b+1}} \quad x \in \mathbb{R}.$$

Similarly, from $\theta_2(y) = h(y)\delta_2(y)$, we obtain the marginal density of X

$$h(y) = \frac{e^a e^{-y/\sigma}}{\left[b - d\log\left(1+e^{-\frac{y}{\sigma}}\right)\right]\left(1+e^{-y/\sigma}\right)^{b+1}} \quad y \in \mathbb{R}.$$

Next, let us derive the normalizing constant e^a. We have the condition that

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} g(x)dx \\ &= e^a \int_{-\infty}^{\infty} \frac{e^{-x/\sigma}}{\left[c - d\log\left(1+e^{-\frac{x}{\sigma}}\right)\right]\left(1+e^{-x/\sigma}\right)^{b+1}} dx \\ &= e^a b d e^{-bc/d} Ei(bc/d), \end{aligned}$$

after some algebraic simplification, and $Ei(x) = \int_x^\infty \frac{e^{-u}}{u} du$.

Hence,

$$e^a = (bd)^{-1} e^{bc/d} \left(Ei(bc/d)\right)^{-1}.$$

So, our joint density will be (from (3.10))

$$f_{X,Y}(x,y) = \frac{(bd)^{-1} e^{bc/d} \left(Ei(bc/d)\right)^{-1} e^{-x/\sigma} e^{-y/\sigma} e^{d\log(1+e^{-x/\sigma})\log(1+e^{-y/\sigma})}}{\left[\left(1+e^{-x/\sigma}\right)^{b+1}\left(1+e^{-y/\sigma}\right)^{c+1}\right]},$$

$$(x,y) \in \mathbb{R}^2. \quad (3.11)$$

3.1. TP_2 Property. Let t_{11}, t_{12}, t_{21} and t_{22} be real numbers with $0 < t_{11} < t_{12}$ and $0 < t_{21} < t_{22}$. Then, (X, Y) is said to have the total positivity of order two (TP_2) property if for any such set of t_{ij}'s,

$$f_{X,Y}(t_{11}, t_{21}) f_{X,Y}(t_{12}, t_{22}) - f_{X,Y}(t_{12}, t_{21}) f_{X,Y}(t_{11}, t_{22}) \geq 0. \tag{3.12}$$

The positive quadrant dependence property (alternatively, the TP_2 property) implies the following:

- $P(X \leq x | Y = y)$ is non-increasing in y for all x,
- $P(Y \leq y | X = x)$ is non-increasing in x for all y,
- $P(Y > y | X > x)$ is non-decreasing in x for all y,
- $P(Y \leq y | X \leq x) \geq P(Y \leq y) P(X \leq x)$,
- $P(Y > y | X > x) \geq P(Y > y) P(X > x)$.

Next, we can rewrite the bivariate density in (3.11) as

$$f(x, y) = C r_1(x) r_2(y) \exp\left(\vec{q}_1(x) M_1 \vec{q}_2(y)\right),$$

where $C = (bd)^{-1} \left(Ei(bc/d)\right)^{-1}$, $r_1(x) = \exp(-x/\sigma)$, $r_2(y) = (1 + \exp(-y/\sigma))^{-1}$, $\vec{q}_1(x) = (1 \;\; \log(1 + \exp(-x/\sigma))$, $\vec{q}_2(y) = (1 \;\; \log(1 - \exp(-y/\sigma))$, and

$$M_1 = \begin{bmatrix} bc/d & -(c+1) \\ -(b+1) & d \end{bmatrix}.$$

Then, according to Arnold et al. (1999), based on the elements of the matrix M_1, we can have the following different scenarios:

- One will observe positive correlation if $bc/d^2 > (b+1)(c+1)$,
- One will observe negative correlation if $bc/d^2 < (b+1)(c+1)$,
- The distributions of X and Y will be independent if $d = 0$.

