Serdica Journal of Computing 17(2), 2023, pp. 79-93, 10.55630/sjc.2023.17.79-93
Published by Institute of Mathematics and Informatics, Bulgarian Academy of Sciences

On a Bivariate Katz's Distribution Constructed by the Trivariate Reduction Method

Michel Koukouatikissa Diafouka[1], Chedly Gélin Louzayadio[2]

[1]Higher Teachers Training College,
Marien Ngouabi University, Republic of Congo
michel.koukouatikissa@umng.cg

[2]Faculty of Economics Sciences,
Marien Ngouabi University, Republic of Congo
gelinlouz@gmail.com

Abstract

In this paper, we propose the bivariate distribution of the Katz distribution [1] constructed by the trivariate reduction method, method developed in [6] and used in [5] to give an equivalent definition of the bivariate Poisson distribution [7, 8]. The constructed distribution includes, in particular, the bivariate Poisson distribution [5] and has interesting properties. For the estimation of the parameters, we used two methods: the method of moments and the maximum likelihood method using the EM algorithm. An application to concrete data has been made in order to carry out a comparative study between bivariate Poisson and Katz distributions, and we discuss the likelihood-ratio test, which assesses the goodness of fit of two competing statistical models.

Keywords: bivariate distributions, trivariate reduction method, Katz's distribution, EM algorithm, likelihood ratio test, multivariate count data
ACM Computing Classification System 2012: Mathematics of computing
Mathematics Subject Classification 2020: 62H10, 62E15, 60E05

Received: November 17, 2022, *Revised:* October 6, 2023
Accepted: November 27, 2023, *Published:* December 14, 2023
Citation: Michel Koukouatikissa Diafouka, Chedly Gélin Louzayadio, On a Bivariate Katz's Distribution Constructed by the Trivariate Reduction Method, Serdica Journal of Computing 17(2), 2023, pp. 79-93, https://doi.org/10.55630/sjc.2023.17.79-93

1 Introduction

Katz [1] has formulated one of the most important families of probability distributions in the analysis and modeling of count data. Defined from the successive probability ratios:

$$\frac{p(z+1)}{p(z)} = \frac{\lambda + \beta z}{z+1}, \quad z = 0, 1, \dots,$$

with $p(0) \neq 0$ and $p(z) = P(Z = z)$, where $\lambda > 0$ and $\beta < 1$, are the canonical and the dispersion' parameters, respectively. It is understood that if $\lambda + \beta z < 0$ then $p(z) = 0$ for $z = 1, 2, \dots$ [2,3]. Its probability mass function (pmf) is given by [4]:

$$p(z) = \begin{cases} \dfrac{\lambda^z}{z!} e^{-\lambda}, & \text{if } \beta = 0, \\[2ex] \dfrac{(\lambda/\beta)_z \beta^z}{z!} (1-\beta)^{\lambda/\beta}, & \text{otherwise,} \end{cases} \qquad z = 0, 1, \dots,$$

where $(\alpha)_z$ is the Pocchammer symbol and defined to be $(\alpha)_z = \alpha(\alpha+1)\dots(\alpha+z-1)$ for $z = 0, 1, \dots,$ and α any real number with $(\alpha)_0 = 1$. In particular, this distribution is reduce to [3]:

- Poisson distribution $\mathcal{P}(\lambda)$, when $\beta = 0$,
- binomial distribution $\mathcal{B}(N, P)$ with $N = -\dfrac{\lambda}{\beta}$ or $\left[-\dfrac{\lambda}{\beta}\right] + 1$ according as $-\dfrac{\lambda}{\beta}$ is or not an integer and $P = \dfrac{-\beta}{1-\beta}$, when $\beta < 0$,
- negative binomial $\mathcal{NB}(r, \beta)$ with $r = \dfrac{\lambda}{\beta}$, when $0 < \beta < 1$.

Each of these last three distributions is the prototype of equi-, under-, and over-dispersed distributions, respectively. So the Katz distribution is a good way to unify Poisson, binomial and negative binomial distributions.

Starting from this univariate Katz distribution, we construct a bivariate Katz distribution using the trivariate reduction method for analysis and modeling of count data. The obtained model is a good way to unify bivariate Poisson, bivariate binomial, and bivariate negative binomial distributions. Since the Katz distribution is a generalization of the Poisson distribution, this work can be viewed as a generalization of [5]. Moreover, this bivariate Katz distribution can be seen as resulting from a combination of the univariate Poisson, binomial, and negative binomial distributions.

