Applications and Applied Mathematics: An International Journal (AAM)

Volume 18 | Issue 1 | Article 6

12-2022

(R2024) A New Weighted Poisson Distribution For Over- and Under-Dispersion Situations

Michel Koukouatikissa Diafouka
Marien Ngouabi University

Gelin Chedly Louzayadio
Marien Ngouabi University

Rodnellin Onéime Malouata
Marien Ngouabi University

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Applications and Applied
Mathematics:
An International Journal
(AAM)

Vol. 18, Issue 1 (June 2023), Article 6, 15 pages

A New Weighted Poisson Distribution For Over- and Under- Dispersion Situations

[1*]Michel Koukouatikissa Diafouka, [2]Chedly Gélin Louzayadio and [3]Rodnellin Onéime Malouata

[1]Higher Teachers Training College
Marien Ngouabi University
Avenue of the First African Games
Brazzaville, Congo
michel.koukouatikissa@umng.cg

[2]Faculty of Economics Sciences
Marien Ngouabi University
Avenue of the First African Games
Brazzaville, Congo
gelinlouz@gmail.com

[3]Higher Institute of Management
Marien Ngouabi University
Avenue of the First African Games
Brazzaville, Congo
onesimero@gmail.com

*Corresponding Author

Received: August 15, 2022; Accepted: April 17, 2023

Abstract

In this paper, we propose a four-parameter weighted Poisson distribution that includes and generalizes the weighted Poisson distribution proposed by Castillo and Pérez-Casany and the Conway-Maxwell-Poisson distribution, as well as other well-known distributions. It is a distribution that is a member of the exponential family and is an exponential combination formulation between the weighted Poisson distribution proposed by Castillo and Pérez-Casany and the Conway-Maxwell-Poisson distribution. This new distribution with an additional parameter of dispersion is more flexible, and the Fisher dispersion index can be greater than, equal to, or less than one. This last property allows it to model over-dispersed data as well as under-dispersed or equi-dispersed data. Many other properties of the new distribution are studied in this article. The parameters are estimated by two methods: the least squares and the maximum likelihood methods. The properties of the estimators are not studied because they follow directly. Two examples of application to real data, taking into account situations of overdispersion and underdispersion, are examined.

Keywords: Count data; Over- and under-dispersion; Exponential family; Conway-Maxwell-Poisson distribution; Weighted Poisson distribution; Estimation

MSC 2010 No.: 62E10, 62E15, 62H10

1. Introduction

The concept of weighted distribution was introduced by Rao (1965) as one of the alternatives to the Poisson distribution for the modeling and statistical analysis of count data. Indeed, the Poisson distribution is still not appropriate for modeling count data although it is a benchmark model. The lack of adequacy of the Poisson distribution is due to the variation observed in the data because of the variance of the sample which can be larger or smaller than the mean. These variation phenomena are called overdispersion and underdispersion, respectively, relative to the Poisson distribution. Thus, one of the most important questions in the modeling and analysis of count data is how to formulate an adequate probabilistic model to describe these observed variations in the data.

Since then, the family of weighted distributions (weighted Poisson distributions in particular) has been widely and variously studied in the literature; we can cite Patil et al. (1986), Patil (2002), Chakraborty and Imoto (2016), Louzayadio et al. (2021) and their references. In particular, Conway and Maxwell (1962) have formulated a family of generalized Poisson distributions of discrete probabilities with two parameters, called the Conway-Maxwell-Poisson distribution (noted COM-Poisson or CMP distribution). It is a weighted Poisson distribution (Gupta et al. (2014)) which has been used as a basis to develop several general families of distributions, and Castillo and Pérez-Casany (1998) proposed a family of weighted Poisson distributions with three parameters, noted WPD, which includes and generalizes the family of Poisson distributions.

On this, in this paper we propose a new family of weighted Poisson distributions with four parameters which includes and generalizes the WPD proposed by Castillo and Pérez-Casany (1998) and COM-Poisson distribution (Conway and Maxwell (1962)). The new family of weighted Poisson distributions is a new exponential family (cf Section 2.2) which also generalizes other well known distributions (cf Section 2.4). If CMP and WPD have only one dispersion parameter each, the new family of weighted Poisson distributions has two (cf Section 2.5). Its Fisher dispersion index can be greater than, equal to or smaller than one; it allows it to describe both over- and under-dispersed data. This new family of weighted Poisson distributions is an exponential combination between the WPD proposed by (Castillo and Pérez-Casany (1998)) and the CMP distribution (Conway and Maxwell (1962)) (cf Section 2.7). All these properties and the mass function are presented in Section 2. In Section 3, we determine the estimators of the parameters. The properties of these estimators will not be studied because the methods used are usual and the asymptotic properties of the estimators derived from these methods follow naturally. In Section 4, we fit the proposed new distribution to real data. The conclusion and perspectives are presented in Section 5.

2. New Weighted Poisson Distribution

In this section, the mass function of the new distribution is presented and studied from a probabilistic and statistical point of view.

