Biom. J. vol. 23 · 1981 · no. 5 · pp. 437–450

Statistical Unit, University of Athens, Greece

Bivariate POISSON Binomial Distributions

CH. A. CHARALAMBIDES and H. PAPAGEORGIOU

Abstract

Three bivariate generalizations of the POISSON binomial distribution are introduced. The probabilities, moments, conditional distributions and regression functions for these distributions are obtained in terms of bipartitional polynomials. Recurrences for the probabilities and moments are also given. Parameter estimators are derived using the methods of moments and zero frequencies and the three distributions are fitted to some ecological data.

Key words: *Bivariate Poisson binomial distribution, bipartitional polynomials, method of moments, method of zero frequencies, ecological data.*

1. Introduction

The univariate POISSON binomial distribution was introduced by SKELLAM (1952) and it has found an increasing number of applications since McGUIRE et al. (1957) used the distribution to represent variation in numbers of corn-borer larvae in a randomly chosen area.

Because of its close relation with the NEYMAN type A distribution and its wide applicability, the POISSON binomial distribution has attracted the attention of several authors (see for example, SPROTT (1958), SHUMWAY and GURLAND (1960a), (1960b), KATTI and GURLAND (1962), MARTIN and KATTI (1965), GURLAND (1965), HINZ and GURLAND (1967), (1970), CHARALAMBIDES (1977a)).

Bivariate NEYMAN type A distributions were introduced by HOLGATE (1966) by considering bivariate versions of the well-known model "egg masses and larvae". GILLINGS (1974) and TALWALKER (1975) suggested that bivariate NEYMAN type A distributions or alternative models developed under similar assumptions may have applications in ecology, health services and toxicology.

In the present paper bivariate POISSON binomial distributions are introduced and various properties are studied including conditional distributions and regression functions.

Parameter estimators are derived using the methods of moments and zero frequencies (ZF) and for comparison purposes the POISSON binomial distributions are fitted to the same set of botanical data, describing the numbers of plants of species Lacistema aggregatum and Protium guianense in each of 100 quadrats of land, used by HOLGATE to illustrate the bivariate NEYMAN type A models.

2. Definitions and preliminary results

The probability generating function (p.g.f.) of a univariate POISSON distribution is given by

(2.1) $$F(t) = \exp\{\lambda(t-1)\}, \quad \lambda > 0$$

and the p.g.f. of a bivariate POISSON distribution is given by

(2.2) $$F(t_1, t_2) = \exp\{\lambda_1(t_1-1) + \lambda_2(t_2-1) + \lambda_{12}(t_1 t_2 - 1)\}, \quad \lambda_i > 0, \quad i = 1, 2, \quad \lambda_{12} > 0 .$$

With these distributions as primary we introduce the following three types of bivariate POISSON binomial distributions.

Type I: If the variable t in (2.1) is replaced by the p.g.f. of a bivariate binomial variable (Z_1, Z_2):

$$g(u, v) = \sum\sum p(r, s)\, u^r v^s = (p_{00} + p_{10}u + p_{01}v + p_{11}uv)^n$$

where

$$p(r, s) = P(Z_1 = r, Z_2 = s), \quad p_{00} + p_{10} + p_{01} + p_{11} = 1, \quad 0 < p_{ij} < 1 ,$$

the resulting p.g.f.

(2.3) $$G(u, v) = \exp[\lambda\{(p_{00} + p_{10}u + p_{01}v + p_{11}uv)^n - 1\}]$$

defines a bivariate distribution which we call bivariate POISSON binomial distribution of type I.

Type II: If the variables t_1 and t_2 in (2.2) are replaced by the p.g.f's of two independent binomial variables W_1, W_2:

$$g_i(\tau) = \sum p_i(r)\, u^r = (q_i + p_i \tau)^{n_i}, \quad i = 1, 2, \quad q_i = 1 - p_i, \quad p_i(r) = P(W_i = r)$$

the resulting p.g.f.

(2.4) $$H(u, v) = \exp[\lambda_1\{(q_1 + p_1 u)^{n_1} - 1\} + \lambda_2\{(q_2 + p_2 v)^{n_2} - 1\} + \lambda_{12}\{(q_1 + p_1 u)^{n_1}(q_2 + p_2 v)^{n_2} - 1\}]$$

defines a bivariate distribution which we call bivariate POISSON binomial distribution of type II.

Type III: If the variable t in (2.1) is replaced by the p.g.f. $g(u) = (q + pu)^n$ the resulting p.g.f. $G(u) = \exp[\lambda\{(q + pu)^n - 1\}]$ defines the well known univariate POISSON binomial distribution. Let $X_1 = X_1' + X$, $X_2 = X_2' + X$ where X_1', X_2', X are independent univariate POISSON binomial r.v's with λ_1, λ_2 and λ_{12} the parameters of the POISSON distribution respectively and n, p the parameters of the binomial distribution. Then the p.g.f. of (X_1, X_2) is given by

$$(2.5) \qquad L(u, v) = \exp\left[\lambda_1\{(q+pu)^n - 1\} + \lambda_2\{(q+pv)^n - 1\} + \lambda_{12}\{(q+puv)^n - 1\}\right].$$

This p.g.f. defines a bivariate distribution which we call bivariate POISSON binomial distribution of type III.

