Bivariate ZIP models

J.F. Walhin

Université Catholique de Louvain

Secura Belgian Re

Abstract

The Zero-Inflated Poisson model is extended into a bivariate form. Three new bivariate models are considered. Parameters are estimated by maximum likelihood. Two numerical examples are discussed.

Keywords

Zero-Inflated Poisson, Bivariate Poisson, Bivariate mixed Poisson distribution, trivariate reduction method, recursion, maximum likelihood.

1. Introduction

In a recent paper Böhning (1998) discussed the use of Zero-Inflated Poisson models. In his paper, Böhning (1998) stated that it is not clear how to proceed with bivariate or multivariate distributional forms. The aim of this paper is to show how to use the Zero-Inflated Poisson model in the bivariate setting.

In section 2 we will recall the concept of Zero-Inflated Poisson models. Section 3 will be devoted to three extensions of the Zero-Inflated Poisson model in the bivariate setting. Section 4 gives two numerical fits.

In the present paper we just analyse the case without covariates while Lambert (1992) studied that subject in the univariate ZIP model. Note that univariate ZIP models have been investigated by van den Broek (1995) where a score test for zero inflation in a Poisson distribution were derived.

2. Univariate Zero-Inflated Poisson model

The Zero-Inflated Poisson model is used when a count data set shows a large proportion of zeroes. The probability function is given by

$$f(y|\lambda, p) = pPo(y, 0) + (1-p)Po(y, \lambda)$$

i.e. when $y = 0$ we find $p + (1-p)Po(0, \lambda)$ whereas when $y > 0$ we have $(1-p)Po(y, \lambda)$. This is a Poisson model with extra mass p at zero.
When a random variable N with values in $0, 1, 2, \ldots$ is ZIP distributed, we will write

$$N \sim ZIP(p, \lambda)$$

As mentioned by Böhning (1998), the Zero-Inflated Poisson model is a particular case of the class of non-parametric mixtures models :

$$p_1 Po(y, \lambda_1) + \cdots + p_k Po(y, \lambda_k), \quad y = 0, 1, 2, \ldots$$

with $p_j \geq 0$ and $p_1 + \cdots + p_k = 1$. The non-parametric aspect refers to the mixing distribution.
In the next section, we propose three possible bivariate extensions for the ZIP model.
An anonymous referee made me aware of the paper by Li et al. (1999) where a multivariate ZIP model is described. If we restrict ourselves to the bivariate case, the model of Li et al. (1999) is encompassing our models 1 and 2. However the extra parameters of that extended model do not provide better fits with respect to the principle of parcimony. We will therefore concentrate on the physical derivation of our models.

3. Bivariate Zero-Inflated Poisson models

Let us define the probability function of the joint distribution of (N, M) :

$$f(n, m) = \mathbb{P}[N = n, M = m]$$

Model 1

Let Λ denote the random variable taking values λ in the parameter space $[0, \infty)$ with mixing cumulative density function U.
We are interested in studying a mixed bivariate distribution

$$f(n, m) = \int_0^\infty \mathbb{P}(N = n, M = m | \Lambda = \lambda) dU(\lambda) \tag{1}$$

Furthermore, we assume that

- Conditionally on $\Lambda = \lambda$ the random variables N and M are independent

- The conditional distributions of N and M are univariate Poisson with parameter respectively λ and $v\lambda$, $\lambda > 0$, $v > 0$

It is easy to deduce the unconditional bivariate probability function :

$$f(n,m) = C^n_{n+m} \frac{v^m}{(1+v)^{n+m}} \int_0^\infty e^{-(1+v)\lambda} \frac{((1+v)\lambda)^{n+m}}{(n+m)!} dU(\lambda)$$

We will assume a non-parametric form of the mixing distribution U for the random variable Λ with two mass points : 0 and λ with probabilities p and $1-p$.
This gives

$$f(n,m) = pPo(n,0)Po(m,0) + (1-p)Po(n,\lambda)Po(m,v\lambda)$$

Using classical results related with Poisson mixtures (see Lindsay (1995)) we know that the maximum likelihood estimate of the mean is given by the empirical mean :

$$\bar{N} + \bar{M} = (1-\hat{p})\hat{\lambda}(1+\hat{v}) \tag{2}$$

where $\bar{X}$ is the experimental mean of X.
Let us assume that we have observed a sample (n_i, m_i) , $1 \le i \le \alpha$ of (N, M).
By maximum likelihood we get

