Construction of a New Multivariate Poisson Distribution through the Conditional Probability with its Variance Covariance Matrix: Application to the Simulation of Poisson Regressions

Réolie F. Mizele Kitoti, R. Bidounga and D. Mizere

Laboratoire de Statistique et d’Analyse des Données
Marien Ngouabi University, Brazzaville, Congo

This article is distributed under the Creative Commons by-nc-nd Attribution License. Copyright © 2019 Hikari Ltd.

Abstract

In this study, we construct a new multivariate Poisson distribution. This new law is the product of the conditional univariate Poisson distribution. So we determine its variance covariance matrix. Through the parameters of Poisson distribution, we construct k-models of Poisson regression. We carry out their estimation and hypothesis tests related to it. As the bivariate data were studied a lot, we want to realize an application of the trivariate case.

Mathematics Subject Classification: 62E10, 62G20, 62E15

Keywords: Multivariate law of Poisson, conditional probabilities, Simulations
1. Introduction

The law we are building is the limit of existing multivariate Poisson law Kokonendji and Puig ([5]). It is considered as a generalization of bivariate Poisson law of Berkhout and Plug ([1]) hence, this new law is the product of the conditional Poisson distribution.

We will provide parameters estimation and simulation studies using associated Poisson regressions will be made in the trivariate case in Section 4.

Let us consider $k + 1$ independent variables $X_1, X_2, ..., X_k$ and $Z$ which follows the univariate Poisson distribution of respective parameters $\mu_1, \mu_2, ..., \mu_k$ and $\lambda$.

One builds the new dependent variables $Y_i = X_i + Z$, $i \in \{1, \cdots, k\}$, $X = (X_1, ..., X_k)^T$ and $Y = (Y_1, ..., Y_k)^T$ of realization $y = (y_1, \cdots, y_k)^T$. The random vector $Y$ has the following probability mass function (pmf).

$$
\mathbb{P}(Y = y) = \sum_{\ell \geq 0} \mathbb{P}(X = y - Z|Z = \ell) \mathbb{P}(Z = \ell) \\
= \sum_{\ell \geq 0} \mathbb{P}(X = y - \ell) \mathbb{P}(Z = \ell) \\
= \sum_{0 \leq \min(y_1, \cdots, y_k)} \mathbb{P}(X = y - \ell) \mathbb{P}(Z = \ell).
$$

since $(X = y - \ell)$ is empty whenever one of the $y_i$, $i \in \{1, \cdots, k\}$, is less that $\ell$. We write $\mathbb{P}(Y = y|(\mu, \lambda))$ with $\mu = (\mu_1, \mu_2, ..., \mu_k)^T$, we have the following result (Kokonendji and Puig ([5])).

$$
\mathbb{P}(Y = y) = \mathbb{P}(Y = y|\mu, \lambda) = \exp \left( -\lambda - \sum_{\ell = 1}^{k} \mu_{\ell} \right) \prod_{\ell = 1}^{k} \frac{\mu_{\ell} y_{\ell}}{y_{\ell}!} \\
\times \sum_{\ell = 0}^{\min(y_1, \cdots, y_k)} \left\{ \prod_{i = 1}^{k} \left( \frac{y_i}{\ell} \right) (\ell!)^{k-1} \left( \frac{\lambda}{\mu_1 \cdots \mu_k} \right)^{\ell} \right\}.
$$

For this law, despite the dependence of variables $Y_i$, $i = 1, ..., k$, no regression equation is formally established. We will construct our new model as follows.

Let us considering a sequence $(Y^{(n)})_{n \geq 1}$ of $k$-random vectors such that $Y^{(n)} = \ldots$
The construction of a new multivariate Poisson distribution $X + Z_n$, where $Z_n$ follows a Poisson law with mean $\lambda_n \geq 0$ with $\lambda_n \to 0$ when $n \to +\infty$. We have

$$\lim_{n \to +\infty} P(Y^{(n)} = y|\mu, \lambda_n) = \prod_{i=1}^{k} \left( \frac{\mu_i^{y_i}}{y_i!} e^{-\mu_i} \right).$$ (1.2)

This latter asymptotic law give birth to the idea of building the new multivariate Poisson distribution we are studying in the next section.

