J COMP SCI
APPL MATH
Journal of Computer Science and Applied Mathematics
Vol. 6, no.2, (2024), 33–43
ISSN: 1857-9582 https://doi.org/10.37418/jcsam.6.2.1

THE GENERALIZED BIVARIATE POISSON DISTRIBUTION ACCORDING TO BERKHOUT AND PLUG.

R.F. Mizelé Kitoti[1], P.C. Batsindila Nganga, E. Nguessolta, R. Bidounga, and D. Mizere

ABSTRACT. In this article, we will construct a new bivariate Poisson distribution through the bivariate law of probabilities of causes highlighted by Bidounga et al. in [2]. This law generalise the bivariate Poisson distribution according to Berkhout and Plug [1].And finally we simulated the data.

1. INTRODUCTION

Several bivariate Poisson laws have been constructed, notably that of Holgate [3], Lakshminarayana [4] and Berkhout Plug [1].

The new bivariate Poisson law that we will construct in this paper through the bivariate law of the probabilities of causes, highlighted by Bidounga et al. [2]. It generalizes the bivariate Poisson law according to Berkhout and Plug [1].

In section 1, we will review the bivariate Poisson distribution according to Berkhout and Plug [1] and the bivariate distribution using the probabilities of causes (Bidounga et al. [2]).

In section 2, we will define the new law and in section 3, we will present a simulation of this model.

[1]*corresponding author*
Key words and phrases. Bivariate law of Poisson, Conditional probability, Distribution, overdispersed variables, Simulations.
Submitted: 09.08.2024; *Accepted:* 24.08.2024; *Published:* 14.10.2024.

2. A REVIEW OF DISTRIBUTION

2.1. **Bivariate Poisson distribution according to Berkhout and Plug [1].** Let $Y_i \quad (i = 1, n)$ a random variable which follows the unvariate Poisson distribution with parameters $\lambda_i \quad (i = 1, 2)$. The vector (Y_1, Y_2) follows the bivariate Poisson distribution according to Berkhout and Plug [1] if its mass function denoted f_{BP} is equal to.

$$f_{BP}(y_1, y_2, \lambda_1, \lambda_2) = \left(\frac{\lambda_1^{y_1} e^{-\lambda_1}}{y_1!}\right)\left(\frac{\lambda_2^{y_2} e^{-\lambda_2}}{y_2!}\right), y_i \in \mathbb{N}, \lambda_i \in \mathbb{R}_+^*(i = 1, 2). \tag{2.1}$$

under the conditions,

$$\ln \lambda_1 = x'\rho_1 \tag{2.2}$$

and

$$\ln \lambda_2 = x'\rho_2 + \eta y_1. \tag{2.3}$$

The bivariate Poisson distribution according to Berkhout and Plug [1] has the following characteristics:

$$E(Y_1) = Var(Y_1) = \lambda_1, \tag{2.4}$$

$$E(Y_2) = e^{x'\rho_2 + c_2 + \lambda_1(e^{\eta} - 1)}, \tag{2.5}$$

where c_2 is the intercept of the model 2.3, and

$$V(Y_2) = E(Y_2) + [E(Y_2)]^2[e^{\lambda_1(e^{\eta}-1)} - 1], \tag{2.6}$$

$$Cov(Y_1, Y_2) = \lambda_1 E(Y_2)(e^{\eta} - 1). \tag{2.7}$$

The expression 2.6 shows that the variable Y_2 is overdispersed. The expression 2.7 confirms the that the variables Y_1 and Y_2 are independent if and only if $\eta = 0$. And the covariance is negative, zero and positive depending on whether η is negative, zero or positive.

2.2. **Bivariate distribution using the probabilities of causes (Bidounga and al. [2]).** Let consider the positive integers random variables Y_1, Y_2 and T. Let $T_1, T_2, \ldots, T_n$ the sample of size n of the variable T.

Definition 2.1. *The bivariate distribution using the probabilities of causes has a mass function equal to*

$$P(Y_1 = y_1, Y_2 = y_2) = [\sum_{i}^{n} p_i P(Y_1 = y_1/T_i = t_i)] \times [P(Y_2 = y_2/Y_1 = y_1)] \tag{2.8}$$

with $\sum_i^n p_i = 1$.