To visualize the first two types of bivariate POISSON binomial distributions in ecological situations let X_1 and X_2 denote the numbers of two different kinds of individuals occurring in a quadrat of land. The individuals are considered as arising from independent clusters, the number of clusters being itself a univariate random variable Y (in case of clusters of only one type) or a bivariate random variable (Y_1, Y_2) (in case of clusters of two types).

Assume that each cluster gives rise to individuals of two kinds and let Z_1, Z_2 denote the number of individuals of each kind.

(i) If (Z_1, Z_2) follows a bivariate binomial, and Y follows a univariate POISSON, then (X_1, X_2) follows a bivariate POISSON binomial distribution of type I.

(ii) If Z_1 and Z_2 are independent univariate binomial r.v's and (Y_1, Y_2) is a bivariate POISSON, then (X_1, X_2) has a bivariate POISSON binomial distribution of type II.

In type III, the contribution to (X_1, X_2) from the variable X can be considered as arising from a distinct type of cluster, which gives rise to equal numbers of individuals of the two kinds (cf. HOLGATE, 1966).

The p.g.f's (2.3), (2.4) and (2.5) are all special cases of

$$(2.6) \qquad G(u, v) = e^{h(u,v)}$$

Hence the problem of deriving the probabilities and moments of the corresponding distribution is a problem of expanding it into a power series. The coefficients of this expansion appeared to be a special case of the bipartitional polynomials introduced and studied by one of the authors (CHARALAMBIDES (1981)). These polynomials denoted by $Y_{m,k} = Y_{m,k}(y_{01}, y_{10}, y_{11}, \ldots, y_{mk})$ are multivariable polynomials defined by a sum over all partitions of their bipartite indexes (mk), and have exponential generating function

$$(2.7) \qquad \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} Y_{m,k} \frac{u^m}{m!} \frac{v^k}{k!} = \exp\{y(u, v) - y(0, 0)\}$$

where

$$(2.8) \qquad y(u, v) = \sum_{s=0}^{\infty} \sum_{r=0}^{\infty} y_{rs} \frac{u^r}{r!} \frac{v^s}{s!}.$$

Moreover they satisfy the recurrence relation

$$(2.9) \qquad Y_{m,k+1} = \sum_{s=0}^{k} \sum_{r=0}^{m} \binom{m}{r} \binom{k}{s} y_{r,s+1} Y_{m-r,k-s}, \quad Y_{0,0} = 1.$$

The corresponding unipartitional polynomials $Y_m = Y_m(x_1, x_2, \ldots, x_m)$ have exponential generating function

$$\sum_{m=0}^{\infty} Y_m \frac{u^m}{m!} = \exp\{x(u) - x(0)\} \tag{2.10}$$

where

$$x(u) = \sum_{r=0}^{\infty} x_r \frac{u^r}{r!}.$$

Since the generating function (2.7) can be written in the form

$$\exp\{y(u, v) - y(0, 0)\} = \exp\{y_0(v) - y_0(0)\} \cdot \exp\left\{\sum_{r=0}^{\infty} y_r(v) \frac{u^r}{r!} - y_0(v)\right\}$$

where

$$y_r(v) = \sum_{s=0}^{\infty} y_{rs} \frac{v^s}{s!}, \quad r = 0, 1, 2, \ldots$$

it follows that

$$\sum_{k=0}^{\infty} Y_{m,k}(y_{01}, y_{10}, y_{11}, \ldots, y_{mk}) \frac{v^k}{k!} = \exp\{y_0(v) - y_0(0)\} \times Y_m(x_1, x_2, \ldots, x_m) \tag{2.11}$$

with

$$x_r = y_r(v), \quad r = 1, 2, \ldots, m.$$

For $x_r = (n)_r\, \vartheta$, $r = 1, 2, \ldots, m$ we have $x(u) = \vartheta\,(1+u)^n$ and therefore

$$Y_m((n)_1 \vartheta, (n)_2 \vartheta, \ldots, (n)_m \vartheta)) \equiv C_{m,n}(\vartheta) = \sum_{k=0}^{m} C(m, k, n)\, \vartheta^k \tag{2.12}$$

where

$$\sum_{m=k}^{\infty} C(m, k, n) = \frac{1}{k!} [(1+u)^n - 1]^k.$$

Properties and applications of the numbers $C(m, k, n)$ have been discussed by one of the authors (CHARALAMBIDES, 1976, 1977b).