2.1. Point probability functions and recursive ratio probabilities

Let X be a positive discrete random variable, we propose a new discrete distribution whose probability mass function (in short pmf) $p(x) = P(X = x)$ is defined by

$$p(x) = \frac{\lambda^x (x+a)^\alpha}{x!^\nu} \frac{1}{C(\lambda, \alpha, \nu, a)}, \quad x \in \mathbb{N}, \tag{1}$$

where

$$C(\lambda, \alpha, \nu, a) = \sum_{j \geq 0} \frac{\lambda^j (j+a)^\alpha}{j!^\nu}. \tag{2}$$

By the quotient criterion, the series (2) converges for all

- $\lambda > 0,\ \alpha \in \mathbb{R},\ \nu > 0,\ a > 0$, or
- $\lambda > 0,\ \alpha \in \mathbb{R}_+,\ \nu > 0,\ a = 0$, or
- $0 < \lambda < 1,\ \alpha \in \mathbb{R},\ \nu = 0,\ a > 0$ ($a = 0$ for $x \in \mathbb{N}^*,$ see Section 2.4).

The recursive ratio probabilities corresponding to pmf defined in (1) is:

$$\frac{p(x+1)}{p(x)} = \frac{\lambda}{(x+1)^\nu} \left(1 + \frac{1}{x+a}\right)^\alpha. \tag{3}$$

As in Castillo and Pérez-Casany (1998), note that the second factor in the right hand of (3) tends to one, for every $a \in \mathbb{R}_+^*$ and $\alpha \in \mathbb{R}$, when x tend to infinity. Then, the tails of the new distribution with parameter λ are similar to the tails of COM-Poisson distribution with parameter λ (see Shmueli and al. (2005)). While, values of ν that are less than 1 correspond to flatter recursive ratio probabilities than the WPD proposed by Castillo and Pérez-Casany (1998) and hence to longer tails.

2.2. Weighted Poisson distribution and exponential family

The pmf given by (1) can be written as follows:

$$p(x) = \frac{(x+a)^\alpha x!^{1-\nu}}{e^{-\lambda} C(\lambda, \alpha, \nu, a)} \frac{\lambda^x}{x!} e^{-\lambda}, \quad x \in \mathbb{N}. \tag{4}$$

Relationship (4) shows that the new distribution as defined is a four-parameter weighted Poisson distribution (Rao (1965); Patil (2002)), of weight function $w(x) = (x+a)^\alpha x!^{1-\nu}$ and normalization constant $E_\lambda\left[(X+a)^\alpha X!^{1-\nu}\right] = e^{-\lambda} C(\lambda, \alpha, \nu, a)$. We denote this new distribution by NWPD $(\lambda, \alpha, \nu, a)$.

The pmf given by (1) can also expressed as

$$p(x) = \exp[x \log \lambda + \alpha \log(x+a) - \nu \log(x!) - \log C(\lambda, \alpha, \nu, a)], \quad x \in \mathbb{N}. \tag{5}$$

And from (5), NWPD $(\lambda, \alpha, \nu, a)$ is, for a fixed, a member of the exponential family of parameters $(\log \lambda, \alpha, \nu)$ with $\log \lambda$ the canonical parameter. Moreover, the parameters $(\lambda, \alpha, \nu, a)$ of NWPD have the same interpretation as the parameters (λ, ν, a) of WPD, except that NWPD has two parameters of dispersion, namely α and ν (see Section 2.5).

2.3. Moments of distribution

Consider a random variable X following the NWPD $(\lambda, \alpha, \nu, a)$. The moment of order r of the variable $X + a$ is given by

$$E[(X+a)^r] = \frac{C(\lambda, \alpha + s, \nu, a)}{C(\lambda, \alpha, \nu, a)}, \quad r \in \mathbb{N}. \tag{6}$$

Thus, we deduce the mathematical expectation and the variance from (6),

$$E(X) = \frac{C(\lambda, \alpha + 1, \nu, a) - aC(\lambda, \alpha, \nu, a)}{C(\lambda, \alpha, \nu, a)},$$

$$Var(X) = \frac{C(\lambda, \alpha + 2, \nu, a)C(\lambda, \alpha, \nu, a) - C^2(\lambda, \alpha + 1, \nu, a)}{C^2(\lambda, \alpha, \nu, a)}.$$

2.4. Some distributions derived of NWPD

The NWPD is a generalization of some well-known discrete distributions, and the distributions derived from them. In particular, we have:

- if $\alpha = 0$ and $\nu = 1$, NWPD genered an ordinary Poisson distribution with parameter λ.
- if $\alpha = 0$ and $\nu = 0$, NWPD is a geometric distribution with zero probability $1 - \lambda$ where $0 < \lambda < 1$.
- if $\alpha = 0$, NWPD is the COM-Poisson model of parameters λ and ν (Conway and Maxwell (1962)).
- if $\nu = 0$ and $a = 0$, NWPD is a new discrete distribution of support $\mathbb{N}^*$ proposed by Kulasekera and Tonkyn (1992).
- if $\nu = 1$, NWPD is WPD (λ, α, a) proposed by Castillo and Pérez-Casany (1998).