3. The generalised bivariate Poisson distribution according to Berkhout and Plug [1]

Assume that the variables Y_1, Y_2 and T follow univariate Poisson distributions of parameters λ_1, λ_2 and λ. We have the following conditional probability

$$\mathbb{P}(Y_1 = y_1/T_i = t_i) = \frac{\lambda_1^{y_1}}{y_1!} e^{-\lambda_1}, \quad i = 1, 2 \ldots, n, \tag{3.1}$$

as we have the model (Mizelé and al.[5]),

$$\lambda_1 = \lambda_1(t_i), \quad i = 1, 2, \ldots, n,$$

and consequently

$$\lambda_1 = \lambda_1(t_1, t_2, \ldots, t_3). \tag{3.2}$$

In the same vein, we have

$$\mathbb{P}(Y_2 = y_2/Y_1 = y_1) = \frac{\lambda_2^{y_1}}{y_2!} e^{-\lambda_2}, \tag{3.3}$$

with the model

$$\lambda_2 = \lambda_2(y_1). \tag{3.4}$$

We will assume that the model 3.2 and 3.4 are Log-linear defined as follows

$$\ln(\lambda_1) = x'\rho_1 + \sum_{i=1}^{n} \alpha_i t_i, \tag{3.5}$$

and

$$\ln(\lambda_2) = x'\rho_2 + \eta y_1, \tag{3.6}$$

where $\rho_1,\ \rho_2,\ \alpha_1,\ \alpha_2, \ldots,\ \alpha_n$ and η are the parameters and x a deterministe variable or factor. The generalized linear model 3.5 has the response variable Y_1 and the model 3.6 has the response variable Y_2. We have the following result.

Proposition 3.1. *When in the expression 3.1 we replace λ_1 by the expression 3.2 then we have*

$$\mathbb{P}(Y_1 = y_1/T_i = t_i) = \mathbb{P}(Y_1 = y_1/T_1 = t_1, \ldots, T_n = t_n), \forall i. \tag{3.7}$$

Corollary 3.1. *The bivariate distribution of the probabilities of the causes is then equal to*

$$\begin{aligned} \mathbb{P}(Y_1 = y_1, Y_2 = y_2) &= \mathbb{P}(Y_1 = y_1/T_1 = t_1, T_2 = t_2, \ldots, T_n = t_n) \\ &\times \mathbb{P}(Y_2 = y_2/Y_1 = y_1). \end{aligned} \tag{3.8}$$

Definition 3.1. *Let Y_1, Y_2 and T the random variables which follows the univariate Poisson distribution of parameters λ_1, λ_2 and λ. Let $T_1, T_2, \ldots, T_n$ the sample of size n of the variable T. The vector (Y_1, Y_2) follows the generalised bivariate Poisson distribution according to Berkhout and Plug [1], if its mass function is equal to (cf. expression 3.8):*

$$\begin{aligned} P(Y_1 = y_1, Y_2 = y_2) = \mathbb{P}(Y_1 = y_1/T_1 = t_1, T_2 = t_2, \ldots, T_n = t_n) \\ \times \mathbb{P}(Y_2 = y_2/Y_1 = y_1). \end{aligned} \tag{3.9}$$

Under the condition 3.5 and 3.6. We have the following properties.

Properties 3.1. *Under the null hypothesis*

(1) $H_0 : \alpha_i = 0 \quad \forall i$, *the variable Y_1 is independent of the variables T_1, $T_2, \ldots, T_n$, then this law is identically equal to the bivariate Poisson distribution according to Berkhout and Plug [1].*
(2) $H_0 : \eta = 0$, *the variables Y_2 and Y_1 are independent.*

We have the following characteristics.

Proposition 3.2.

$$E(Y_1) = e^{x'\rho_1} \prod_i e^{\lambda(e^{\alpha_i}-1)} \tag{3.10}$$

$$V(Y_1) = E(Y_1) + e^{2x'\rho_1}[\prod_i e^{\lambda(e^{2\alpha_i}-1)} - \prod_i e^{2\lambda(e^{\alpha_i}-1)}] \tag{3.11}$$

$$E(Y_2) = e^{x'\rho_2} e^{\lambda_1(e^{\eta}-1)} \tag{3.12}$$

$$V(Y_2) = E(Y_2) + e^{2x'\rho_2}[e^{\lambda_1(e^{2\eta}-1)} - e^{2\lambda_1(e^{\eta}-1)}] \tag{3.13}$$

$$Cov(Y_1, Y_2) = e^{x'\rho_2} e^{\lambda_1(e^{\eta}-1)}[\lambda_1 e^{\eta} - e^{x'\rho_1} \prod_i e^{\lambda(e^{\alpha_i}-1)}] \tag{3.14}$$