$$\begin{aligned} l_{n,m}(v,p,\lambda) &= \ln \prod_{i=1}^{\alpha} f(n_i, m_i) \\ &= C + v\sum_{i=1}^{\alpha} m_i - (1+v)\sum_{i=1}^{\alpha}(n_i+m_i) + l_{n+m}(v,p,\lambda) \end{aligned}$$

where $l_{n+m}(v,p,\lambda)$ is the loglikelihood for the univariate distribution $ZIP(p,(1+v)\lambda)$ with the sample $(n_i+m_i)\,, 1 \le i \le \alpha$ and $l_{n,m}(v,p,\lambda)$ is the loglikelihood for the bivariate ZIP distribution with parameters p, v and λ with the sample $(n_i, m_i)\,, 1 \le i \le \alpha$.
As shown in Besson and Partrat (1992), $\frac{d}{dv} l_{n+m}(v,p,\lambda) = 0$ at the maximum likelihood estimate. So we immediately get

$$\hat{v} = \frac{\bar{M}}{\bar{N}} \tag{3}$$

Moreover it is easy to show that $\frac{\partial l_{n,m}(v,p,\lambda)}{\partial p} = 0$ implies

$$\frac{\alpha_{0,0}}{\alpha} = \hat{p} + (1-\hat{p})e^{-(1+\hat{v})\hat{\lambda}} \tag{4}$$

where $\alpha_{i,j}$ denotes the number of observations (i,j) in the sample of size α.
Isolating $\hat{\lambda}$ in (2) and inserting $\hat{\lambda}$ and $\hat{v}$ (see (3)) in (4) allows to solve numerically (4) for $\hat{p}$. Therefore, only one parameter has to be estimated on a numerical basis when looking for the maximum likelihood.
Obviously, as is stated in section 2, this model is a particular case of the general non-parametric mixture. Note that in the case where Λ is gamma distributed we find a bivariate negative binomial model.

Model 2

The classical bivariate Poisson distribution is modelled by its probability generating function (see Kocherlakota and Kocherlakota (1992)) :

$$Q(u,v) = e^{\lambda_1(u-1)+\lambda_2(v-1)+\lambda_3(uv-1)}$$

It is not difficult to find a more general bivariate distribution by mixing :

$$Q(u,v|\Lambda=\lambda) = e^{\lambda\lambda_1(u-1)+\lambda\lambda_2(v-1)+\lambda\lambda_3(uv-1)}$$

By choosing the non-parametric structure for the random variable Λ with two mass points : 0 and 1 with probabilities p and $1-p$ we have

$$P(u,v) = p + (1-p)e^{\lambda_1(u-1)+\lambda_2(v-1)+\lambda_3(uv-1)}$$

This is clearly a zero-inflated model in the bivariate setting.
Derivating $P(u,v)$ with respect to u on the one hand and with respect to v on the other hand gives

$$\frac{\partial P(u,v)}{\partial u} = \lambda_1 P(u,v) + \lambda_3 v P(u,v) - p\lambda_1 - p\lambda_3 v \tag{5}$$

$$\frac{\partial P(u,v)}{\partial v} = \lambda_2 P(u,v) + \lambda_3 u P(u,v) - p\lambda_2 - p\lambda_3 u \tag{6}$$

Let us recall that $P(u,v)$ is the probability generating function of our model 2, that is

$$P(u,v) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} f(n,m)u^n v^m$$

Then equating the coefficients of the terms in $u^n v^m$ in the equations (5) and (6) gives the probability function recursively :

$$\begin{array}{lrcl}
 & f(0,0) & = & p + (1-p)e^{-\lambda_1-\lambda_2-\lambda_3} \\
 & f(1,0) & = & \lambda_1 f(0,0) - p\lambda_1 \\
 & f(1,1) & = & \lambda_1 f(0,1) + \lambda_3 f(0,0) - p\lambda_3 \\
(7) & f(n,m) & = & \frac{1}{n}\left(\lambda_1 f(n-1,m) + \lambda_3 f(n-1,m-1)\right) \quad n>0, (n,m) \neq (1,0),(1,1) \\
 & f(0,1) & = & \lambda_2 f(0,0) - p\lambda_2 \\
 & f(1,1) & = & \lambda_2 f(1,0) + \lambda_3 f(0,0) - p\lambda_3 \\
(8) & f(n,m) & = & \frac{1}{m}\left(\lambda_2 f(n,m-1) + \lambda_3 f(n-1,m-1)\right) \quad m>0, (n,m) \neq (0,1),(1,1)
\end{array}$$