The trivariate reduction method was developed in [6] and used in [5] to give an equivalent definition of the bivariate Poisson distribution [7–9]. The basic idea is to construct a pair of dependent but also independent random variables from three independent random variables. Let be consider three independent random variables Z_1, Z_2 and Z_3. The random pair (X, Y) defined as follow:

$$X = Z_1 + Z_3, \quad Y = Z_2 + Z_3, \tag{1}$$

admits for pmf $p(x, y)$, the joint probability $P(X = x, Y = y) = p(x, y)$, given by [10]

$$p(x, y) = \sum_{k=0}^{\min(x,y)} p(x-k)p(y-k)p(k),$$

where $p(t)$ is the pmf of univariate distribution.

On this basis, the rest of the paper is presented as follows. First, we present successively the following notions: probability mass function, probability generating function, moments, correlation and independence, marginal distributions, and recurrent relations. Secondly, we estimate the parameters and we discuss the likelihood-ratio test. For parameter estimation, we use two methods: the method of moments and the maximum likelihood method using the Expectation-Maximization (EM) algorithm. In this work, we do not discuss the existence and uniqueness of the estimators. Finally, we make an application to concrete data in order to carry out a comparative study between bivariate Poisson and Katz distributions. Two criteria were used, Akaike information criterion (AIC) and bayesian information criterion (BIC), for a better choice of distribution.

2 Bivariate Katz's distribution

In this section, we present and study the bivariate Katz distribution from a probabilistic point of view.

2.1 Probability mass function

Let be consider three independent univariate Katz random variables Z_1, Z_2 and Z_3 with parameters (λ_1, β_1), (λ_2, β_2) and (λ_3, β_3), respectively. From (1), the random pair (X,Y), such that $X = Z_1 + Z_3$ and $Y = Z_2 + Z_3$, follows a bivariate Katz distribution which the pmf $p_{BK}(x,y)$ is:

$$p_{BK}(x,y) = (1-\beta_1)^{\lambda_1/\beta_1}(1-\beta_2)^{\lambda_2/\beta_2}(1-\beta_3)^{\lambda_3/\beta_3} \times \sum_{k=0}^{\min(x,y)} \frac{(\lambda_1/\beta_1)_{x-k}(\lambda_2/\beta_2)_{y-k}(\lambda_3/\beta_3)_k \beta_1^{x-k}\beta_2^{y-k}\beta_3^k}{(x-k)!(y-k)!k!}. \quad (2)$$

In particular, we have the following bivariate distributions:

- bivariate Poisson distribution [7]:

$$p(x,y) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{k=0}^{\min(x,y)} \frac{\lambda_1^{x-k}\lambda_2^{y-k}\lambda_3^k}{(x-k)!(y-k)!k!},$$

- bivariate binomial distribution [11]:

$$p(x,y) = N_1!N_2!N_3! \times \sum_{k=0}^{\min(x,y)} \frac{P_1^{x-k}P_2^{y-k}P_3^k(1-P_1)^{N-x+k}(1-P_2)^{N-y+k}(1-P_3)^{N-k}}{(x-k)!(y-k)!k!(N-x+k)!(N-y+k)!(N-k)!},$$

  with $x, y = 0, 1, \ldots, N$,

- bivariate negative binomial distribution [12]:

$$p(x,y) = \frac{\beta_1^{r_1}\beta_2^{r_2}\beta_3^{r_3}}{\Gamma(r_1)\Gamma(r_2)\Gamma(r_3)} \sum_{k=0}^{\min(x,y)} \frac{\Gamma(r_1+x-k)\Gamma(r_2+y-k)\Gamma(r_2+k)}{(x-k)!(y-k)!k!} \times (1-\beta_1)^{x-k}(1-\beta_2)^{y-k}(1-\beta_3)^k.$$

2.2 Probability generating function

The probability generating function G_Z for univariate Katz random variable Z with parameters (λ, β) is given by [2,3]:

$$G_Z(t) = \left[\frac{1-\beta t}{1-\beta}\right]^{-\lambda/\beta}.$$

Let us consider the Katz random pair (X, Y) with parameters $(\lambda_1, \lambda_2, \lambda_3, \beta_1, \beta_2, \beta_3)$, the probability generating function $G_{X,Y}$ of (X, Y) is:

$$\begin{aligned} G_{X,Y}(t_1, t_2) &= E\left[t_1^X t_2^Y\right] \\ &= E\left[t_1^{Z_1}\right] E\left[t_2^{Z_2}\right] E\left[(t_1 t_2)^{Z_3}\right] \\ &= G_{Z_1}(t_1) G_{Z_2}(t_2) G_{Z_3}(t_1 t_2) \\ &= \left[\frac{1-\beta_1 t_1}{1-\beta_1}\right]^{-\lambda_1/\beta_1} \left[\frac{1-\beta_2 t_2}{1-\beta_2}\right]^{-\lambda_2/\beta_2} \left[\frac{1-\beta_3 t_1 t_2}{1-\beta_3}\right]^{-\lambda_3/\beta_3}, \end{aligned}$$

where Z_1, Z_2, Z_3 are independent.