Proof. The moment generating function of the variable Y_1 is equal to

$$M_{Y_1}(z) = \mathbb{E}(e^{zY_1}) = e^{\lambda_1(e^z-1)}.$$

$$\begin{aligned} \mathbb{E}(Y_1) &= \mathbb{E}[\mathbb{E}(Y_1/T_1, \ldots, T_n)] \\ &= \mathbb{E}[\lambda_1(T_1, \ldots, T_n)] \\ &= \mathbb{E}[e^{x'\rho_1} e^{\sum_i \alpha_i T_i}] \\ &= e^{x'\rho_1} \mathbb{E}\left(e^{\sum_i \alpha_i T_i}\right) \\ &= e^{x'\rho_1} \mathbb{E}\left(\prod_i e^{\alpha_i T_i}\right) \\ &= e^{x'\rho_1} \prod_i \mathbb{E}\left(e^{\alpha_i T_i}\right) \\ &= e^{x'\rho_1} \prod_i e^{\lambda(e^{\alpha_i}-1)} \end{aligned}$$

$$\begin{aligned} \mathbb{E}(Y_1^2) &= \mathbb{E}[\mathbb{E}(Y_1^2/T_1, \ldots, T_n)] \\ &= \mathbb{E}[Var(Y_1/T_1, \ldots, T_n)] + \mathbb{E}\left\{[\mathbb{E}(Y_1/T_1, \ldots, T_n)]^2\right\} \\ &= \mathbb{E}[\lambda_1(T_1, \ldots, T_n)] + \mathbb{E}(\lambda_1(T_1, \ldots, T_n))^2 \\ &= \mathbb{E}(Y_1) + \mathbb{E}\left[(e^{x'\rho_1} \prod_i e^{\alpha_i T})^2\right] \\ &= \mathbb{E}(Y_1) + e^{2x'\rho_1} \mathbb{E}(\prod_i e^{2\alpha_i T}]) \\ &= \mathbb{E}(Y_1) + e^{2x'\rho_1} \prod_i \mathbb{E}(e^{2\alpha_i T}) \\ &= \mathbb{E}(Y_1) + e^{2x'\rho_1} \prod_i e^{\lambda(e^{2\alpha_i}-1)} \end{aligned}$$

$$\begin{aligned}
\mathbb{V}(Y_1) &= \mathbb{E}(Y_1^2) - [\mathbb{E}(Y_1)]^2 \\
&= \mathbb{E}(Y_1) + e^{2x'\rho_1} \prod_i e^{\lambda(e^{2\alpha_i}-1)} - e^{2x'\rho_1} \prod_i e^{2\lambda(e^{\alpha_i}-1)} \\
&= \mathbb{E}(Y_1) + e^{2x'\rho_1} [\prod_i e^{\lambda(e^{2\alpha_i}-1)} - \prod_i e^{2\lambda(e^{\alpha_i}-1)}]
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(Y_2) &= \mathbb{E}[\mathbb{E}(Y_2/Y_1)] \\
&= \mathbb{E}[\mathbb{E}(\lambda_2(Y_1)] \\
&= \mathbb{E}[e^{x'\rho_2} e^{\eta Y_1}] \\
&= e^{x'\rho_2} \mathbb{E}[e^{\eta Y_1}] \\
&= e^{x'\rho_2} e^{\lambda_1(e^{\eta}-1)}
\end{aligned}$$

$$\mathbb{V}(Y_2) = \mathbb{E}(Y_2^2) - [\mathbb{E}(Y_2)]^2$$

$$\begin{aligned}
\mathbb{E}(Y_2^2) &= \mathbb{E}[\mathbb{E}(Y_2^2/Y_1)] \\
&= \mathbb{E}[Var(Y_2/Y_1) + \mathbb{E}(Y_2/Y_1)^2] \\
&= \mathbb{E}[Var(Y_2/Y_1)] + \mathbb{E}[\mathbb{E}(Y_2/Y_1)^2] \\
&= \mathbb{E}[\lambda_2(Y_1)] + \mathbb{E}[\lambda_2^2(Y_1)] \\
&= \mathbb{E}(Y_2) + \mathbb{E}[e^{2x'\rho_2 + 2\eta Y_1}] \\
&= \mathbb{E}(Y_2) + e^{2x'\rho_2} \mathbb{E}(e^{2\eta Y_1}) \\
&= \mathbb{E}(Y_2) + e^{2x'\rho_2} e^{\lambda_1(e^{2\eta}-1)}
\end{aligned}$$

