International Journal of Mathematics and Computer Science
Volume **20,** Issue no. 4, (2025), 1057–1073
DOI: https://doi.org/10.69793/ijmcs/04.2025/chedly

The bivariate extended Poisson distribution of type 3 and applications

Chedly Gélin Louzayadio[1], Emmanuel Nguessolta[2], Michel Koukouatikissa Diafouka[1], Dominique Mizère[1]

[1]Department of Mathematics
Marien NGOUABI University
BP. 69
Brazzaville, Congo

[2]Department of Mathematics
University of Moundou
BP. 206
Moundou, Tchad

email: gelinlouz@gmail.com, nguessolta.emman@gmail.com, michel.koukouatikissa@umng.cg, domizere@gmail.com

(Received September 18, 2025, Revised October 4, 2025, Accepted October 14, 2025, Published October 18, 2025)

Abstract

In this paper, we introduce a new bivariate count distribution, called the Bivariate Extended Poisson distribution of type 3 (BEP3 distribution), constructed using the trivariate reduction method, which is widely used in the statistical literature to model positive dependence between discrete variables. This distribution accommodates both overdispersion or underdispersion and strictly positive dependence between two count variables. We study its main theoretical properties and estimate its parameters using the maximum likelihood method, based on an Expectation-Maximization (EM) algorithm. We compare the goodness-of-fit of the new distribution to that of other bivariate models for count data on two real-world datasets. The BEP3

Key words and phrases: Bivariate count data, extended Poisson distribution, trivariate reduction method, EM algorithm, overdispersion.
AMS (MOS) Subject Classifications: 62H10, 62E15, 60E05.
ISSN 1814-0432, 2025, https://future-in-tech.net

distribution was found to provide a better fit than some of the models considered.

1 Introduction

The bivariate Poisson distribution, derived from the univariate Poisson distribution, is a classical model for the study of correlated count data [15]. Several forms of this distribution can be found in the statistical literature, constructed using various methods depending on the type of dependence desired (negative, zero, or positive) [13]. A well-known form offers a simple formulation in which the dependence between the two variables is positive and induced by a common latent component, often referred to as the common variable model [11, 14]. However, its flexibility remains limited in certain applications due to structural constraints, particularly when modeling data exhibiting overdispersion or underdispersion [23].

To overcome these limitations, several alternative distributions have been proposed, such as the bivariate generalized Poisson distribution [9], the bivariate Poisson-Lindley distribution [8], and the bivariate Katz distribution [5]. The bivariate extended Poisson (BEP) distribution, derived from the univariate extended Poisson distribution introduced by [6], constitutes one of the most promising alternatives to the bivariate Poisson distribution, particularly for modeling overdispersed or underdispersed data.

Recent studies have introduced two distinct types of BEP distributions to enhance the modeling flexibility of bivariate count data. For example, [2] proposed the type 1 BEP distribution, obtained by multiplying the extended Poisson (EP) marginals by a dependence factor, following a method inspired by [16]. This construction retains flexible EP marginals (allowing under or over-dispersion) and permits negative, zero or positive correlation via a single parameter, yet the multiplicative form imposes positivity constraints that bound the attainable correlation and restrict the shape of dependence. Subsequently, [19] developed the type 2 BEP distribution using the conditional marginal approach, explicitly introduced by [1] and used by [7]. It preserves EP flexibility while offering a regression like structure $Y_2 \mid Y_1$ that integrates covariates naturally and yields interpretable dependence, although the induced dependence is asymmetric (non-exchangeable) and, as with type 1, the magnitude of correlation remains bounded, which can complicate estimation and calibration in practice.

In this work, we propose a third version of the bivariate extended Poisson (BEP3) distribution, based on the trivariate reduction method [11]. This ap-

proach models dependence by introducing a common latent variable, which is added to each of the marginal count variables, each following a univariate extended Poisson (EP) distribution. However, a limitation of this method is that it can generate only positive correlation, which may restrict its applicability in contexts where negative or zero dependence is observed.

The remainder of this article is organized as follows. In Section 2, we review univariate extended Poisson distribution. Section 3 is devoted to the construction of the BEP3 distribution using the trivariate reduction approach, detailing its probability mass function and marginal distributions. We study its properties, in particular, the joints probability and moment generating functions, factorial moments and ordinary moments. In Section 4, we estimate the parameters of the BEP3 distribution using the maximum likelihood method, implemented via the EM algorithm [4]. Finally, in Section 5, we compare the goodness-of-fit of the BEP3 distribution relative to the bivariate Poisson distribution as well as to the BEP1 and BEP2 distributions, on two real-world datasets: healthcare demand in Australia [11, 14] and soccer [17]. The model selection is performed using Akaike's information criterion (AIC) and Bayesian information criterion (BIC); the model with the lowest AIC and BIC values is considered the best fit for the data.