4 Bivariate Type II Generalized Logistic Distribution

Suppose we want for each fixed x,

$$f(y | X = x) = \frac{b_1(x) e^{-b_1(x) y}}{\left(1 + e^{-b_1(x) y}\right)^{b_1(x)+1}}, \quad y \in \mathbb{R},$$

and also for each fixed y,

$$f(x|Y=y) = \frac{b_2(y)e^{-b_2(y)x}}{\left(1+e^{-b_2(y)x}\right)^{b_2(y)+1}}, \quad y \in \mathbb{R},$$

where $b_1(x)$ is a function depending on x and $b_2(y)$ is a function depending on y and that both are unknown. We are now interested in identifying the class of all bivariate distributions for which both conditionals are Type II generalized logistic of the forms mentioned above. Now, let $g(x)$ and $h(y)$ be the marginal densities of X and Y, respectively. Then, one can write the joint density of (X, Y) as a product of marginals and conditionals, that is,

$$g(x)f(y|X=x) = h(y)f(x|Y=y), \tag{4.1}$$

or equivalently,

$$g(x)\frac{b_1(x)e^{-b_1(x)y}}{\left(1+e^{-b_1(x)y}\right)^{b_1(x)+1}} = h(y)\frac{b_2(y)e^{-b_2(y)x}}{\left(1+e^{-b_2(y)x}\right)^{b_2(y)+1}}. \tag{4.2}$$

Following the same steps as in the preceding section, the bivariate density in this case can be derived to be

$$f(x,y) = \frac{e^a e^{-bx-cy+dxy} e^{d\log(1+e^{-x})\log(1+e^{-y})}}{(1+e^{-x})^{b+1-dy}\,(1+e^{-y})^{c+1-dx}}, \quad (x,y) \in \mathbb{R}^2, \tag{4.3}$$

with the condition that $b > 0$, $c > 0$ and $d \le 0$. As before, a is the normalizing constant.

The corresponding marginals of X and Y are respectively given by

$$g(x) = \frac{e^a e^{-bx}}{\left[c - d\left(x + \log(1+e^{-x})\right)\right](1+e^{-x})^{b+1}} \quad x \in \mathbb{R}.$$

$$h(y) = \frac{e^a e^{-cy}}{\left[b - d\left(y + \log(1+e^{-y})\right)\right](1+e^{-y})^{c+1}} \quad y \in \mathbb{R}.$$

THEOREM 2. *The bivariate Type II generalized logistic distribution in* (4.3) *possesses the* TP_2 *property.*

PROOF. Let us consider different cases separately. If $0 < t_{11} < t_{21} < t_{12} < t_{22}$, then for the density function in (4.3), we can easily show that the condition in (3.12) is equivalent to

$$\exp\Big[d(t_{12}-t_{11})(t_{22}-t_{21}) + \left(\log(1+e^{-t_{12}}) - \log(1+e^{-t_{11}})\right) \\ \times \left(\log(1+e^{-t_{22}}) - \log(1+e^{-t_{21}})\right)\Big] \ge 0. \tag{4.4}$$

Observe that, since $0 < t_{11} < t_{12}$, the expression in (4.4) is true. Hence, the inequality in (3.12) is true if $0 < t_{11} < t_{12}$ and $0 < t_{21} < t_{22}$. All other cases can be similarly handled, and hence the theorem.

5 Bivariate Type III Generalized Logistic Distribution

In this case, by following the same lines as in the case of Type I, the bivariate Type III generalized logistic distribution (with both conditionals being of Type III) can be shown to be of the form

$$f(x,y) = \frac{e^{a} e^{-bx-cy+dxy} \left(1+e^{-x}\right)^{2dy}}{\left(1+e^{-2x}\right)^{2b} \left(1+e^{-y}\right)^{2c-2dx-4d\log(1+e^{-x})}}, (x,y) \in \mathbb{R}^2. \qquad (5.1)$$

THEOREM 3. *The bivariate Type III generalized logistic distribution in* (5.1) *possesses the* TP_2 *property.*

PROOF. Let us consider different cases separately. If $0 < t_{11} < t_{21} < t_{12} < t_{22}$, then for the density function in (5.1), we can easily show that the condition in (3.12) is equivalent to

$$\exp\left[d(t_{12}-t_{11})(t_{22}-t_{21})\right] \times \left[(e^{-t_{11}} - e^{-t_{12}})(e^{-t_{21}} - e^{-t_{22}})\right] \geq 0. \qquad (5.2)$$

Observe that, for $0 < t_{11} < t_{21} < t_{12} < t_{22}$, the expression in (5.2) is true. Hence, the inequality in (3.12) is true if $0 < t_{11} < t_{12}$ and $0 < t_{21} < t_{22}$. All other cases can be similarly handled, and hence the theorem.