2.3 Moments

The expressions for the first three moments of the Katz are as follows [2,3]:

$$\begin{cases} E(Z) = \dfrac{\lambda}{1-\beta}, \\ V(Z) = \dfrac{\lambda}{(1-\beta)^2}, \\ E(Z^3) = \dfrac{\lambda(1+\lambda)}{(1-\beta)^2}. \end{cases} \tag{3}$$

Since Z_1, Z_2 and Z_3 are independent Katz random variables, and from (1) and (3), we have:

$$\begin{cases} E(X) = \dfrac{\lambda_1}{1-\beta_1} + \dfrac{\lambda_3}{1-\beta_3}, \\ V(X) = \dfrac{\lambda_1}{(1-\beta_1)^2} + \dfrac{\lambda_3}{(1-\beta_3)^2}, \\ E(Y) = \dfrac{\lambda_2}{1-\beta_2} + \dfrac{\lambda_3}{1-\beta_3}, \\ V(Y) = \dfrac{\lambda_2}{(1-\beta_2)^2} + \dfrac{\lambda_3}{(1-\beta_3)^2}. \end{cases} \tag{4}$$

Now let $\mu_{r,s} = E\left[(X - E(X))^r (Y - E(Y))^s\right]$ be the $(r, s)^{th}$ central moment of (X, Y). For $r = s = 1$,

$$\mu_{11} = cov(X, Y) = V(Z_3) = \frac{\lambda_3}{(1-\beta_3)^2}, \tag{5}$$

and for $r = 1$ and $s = 2$ or $r = 2$ and $s = 1$,

$$\mu_{21} = \mu_{12} = E\left(Z_3^3\right) - 3E\left(Z_3^2\right)E\left(Z_3\right) + 2\left[E\left(Z_3\right)\right]^3 = \frac{\lambda_3(1+\beta_3)}{(1-\beta_3)^3}.$$

2.4 Correlation and independence

From (4) and (5), the correlation coefficient of X and Y is:

$$\rho_{XY} = \frac{\lambda_3}{(1-\beta_3)^2\left[\frac{\lambda_1}{(1-\beta_1)^2} + \frac{\lambda_3}{(1-\beta_3)^2}\right]^{1/2}\left[\frac{\lambda_2}{(1-\beta_2)^2} + \frac{\lambda_3}{(1-\beta_3)^2}\right]^{1/2}}.$$

Since $\lambda_3 > 0$ then $\rho_{XY} > 0$, i.e., the correlation coefficient of this model is strictly positive. This shows that the condition of zero correlation is not satisfied and the random variables X and Y can't be independent. As in [13], for the variables X and Y to be independent, it is necessary and sufficient that $X_3 = 0$, which would imply $\lambda_3 = 0$.

2.5 Marginal distributions

The marginal distributions are:

$$P(X = x) = (1-\beta_1)^{\lambda_1/\beta_1}(1-\beta_3)^{\lambda_3/\beta_3}\sum_{k=0}^{x}\frac{(\lambda_1/\beta_1)_{x-k}(\lambda_3/\beta_3)_k\beta_1^{x-k}\beta_3^k}{(x-k)!k!}, \quad (6)$$

and

$$P(Y = y) = (1-\beta_2)^{\lambda_2/\beta_2}(1-\beta_3)^{\lambda_3/\beta_3}\sum_{k=0}^{y}\frac{(\lambda_2/\beta_2)_{y-k}(\lambda_3/\beta_3)_k\beta_2^{y-k}\beta_3^k}{(y-k)!k!}. \quad (7)$$

In particular, if $\beta_1 = \beta_2 = \beta_3 = \beta$, then since $\sum_{k=0}^{n}\frac{(a)_{n-k}(b)_k}{(n-k)!k!} = \frac{(a+b)_n}{n!}$, equations (6) and (7) reduce to:

$$P(X = x) = \frac{((\lambda_1+\lambda_3)/\beta)_x\beta^x}{x!}(1-\beta)^{(\lambda_1+\lambda_3)/\beta},$$

and

$$P(Y = y) = \frac{((\lambda_2+\lambda_3)/\beta)_y\beta^y}{y!}(1-\beta)^{(\lambda_2+\lambda_3)/\beta},$$

i.e., X and Y follow univariate Katz distributions with parameters $(\lambda_1 + \lambda_3, \beta)$ and $(\lambda_2 + \lambda_3, \beta)$, respectively. We have the same result as Theorem 3 in [4].