2 Review of extended Poisson (EP) distribution

The univariate extended Poisson distribution, introduced by [6], is an extension of the classical Poisson distribution for modeling over- or under-dispersed count data (variance greater than or less than the mean). Its probability mass function (PMF) is defined as follows:

$$P(Y=y)=\begin{cases}\dfrac{1-e^{-\theta}}{\beta}, & \text{if } y=0,\\ \dfrac{e^{-\theta}\theta^{y}}{y!}\cdot\dfrac{1}{\beta}\left(\dfrac{\beta y}{\theta}-1\right), & \text{if } y=1,2,\ldots\end{cases}\qquad \text{with } \theta>0 \text{ and } \beta>\theta. \tag{2.1}$$

The PMF of the univariate extended Poisson distribution can also be written as, see [18]:

$$P(Y=y)=\frac{\theta^{y}}{y!}e^{-\theta}\left(\frac{\beta y}{\theta}-1\right)\cdot\frac{1}{\beta}\left[\frac{1-e^{-\theta}}{\left(\frac{\beta y}{\theta}-1\right)e^{-\theta}}\right]^{\delta_0(y)}; \tag{2.2}$$

where

$$\delta_0(y) = \begin{cases} 1 & if \quad y = 0 \\ 0 & else \end{cases}$$

is the Dirac delta function at 0.
In [6], the probability generating function (PGF) of a random variable Y is given by

$$G_Y(t) = \mathbb{E}(t^Y) = \frac{1 + (\beta t - 1)\, e^{\theta(t-1)}}{\beta}, \qquad |t| \leq 1. \tag{2.3}$$

And the moment generating function (MGF) is given by $M_Y(s) = G_Y(e^s)$. We deduce the factorial moments of order $r \geq 1$

$$\mu_{[r]}^{(Y)} = \mathbb{E}\big[Y_{[r]}\big] = \frac{(\beta - 1)\theta^r + r\beta\, \theta^{r-1}}{\beta} = r\, \theta^{r-1} + \Big(1 - \frac{1}{\beta}\Big)\theta^r, \tag{2.4}$$

where $Y_{[r]} = Y(Y-1)\cdots(Y-r+1)$.
And then the ordinary moments via Stirling numbers of the second kind $S(r,k)$ [20]

$$m_r^Y = \mathbb{E}(Y^r) = \sum_{k=0}^{r} S(r,k)\, \mu_{[k]}^{(Y)}, \tag{2.5}$$

avec

$$S(r,k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^{\,k-j} \binom{k}{j}\, j^{\,r} \qquad (r, k \in \mathbb{N},\ 0 \leq k \leq r).$$

In particular,

$$\mathbb{E}(Y) = 1 + \Big(1 - \frac{1}{\beta}\Big)\theta \tag{2.6}$$

and

$$\mathrm{Var}(Y) = \mu_{[2]}^{(Y)} + \mu_{[1]}^{(Y)} - \big(\mu_{[1]}^{(Y)}\big)^2 = \frac{\theta\big(\beta^2 + \beta\theta + \beta - \theta\big)}{\beta^2}. \tag{2.7}$$

This distribution is highly flexible and adapts well to real discrete data where the variance differs significantly from the mean. Unlike the classical Poisson distribution, it can account for both overdispersion and underdispersion, depending on the range of the parameter values [2].

3 Construction of the BEP3 distribution

In this section, we propose the third type of bivariate extended Poisson distribution, which we denote BEP3 distribution. It is constructed using the trivariate reduction approach, more commonly called the latent variable summation method used by [11, 14] to construct the bivariate Poisson distribution. Essentially, it links two discrete variables through an underlying latent variable.

3.1 Probability mass function

Definition 3.1. *Let X_1, X_2 and U be three independent univariate extended Poisson random variables with parameters (θ_1, β_1), (θ_2, β_2) and (θ, β), respectively. We construct two new variables*

$$Y_1 = X_1 + U \text{ and } Y_2 = X_2 + U. \tag{3.8}$$

The random pair (Y_1, Y_2) follows a BEP3 distribution of the parameters vector $\Lambda = (\theta_1, \theta_2, \theta, \beta_1, \beta_2, \beta)$ whose PMF f_{BEP3} is

$$f_{BEP3}(y_1, y_2; \Lambda) = P(Y_1 = y_1, Y_2 = y_2),$$

with

$$P(Y_1 = y_1, Y_2 = y_2) = \sum_{l=0}^{t} P(X_1 = y_1 - l)P(X_2 = y_2 - l)P(U = l)$$

and $t = \min(y_1, y_2)$, *with* $\beta_j \geq \theta_j$ $(j = 1, 2)$ *and* $\beta \geq \theta$.