2. A new multivariate Poisson distribution

2.1. The bivariate Berkhout model.

Let us reconsider the Berkhout model again. From a probabilistic point of view, we begin by an unconditional Poisson law $Y_1 \sim \mathcal{P}(\mu_1)$. Given a real-valued positive function $\mu_2(r)$ of $r \in \mathbb{N}$, we may use the Kolmogorov existence theorem (e.g., Lo ([6])) to construct a pair of random variables $(Y_1, Y_2)$ of mpf

$$\forall (y_1, y_2) \in \mathbb{N}^2, \; \mathbb{P}(Y_1 = y_1, Y_2 = y_2) = \frac{\mu_1^{y_1} e^{-\mu_1} \mu_2(y_1)^{y_2} e^{-\mu_2(y_1)}}{y_1! y_2!}. \quad (2.1)$$

In such a case,

$$\forall (y_1, y_2) \in \mathbb{N}^2, \; \mathbb{P}(Y_2 = y_2/Y_1 = y_1) = \frac{\mu_2(y_1)^{y_2} e^{-\mu_2(y_1)}}{y_2!}.$$ 

In other terms $Y_2$ is constructed such that the probability law of $Y_2$ given $Y_1 = y_1$ follows a Poisson law $\mathcal{P}(\mu_2(y_1))$, what we can write in term of conditional expectation as

$$\forall y_2 \in \mathbb{N}, \; \mathbb{P}(Y_2 = y_2/Y_1) = \frac{\mu_2(Y_1)^{y_2} e^{-\mu_2(Y_1)}}{y_2!}$$

and we find again (2.1) by using the conditional projection principle through $\mathbb{E}(\mathbb{P}(Y_2 = y_2/Y_1))$.

The log-linear associated models:

$$\ln \mu_1 = x^T \beta_1$$
$$\ln \mu_2 = x^T \beta_2 + \eta y_1$$

Where $x^T = (x_1, x_2, ..., x_p)$ be a vector of explanatory variables or factors.
2.2. The new multivariate model.

We have highlighted this model which is a generalization of the bivariate Poisson distribution of Berkhout and Plug ([1]). From there, it is possible to go to a $k$-multivariate model, $k \geq 3$, by considering real-valued positive functions $\mu_i$, $2 \leq i \leq k$ such that each $m_i$ is a function $i - 1$ arguments in $\mathbb{N}$ and to get

$$\forall i \in \{3, \ldots, k\}, \text{for } y_i \in \mathbb{N}, \mathbb{P}(Y_i = y_i/(Y_1, \ldots, Y_{i-1})) = \frac{[\mu_i(Y_1, \ldots, Y_{i-1})]^{y_i} e^{-\mu_i(Y_1, \ldots, Y_{i-1})}}{y_i!}.\quad (2.2)$$

which in turn leads to : for $(y_1, \ldots, y_k)$,

$$\forall i \in \{3, \ldots, k\}, \forall y_i \in \mathbb{N}, \mathbb{P}(Y_i = y_i/(Y_1 = y_1, \ldots, Y_{i-1} = y_{i-1})) = \frac{[\mu_i(y_1, \ldots, y_{i-1})]^{y_i} e^{-\mu_i(y_1, \ldots, y_{i-1})}}{y_i!}. \quad (2.3)$$

Finally, by successive conditioning, we get the general model

$$\forall (y_1, \ldots, y_k) \in \mathbb{N}^k, \mathbb{P}(Y_1 = y_1, \ldots, Y_k = y_k) = \prod_{i=1}^k \frac{[\mu_i(y_1, \ldots, y_{i-1})]^{y_i} e^{-\mu_i(y_1, \ldots, y_{i-1})}}{y_i!},$$

where we use the convention that for $i = 1$, $\mu_1(y_0) = \mu_1$.