2.6 Recurrent relations

Following [14], the terms in the first row and column can be computed using the univariate Katz distribution, as is seen from

$$\begin{aligned} p(0,0) &= (1-\beta_1)^{\lambda_1/\beta_1}(1-\beta_2)^{\lambda_2/\beta_2}(1-\beta_3)^{\lambda_3/\beta_3}, & \\ p(x,0) &= (1-\beta_2)^{\lambda_2/\beta_2}(1-\beta_3)^{\lambda_3/\beta_3}p_1(x), & x = 1,2,\ldots, \\ p(0,y) &= (1-\beta_1)^{\lambda_1/\beta_1}(1-\beta_3)^{\lambda_3/\beta_3}p_2(y), & y = 1,2,\ldots, \end{aligned}$$

where p_1 and p_2 are the pmf of univariate Katz distributions with parameters (λ_1, β_1) and (λ_2, β_2), respectively. The probabilities for $x = 1, 2, \ldots,\ y = 1, 2, \ldots$ can be computed recursively as:

$$\begin{aligned} p(x,y) = {} & (1-\beta_1)^{-\lambda_1/\beta_1}(1-\beta_2)^{-\lambda_2/\beta_2}(1-\beta_3)^{-\lambda_3/\beta_3} \\ & \times \sum_{k=0}^{\min(x,y)} p(x-k,0)p(0,y-k)\frac{(\lambda_3/\beta_3)_k \beta_3^k}{k!}. \end{aligned}$$

3 Statistical study

In this section, we study the bivariate Katz distribution from a statistical point of view. Specifically, we estimate the parameters and discuss the adequacy test.

3.1 Parameters estimating

In this subsection, we are interested in the estimation of the parameters, and we use two methods of estimation: the method of moments and the maximum likelihood method. In practice, the moment estimators (in short, MME) can be used as initial values in the algorithm for determining the maximum likelihood estimators (in short, MLE). We then use the EM algorithm to maximize the likelihood function of the bivariate Katz distribution.

3.1.1 Method of moments

The method of moments consists to equal the theoretical moments and the empirical moments in order to determine the estimators. Let be consider a $n-$sample $(x_i, y_i),\ i = 1, 2, \ldots, n$ and note

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \widehat{\sigma}_X^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2, \quad \widehat{\sigma}_Y^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i-\overline{y})^2,$$

and

$$\widehat{\mu}_{j1} = \frac{1}{n}\sum_{x,y}^{n} n_{xy}(x-\overline{x})(y-\overline{y})^j \text{ for } j=1,2,$$

where $n_{x,y}$ is the frequency of the pair (x, y) for $x = 0, 1, \ldots,\ y = 0, 1, \ldots$, and $\sum_{x,y} n_{xy} = n$.

The system equalizing the theoretical moments to the empirical moments is:

$$\begin{cases} \dfrac{\lambda_1}{1-\beta_1} + \dfrac{\lambda_3}{1-\beta_3} = \overline{x}, \\ \dfrac{\lambda_1}{(1-\beta_1)^2} + \dfrac{\lambda_3}{(1-\beta_3)^2} = \widehat{\sigma}_X^2, \\ \dfrac{\lambda_2}{1-\beta_2} + \dfrac{\lambda_3}{1-\beta_3} = \overline{y}, \\ \dfrac{\lambda_2}{(1-\beta_2)^2} + \dfrac{\lambda_3}{(1-\beta_3)^2} = \widehat{\sigma}_Y^2, \\ \dfrac{\lambda_3}{(1-\beta_3)^2} = \widehat{\mu}_{11}, \\ \dfrac{\lambda_3(1+\beta_3)}{(1-\beta_3)^3} = \widehat{\mu}_{21}. \end{cases} \tag{8}$$

We can derive the moment estimators from (8):