Parameters θ_1, θ_2, and θ correspond to the marginal and shared components of the latent Poisson structure, while β_1, β_2, and β capture overdispersion or underdispersion effects in the BEP3 distribution.

Proposition 3.2. *The joint distribution of the random variables Y_1 and Y_2 is given by*

$$f_{BEP3}(y_1, y_2; \Lambda) = \left(\frac{\theta_1^{y_1}}{y_1!}e^{-\theta_1}\right)\left(\frac{\theta_2^{y_2}}{y_2!}e^{-\theta_2}\right) \times h(y_1, y_2; \Lambda), \tag{3.9}$$

where

$$\begin{aligned} h(y_1,y_2;\Lambda) &= \frac{e^{-\theta}}{\beta_1\beta_2\beta}\left[\left(\frac{\beta_1}{\theta_1}y_1-1\right)\left(\frac{\beta_2}{\theta_2}y_2-1\right)\left(\frac{1-e^{-\theta}}{e^{-\theta}}\right)+\right.\\ &\sum_{l=1}^{t-1}(-y_1)_l(-y_2)_l\left(\frac{\beta_1}{\theta_1}(y_1-l)-1\right)\left(\frac{\beta_2}{\theta_2}(y_2-l)-1\right)\times\\ &\left(\frac{\beta}{\theta}l-1\right)\frac{(\theta/\theta_1\theta_2)^l}{l!}+(-y_1)_t(-y_2)_t\frac{(\theta/\theta_1\theta_2)^t}{t!}\times\\ &\left\{\left(\frac{\beta_2}{\theta_2}(y_2-y_1)-1\right)\left(\frac{1-e^{-\theta_1}}{e^{-\theta_1}}\right)\right\}^{\delta_{y_1}(t)}\times\\ &\left.\left\{\left(\frac{\beta_1}{\theta_1}(y_1-y_2)-1\right)\left(\frac{1-e^{-\theta_2}}{e^{-\theta_2}}\right)\right\}^{\delta_{y_2}(t)}\left(\frac{\beta}{\theta}t-1\right)\right] \quad (3.10) \end{aligned}$$

with

$$(-y_i)_l=\begin{cases}(-1)^l\frac{y_i!}{(y_i-l)!} & if \quad l\le y_i\\ 0 & else\end{cases}$$

is the Pocchammer symbol and $\delta_{y_i}(t)$ Dirac function at y_i $(i=1,2)$. Since $h(y_1,y_2;\Lambda)\neq 1$ in general, this confirms that Y_1 and Y_2 are dependent.

Proof.

$$\begin{aligned} P(Y_1=y_1,Y_2=y_2) &= \sum_{l=0}^{t}\left(\frac{\theta_1^{(y_1-l)}}{(y_1-l)!}e^{-\theta_1}\right)\left(\frac{\theta_2^{(y_2-l)}}{(y_2-l)!}e^{-\theta_2}\right)\left(\frac{\beta_1}{\theta_1}(y_1-l)-1\right)\times\\ &\left\{\frac{1-e^{-\theta_1}}{\left(\frac{\beta_1}{\theta_1}(y_1-l)-1\right)e^{-\theta_1}}\right\}^{\delta_0(y_1-l)}\times\left(\frac{\beta_2}{\theta_2}(y_2-l)-1\right)\times\\ &\left\{\frac{1-e^{-\theta_2}}{\left(\frac{\beta_2}{\theta_2}(y_2-l)-1\right)e^{-\theta_2}}\right\}^{\delta_0(y_2-l)}\times\frac{\theta^l}{l!}e^{-\theta}\left(\frac{\beta}{\theta}l-1\right)\times\\ &\left\{\frac{1-e^{-\theta}}{\left(\frac{\beta}{\theta}l-1\right)e^{-\theta}}\right\}^{\delta_0(l)}\times\frac{1}{\beta_1\beta_2\beta}\\ &= \left(\frac{\theta_1^{y_1}}{y_1!}e^{-\theta_1}\right)\left(\frac{\theta_2^{y_2}}{y_2!}e^{-\theta_2}\right)\times h(y_1,y_2;\Lambda), \end{aligned}$$