As we are interested in the maximum likelihood estimation, we write the loglikelihood :

$$l_{n,m}(p,\lambda_1,\lambda_2,\lambda_3) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \alpha_{n,m}\ln(f(n,m))$$

The maximum likelihood estimates of model 2 are given as the solution of the following system:

$$
\begin{aligned}
\frac{\alpha_{0,0}}{\alpha} &= \hat{p} + (1-\hat{p})e^{-\hat{\lambda}_1-\hat{\lambda}_2-\hat{\lambda}_3} \\
\bar{N} &= (\hat{\lambda}_1+\hat{\lambda}_3)(1-\hat{p}) \\
\bar{M} &= (\hat{\lambda}_2+\hat{\lambda}_3)(1-\hat{p}) \\
0 &= -\alpha + \frac{\alpha_{0,0}p}{f(0,0)} - \frac{\alpha_{1,1}p}{f(1,1)} + \frac{\alpha_{1,1}f(0,0)}{f(1,1)} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\alpha_{n,m}\frac{f(n-1,m-1)}{f(n,m)}
\end{aligned}
$$

The proof of this result is given in the appendix.

Model 3

A classical way of obtaining a bivariate distribution is to use the trivariate reduction method:

$$
\begin{aligned}
N &= N_0 + N_1 \\
M &= N_0 + N_2
\end{aligned}
$$

where the N_i , $i = 0, 1, 2$ are independent random variables with values in $0, 1, 2, \ldots$. If the N_i are Poisson distributed, we find the classical bivariate Poisson distribution. We will now assume that the N_i are ZIP distributed. This model is certainly less intuitive than the previous two in the sense that the construction of the bivariate dependence is more technical. However, as we will see in the next section, it gives good fits.
Let us assume that

$$
\begin{aligned}
\mathbb{P}[N_i = 0] &= p_i + (1-p_i)e^{-\lambda_i} \\
\mathbb{P}[N_i = y] &= (1-p_i)\frac{\lambda_i^y}{y!}e^{-\lambda_i} \quad , \quad y > 0
\end{aligned}
$$

A particular case would be to assume that $p_1 = p_2 = 0$. This would be the situation where the zero inflation is assumed to come only from the common N_0 (this situation is denoted Model 3 bis).
The probability function of (N, M) is given by

$$
f(n,m) = \sum_{k=0}^{\min(n,m)} \mathbb{P}[N_0 = k]\mathbb{P}[N_1 = n-k]\mathbb{P}[N_2 = m-k]
$$

Due to the form of the probability generating function of this bivariate model :

$$
P(u,v) = (p_0 + (1-p_0)e^{-\lambda_0(1-uv)})(p_1 + (1-p_1)e^{-\lambda_1(1-u)})(p_2 + (1-p_2)e^{-\lambda_2(1-v)})
$$

it does not seem to be possible to find an easy recursion in order to find the probability function. Two parameters can easily be eliminated from the numerical maximisation. Indeed we have the moment equations :

$$
\begin{aligned}
\bar{N}_0 + \bar{N}_1 &= (1-\hat{p}_0)\hat{\lambda}_0 + (1-\hat{p}_1)\hat{\lambda}_1 \\
\bar{N}_0 + \bar{N}_2 &= (1-\hat{p}_0)\hat{\lambda}_0 + (1-\hat{p}_2)\hat{\lambda}_2
\end{aligned}
$$

4. Numerical example

We will fit two data sets. The first one is taken from Cresswell and Frogatt (1963) and gives the number of accidents caused by drivers in two non-overlapping periods of time. The second one is taken from Kocherlakota and Kocherlakota (1992) and gives the number of plants of the species Lacistema aggregatum and Protium guianense in each of 100 systematically laid and contiguous quadrats.
Models 1,2, 3 and 3 bis will be adjusted. For the point of comparison, the bivariate Poisson as well as the bivariate independent Poisson models will be adjusted. The maximum likelihood estimation is used.
In order to be able to compare the loglikelihoods (denoted by l), let us note that model 3 bis is a particular case of model 3 ($p_1 = p_2 = 0$). The bivariate Poisson distribution is a particular case of model 3 bis ($p_0 = 0$) and of of model 2 ($p = 0$). The bivariate independent Poisson distribution is a particular case of model 1 ($p = 0$) and of the bivariate Poisson distribution ($\lambda_3 = 0$). Note also that model 1 is a particular case of model 2 ($\lambda_3 = 0$). Model 1 and model 2 are particular cases of the model of Li et al. (1999).
χ^2 tests are provided as well as the AIC and BIC criteria.