3. Properties of the distribution

3.1. *Probabilities and moments*

The probability function (p. f.), say $P(m, k; h) = P(X_1 = m,\ X_2 = k)$, with p.g.f. (2.6), on using (2.7), may be obtained as

$$P(m, k; h) = e^{h(0,0)} Y_{m,k}(h_{01}, h_{10}, h_{11}, \ldots, h_{m,k})/m!k! \tag{3.1}$$

where

$$h_{rs}=\left[\frac{\partial^s}{\partial v^s}\,\frac{\partial^r}{\partial u^r}\,h(u,\,v)\right]_{u=0,v=0}$$

Similarly the factorial moments $\mathrm{M}_{(m,k)}=\mathrm{E}\,[(X_1)_m(X_2)_k]$ may be obtained as

(3.2) $$\mathrm{M}_{(m,k)}=Y_{m,k}(c_{01},\,c_{10},\,c_{11},\,\ldots\,,\,c_{mk})$$

where

$$c_{rs}=\left[\frac{\partial^s}{\partial v^s}\,\frac{\partial^r}{\partial u^r}\,h(u,\,v)\right]_{u=1,v=1}$$

From (3.1), (3.2) and using (2.9) we get for the probabilities and moments the following recurrence relations

(3.3) $$P\,(m,\,k+1\,;\,h)=\frac{1}{k+1}\sum_{s=0}^{k}\sum_{r=0}^{m}\frac{h_{r,s+1}}{r!s!}\,P\,(m-r,\,k-s\,;\,h)$$

and

(3.4) $$\mathrm{M}_{(m,k+1)}=\sum_{s=0}^{k}\sum_{r=0}^{m}\binom{m}{r}\binom{k}{s}c_{r,s+1}\mathrm{M}_{(m-r,k-s)}\;.$$

3.2. *Conditional distributions and regression functions*

Summing (3.1) for all k we get on using (2.11) with $v=1$, the marginal p.f. of X_1

(3.5) $$P(m\,;\,h)=P\,(X_1=m)=\mathrm{e}^{h(0,1)}Y_m(h_1,\,h_2,\,\ldots\,,\,h_m)/m!$$

where

$$h_r=\left[\frac{\partial^r}{\partial u^r}\,h(u,\,1)\right]_{u=0},\quad r=1,\,2,\,\ldots\,,\,m\,.$$

Therefore the conditional p.f. of X_2 given $X_1=m$ is given by

(3.6) $$P(k\mid m\,;\,h)=\mathrm{e}^{h(0,0)-h(0,1)}\,\frac{Y_{m,k}(h_{01},\,h_{10},\,h_{11},\,\ldots\,,\,h_{m,k})}{k!\,Y_m(h_1,\,h_2,\,\ldots\,,\,h_m)}\,.$$

Note that the p.g.f. of (3.6) is given by

(3.7) $$G_{X_2|X_1}(v)=\mathrm{e}^{h(0,v)-h(0,1)}\,\frac{Y_m(h_1(v),\,h_2(v),\,\ldots\,,\,h_m(v))}{Y_m(h_1,\,h_2,\,\ldots\,,\,h_m)}$$

where

$$h_r(v)=\left[\frac{\partial^r}{\partial u^r}\,h(u,\,v)\right]_{u=0},\quad r=1,\,2,\,\ldots\,,\,m$$

and therefore the regression of X_2 on $X_1=m$ may be obtained as

(3.8) $$\mathrm{E}[X_2\mid X_1=m]=\frac{Y_{m,1}(c_{(1|0)},\,h_1c_{(0|1)},\,h_1c_{(1|1)},\,\ldots\,,\,h_mc_{(1|m)})}{Y_m(h_1,\,h_2,\,\ldots\,,\,h_m)}$$

where

$$h_r c_{(s|r)} = \left[\frac{\partial^r}{\partial u^r}\frac{\partial^s}{\partial v^s} h(u, v)\right]_{\substack{u=0\\ v=1}}$$

3.3. *Bivariate* POISSON *binomial distributions*

Type I: The p.g.f. (2.3) of this distribution is of the form (2.6) with

$$h(u, v) = \lambda\{(p_{00} + p_{10}u + p_{01}v + p_{11}uv)^n - 1\}.$$

Hence

$$h_{rs} = \lambda r!s!p(r, s), \quad c_{rs} = \lambda\mu_{(r,s)} \equiv \lambda \mathrm{E}\,[(Z_1)_r\,(Z_2)_s]$$

so that

$$P(m, k) = \mathrm{e}^{-\lambda(1-p_{00}^n)} Y_{m,k}(\lambda p(0, 1), \lambda p(1, 1), \ldots, \lambda m!k!p(m, k))$$

$$\mathrm{M}_{(m,k)} = Y_{m,k}(\lambda\mu_{(0,1)}, \lambda\mu_{(1,0)}, \ldots, \lambda\mu_{(m,k)})$$

$$P(m, k+1) = \frac{\lambda}{k+1}\sum_{s=0}^{k}\sum_{r=0}^{m}(s+1)\,p\,(r, s+1)\,P\,(m-r, k-s)$$

$$\mathrm{M}_{(m,k+1)} = \lambda\sum_{s=0}^{k}\sum_{r=0}^{m}\binom{m}{r}\binom{k}{s}\mu_{(r,s+1)}\mathrm{M}_{(m-r,k-s)}.$$