The marginal density of X is given by

$$g(x) = \frac{e^{a} e^{-bx}}{\left(1+e^{-2x}\right)^{2b}}, \quad x \in \mathbb{R}.$$

Similarly, the marginal density of Y will be

$$h(y) = \frac{e^{a} e^{-cy}}{\left(1+e^{-2y}\right)^{2c}}, \quad y \in \mathbb{R}.$$

5.1. Shape of the Distribution. A critical point of a function with two variables is a point where the partial derivatives of first order are equal to zero. There are two main reasons as to why one would be interested in finding the critical points of a bivariate probability distribution: (1) To determine the shape of the distribution in order to examine its flexibility in fitting an observed phenomenon exhibiting same pattern, and (2) in bivariate distributions, the property of peakedness is important. Peakedness is a descriptive index of a distribution that provides an indication of concentration. It is effectively a measure of the fatness of the tails of the density function. We want to identify the location of the unique mode function for the given bivariate probability distribution. Let us now consider the shape of the bivariate Type I generalized logistic density in (3.11). The first derivative of $\log f(x,y)$ with respect to x and y are given by

$$\frac{\partial f(x,y)}{\partial x} = -\frac{\left(e^{-\frac{x}{\sigma}}+1\right)^{-b}\left(e^{-\frac{y}{\sigma}}+1\right)^{-c} e^{a+d\log\left(e^{-\frac{x}{\sigma}}+1\right)\log\left(e^{-\frac{y}{\sigma}}+1\right)}\left(-b+d\log\left(e^{-\frac{y}{\sigma}}+1\right)+e^{x/\sigma}\right)}{\sigma\left(e^{x/\sigma}+1\right)^2\left(e^{y/\sigma}+1\right)}.$$

$$\frac{\partial f(x,y)}{\partial y} = -\frac{\left(e^{-\frac{x}{\sigma}}+1\right)^{-b}\left(e^{-\frac{y}{\sigma}}+1\right)^{-c} e^{a+d\log\left(e^{-\frac{x}{\sigma}}+1\right)\log\left(e^{-\frac{y}{\sigma}}+1\right)}\left(-c+d\log\left(e^{-\frac{x}{\sigma}}+1\right)+e^{y/\sigma}\right)}{\sigma\left(e^{x/\sigma}+1\right)\left(e^{y/\sigma}+1\right)^2}$$

So, for this model, there could be several critical points.

Next, for the Type II bivariate logistic distribution in (4.3), the first derivative of $\log f(x,y)$ with respect to x and y, yield

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= \left(e^{-x}+1\right)^{dy-b}\left(e^{-y}+1\right)^{dx-c}\left(b\left(-e^{x}\right)+de^{x}y+de^{x}\log\left(e^{-y}+1\right)+1\right)\\ &\quad\times\left[\exp\left(a-bx-cy+dxy+d\log\left(e^{-x}+1\right)\log\left(e^{-y}+1\right)+x+y\right)\right]\\ &\quad\times\left[\left(e^{x}+1\right)^2\left(e^{y}+1\right)\right]^{-1},\end{aligned}$$

$$\begin{aligned}\frac{\partial f(x,y)}{\partial y} &= \left(e^{-x}+1\right)^{dy-b}\left(e^{-y}+1\right)^{dx-c}\left(c\left(-e^{y}\right)+dxe^{y}+de^{y}\log\left(e^{-x}+1\right)+1\right)\\ &\quad\times\left[\exp\left(a-bx-cy+dxy+d\log\left(e^{-x}+1\right)\log\left(e^{-y}+1\right)+x+y\right)\right]\\ &\quad\times\left[\left(e^{x}+1\right)\left(e^{y}+1\right)^2\right]^{-1}.\end{aligned}$$

Similarly, for this model, there could be several critical points.