$$\begin{cases} \widehat{\lambda}_1 = \dfrac{(\widehat{\mu}_{11}+\widehat{\mu}_{21})^2\overline{x}^2 + 4\left[\widehat{\mu}_{11}^2 - \overline{x}(\widehat{\mu}_{21}+\widehat{\mu}_{11})\right]\widehat{\mu}_{11}^2}{(\widehat{\mu}_{11}+\widehat{\mu}_{21})\left[(\widehat{\mu}_{11}+\widehat{\mu}_{21})\widehat{\sigma}_X^2 - \widehat{\mu}_{11}^2 - \widehat{\mu}_{11}\widehat{\mu}_{21}\right]}, \\ \widehat{\lambda}_2 = \dfrac{(\widehat{\mu}_{11}+\widehat{\mu}_{21})^2\overline{y}^2 + 4\left[\widehat{\mu}_{11}^2 - \overline{y}(\widehat{\mu}_{21}+\widehat{\mu}_{11})\right]\widehat{\mu}_{11}^2}{(\widehat{\mu}_{11}+\widehat{\mu}_{21})\left[(\widehat{\mu}_{11}+\widehat{\mu}_{21})\widehat{\sigma}_Y^2 - \widehat{\mu}_{11}^2 - \widehat{\mu}_{11}\widehat{\mu}_{21}\right]}, \\ \widehat{\lambda}_3 = \dfrac{4\widehat{\mu}_{11}^3}{(\widehat{\mu}_{11}+\widehat{\mu}_{21})^2}, \\ \widehat{\beta}_1 = \dfrac{(\overline{x}-\widehat{\sigma}_X^2)(\widehat{\mu}_{11}+\widehat{\mu}_{21}) + \widehat{\mu}_{11}\widehat{\mu}_{21} - \widehat{\mu}_{11}^2}{\widehat{\mu}_{11}^2 + \widehat{\mu}_{11}\widehat{\mu}_{21} - (\widehat{\mu}_{11}+\widehat{\mu}_{21})\widehat{\sigma}_X^2}, \\ \widehat{\beta}_2 = \dfrac{(\overline{y}-\widehat{\sigma}_Y^2)(\widehat{\mu}_{11}+\widehat{\mu}_{21}) + \widehat{\mu}_{11}\widehat{\mu}_{21} - \widehat{\mu}_{11}^2}{\widehat{\mu}_{11}^2 + \widehat{\mu}_{11}\widehat{\mu}_{21} - (\widehat{\mu}_{11}+\widehat{\mu}_{21})\widehat{\sigma}_Y^2}, \\ \widehat{\beta}_3 = \dfrac{\widehat{\mu}_{21}-\widehat{\mu}_{11}}{\widehat{\mu}_{11}+\widehat{\mu}_{21}}. \end{cases}$$

3.1.2 Maximum likelihood estimation using the EM algorithm

The EM algorithm [15] is an iterative parametric estimation algorithm within the general maximum likelihood framework.

Consider three independent random variables $Z_j,\ j = 1, 2, 3$, which are Katz distributed with parameters (λ_j, β_j), $j = 1, 2, 3$, respectively.

Define $\theta = (\lambda_1, \lambda_2, \lambda_3, \beta_1, \beta_2, \beta_3)$ as the vector of parameters. The joint pmf of (Z_1, Z_2, Z_3) is given by [16]:

$$p(z_1, z_2, z_3|\theta) = \prod_{j=1}^{3} \frac{(1-\beta_j)^{\lambda_j/\beta_j}}{z_j!} \prod_{k=1}^{z} [\lambda_j + \beta_j(k-1)]. \tag{9}$$

Now consider the diffeomorphism φ defined as follows:

$$\begin{aligned} \varphi:\ \mathbb{N}^3 &\longrightarrow \mathbb{N}^3 \\ (z_1, z_2, z_3) &\longmapsto (x, y, t) \end{aligned}$$

such as: $x = z_1 + z_3,\ y = z_2 + z_3$ and $t = z_3$.

The joint pmf of (X, Y, T), using (9), is given by:

$$\begin{aligned} p(x, y, t|\theta) &= p \circ \varphi^{-1}(x, y, t) \left|J_{\varphi^{-1}}\right| = p(z_1, z_2, z_3|\theta) = p(x-t, y-t, t|\theta), \\ &= \frac{(1-\beta_1)^{\lambda_1/\beta_1}}{(x-t)!} \prod_{k=1}^{x-t} [\lambda_1 + \beta_1(k-1)] \frac{(1-\beta_2)^{\lambda_2/\beta_2}}{(y-t)!} \prod_{k=1}^{y-t} [\lambda_2 + \beta_2(k-1)] \\ &\times \frac{(1-\beta_3)^{\lambda_3/\beta_3}}{t!} \prod_{k=1}^{t} [\lambda_3 + \beta_3 t], \end{aligned}$$

where $t \le \min(x, y)$ and $\left|J_{\varphi^{-1}}\right|$ is the absolute value of the determinant of the Jacobian of φ. The log-likelihood corresponding is given by:

$$\begin{aligned} l(\theta) = \sum_{j=1}^{3} \frac{n\lambda_j}{\beta_j} \log(1-\beta_j) &+ \sum_{i=1}^{n} \sum_{k=1}^{x_i - t_i} \log\left[\lambda_1 + \beta_1(k-1)\right] - n\overline{\log(x-t)!} \\ &+ \sum_{i=1}^{n} \sum_{k=1}^{y_i - t_i} \log\left[\lambda_2 + \beta_2(k-1)\right] - n\overline{\log(y-t)!} \\ &+ \sum_{i=1}^{n} \sum_{k=1}^{t_i} \log\left[\lambda_3 + \beta_3(k-1)\right] - n\overline{\log t!}, \end{aligned} \tag{10}$$

where $\overline{\log t!} = \frac{1}{n}\sum_{i=1}^{n} \log t_i!$, with convention $\sum_{k=1}^{t_i} = 0$ for $t_i = 0$.

Following [17], the EM algorithm for the bivariate Katz model (2) is given by:

- **E-step:** Using the current parameter values of r iteration noted by $\lambda_1^{(r)}$, $\lambda_2^{(r)}$, $\lambda_3^{(r)}$, $\beta_1^{(r)}$, $\beta_2^{(r)}$ and $\beta_3^{(r)}$, calculate the conditional expected values of t_i^* for $i = 1, \dots, n$:

$$t_i^* = Q(\theta|\theta^{(r)}) = E[T_i|X_i, Y_i, \theta^{(r)}]$$

$$= \begin{cases} \lambda_3 \frac{p_{BK}\left(x_i - 1, y_i - 1|\theta^{(r)}\right)}{p_{BK}\left(x_i, y_i|\theta^{(r)}\right)} + \beta_3 \left[\lambda_3 \frac{p_{BK}\left(x_i - 2, y_i - 2|\theta^{(r)}\right)}{p_{BK}\left(x_i, y_i|\theta^{(r)}\right)} + \dots \right. \\ \quad + \beta_3 \left[\lambda_3 \frac{p_{BK}\left(x_i - 3, y_i - 3|\theta^{(r)}\right)}{p_{BK}\left(x_i, y_i|\theta^{(r)}\right)} \right. \\ \quad \left.\left. + \ \beta_3 \left[\lambda_3 \frac{p_{BK}\left(x_i - \min(x_i, y_i), y_i - \min(x_i, y_i)|\theta^{(r)}\right)}{p_{BK}\left(x_i, y_i|\theta^{(r)}\right)}\right] \dots \right]\right], \\ \hfill \text{if } \min(x_i, y_i) > 0, \\ 0, \hfill \text{if } \min(x_i, y_i) = 0, \end{cases}$$

where $p_{BK}(x, y|\theta)$ is given in (2).

- **M-step:** By replacing t_i by t_i^* in (10), we obtain:

$$\begin{aligned} l(\theta|\theta^{(r)}) = & \sum_{j=1}^{3} \frac{n\lambda_j}{\beta_j} \log(1 - \beta_j) + \sum_{i=1}^{n} \sum_{k=1}^{x_i - t_i^*} \log\left[\lambda_1 + \beta_1(k-1)\right] - n\overline{\log(x - t^*)!} \\ & + \sum_{i=1}^{n} \sum_{k=1}^{y_i - t_i^*} \log\left[\lambda_2 + \beta_2(k-1)\right] - n\overline{\log(y - t^*)!} \\ & + \sum_{i=1}^{n} \sum_{k=1}^{t_i^*} \log\left[\lambda_3 + \beta_3(k-1)\right] - n\overline{\log t^*!}. \end{aligned} \tag{11}$$

Next, we maximize (11). The two steps are repeated iteratively until the difference between two successive iterations is less than ε, for all $\varepsilon > 0$ quite small. We use a numerical algorithm as the Newton-Raphson procedure for iteratively computing $\theta^{(r)}$, because there is no closed form solution of the M-step [18].

3.2 Hypotheses testing

In this subsection, we perform a test on the dispersion parameters of the bivariate Katz distribution. Since the bivariate Katz distribution reduces to the bivariate Poisson distribution for $\beta_1 = \beta_2 = \beta_3 = 0$, we use the likelihood ratio test to test the bivariate Poisson distribution (constrained parametric model) against the bivariate Katz distribution (unconstrained parametric model).

Following [19], the test hypotheses are reformulated as follows:

$$H_0 \ : \ \beta_1 = \beta_2 = \beta_3 = 0 \text{ versus } H_0 \text{ isn't true.}$$

The null hypothesis H_0 signifies the bivariate Poisson model is reasonable, and the alternative hypothesis $(H_1 : \exists j \in \{1, 2, 3\}, \beta_j \neq 0)$, the bivariate Katz distribution is more appropriate.

