CHAPTER 4

BIVARIATE DISTRIBUTION EXTENSION

4.0 Introduction

Numerous bivariate discrete distributions have been defined and studied (see Mardia, 1970 and Kocherlakota and Kocherlakota, 1992) based on various methods of construction. For example, Famoye and Consul (1995) proposed a bivariate generalized Poisson distribution (BGPD) by using the method of trivariate reduction. However, the range of the correlation is restricted and it only permits positive correlation between the two random variables, X_1 and X_2. The bivariate negative binomial distribution (BNBD) defined by Lee (1999) based on copula function is very complicated and only be used for over dispersed data.

A bivariate Poisson distribution was proposed by Lakshaminaraya *et al.* (1999) with a flexible correlation that varies over the full range $(-1,\ 1)$. The model is formulated as a product of Poisson marginals with a multiplicative factor. Famoye (2010) adopted the approach of Lakshaminaraya *et al.* to construct a new bivariate generalized Poisson distribution (BGPD) which allows a more flexible correlation structure compared to the existing BGPD.

This chapter is an extension of Chapter 3 where two new bivariate $\mathrm{GIT}_{3,1}$ distributions are defined and examined. We shall introduce the new bivariate distributions as the alternative bivariate discrete distributions that allow more flexibility in modelling and less limitation on the correlation between the two random variables.

In section 2, we define two new bivariate $\mathrm{GIT}_{3,1}$ distribution (BGITD) based on the convolution of two bivariate distributions and the classical trivariate reduction method. Then some properties of these two BGITD are presented in section 3, and in section 4 the characteristic of the distributions are studied. Next, we discuss the estimation of the parameters of BGITD by maximum likelihood estimation and the estimation based on the probability generating function (pgf) in section 5. In section 6, we illustrate these two methods of estimation with three real life data together with test of goodness-of-fit.

4.1 Formulation of Bivariate $\mathrm{GIT}_{3,1}$ distributions

Two formulations of the bivariate $\mathrm{GIT}_{3,1}$ distributions are considered: (a) convolution of bivariate binomial and bivariate negative binomial distributions (Type I) and (b) the method of trivariate reduction (Type II).

4.1.1 Convolution (Type I)

The pgf of $\mathrm{GIT}_{3,1}$ distribution can be written as

$$\varphi(t) = \left(\frac{p_1 + p_2 t}{1 - p_3 t}\right)^n = \left(\frac{p_1}{p_1 + p_2} + \frac{p_2 t}{p_1 + p_2}\right)^n \left(\frac{1 - p_3}{1 - p_3 t}\right)^n \tag{4.1}$$

If $p_3 = 0$, it reduces to the pgf of the binomial distribution and when $p_2 = 0$, it is the negative binomial pgf. Thus, the $\mathrm{GIT}_{3,1}$ distribution is a convolution of the binomial $\mathrm{B}\left(n, \frac{p_1}{p_1 + p_2}\right)$ and negative binomial $\mathrm{NB}(n, p_3)$ distributions. In a similar manner, we define a BGITD (Type I BGITD) as the mixture of bivariate binomial and bivariate negative binomial distributions.

Joint probability generating function of Type I BGITD

The joint pgf is given by

$$\varphi(t_1,t_2)=\left(p_3+p_1t_1+p_2t_2\right)^{n_1}\left(\frac{q_3}{1-q_1t_1-q_2t_2}\right)^{n_2} \tag{4.2}$$

where $p_1+p_2+p_3=q_1+q_2+q_3=1$ and n_1,n_2 = positive integer.

Joint probability mass function

From (4.2), we solve for the joint pmf as mentioned in (2.5)

$$f(x_1,x_2)=q_3^{n_2}p_1^{x_1}p_2^{x_2}p_3^{n_1-x_1-x_2}\sum_{u=0}^{x_1}\sum_{v=0}^{x_2}\frac{n_1!}{(x_1-u)!(x_2-v)!(n_1-x_1-x_2+u+v)!}\frac{(u+v+n_2-1)!}{u!v!(n_2-1)!}\left(\frac{q_1}{p_1}\right)^u\left(\frac{q_2}{p_2}\right)^v p_3^{u+v} \tag{4.3}$$

where $n_1\geq\max(x_1+x_2-u-v)$, $n_2\geq 1$, n_1,n_2 = positive integer and

$p_1+p_2+p_3=q_1+q_2+q_3=1$

4.1.2 Trivariate Reduction (Type II)

The Type II BGITD is formulated by the method of trivariate reduction (Mardia, 1970). As explained in section (2.4.2), we have

$$X_1=(Y_1+Y_2)$$
$$X_2=(Y_1+Y_3)$$

where Y_1, Y_2 and Y_3 are independent $\mathrm{GIT}_{3,1}$ random variables. Now let φ_1, φ_2 and φ_3 be the pgf of Y_1,Y_2 and Y_3 respectively, that is

$$\varphi_1=\left(\frac{p_1+p_2t}{1-p_3t}\right)^{n_1};\varphi_2=\left(\frac{q_1+q_2t}{1-q_3t}\right)^{n_2};\varphi_3=\left(\frac{r_1+r_2t}{1-r_3t}\right)^{n_3}.$$

Joint probability generating function

By definition, the joint pgf of (X_1, X_2) is written as

$$\varphi(t_1, t_2) = \left(\frac{p_1 + p_2(t_1 t_2)}{1 - p_3(t_1 t_2)}\right)^{n_1} \left(\frac{q_1 + q_2 t_1}{1 - q_3 t_1}\right)^{n_2} \left(\frac{r_1 + r_2 t_2}{1 - r_3 t_2}\right)^{n_3} \tag{4.4}$$

where $p_1 + p_2 + p_3 = q_1 + q_2 + q_3 = r_1 + r_2 + r_3 = 1$ and n_1, n_2, n_3 = positive integer.