Next, for the bivariate Type III generalized logistic density in (5.1), the first derivative of $\log f(x,y)$ with respect to x and y, yield

$$\begin{aligned}\frac{\partial f(x,y)}{\partial x} &= \left(e^{-2x}+1\right)^{-2b}\left(e^{-x}+1\right)^{2dy}\\ &\quad\left(-b\left(-3e^{x}+e^{2x}+e^{3x}-3\right)+d\left(e^{x}-e^{2x}+e^{3x}-1\right)y\right.\\ &\quad\left.+d\left(e^{x}+5e^{2x}+e^{3x}+5\right)\log\left(e^{-y}+1\right)\right)\\ &\quad\times\left(e^{-y}+1\right)^{-2c+dx-4d\log\left(e^{-x}+1\right)}e^{a-bx+y(-c+dx+2)}\\ &\quad\times\left(e^{x}+1\right)\left(e^{2x}+1\right)\left(e^{y}+1\right)^{-2}.\end{aligned}$$

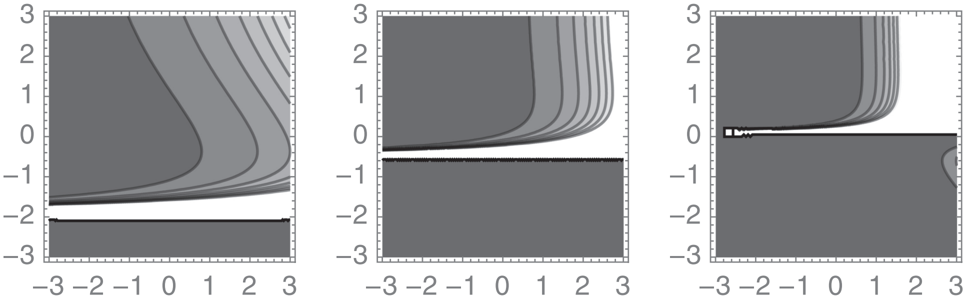

Figure 1: The Bivariate Type I generalized logistic contour plot for $\sigma = (4, 0.5, 1), a = (3, 0.5, 1), b = (5, 0.5, 1), c = (4, 0.6, 1), d = (3.5, 0.8, 1)$

$$\frac{\partial f(x,y)}{\partial y} = \left(e^{-2x}+1\right)^{-2b}\left(e^{-x}+1\right)^{2dy}\left(e^{-y}+1\right)^{-2c+dx-4d\log\left(e^{-x}+1\right)}$$
$$\times\left(c\left(-e^{y}\right)+c+dxe^{y}+2d\left(e^{y}+3\right)\log\left(e^{-x}+1\right)+2\right)e^{a-bx+y(-c+dx+2)}$$
$$\times\left(e^{y}+1\right)^{-3}.$$

In this case also, there will be several critical points. Some contour plots are provided in Figs. 1, 2, and 3 for the three bivariate generalized logistic distributions.

6 Likelihood Inference

In this section, we consider the maximum likelihood estimation of parameters for the three models derived in the preceding sections. The maximum likelihood estimators can be obtained by direct maximization of the

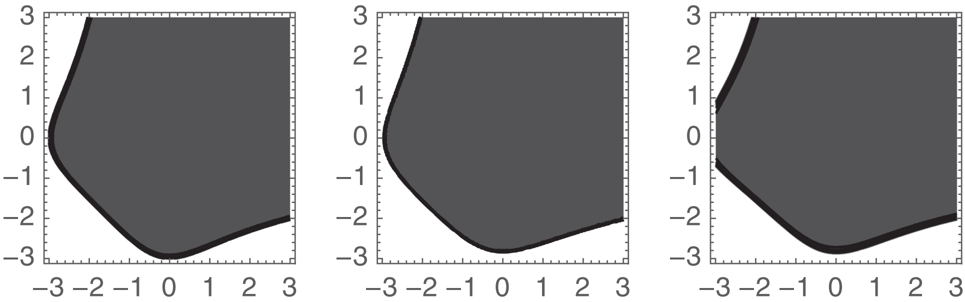

Figure 2: The Bivariate Type II generalized logistic contour plot for $a = (1, 1.5, 0.7), b = (1, 1, 2, 0.8), c = (1, 1.5, 0.5), d = (1, 2.5, 0.6)$