Since

$$h_r(v) = (n)_r\,(p_{01} + p_{11}v)^r\,(p_{00} + p_{01}v)^{n-r}$$

it follows from (2.12) that the p.g.f. (3.7) may be written in the form

$$G_{X_2|X_1=m}(v) = \exp\left[\lambda\{(p_{00} + p_{01}v)^n - (p_{00} + p_{01})^n\}\right]\left(\frac{p_{01} + p_{11}v}{p_{01} + p_{11}}\right)^m$$
$$\times\left(\frac{p_{00} + p_{01}v}{p_{00} + p_{01}}\right)^{-m}\frac{C_{m,n}([p_{00} + p_{01}v]^n)}{C_{m,n}([p_{00} + p_{01}]^n)}.$$

Therefore the conditional distribution is a convolution of

(i) Univariate POISSON binomial
(ii) a binomial
(iii) a mixture of binomials

with weights a combinatorial distribution (in the sence of HARPER (1967); see also CHARALAMBIDES (1976)) with p.f.

$$(3.9)\qquad p(k; \vartheta, m, n) = \frac{C(m, k, n)\,\vartheta^k}{C_{m,n}(\vartheta)}, \quad k = 0, 1, 2, \ldots, m, \quad \vartheta = (p_{00} + p_{01})^n.$$

Using the recurrence relation (see CHARALAMBIDES (1977b))

$$C\,(m+1, k, n) = (nk-m)\,C(m, k, n) + nC\,(m, k-1, n)$$

we get the expectation of (3.9) in the form

$$\sum_{k=0}^{m} kp(k;\vartheta, m, n)=\frac{C_{m+1,n}(\vartheta)}{nC_{m,n}(\vartheta)}-\frac{n\vartheta-m}{n}.$$

Summing the expectations of the convolutes we get the regression of X_2 on $X_1=m$ in the form

$$\begin{aligned}
\mathrm{E}\,(X_2 \mid X_1=m)&=\frac{\lambda np_{01}}{p_{00}+p_{01}}+\frac{mp_{11}}{p_{01}+p_{11}}-\frac{mp_{01}}{p_{00}+p_{01}}+\frac{np_{01}}{p_{00}+p_{01}}\\
&\quad\times\left[\frac{C_{m+1,n}([p_{00}+p_{01}]^n)}{nC_{m,n}([p_{00}+p_{01}]^n)}-(p_{00}+p_{01})^n\right]+\frac{np_{01}}{p_{00}+p_{01}}\,\frac{m}{n}\\
&=\frac{np_{01}}{p_{00}+p_{01}}\left[\frac{C_{m+1,n}([p_{00}+p_{01}]^n)}{nC_{m,n}([p_{00}+p_{01}]^n)}-(p_{00}+p_{01})^n+\lambda\right]\\
&\quad+\frac{mp_{11}}{p_{01}+p_{11}}.
\end{aligned}$$

Type II. In this case

$$\begin{aligned}
h(u, v)&=\lambda_1\{(q_1+p_1u)^{n_1}-1\}+\lambda_2\{(q_2+p_2v)^{n_2}-1\}\\
&\quad+\lambda_{12}\{(q_1+p_1u)^{n_1}(q_2+p_2v)^{n_2}-1\}.
\end{aligned}$$

Hence

$$h_{r0}=\{\lambda_1+\lambda_{12}q_2^{n_2}\}(n_1)_r p_1^r q_1^{n_1-r},\quad r>0,\quad h_{0s}=\{\lambda_2+\lambda_{12}q_1^{n_1}\}(n_2)_s p_2^s q_2^{n_2-s},\quad s>0,$$

$$h_{rs}=\lambda_{12}(n_1)_r(n_2)_r p_1^r q_1^{n_1-r} p_2^s q_2^{n_2-s},\quad r, s>1.$$

$$c_{r0}=(\lambda_1+\lambda_{12})(n_1)_r p_1^r,\quad r>0,\quad c_{0s}=(\lambda_2+\lambda_{12})(n_2)_s p_1^s,\quad s>0$$

$$c_{rs}=\lambda_{12}(n_1)_r(n_2)_s p_1^r p_2^s,\quad r, s>1.$$

Introducing these values in (3.1) and (3.2) we get the probabilities and moments of the distribution. The recurrence relations (3.3) and (3.4) reduce to the following