where,

$$\begin{aligned} h(y_1,y_2;\Lambda) &= \frac{e^{-\theta}}{\beta_1\beta_2\beta}\sum_{l=0}^{t}\frac{(-1)^l y_1!}{(y_1-l)!}\frac{(-1)^l y_2!}{(y_2-l)!}\frac{(\theta/\theta_1\theta_2)^l}{l!}\times \\ &\left(\frac{\beta_1}{\theta_1}(y_1-l)-1\right)\left\{\frac{1-e^{-\theta_1}}{\left(\frac{\beta_1}{\theta_1}(y_1-l)-1\right)e^{-\theta_1}}\right\}^{\delta_0(y_1-l)}\times \\ &\left(\frac{\beta_2}{\theta_2}(y_2-l)-1\right)\left\{\frac{1-e^{-\theta_2}}{\left(\frac{\beta_2}{\theta_2}(y_2-l)-1\right)e^{-\theta_2}}\right\}^{\delta_0(y_2-l)}\times \\ &\left(\frac{\beta}{\theta}l-1\right)\left\{\frac{1-e^{-\theta}}{\left(\frac{\beta}{\theta}l-1\right)e^{-\theta}}\right\}^{\delta_0(l)} \\ &= \frac{e^{-\theta}}{\beta_1\beta_2\beta}\sum_{l=0}^{t}A_l = \frac{e^{-\theta}}{\beta_1\beta_2\beta}\left[A_0+\sum_{l=1}^{t-1}A_l+A_t\right]. \end{aligned}$$

This is calculated as follows

$$A_0=\left(\frac{\beta_1}{\theta_1}y_1-1\right)\left(\frac{\beta_2}{\theta_2}y_2-1\right)\left(\frac{1-e^{-\theta}}{e^{-\theta}}\right),$$

$$\sum_{l=1}^{t-1}A_l=\sum_{l=1}^{t-1}(-y_1)_l(-y_2)_l\left(\frac{\beta_1}{\theta_1}(y_1-l)-1\right)\left(\frac{\beta_2}{\theta_2}(y_2-l)-1\right)\left(\frac{\beta}{\theta}l-1\right)\frac{(\theta/\theta_1\theta_2)^l}{l!}$$

and

$$\begin{aligned} A_t &= (-y_1)_t(-y_2)_t\frac{(\theta/\theta_1\theta_2)^t}{t!}\left\{\left(\frac{\beta_2}{\theta_2}(y_2-y_1)-1\right)\left(\frac{1-e^{-\theta_1}}{e^{-\theta_1}}\right)\right\}^{\delta_{y_1}(t)} \\ &\times \left\{\left(\frac{\beta_1}{\theta_1}(y_1-y_2)-1\right)\left(\frac{1-e^{-\theta_2}}{e^{-\theta_2}}\right)\right\}^{\delta_{y_2}(t)}\left(\frac{\beta}{\theta}t-1\right), \end{aligned}$$

we are assured of the result. □

Corollary 3.3. *We have the following result*

$$f_{BEP3}(y_1, y_2; \Lambda) = h(y_1, y_2; \Lambda) \times f_{BP}(y_1, y_2; \theta_1, \theta_2) \tag{3.11}$$

Expression 3.11 shows that the BEP3 distribution is a member of the family of bivariate Poisson distributions[1].

3.2 Marginal distributions

For $i \in \{1, 2\}$, since $Y_i = X_i + U$, where (X_1, X_2, U) are mutually independent, the marginal distribution of Y_i is the convolution of the distributions of X_i and U. Thus, for $y_i \in \mathbb{N}$ (i=1,2),

$$P(Y_i = y_i) = \sum_{l=0}^{y_i} P\big(X_i = y_i - l\big)\, P(U = l).$$

We can write

$$P(Y_i = y_i) = \sum_{l=0}^{y_i} f_{\text{EP}}\big(y_i - l; \theta_i, \beta_i\big)\, f_{\text{EP}}\big(l; \theta, \beta\big). \tag{3.12}$$

Noting $f_{\text{EP}}(k; \theta, \beta)$ as the probability mass function of the extended Poisson distribution $\text{EP}(\theta, \beta)$.