In the course of the construction, we have to choose the regression function $\mu_i$, $2 \leq i \leq k$.

In this paper, we use a log-linear model in the form

$$\forall i \in \{2, i\}, ln(\mu_i) = x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_i y_\ell, \quad (2.5)$$

$\mu_1$ being unconditional. We conclude our construction as follows.

**Theorem 1.** Let $k \geq 1$. Given $\mu_1 > 0$ and $k-1$ real-valued positive functions $\mu_i$, $2 \leq i \leq k$ such that each $m_i$ is a function $i - 1$ arguments in $\mathbb{N}$, there exists a $k$-dimensional discrete random variable $Y = (Y_1, \ldots, Y_k)$ with values in $\mathbb{N}^k$ such that
\begin{align*}
\forall i \in \{2, \cdots, k\}, \forall y_i \in \mathbb{N}, P(Y_i = y_i/(Y_1, \cdots, Y_{i-1}) & = \left[\mu_i(Y_1, \cdots, Y_{i-1})\right]^{y_i} e^{-\mu_i(Y_1, \cdots, Y_{i-1})} y_i! \tag{2.6} \\
\forall \forall y_i \in \mathbb{N}, P(Y_i = y_i/(Y_1, \cdots, Y_{i-1}) & = \left[\mu_i(Y_1, \cdots, Y_{i-1})\right]^{y_i} e^{-\mu_i(Y_1, \cdots, Y_{i-1})} y_i! \tag{2.7}
\end{align*}

and whose mpf is
\[
\forall (y_1, \cdots, y_k) \in \mathbb{N}^k, P(Y_1 = y_1, \cdots, Y_k = y_k) = \prod_{i=1}^{k} \left[\mu_i(y_1, \cdots, y_{i-1})\right]^{y_i} e^{-\mu_i(y_1, \cdots, y_{i-1})} \frac{y_i!}{y_i!}, \tag{2.8}
\]

where we use the convention that for $i = 1$, $\mu_1(y_0) = \mu_1$.

We define

**Definition 1.** By $k$-dimensional Berkhout model, we mean a random vector whose mpf is given by 2.8 and the following regression holds
\[
\forall i \in \{2, i\}, ln(\mu_i(Y_1, \cdots, Y_{i-1})) = x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_i \ell y_\ell, \tag{2.9}
\]

where $x$ is some covariate, the $\beta_i$’s and $\eta_\ell$’s are the parameters of the model.

As we said above, and taking into account theorem 1 and definition 1, this new law is a generalization of the bivariate Poisson law of Berkhout and Plug ([1]) , in that it highlights the regressions equations.

2.3. Variance-covariance matrix. Let us recall that for the consecutive variable $y_i$, $y_{i+1}$ and $y_{i+2}$ of the vector $y = (y_1, y_2, \cdots, y_k)$, we have:
\[
ln(\mu_i) = x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_i \ell y_\ell, \quad \beta_i, \quad \eta_i \ell \in \mathbb{R} \quad i = (1, 2, \cdots, k)
\]
\[
ln(\mu_{i+1}) = x^T \beta_{i+1} + \sum_{\ell=1}^{i} \eta_{ii} y_\ell + \eta_{ii} y_i. \tag{2.10}
\]
\[
ln(\mu_{i+2}) = x^T \beta_{i+2} + \sum_{\ell=1}^{i+1} \eta_{ii} y_\ell + \eta_{ii} y_i + \eta_{(i+1)i} y_{i+1}. \tag{2.11}
\]
By setting
\[ z = (x^T, y_1, ..., y_i) \], \( \theta_1 = (\beta_i, \eta_{i_1}, ..., \eta_{(i-1)i})^T \),
\[ \theta_2 = (\beta_{i+1}, \eta_{i_1}, \eta_{2i}, ..., \eta_{(i-1)i})^T \]
and
\[ h = (z^T, y_{i+1}) \] and \( \theta_3^T = (\theta_2^T, \eta_{(i+1)i})^T \),
we have the following results.

