Gujarat Journal of Statistics and Data Science
GUJARAT STATISTICAL ASSOCIATION

Univariate and Bivariate Discrete Laplace and Generalized Discrete Laplace Distributions

Barry C. Arnold[1] and Matthew A. Arvanitis[2]

1. *University of California, Riverside*
2. *USDA Forest Products Laboratory. Email: mathew@umbc.edu*

Received: 15 Jan 2022 / Revised: 22 Apr 2022 / Accepted: 05 Jun 2022

Abstract

Univariate discrete Laplace models are described and investigated. A more general mixture model is represented. Bivariate extensions of these models are discussed in some detail, with particular emphasis on associated parameter estimation strategies. Multivariate versions of the models are briefly introduced.

Keywords: Discrete Laplace, Asymmetric Discrete Laplace, Generalized Asymmetric Discrete Laplace, Bivariate Discrete Laplace, likelihood, method of moments.
AMS Subject Classification: 62E10

1 Introduction to univariate discrete Laplace models

Inusah and Kozubowski [4] define the univariate asymmetric discrete Laplace distribution to be one with characteristic function of the form:

$$\phi_X(t) = p_1p_2(1-(1-p_1)e^{it})^{-1}(1-(1-p_2)e^{-it})^{-1}, \tag{1}$$

where $p_1, p_2 \in (0,1)$.

An alternative description of the model is available and will be used. For it we begin with two independent geometric random variables, V_1 and V_2 with $V_i \sim geo(p_i)$, $i = 1,2$. We then define

$$X = V_1 - V_2. \tag{2}$$

If X has a representation of this form, we write $X \sim ADL(p_1,p_2)$. Note that in this paper, geometric random variables are defined to have support $\{0,1,2,...\}$, and can be thought of as representing the number of failures preceding the first success in series of Bernoulli trials.

The possible values of X are $\{..., -3. -2. -1, 0, 1, 2, 3, ...\}$.

For $x \geq 0$ we have

$$\begin{aligned} P(X=x) &= \sum_{v_2=0}^{\infty} P(V_1 = x+v_2, V_2 = v_2) \\ &= \sum_{v_2=0}^{\infty} p_1q_1^{x+v_2}p_2q_2^{v_2} \\ &= p_1p_2q_1^x(1-q_1q_2)^{-1} \end{aligned}$$

Analogously, for $x < 0$,

$$\begin{aligned} P(X = x) &= \sum_{v_1=0}^{\infty} P(V_1 = v_1, V_2 = v_1 - x) \\ &= \sum_{v_2=0}^{\infty} p_1 q_1^{v_1} p_2 q_2^{v_1 - x} \\ &= p_1 p_2 q_2^{-x} (1 - q_1 q_2)^{-1}. \end{aligned}$$

It is then not difficult to confirm that $\sum_{-\infty}^{\infty} P(X = x) = 1$

This distribution will be called the asymmetric discrete Laplace distribution (ADL). It is a discrete parallel to the asymmetric Laplace model $Y = U_1 - U_2$ where the U_i's are independent with $U_i \sim exp(\lambda_i), \quad i = 1, 2$.

A mixture alternative is suggested by consideration of available mixture representations of continuous asymmetric Laplace models. It includes an additional parameter for flexibility. See [5], [7], and [6] for relevant discussion on analogous asymmetric Laplace models. We thus will consider

$$Y = IV_1 + (1 - I)(-V_2), \tag{3}$$

where $V_i \sim geo(p_i)$, $i = 1, 2$, and I is an independent Bernoulli random variable with $P(I = 1) = \pi$.

It is readily verified that the characteristic function of such a generalized asymmetric discrete Laplace (GADL) variable of the form (3) is given by

$$\phi_Y(t) = \frac{1 - \pi(1 - p_1)e^{it} - (1 - \pi)(1 - p_2)e^{-it}}{1 - (1 - p_1)e^{it} - (1 - p_2)e^{-it} + (1 - p_1)(1 - p_2)}. \tag{4}$$

The density of this GADL distribution is of the following form.

$$\begin{aligned} P(Y = y) &= \pi p_1 q_1^{y}, \quad y = 1, 2, \ldots, \\ &= \pi p_1 + (1 - \pi) p_2, \quad y = 0, \\ &= (1 - \pi) p_2 q_2^{-y}, \quad = -1, -2, -3, \ldots, \end{aligned} \tag{5}$$

where p_1, p_2 and π are parameters ranging over the interval $(0, 1)$.

From the characteristic function, or from the mixture representation (3) we find

$$E(Y) = \frac{\pi(1 - p_1)}{p_1} - \frac{(1 - \pi)(1 - p_2)}{p_2}$$

and

$$var(Y) = \left\{ \frac{\pi(1 - p_1)(2 - p_1)}{p_1^2} + \frac{(1 - \pi)(1 - p_2)(2 - p_2)}{p_2^2} \right\} - \left\{ \frac{\pi(1 - p_1)}{p_1} - \frac{(1 - \pi)(1 - p_2)}{p_2} \right\}^2.$$

In the case in which $X \sim ADL(p_1, p_2)$ the moments simplify to become

$$E(X) = (1 - p_1)/p_1 - (1 - p_2)/p_2$$

and

$$var(X) = (1 - p_1)/p_1^2 + (1 - p_2)/p_2^2.$$

2 Estimation for the ADL distribution

Maximum likelihood

Suppose we have a sample $X_1, X_2, ..., X_n$ from an asymmetric discrete Laplace distribution defined in (2). We wish to estimate the parameters using maximum likelihood.