3.3 Fundamental statistical properties

The joint probability generating function (PGF) of the BEP3 distribution is given by

$$\begin{aligned} G_{Y_1,Y_2}(t_1, t_2) &= E\left(t_1^{Y_1} t_2^{Y_2}\right) = E\left(t_1^{X_1+U} t_2^{X_2+U}\right) \\ &= E\left(t_1^{X_1}\right) E\left(t_2^{X_2}\right) E\left[(t_1 t_2)^U\right] \ (\textit{by i.i.d of } X_i,\ (i = 1, 2) \textit{ and } U) \\ &= \left(\frac{1 + (\beta_1 t_1 - 1)\, e^{\theta_1 (t_1 - 1)}}{\beta_1}\right) \left(\frac{1 + (\beta_2 t_2 - 1)\, e^{\theta_2 (t_2 - 1)}}{\beta_2}\right) \times \\ &\quad \left(\frac{1 + (\beta t_1 t_2 - 1)\, e^{\theta (t_1 t_2 - 1)}}{\beta}\right). \end{aligned}$$

The joint moment generating function (MGF) is defined as

$$M_{Y_1,Y_2}(s_1, s_2) = G_{Y_1,Y_2}(e^{s_1}, e^{s_2}) = M_{X_1}(s_1)\, M_{X_2}(s_2)\, M_U(s_1 + s_2).$$

From the convolution formula for falling factorials, the marginal factorial moments are derived.

$$\mu_{[r]}^{(Y_i)} = \sum_{k=0}^{r} \binom{r}{k} \mu_{[k]}^{(X_i)} \mu_{[r-k]}^{(U)} \quad (i = 1, 2).$$

The ordinary marginal moments are given by

$$\mathbb{E}(Y_i^r) = \sum_{k=0}^{r} S(r,k)\, \mu_{[k]}^{(Y_i)} \quad (i = 1, 2).$$

Using the identity [10]

$$(x+u)_{[r]} = \sum_{k=0}^{r} \binom{r}{k} x_{[r-k]} u_{[k]}$$

and [3, 22]

$$u_{[k]}\, u_{[\ell]} = \sum_{m=0}^{\min(k,\ell)} \binom{k}{m} \binom{\ell}{m} m!\, u_{[k+\ell-m]},$$

we obtain the joint factorial moments, for $r_1, r_2 \geq 0$,

$$\mu_{[r_1,r_2]}^{(Y_1,Y_2)} = \sum_{k=0}^{r_1} \sum_{\ell=0}^{r_2} \binom{r_1}{k} \binom{r_2}{\ell} \mu_{[r_1-k]}^{(X_1)}\, \mu_{[r_2-\ell]}^{(X_2)} \sum_{m=0}^{\min(k,\ell)} \binom{k}{m} \binom{\ell}{m} m!\, \mu_{[k+\ell-m]}^{(U)},$$

where all $\mu_{[\cdot]}^{(\cdot)}$ refer to those of the EPD (see formula 2.4).
The ordinary joint moments are obtained from this by

$$\mathbb{E}(Y_1^{r_1} Y_2^{r_2}) = \sum_{i=0}^{r_1} \sum_{j=0}^{r_2} S(r_1,i)\, S(r_2,j)\, \mu_{[i,j]}^{(Y_1,Y_2)}.$$

It follows that the expectations, variances, covariance, and correlation coefficient are respectively given by

$$\mathbb{E}(Y_i) = 2 + \Big(1 - \frac{1}{\beta_i}\Big)\theta_i + \Big(1 - \frac{1}{\beta}\Big)\theta, \qquad (i = 1, 2)$$

$$\mathrm{Var}(Y_i) = \frac{\theta_i(\beta_i^2 + \beta_i\theta_i + \beta_i - \theta_i)}{\beta_i^2} + \frac{\theta(\beta^2 + \beta\theta + \beta - \theta)}{\beta^2}, \qquad (i = 1, 2)$$

$$\mathrm{Cov}(Y_1, Y_2) = \mathrm{Var}(U) = \frac{\theta(\beta^2 + \beta\theta + \beta - \theta)}{\beta^2},$$

and

$$\rho_{(Y_1,Y_2)} = \frac{\mathrm{Var}(U)}{\sqrt{\mathrm{Var}(Y_1)\,\mathrm{Var}(Y_2)}}.$$

4 Parameters estimation

In this subsection, we present the estimation for the parameter vector $\Lambda = (\theta_1, \theta_2, \theta, \beta_1, \beta_2, \beta)$ of the BEP3 distribution using the maximum likelihood (ML) method based on the EM (Expectation-Maximization) algorithm [4]. Let $(Y_{1i},\, Y_{2i},\, Z_i)$ be a sample of size n of random variables such that: $Y_{1i} = X_{1i} + U_i$, $Y_{2i} = X_{2i} + U_i$ and $Z_i = U_i$ (latent variable), where $(X_{1i},\, X_{2i},\, U_i)$ is a sample of n independent random variables X_1, X_2 and U, each following an univariate extended Poisson distribution of parameters (θ_1, β_1), (θ_2, β_2) and (θ, β) respectively. The joint mass function of (Y_1, Y_2, Z) is written:

$$P(Y_1 = y_1, Y_2 = y_2, Z = z) = P(X_1 = y_1 - z) \cdot P(X_2 = y_2 - z) \cdot P(U = z)$$

Consequently, the complete log-likelihood function is given by

$$\ell(\Lambda) = \sum_{i=1}^{n} \left[\ln P(X_{1i} = y_{1i} - z_i) + \ln P(X_{2i} = y_{2i} - z_i) + \ln P(U_i = z_i) \right] \quad (4.13)$$

under the conditions $y_{1i} \geq z_i \geq 0$ and $y_{2i} \geq z_i \geq 0$ for all $i = 1, \ldots, n$.