To simplify notation, define $\theta = (\lambda_1, \lambda_2, \lambda_3, \beta_1, \beta_2, \beta_3)$ and $\theta_0 = (\lambda_1, \lambda_2, \lambda_3)$. The likelihood ratio test statistic LR is given by [20]:

$$LR = -2\log\left[\frac{L(\theta_0)}{L(\theta)}\right] = -2[l(\theta_0) - l(\theta)], \tag{12}$$

where L and l denote the maximum likelihood and the maximum log-likelihood, respectively.

The LR test statistic follows approximately, under the null hypothesis, the chi-square distribution (χ^2) with degrees of freedom df $= dim(\theta) - dim(\theta_0)$, where dim designates the dimension [21].

4 Applications

In this section, we carry out a comparative study between the bivariate Katz distribution (BKD) and the Poisson distribution (BPD). On this basis, we consider the same real data used in [17] to present an R package called bivpois. This data concerns the demand for Health Care in Australia (Data set 1) and the Italian football championship (Serie A) for season 1991-92 (Data set 2), presented in Tables 1 and 2, respectively. See [13, 17] for more details on the description of these data sets.

For the bivariate Poisson distribution, we used the simple.bp EM function from the bivpois package, which provides a number of fitting statistics. For the bivariate Katz distribution, we wrote an EM algorithm program in the R language environment [22] similar to the simple.bp EM function. Table 3 contains the parameter estimates of the two distributions for the two data sets,

Number of Doctor Consultations (X)	**Number of Prescribed medications** (Y)								
	0	1	2	3	4	5	6	7	8
0	2789	726	307	171	76	32	16	15	9
1	224	212	149	85	50	35	13	5	9
2	49	34	38	11	23	7	5	3	4
3	8	10	6	2	1	1	2	0	0
4	8	8	2	2	3	1	0	0	0
5	3	3	2	0	1	0	0	0	0
6	2	0	1	3	1	2	2	0	1
7	1	0	3	2	1	2	1	0	2
8	1	1	1	0	1	0	1	0	0
9	0	0	0	0	0	0	0	0	1

Table 1: Cross-tabulation of data from the Australian health survey [5].

Goals scored by the home team (X)	**Goals scored by the away team** (Y)					
	0	1	2	3	4	8
0	38	23	13	0	1	0
1	41	58	12	10	3	0
2	28	19	10	3	0	1
3	6	11	4	4	1	0
4	7	5	1	0	1	0
5	2	2	2	0	0	0

Table 2: Cross-tabulation of data for Italian football championship (Serie A) for season 1991-92.

and Table 4 contains the fitting statistics: values of the loglikelihood, AIC, BIC, and the number of iterations (Iter). The moment estimates were used as initial values for the EM algorithm for the bivariate Katz distribution. The package maxLik for the R statistical environment [23] was used for the optimization of (11).

Table 3 shows that the dispersion parameters are positive (or null), and as the bivariate Katz distribution can result from a combination of the univariate Poisson, binomial, and negative binomial distributions, for these two data sets, the bivariate Katz distribution is seen as a combination of a univariate Poisson distribution and two negative binomials, i.e., the bivariate Katz is none other than the bivariate Poisson-negative binomial distribution. And from Table 4,

Data set	Distri-butions	Parameters					
		λ_1	λ_2	λ_3	β_1	β_2	β_3
1	**BPD**	0.176	0.737	0.125			
	BKD	0.173	0.362	0.097	7.2×10^{-24}	0.505	0.241
2	**BPD**	1.242	0.834	0.096			
	BKD	0.967	0.620	0.180	0.165	0.174	8.1×10^{-25}

Table 3: Parameters estimation.

Data set	Distributions	loglikelihood	AIC	BIC	Iter.
1	**BPD**	-11268.36	22542.71	22564.46	14
	BKD	-9382.162	18776.32	18819.81	22
2	**BPD**	-845.4001	1696.800	1710.050	74
	BKD	-789.4036	1590.807	1617.308	147

Table 4: Goodness-of-fit.

depending on the values of AIC and BIC, the bivariate Katz distribution would be preferable to the Poisson one, despite the number of additional parameters. That is confirmed by the ratio test. Indeed, the values of the $LR-$statistic corresponding to two data sets are 3772.396 and 111.993, respectively, and for df $= 3$, $\chi^2 = 7.8147$ at the significance level of 5%. In other terms, the bivariate Katz model is more appropriate than the Poisson model.

5 Conclusion

In this paper, we have proposed the bivariate Katz distribution constructed by the trivariate reduction technique. This model has some interesting properties, and is a good way to unify bivariate Poisson, bivariate binomial, and bivariate negative binomial distributions. Katz's bivariate distribution is therefore a generalization of Poisson's and is as flexible and competitive as other bivariate distributions in the literature, as the results obtained in the application clearly illustrate.