2.5. Characterization : overdispersion and underdispersion

The study of the dispersion of NWPD $(\lambda, \alpha, \nu, a)$ follows from the Theorem 4 of Kokonendji et al. (2008). The second derivative of the logarithm of the weight function $x \longmapsto w(x)$ of NWPD $(\lambda, \alpha, \nu, a)$ is given by

$$\frac{d^2}{dx^2} \log w(x) = \frac{-\alpha}{(x+a)^2} + (1-\nu) \sum_{k \geq 1} \frac{1}{(x+k)^2}.$$

This means that, NWPD $(\lambda, \alpha, \nu, a)$ is

- overdispersed if one of the following two conditions apply:
  - $\alpha < 0$ and $\nu \leq 1$ for all $a > 0$,
  - $\alpha < 0$ and $\nu > 1$ when $0 < \nu - 1 < -\alpha$ for all $a > 0$.
- underdispersed if one of the following three conditions apply:
  - $\alpha > 0$ and $\nu \geq 1$ for all $a > 0$,
  - $\alpha > 0$ and $\nu < 1$ when $-\alpha < \nu - 1 < 0$ or $\nu - 1 < -\alpha < 0$ for all $a > 0$,
  - $\alpha < 0$ and $\nu > 1$ when $0 < -\alpha < \nu - 1$ for all $a > 0$.

2.6. Sufficient statistics

The likelihood of NWPD $(\lambda, \alpha, \nu, a)$ for a set of n independent and identically distributed observations $x_1, x_2, \ldots, x_n$ is

$$L(x_1, \ldots, x_n | \lambda, \alpha, \nu) = \lambda^{t_1} \exp\left[\alpha t_2 - \nu t_3\right] C^{-n}(\lambda, \alpha, \nu, a),$$

where $t_1 = \sum_{i=1}^{n} x_i$, $t_2 = \sum_{i=1}^{n} \log(x_i + a)$ and $t_3 = \sum_{i=1}^{n} \log(x_i!)$. By the Factorization Theorem, for a fixed, (t_1, t_2, t_3) are sufficient statistics for $x_1, \ldots, x_n$ (see Conway and Maxwell (1962); Meena and Gangopadhyay (2020)).

2.7. Approximation of the normalizing constant $C(\lambda, \alpha, \nu, a)$ using truncation of the series

$C(\lambda, \alpha, \nu, a)$ being the sum of a convergent series, then there exists a natural number k such that

$$C(\lambda, \alpha, \nu, a) = \sum_{j=0}^{k} \frac{\lambda^j (j+a)^\alpha}{j!^\nu} + \mathcal{R}_k,$$

where $\mathcal{R}_k = \sum_{j \geq k+1} \frac{\lambda^j (j+a)^\alpha}{j!^\nu}$ is the absolute truncation error. Thus, there exist $0 < \varepsilon_k < 1$ for all $j > k$ so that

$$\mathcal{R}_k < \frac{\lambda^{k+1}(k+a+1)^\alpha}{(k+1)!^\nu} \sum_{j \geq 0} \varepsilon_k^j = \frac{\lambda^{k+1}(k+a+1)^\alpha}{(k+1)!^\nu (1-\varepsilon_k)}.$$

The relative truncation error is

$$\mathcal{R}_k \Big/ \sum_{j=0}^{k} \frac{\lambda^j (j+a)^\alpha}{j!^\nu},$$

and can be bounded by

$$\frac{\lambda^{k+1}(k+a+1)^{\alpha}}{(k+1)!^{\nu}(1-\varepsilon_k)} \frac{1}{\sum_{j=0}^{k} \frac{\lambda^j (j+a)^{\alpha}}{j!^{\nu}}}.$$

Using this approximation, we have represented pmf for different values of the parameters in Figure 1