The EM algorithm for the BEP3 distribution is composed of two iterative steps:

Step E (Expectation)

At iteration k, for each observation $i = 1, \ldots, n$, we calculate the conditional expectation of the latent variable Z_i:

$$\mathbb{E}\left[Z_i \mid Y_{1i}, Y_{2i}, \Lambda^{(k)}\right] = \sum_{z=0}^{\min(y_{1i}, y_{2i})} z \cdot w_i^{(k)}(z) \quad (4.14)$$

where the weights $w_i^{(k)}(z)$ are given by :

$$w_i^{(k)}(z) = \frac{f_{EP}(y_{1i} - z; \theta_1^{(k)}, \beta_1^{(k)}) \cdot f_{EP}(y_{2i} - z; \theta_2^{(k)}, \beta_2^{(k)}) \cdot f_{EP}(z; \theta^{(k)}, \beta^{(k)})}{\sum_{z'=0}^{\min(y_{1i}, y_{2i})} f_{EP}(y_{1i} - z'; \theta_1^{(k)}, \beta_1^{(k)}) \cdot f_{EP}(y_{2i} - z'; \theta_2^{(k)}, \beta_2^{(k)}) \cdot f_{EP}(z'; \theta^{(k)}, \beta^{(k)})} \tag{4.15}$$

with $\Lambda^{(k)} = (\theta_1^{(k)}, \theta_2^{(k)}, \theta^{(k)}, \beta_1^{(k)}, \beta_2^{(k)}, \beta^{(k)})$ the vector of parameters at iteration k.

Step M (Maximization)

We maximize the expected complete log-likelihood:

$$\Lambda^{(k+1)} = \arg\max_{\Lambda} Q(\Lambda | \Lambda^{(k)}) \tag{4.16}$$

with the objective function defined by:

$$\begin{aligned} Q(\Lambda \mid \Lambda^{(k)}) \;=\; & \sum_{i=1}^{n} \sum_{z=0}^{t_i} w_i^{(k)}(z) \, [\log f_{EP}(y_{1i} - z; \theta_1, \beta_1) \\ & + \log f_{EP}(y_{2i} - z; \theta_2, \beta_2) + \log f_{EP}(z; \theta, \beta)] \,, \end{aligned} \tag{4.17}$$

where $t_i = \min(Y_{1i}, Y_{2i})$.
This maximization is performed numerically by using the optim function in R's stats package [21].

Stop criterion

The algorithm is repeated until convergence, determined according to the following criterion:

$$\frac{|\ell(\Lambda^{(k+1)}) - \ell(\Lambda^{(k)})|}{|\ell(\Lambda^{(k)})|} < \epsilon, \tag{4.18}$$

where ϵ is a predefined tolerance (typically $\epsilon = 10^{-6}$).

5 Applications to real data

In this section, we compare the goodness-of-fit of the BEP3 distribution relative to the bivariate Poisson (BP) distribution [14] as well as to the BEP1 and BEP2 distributions [2, 19], on two real-world datasets: healthcare demand

in Australia [11, 14] and football [17]. The model selection is performed using Akaike's information criterion (AIC) and Bayesian information criterion (BIC), where

$$AIC = 2k - 2\ell \qquad and \qquad BIC = kln(n) - 2\ell,$$

k is the number of parameters to be estimated and ℓ is the maximum of the log likelihood function. The model with the lowest AIC and BIC values is considered the best fit for the data.

The BP distribution fit is performed using the simple.bp EM R function from the `bivpois` R package [14]. For the BEP3 distribution, we developed and implemented an EM algorithm in R [21]. The parameters of the BEP1 and BEP2 distributions are estimated by direct maximum likelihood, without using the EM algorithm, as they do not involve latent variables. Conversely, the BP and BEP3 distributions, which are based on latent variables, are estimated by maximum likelihood using the EM algorithm.