Joint probability mass function

From (4.4), the joint pmf is derived as

$$f(x_1, x_2) = p_1^{n_1} q_1^{n_2} r_1^{n_3} q_3^{x_1} r_3^{x_2} \sum_{k=0}^{\min(x_1, x_2)} \left(\begin{array}{l} \binom{n_1+k-1}{k} \binom{n_2+x_1-k-1}{x_1-k} \binom{n_3+x_2-k-1}{x_2-k} p_3^k q_3^{-k} r_3^{-k} \, {}_2F_1\left(-n_1, -k; -n_1-k+1; -\frac{p_2}{p_1 p_3}\right) \\ \times {}_2F_1\left(-n_2, -(x_1-k); -n_2-(x_1-k)+1; -\frac{q_2}{q_1 q_3}\right) {}_2F_1\left(-n_3, -(x_2-k); -n_2-(x_2-k)+1; -\frac{r_2}{r_1 r_3}\right) \end{array} \right)$$

where $p_1 + p_2 + p_3 = q_1 + q_2 + q_3 = r_1 + r_2 + r_3 = 1$ and n_1, n_2, n_3 = positive integer.

(4.5)

4.2 Properties of the Bivariate Distributions

4.2.1 Properties of Type I BGITD

Marginal distributions

Using (2.12) and (2.13), the marginal pgf of X_i, $i = 1, 2$ can easily be found as

$$\varphi_{X_i}(t) = \left(1 - p_i + p_i t\right)^{n_1} \left(\frac{q_3}{q_3 + q_i(1-t)}\right)^{n_2} \tag{4.6}$$

Marginal means and variances

Using (4.2), we can find the expected means and variances of the marginal X_1 and X_2 as follows. For i =1, 2

$$E(X_i) = \frac{n_2 q_i}{q_3} + n_1 p_i$$

$$Var(X_i) = \frac{n_1 q_i (q_i + q_3)}{q_3^{\ 2}} + n_1 p_i (1 - p_i)$$

Covariance and correlation

The marginal means and variances are used to yield the covariance, $Cov(X_1, X_2)$ and the correlation, $\rho(X_1, X_2)$ of X_1 and X_2.

$$Cov(X_1, X_2) = \frac{n_2 q_1 q_2}{q_3^{\ 2}} - n_1 p_1 p_2 \tag{4.7}$$

$$\rho(X_1, X_2) = \frac{n_2 q_1 q_2 - n_1 p_1 p_2 q_3^2}{\sqrt{\left(\left(n_1 q_1 (1 - q_2) + n_1 p_1 (1 - p_1) q_3^{\ 2}\right)\left(n_2 q_2 (1 - q_1) + n_1 p_2 (1 - p_2) q_3^{\ 2}\right)\right)}} \tag{4.8}$$

To further investigate the range of correlation, we look into different parameter sets of Type I BGITD. As illustrated in Table 4.1, five cases are considered. Note that from Table 4.2, the covariance and the correlation of Type I BGITD can be either positive or negative. The correlation provides a full range value which is from -1 to 1.

Table 4.1: Numerical examples

case	1	2	3	4	5
p_1	0.84	0.59	0.01	0.84	0.01
p_2	0.15	0.40	0.01	0.15	0.01
p_3	0.01	0.01	0.98	0.01	0.98
q_1	0.84	0.84	0.84	0.01	0.01
q_2	0.15	0.15	0.15	0.01	0.01
q_3	0.01	0.01	0.01	0.98	0.98

Table 4.2: Covariance and correlation of Type I BGITD

case	1	2	3	4	5
			$n_1 = n_2 = 1$		
$Cov(X_1, X_2)$	1259.87	1259.76	1260.00	-0.1260	0.000004
$\rho(X_1, X_2)$	0.9600	0.9620	0.9630	-0.8920	0.000204
			n_1=5 > n_2=1		
$Cov(X_1, X_2)$	1259.37	1258.82	1260.00	-0.6300	-0.00040
$\rho(X_1, X_2)$	0.9610	0.9590	0.9620	-0.9470	-0.00662
			n_1=1 < n_2=5		
$Cov(X_1, X_2)$	6299.87	6299.76	6300.00	-0.1250	0.0004
$\rho(X_1, X_2)$	0.9620	0.9620	0.9630	-0.6880	0.0068

From (4.1), the first few probabilities are

$$f(0,0) = p_3^{n_1} q_3^{n_2}$$

$$f(1,0) = f(0,0)\frac{(n_1 p_1 + n_2 p_3 q_1)}{p_3}$$

$$f(0,1) = f(0,0)\frac{(n_1 p_2 + n_2 p_3 q_2)}{p_3}$$

$$f(1,1) = f(0,0)\left(n_2(1+n_2)q_1 q_2 + \frac{(n_1(n_2 p_2 p_3 q_1 + p_1(-p_2 + n_2 p_3 q_2))}{p_3^2} + \frac{n_1^2 p_1 p_2}{p_3} \right)$$

Conditional distribution

Next, we determine the conditional distribution on X_1 given $X_2 = x_2$ by

(1) Applying the derivative product rule

(2) Successive differentiation.

Firstly, setting $\varphi(t_1, t_2) = j(t_1, t_2) g(t_1, t_2)$ as a convolution of two independent bivariate random variables, where $j(t_1, t_2)$ and $g(t_1, t_2)$ are respectively the pgfs of the binomial and negative binomial distributions.

By definition, the pgf of the conditional distribution of X_1 given $X_2 = x_2$ is

$$\varphi_{X_1}(t|\,y) = \frac{\varphi^{(0,x_2)}(t,0)}{\varphi^{(0,x_2)}(1,0)}$$

where $\varphi^{(x_1,x_2)}(t_1,t_2) = \dfrac{\partial^{x_1+x_2}}{\partial t_1^{\,x_1}\partial t_2^{\,x_2}}\varphi(t_1,t_2)\Big|_{t_1=u,t_2=v}$