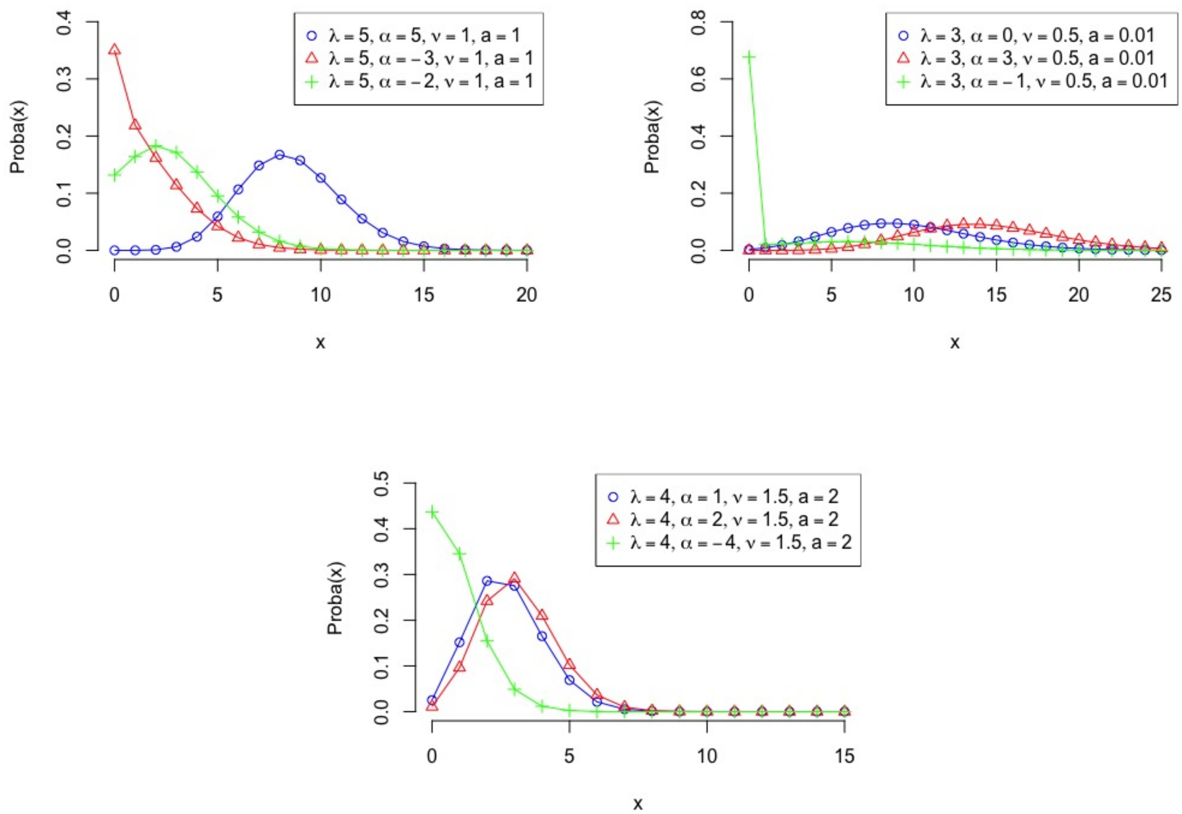

Figure 1. pmfs for different values of parameters

2.8. Other properties

2.8.1. *NWPD as exponential combination formulation*

From the exponential combination formulation given by Chakraborty and Imoto (2016), we have the following result.

Proposition 2.1.

The NWPD $(\lambda, \alpha, \nu, a)$ is an exponential combination WPD (μ, r, a) proposed by Castillo and

Pérez-Casany (1998) and CMP (θ,γ) distribution distributions, with $\lambda=\theta^{\beta}\mu^{1-\beta},\ \alpha=(1-\beta)r$ and $\nu=(1-\beta)\gamma$.

Proof:

The pfm's CMP (θ,γ) distribution and WPD (μ,r,a) are given respectively by (Shmueli and al. (2005), Castillo and Pérez-Casany (1998))

$$p_2(x,\theta,\gamma)=\frac{\theta^x}{x!^{\gamma}Z(\theta,\gamma)},$$

and

$$p_1(x,\mu,r,a)=\frac{(x+a)^r\mu^xe^{-\mu}}{E_\lambda[(X+a)^r]x!},$$

where $Z(\theta,\gamma)$ and $E_\lambda[(X+a)^r]$ are the normalizing constants. So, the probability function resulting from the exponential combination of this two distributions is given by

$$[p_1(x,\theta,\gamma)]^{\beta}[p_2(x,\mu,r,a)]^{1-\beta}=\frac{\left(\theta^{\beta}\mu^{1-\beta}\right)^x(x+a)^{(1-\beta)r}}{x!^{(1-\beta)\gamma}}.$$

Substituting $\theta^{\beta}\mu^{1-\beta}$ by λ, $(1-\beta)r$ by α and $(1-\beta)\gamma$ by ν, we have the pmf of NWPD (λ,α,ν,a). ∎

2.8.2. *Failure rate function and log-concavity*

The failure rate function $r(t)$ of NWPD (λ,α,ν,a) is given by

$$\begin{aligned}\frac{1}{r(t)}&=\frac{P(X\geq t)}{P(X=t)}\\&=1+\sum_{i\geq 0}\prod_{u=t}^{t+i}\left[\frac{\lambda}{(x+1)^{\nu}}\left(1+\frac{1}{x+a}\right)^{\alpha}\right],\\&={}_{\alpha+1}F_{\nu}\left(1,1+\frac{1}{t+a},1+\frac{1}{t+a},\ldots,1+\frac{1}{t+a};t+1,t+1,\ldots,t+1;\lambda\right),\end{aligned}$$

where α, ν are positive integers in expression of hypergeometric function (Gupta et al. (1997); Qureshi and Shadab (2020)).

For the log-concavity of NWPD, we have the following result.

Proposition 2.2.

NWPD (λ,α,ν,a) has a

- log-concave pmf if

- $\alpha \geq 0,\ \nu > 0,\ \forall\, a > 0$, or
- $\alpha > 0,\ \nu = 0\ \forall\, a > 0$, or
- $0 < -\alpha < \nu\ \forall\, a > -1 + \sqrt{2}$.

- log-convex pmf for $\alpha < 0,\ \nu = 0,\ \forall\, a > 0$.

Proof:

Let put $\eta(t) = 1 - \dfrac{p(t+1)}{p(t)}$ and $\Delta\eta(t) = \dfrac{p(t+1)}{p(t)} - \dfrac{p(t+2)}{p(t+1)}$. A distribution is said log-concave if $\Delta\eta(t) > 0$ and log-convex if $\Delta\eta(t) < 0$ (see Gupta et al. (1997)).

We have for NWPD $(\lambda, \alpha, \nu, a)$,

$$\Delta\eta(t) = \frac{\lambda}{(t+1)^\nu}\left(\frac{t+a+1}{t+1}\right)^\alpha \left[1 - \left(\frac{t+1}{t+2}\right)^\nu \left[\frac{(t+a+1)^2-1}{(t+a+1)^2}\right]^\alpha\right].$$

The sign of $\Delta\eta(t)$ depends only on $1 - \left(\dfrac{t+1}{t+2}\right)^\nu \left[\dfrac{(t+a+1)^2-1}{(t+a+1)^2}\right]^\alpha$.