5.1 Example 1

In accordance with [14], we reexamine the dataset on healthcare demand in Australia (see Table 1), where Y_1 denotes the Number of Doctor Consultations and Y_2 the Number of Prescribed Medications. These data are overdispersed and positively correlated; for more details, see [14]. This dataset is also available in the faraway R package, known as dvisits data. In Tables 2 and 3, we present the results of estimating the parameters of the BP(θ_1, θ_2, θ), BEP1(θ_1, θ_2, β_1, β_2, β), BEP2(θ_1, θ_2, β_1, β_2, β), and BEP3(θ_1, θ_2, θ, β_1, β_2, β) distributions, along with their goodness-of-fit on the dvisits data. In Table 3, the results indicate that the BEP3 distribution provides a substantially better fit than the BP, $BEP1$, and $BEP2$ distributions. In particular, the log-likelihood (ℓ) for the BEP3 distribution is considerably higher than that for the BP, $BEP1$, and $BEP2$ distributions. Moreover, both information criteria (AIC and BIC) are systematically lower for the BEP3 distribution, confirming its improved fit while accounting for its greater model complexity.

5.2 Example 2

In this example, we draw on the dataset analyzed by [17]: the scores from 26 soccer matches between ACF Fiorentina and Juventus played between 1996 and 2011. Table 4 reports these data, with Y_1 denoting the number of

Table 1: Cross-tabulation of dvisits data [11].

Y_1 \ Y_2	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	2	1	1	1	0	1	0	1	0
9	0	0	0	0	0	0	0	0	1

Table 2: Parameters estimation for BP, $BEP1$, $BEP2$ and $BEP3$ distributions with dvisits data.

Distributions	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}$
BP	0.176	0.737	0.126			
$BEP1$	0.196	0.6996		0.239	0.973	2.6
$BEP2$	0.173437	0.676073		0.207413	0.941704	0.028
$BEP3$	0.107	0.578	0.046	0.117	0.75	0.05

Table 3: Goodness-of-fit of the BP, $BEP1$, $BEP2$ and $BEP3$ distributions on dvisits data.

Distributions	ℓ	AIC	BIC
BP	-11268.36	22542.71	22564.46
$BEP1$	-10997.27	22004.53	22037.31
$BEP2$	-11300.92	22611.84	22644.61
$BEP3$	-10792.775	21585.55	21636.88

Table 4: Cross-tabulation of football data [17].

Y_1 \ Y_2	0	1	2	3
0	0	6	0	0
1	1	7	6	1
2	0	0	0	0
3	1	0	1	2

Table 5: Parameters estimation for BP, $BEP1$, $BEP2$ and $BEP3$ distributions with football data.

Distributions	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}$
BP	0.446	0.806	0.634			
$BEP1$	0.37	0.546		1.38	5.334	1.4
$BEP2$	0.367	0.188		1.349	4.131	0.755
$BEP3$	0.171	0.081	0.11	0.309	0.717	0.172

goals scored by ACF Fiorentina and Y_2 the number scored by Juventus. The same overdispersed dataset was also used by [12] to investigate new classes of bivariate discrete distributions.

Tables 5 and 6 report the parameter estimates for the BP$(\theta_1, \theta_2, \theta)$, BEP1$(\theta_1, \theta_2, \beta_1, \beta_2, \beta)$, BEP2$(\theta_1, \theta_2, \beta_1, \beta_2, \beta)$ and BEP3$(\theta_1, \theta_2, \theta, \beta_1, \beta_2, \beta)$ distributions, together with their goodness-of-fit metrics on the football data.

In Table 6, all variants of the BEP distributions improve the fit compared to the BP distribution. The BEP2 distribution offers the best overall fit with a higher log-likelihood and lower AIC/BIC values. The BEP3 distribution also outperforms the BP and BEP1 distributions with a higher log-likelihood and lower AIC/BIC values than those of the BP and BEP1 distributions.

Table 6: Goodness-of-fit of the BP, $BEP1$, $BEP2$ and $BEP3$ distributions on football data.

Distributions	ℓ	AIC	BIC
BP	-63.361	132.723	138.459
$BEP1$	-61.084	132.168	138.263
$BEP2$	-56.399	122.797	128.892
$BEP3$	-57.183	126.366	133.679

6 Conclusion

In this paper, we introduced a new bivariate distribution for count data, called the Bivariate Extended Poisson distribution of type 3 (BEP3 distribution), constructed using the trivariate reduction method. This approach makes it possible to model overdispersed or underdispersed and positively correlated count data. We studied several theoretical properties of the BEP3 distribution, including its joint probability mass function, marginal distributions, probability and moment generating function, moments, covariance and correlation coefficient. The model parameters were estimated using the maximum likelihood method via EM algorithm. The proposed distribution was applied to two reals-world datasets overdispersed: dvisits and football data. According to the AIC and BIC criteria, based on the dvisits data, the BEP3 distribution offers the best fit, ahead of the BP, BEP1 and BEP2 distributions. For football data, the BEP2 distribution is the best fit, while the BEP3 distribution remains superior to the BP and BEP1 distributions. These results confirm the flexibility and relevance of the BEP3 distribution for analyzing bivariate, overdispersed and positively correlated count data.