Define

$$U = \sum_1^n X_i I(X_i \geq 0) \text{ and } V = -\sum_1^n X_i I(X_i < 0).$$

Taking partial derivatives of the log-likelihood and equating to 0, yields the following likelihood equations

$$\frac{n}{p_1} = \frac{n(1-p_2)}{p_1+p_2-p_1p_2} + \frac{U}{1-p_1}, \tag{6}$$

$$\frac{n}{p_2} = \frac{n(1-p_1)}{p_1+p_2-p_1p_2} + \frac{V}{1-p_2}. \tag{7}$$

In order to solve these equations it is convenient to rewrite them in terms of the means of the V_i's, thus we define $\mu_i = (1-p_i)/p_i, \quad i = 1,2$. The likelihood equations then become

$$\mu_1 = \frac{\mu_1\mu_2}{1+\mu_1+\mu_2} + \frac{U}{n}, \tag{8}$$

$$\mu_2 = \frac{\mu_1\mu_2}{1+\mu_1+\mu_2} + \frac{V}{n}. \tag{9}$$

From these equations we deduce that

$$\mu_1 - \mu_2 = \frac{U}{n} - \frac{V}{n}. \tag{10}$$

Next, using (10), express μ_2 as a linear function of μ_1 and substitute this in equation (8). This upon rearranging is a quadratic equation in μ_1 which can be solved to yield the maximum likelihood estimate of μ_1, i.e.,

$$\widehat{\mu}_1 = \frac{U}{n} - \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{UV}{n^2}}, \tag{11}$$

and then

$$\widehat{\mu}_2 = \widehat{\mu}_1 - \frac{U}{n} + \frac{V}{n.}. \tag{12}$$

Method of moments

Suppose we have a sample $X_1, X_2, ..., X_n$ from an asymmetric discrete Laplace distribution represented in the form

$$X = V_1 - V_2,$$

where the V_i's are i.i.d. with $V_i \sim geo(p_i), \quad i = 1,2$.

We will equate the first two moments of X to the corresponding sample moments. Elementary computations yield

$$E(X) = \frac{1-p_1}{p_1} - \frac{1-p_2}{p_2}$$

and

$$E(X^2) = \frac{1-p_1}{p_1^2} + \frac{1-p_2}{p_2^2} + (E(X))^2$$

Denote the first two sample moments, based on a sample of size n, by $M_1 = (1/n)\sum_{i=1}^n X_i$ and $M_2 = (1/n)\sum_{i=1}^n X_i^2$, and define $S^2 = M_2 - M_1^2$ and set up the equations $M_1 = E(X)$ and $S^2 = var(X)$ which may be solved to obtain method of moments estimates of the p_i's. Here, as in the maximum likelihood case, it is convenient to rewrite the equations in terms of the means of the V_i's (i.e., $\mu_i = (1-p_i)/p_i, \quad i = 1, 2$). The moment equations are of the form

$$M_1 = \mu_1 - \mu_2, \tag{13}$$

$$S^2 = \mu_1(1+\mu_1) + \mu_2(1+\mu_2). \tag{14}$$

From equation (13) we can write $\mu_2 = \mu_1 - M_1$ and substitute this into (16). This then can be rearranged into a quadratic function of μ_1 which is readily solved. In this way we obtain method of moments estimates of the following form.

$$\widetilde{\mu}_1 = \frac{1}{2}\left[M_1 - 1 + \sqrt{1 - M_1^2 + 2M_1S^2}\right], \tag{15}$$

and

$$\widetilde{\mu}_2 = \widetilde{\mu}_1 - M_1. \tag{16}$$

Bayesian Method

Suppose we have a sample $X_1, X_2, ..., X_n$ from an asymmetric discrete Laplace distribution with density

$$f_X(x; p_1, p_2) = \left(\frac{p_1p_2}{p_1 + p_2 - p_1p_2}\right)(1-p_1)^{xI(x\geq 0)}(1-p_2)^{-xI(x<0)}$$

Define

$U = \sum_1^n X_i I(X_i > 0)$, $V = -\sum_1^n X_i I(X_i < 0)$ and $W = \sum_1^n I(X_i > 0)$

We wish to estimate the parameters from a Bayesian viewpoint.

The likelihood for the sample is given by

$$L(p_1, p_2) = \left(\frac{p_1p_2}{p_1 + p_2 - p_1p_2}\right)^n (1-p_1)^u(1-p_2)^v.$$

If we take independent beta priors for p_1 and p_2, i.e.,

$$\widetilde{p}_i \sim Beta(\alpha_i, \beta_i), \quad i = 1, 2,$$

then the posterior will be of the form

$$f(p_1, p_2|\underline{X} = \underline{x}) \propto \left(\frac{p_1^{n+\alpha_1}p_2^{n+\alpha_2}}{(p_1 + p_2 - p_1p_2)^n}\right)(1-p_1)^{u+\beta_1}(1-p_2)^{v+\beta_2}.$$

From this joint posterior density the usual Bayes estimates of p_1 and p_2, namely $E(p_i|\underline{X} = \underline{x}), \quad i = 1, 2$, will be obtained by numerical integration.

3 Estimation for the GADL distribution

Maximum likelihood

Suppose instead we have a sample $X_1, X_2, ..., X_n$ from a generalized asymmetric discrete Laplace distribution of the form

$$X = IW_1 + (1 - I)(-W_2)$$

where $I \sim Bernoulli(\pi)$ and the W_i's are independent with $W_i \sim geo(p_i)$, $i = 1, 2$.