Data set 1 : N and M give the number of accidents in two non-overlapping periods of time:

INSERT TABLE 1 HERE

The grouping rule for the χ^2 statistic has been chosen so that the expected frequencies are most of the time greater 5 than and always greater than 1. We adopt the following notation : $(i, \geq j)$ means the cells corresponding to $N = i$ and $M \geq j$. Furthermore $(i, [j, k])$ means the cells corresponding to $N = i$ and $j \leq M \leq k$.
The following cells have been grouped : $(0; \geq 5), (1; \geq 5), (2; \geq 5), (3; \geq 4), (4; [0, 1, 2])$, the remaining cells with frequency lower than 1.
The characteristics of the fits are given in the next table :

INSERT TABLE 2 HERE

The maximum likelihood estimates for each model are given in the following table

INSERT TABLE 3 HERE

The fits show that the classical Bivariate Poisson model has to be rejected. Model 3 is acceptable mainly due to its large number of parameters. Model 2 is also acceptable and according to the BIC, it is the best model.
Data set 2 : N and M the number of plants of the species Lacistema aggregatum and Protium guianense :

INSERT TABLE 4 HERE

The cells have been grouped for the calculation of the χ^2 statistic as :
$(0;0),(0;1),(0;\geq 2),(1;0),(1;1),(1;\geq 2),(2;0),(2;1),(3;[0,1])$, the remaining cells.
The characteristics of the fits are given in the next table :

INSERT TABLE 5 HERE

The maximum likelihood estimates for each model are given in the following table

INSERT TABLE 6 HERE

Once again the classical bivariate Poisson model fails to provide an adequate fit. Models 1,2 and 3 give nice fits. Model 1, however, has the least number of parameters and should therefore be retained according to the parsimony principle.

REFERENCES

Besson, J.L. and Partrat, C. (1992). Trend et Systèmes de Bonus-Malus. *Astin Bulletin*, 22:11–31.

Böhning, D. (1998). Zero-Inflated Poisson Models and C.A.MAN : A Tutorial Collection of Evidence. *Biometrical Journal*, 7:833–843.

Cresswell, W.L. and Frogatt, P. (1963). *The Causation of Bus Driver Accidents : An Epidemiological Study.* Oxford University Press, London.

Kocherlakota, S. and Kocherlakota, K. (1992). *Bivariate Discrete Distributions.* Marcel Dekker, New-York.

Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics*, 34:1–14.

Li, C.S, Lu, J.C., Park, J., Kim, K., Brinkley, P.A., and Peterson, J. (1999). Multivariate Zero-Inflated Poisson Models and their Applications. *Technometrics*, 41:29–38.

Lindsay, B. (1995). *Mixture Models : Theory, Geometry and Applications.* NSI-CBMS Regional Conference Series in Probability and Statistics, Vol.5, Institute of Mathematical Statistics, Hayward, California.

van den Broek, J. (1995). A Score Test for Zero Inflation in a Poisson Distribution. *Biometrics*, 51:738–743.

Jean-François Walhin
Secura Belgian Re
Rue Montoyer, 12
B-1000 Bruxelles
Belgium
e-mail : jfw@secura-re.com

APPENDIX

In this appendix we derive the maximum likelihood estimates for model 2.
The score equations will require the knowledge of $\frac{\partial \ln(f(n,m))}{\partial \theta}$ where θ is any one of the four parameters of model 2.
Therefore we study the derivative of the probability generating function with respect to the four parameters :

$$\begin{aligned}
\frac{\partial P(u,v)}{\partial \lambda_1} &= uP(u,v) - P(u,v) - pu + p \\
\frac{\partial P(u,v)}{\partial \lambda_2} &= vP(u,v) - P(u,v) - pv + p \\
\frac{\partial P(u,v)}{\partial \lambda_3} &= uvP(u,v) - P(u,v) - puv + p \\
\frac{\partial P(u,v)}{\partial p} &= \frac{1 - P(u,v)}{1-p}
\end{aligned}$$