Using the derivative product rule, we have

$$\varphi^{(0,x_2)}(t_1,t_2) = \frac{\partial^{x_2}}{\partial t_2^{\,x_2}}\varphi(t_1,t_2)\Big|_{t_2=0} = \sum_{k=0}^{x_2}\binom{x_2}{k} j^k g^{x_2-k}\Big|_{t_2=0} \tag{4.9}$$

where

$$j^0\Big|_{t_2=0} = \left(p_3 + p_1 t\right)^{n_1} \tag{4.10}$$

$$j^i\Big|_{t_2=0} = p_2^{\,i}\left(p_3 + p_1 t\right)^{n_1-i}\left(\prod_{j=0}^{i-1}(n_1 - j)\right) \quad , \; i \ge 1 \tag{4.11}$$

$$g^0\Big|_{t_2=0} = \left(\frac{q_3}{1-q_1 t}\right)^{n_2} \tag{4.12}$$

$$g^i\Big|_{t_2=0} = q_3^{\,n_2} q_2^{\,i}\left(\frac{1}{1-q_1 t}\right)^{n_2+i}\left(\prod_{j=0}^{i-1}(n_2 + j)\right) \quad , \; i \ge 1 \tag{4.13}$$

With the substitution of equations (4.10), (4.11), (4.12) and (4.13) into equation (4.9), the numerator part of equation (2.14) can be obtained. A further substitution of t=1 into equation (4.9) will lead us to solve for the denominator part of equation (2.14). Then, the conditional distribution of X_1 given $X_2 = x_2$ can be determined.

Secondly, by successive differentiation with respect to t_1 and setting $t_2 = 0$, we have

$$\varphi_{X_1}(t|\,x_2) = \left(\frac{p_3+p_1 t}{p_1+p_3}\right)^{n_1}\left(\frac{1-q_1}{1-q_1 t}\right)^{n_2}\left\{\frac{\left(\frac{q_2}{1-q_1 t}\right)^{x_2}\left(\prod_{i=0}^{x_2-1}(n_2+i)\right)+\left(\frac{p_2}{p_3-p_1 t}\right)^{x_2}\left(\prod_{j=0}^{x_2-1}(n_1-j)\right) + \sum_{k=1}^{y-1}\binom{y}{k}\left(\frac{p_2}{p_3-p_1 t}\right)^{k}\left(\frac{q_2}{1-q_1 t}\right)^{x_2-k}\left(\prod_{j=0}^{k-1}(n_1-j)\right)\left(\prod_{i=0}^{x_2-k-1}(n_2+i)\right)\Bigg|_{x_2\ge 2}}{\left(\frac{q_2}{1-q_1}\right)^{x_2}\left(\prod_{i=0}^{x_2-1}(n_2+i)\right)+\left(\frac{p_2}{p_3-p_1}\right)^{x_2}\left(\prod_{j=0}^{x_2-1}(n_1-j)\right) + \sum_{k=1}^{x_2-1}\binom{x_2}{k}\left(\frac{p_2}{p_3-p_1}\right)^{x_2}\left(\frac{q_2}{1-q_1}\right)^{x_2-k}\left(\prod_{j=0}^{k-1}(n_1-j)\right)\left(\prod_{i=0}^{x_2-k-1}(n_2+i)\right)\Bigg|_{x_2\ge 2}}\right\} \quad , \; x_2 \ge 1$$

and for $x_2 = 0$, we have $\varphi_{X_1}(t|\ x_2) = \left(\frac{p_3 + p_1 t}{p_1 + p_3}\right)^{n_1} \left(\frac{1-q_1}{1-q_1 t}\right)^{n_2}$

4.2.2 Properties of Type II BGITD

Marginal distributions

Similarly, using (2.12) and (2.13), the marginal pgf of X_1 and X_2 are given as

$$\varphi_{X_1}(t) = \left(\frac{p_1 + p_2 t}{1 - p_3 t}\right)^{n_1} \left(\frac{q_1 + q_2 t}{1 - q_3 t}\right)^{n_2} \left(\frac{r_1 + r_2}{1 - r_3}\right)^{n_3}$$

$$\varphi_{X_2}(t) = \left(\frac{p_1 + p_2 t}{1 - p_3 t}\right)^{n_1} \left(\frac{q_1 + q_2}{1 - q_3}\right)^{n_2} \left(\frac{r_1 + r_2 t}{1 - r_3 t}\right)^{n_3}$$

Marginal means and variances

From (4.4), the means and variances of the marginal X_1 and X_2 are obtained as

$$E(X_1) = n_1\left(\frac{1-p_1}{1-p_3}\right) + n_2\left(\frac{1-q_1}{1-q_3}\right)$$

$$E(X_2) = n_1\left(\frac{1-p_1}{1-p_3}\right) + n_3\left(\frac{1-r_1}{1-r_3}\right)$$

$$Var(X_1) = n_1 \frac{(1-p_1)(1-p_2)}{(1-p_3)^2} + n_2 \frac{(1-q_1)(1-q_2)}{(1-q_3)^2}$$

$$Var(X_2) = n_1 \frac{(1-p_1)(1-p_2)}{(1-p_3)^2} + n_3 \frac{(1-r_1)(1-r_2)}{(1-r_3)^2}$$

Covariance and correlation

The covariance and the correlation of X_1 and X_2 are given as

$$Cov(X_1, X_2) = \frac{n_1(1-p_1)(1-p_2)}{(1-p_3)^2} \tag{4.14}$$

$$\rho(X_1, X_2) = \frac{n_1(1-p_1)(1-p_2)(q_1+q_2)(r_1+r_2)}{\sqrt{\begin{array}{l}\left(n_1(1-p_1)(1-p_2)(q_1+q_2)^2 + n_2(1-q_1)(1-q_2)(p_1+p_2)^2\right) \\ \times\left(n_1(1-p_1)(1-p_2)(r_1+r_2)^2 + n_3(1-r_1)(1-r_2)(p_1+p_2)^2\right)\end{array}}} \tag{4.15}$$

Since $n_1(1-p_1)(1-p_2)$ >0, it follows that for this model, the correlation of X_1 and X_2 are always positive.

The first few probabilities are given by

$$f(0,0) = p_1^{n_1} q_1^{n_2} r_1^{n_3}$$

$$f(1,0) = f(0,0)\frac{n_2(1-q_1)(1-q_3)}{q_1}$$

$$f(0,1) = f(0,0)\frac{n_3(1-r_1)(1-r_3)}{r_1}$$

$$f(1,1) = f(0,0)\left(\frac{n_1(p_2+p_1p_3)}{p_1} + \frac{n_2n_3(q_2+q_1q_3)(r_2+r_1r_3)}{q_1r_1}\right)$$

Conditional distribution

The conditional distribution on X_1 given $X_2 = x_2$ is determined by using the derivative product rule as the successive differentiation of equation (4.4) is tedious.