We wish to estimate the parameters using maximum likelihood. Here too, we will define

$$U = \textstyle\sum_1^n X_i I(X_i \geq 0) \text{ and } V = -\sum_1^n X_i I(X_i < 0),$$

and also define $N_0 = \sum_1^n I(X_i = 0)$, $N_1 = \sum_1^n I(X_i > 0)$ and $N_2 = \sum_1^n I(X_i < 0)$

Taking partial derivatives of the log-likelihood and equating to 0, yields the following likelihood equations

$$\frac{N_1}{\pi} = \frac{N_2}{1-\pi} - \frac{N_0(p_1 - p_2)}{\pi p_1 + (1-\pi)p_2} \tag{17}$$

$$\frac{N_1}{p_1} = \frac{U}{1-p_1} - \frac{N_0 \pi}{\pi p_1 + (1-\pi)p_2} \tag{18}$$

$$\frac{N_2}{p_2} - \frac{V}{1=p_2} - \frac{N_0(1-\pi)}{\pi p_1 + (1-\pi)p_2} \tag{19}$$

Note these equations are particularly easy to solve if $N_0 = 0$. In other cases an iterative solution may be obtained. Note that if p_1 and p_2 are known, then equation (17) is equivalent to a quadratic equation in π. If p_1 and π are known, then equation (18) is equivalent to a linear equation in p_2. And, finally, if p_2 and π are known, then equation (19) is equivalent to a linear equation in p_1.

Method of moments

Suppose we have a sample $X_1, X_2, ..., X_n$ from a generalized asymmetric discrete Laplace distribution of the form

$$X = IW_1 + (1 - I)(-W_2)$$

where $I \sim Bernoulli(\pi)$ and the W_i's are independent with $W_i \sim geo(p_i)$, $i = 1, 2$.

We wish to estimate the parameters using the method of moments.

The first three moments about zero of X are:

$$\begin{aligned}
E(X) &= \pi \frac{1-p_1}{p_1} - (1-\pi)\frac{1-p_2}{p_2} \\
E(X^2) &= \pi \left\{ 2\left(\frac{1-p_1}{p_1}\right)^2 + \frac{1-p_1}{p_1} \right\} + (1-\pi)\left\{ 2\left(\frac{1-p_2}{p_2}\right)^2 + \frac{1-p_2}{p_2} \right\} \\
E(X^3) &= \pi \left\{ 6\left(\frac{1-p_1}{p_1}\right)^3 + 6\left(\frac{1-p_1}{p_1}\right)^2 + \left(\frac{1-p_1}{p_1}\right) \right\} \\
&\quad -(1-\pi)\left\{ 6\left(\frac{1-p_2}{p_2}\right)^3 + 6\left(\frac{1-p_2}{p_2}\right)^2 + \left(\frac{1-p_2}{p_2}\right) \right\}
\end{aligned}$$

If we denote the first three sample moments, based on a sample of size n, by $M_1 = (1/n)\sum_{i=1}^{n} X_i$, $M_2 = (1/n)\sum_{i=1}^{n} X_i^2$, and $M_3 = (1/n)\sum_{i=1}^{n} X_i^3$, and set up the equations $M_j = E(X^j)$, $\quad j = 1, 2, 3$, then our method of moments estimates will be the solution to these three equations.

Note that, if p_1 and p_2 are known then the equation $M_3 = E(X^3)$ is a linear equation in π. Also, if p_2 and π are known, then the equation $M_2 = E(X^2)$ is equivalent to a quadratic equation in p_1, and finally, if π and p_1 are known then the equation $M_1 = E(X)$ is equivalent to a linear equation in p_2. Using these observations, an iterative scheme for identifying the method of moments estimates is readily set up.

Bayesian Method

Suppose we have a sample $X_1, X_2, ..., X_n$ from a generalized asymmetric discrete Laplace distribution of the form

$$X = IW_1 + (1 - I)(-W_2)$$

where $I \sim Bernoulli(\pi)$ and the W_i's are independent with $W_i \sim geo(p_i)$, $i = 1, 2$.

Recall that we defined

$$U = \textstyle\sum_1^n X_i I(X_i \geq 0) \text{ and } V = -\sum_1^n X_i I(X_i < 0),$$

and also defined $N_0 = \sum_1^n I(X_i = 0)$, $N_1 = \sum_1^n I(X_i > 0)$ and $N_2 = \sum_1^n I(X_i < 0)$

We wish to estimate the parameters from a Bayesian viewpoint.

In this case the likelihood will be

$$L(\pi, p_1, p_2) = \pi^{n_1} p_1^{n_1} (1 - p_1)^u (1 - \pi)^{n_2} p_2^{n_2} (1 - p_2)^v [\pi p_1 + (1 - \pi)p_2]^{n_0}$$

A plausible prior with independent marginals will be of the form

$$\widetilde{\pi} \sim Beta(\tau_1, \tau_2), \text{ and } \widetilde{p_i} \sim Beta(\alpha_i, \beta_i), \quad i = 1, 2.$$

The corresponding joint posterior density will be

$$f(\pi, p_1, p_2 | \underline{X} = \underline{x}) \propto \pi^{n_1+\tau_1-1} p_1^{n_1+\alpha_1-1} (1 - p_1)^{u+\beta_1-1} (1 - \pi)^{n_2+\tau_2-1} p_2^{n_2+\alpha_2-1} (1 - p_2)^{v+\beta_2-1} \times [\pi p_1 + (1 - \pi)p_2]^{n_0}. \quad (20)$$

If $n_0 = 0$ then the posterior density will have independent beta distributed marginals. If $n_0 > 0$ then a posterior density that is a mixture of distributions with independent marginals will be encountered.