**Proposition 1.** We have the following facts:

1. \( \ln(\mu_i) = z^T \theta_1 \) and \( \ln(\mu_{i+1}) = z^T \theta_2 + \eta_{ii}y_i \)
2. \( \mathbb{E}(Y_i) = \text{Var}(Y_i) = \mu_i \)
3. \( \mathbb{E}(Y_{i+1}) = e^{z^T \theta_2 + \mu_i (\exp(\eta_i) - 1)} \)
4. \( \text{Var}(Y_{i+1}) = \mathbb{E}(Y_{i+1}) + (\mathbb{E}(Y_{i+1}))^2 (e^{\mu_i (\exp(\eta_i) - 1)})^2 - 1 \)
5. \( \text{Cov}(Y_i, Y_{i+1}) = \mu_i \mathbb{E}(Y_{i+1})(e^{\eta_{ii}} - 1), \quad i = 1, 2, ..., k-2. \)

**Proof.** The proof is trivial. That is enough, in Elion and al.(2016) concerning the calculation of the factorial moments of order \( (r,s) \) to put \( Z_1 = Y_i \), \( Z_2 = Y_{i+1} \) and \( \eta = \eta_{ii}. \) □

**Corollary 1.** We also have :

1. \( \ln(\mu_{i+2}) = z^T \theta_2 + \eta_{ii}y_i + \eta_{(i+1)i}y_{i+1} = h^T \theta_3 + \eta_{ii}y_i \)
2. \( \mathbb{E}(Y_{i+2}) = e^{h^T \theta_3 + \mu_i (\exp(\eta_i) - 1)} \)
3. \( \text{Var}(Y_{i+2}) = \mathbb{E}(Y_{i+2}) + (\mathbb{E}(Y_{i+2}))^2 (e^{\mu_i (\exp(\eta_i) - 1)})^2 - 1 \)
4. \( \text{Cov}(Y_i, Y_{i+2}) = \mu_i \mathbb{E}(Y_{i+2})(e^{\eta_{ii}} - 1), \quad i = 1, 2, ..., k-2. \)

The results of Proposition 1 and its corollary allow us to build all the matrix of variance covariance \( Y = (Y_1, Y_2, ..., Y_k). \)

The computed covariances can be negative, null or positive following the values of the parameter \( \eta_{ii}. \)

The variables \( Y_i \) are independent if the coefficients \( \eta_{ii} = 0 \) are all zero.

3. **Log likelihood estimation of the parameters and statistical hypothesis tests**

3.1. **Log likelihood estimation of the parameters \( \beta_i, \eta_{ii}. \)**

**Proposition 2.** We have the following results:

1. \( \frac{\partial \ln f}{\partial \beta_i} = x^T (y_i - \mu_i) \) with \( \frac{\partial^2 \ln f}{\partial \beta_i \partial \beta_i} = -\|x\|^2 \mu_i \)
2. \( \frac{\partial \ln f}{\partial \eta_{ii}} = y_i (y_i - \mu_i) \) with \( \frac{\partial^2 \ln f}{\partial \eta_{ii} \partial \eta_{ii}} = -y_i^2 \mu_i. \)
Corollary 2. The maximum likelihood estimator $\hat{\beta}_i(\hat{\eta}_{il})$ of the parameter $\beta_i(\eta_{il})$ is the coefficient of the generalized model (2.9)

Proof of proposition 2: Let consider the probability density:

$$
P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_{k-1} = y_{k-1}, Y_k = y_k) = \prod_{i=1}^{k} \left( \frac{\mu_i^{y_i}}{y_i!} e^{-\mu_i} \right)
= f(y_1, y_2, \ldots, y_{k-1}, y_k)
$$

under the condition

$$
\ln(\mu_i) = x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_{il} y_\ell, \quad \beta_i, \quad \eta_{il} \in \mathbb{R} \quad i = (1, 2, \ldots, k),
$$