$$\begin{aligned}
P(m, k+1)&=\frac{1}{k+1}\left\{(\lambda_2+\lambda_{12}q_1^{n_1})\,q_2^{n_2}\sum_{s=1}^{k+1}\binom{n_2}{s}s\left(\frac{p_2}{q_2}\right)^s P(m, k-s+1)\right.\\
&\left.\quad+\lambda_{12}q_1^{n_1}q_2^{n_2}\sum_{s=1}^{k+1}\sum_{r=1}^{m}\binom{n_1}{r}\binom{n_2}{s}s\left(\frac{p_1}{q_1}\right)^r\left(\frac{p_2}{q_2}\right)^s P(m-r, k-s+1)\right\}
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{M}_{(m,k+1)}&=(\lambda_1+\lambda_{12})\sum_{s=0}^{k}\binom{k}{s}\binom{n_2}{s+1}(s+1)!\,p_1^{s+1}\mathrm{M}_{(m,k-s)}\\
&\quad+\lambda_{12}\sum_{s=0}^{k}\sum_{r=1}^{m}\binom{m}{r}\binom{k}{s}(n_1)_r(n_2)_{s+1}\,p_1^r p_2^{s+1}\mathrm{M}_{(m-r,k-s)}.
\end{aligned}$$

Since

$$h_r(v)=[\lambda_1+\lambda_{12}(q_2+p_2v)^{n_2}]\,(n_1)_r\,p_1^r q_1^{n_1-r}$$

it follows from (2.12) that the p.g.f. (3.7) reduce to

$$G_{X_2|X_1=m}(v)=\exp[(\lambda_2+\lambda_{12}q_1^{n_1})\{(q_2+p_2v)^{n_2}-1\}]$$
$$\times\frac{C_{m,n_1}(q_1^{n_1}[\lambda_1+\lambda_{12}(q_2+p_2v)]^{n_2})}{C_{m,n_1}(q_1^{n_1}[\lambda_1+\lambda_{12}])}$$

from which we conclude that one of the convolutes is a univariate POISSON binomial.

The regression of X_2 on $X_1=m$ is given by

$$\mathrm{E}(X_2\mid X_1=m)=\lambda_2n_2p_2+\frac{n_2p_2\lambda_{12}}{n_1(\lambda_1+\lambda_{12})}\left[\frac{C_{m+1,n_1}(q^{n_1}[\lambda_1+\lambda_{12}])}{C_{m,n_1}(q_1^{n_1}[\lambda_1+\lambda_{12}])}+m\right].$$

Type III. The p.g.f. (2.5) is of the form (2.6) with

$$h(u,v)=\lambda_1\{(q+pu)^n-1\}+\lambda_2\{(q+pv)^n-1\}+\lambda_{12}\{(q+puv)^n-1\}.$$

Therefore the p.f. and factorial moments of this distribution are given by (3.1) and (3.2) respectively with

$$h_{r0}=(n)_r\lambda_1p^rq^{n-r},\quad r\geqq1,\quad h_{0s}=(n)_s\lambda_2p^sq^{n-s},\quad s\geqq1$$
$$h_{rr}=(n)_r r!\lambda_{12}p^rq^{n-r},\quad h_{rs}=0,\quad r\neq s,\quad r,s\geqq1\,.$$
$$c_{r0}=(\lambda_1+\lambda_{12})(n)_r p^r,\quad r\geqq1,\quad c_{0s}=(\lambda_2+\lambda_{12})(n)_r p^r,\quad s\geqq1\,.$$
$$c_{rs}=\lambda_{12}(n)_r\sum_{k=r}^{n}\binom{n-r}{k-r}(k)_s p^kq^{n-k},\quad r,s\geqq1\,.$$

The p.g.f. of the conditional distribution of X_2 given $X_1=m$ may be obtained in the form

$$G_{X_2|X_1=m}(v)=\exp[\lambda_2\{(q+pv)^n-1\}]$$
$$\times\frac{Y_m(nq^n[\lambda_1+\lambda_{12}v],\ldots(n)_m q^n[\lambda_1+\lambda_{12}v^m])}{C_{m,n}(q^n[\lambda_1+\lambda_{12}])}.$$

Therefore one of the convolutes is again a univariate POISSON binomial.

The regression of X_2 on $X_1=m$ may be obtained as

$$\mathrm{E}(X_2\mid X_1=m)=\lambda_2np+m\sum_{k=1}^{m}\binom{m-1}{k-1}\frac{C_{m-k,n}(\lambda_1q^n)\,C_{k,n}(\lambda_{12}q^n)}{C_{m,n}([\lambda_1+\lambda_{12}]\,q^n)}.$$

4. Estimation

Although various estimation procedures for the univariate POISSON binomial distribution have been discussed, attempts to fit this distribution for known exponent n greater than two have not been made very often, because, apart from inherent difficulties involved in the fitting procedures, it is frequently the case that the distribution rapidly approaches the NEYMAN type A distribution as n increases.

By analogy to the univariate case, assuming that in bivariate POISSON binomial models the exponent(s) n is small and known, moment and zero frequency estimators of the parameters are derived.