Setting

$$\varphi(t_1,t_2) = e(t_1,t_2)h(t_1,t_2)\left(\frac{q_1+q_2t_1}{1-q_3t_1}\right)^{n_2} \tag{4.16}$$

where $e(t_1,t_2) = \left(\frac{p_1+p_2t_1t_2}{1-p_3t_1t_2}\right)^{n_1}$ and $h(t_1,t_2) = \left(\frac{r_1+r_2t_2}{1-r_3t_2}\right)^{n_3}$

By definition,

$$\begin{aligned}
\varphi^{(0,x_2)}(t_1,t_2) &= \frac{\partial^{x_2}}{\partial t_2^{\,x_2}}\varphi(t_1,t_2)\Bigg|_{t_2=0} \\
&= \left(\frac{q_1+q_2t_1}{1-q_3t_1}\right)^{n_2}\frac{\partial^{x_2}}{\partial t_2^{\,x_2}}e(t_1,t_2)h(t_1,t_2)\Bigg|_{t_2=0} \\
&= \left(\frac{q_1+q_2t_1}{1-q_3t_1}\right)^{n_2}\sum_{k=0}^{x_2}\binom{x_2}{k}e^k h^{x_2-k}\Bigg|_{t_2=0}
\end{aligned}$$

The successive differentiation of $e(t_1,t_2)$ and $h(t_1,t_2)$ with respect to t_2 gives us

$$e^0\Big|_{t_2=0} = \left(p_1\right)^{n_1} \tag{4.17}$$

$$e^i\Big|_{t_2=0} = t^i\left[\sum_{k=1}^{i}\binom{i}{k}\frac{(i-1)!}{(k-1)!}p_1^{n_1-k}p_3^{i-k}(p_2+p_1p_3)^k\left(\prod_{j=0}^{k-1}(n_1-j)\right)\right] \quad , \ i \geq 1 \tag{4.18}$$

$$h^0\Big|_{t_2=0} = \left(r_1\right)^{n_3} \tag{4.19}$$

$$h^i\Big|_{t_2=0} = \left[\sum_{k=1}^{i}\binom{i}{k}\frac{(i-1)!}{(k-1)!}r_1^{n_3-k}r_3^{i-k}(r_2+r_1r_3)^k\left(\prod_{j=0}^{k-1}(n_3-j)\right)\right] \quad , \ i \geq 1 \tag{4.20}$$

Similarly, by using equations (4.17), (4.18), (4.19) and (4.20), $\varphi^{(0,x_2)}(t_1,t_2)$ can be solved. The conditional distribution of X_1 given $X_2 = x_2$ can then be easily derived.

4.3 Characteristics of the Distributions

To explore the structure of the distributions graphically, a few figures are presented by varying the parameters of the Type I and Type II BGITD. Note that all the figures are scaled by multiplying the y-axis with the value 100. Let the joint pmf be $pr(x_1,x_2)$ with $x_1 = 0,1,2,...,r$ and $x_2 = 0,1,2,...,s$.

4.3.1 Characteristics of Type I BGITD

To study the effect of changes of the parameters n_1 and n_2 of the Type I BGITD, we set r =8 and s =8 and n_1 = r + s =16 (as $n_1 \geq \max(x_1+x_2)$ from equation (4.2)) and n_2 = 1 (Figure 4.1). In Figure 4.2, the value of n_1 is increased with n_2 fixed whereas in Figure 4.3 we increase the value of n_2 but fix the value of n_1 (same as the value of n_1 in Figure 4.2).

It is clear from Figures 4.1 and 4.2 that increasing the value of n_1 shifts the local mode away from the origin, (0, 0) and decreases the height of the peak. By comparing Figures 4.2 and 4.3, it is seen that the local modes are flatter and located further away.

In conclusion, if the data are located close to the origin and the local mode is sharp, small positive integers value of n_1 and n_2 are suggested. Larger value of n_1 is suggested if the value of the mode is low and situated slightly away from the origin. Meanwhile, large positive integers for both n_1 and n_2 are recommended only when the local mode and the data are placed quite far away from the origin and the data form a nice bell-shaped. This observation will be useful in the empirical modelling of data.

A local mode at (1,1) with value 94.

Figure 4.1: p_1=0.1, p_2=0.1 q_1=0.1, q_2=0.1, n_1= 16, n_2=1

A local mode at (4,4) with value 41.

Figure 4.2: p_1=0.1, p_2=0.1 q_1=0.1, q_2=0.1, n_1=42, n_2=1

Local modes are (7, 7), (7,8) and (8,7) with value 20.

Figure 4.3: p_1=0.1, p_2=0.1 q_1=0.1, q_2=0.1, n_1=42, n_2=30

4.3.2 Characteristics of Type II BGITD

Figures 4.4 to 4.10 are displayed in order to understand the behaviour of Type II BGIT with different combination of the positive integers n_1, n_2 and n_3. We examine the changes of n_1, n_2 and n_3 as given in Table 4.3. It is easily seen that by varying n_1, n_2 and n_3, different characteristics of Type II BGIT models are displayed.

In Figure 4.4, where n_1, n_2 and n_3 are set at 1, we achieve a very sharp and high value of the local mode. However, when n_1 is increasing (n_2 and n_3 fixed) as in Figure 4.5, the contours change to ellipses. Furthermore, the height of its local mode is decreased and situated further from the origin in comparison with Figure 4.4. In Figure 4.6, n_2 is increased but $n_1 = n_3 = 1$ and it is found that a flatter mode is formed; the contour plots show a shift to the *x*-axis. On the other hand, when $n_1 = n_2 = 1$ and n_3 is increased, Figure 4.7 is seen to be a reflection of Figure 4.6 through the origin, with the heights of the modes unchanged.

Figures 4.8 to 4.10 show the effect of increasing two parameters and fixing the remaining one. It is observed that in Figure 4.8 the local mode has the same location as the one in Figure 4.5 but the peak of the local mode is much lower and the elliptic

contours are broader. The plot in Figure 4.9 is similar with Figure 4.6 while Figures 4.10 and 4.7 are similar (with reflection).