For the variable $y_i$, the log-likelihood of $f$ is equal to:

$$
\ln(f) = \sum_{i=1}^{k} [y_i ln \mu_i - \mu_i - ln(y_i)!]
= \sum_{i=1}^{k} [y_i x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_{il} y_\ell - e^{x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_{il} y_\ell}]
$$

(1) we have:

$$
\frac{\partial \ln f}{\partial \beta_i} = y_i x^T - x^T e^{x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_{il} y_\ell}
= x^T (y_i - e^{x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_{il} y_\ell})
$$

However $\mu_i = e^{x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_{il} y_\ell}$, we have

$$
\frac{\partial \ln f}{\partial \beta_i} = x^T (y_i - \mu_i).
$$

Otherwise, we have

$$
\frac{\partial^2 \ln f}{\partial \beta_i \partial \beta_i^T} = -x^T x e^{x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_{il} y_\ell}
= -\|x\|^2 \mu_i.
$$

(2) We have:

$$
\frac{\partial \ln f}{\partial \eta_{il}} = y_i y_{il} - y_{il} e^{x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_{il} y_\ell}
= y_i y_{il} - y_{il} \mu_i
= y_{il} (y_i - \mu_i).
$$
Hence,
\[
\frac{\partial^2 \ln f}{\partial \eta_i \partial \eta_l} = \frac{\partial}{\partial \eta_l} \left( x^T \beta_i - e^{x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_i y_{\ell} + \eta_i y_i} \right)
= -y_{\ell} e^{x^T \beta_i + \sum_{\ell=1}^{i-1} \eta_i y_{\ell}}
= -y_{\ell}^2 \mu_i.
\]

3.2. **Statistical hypothesis tests.** Let \( \beta_i^T = (\beta_{i1}, \beta_{i2}, ..., \beta_{ij}, ..., \beta_{ip}) \) as \( x^T = (x_1, x_2, ..., x_j, ..., x_p) \) we have,
\[
x^T \beta_i = \sum_{j=1}^{p} x_j \beta_{ij}.
\]

For the generalized linear model (2.2), we finish the regression on the statistical hypotheses

- \( H_{01} : \beta_{ij} = 0 \) (the factor \( x_j \) has no effect on the response variable \( Y_i \))
- \( H_{11} : \beta_{ij} \neq 0 \) (the factor \( x_j \) has an effect on the response variable \( Y_i \))
- \( H_{02} : \eta_i = 0 \) (the variable \( Y_l(l = 1, ..., i-1) \) and the response variable \( Y_i \) are independant)
- \( H_{12} : \eta_i \neq 0 \) (the variable \( Y_l(l = 1, ..., i-1) \) and the response variable \( Y_i \) are dependant).

4. **Simulation of the Poisson regressions in the trivariate case**

When a statistical series \( y_1, y_2, ..., y_k \) of a variable \( Y \) is available, it is not always obvious by the use of the chi-square adequacy test that the data are Poisson-free.

Since we are talking about using Poisson data in this work, it seemed to us opportune to use simulated Poisson data.

We have the following regressions to treat :

1. \( \ln(\mu_1) = x^T \beta_1 \)
2. \( \ln(\mu_2) = x^T \beta_2 + \eta_{21} y_1 \)
3. \( \ln(\mu_3) = x^T \beta_3 + \eta_{31} y_1 + \eta_{32} y_2. \)

Tables 1, 2, 3 below represent simulated data of size 82 and average 2 of the Poisson variables \( y_1, y_2 \) and \( y_3 \). We got the values of the \( x \) factor, by simulated the normal standard law with the same size 82.
Table 1. Table of distribution of \( y_1 \)

<table>
<thead>
<tr>
<th>( y_1 )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eff. observed</td>
<td>15</td>
<td>21</td>
<td>15</td>
<td>19</td>
<td>7</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. Table of distribution of \( y_2 \)