4.1. *Method of Moments*

Let $(\bar{x}_1, \bar{x}_2)$ be the marginal means and $(s_{x_1x_1}, s_{x_2x_2}, s_{x_1x_2})$ the unbiased estimates of the second order moments. Then moment estimators are simply derived as:
For type I:

$$\hat{\lambda}=\frac{(n-1)\,(\bar{x}_1^2+\bar{x}_2^2)}{n\,(s_{x_1x_1}+s_{x_2x_2}-\bar{x}_1-\bar{x}_2)}\,,$$

$$\hat{p}_{10}=\frac{n\hat{\lambda}\,(\bar{x}_1-s_{x_1x_2})+(n-1)\,\bar{x}_1\bar{x}_2}{(n\hat{\lambda})^2}\,,$$

$$\hat{p}_{01}=\frac{\bar{x}_2-\bar{x}_1}{n\hat{\lambda}}+\hat{p}_{10}\,,$$

$$\hat{p}_{11}=\frac{\bar{x}_1}{n\hat{\lambda}}-\hat{p}_{10}\,.$$

For type II:

$$\hat{p}_1=\frac{s_{x_1x_1}-\bar{x}_1}{(n_1-1)\,\bar{x}_1}\,,$$

$$\hat{p}_2=\frac{s_{x_2x_2}-\bar{x}_2}{(n_2-1)\,\bar{x}_2}\,,$$

$$\hat{\lambda}_{12}=\frac{s_{x_1x_2}}{n_1n_2\hat{p}_1\hat{p}_2}\,,$$

$$\hat{\lambda}_1=\frac{\bar{x}_1}{n_2\hat{p}_2}-\hat{\lambda}_{12}\,,$$

$$\hat{\lambda}_2=\frac{\bar{x}_2}{n_2\hat{p}_2}-\hat{\lambda}_{12}\,.$$

For type III:

$$\hat{p}=\frac{1}{n-1}\left[\frac{s_{x_1x_1}+s_{x_2x_2}}{\bar{x}_1+\bar{x}_2}-1\right],$$

$$\hat{\lambda}_{12}=\frac{s_{x_1x_2}}{n\hat{p}\,[(n-1)\,\hat{p}+1]}\,,$$

$$\hat{\lambda}_1=\frac{x_1}{n\hat{p}}-\hat{\lambda}_{12}\,,$$

$$\hat{\lambda}_2=\frac{\bar{x}_2}{n\hat{p}}-\hat{\lambda}_{12}\,.$$

4.2. *Method of Zero Frequencies*

It was pointed out by PARAGEORGIOU (1979) that for bivariate discrete distributions parameter estimators can be obtained using the marginal means $(\bar{x}_1, \bar{x}_2)$, the proportion of observations $(f_{0,0})$ in the (0, 0) cell, the proportion of zeros $(f_{0.})$ in the X_1 margin and/or the proportion of zeros $(f_{.0})$ in the X_2 margin.

Consequently, for the three bivariate POISSON binomial distributions parameter estimators can be easily derived from the following systems of equations.

For type I:

$$p_{10}+p_{01}+p_{11}=1-\left\{\frac{1}{\lambda}\log(f_{0,0})+1\right\}^{1/n},$$

$$\lambda\{(1-p_{10}-p_{11})^n-1\}=\log(f_{0.}), \tag{4.1}$$

$$n\lambda(p_{10}+p_{11})=\bar{x}_1, \tag{4.2}$$

$$n\lambda(p_{01}+p_{11})=\bar{x}_2.$$

λ may be eliminated between equations (4.1) and (4.2) giving the equation

$$\frac{\{1-(p_{10}+p_{11})\}^n-1}{p_{10}+p_{11}}=\frac{n\log(f_{0.})}{\bar{x}_1}, \tag{4.3}$$

which can be solved iteratively for $(p_{10}+p_{11})$. In particular for $n=2$

$$p_{10}+p_{11}=\frac{2\log(f_{0.})}{\bar{x}_1}+2.$$

For type II:

$$\lambda_1\{(1-p_1)^{n_1}-1\}+\lambda_2\{(1-p_2)^{n_2}-1\}+\lambda_{12}\{(1-p_1)^{n_1}(1-p_2)^{n_2}-1\}$$
$$=\log(f_{0,0}),$$
$$(\lambda_1+\lambda_{12})\{(1-p_1)^{n_1}-1\}=\log(f_{0.}),$$
$$(\lambda_2+\lambda_{12})\{(1-p_2)^{n_2}-1\}=\log(f_{.0}),$$
$$np_1(\lambda_1+\lambda_{12})=\bar{x}_1,$$
$$np_2(\lambda_2+\lambda_{12})=\bar{x}_2.$$

For type III:

$$\{(1-p)^n-1\}(\lambda_1+\lambda_2+\lambda_{12})=\log(f_{0,0}),$$
$$\{(1-p)^n-1\}(\lambda_1+\lambda_{12})=\log(f_{0.}),$$
$$np(\lambda_1+\lambda_{12})=\bar{x}_1,$$
$$np(\lambda_2+\lambda_{12})=\bar{x}_2.$$