Equating the coefficients of the terms in $u^n v^m$ gives

$$\begin{aligned}
\frac{\partial f(0,0)}{\partial \lambda_1} &= p - f(0,0) \\
\frac{\partial f(1,0)}{\partial \lambda_1} &= f(0,0) - f(1,0) - p \\
\frac{\partial f(n,m)}{\partial \lambda_1} &= f(n-1,m) - f(n,m) \quad , \quad (n,m) \neq (0,0),(1,0) \\
\frac{\partial f(0,0)}{\partial \lambda_2} &= p - f(0,0) \\
\frac{\partial f(0,1)}{\partial \lambda_2} &= f(0,0) - f(0,1) - p \\
\frac{\partial f(n,m)}{\partial \lambda_2} &= f(n-1,m) - f(n,m) \quad , \quad (n,m) \neq (0,0),(0,1) \\
\frac{\partial f(0,0)}{\partial \lambda_3} &= p - f(0,0)
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial f(1,0)}{\partial \lambda_3} &= -f(1,0) \\
\frac{\partial f(0,1)}{\partial \lambda_3} &= -f(0,1) \\
\frac{\partial f(1,1)}{\partial \lambda_3} &= f(0,0) - f(1,1) - p \\
\frac{\partial f(n,m)}{\partial \lambda_3} &= f(n-1,m) - f(n,m) \quad , \quad (n,m) \neq (0,0),(0,1),(1,0),(1,1) \\
\frac{\partial f(0,0)}{\partial p} &= \frac{1-f(0,0)}{1-p} \\
\frac{\partial f(1,0)}{\partial p} &= \frac{-f(n,m)}{1-p}
\end{aligned}
$$

The score equation for p writes :

$$
\begin{aligned}
\frac{\partial l_{n,m}(\lambda_1,\lambda_2,\lambda_3,p)}{\partial p} &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \alpha_{n,m} \frac{\frac{\partial f(n,m)}{\partial p}}{f(n,m)} = 0 \\
&= \frac{\alpha_{0,0}}{(1-p)f(0,0)} - \frac{\alpha}{1-p} = 0
\end{aligned}
$$

implying

$$
\hat{p} + (1-\hat{p})e^{-\hat{\lambda}_1-\hat{\lambda}_2-\hat{\lambda}_3} = \frac{\alpha_{0,0}}{\alpha}
$$

The score equation for λ_1, λ_2 and λ_3 write, after some calculations :

$$
\begin{aligned}
(9) \quad \frac{\partial l_{n,m}}{\partial \lambda_1} &= 0 = \frac{\alpha_{0,0}p}{f(0,0)} + \frac{\alpha_{1,1}f(0,1)}{f(1,1)} + \frac{\alpha_{1,0}}{\lambda_1} - \alpha + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{n,m}\frac{f(n-1,m)}{f(n,m)} \\
(10) \quad \frac{\partial l_{n,m}}{\partial \lambda_2} &= 0 = \frac{\alpha_{0,0}p}{f(0,0)} + \frac{\alpha_{1,1}f(1,0)}{f(1,1)} + \frac{\alpha_{0,1}}{\lambda_2} - \alpha + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{n,m}\frac{f(n,m-1)}{f(n,m)} \\
(11) \quad \frac{\partial l_{n,m}}{\partial \lambda_3} &= 0 = -\alpha + \frac{\alpha_{0,0}p}{f(0,0)} - \frac{\alpha_{1,1}p}{f(1,1)} + \frac{\alpha_{1,1}f(0,0)}{f(1,1)} + \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{n,m}\frac{f(n-1,m-1)}{f(n,m)}
\end{aligned}
$$

According to equations (7) and (8) we are able to write

$$
\begin{aligned}
(12) \qquad \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{n,m}\frac{f(n-1,m)}{f(n,m)} &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{n,m}\left[\frac{n}{\lambda_1} - \frac{\lambda_3}{\lambda_1}\frac{f(n-1,m-1)}{f(n,m)}\right] \\
(13) \qquad \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{n,m}\frac{f(n,m-1)}{f(n,m)} &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{n,m}\left[\frac{m}{\lambda_2} - \frac{\lambda_3}{\lambda_2}\frac{f(n-1,m-1)}{f(n,m)}\right]
\end{aligned}
$$

The score equation in λ_3 can be rewritten as

$$
(14) \qquad \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \alpha_{n,m}\frac{f(n-1,m-1)}{f(n,m)} = \alpha - \frac{\alpha_{0,0}p}{f(0,0)} + \frac{\alpha_{1,1}p}{f(1,1)} - \frac{\alpha_{1,1}f(0,0)}{f(1,1)}
$$