The plots indicate the flexibility of the shape of the Type II BGIT and small n_1, n_2 and n_3 give local modes situated very close to the origin. To apply Type II BGIT in data analysis the modes of the data will indicate the values of n_1, n_2 and n_3.

Table 4.3: Change of the positive integers n_1, n_2 and n_3

Figure	4.4	4.5	4.6	4.7	4.8	4.9	4.10
n_1	1	5	1	1	1	5	5
n_2	1	1	5	1	5	5	1
n_3	1	1	1	5	5	1	5

A local mode at (2,2) with value 98.

Figure 4.4: $p_1 = 0.1, p_2 = 0.4, q_1 = 0.1, q_2 = 0.4, r_1 = 0.1, r_2 = 0.4, n_1 = n_2 = n_3 = 1$

A local mode at (9,9) with value 36.

Figure 4.5: $p_1 = 0.1, p_2 = 0.4, q_1 = 0.1, q_2 = 0.4, r_1 = 0.1, r_2 = 0.4, n_1 = 5, n_2 = n_3 = 1$

Local modes at (8,2) and (9,2) with value 32.

Figure 4.6: $p_1 = 0.1, p_2 = 0.4, q_1 = 0.1, q_2 = 0.4, r_1 = 0.1, r_2 = 0.4, n_1 = 1, n_2 = 5, n_3 = 1$

Local modes at (2,8) and (2,9) with value 32.

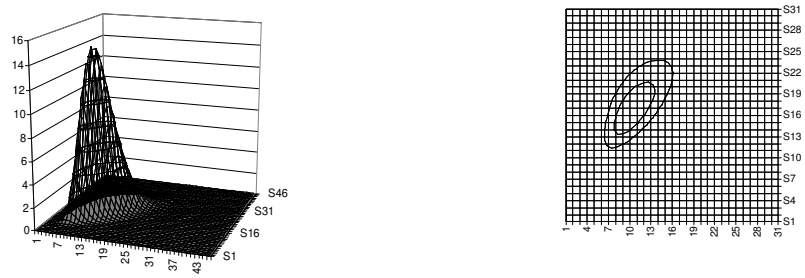
Figure 4.7: $p_1 = 0.1, p_2 = 0.4, q_1 = 0.1, q_2 = 0.4, r_1 = 0.1, r_2 = 0.4, n_1 = 1, n_2 = 1, n_3 = 5$

A local mode at (9,9) with value 15 .

Figure 4.8: $p_1 = 0.1, p_2 = 0.4, q_1 = 0.1, q_2 = 0.4, r_1 = 0.1, r_2 = 0.4, n_1 = 1, n_2 = 5, n_3 = 5$

Local modes at (15,9), (16,9) and (16,10) with value 15.

Figure 4.9: $p_1 = 0.1, p_2 = 0.4, q_1 = 0.1, q_2 = 0.4, r_1 = 0.1, r_2 = 0.4, n_1 = 5, n_2 = 5, n_3 = 1$



Local modes at (9,15), (9,16) and (10,16) with value 15 .

Figure 4.10: $p_1 = 0.1, p_2 = 0.4, q_1 = 0.1, q_2 = 0.4, r_1 = 0.1, r_2 = 0.4, n_1 = 5, n_2 = 1, n_3 = 5$

4.4 Parameter Estimation

In this section, two methods of parameter estimation are considered. Due to the complicated joint pmf, simulated annealing is proposed for the numerical optimization under these two estimations. For the Type I BGITD distribution, we shall consider the case where the parameters n_1 and n_2 are fixed and for the Type II BGITD distribution, the parameters n_1 , n_2 and n_3 are fixed.

4.4.1 Maximum Likelihood Estimation (MLE)

For the probability function based estimation method, we consider maximum likelihood estimation where the log-likelihood function

$$\ln L = \sum_{i=0}^{r}\sum_{j=0}^{s}\pi_{ij}\ln f(i,j)$$

is maximized, where π_{ij} is the observed frequency in the (i, j) cell for $i = 0,1,2,\ldots, r$; $j = 0,1,2,\ldots, s$.

4.4.2 The Pgf-based Minimum Hellinger-type Distance Estimation (PGFBE)

The use of pgf is popular as generally it has a simpler form than the probability mass function and this leads to simpler statistical inference procedures which shorten computation times. From Sim and Ong (2010), the univariate pgf-based estimator is extended to bivariate distribution as follows

$$T_2 = \int_0^1\int_0^1 \left(\psi(t_1,t_2) - \varphi(t_1,t_2)\right)^2 dt_1 dt_2$$

where $\psi(t_1,t_2) = \sum_{i}^{k_1}\sum_{j}^{k_2} n_{ij} t_1^i t_2^j$ is the empirical pgf and $\varphi(t_1,t_2)$ is the pgf of the distribution.

4.5 Numerical examples

To illustrate the application in data fitting, three real life data sets are considered:

a) Number of accidents sustained by 122 experienced shunters over 2 successive periods of time. (Arbous and Kerrich, 1951)

b) Number of patients in two boxes in a room of the critical care and emergency service in the San Agustin Hospital (Linares, Spain). (Rodriguez *et al.*, 2006).

c) Number of times bacon and eggs were purchased on four consecutive shopping trips. (Peter and Bruce, 2005)

These data are fitted by the bivariate negative binomial distribution (Subrahmaniam, 1966), Type I and Type II BGITD by MLE and the pgf-based

minimum Hellinger-type distance estimation. The performance of these parameter estimation methods will be discussed and the expected frequencies are displayed in Tables 4.4, 4.6 and 4.8 respectively. No grouping has been done and the chi-square goodness-of-fit (GOF) tests are carried out to investigate how well the observed distribution fits the estimated distribution. The parameters n_1 , n_2 and n_3 are chosen which provide the lowest chi-square value for the fit. For Type I BGITD, the initial parameter estimate of n_2 corresponding to the bivariate negative binomial component is first treated as a real number. Then, we choose the positive integers of parameters n_1 and n_2 based on the lowest chi-square value as mentioned above. For type II BGITD, Table 4.3 is used to assist us to choose the suitable positive integers.