For log-concavity, we have

- if $\alpha \geq 0$ and $\nu > 0$, $\left(\dfrac{t+1}{t+2}\right)^\nu \left[\dfrac{(t+a+1)^2-1}{(t+a+1)^2}\right]^\alpha < 1\ \forall\, a > 0$;
- if $\alpha > 0$ and $\nu = 0$, $\left[\dfrac{(t+a+1)^2-1}{(t+a+1)^2}\right]^\alpha < 1\ \forall\, a > 0$,
- for all $t \geq 0$, $(t+a)(t+a+2) > t+1\ \forall\, a > -1+\sqrt{2}$; that imply $1 + \dfrac{1}{(t+a)(t+a+2)} < 1 + \dfrac{1}{t+1}\ \forall\, a > -1+\sqrt{2}$, i.e., $1 < \dfrac{(t+a+1)^2}{(t+a+1)^2-1} < \dfrac{t+2}{t+1}\ \forall\, a > -1+\sqrt{2}$. Taking the logarithm, we have $0 < \log\left[\dfrac{(t+a+1)^2}{(t+a+1)^2-1}\right] < \log\left[\dfrac{t+2}{t+1}\right]\ \forall\, a > -1+\sqrt{2}$. Finally, if $0 < -\alpha < \nu$ we have $-\alpha\log\left[\dfrac{(t+a+1)^2}{(t+a+1)^2-1}\right] < \nu\log\left[\dfrac{t+2}{t+1}\right]\ \forall\, a > -1+\sqrt{2}$, i.e., $\left(\dfrac{t+1}{t+2}\right)^\nu \left[\dfrac{(t+a+1)^2-1}{(t+a+1)^2}\right]^\alpha < 1\ \forall\, a > -1+\sqrt{2}$.

These three cases imply $\Delta\eta(t) > 0$.

And for log-convex, if $\nu = 0$ and $\alpha < 0$, we have $\left[\dfrac{(t+a+1)^2-1}{(t+a+1)^2}\right]^\alpha > 1$. Thus $\Delta\eta(t) < 0$.

Note for $\alpha = 0$ and $\nu = 0$, $\Delta\eta(t) = 0$. This case corresponding to geometric distribution (see Section 2.4). ■

The WPD (λ, α, a) and CMP (λ, ν) distributions can be obtained as particular cases from the above results. WPD (λ, α, a) have log-concave pmf if $\alpha \geq 0\ \forall\, a > 0$ or $0 < -\alpha < 1\ \forall\, a > -1+\sqrt{2}$

and CMP (λ, ν) log-concave pmf for $\nu > 0$.

The following are important results as a consequence of log-concavity (Gupta et al. (1997)).

Corollary 2.1.

If $\alpha \geq 0,\ \nu > 0,\ \forall\, a > 0$ or $\alpha > 0,\ \nu = 0,\ \forall\, a > 0$ or $0 < -\alpha < \nu,\ \forall\, a > -1 + \sqrt{2}$, the NWPD $(\lambda, \alpha, \nu, a)$ has

- an increasing failure rate function,
- a strongly unimodal distribution (Kim and al. (2018); Weller and Martin (2020)).

And the following is important result as a consequence of log-convexity (Gupta et al. (1997)).

Corollary 2.2.

If $\alpha < 0,\ \nu = 0,\ a > 0$, the NWPD $(\lambda, \alpha, \nu, a)$ has a decreasing failure rate function.

Remark 2.1.

Note when NWPD $(\lambda, \alpha, \nu, a)$ generate the geometric distribution for $\alpha = 0$ and $\nu = 0$, we have a constant failure rate (see Gupta et al. (1997)).

2.8.3. *Stochastic order relations and reliability characteristics*

We are interested in three types of stochastic orders: the likelihood ratio order, the hazard rate order, and the mean residual life order. Consider two discrete random variables X and Y, the notation $X \leq_{lr} Y$ ($X \leq_{hr} Y$ or $X \leq_{MRL} Y$) means X is smaller than Y in the likelihood ratio (hazard rate or mean residual life, respectively) order. Note that the likelihood ratio order implies the hazard rate order and subsequently the mean residual life order (Raeisi and Yari (2019); Huang and Mi (2018)).

Proposition 2.3.

Let be X WPD (λ, α, a), Y CMP (λ, ν) and Z NWPD $(\lambda, \alpha, \nu, a)$ random variables.

- $Z \leq_{lr} X$ if $\nu > 1$.
- $Z \leq_{lr} Y$ when $\alpha < 0$.

Proof:

If X has WPD (λ, α, a) as pmf and Z WPD $(\lambda, \alpha, \nu, a)$, then

$$\frac{P(X = k)}{P(Z = k)} = k!^{\nu - 1} \frac{C(\lambda, \alpha, \nu, a)}{C(\lambda, \alpha, a)},$$

is increasing in k if $\nu > 1$.

Similarly, if Y has CMP (λ, ν) as pmf, then

$$\frac{P(Y=k)}{P(Z=k)} = (k+a)^{-\alpha}\frac{C(\lambda,\alpha,\nu,a)}{Z(\lambda,\nu)},$$

is increasing in k if $\alpha < 0$.