Inserting (14) and (12) in (9) we find

$$
\frac{\alpha_{0,0}p}{f(0,0)} + \frac{\alpha_{1,1}f(0,1)}{f(1,1)} - \frac{\alpha_{1,1}}{\lambda_1} - \alpha + \frac{\bar{N}}{\lambda_1} - \frac{\lambda_0}{\lambda_1}\left[\alpha - \frac{\alpha_{0,0}p}{f(0,0)} - \frac{\alpha_{1,1}f(0,0)}{f(1,1)} + \frac{\alpha_{1,1}p}{f(1,1)}\right] = 0
$$

Some algebra leads immediately to the following equation :

$$\bar{N} = (\hat{\lambda}_1 + \hat{\lambda}_3)(1 - \hat{p})$$

Similarly we have the other moment equation :

$$\bar{M} = (\hat{\lambda}_2 + \hat{\lambda}_3)(1 - \hat{p})$$

In conclusion the maximum likelihood estimates of model 2 are given as the solution of the following system :

$$\begin{aligned}
\frac{\alpha_{0,0}}{\alpha} &= \hat{p} + (1 - \hat{p})e^{-\hat{\lambda}_1 - \hat{\lambda}_2 - \hat{\lambda}_3} \\
\bar{N} &= (\hat{\lambda}_1 + \hat{\lambda}_3)(1 - \hat{p}) \\
\bar{M} &= (\hat{\lambda}_2 + \hat{\lambda}_3)(1 - \hat{p}) \\
0 &= -\alpha + \frac{\alpha_{0,0} p}{f(0,0)} - \frac{\alpha_{1,1} p}{f(1,1)} + \frac{\alpha_{1,1} f(0,0)}{f(1,1)} + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \alpha_{n,m} \frac{f(n-1, m-1)}{f(n,m)}
\end{aligned}$$

N/M		0	1	2	3	4	5	6	7
Obs	0	117	96	55	19	2	2	0	0
Model 1		117.00	72.50	51.20	24.10	8.51	2.40	0.57	0.11
Model 2		117.00	80.67	48.91	19.77	5.99	1.45	0.29	0.05
Model 3		117.05	90.65	54.36	21.73	6.52	1.56	0.31	0.05
Model 3 bis		91.15	99.16	53.93	19.56	5.32	1.16	0.21	0.03
Model Biv. Pois.		89.95	95.51	50.71	17.95	4.76	1.01	0.18	0.03
Model Biv. Ind. Pois.		71.53	92.34	59.60	25.65	8.28	2.14	0.46	0.08
Obs	1	61	69	47	27	8	5	1	0
Model 1		57.21	80.80	57.05	26.86	9.48	2.68	0.63	0.13
Model 2		59.93	84.47	58.37	26.49	8.91	2.37	0.52	0.10
Model 3		62.92	74.41	49.11	23.61	8.27	2.27	0.51	0.10
Model 3 bis		72.77	90.84	55.76	22.52	6.75	1.61	0.32	0.05
Model Biv. Pois.		69.46	94.37	61.04	25.48	7.79	1.87	0.37	0.06
Model Biv. Ind. Pois.		71.63	92.47	59.69	25.68	8.29	2.14	0.46	0.08
Obs	2	34	42	31	13	7	2	3	0
Model 1		31.88	45.02	31.79	14.97	5.28	1.49	0.35	0.07
Model 2		27.00	43.37	33.79	17.11	6.37	1.86	0.45	0.09
Model 3		30.30	37.27	30.12	16.19	6.74	2.17	0.56	0.12
Model 3 bis		29.04	40.92	30.98	15.72	5.85	1.70	0.40	0.08
Model Biv. Pois.		26.82	44.40	34.38	16.83	5.93	1.62	0.36	0.07
Model Biv. Ind. Pois.		35.86	46.30	29.89	12.86	4.15	1.07	0.23	0.04
Obs	3	7	15	16	7	3	1	0	0
Model 1		11.84	16.72	11.81	5.56	1.96	0.55	0.13	0.03
Model 2		8.11	14.62	12.71	7.14	2.92	0.93	0.24	0.05
Model 3		9.73	14.18	12.55	7.87	3.69	1.38	0.41	0.10
Model 3 bis		7.73	12.13	11.53	7.79	3.80	1.39	0.40	0.09
Model Biv. Pois.		6.90	13.48	12.24	6.96	2.81	0.87	0.22	0.04
Model Biv. Ind. Pois.		11.97	15.45	9.98	4.29	1.39	0.36	0.08	0.01
Obs	4	3	3	1	1	2	1	1	1
Model 1		3.30	4.66	3.29	1.55	0.55	0.15	0.04	0.01
Model 2		1.83	3.65	3.51	2.17	0.97	0.34	0.10	0.02
Model 3		2.34	3.35	4.13	2.90	1.56	0.65	0.22	0.06
Model 3 bis		1.54	2.67	3.15	2.79	1.77	0.81	0.28	0.08
Model Biv. Pois.		1.33	3.00	3.14	2.04	0.94	0.33	0.09	0.02
Model Biv. Ind. Pois.		3.00	3.87	2.50	1.07	0.35	0.09	0.02	0.00
Obs	5	2	1	0	0	0	0	0	0
Model 1		0.74	1.04	0.73	0.35	0.13	0.03	0.01	0.00
Model 2		0.33	0.72	0.76	0.52	0.25	0.10	0.03	0.01
Model 3		0.05	0.86	1.05	0.86	0.52	0.25	0.09	0.03
Model 3 bis		0.25	0.47	0.70	0.75	0.60	0.34	0.15	0.05
Model Biv. Pois.		0.21	0.52	0.62	0.46	0.24	0.09	0.03	0.01
Model Biv. Ind. Pois.		0.60	0.77	0.50	0.22	0.07	0.02	0.00	0.00
Obs	6	0	0	0	0	1	0	0	0
Model 1		0.14	0.19	0.14	0.06	0.02	0.01	0.00	0.00
Model 2		0.05	0.12	0.14	0.10	0.05	0.02	0.01	0.00
Model 3		0.07	0.16	0.22	0.20	0.14	0.07	0.03	0.01
Model 3 bis		0.03	0.07	0.12	0.16	0.15	0.11	0.05	0.02
Model Biv. Pois.		0.03	0.08	0.10	0.08	0.05	0.02	0.01	0.00
Model Biv. Ind. Pois.		0.10	0.13	0.08	0.04	0.01	0.00	0.00	0.00
Obs	7	0	0	0	1	0	0	0	0
Model 1		0.02	0.03	0.02	0.01	0.00	0.00	0.00	0.00
Model 2		0.01	0.02	0.02	0.02	0.01	0.00	0.00	0.00
Model 3		0.01	0.02	0.04	0.04	0.03	0.02	0.01	0.00
Model 3 bis		0.00	0.01	0.02	0.03	0.03	0.03	0.02	0.01
Model Biv. Pois.		0.00	0.01	0.01	0.01	0.01	0.00	0.00	0.00
Model Biv. Ind. Pois.		0.01	0.02	0.01	0.00	0.00	0.00	0.00	0.00