In Table 4.4 (Arbous and Kerrich, 1951) X_1 represents the number of accidents in 5- year period 1943-47 and X_2 is the number of accidents in 6- year period 1937-42. From the data, we find that the sample mean and variance for X_1 are 0.9754 and 1.2969 respectively. Meanwhile, for X_2 , the sample mean is $\bar{x}_2$=1.2705 and sample variance is $\hat{\sigma}^2_{x_2}$=1.6535. The sample correlation is found to be 0.2585.

In Table 4.6, the data were obtained from a room of Critical Care and Emergency Service in San Agustín Hospital (Linares, Spain) where the reference population is more than 135,000 people. Patients are dealt with in one of two boxes where they can stay a maximum of 24 hours without preference (random). The handling of patients in shifts worked in 2000 and 2001 is considered and the observed frequencies of patients dealt with in each box are shown in Table 4.7. Similarly, we compute the sample mean and variance for X_1 and X_2 : $\bar{x}_1$ =3.0066, $\hat{\sigma}^2_{x_1}$ =1.4768,

$\bar{x}_2$=2.9868 and $\hat{\sigma}^2_{x_2}$=1.4170. The sample correlation is found to be slightly negative which is -0.0091. Hence, we fit the data with the Type I BGITD.

Table 4.8 exhibits the purchase of bacon and eggs from the Information Resources, Inc., a consumer panel based in a large U.S. city. A sample of 548 households over four consecutive store trips was tracked. For each household, a total number of bacon purchases in their four shopping trips and the total number of egg purchases for the same trip were counted. From the data, the sample mean and variance for X_1 and X_2 are $\bar{x}_1$=0.2956, $\hat{\sigma}^2_{x_1}$=0.4243, $\bar{x}_2$=0.7147 and $\hat{\sigma}^2_{x_2}$=0.5074. In addition, the sample correlation is 0.3101.

Example 1

Table 4.4: Number of accidents sustained by 122 experienced shunters over 2 successive periods of time

	X_2							
X_1	**0**	**1**	**2**	**3**	**4**	**5**	**6**	Total
0	**21**	**18**	**8**	**2**	**1**	**0**	**0**	**50**
BNBD*	21.90	16.67	7.98	3.07	1.04	0.32	0.09	51.08
BNBD^	21.32	16.97	8.04	2.95	0.92	0.26	0.07	50.53
BGITD1*	20.45	18.28	8.89	3.29	1.08	0.34	0.11	52.42
BGITD1^	20.79	17.77	8.48	3.17	1.07	0.35	0.12	51.74
BGITD2*	21.18	17.14	6.89	2.77	1.11	0.45	0.18	49.72
BGITD2^	21.56	16.86	7.89	2.87	0.89	0.25	0.06	50.38
1	**13**	**14**	**10**	**1**	**4**	**1**	**0**	**43**
BNBD*	12.52	13.18	8.06	3.77	1.50	0.53	0.18	39.75
BNBD^	12.52	14.01	8.56	3.84	1.42	0.46	0.14	40.94
BGITD1*	13.25	13.16	7.47	3.34	1.34	0.50	0.18	39.24
BGITD1^	13.20	12.80	7.29	3.34	1.39	0.55	0.21	38.78
BGITD2*	11.32	15.61	8.91	3.58	1.44	0.58	0.23	41.67
BGITD2^	11.61	15.10	8.96	3.75	1.28	0.38	0.10	41.18
2	**4**	**5**	**4**	**2**	**1**	**0**	**1**	**17**
BNBD*	4.50	6.06	4.54	2.52	1.16	0.47	0.17	19.44
BNBD^	4.38	6.31	4.79	2.58	1.12	0.42	0.14	19.74
BGITD1*	4.88	5.67	3.88	2.10	1.01	0.44	0.19	18.18
BGITD1^	4.84	5.60	3.90	2.19	1.09	0.50	0.22	18.33
BGITD2*	3.77	6.50	5.04	2.44	0.98	0.40	0.16	19.29
BGITD2^	4.05	6.41	4.90	2.41	0.92	0.30	0.09	19.08

Table 4.4, continued.

3	**2**	**1**	**3**	**2**	**0**	**1**	**0**	**9**
BNBD*	1.30	2.13	1.90	1.23	0.65	0.30	0.12	7.62
BNBD^	1.18	2.09	1.91	1.21	0.61	0.26	0.10	7.35
BGITD1*	1.44	2.01	1.66	1.07	0.60	0.30	0.14	7.22
BGITD1^	1.43	2.03	1.72	1.15	0.67	0.35	0.17	7.53
BGITD2*	1.25	2.16	1.88	1.09	0.49	0.20	0.08	7.16
BGITD2^	1.42	2.24	1.88	1.06	0.45	0.16	0.05	7.26
4	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**2**
BNBD*	0.33	0.64	0.66	0.49	0.29	0.15	0.07	2.61
BNBD^	0.27	0.57	0.61	0.45	0.26	0.12	0.05	2.33
BGITD1*	0.39	0.65	0.64	0.48	0.31	0.17	0.09	2.74
BGITD1^	0.39	0.68	0.68	0.53	0.35	0.21	0.12	2.97
BGITD2*	0.42	0.72	0.63	0.39	0.19	0.08	0.03	2.45
BGITD2^	0.49	0.78	0.66	0.39	0.18	0.07	0.02	2.59
5	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
BNBD*	0.08	0.17	0.20	0.17	0.11	0.06	0.03	0.82
BNBD^	0.06	0.14	0.17	0.14	0.09	0.05	0.02	0.66
BGITD1*	0.10	0.20	0.23	0.20	0.14	0.09	0.05	1.02
BGITD1^	0.11	0.22	0.25	0.23	0.17	0.11	0.07	1.15
BGITD2*	0.14	0.24	0.21	0.13	0.07	0.03	0.01	0.82
BGITD2^	0.17	0.27	0.23	0.14	0.06	0.02	0.01	0.91
6	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
BNBD*	0.02	0.04	0.06	0.05	0.04	0.02	0.01	0.24
BNBD^	0.01	0.03	0.04	0.04	0.03	0.02	0.01	0.17
BGITD1*	0.03	0.06	0.08	0.08	0.06	0.04	0.03	0.38
BGITD1^	0.03	0.07	0.09	0.09	0.07	0.05	0.04	0.44
BGITD2*	0.05	0.08	0.07	0.04	0.02	0.01	0.00	0.27
BGITD2^	0.06	0.10	0.08	0.05	0.02	0.01	0.00	0.32
7	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**1**
BNBD*	0.00	0.01	0.01	0.01	0.01	0.01	0.38	0.45
BNBD^	0.00	0.01	0.01	0.01	0.01	0.00	0.24	0.28
BGITD1*	0.01	0.02	0.03	0.03	0.03	0.02	0.68	0.81
BGITD1^	0.01	0.02	0.03	0.03	0.03	0.02	0.90	1.05
BGITD2*	0.02	0.03	0.02	0.01	0.01	0.00	0.52	0.61
BGITD2^	0.02	0.03	0.03	0.02	0.01	0.00	0.18	0.29