<table>
<thead>
<tr>
<th>( y_2 )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eff. observed</td>
<td>14</td>
<td>24</td>
<td>27</td>
<td>11</td>
<td>4</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 3. Table of distribution of \( y_3 \)

<table>
<thead>
<tr>
<th>( y_3 )</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eff. observed</td>
<td>13</td>
<td>15</td>
<td>22</td>
<td>19</td>
<td>8</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4. Coefficients of regressions 1

| Variable | \( \beta \) | \( S_\beta \) | \( t_i \) | \( P(>|t|) \) |
|----------|-------------|-------------|--------|------------|
| Intercept | 0.73391 | 0.07863 | 9.333 | < 2e-16 *** |
| \( x \) | -0.01802 | 0.06756 | -0.267 | 0.79 |

AIC = 290.71

Since the data have been simulated, the interpretation of the statistical treatments has no interest. The interest here resides in checking the conformity with the chosen model. Here the coefficient \( \eta_{21} \), estimated by \( \hat{\eta}_{21} = 0.04879 \), is significantly null since \( p-value = 0.35595 \) is high. Hence, the variables \( Y_1 \) and \( Y_2 \) are effectively independent. By the choice of the coefficients \( \eta_{ij} \), a mass pdf of multivariate Poisson laws can be proposed.

5. Conclusion

In this study, we have developed a new multivariate Poisson distribution based on his conditional laws. This new law is the k-variate extension of the bivariate Poisson distribution according to Berkhout and Plug(2004). It is also the asymptotic law of Holgate (1964). Also we have calculate the variance covariance matrix associated to this new distribution.
Table 5. Coefficients of regressions 2

| Variable | $\hat{\beta}$ | $S_{\hat{\beta}}$ | $t_i$ | $P(> |t|)$ |
|----------|----------------|-------------------|-------|-------------|
| Intercept | 0.42041        | 0.14328           | 2.934 | 0.00335 ** |
| x        | -0.03260       | 0.07587           | -0.430| 0.66740     |
| $y_1$    | 0.04879        | 0.05285           | 0.923 | 0.35595     |
|          |                |                   |       | AIC= 259.74 |

Table 6. Coefficients of the regression 3

| Variable | $\hat{\beta}$ | $S_{\hat{\beta}}$ | $t_i$ | $P(> |t|)$ |
|----------|----------------|-------------------|-------|-------------|
| Intercept | 0.69613        | 0.15881           | 4.383 | 1.17e-05 ***|
| x        | -0.02996       | 0.06594           | -0.454| 0.650       |
| $y_1$    | -0.03645       | 0.04848           | -0.752| 0.452       |
| $y_2$    | 0.08525        | 0.06217           | 1.371 | 0.170       |
|          |                |                   |       | AIC= 297.04 |

x-factor TABLE.

x-factor TABLE

<table>
<thead>
<tr>
<th>x-factor TABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1] -0.2234274287 -1.6307471735 -2.0365045043 1.7302872565 0.0424844894</td>
</tr>
<tr>
<td>[6] -1.7692908059 1.0288525742 0.0007483475 1.5427671012 0.3284908516</td>
</tr>
<tr>
<td>[11] -1.8513436668 1.4115969399 2.0357701461 0.422646764 -1.3270508650</td>
</tr>
<tr>
<td>[16] 0.4250898111 1.9588314535 1.2017149653 -0.4754811766 2.3086570508</td>
</tr>
<tr>
<td>[21] 1.5614411801 -0.4724178408 -0.9367822196 1.6709703732 0.2493518755</td>
</tr>
<tr>
<td>[26] 1.3601494466 1.2887133396 0.2922499219 0.5365834280 -0.3136921186</td>
</tr>
<tr>
<td>[31] 1.0294728866 -0.0902694769 0.4830329282 -2.0420793999 0.6096617749</td>
</tr>
<tr>
<td>[36] 0.3256575397 0.2340216897 -0.3085852627 0.6435525271 -0.8841832703</td>
</tr>
</tbody>
</table>
References


Received: October 14, 2019; Published: November 9, 2019