Table 1: Data set 1 : Observed and fitted distribution

	l	χ^2	df	$p-value$	AIC	BIC
Model 1	-2012.11	36.74	21	0.017	-2015.11	-2021.95
Model 2	-2004.47	24.33	20	0.228	-2008.47	-2017.59
Li et al model	-1998.47	16.24	18	0.576	-2004.47	-2018.19
Model 3	-1998.51	14.89	18	0.669	-2004.51	-2018.19
Model 3 bis	-2007.89	35.96	20	0.016	-2017.89	-2021.01
Model Biv. Pois.	-2012.20	40.82	21	0.006	-2015.20	-2022.04
Model Biv. Ind. Pois.	-2031.60	84.66	22	0.000	-2033.64	-2038.20

Table 2: Data set 1 : Characteristics of the fit

Model 1	$p = 0.0927$	$\lambda_1 = 1.1037$	$\lambda_2 = 1.4229$	
Model 2	$p = 0.0712$	$\lambda_1 = 0.9008$	$\lambda_2 = 1.2126$	$\lambda_3 = 0.1774$
Model 3	$p_0 = 0.3728$	$p_1 = 0.2321$	$p_2 = 0.1418$	
	$\lambda_0 = 0.4173$	$\lambda_1 = 0.9632$	$\lambda_2 = 1.1993$	
Model 3 bis	$p_0 = 0.6746$	$\lambda_0 = 0.6244$	$\lambda_1 = 0.7982$	$\lambda_2 = 1.0878$
Model Biv. Pois.	$\lambda_1 = 0.7722$	$\lambda_2 = 1.0618$	$\lambda_3 = 0.2291$	
Model Biv. Ind. Pois.	$\lambda_1 = 1.0014$	$\lambda_2 = 1.2909$		