Hence, the result is proved. ■

Using the above results, it follows that (Perrakis (2019); Shaked and Shanthikumar (2007)),

Corollary 2.3.

- $Z \leq_{hr} X$ if $\nu > 1$ and subsequently $Z \leq_{MRL} X$ if $\nu > 1$.
- $Z \leq_{hr} Y$ if $\alpha < 0$ and subsequently $Z \leq_{MRL} Y$ if $\alpha < 0$.

3. Estimation of Parameters

In this section, we are interested in the estimation of the parameters of NWPD $(\lambda, \alpha, \nu, a)$. Two estimation methods are presented. The first method is simple and is based on the ratio recursive probabilities using the method of least squares to estimate the parameters (see Kulasekera and Tonkyn (1992); Castillo and Pérez-Casany (1998)). The second method is the maximum likelihood method. Just three parameters (λ, ν, α) are estimated; the parameter a is fixed as in Castillo and Pérez-Casany (1998).

3.1. Alternative method of estimation: least squares method

Consider the recursive ratio probabilities (3) and replace $x+1$ and x by $x+2$ and $x+1$, respectively. We have,

$$\frac{p(x+2)}{p(x+1)} = \frac{\lambda}{(x+2)^{\nu}}\left(1+\frac{1}{x+a+1}\right)^{\alpha}, \tag{7}$$

and dividing (7) by (3), we have,

$$\frac{p(x+2)p(x)}{p^2(x+1)} = \left(\frac{x+1}{x+2}\right)^{\nu}\left[\frac{(x+a+2)(x+a)}{(x+a+1)^2}\right]^{\alpha}. \tag{8}$$

So, we can get a linear relation, for a fixed, between the two parameters α and ν taking the logarithms of (8) and then dividing by $\log\left(\frac{x+1}{x+2}\right)$ as follows,

$$y = \nu + \alpha x_a, \tag{9}$$

where

$$y = \log\left[\frac{p(x+2)p(x)}{p^2(x+1)}\right] / \log\left(\frac{x+1}{x+2}\right),$$

and

$$x_a = \log\left[\frac{(x+a+2)(x+a)}{(x+a+1)^2}\right] / \log\left(\frac{x+1}{x+2}\right).$$

For $x_1, \dots, x_k$ a given sample, $p(x_i),\ i = 1 \dots k$ is estimated and replaced by proportion observed, i.e., by $f_i = \frac{n_i}{n},\ i = 1 \dots k$ where n_i is the frequency observed for individual i and n the size of the sample (see Shmueli and al. (2005); Lakshmi et al. (2021)). Once the parameters α and ν are estimated by the regression of (9), we can obtain the estimate of the parameter λ by the equation (7).

3.2. Maximum likelihood estimation

Let us denote $\theta = (\log\lambda, \alpha, \nu)$ the vector of the parameters of interest. The log-likelihood of $NWPD_a(\theta)$ is given by

$$l(\theta) = n\overline{x}\log\lambda + n\alpha\overline{\log(x+a)} - n\nu\overline{\log(x!)} - n\log C(\lambda, \alpha, \nu, a),$$

where $\overline{x} = \frac{t_1}{n}, \quad \overline{\log(x+a)} = \frac{t_2}{n}, \quad and \quad \overline{\log(x!)} = \frac{t_3}{n}$. And the maximum likelihood estimator $\widehat{\theta}$ of θ is solution of the equations system

$$\left.\begin{aligned} \frac{\partial}{\partial\log\lambda}\log C(\lambda, \alpha, \nu, a) = E(X) = \overline{x}, \\ \frac{\partial}{\partial\alpha}\log C(\lambda, \alpha, \nu, a) = E[\log(X+a)] = \overline{\log(x+a)}, \\ -\frac{\partial}{\partial\nu}\log C(\lambda, \alpha, \nu, a) = E[\log(X!)] = \overline{\log(x!)}. \end{aligned}\right\}. \tag{10}$$

The system (10) cannot be solved analytically; an alternative method such as Newton-Raphson can be use (see Gelman et al. (1995), pages 312-313). Following Minka et al. (2003), the gradient of the log-likelihood is,

$$\nabla l(\theta) = n\begin{bmatrix} \overline{x} - E(X) \\ \overline{\log(x+a)} - E[\log(X+a)] \\ -\overline{\log(x!)} + E[\log(X!)] \end{bmatrix},$$

and the seconde matrix is,

$$\nabla^2 l(\theta) = n \begin{bmatrix} -Var(X) & -Cov(X, \log(X+a)) & Cov(X, \log(X!)) \\ -Cov(X, \log(X+a)) & -Var[\log(X+a)] & Cov(\log(X+a), \log(X!)) \\ Cov(X, \log(X!)) & Cov(\log(X+a), \log(X!)) & -Var[\log(X!)] \end{bmatrix}.$$

So, the Newton's method update is,

$$\theta^{new} = \theta - (\nabla^2 l(\theta))^{-1} \nabla l(\theta).$$

A reasonable starting point for the iteration is the ordinary Poisson MLE, $(\lambda = \overline{x}, \alpha = 0, \nu = 1)$, i.e., $\theta = (\log(\overline{x}), 0, 1)$ (Minka et al. (2003)) for a fixed. And using the package maxLik for the R statistical environment, we obtain the estimator $\widehat{\theta}$ of θ easly (see Henningsen and Toomet (2011)).