4 Bivariate models

In [1], two bivariate asymmetric Laplace models are described. The first bivariate asymmetric Laplace model was introduced by [3] and we refer the reader to that source for detailed discussion of the model. Construction of the model begins with the components used in developing the general bivariate beta model introduced in [2]. Thus we begin with 8 independent gamma variables $U_1, U_2, ..., U_8$ with $U_j \sim \Gamma(\delta_j, 1)$, $\quad j = 1, 2, ..., 8$.. We then define (X, Y) by

$$\begin{aligned} X &= \lambda_{11}^{-1}(U_1 + U_5 + U_7) - \lambda_{12}^{-1}(U_3 + U_6 + U_8), \\ Y &= \lambda_{21}^{-1}(U_2 + U_6 + U_7) - \lambda_{22}^{-1}(U_4 + U_5 + U_8), \end{aligned} \quad (21)$$

where it is assumed that the constraints, $\delta_1 + \delta_5 + \delta_7 = 1$, $\delta_3 + \delta_6 + \delta_8 = 1$, $\delta_2 + \delta_6 + \delta_7 = 1$, and $\delta_4 + \delta_5 + \delta_8 = 1$, have been imposed to ensure that the distribution has asymmetric Laplace marginals. This model will be called the bivariate asymmetric Laplace model of the first kind and if (X, Y) is as defined in (21) we will write $(X, Y) \sim BAL(1)(\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22}, \underline{\delta})$. Since there were four constraints on the δ_j's, this is an 8 parameter model. The marginal distributions depend only on the four λ parameters, thus:

$$X \sim AL(\lambda_{11}, \lambda_{12}), \quad Y \sim AL(\lambda_{21}, \lambda_{22}). \tag{22}$$

The second bivariate asymmetric Laplace model utilizes the closure under minimization property of the exponential distribution. For it we again begin with 8 independent random variables, $V_1, V_2, ..., V_8$ but this time we assume that they are exponentially distributed, thus $V_j \sim exp(\lambda_j), \quad j = 1, 2, ..., 8$. We then define

$$\begin{aligned} X &= min\{V_1, V_5, V_7\} - min\{V_3, V_6, V_8\}, \\ Y &= min\{V_2, V_6, V_7\} - min\{V_4, V_5, V_8\}. \end{aligned} \tag{23}$$

If (X, Y) has the structure shown in (23) then we will write $(X, Y) \sim BAL(II)(\underline{\lambda})$ and say that it has a bivariate asymmetric Laplace distribution of the second kind with parameter vector $\underline{\lambda}$. Note that both the first kind and the second kind bivariate asymmetric Laplace distributions have an 8 dimensional parameter space. The marginal distributions of the BAL(II) distribution are by construction of the asymmetric Laplace form. Thus:

$$X \sim AL(\lambda_1 + \lambda_5 + \lambda_7, \lambda_3 + \lambda_6 + \lambda_8), \tag{24}$$
$$Y \sim AL(\lambda_2 + \lambda_6 + \lambda_7, \lambda_4 + \lambda_5 + \lambda_8), \tag{25}$$

Discrete versions of the BAL(I)-(II) distributions will now be constructed using negative binomial and geometric components instead of gamma and exponential distributed components. But, before discussing such distributions, we will consider bivariate versions of the generalized asymmetric Laplace (GAL) distribution. This can be achieved by modifying the BAL(I) or the BAL(II) models by the introduction of two additional probability parameters.

The generalized version of the BAL(I) model may be defined as follows

$$\begin{aligned} X &= I_1\lambda_{11}^{-1}(U_1 + U_5 + U_7) - (1 - I_1)\lambda_{12}^{-1}(U_3 + U_6 + U_8), \\ Y &= I_2\lambda_{21}^{-1}(U_2 + U_6 + U_7) - (1 - I_2)\lambda_{22}^{-1}(U_4 + U_5 + U_8), \end{aligned} \tag{26}$$

where it is assumed that the constraints, $\delta_1 + \delta_5 + \delta_7 = 1$, $\delta_3 + \delta_6 + \delta_8 = 1$, $\delta_2 + \delta_6 + \delta_7 = 1$, and $\delta_4 + \delta_5 + \delta_8 = 1$, have been imposed, and where the I_j's are independent with $I_j \sim Bernoulli(p_j), \ j = 1, 2$.

Means variances and covariance of the coordinates of this random vector are not difficult to evaluate, or could be evaluated by simulation.

The generalized version of the BAL(II) model may be defined as follows

$$\begin{aligned} X &= I_1[min\{(V_1, V_5, V_7\}] - (1 - I_1)[min\{V_3, V_6, V_8\}], \\ Y &= I_2[min\{(V_2, V_6, V_7\}] - (1 - I_2)[min\{V_4, V_5, V_8)\}], \end{aligned} \tag{27}$$

where the δ_j's are positive parameters, and where the I_j's are independent with $I_j \sim Bernoulli(p_j)$, $j = 1, 2$.

Means variances and covariance of the coordinates of this random vector are also not difficult to evaluate, or could be evaluated by simulation.

We now will define analogous discrete versions of these bivariate models. To construct the bivariate symmetric discrete Laplace model of the first kind (BDL(I)) we begin with a set of 8 independent random variables $U_1, U_2, ..., U_8$ with $U_i \sim neg.bin.(\delta_j, p)$, $j = 1, 2, ..., 8$. Note that all of the U_j's share a common value for p. This results in a construction of a bivariate distribution with symmetric discrete Laplace marginals. It will be seen that it is not possible to use this kind of construction to yield asymmetric marginals. To continue, we now define (X, Y) by

$$\begin{aligned} X &= (U_1 + U_5 + U_7) - (U_3 + U_6 + U_8), \\ Y &= (U_2 + U_6 + U_7) - (U_4 + U_5 + U_8), \end{aligned} \tag{28}$$

where it is assumed that the constraints, $\delta_1 + \delta_5 + \delta_7 = 1$, $\delta_3 + \delta_6 + \delta_8 = 1$, $\delta_2 + \delta_6 + \delta_7 = 1$, and $\delta_4 + \delta_5 + \delta_8 = 1$, have been imposed. This model will be called the bivariate discrete Laplace model of the first kind and if (X, Y) is as defined in (28) we will write $(X, Y) \sim BDL(1)(\underline{\delta})$. Since there were four constraints on the δ_j's, this is a 5 parameter model. The marginal distributions are differences of independent $geometric(p)$ variables and thus have discrete Laplace densities. Moments are obtainable from the representation (28).