Table 3: Data set 1 : Maximum likelihood estimates of the parameters

N/M		0	1	2	3	4
Obs	0	34	8	3	1	0
Model 1		34.00	7.83	3.33	0.94	0.20
Model 2		34.00	7.96	2.98	0.75	0.14
Model 3		34.32	9.35	3.09	0.68	0.11
Model 3 bis		24.60	12.50	3.18	0.54	0.07
Model Biv. Pois.		25.12	11.09	2.45	0.36	0.04
Model Biv. Ind. Pois.		20.60	13.18	4.22	0.90	0.14
Obs	1	12	13	6	1	0
Model 1		11.50	9.79	4.16	1.18	0.25
Model 2		12.11	9.99	4.09	1.11	0.22
Model 3		10.50	11.75	3.37	1.01	0.21
Model 3 bis		19.88	11.37	3.21	0.54	0.07
Model Biv. Pois.		18.63	13.21	4.02	0.75	0.10
Model Biv. Ind. Pois.		19.36	12.39	3.97	0.85	0.14
Obs	2	4	3	1	0	0
Model 1		7.19	6.12	2.60	0.74	0.16
Model 2		6.91	6.22	2.76	0.81	0.17
Model 3		6.74	4.56	2.50	0.69	0.18
Model 3 bis		8.04	5.10	2.20	0.64	0.13
Model Biv. Pois.		6.91	6.75	2.80	0.68	0.11
Model Biv. Ind. Pois.		9.10	5.82	1.86	0.40	0.06
Obs	3	5	3	2	1	0
Model 1		3.00	2.55	1.08	0.31	0.07
Model 2		2.63	2.56	1.23	0.39	0.09
Model 3		2.88	2.55	1.09	0.41	0.10
Model 3 bis		2.17	1.51	1.01	0.59	0.20
Model Biv. Pois.		1.71	2.12	1.14	0.35	0.07
Model Biv. Ind. Pois.		2.85	1.82	0.58	0.12	0.02
Obs	4	2	0	0	0	0
Model 1		0.94	0.80	0.34	0.10	0.02
Model 2		0.75	0.79	0.40	0.14	0.03
Model 3		0.93	1.00	0.51	0.18	0.05
Model 3 bis		0.44	0.33	0.32	0.31	0.18
Model Biv. Pois.		0.32	0.48	0.32	0.12	0.03
Model Biv. Ind. Pois.		0.67	0.43	0.14	0.03	0.00
Obs	5	0	0	0	0	1
Model 1		0.23	0.20	0.08	0.02	0.00
Model 2		0.17	0.19	0.11	0.04	0.01
Model 3		0.24	0.30	0.18	0.07	0.02
Model 3 bis		0.07	0.06	0.08	0.11	0.09
Model Biv. Pois.		0.05	0.08	0.07	0.03	0.01
Model Biv. Ind. Pois.		0.13	0.08	0.03	0.01	0.00

Table 4: Data set 2 : Observed and fitted distribution

	l	χ^2	df	$p-value$	AIC	BIC
Model 1	-236.78	5.65	6	0.463	-239.77	-243.68
Model 2	-236.47	5.95	5	0.311	-240.47	-245.68
Li et al model	-236.47	5.95	3	0.114	-242.47	-250.28
Model 3	-235.39	5.02	3	0.171	-241.39	-249.20
Model 3 bis	-240.03	19.03	5	0.002	-244.02	-249.23
Model Biv. Pois.	-241.62	15.59	6	0.016	-244.61	-248.52
Model Biv. Ind. Pois.	-245.65	22.55	7	0.002	-247.64	-250.25

Table 5: Data set 2 : Characteristics of the fit

Model 1	$p = 0.2480$	$\lambda_1 = 1.2500$	$\lambda_2 = 0.8511$	
Model 2	$p = 0.2339$	$\lambda_1 = 1.1415$	$\lambda_2 = 0.7499$	$\lambda_3 = 0.0854$
Model 3	$p_0 = 0$	$p_1 = 0.4694$	$p_2 = 0.4245$	
	$\lambda_0 = 0.2590$	$\lambda_1 = 1.2834$	$\lambda_2 = 0.6620$	
Model 3 bis	$p_0 = 0.8721$	$\lambda_0 = 1.0300$	$\lambda_1 = 0.8083$	$\lambda_2 = 0.5083$
Model Biv. Pois.	$\lambda_1 = 0.7414$	$\lambda_2 = 0.4414$	$\lambda_3 = 0.1985$	
Model Biv. Ind. Pois.	$\lambda_1 = 0.94$	$\lambda_2 = 0.64$		

Table 6: Data set 2 : Maximum likelihood estimates of the parameters