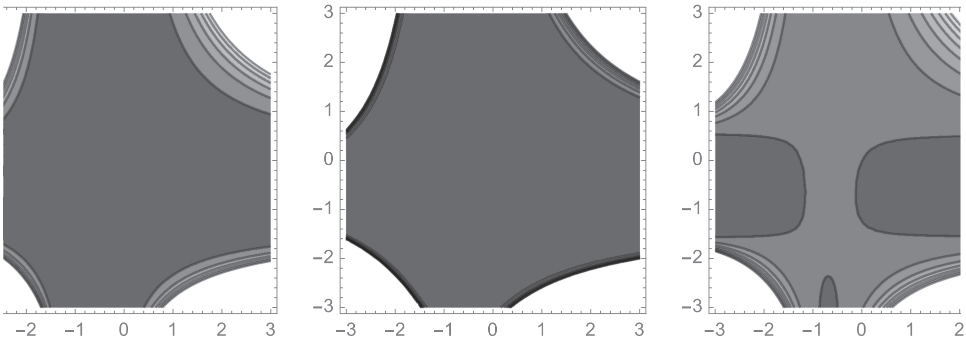

Figure 3: The Bivariate Type III generalized logistic contour plot for $a = (1, 1.5, 0.7), b = (1, 1, 2, 0.8), c = (1, 1.5, 0.5), d = (1, 2.5, 0.6)$

likelihood functions given below. In simulations and real life data applications described later on, we maximized the log-likelihood function using SAS PROC NLMIXED. For each maximization, the SAS PROC NLMIXED function was executed for a wide range of initial values, and the maximum likelihood estimates were determined as the ones that corresponds to the largest of the maxima.

6.1. Simulation Study for Bivariate Type I Generalized Logistic Distribution. In this case, the log likelihood function (after adding different location and different scale parameters for both X and Y) is given by (from (3.11))

$$\ell = \sum_{i=1}^{n} \log \left(\frac{(bd)^{-1} e^{bc/d} \left(Ei(bc/d)\right)^{-1} e^{-(x_i-\mu_1)/\sigma_1} e^{-(y_i-\mu_2)/\sigma_2} e^{d \log(1+e^{-(x_i-\mu_1)/\sigma_1}) \log(1+e^{-(y_i-\mu_2)/\sigma_2})}}{\left[\left(1+e^{-(x_i-\mu_1)/\sigma_1}\right)^{b+1} \left(1+e^{-(y_i-\mu_2)/\sigma_2}\right)^{c+1} \right]} \right). \tag{6.1}$$

Next, to demonstrate the feasibility of the suggested estimation strategy, a small simulation study was undertaken. The simulation study was carried out with $(\mu_1, \mu_2, \sigma_1, \sigma_2, a, b, c, d)$ =$(2, 3, 1.5, 1.7, 1.5, 1.5, 1.5, 1.5)$, respectively and the process was repeated 10000 times. Three different sample sizes $n = 50$, 100 and 200 were considered. The bias (actual-estimate) and the standard deviation of the parameter estimates for the maximum likelihood estimates were determined from this simulation study and are as presented in Table 1.

Table 1: Bias and standard deviation of the parameter estimates

Parameter	Sample size ($n = 50$)	Sample size ($n = 100$)	Sample size ($n = 200$)
μ_1	0.1211(0.3548)	0.0769(0.2271)	0.0651(0.1718)
μ_2	0.2829(0.5673)	0.1743(0.3496)	0.0927(0.1154)
σ_1	0.0189(0.03943)	−0.0134(0.02723)	0.0135(0.01515)
σ_2	0.0962(0.0452)	0.08782(0.0389)	0.0545(0.0281)
a	0.00145(0.10916)	0.00114(0.05328)	0.00089(0.04132)
b	0.0132(0.05616)	−0.0117(0.03831)	0.0014(0.02014)
c	0.01371(0.0247)	0.00138(0.00109)	0.00062(0.0076)
d	0.00161(0.00282)	0.00118(0.00134)	0.00107(0.00108)