The second bivariate asymmetric discrete Laplace model that we will consider will utilize the closure under minimization property of the geometric distribution. For it we again begin with 8 independent random variables, $V_1, V_2, ..., V_8$ but this time we assume that they are geometrically distributed, thus $V_j \sim geo(\tau_j)$, $j = 1, 2, ..., 8$, where $\tau_j \in (0, 1)$, $j = 1, 2, ..., 8$, We then define

$$\begin{aligned} X &= min\{V_1, V_5, V_7\} - min\{V_3, V_6, V_8\}, \\ Y &= min\{V_2, V_6, V_7\} - min\{V_4, V_5, V_8\}, \end{aligned} \tag{29}$$

using a construction parallel to that used in the construction of the BAL(II) model earlier described in this paper. If (X, Y) has the structure shown in (29) then we will write $(X, Y) \sim BADL(II)(\underline{\tau})$ and say that it has a bivariate asymmetric discrete Laplace distribution of the second kind with parameter vector $\underline{\tau}$. Note that the second kind bivariate asymmetric discrete Laplace distributions has an 8 dimensional parameter space. The marginal distributions of the BADL(II) distribution are of the asymmetric discrete Laplace form. Thus:

$$X \sim ADL(1 - (1 - \tau_1)(1 - \tau_5)(1 - \tau_7), 1 - (1 - \tau_3)(1 - \tau_6)(1 - \tau_8)), \tag{30}$$

$$Y \sim ADL(1 - (1 - \tau_2)(1 - \tau_6)(1 - \tau_7), 1 - (1 - \tau_4)(1 - \tau_5)(1 - \tau_8)), \tag{31}$$

The marginal moments of the BADL(II) distribution are thus readily identified. However $cov(X, Y)$ is quite complicated and will usually be approximated by simulation. using the definition (29).

Generalized versions of these bivariate discrete Laplace models can be formulated in a manner parallel to that used to generalize the bivariate asymmetric Laplace models.

The generalized version of the BDL(I) model may be defined as follows

$$X = I_1(U_1 + U_5 + U_7) - (1 - I_1)(U_3 + U_6 + U_8),$$
$$Y = I_2(U_2 + U_6 + U_7) - (1 - I_2)(U_4 + U_5 + U_8), \tag{32}$$

where it is assumed that the constraints, $\delta_1 + \delta_5 + \delta_7 = 1$, $\delta_3 + \delta_6 + \delta_8 = 1$, $\delta_2 + \delta_6 + \delta_7 = 1$, and $\delta_4 + \delta_5 + \delta_8 = 1$, have been imposed, and where the I_j's are independent with $I_j \sim Bernoulli(\pi_j)$, $j = 1, 2$, and the U_j's are independent with $U_j \sim neg.bin.(\delta_j, p)$, $j = 1, 2, 3..., 8$. This model will be called the generalized bivariate discrete Laplace model of the first kind and if (X, Y) is as defined in (32) we will write $(X, Y) \sim GBDL(1)(\pi_1, \pi_2, \underline{\delta})$. Since there were four constraints on the δ_j's, this is a 7 parameter model.

Means variances and covariance of the coordinates of this random vector are not difficult to evaluate, or could be evaluated by simulation.

The generalized version of the BADL(II) model may be defined as follows

$$X = I_1[min\{(V_1, V_5, V_7\}] - (1 - I_1)[min\{V_3, V_6, V_8\}],$$
$$Y = I_2[min\{V_2, V_6, V_7\}] - (1 - I_2)[min\{V_4, V_5, V_8)\}], \tag{33}$$

where the I_j's are independent with $I_j \sim Bernoulli(\pi_j)$, $j = 1, 2$. and the V_j's are independent with $V_j \sim geo(\tau_j)$, $j = 1, 2, ..., 8$. This is a 10 parameter model.

Means variances and covariance of the coordinates of this random vector are also not difficult to evaluate, or could be evaluated by simulation.

Remark Even more general distribution than (32) can be constructed in which the I_i's are dependent indicators.

For this more general model we begin with a random vector (J_1, J_2, J_3, J_4) with 4 possible values $(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0)$ and $(0, 0, 0, 1)$ with associated probabilities π_1, π_2, π_3 and π_4, and then define $I_1 = max\{J_1, J_3\}$ and $I_2 = max\{J_2, J_3\}$.

The random variables (X, Y) are then defined as in (32) using the dependent I_i's just defined.

The same modification can be made to generalize (33).

Parameter estimation

None of the full models described in this paper are expected to be useful for practical purposes. Instead the authors expect that they will be used as a source for smaller, more manageable, submodels. For all four of the bivariate models discussed, only a small collection of specific cases may be described which have explicit discrete densities, and even these densities are rather complex. Therefore, unconventional methods for parameter estimation may be called for.

For illustrative purposes, we will consider parameter estimation for some submodels with significantly reduced parameter spaces. In describing submodels we adhere to the following conventions : (1) If $U \sim neg.bin(\delta, p)$ with $\delta = 0$ then $U = 0$ with probability 1. Consequently, such U's can be deleted from the description of the bivariate models BDL(I) and GBDL(I) as described in this paper.

(2) If $V \sim geo(\tau)$ with $\tau = 0$ then $V = \infty$ with probability 1. Consequently such V's can be deleted from the description of the bivariate models BADL(II) and GBADL(II) as described in this paper.

Some two-parameter submodels

In this subsection, we will consider parameter estimation for a pair of simple submodels of the $BDL(I)$ and $BADL(II)$ models.