So,

$$\begin{aligned}
\mathbb{V}(Y_2) &= \mathbb{E}(Y_2) + e^{2x'\rho_2} e^{\lambda_1(e^{2\eta}-1)} - e^{2x'\rho_2} e^{2\lambda_1(e^{\eta}-1)} \\
&= \mathbb{E}(Y_2) + e^{2x'\rho_2} [e^{\lambda_1(e^{2\eta}-1)} - e^{2\lambda_1(e^{\eta}-1)}]
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(Y_1Y_2) &= \mathbb{E}[\mathbb{E}(Y_1Y_2/Y_1)] \\
&= \mathbb{E}(Y_1\mathbb{E}(Y_2/Y_1) \\
&= \mathbb{E}[Y_1 e^{x'\rho_2} e^{\eta Y_1}] \\
&= e^{x'\rho_2}\mathbb{E}[Y_1 e^{\eta Y_1}] \\
&= e^{x'\rho_2}\mathbb{E}\left[\frac{d}{d\eta}e^{\eta Y_1}\right] \\
&= e^{x'\rho_2}\frac{d}{d\eta}\mathbb{E}[e^{\eta Y_1}] \\
&= e^{x'\rho_2}\frac{d}{d\eta}e^{\lambda_1(e^{\eta}-1)} \\
&= \lambda_1 e^{\eta} e^{x'\rho_2} e^{\lambda_1(e^{\eta}-1)}
\end{aligned}$$

$$\begin{aligned}
Cov(Y_1Y_2) &= \mathbb{E}(Y_1Y_2) - \mathbb{E}(Y_1)\mathbb{E}(Y_2) \\
&= \lambda_1 e^{\eta} e^{x'\rho_2} e^{\lambda_1(e^{\eta}-1)} - e^{x'\rho_1}\prod_i e^{\lambda(e^{\alpha_i}-1)} e^{x'\rho_2} e^{\lambda_1(e^{\eta}-1)} \\
&= e^{x'\rho_2} e^{\lambda_1(e^{\eta}-1)}[\lambda_1 e^{\eta} - e^{x'\rho_1}\prod_i e^{\lambda(e^{\alpha_i}-1)}]
\end{aligned}$$

□

Properties 3.2. *Under the alternative hypothesis (cf. properties* 1*)* $H_1 : \exists i_0$ *such as* $\alpha_{i_0} \neq 0$, *and* $H_1 : \eta \neq 0$, *the expressions* 3.11 *and* 3.13 *shows that the marginal variables* Y_1 *and* Y_2 *are overdispersed.*

4. Estimation of the parameters

Given that the variables Y_1, Y_2 and T, follow Poisson distributions, their respective parameters λ_1, λ_2 and λ will be estimated by the empirical averages of the samples taken.

Concerning, the parameters $\rho_1, \rho_2, \alpha_1, \alpha_2, \alpha_n$ and η it will be enough to solve the linear models 3.5 and 3.6.

5. Simulation of the Poisson regressions

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) $\ln(\lambda_1) = x^T\rho_1 + \sum_{i=1}^{n} \alpha_i t_i$.
(2) $\ln(\lambda_2) = x^T\rho_2 + \eta y_1$.

Tables $1, 2, 3, 4$ below represent simulated data of size 82 and average 2 of the Poisson variables Y_1, Y_2, T_1 and T_2. 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

y_1	0	1	2	3	4	5
Eff. observed	10	19	22	18	6	7

TABLE 2. Table of distribution of y_2

y_2	0	1	2	3	4	5	7
Eff. observed	13	25	21	14	5	3	1

TABLE 3. Table of distribution of t_1

t_1	0	1	2	3	4	5	6
Eff. observed	17	17	18	10	12	4	4

TABLE 4. Table of distribution of t_2

t_2	0	1	2	3	4	5	6
Eff. observed	5	27	17	18	8	6	1

TABLE 5. Table of x

x	0	1	2	3	4
[1]	-0.2234274287	-1.6307471735	-2.0365045043	1.7302872565	0.0424844894
[6]	-1.7692908059	1.0288525742	0.0007483475	1.5427671012	0.3284908516
[11]	-1.8513436668	1.4115969399	2.0357701461	0.4292646764	-1.3270508650
[16]	0.4250898111	1.9588314535	1.2017149653	-0.4754811766	2.3086570508
[21]	1.5614411801	-0.4724178408	-0.9367822196	1.6709703732	0.2493518755
[26]	1.3601494466	1.2887133396	0.2922499219	0.5365834280	-0.3136921186
[31]	1.0294728866	-0.0902694769	0.4830332982	-2.0420793999	0.6096617749
[36]	0.3256575397	0.2340216897	-0.3085852627	0.6435525271	-0.8841832703
[41]	-1.5655128770	-0.7719396725	0.0136725487	-0.3170027173	0.6337965468
[46]	2.4052146745	1.1634279755	0.4357291315	-0.1733713775	-0.0660839121
[51]	0.6065929018	0.7459058098	0.5424165314	-0.5778619427	-0.2027294642
[56]	-0.5063662545	-0.6029991346	0.7763432964	0.8853166138	0.4386325334
[61]	1.2833044167	1.7612946606	2.3917544084	0.9032137340	1.4822505527
[66]	-1.0846848199	-0.3231669389	-0.8289972798	0.5891051748	0.2201731547
[71]	2.4143140831	1.9821262243	-1.0123536855	-0.7008883823	0.8334672265
[76]	0.4995551955	-2.2923028477	0.4581504401	0.9609696613	0.4543416685
[81]	0.8822026461	-2.1588372515			