4.3. *An Illustrative Example*

For comparison purposes the POISSON binomial distributions are fitted to the same set of botanical data used by HOLGATE. The exponent(s) is assumed known and equal to two. In this case the recurrences for the probabilities are simplified

and given by:
For type I

$$P(m, k+1) = \frac{2\lambda}{k+1} \{p_{01}p_{00}P(m, k) + p_{01}^2 P(m, k-1)$$
$$+ (p_{10}p_{01} + p_{11}p_{00}) P(m-1, k)$$
$$+ p_{11}p_{10}P(m-2, k) + 2p_{11}p_{01}P(m-1, k-1)$$
$$+ p_{11}^2 P(m-2, k-1)\} .$$

A similar expression for $P(m+1, k)$ can be obtained.
For type II:

$$P(m, k+1) = \frac{2}{k+1} \{(\lambda_2 + \lambda_{12}q_1^2) p_2 q_2 P(m, k) + (\lambda_2 + \lambda_{12}q_1^2) p_2^2 P(m, k-1)$$
$$+ 2\lambda_{12}p_1p_2q_1q_2P(m-1, k) + \lambda_{12}p_1^2p_2q_2P(m-2, k)$$
$$+ 2\lambda_{12}p_1p_2^2q_1P(m-1, k-1) + \lambda_{12}p_1^2p_2^2P(m-2, k-1)\} .$$

For type III:

$$P(m, k+1) = \frac{2p}{k+1} \{\lambda_2 q P(m, k) + \lambda_2 p P(m, k-1)$$
$$+ \lambda_{12} q P(m-1, k) + \lambda_{12} p P(m-2, k-1)\} .$$

Tables 1, 2 and 3 give the moments and ZF estimates of the parameters for the bivariate POISSON binomial distributions.

Table 1

POISSON binomial type I	λ	p_{10}	p_{01}	p_{11}
Moments	1.1725	0.4614	0.3122	−0.0563(0.00)
ZF	1.3006	0.3564	0.2218	0.0088

The moment negative estimate of p_{11} was revised to satisfy the inequalities imposed on the parameter (cf. HOLGATE, 1966). The revised value is given in brackets after the original moment estimate.

Table 2

POISSON binomial type II	λ_1	λ_2	λ_{12}	p_1	p_2
Moments	0.4535	2.1955	0.5032	0.4965	0.1112
ZF	0.0868	0.4144	1.2139	0.3652	0.1842

Table 3

POISSON binomial type III	λ_1	λ_2	λ_{12}	p
Moments	1.2488	0.7450	0.1187	0.3474
ZF	0.9855	0.5063	0.3152	0.3652

Table 4
Bivariate POISSON binomial distributions fitted to botanical data by the method of moments

x_2	x_1	observed	Type I	Type II	Type III
0	0	34	32.88	30.06	29.74
0	1	12	8.05	12.79	16.84
0	2	4	9.19	9.03	9.25
0	3	5	2.09	3.07	3.44
0	4	2	1.28	1.27	1.18
1	0	8	5.45	13.80	10.05
1	1	13	12.44	7.37	7.29
1	2	3	4.24	5.52	4.03
1	3	3	3.45	2.17	1.66
1	4	0	0.92	0.96	0.58
2	0	3	4.21	4.03	4.37
2	1	6	2.87	2.49	3.01
2	2	1	3.50	1.96	2.13
2	3	2	1.24	0.86	0.94
2	4	0	0.80	0.40	0.38
3	0	1	0.65	0.88	1.09
3	1	1	1.58	0.62	0.85
3	2	0	0.84	0.51	0.63
3	3	1	0.73	0.25	0.31
3	4	0	0.25	0.12	0.13
$x_2 \geqq 4$	$x_1 \leqq 4$	0	1.60	0.57	1.09
	$x_1 \geqq 5$	1	1.73	1.27	0.99
χ^2		–	17.31	18.15	15.36
D. F.		–	8	7	8

Table 5
Bivariate POISSON binomial distributions fitted to botanical data by the method of ZF

x_2	x_1	observed	Type I	Type II	Type III
0	0	34	34.00	34.00	34.00
0	1	12	13.02	14.10	15.54
0	2	4	8.11	6.98	8.02
0	3	5	2.47	2.09	2.58
0	4	2	0.91	0.63	0.82
1	0	8	8.10	9.23	7.98
1	1	13	10.42	9.58	8.62
1	2	3	5.01	5.94	4.15
1	3	3	2.44	2.43	1.78
1	4	0	0.81	0.86	0.57
2	0	3	3.14	2.30	3.23
2	1	6	3.12	3.16	2.64
2	2	1	2.34	2.51	3.09
2	3	2	1.07	1.34	1.34
2	4	0	0.45	0.57	0.59
3	0	1	0.60	0.40	0.61
3	1	1	0.94	0.73	0.75
3	2	0	0.67	0.72	0.78
3	3	1	0.38	0.47	0.58
3	4	0	0.16	0.24	0.25
$x_2 \geqq 4$	$x_1 \leqq 4$	0	0.82	0.68	1.05
	$x_1 \geqq 5$	1	1.02	1.02	1.02
χ^2		–	10.48	14.42	14.70
D. F.		–	8	7	8

Tables 4 and 5 show the observed and expected values when the distributions are fitted by the methods of moments and ZF. Values of χ^2 were computed after the grouping of cells indicated in the table (same groups used by HOLGATE).