* = MLE, ^ = PGF based estimator, BGITD1 = type I BGITD, BGITD2 = type II BGITD

ML estimates:

BNBD: $\hat{p}_1$=0.1964, $\hat{p}_2$=0.1475, $\hat{p}_3$=0.0140, $\hat{v}$=3.8763.

Type I BGITD: $\hat{p}_1$=0.0418, $\hat{p}_2$=0.0275, $\hat{q}_1$=0.3101, $\hat{q}_2$=0.2634, n_1=13, n_2=1.

Type II BGITD: $\hat{p}_1$=0.9855, $\hat{p}_2$=0.0010, $\hat{q}_1$=04249, $\hat{q}_2$=0.1730, $\hat{r}_1$=0.5552, $\hat{r}_2$= 0.1119,

n_1=21, n_2=1, n_3=1.

PGF-based estimates:

BNBD: $\hat{p}_1$=01519, $\hat{p}_2$=0.1121, $\hat{p}_3$=0.0192, $\hat{v}$=5.2389

Type I BGITD: $\hat{p}_1$=0.0380, $\hat{p}_2$=0.0264, $\hat{q}_1$=0.3263, $\hat{q}_2$=0.2689, n_1=13, n_2=1.

Type II BGITD: $\hat{p}_1$=0.9872, $\hat{p}_2$=0.0010, $\hat{q}_1$=04255, $\hat{q}_2$=0.1519, $\hat{r}_1$=0.5439, $\hat{r}_2$= 0.1254,

n_1=21, n_2=1, n_3=1.

By using the estimated parameters as shown above, we compute the correlation for the estimated distributions. All of the estimated correlation values are close to the sample correlation which is 0.2585.

MLE

Model	$\rho(X_1, X_2)$
BNBD*	0.28
BGITD1*	0.28
BGITD2*	0.22

PGFBE

Model	$\rho(X_1, X_2)$
BNBD^	0.28
BGITD1^	0.31
BGITD2^	0.22

Chi-square GOF test

A summary of the χ^2 values, the full model log-likelihood values and the P-values are shown in the following table.

Table 4.5: Summary statistics for Table 4.4

Model	χ^2	Log-likelihood value	P-value	Degree of freedom
BNBD*	121.70	-341.61	0.0000	51
BGITD1*	75.04	-342.05	0.0159	51
BGITD2*	60.40	-341.48	0.1274	49

It is to be observed that in Table 4.5, the fit by Type I and Type II BGITD are significantly better than the BNBD based upon the χ^2 values. The Type II BGITD is preferred over the Type I BGITD as its P-value is higher and the result is significant.

Example 2

Table 4.6: Number of patients in two boxes in a room of the critical care and emergency service in the San Agustin Hospital (Linares, Spain)

Box 1, X_1	Box 2, X_2							
	0	**1**	**2**	**3**	**4**	**5**	**6**	**Total**
0	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**1**
BGITD1*	0.04	0.22	0.61	1.02	1.14	0.91	0.53	4.46
BGITD1^	0.02	0.15	0.45	0.80	0.96	0.83	0.52	3.73
1	**0**	**1**	**4**	**5**	**3**	**1**	**0**	**14**
BGITD1*	0.23	1.24	3.07	4.58	4.56	3.17	1.58	18.42
BGITD1^	0.15	0.90	2.42	3.89	4.17	3.14	1.68	16.35
2	**0**	**3**	**10**	**13**	**11**	**3**	**0**	**40**
BGITD1*	0.62	3.09	6.92	9.17	7.98	4.76	1.97	34.53
BGITD1^	0.45	2.43	5.88	8.41	7.90	5.09	2.28	32.45
3	**1**	**4**	**12**	**11**	**11**	**5**	**1**	**45**
BGITD1*	1.04	4.64	9.23	10.71	7.99	3.97	1.32	38.91
BGITD1^	0.82	3.95	8.48	10.62	8.55	4.59	1.64	38.66
4	**1**	**3**	**10**	**10**	**7**	**3**	**1**	**35**
BGITD1*	1.17	4.65	8.09	8.04	5.00	1.99	0.50	29.43
BGITD1^	1.00	4.27	8.03	8.62	5.79	2.49	0.67	30.86
5	**0**	**1**	**4**	**4**	**4**	**1**	**0**	**14**
BGITD1*	0.94	3.26	4.86	4.03	2.00	0.60	0.10	15.78
BGITD1^	0.86	3.24	5.21	4.67	2.51	0.81	0.15	17.44
6	**0**	**0**	**1**	**1**	**1**	**0**	**0**	**3**
BGITD1*	0.55	1.63	2.03	1.34	0.50	0.10	4.32	10.47
BGITD1^	0.54	1.75	2.35	1.68	0.68	0.15	5.35	12.50

*= MLE, ^= PGF based estimator, BGITD1=type I BGITD, BGITD2=type II BGITD

ML estimates:

Type I BGITD: $\hat{p}_1$=0.2488, $\hat{p}_2$=0.2505, $\hat{q}_1$=0.0010, $\hat{q}_2$=0.0010, n_1=12, n_2=1.