TABLE 6. Coefficients of regressions 1.

Variable	parameters	$S_{\hat{\rho}}$	t_i	$P(>\|t\|)$
Intercept	0.94687	0.16853	5.619	$<$ 1.93e-08 $***$
x	$\hat{\rho}_1 =$ -0.02567	0.07067	-0.363	0.716
t_1	$\hat{\alpha}_1 =$ -0.04521	0.04502	-1.004	0.315
t_2	$\hat{\alpha}_2 =$ -0.03949	0.05467	-0.722	0.470
		AIC= 289.71		

TABLE 7. Coefficients of regressions 2.

Variable	parameters	$S_{\hat{\rho}}$	t_i	$P(>\|t\|)$
Intercept	0.81896	0.14262	5.742	$<$ 9.34e-09 $***$
x	$\hat{\rho}_2 =$ 0.01929	0.07665	0.252	0.8013
y_1	$\hat{\eta} =$ -0.10315	0.05965	-1.729	0.0838
		AIC= 277.29		

It is evident from the table 6, that at the level $\alpha = 5\%$ of significance, the intercept is not null significantly because is p$-$ value is smaller than α. It also evident from this table that to the same level of significance, the coefficients

ρ_1 α_1 and α_2 are nulls significantly because their p$-$ value are higher than α.

It is evident from the table 7, that at the level $\alpha = 5\%$ of significance, the intercept is not null significantly because is p$-$ value is smaller than α. It also evident from this table that to the same level of significance, the coefficients ρ_2 and η are nulls significantly because their p$-$ value are higher than α.

6. CONCLUSION

This paper allowed us to construct the generalized bivariate Poisson distribution according to Berkhout and Plug [1] through the bivariate law of the probabilities of causes highlighted by Bidounga and al.[2]. We have also calculate their characteristics and finally we simulated this model.

REFERENCES

[1] P. BERKHOUT, E. PLUG: *A bivariate Poisson count data model using conditional probabilities*, Statist. Neerlandica, **58** (3)(2004), 349-364.

[2] R. BIDOUNGA, E.G.B. MANDANGUI MALOUMBI, R.F. MIZELÉ KITOTI, P.C. BATSINDILA NGANGA, D. MIZÈRE: *Construction of a family of bivariate distribution using the probabilities of cause*, Far East Journal of Theorical Statistics, **61** no 1(2021), 35-48.

[3] P. HOLGATE: *Estimation for the bivariate Poisson distribution*, Biometrika, **51**(1964), 241-245.

[4] J. LAKSHMINARAYANA, S.N.N. PANDIT, K. SRINIVASA RAO: *On a Bivariate Poisson distribution*, Communication in Statistics Theory and Methods, **28**(2)(1999), 267-276.

[5] R. F. MIZELÉ KITOTI, R. BIDOUNGA, D. MIZÈRE:*Construction of a New Multivariate Poisson Distribution through the conditional Probability with its variance covariance matrix: Application to the Simulation of Poisson Regressions*, Hikari Ltd, Applied Mathematical Sciences, **13** no. 23(2019), 1117-1128.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF DENIS SASSOU N'GUESSO
KINTELE,
REPUBLIC OF CONGO.
Email address: foxie_reolie2000@yahoo.fr

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF DENIS SASSOU N'GUESSO
KINTELE,
REPUBLIC OF CONGO.
Email address: prevot.batsindila@gmail.com

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MOUNDOU
NDJAMENA,
TCHAD.
Email address: nguessolta.emman@gmail.com

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARIEN N'GOUABI
BRAZZAVILLE,
REPUBLIC OF CONGO.
Email address: rufid@yahoo.fr

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARIEN N'GOUABI
BRAZZAVILLE,
REPUBLIC OF CONGO.
Email address: domizere@gmail.com