The work carried out presents several avenues for future research. More precisely, we are working to propose goodness-of-fit tests for the bivariate Katz distribution, generalizing the goodness-of-fit tests for the bivariate Poisson distribution [24], and to realize a regression study of the uni- and bi-variate Katz models that will enable us to carry out an in-depth study of the data processed in Section 3. We will also carry out a discussion on the existence and uniqueness of the estimators.

Acknowledgements

The authors thank the reviewers who helped improve the work and the Editor.

References

[1] L. Katz, Unified treatment of a broad class of discrete probability distributions, *Classical and contagious discrete distributions*, 1:175–182, 1965.

[2] M. A. B. Ghahfarokhi, H. Iravani, M. R. Sepehri, Application of Katz family of distributions for detecting and testing overdispersion in Poisson regression models, *World Academy of Science, Engineering and Technology*, 42:514–519, 2008.

[3] M. A. B. Ghahfarokhi, P. B. Ghahfarokhi, R. Bahramian, Application of Katz family of distributions for detecting and testing overdispersion in Poisson models, *Journal of Statistical Modeling and Analytics*, 1:29–44, 2010.

[4] A. R. Washburn, Katz distributions, with applications to minefield clearance, Naval Postgraduate School, Monterey, California, 1996.

[5] P. Holgate, Estimation for the bivariate Poisson distribution, *Biometrika*, 51:241–245, 1964.

[6] G. K. Eagleson, Polynomial expansions of bivariate distributions, *Annals of Mathematical Statistics*, 35:1208–1215, 1964.

[7] J. T. Campbell, The Poisson correlation function, *Proceedings of the Edinburgh Mathematical Society*, 4:18–26, 1934.

[8] H. Teicher, On the multivariate Poisson distribution, *Scandinavian Actuarial Journal*, 1954:1–9, 1954.

[9] N. L. Johnson, S. Kotz, *Distributions in statistics: discrete distributions*, John Wiley and Sons, New Jersey, 1969.

[10] S. Kocherlakota, K. Kocherlakota, *Bivariate discrete distributions*, Boca Raton, CRC Press, 1992.

[11] A. W. Marshall, I. Olki, A family of bivariate distributions generated by the bivariate Bernoulli distribution, *Journal of the American Statistical Association*, 80:332–338, 1985.

[12] F. Liu, D. Pitt, Application of bivariate negative binomial regression model in analysing insurance count data, *Annals of Actuarial Science*, 11:390–411, 2017.

[13] D. Karlis, I. Ntzoufras, Analysis of sports data by using bivariate Poisson models, *Journal of the Royal Statistical Society D (The Statistician)*, 52:381–393, 2003.

[14] R. Vernic, On the bivariate generalized Poisson distribution, *Astin Bulletin*, 27:23–31, 1997.

[15] A. P. Dempster, N. M. Laird, D. B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)*, 39:1–38, 1977.

[16] M. Koukouatikissa Diafouka, C. G. Louzayadio, R. O. Malouata, N. R. Ngabassaka, R. Bidounga, On a bivariate Katz's distribution, *Advances in Mathematics: Scientific Journal*, 11:955–968, 2022.

[17] D. Karlis, I. Ntzoufras, Bivariate Poisson and Diagonal Inflated Bivariate Poisson Regression Models in R, *Journal of Statistical Software*, 14(10):1–36, 2005.

[18] R. E. Bedar, Parameters Estimation of the Bivariate Weibull Distribution by EM Algorithm Based on Censored Samples, *Asian Journal of Applied Sciences*, 5(6):1145–1152, 2017.

[19] S. H. Ong, R. C. Gupta, T. Ma, S. Z. Sim, Bivariate Conway-Maxwell Poisson Distributions with Given Marginals and Correlation, *Journal of Statistical Theory and Practice*, 15:10, 2021.

[20] S. D. Silvey, *Statistical Inference*, Chapman and Hall, London, 1970.

[21] K.-R. Koch, *Parameter Estimation and Hypothesis Testing in Linear Models*, Springer, New York, 1988.

[22] R Core Team, R: A language and environment for statistical computing, *R Foundation for Statistical Computing*, Vienna, Austria, 2022.

[23] A. Henningsen, O. Toomet, maxlik: A package for maximum likelihood estimation in R, *Computational Statistics*, 26:443–458, 2011.

[24] F. Novoa-Muñoz, M. D. Jiménez-Gamero, Testing for the bivariate Poisson distribution, *Metrika*, 77:771–793, 2014.