6.2. Simulation for Bivariate Type II Generalized Logistic Distribution. In this case the log-likelihood function is (from (4.3))

$$\ell = \sum_{i=1}^{n} \log \left(\frac{e^a e^{-b\left(\frac{x_i-\mu_1}{\sigma_1}\right)-c\left(\frac{y_i-\mu_2}{\sigma_2}\right)+d\left(\frac{x_i-\mu_1}{\sigma_1}\right)\left(\frac{y_i-\mu_2}{\sigma_1}\right)} e^{d\log\left[1+e^{-\left(\frac{x_i-\mu_1}{\sigma_1}\right)}\right]\log\left[1+e^{-\left(\frac{y_i-\mu_2}{\sigma_2}\right)}\right]}}{\left(1+e^{-\left(\frac{x_i-\mu_1}{\sigma_1}\right)}\right)^{b+1-d\left(\frac{y_i-\mu_2}{\sigma_2}\right)}\left(1+e^{-\left(\frac{y_i-\mu_2}{\sigma_2}\right)}\right)^{c+1-d\left(\frac{x_i-\mu_1}{\sigma_1}\right)}} \right). \tag{6.2}$$

Next, to illustrate the feasibility of the suggested estimation strategy, a small simulation study was undertaken. The simulation study was carried out for one representative set of parameters $(\mu_1, \mu_2, \sigma_1, \sigma_2, a, b, c, d)$ $=(2, 3, 1.5, 1.7, 2, 2.5, 1.4, 2.5)$ and the process was repeated 10000 times. Three different sample sizes $n = 50$, 100 and 200 were considered. The bias (actual-estimate) and the standard deviation of the parameter estimates for the

Table 2: Bias and standard deviation of the parameter estimates

Parameter	Sample size ($n = 50$)	Sample size ($n = 100$)	Sample size ($n = 200$)
μ_1	0.1463(0.5382)	0.0614(0.2345)	0.0437(0.1139)
μ_2	0.1543(0.4628)	−0.1321(0.1894)	0.0672(0.0933)
σ_1	0.0445(0.4321)	0.1138(0.2467)	0.0946(0.1264)
σ_2	0.0853(0.0643)	0.0779(0.0627)	0.0617(0.0358)
a	0.01791(0.05138)	0.01146(0.0493)	0.00774(0.01091)
b	0.05057(0.16442)	0.04085(0.09045)	0.02942(0.04345)
c	−0.6818(0.0849)	0.5902(0.0668)	0.3312(0.03087)
d	0.0877(0.0068)	−0.0679(0.00405)	0.0419(0.04892)

maximum likelihood estimates were determined from this simulation study and are presented in Table 2.

6.3. Simulation for bivariate Type III generalized logistic distribution. In this case, the log-likelihood function (after location and scale parameters for both X and Y) is (from (5.1))

$$\ell = \sum_{i=1}^{n} \log \left(\frac{e^a e^{-b\left(\frac{x_i-\mu_1}{\sigma_1}\right)-c\left(\frac{y_i-\mu_2}{\sigma_2}\right)+d\left(\frac{x_i-\mu_1}{\sigma_1}\right)\left(\frac{y_i-\mu_2}{\sigma_2}\right)} \left(1+e^{-\left(\frac{x_i-\mu_1}{\sigma_1}\right)}\right)^{2d\left(\frac{y_i-\mu_2}{\sigma_2}\right)}}{\left(1+e^{-2\left(\frac{x_i-\mu_1}{\sigma_1}\right)}\right)^{2b} \left(1+e^{-\left(\frac{y_i-\mu_2}{\sigma_2}\right)}\right)^{2c-2d\left(\frac{x_i-\mu_1}{\sigma_1}\right)-4d\log\left[1+e^{-\left(\frac{x_i-\mu_1}{\sigma_1}\right)}\right]}} \right). \tag{6.3}$$

Next, to illustrate the feasibility of the suggested estimation strategy, a small simulation study was undertaken. The simulation study was carried out for one representative set of parameters $(\mu_1, \mu_2, \sigma_1, \sigma_2, a, b, c, d)$ $=(2, 3, 1.5, 1.7, 2, 2.5, 1.4, 2.5)$ respectively and the process was repeated 10000 times. Three different sample sizes $n = 50$, 100 and 200 were considered. The bias (actual-estimate) and the standard deviation of the parameter estimates for the maximum likelihood estimates were determined from this simulation study and are presented in Table 3.