Example 1. *To begin, consider the sub-model of the $BDL(I)$ model given by restricting its parameter space as follows:*

$$\begin{aligned} &p \in (0,1) \\ &\delta_1 = \delta_4 = \alpha \in (0,1) \\ &\delta_2 = \delta_3 = 1 \\ &\delta_5 = 1 - \alpha \\ &\delta_6 = \delta_7 = \delta_8 = 0. \end{aligned} \tag{34}$$

We will call this model M1. We are adopting the convention that if $U \sim neg.bin.(\delta, p)$ with $\delta = 0$ then $U = 0$. The full parameter vector for this two parameter model is thus

$$(p, \underline{\delta}) = (p, (\alpha, 1, 1, \alpha, 1-\alpha, 0, 0, 0)).$$

Suppose that we have a sample of size n, i.e., $\{(X_j, Y_j) : j = 1, 2, ..., n\}$ and we wish to estimate the parameters p and α. Upon writing this model in terms of the U_i variables, thus

$$\begin{aligned} X &= (U_1 + U_5) - (U_3), \\ Y &= (U_2) - (U_4 + U_5), \end{aligned} \tag{35}$$

it may be verified that $E(X) = E(Y) = 0$, $var(X) = var(Y) = 2(1-p)/p^2$ and $cov(X,Y) = -(1-\alpha)(1-p)/p^2$. Consequently, if we define

$$T_1 = (1/n)[\sum_{i=1}^{n}[(X_i - \overline{X})^2 + (Y_i - \overline{Y})^2]$$

and

$$T_2 = (1/n)\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}),$$

we can set up the following moment equations

$$T_1 = E(T_1) = 4(1-p)/p^2, \tag{36}$$

$$T_2 = E(T_2) = -(1-\alpha)(1-p)/p^2. \tag{37}$$

These are readily solved to yield the following consistent moment estimates of the parameters of the model.

$$\widetilde{p} = [1 + \sqrt{1+T_1}]/2$$

and

$$\widetilde{\alpha} = 1 + (4T_2/T_1).$$

Example 2. *This time we will consider the two-parameter sub-model of the $BADL(II)$ model given by restricting its parameter space as follows:*

$$\begin{aligned} \tau_2 &= \tau_3 = \alpha \in (0,1) \\ \tau_1 &= \tau_4 = \tau_5 = \beta \in (0,1) \\ \tau_6 &= \tau_7 = \tau_8 = 0. \end{aligned} \tag{38}$$

We will call this model M2. We are adopting the convention that if $V \sim geo(\tau)$ with $\tau = 0$ then $V = \infty$. The full parameter vector for this two parameter model is thus

$$\underline{\tau} = (\beta, \alpha, \alpha, \beta, \beta, 0, 0, 0),$$

and the model, in terms of the V_i variables is given by

$$\begin{aligned} X &= min\{V_1, V_5\} - V_3, \\ \\ Y &= V_2 - min\{V_4, V_5\}. \end{aligned} \tag{39}$$

Elementary computations yield

$$E(X) = -E(Y) = \frac{(1-\beta)^2}{1-(1-\beta)^2} - \frac{1-\alpha}{\alpha}$$

and

$$var(X) = var(Y) = \frac{(1-\beta)^2}{[1-(1-\beta)^2]^2} + \frac{1-\alpha}{\alpha^2}.$$

Let us define

$$T_1 = (1/2n)[\sum_{i=1}^{n}[X_i - Y_i]$$

and

$$T_2 = (1/2n)[\sum_{i=1}^{n}[(X_i - \overline{X})^2 + (Y_i - \overline{Y})^2].$$

We can then set up the following moment equations to be solved for (α, β).

$$T_1 = E(T_1) = \frac{(1-\beta)^2}{1-(1-\beta)^2} - \frac{1-\alpha}{\alpha}, \tag{40}$$

$$\tag{41}$$

$$T_2 = E(T_2) = \frac{(1-\beta)^2}{[1-(1-\beta)^2]^2} + \frac{1-\alpha}{\alpha^2}. \tag{42}$$

These can be solved by iterative substitution. Thus choose an initial value for α , perhaps $\alpha = 0.5$, and then with this value of alpha solve for β in equation (40). Then substitute this value of β in equation (42) and solve for α, then use equation (40) once more, etc.

Multi-parameter submodels
As the two-parameter models discussed in Section 4 would readily suggest, freeing more of the δ's in $BDL(I)$ or τ's in $BADL(II)$ can lead to more complex models with multiple tail dependencies. Due to the complexity of the models, more creative computer intensive approaches, will need to be applied for parameter estimation.

Remark Higher dimensional versions of the BDL(I), BADL(II), GBDL(I), GBADL(II) models are readily described. Since, for example, the completely general trivariate version of the asymmetric discrete Laplace model of Type II will involve 26 independent geometric components each with its own τ parameter, only submodels including just a limited number of components will be tractable and useful for modeling purposes. One very simple trivariate submodel model that might find application is the following.

$$
\begin{aligned}
X &= \min\{V_1, V_7\} - V_{19} \\
Y &= \min\{V_2, v_7\} - V_{19} \\
Z &= V_3 - V_{19}
\end{aligned}
$$

which involves only 5 of the 26 components in the general model, but still has a non-trivial dependence structure..

5 Conclusion

In this paper, two discrete Laplace models with new methods of construction are detailed. Standard parameter estimation techniques are exhibited for both. Further, new bivariate variants of this family of distributions are discussed, and parameter estimation techniques are outlined. We finished with a brief discussion of higher-dimensional models.

References

[1] Barry C Arnold. "Some bivariate and multivariate models involving independent gamma distributed components." In: *Methodology and Applications of Statistics*. Springer, 2021, pp. 115–128.

[2] Barry C Arnold and Hon Keung Tony Ng. "Flexible bivariate beta distributions." In: *Journal of Multivariate Analysis* 102.8 (2011), pp. 1194–1202.

[3] Matthew Alexander Arvanitis. *Likelihood-free estimation for some flexible families of distributions*. University of California, Riverside, 2018.