4. Fitting the New Weighted Poisson Distribution: Application Examples

In this section, we give two examples of fitting practical data by the NWPD $(\lambda, \alpha, \nu, a)$. In order to make the comparison between WPD (λ, ν, a) and NWPD $(\lambda, \alpha, \nu, a)$, we consider the same data as in Castillo and Pérez-Casany (1998) which allowed them to introduce WPD (λ, ν, a). Thus, we fixed the same values of a as WPD (λ, ν, a) to estimate the parameter $\theta = (\log(\lambda), \alpha, \nu)$ of NWPD $(\lambda, \alpha, \nu, a)$. In the Example 1 the data provide of Greenwood and Yule (1920) and are overdispersed; in the Example 2 the date provide of Kendall (1961) and are underdispersed. These data were used by X to fit the W distribution, and we had added the new distribution to these tables. The differences between the observed and expected values are calculated by Pearson's χ^2 test.

Example 4.1.

Greenwood and Yule (1920) used the data set of the Table 1 to introduce the negative binomial distribution, a distribution that is considered the prototype of overdispersed distributions. The data show the distribution of the number of accidents among 647 machine operators in a fixed period of time. The basic statistics are

$$\overline{x} = 0.46522 \quad s^2 = 0.6919,$$

and the ratio of the sample variance to the sample mean is greater than 1.

Table 1 shows that both WPD and NWPD fit the data and have to within 10^{-1} the same values of χ^2 and p-value.

Example 4.2.

The statistical data, Table 2, are taken from Kendall (1961) and correspond to the observed data on the number of outbreaks of strikes in 4-week periods, in a coal mining industry in the United Kingdom during 1948-1959. The basic statistics are

$$\overline{x} = 0.99359 \quad s^2 = 0.741894,$$

and the ratio of the sample variance to the sample mean is small than 1.

Table 2 shows that for the same two values of a, 0.5 and 0.00001, NWPD fits slightly better than WPD.

Table 1. Number of accidents for machine operators

No. of accidents	Obs	WPD $a = 0.8$ $\widehat{\lambda} = 2.16572$ $\widehat{r} = -2.47553$	NWPD $a = 0.8$ $\widehat{\lambda} = 3.00571421$ $\widehat{\alpha} = -2.888255$ $\widehat{\nu} = 1.183231$	WPD $a = 1$ $\widehat{\lambda} = 2.46942$ $\widehat{r} = -3.07412$	NWPD $a = 1$ $\widehat{\lambda} = 6.27464266$ $\widehat{\alpha} = -4.442394$ $\widehat{\nu} = 1.486904$
0	447	447.158	447.077	446.874	447.724
1	132	130.084	129.400	131.032	129.400
2	42	47.182	47.878	46.517	47.878
3	21	15.993	16.175	15.813	16.175
4	3	4.856	4.529	4.916	4.529
5	2	1.725	1.941	1.848	1.294
n	467				
χ^2		2.917	2.957	2.906	2.926
$p - value$		0.712	0.706	0.714	0.711

Table 2. Number of outbreaks strikes

No. of accidents	Obs	WPD $a = 0.5$ $\widehat{\lambda} = 0.454975$ $\widehat{r} = 1.12$	NWPD $a = 0.5$ $\widehat{\lambda} = 0.08983194$ $\widehat{\alpha} = 2.62968$ $\widehat{\nu} = 0.08526064$	WPD $a = 0.00001$ $\widehat{\lambda} = 0.70107$ $\widehat{r} = 0.07268$	NWPD $a = 0.00001$ $\widehat{\lambda} = 0.5675733$ $\widehat{\alpha} = 0.09180$ $\widehat{\nu} = 0.77445$
0	46	46.607	46.020	46.030	46.020
1	76	72.572	74.568	74.514	75.192
2	24	29.252	27.300	27.469	26.520
3	9	6.466	6.552	6.611	6.708
≥ 4	1	1.103	1.560	1.374	1.560
n	156				
χ^2		2.115	1.555	1.432	1.279
$p - value$		0.714	0.816	0.838	0.864

5. Conclusion

In this paper, we proposed a new weighted Poisson distribution for modeling and statistical analysis of count data. It is a four-parameter distribution that unifies the weighted Poisson distribution proposed by Castillo and Pérez-Casany and the Conway-Maxwell Poisson distribution. This new distribution is an exponential combination of the latter two distributions. It is a distribution with interesting properties. From a practical point of view, its Fisher dispersion index, which can be greater than, equal to, or less than one, allows it to describe both over- and under-dispersed data. This property makes it as robust and competitive as the extended Conway-Maxwell-Poisson family of four-parameter distributions proposed by Chakraborty and Imoto in 2016.

In addition, we plan to approximate the normalization constant $C(\lambda, \alpha, \nu, a)$ using the Laplace method and to carry out a comparative study with another four-parameter extended Conway-Maxwell-Poisson distribution in order to explore other performances of the new distribution.

Acknowledgment:

We would like to thank Professors Marta Pérez-Casany and Joan del Castillo Franquet for permission to use the data tables and the Editor-in-Chief, Professor Aliakbar Montazer Haghighi for useful comments and suggestions to improve the paper.