Table 3: Bias and standard deviation of the parameter estimates

Parameter	Sample size ($n = 50$)	Sample size ($n = 100$)	Sample size ($n = 200$)
μ_1	$-0.2418(0.3316)$	$-0.1618(0.2754)$	$0.0638(0.2147)$
μ_2	$0.1876(0.6765)$	$0.1428(0.4898)$	$0.0939(0.2785)$
σ_1	$-0.1130(0.2318)$	$0.0849(0.1793)$	$0.0623(0.1135)$
σ_2	$0.0693(0.0885)$	$0.0427(0.0591)$	$0.0218(0.0426)$
a	$0.-0.4569(0.55358)$	$-0.4250(0.54217)$	$0.1783(0.50643)$
b	$0.33512(0.30734)$	$0.29432(0.22952)$	$0.24597(0.13618)$
c	$0.27602(0.87672)$	$0.25763(0.72941)$	$-0.1392(0.68934)$
d	$-0.2882(1.0321)$	$0.2077(0.9867)$	$-0.1851(0.4522)$

7 Application

For fitting purposes, we consider the full model for all the three forms of bivariate generalized logistic models as follows:

1. Bivariate Type I generalized logistic distribution:

$$f_{X,Y}(x,y) = \frac{(bd)^{-1} e^{bc/d} \left(Ei(bc/d)\right)^{-1} e^{-(x-\mu_1)/\sigma_1} e^{-(y-\mu_2)/\sigma_2} e^{d \log(1+e^{-(x-\mu_1)/\sigma_1}) \log(1+e^{-(y-\mu_2)/\sigma_2})}}{\left[(1+e^{-(x-\mu_1)/\sigma_1})^{b+1} (1+e^{-(y-\mu_2)/\sigma_2})^{c+1}\right]},$$

$(x,y) \in \mathbb{R}^2$.

2. Bivariate Type II generalized logistic distribution:

$$f(x,y) = \frac{e^a e^{-b\left(\frac{x-\mu_1}{\sigma_1}\right) - c\left(\frac{y-\mu_2}{\sigma_2}\right) + d\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_1}\right)} e^{d \log\left[1+e^{-\left(\frac{x-\mu_1}{\sigma_1}\right)}\right] \log\left[1+e^{-\left(\frac{y-\mu_2}{\sigma_2}\right)}\right]}}{\left(1+e^{-\left(\frac{x-\mu_1}{\sigma_1}\right)}\right)^{b+1-d\left(\frac{y-\mu_2}{\sigma_2}\right)} \left(1+e^{-\left(\frac{y-\mu_2}{\sigma_2}\right)}\right)^{c+1-d\left(\frac{x-\mu_1}{\sigma_1}\right)}},$$

$(x,y) \in \mathbb{R}^2$.

3. Bivariate Type III generalized logistic distribution:

$$f(x,y) = \frac{e^a e^{-b\left(\frac{x-\mu_1}{\sigma_1}\right) - c\left(\frac{y-\mu_2}{\sigma_2}\right) + d\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)} \left(1+e^{-\left(\frac{x-\mu_1}{\sigma_1}\right)}\right)^{2d\left(\frac{y-\mu_2}{\sigma_2}\right)}}{\left(1+e^{-2\left(\frac{x-\mu_1}{\sigma_1}\right)}\right)^{2b} \left(1+e^{-\left(\frac{y-\mu_2}{\sigma_2}\right)}\right)^{2c-2d\left(\frac{x-\mu_1}{\sigma_1}\right)-4d\log\left[1+e^{-\left(\frac{x-\mu_1}{\sigma_1}\right)}\right]}},$$

$(x,y) \in \mathbb{R}^2$.