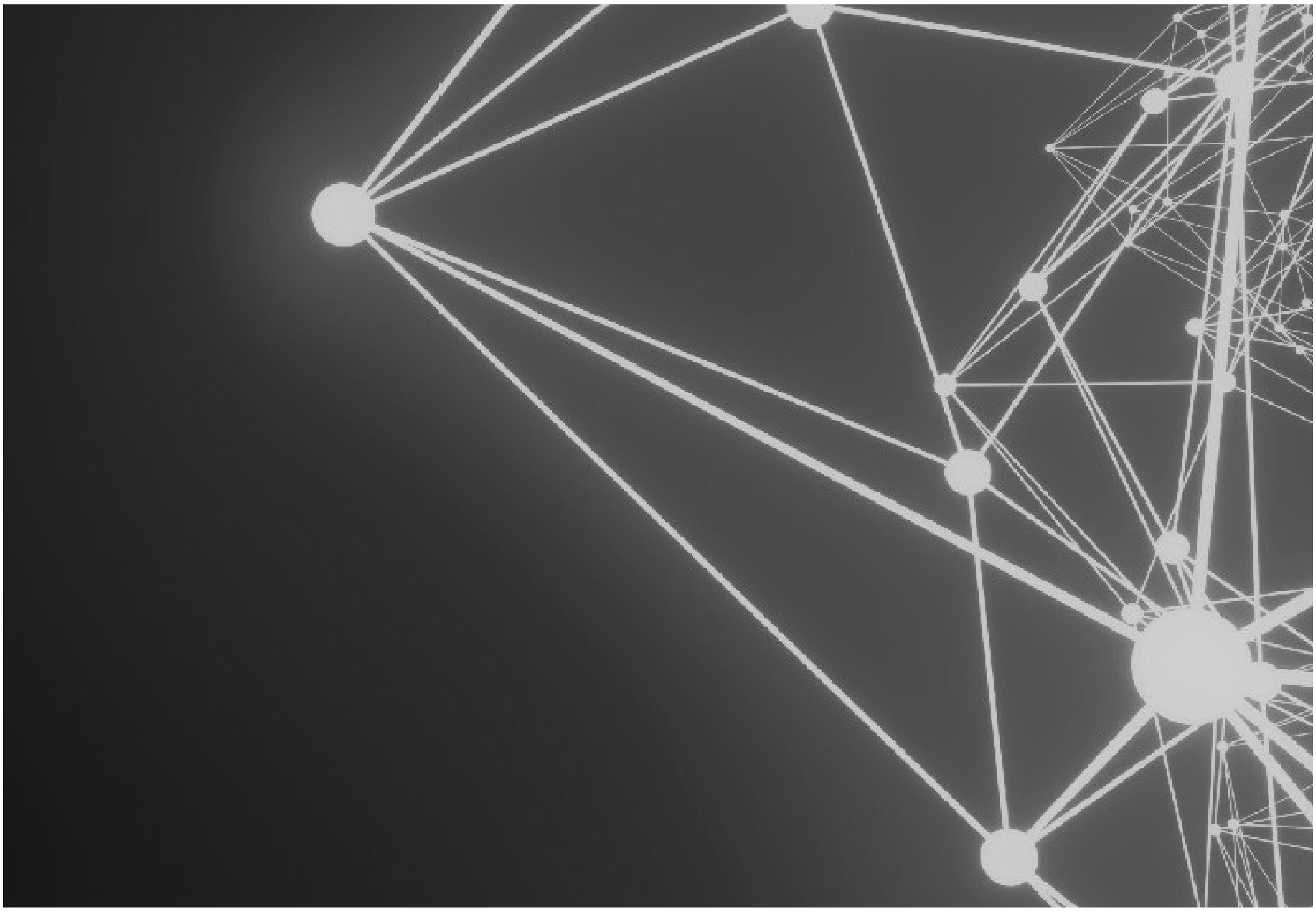
[4] Seidu Inusah and Tomasz J Kozubowski. "A discrete analogue of the Laplace distribution." In: *Journal of statistical planning and inference* 136.3 (2006), pp. 1090–1102.

[5] MC Jones. "On a characteristic property of distributions related to the Laplace." In: *South African statistical journal* 53.1 (2019), pp. 31–34.

[6] Tomasz J Kozubowski and Krzysztof Podgórski. "A multivariate and asymmetric generalization of Laplace distribution." In: *Computational Statistics* 15.4 (2000), pp. 531–540.

[7] Tomasz J Kozubowski and Krzysztof Podgórski. "Asymmetric Laplace distributions." In: *Mathematical Scientist* 25.1 (2000), pp. 37–46.

Gujarat Journal of Statistics and Data Science



Volume: 38 **Issue: 1** **July 2022**

Published by

Gujarat Statistical Association

GUJARAT JOURNAL OF STATISTICS AND DATA SCIENCE
(Formerly GUJARAT STATISTICAL REVIEW)

Gujarat Journal of Statistics and Data Science
GUJARAT STATISTICAL ASSOCIATION

Vol – 38 Issue – 1 July, 2022

CONTENTS PAGE

Gujarat Journal of Statistics and Data Science
GUJARAT STATISTICAL ASSOCIATION

Vol – 38 Issue – 1 July, 2022

Gujarat Journal of Statistics and Data Science
(Formerly Gujarat Statistical Review)

Editor in Chief: message

The special volume of Gujarat Journal of Statistics and Data Science is being published as a memorial volume for Late Professor C. G. Khatri, Retired Professor and Head, Department of Statistics, Gujarat University, Gujarat, India. Dr. Khatri was one of the stalwart experts in the field of Multivariate analysis, eminent researcher in Matrix Algebra, specialised in generalized inverse of matrices. He had jointly published several research papers on generalized inverse of matrices and Multivariate distributions in various national and international journals, co - author with the legendary figures in the world of Statistics like Professor C. R. Rao and Late Professor S. K. Mitra. He was frequent visitor at Indian Statistical Institute, Kolkata, Delhi and at several universities in Canada and USA.