REFERENCES

Castillo, J.D. and Pérez-Casany, M. (1998). Weighted Poisson distributions for overdispersion and underdispersion situations, Ann. Inst. Statist. Math., Vol. 50, No. 3, pp. 567-585.

Chakraborty, S. and Imoto, T. (2016). Extended Conway-Maxwell-Poisson distribution and its properties and applications, Journal of Statistical Distributions and Applications, Vol. 3, No. 5, pp. 1–19. doi:10.1186/s40488-016-0044-1

Conway, R. and Maxwell, W. (1962). A queuing model with state dependent service rates, Journal of Industrial Engineering, Vol. 12, pp.132–136.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B. (1995). *Bayesian Data Analysis*, Third Edition, Chapman and Hall Book.

Greenwood, M. and Yule, U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents, J. Roy Statist. Soc. Ser. A, Vol. 83, pp. 255–279.

Gupta, P.L., Gupta, R.C. and Tripathi, R.C. (1997). On the monotonic properties of discrete failure rates, Journal of Statistical Planning and Inference, Vol. 65, pp. 255–268.

Gupta, R.C., Sim, S.Z. and Ong, S.H. (2014). Analysis of discrete data by Conway-Maxwell Poisson distribution, AStA Adv Stat Anal, Vol. 98, pp. 327–343.

Henningsen, A. and Toomet, O. (2011). maxLik: A package for maximum likelihood estimation in R, Comput Stat, No. 26, pp. 443–458.

Huang, K. and Mi, J. (2018). Applications of likelihood ratio order in Bayesian inferences, Probability in the Engineering and Informational Sciences, Vol. 34, No. 1, pp.1-13.

Kendall, M.G. (1961). Natural law in social sciences, Journal of the Royal Statistical Society Series A., Vol. 124., pp. 1–19. https://doi.org/10.2307/2984139

Kim, B., Kim, J., and Lee, S. (2018). Strong unimodality of discrete order statistics, Statistics & Probability Letters, Vol. 140, pp. 48–52.

Kokonendji, C.C., Mizère, D. and Balakrishnan, N. (2008). Connections of the Poisson weight function to overdispersion and underdispersion, Journal of Statistical Planning and Inference, No. 138, pp. 1287–1296.

Kulasekera, K.B. and Tonkyn, D.W. (1992). A new discrete distribution, with applications to survival, dispersal and dispersion, Communications in Statistics - Simulation and Computation, Vol. 21, No. 2, pp 499–518.

Lakshmi, K., Mahaboob, B., Rajaiah, M. and Narayana, C. (2021). Ordinary least squares estimation of parameters of linear model, J. Math. Comput. Sci., Vol. 11, No. 2, pp. 2015–2030.

Louzayadio, C.G., Malouata, R.O. and Koukouatikissa Diafouka, M. (2021). A weighted Poisson distribution for underdispersed count data, International Journal of Statistics and Probability, Vol. 10, No. 4, pp. 157–165.

Meena, K.R. and Gangopadhyay, A.K. (2020). Estimating parameter of the selected uniform population under the generalized stein loss function, Applications and Applied Mathematics: An International Journal (AAM), Vol. 15, Iss. 2, 894–915.

Minka, T.P., Shmueli, G., Kadane, J.B., Borle, S. and Boatwright, P. (2003). Computing with the COM-Poisson distribution, *Technical Report 77*. Department of Statistics, Carnegie Mellon University, Pittsburgh. (Available from http://www.stat.cmu.edu/tr/

Patil, G.P. (2002). Weighted distributions, Encyclopedia of Environmetrics (eds. A.H. ElShaarawi and W.W. Piegorrsch), Vol. 4, pp. 2369–2377. Chichester: John Wiley & Sons. https://doi.org/10.1002/9781118445112.stat07359

Patil, G.P., Rao, C.R. and Ratnaparkhi, M.V. (1986). On discrete weighted distributions and their use in model choice for observed data, Communications in Statistics-Theory and Methods, Vol. 15, No. 3, pp. 907–918. http://dx.doi.org/10.1080/03610928608829159

Perrakis, S. (2019). Stochastic dominance option pricing: An alternative approach to option market research, Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-11590-6_1

Qureshi, M.I. and Shadab, M. (2020). Some summation theorems for Clausen's hypergeometric functions with unit argument, Applications and Applied Mathematics: An International Journal (AAM), Vol. 15, Iss. 1, pp. 469–478.

Raeisi, M. and Yari, G. (2019). A further result on the aging properties of an extended additive hazard model, Applications and Applied Mathematics: An International Journal (AAM), Vol. 14, Iss. 2, pp. 777–783.

Rao, C.R. (1965). On discrete distributions arising out of methods of ascertainment, Sankhya Sr. A, pp. 311–324.

Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*, New York, NY: Springer New York.

Shmueli, G., Minka, T.P., Kadane, J.B., Borles, S. and Batwright, P. (2005). A useful distribution for fitting discrete data: Revival of the Conway-Maxwell-Poisson distribution, Appl. Statist., No. 54, pp. 127–142.

Weller, S.R. and Martin, J.H. (2020). On strongly unimodal third-order SISO linear systems with applications to pharmacokinetics, IFAC-PapersOnLine, Vol. 53, No. 2, pp. 4654–4661.