For the purpose of illustration of the fitting of the proposed models, we consider the data set given in Table 5.2 in Johnson and Wichern (2007, p.228). It represents the scores obtained by $n = 87$ students on the College

Table 4: Parameter estimates for the college test data

Distribution	Bivariate Type I logistic	Bivariate Type II logistic	Bivariate Type III logistic
Parameter estimates	$\hat{\mu_1} = 538.191(7.421)$	$\hat{\mu_1} = 548.027(8.534)$	$\hat{\mu_1} = 532.237(11.667)$
	$\hat{\mu_2} = 59.513(11.163)$	$\hat{\mu_2} = 58.782(23.39)$	$\hat{\mu_2} = 61.174(29.56)$
	$\hat{\sigma_1} = 15.126(8.679)$	$\hat{\sigma_1} = 14.583(3.6504)$	$\hat{\sigma_1} = 12.977(3.284)$
	$\hat{\sigma_2} = 21.426(18.345)$	$\hat{\sigma_2} = 26.412(13.442)$	$\hat{\sigma_2} = 25.689(21.981)$
	$\hat{a} = 2.311(4.132)$	$\hat{a} = 6.927(17.245)$	$\hat{a} = 7.893(32.456)$
	$\hat{b} = 8.29(2.145)$	$\hat{b} = 43.372(5.764)$	$\hat{b} = 45.377(5.859)$
	$\hat{c} = 4.69(2.064)$	$\hat{c} = 22.666(3.956)$	$\hat{c} = 20.077(3.898)$
	$\hat{d} = 5.13(1.978)$	$\hat{d} = 16.363(8.678)$	$\hat{d} = 16.081(9.326)$
Log- likelihood	-673.441	-686.172	-681.556
AIC	867.195	872.789	875.0013
χ^2 p-value	0.7321	0.6337	0.6548

Level Examination Program (CLEP) subtest X and the College Qualification Test (CQT) subtest Y. Specifically, X represents score on social science and history and Y on scores on verbal ability. The maximum likelihood estimates of the parameters for each the above three bivariate generalized logistic models were obtained by the numerical method discussed earlier. The estimates, standard errors, AIC, K-S and the p-value of the χ^2 goodness of fit were all determined for all three models and these are provided in Table 4.

The obtained results reveal that bivariate Type I generalized logistic model provides the best fit for the data among the three models considered.

8 Concluding Remarks

The concept of conditional specification of bivariate distributions is not new but, except in normal and exponential families (see Arnold et al., 1999), it has not been well developed for other distributions in the literature. Computational difficulties, absence of analytically tractable forms of the marginal as well as bivariate densities, without a doubt discouraged further work in this direction. In this paper, we have addressed the characterization of three different types of bivariate generalized logistic distributions (Type I, Type II and Type III), starting from two given conditional densities, namely, of Y given X and X given Y, both belonging to the same family of generalized logistic distributions. One can extend the conditional specification idea to the multivariate construction case, but, the analytical tractability may need to be examined carefully. Moreover, their properties and methods of fit may also need to be studied in detail. We are hoping to take this up as our future research.

References

ARNOLD, B.C., CASTILLO, E. and SARABIA, J.M. (1999) *Conditional Specification of Statistical Models.* Springer, New York.

BALAKRISHNAN, N. (ed.) (1992) *Handbook of the Logistic Distribution.* Marcel Dekker, New York.

BALAKRISHNAN, N. and LAI, C.D. (2009) *Continuous Bivariate Distributions,* Second edition. Springer, New York.

GUMBEL, E.J. (1961) Bivariate logistic distributions. *Journal of the American Statistical Association*, 56, 335-349.

JOHNSON, R.A. and WICHERN, W.D. (2007) *Applied Multivariate Statistical Analysis*, Sixth edition. Prentice Hall, Englewood Cliffs, New Jersey.

KOTZ, S., BALAKRISHNAN, N. and JOHNSON, N.L. (2000) *Continuous Multivariate Distributions -Vol. 1,* Second edition. John Wiley & Sons, New York.

Indranil Ghosh
University of North Carolina,
Wilmington, NC, USA
E-mail: ghoshi@uncw.edu

N. Balakrishnan
McMaster University, Hamilton,
Ontario, Canada
E-mail: bala@mcmaster.ca

Paper received: 21 September 2015.