PGF-based estimates:

Type I BGITD: $\hat{p}_1$=0.2583, $\hat{p}_2$=0.2603, $\hat{q}_1$=0.0010, $\hat{q}_2$=0.0010, n_1=12, n_2=1.

The correlation for the estimated distribution is compute as follows.

Model	$\rho(X_1, X_2)$
BGITD1*	-0.33
BGITD1^	-0.35

By comparing the sample correlation, the estimated correlation from type I BGITD seems to overestimate the correlation of the data in Table 4.6.

Chi square GOF test

A summary of the χ^2 values, log-likelihood values and the P-values are shown below.

Table 4.7: Summary statistics for Table 4.6

Model	χ^2	Log-likelihood value	P-value	Degree of freedom
BGIT1*	33.46	-502.75	0.8761	44

Table 4.7 shows that the fit of type I BGIT is significant where it provides low value of χ^2 and a high P-value.

Example 3:

Table 4.8: Number of times bacon and eggs were purchased

on four consecutive shopping trips

	0	1	2	3	4	Total
0	**254**	**115**	**42**	**13**	**6**	**430**
BNBD*	242.83	115.51	40.60	12.63	3.68	415.25
BNBD^	253.29	115.84	39.59	12.01	3.42	424.15
BGITD1*	245.31	121.92	39.17	11.23	3.15	420.79
BGITD1^	252.84	116.60	40.06	13.14	4.29	426.92
BGITD2*	242.83	115.51	40.60	12.63	3.68	415.25
BGITD2^	253.29	115.84	39.59	12.01	3.42	424.15
1	**34**	**29**	**16**	**6**	**1**	**86**
BNBD*	40.65	40.49	16.86	5.65	1.72	105.37
BNBD^	34.42	41.55	17.18	5.67	1.69	100.50
BGITD1*	41.54	31.65	15.20	6.06	2.20	96.66
BGITD1^	37.98	29.20	15.23	6.84	2.84	92.08
BGITD2*	40.65	40.49	16.86	5.65	1.72	105.37
BGITD2^	34.42	41.55	17.18	5.67	1.69	100.50
2	**8**	**8**	**3**	**3**	**1**	**23**
BNBD*	6.76	6.76	4.63	1.81	0.59	20.55
BNBD^	4.64	5.63	4.93	1.96	0.64	17.79
BGITD1*	6.70	6.98	4.40	2.21	0.97	21.26
BGITD1^	5.40	5.91	4.09	2.31	1.16	18.87
BGITD2*	6.76	6.76	4.63	1.81	0.59	20.55
BGITD2^	4.64	5.63	4.93	1.96	0.64	17.79
3	**0**	**0**	**4**	**1**	**1**	**6**
BNBD*	1.12	1.12	0.77	0.46	0.17	3.65
BNBD^	0.62	0.76	0.67	0.53	0.21	2.78
BGITD1*	1.08	1.43	1.11	0.67	0.34	4.62
BGITD1^	0.77	1.09	0.94	0.63	0.37	3.80
BGITD2*	1.12	1.12	0.77	0.46	0.17	3.65
BGITD2^	0.62	0.76	0.67	0.53	0.21	2.78
4	**1**	**1**	**1**	**0**	**0**	**3**
BNBD*	0.19	0.19	0.13	0.08	2.61	3.19
BNBD^	0.08	0.10	0.09	0.07	2.43	2.77
BGITD1*	0.17	0.28	0.26	0.18	3.78	4.67
BGITD1^	0.11	0.19	0.20	0.15	5.68	6.33
BGITD2*	0.19	0.19	0.13	0.08	2.61	3.19
BGITD2^	0.08	0.10	0.09	0.07	2.43	2.77

*= MLE, ^= PGF based estimator, BGITD1=type I BGITD, BGITD2=type II BGITD

ML estimates:

BNBD, $\hat{p}_1$=0.2659, $\hat{p}_2$=0.1078, $\hat{p}_3$=0.0044, $\hat{v}$=1.6374.

Type I BGITD, $\hat{p}_1$=0.0264, $\hat{p}_2$=0.0010, $\hat{q}_1$=0.2798, $\hat{q}_2$=0.1611, n_1=8, n_2=1.

Type II BGITD, $\hat{p}_1$=0.9130, $\hat{p}_2$=0.0010, $\hat{q}_1$=07634, $\hat{q}_2$=0.0040, $\hat{r}_1$=0.8328, $\hat{r}_2$= 0.0010,

n_1=1, n_2=2, n_3=1.

PGF-based estimates:

BNBD, $\hat{p}_1$=0.2699, $\hat{p}_2$=0.0895, $\hat{p}_3$=0.0113, $\hat{v}$=1.6639.

Type I BGITD, $\hat{p}_1$=0.0166, $\hat{p}_2$=0.0010, $\hat{q}_1$=0.3263, $\hat{q}_2$=0.1421, n_1=8, n_2=1.

Type II BGITD, $\hat{p}_1$=0.8982, $\hat{p}_2$=0.0010, $\hat{q}_1$=0.7716, $\hat{q}_2$=0.0010, $\hat{r}_1$=0.8643, $\hat{r}_2$= 0.0010,

n_1=1, n_2=2, n_3=1.

The correlations for the estimated distributions are compute as follows.

MLE

Model	$\rho(X_1, X_2)$
BNBD*	0.23
BGITD1*	0.24
BGITD2*	0.19

PGFBE

Model	$\rho(X_1, X_2)$
BNBD^	0.26
BGITD1^	0.26
BGITD2^	0.24

All of the estimated correlation values appear to be slightly lower than the sample correlation which is 0.3101.

Chi-square GOF test

The summary of the χ^2, the log-likelihood and the P- values are shown below.

Table 4.9: Summary statistics for Table 4.8

Model	χ^2	Log-likelihood value	P-value	Degree of freedom
BNBD*	42.54	-1001.43	0.0024	20
BGITD1*	30.57	-999.28	0.0611	20
BGITD2*	44.87	-1001.39	0.0004	18

All of the estimated distributions performed well as the χ^2 values are low. Based on the P-value, Type I BGITD fits better than BNBD and Type II BGITD.