My memory goes back to 1984 when for the first time I met Professor C. G. Khatri at the Vice- Chancellor's office, South Gujarat University, Surat, while I was facing him for an interview for the post of Associate Professor. He asked a few questions on variance balanced design as he had published one paper on variance balanced design in 1982 and my Ph.D. thesis was containing one chapter on VB design. The next meeting with him was at the Department of Statistics, Gujarat University, where he was conducting twenty-one days course on Linear Models, funded by UGC, New Delhi. Twenty-one days of academic learning under his brilliant guidance made me a good research worker in Linear Models and Design of Experiments. He was pioneer in organising the annual conference of Gujarat Statistical Association annually at various colleges/universities in Gujarat and we used to meet him.

It gives me utmost heartiest pleasure in being associated with the noble task of publishing a journal as a memorial volume for a person who had earned name and fame as a Statistician in India and in the world. Since the new name of the journal is Gujarat Journal of Statistics and Data Science, hence we included the first paper on Data science authored by Professor B.L.S. Prakasa Rao. I was entrusted with the responsibility of Editor in Chief of Gujarat Journal of Statistics and Data Science and to edit this volume in the meeting of the Gujarat Statistical Association held on August 2021. I am grateful to the committee members of the Gujarat Statistical Association for having confidence and faith on me to handle and complete this mighty task. Moreover, with the full support from all office bearers of GSA and in particular the President, Dr N D Shah, we could restart the publication of this prestigious journal. I am really happy in bringing up the journal in time with research papers of the National and International reputed authors. Professor Bikas Kumar Sinha, Managing Editor and Professor Ashis SenGupta, Editor of this journal are extremely helpful at every stage with extreme courtesy for having the publication of this volume from beginning to end. We used to discuss various queries and problems with regard to this volume either by email or by mobile. I was able to solve the problems with their advices and suggestions. It has been a wonderful and truly rewarding experience to work with them. I am also extremely grateful to Professor B.L.S. Prakasa Rao, former Director and Emeritus Scientist of I S I Kolkata who suggested the new name of this journal as Gujarat Journal of Statistics and Data Science. One more person Dr. Parag Shah deserves special mention because he has followed thru with all the editorial works and looked at stage by stage progress with extreme interest. He has been quite helpful in compiling the papers with manuscripts number and then making correspondence with authors, Editor, Managing Editor and Editor - in - Chief from time to time.

All the contributory authors and referees were seriously involved in their respective academic activities for which we have passed thru a lengthy editorial process over the last several months. I am extremely thankful to them for their academic and scientific interest in completing this task.

With these few words, I place the special volume of Gujarat Journal of Statistics and Data Science before our readers at large. We fondly hope they will not be disappointed with this volume.

Dilip Kumar Ghosh

Editor in Chief

Preface

My recollection of meeting Late Professor Chinubhai Ghelabhai *Khatri* [CGKhatri] goes back to 1970 during 57th Indian Science Congress at IIT, Kharagpur. As research scholars at CUDS, we [twin brothers] attended the same. Among others, Professors C R Rao and C G Khatri were both prominently visible in all sessions. After almost every presentation of young budding scholars, CGK would throw Qs like a roaring tiger but CRR was remarkably cool and would make observations in his characteristic smiling style. That was CGK – all along very much vocal and loud.

Over time … like many other seniors in our profession, Khatri became our Dear Chinu-da. We met him many times within India and abroad mainly at conferences. I used to meet him at ISI, Kolkata as well. He had a special liking for Sinha Brothers.

When DKG approached me, I did not give any second thought. I am delighted to have this opportunity to serve this Journal as its Managing Editor. With DKG as Chief Editor and ASG as the Editor and with a strong group of Advisory Board Members and Editorial Board Members, I am confident this revival of the journal will reflect on the passion of Late Prof. Khatri for the subject in terms of Teaching and Research in his favourite topics.

This first issue of GJSDS has been carefully crafted to reflect on the topics most liked by CGK. We fondly hope it will attract attention and satisfaction of our readers at large.

Bikas K Sinha

Kolkata

July 12, 2022

Forewords

It was the Summer (June) of 1979, I was a Ph.D. student and it was my very first presentation of a research paper in an international conference, Symposium on Variance Components, g-inverses and Applications, organized by Prof. C.R. Rao at The Ohio State University, Columbus, USA. Hardly had I completed, when someone from the front voiced very strongly that the original result, I had generalized was in fact established by him and not by CR Rao as I had stated. I politely answered that I had seen Prof. Rao's work and was ignorant about his result. Prof. C.R. Rao had his characteristic soft smile! Later I was told that the person was no other than Prof. C.G. Khatri! His passion with the research paper had left a lasting impression on me. Later in my research and teaching, I have used the classic book, *An Introduction to Multivariate Statistics*, by M. S. Srivastava, C. G. Khatri, North-Holland/New York, 1979, and I obviously came to greatly appreciate the profound impact Prof. Khatri has made in multivariate statistical inference and matrix theory. It was thus with reverence and pleasure that I accepted to be the Editor of this revival of Gujarat Statistical Review of which Prof. Khatri was the Founder-Editor. This new version will certainly progress in the footsteps of the earlier one as per the aims of Prof. Khatri. Moreover, we will also strive to fulfil our vision to enrich this journey with the emerging faces in theory and methodology, as well as with the challenges in real-data problems, which are surfacing in the arena of statistical science. With this opening issue, I invite senior and young researchers alike to submit their scholarly papers to our journal to take it to greater heights.

Ashis SenGupta

Editor