4.4. *Discussion and Conclusions*

Although the bivariate POISSON binomial distributions were not fitted satisfactorily by the method of moments, ZF estimators provided an acceptable fit for all three types. Moreover, the χ^2 value for the POISSON binomial type I fitting based on ZF was smaller than the corresponding χ^2 values for the NEYMAN A type I fitting, based on moments, maximum likelihood or minimum chi-square, as calculated by GILLINGS (1974).

Concluding, the bivariate POISSON binomial distributions can be regarded as the natural complement to the bivariate NEYMAN type A distributions and a useful alternative in studying bivariate discrete data.

Zusammenfassung

Es werden drei zweidimensionale Verallgemeinerungen der POISSONschen binomialen Verteilung eingeführt. Die Wahrscheinlichkeiten, Momente, bedingten Verteilungen und Regressionsfunktionen werden als Polynomterme angegeben sowie auch rekursive Relationen für die Wahrscheinlichkeiten und Momente erhalten. Mit Hilfe der Momentenmethode und der Nullfrequenzen werden Schätzfunktionen für die Parameter errechnet. Alle drei Verteilungen werden bestimmten ökologischen Daten angepaßt.

References

CHARALAMBIDES, CH. A., 1976: The asymptotic normality of certain combinatorial distributions. Ann. Inst. Statist. Math. **28**, 499–506.

CHARALAMBIDES, CH. A., 1977a: On the generalized discrete distributions and the Bell polynomials. Sankhyā **B39**, 36–44.

CHARALAMBIDES, CH. A., 1977b: A new kind of numbers appearing in the n-fold convolution of truncated binomial and negative binomial distributions. Siam J. Appl. Math. **33**, 279–288.

CHARALAMBIDES, CH. A., 1981: Bipartitional polymonials and their applications in combinatorics and statistics. Discrete Mathematics **34**, 81–84.

GILLINGS, D. B., 1974: Some further results for bivariate generalizations of the NEYMAN type A distribution. Biometrics **30**, 619–628.

GURLAND, J., 1965: A method of estimation for some generalized Poisson distributions. Proceedings of the International Symposium on classical and contagious discrete distributions, Montreal, Pergamon Press, 141–158.

HARPER, L. H., 1967: Stirling behavior is asymptotically normal. Ann. Math. Statist. **38**, 410–414.

HINZ, P., and J. GURLAND, 1967: Simplified techniques for estimating parameters of some generalized POISSON distributions. Biometrika **54**, 555–566.

HINZ, P., and J. GURLAND, 1970: A test of fit for the negative binomial and other contagious distributions. J. Amer. Statist. Assoc. **65**, 887–903.

HOLGATE, P., 1966: Bivariate generalizations of the NEYMAN's type A distribution. Biometrika **53**, 241–245.

KATTI, S. K., and J. GURLAND, 1962: Some methods of estimation for the POISSON binomial distribution. Biometrics **18**, 42–51.

MARTIN, D. C., and S. K. KATTI, 1965: Fitting of certain contagious distributions to some available data by the maximum likelihood method. Biometrics **21**, 34–48.

McGUIRE, J. U., T. A. BRINDLEY, and T. A. BRANCROFT, 1957: The distribution of European corn borer larvae pyrausta nubilalis (H.B.H.) in field corn. Biometrics 17, 65–78.

PAPAGEORGIOU, H., 1979: Zero frequency estimation for bivariate generalized Poisson distributions. Proceedings of the 42nd session of the International Statistical Institute, Contributed Papers, 397–400.

SHUMWAY, R., and J. GURLAND, 1960a: Fitting the Poisson binomial distribution. Biometrics 16, 522–533.

SHUMWAY, R., and J. GURLAND, 1960b: A fitting procedure for some generalized Poisson distributions. Skand. Aktuar. 43, 87–108.

SKELLAM, J. C., 1952: Studies in statistical ecology. 1. Spatial pattern. Biometrika 39, 346–362.

SPROTT, D. A., 1958: The method of maximum likelihood applied to the Poisson binomial distribution. Biometrics 14, 97–106.

TALWALKER, S., 1975: Models in medicine and toxicology. Statistical distributions in scientific work, Vol. 2, 263–274, Reidel Dordrecht, Holland.

Manuscript received: 31. 1. 1980

Authors' address:

Dr. CH. A. CHARALAMBIDES, Dr. H. PAPAGEORGIOU
Statistical Unit, University of Athens,
Panepistemiopolis, Athens (621)
Greece