Acknowledgment. The authors would like to thank the editorial team and the reviewers for their comments and suggestions.

References

[1] P. Berkhout, E. Plug, A bivariate Poisson count data model using conditional probabilities, Statist Neerlandica, **58**, no. 3, (2004), 349–364.

[2] R. Bidounga, M. Koukouatikissa Diafouka, R.F. Mizélé Kitoti, D. Mizère, The Bivariate Extended Poisson Distribution of Type 1, European Journal of Pure and Applied Mathematics, **14**, no. 4, (2021), 1517–1529.

[3] C. A. Charalambides, Enumerative combinatorics. Chapman and Hall/CRC, 2018.

[4] 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**, no. 1, (1977), 1–38.

[5] M. Koukouatikissa Diafouka, C. G. Louzayadio, On a bivariate katz's distribution constructed by the trivariate reduction method, Serdica Journal of Computing, **17**, no. 2, (2023), 79–93.

[6] B. Dimitrov, N. Kolev, Beta transformation. beta type self-decomposition and related characterizations, Brezilian Journal of Probability and Statistics, **14**, (2000), 123-140.

[7] L. L. Elion, M. Koukouatikissa Diafouka, R. Bidounga, C. G. Louzayadio, D. Mizère, R. A. Makany, G. Kissita, The bivariate weighted Poisson distribution: statistical study of the arterial hypertention data according to the weight and blood sugar level, Far East Journal of Theoretical Statistics, **52**, no. 5, (2016), 365–393.

[8] F. Famoye, A new bivariate generalized Poisson distribution, Statistica Neerlandica, **64**, no. 1, (2017), 112–124.

[9] F. Famoye, P. C. Consul, Bivariate generalized Poisson distribution with some applications, Metrika, **42**, (1995), 127–138.

[10] R. L. Graham, D. E. Knuth, O. Patashnik, Concrete Mathematics, Addison-Wesley, second ed., 1994.

[11] P. Holgate, Estimation for the bivariate Poisson distribution, Biometrika, **51**, (1964), 241–245.

[12] X. Jiang, J. Chu, S. Nadarajah, New classes of discrete bivariate distributions with application to football data, Communications in statistics-theory and methods, **46**, no. 16, (2017), 8069–8085.

[13] N. L. Johnson, S. Kotz, N. Balakrishnan, Discrete Multivariate Distributions, John Wiley & Sons, New York, 1997.

[14] D. Karlis, I. Ntzoufras, Analysis of Sports Data by Using Bivariate Poisson Models, Journal of the Royal Statistical Society D (The Statistician), **52**, no. 3, (2003), 381–393.

[15] S. Kocherlakota, K. Kocherlakota, Discrete Bivariate Distributions, Marcel Dekker, New York, 1992.

[16] J. Lakshminarayana, S.N. Pandit, K. Srinivasa Rao, On a bivariate Poisson distribution, Communications in Statistics-Theory and Methods, **28**, no. 2, (1999), 267–276.

[17] H. Lee, J. H. Cha, On two general classes of discrete bivariate distributions. The American Statistician, **69**, (2015). 221-230.

[18] C. G. Louzayadio, R. O. Malouata, M. Koukouatikissa Diafouka, A weighted Poisson distribution for underdispersed count data. International Journal of Statistics and probability, **10**, no. 4, (2021), 157–165.

[19] C. G. Louzayadio, E. Nguessolta, M. Koukouatikissa Diafouka, R. Bidounga, D. Mizère, The bivariate extended Poisson distribution of type 2, Journal of Computer Science and Applied Mathematics, **5**, no. 2, (2023), 89–102.

[20] H. Papageorgiou, M. D. Katerina, On a class of bivariate compounded Poisson distributions, Statistics & probability letters **23**, no. 1, (1995), 93–104.

[21] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2025.

[22] Y. Simsek, Formulas for p-adic q-integrals including falling-rising factorials, combinatorial sums and special numbers. Revista de la Real Academia de Ciencias Exactas, Físicas y Naturales. Serie A. Matemáticas, Springer, **118**, no. 3, (2024), 92.

[23] K. S. Weems, K. F. Sellers, T. Li, A flexible bivariate distribution for count data expressing data dispersion, Communications in Statistics-Theory and Methods, **52**, no. 13, (2023), 4692–4718.