Statistics and Applications {ISSN 2454-7395 (online)}
Volume 21, No. 1, 2023 (New Series), pp 217–230

Conway-Maxwell Poisson Distribution: Some New Results and Minimum Variance Unbiased Estimation

Jahnavi Merupula and V.S.Vaidyanathan
Department of Statistics
Pondicherry University, Puducherry, India

Received: 02 April 2022; Revised: 06 July 2022; Accepted: 12 July 2022

Abstract

Probability distributions for count data have potential applications in medical, epidemiological, and actuarial studies. Conway-Maxwell Poisson (CMP) distribution is a two-parameter Poisson distribution that can handle over- and under-dispersed data. In this paper, some new results on the distributional properties of CMP distribution are presented. Also, minimum variance unbiased (MVU) estimator of the location parameter is derived using the complete sufficient statistic. The primary advantage of the MVU estimator is that it has a closed-form expression, unlike other existing estimators. An approximate expression for the variance of the MVU estimator is obtained, and the performance of the MVU estimator is compared with that of the ML estimator in terms of relative efficiency through simulated and real-life datasets.

Key words: CMP distribution; Generalized hypergeometric series; Minimum variance unbiased estimation; Power series family; Relative efficiency; Sufficient statistic.

AMS Subject Classifications: 60E05, 62F10

1. Introduction

Poisson distribution is often a natural choice among researchers to model count data. However, its applicability is restricted to situations where there is equi-dispersion of data, i.e., the mean is equal to the variance. Often, in reality, count data are over- or under-dispersed. For example, data on the word lengths in a dictionary or the number of infected spots in the leaves of a plant is under-dispersed. Alternative distributions to Poisson are available in the literature to model over- or under-dispersed data. These include mixtures of Poisson, weighted Poisson and generalized Poisson distributions. However, these distributions have more parameters and involve mathematical intricacies which limit their usage. For example, the generalized Poisson distribution does not model under-dispersion effectively due to parameter constraints. Hence, probability models having fewer parameters that can address the problem of over- or under-dispersion are of interest to study both from a theoretical and application perspective.

A two-parameter Poisson distribution capable of handling over- and under-dispersion is Conway-Maxwell Poisson (CMP) distribution introduced by Conway and Maxwell (1961).

Corresponding Author: V. S.Vaidyanathan
Email: vaidya.stats@gmail.com

The probability mass function (pmf) of CMP distribution is

$$P(X=x)=\frac{\lambda^x}{(x!)^\nu}\frac{1}{Z(\lambda,\nu)},\quad x=0,1,2,\ldots,\quad \lambda>0,\quad \nu\geq 0 \tag{1}$$

where

$$Z(\lambda,\nu)=\sum_{j=0}^{\infty}\frac{\lambda^j}{(j!)^\nu}$$

is the normalizing constant. Here, λ denotes the location parameter and ν denotes the dispersion parameter that captures the degree of over- or under-dispersion. The CMP distribution is over-dispersed for $\nu<1$, under-dispersed for $\nu>1$, and equi-dispersed for $\nu=1$. The pmf is not defined for $\nu=0$ and $\lambda\geq 1$.

Shmueli *et al.* (2005) have revisited this distribution to study its properties. A review of CMP distribution, its characterizations and applications can be found in Sellers *et al.* (2012). There has been an increased interest in research about extensions and generalizations of CMP distribution in the recent past. Cordeiro *et al.* (2012) introduced exponential-CMP distribution as a lifetime distribution by compounding an exponential distribution with a CMP distribution and explored its properties. Chakraborty and Imoto (2016) proposed a flexible four-parameter extension of CMP distribution, which encompasses Conway-Maxwell negative binomial and generalized CMP distributions, and also derived its properties. Roy *et al.* (2020) developed Conway-Maxwell negative hypergeometric distribution as a modification to negative hypergeometric distribution along with its characterizations.

Although some extensions and characterizations of CMP distribution are available, properties in terms of differential equation involving recurrent probabilities have not yet been addressed. Such properties are available for popular discrete distributions, see, for example, Boswell and Patil (1973), and the same is discussed for CMP distribution in this paper. Also, a new representation of the CMP distribution in terms of generalized hypergeometric series is given.

From an inferential point of view, existing estimators of the parameters do not have closed-form expressions and have to be computed using iterative methods. Since the CMP distribution belongs to the exponential family of distributions, in the present work, minimum variance unbiased (MVU) estimation of the location parameter is carried out using the distribution of the complete sufficient statistic. Also, an approximate expression for the variance of the estimator is obtained. The merit of using the proposed MVU estimator is highlighted through numerical illustration.

The paper is organized as follows. In Section 2, some properties of CMP distribution are listed, and two new results involving recurrent probabilities and probability generating function (pgf) are presented. The methodology to obtain MVU estimator of the location parameter is explained in Section 3. Numerical illustration to compare the performance of the MVU estimation with likelihood estimation in terms of mean absolute bias (MAB) and relative efficiency (RE) is provided in Section 4 through simulated and real-life data. Concluding remarks are given in Section 5.

2. Distributional properties

As mentioned in Section 1, CMP distribution can model both over- and under-dispersed data. CMP distribution encapsulates well-known distributions, including Poisson distribution ($\nu = 1$), geometric distribution ($\nu = 0, \lambda < 1$), and Bernoulli distribution ($\nu \to \infty$). From equation (1), based on n independent and identically distributed (iid) samples on X, it can be seen that CMP distribution belongs to the exponential family of distributions. Also, CMP distribution is a member of the two-parameter power series family of distributions with pgf of the form

$$P_X(z) = \frac{Z(\lambda z, \nu)}{Z(\lambda, \nu)} \tag{2}$$

The pgf of CMP distribution can be expressed in terms of generalized hypergeometric series as (See Nadarajah, 2009)

$$P_X(z) = \frac{{}_0F_{\nu-1}(;1,\ldots,1;\lambda z)}{{}_0F_{\nu-1}(;1,\ldots,1;\lambda)} \tag{3}$$

Comparing equations (2) and (3), we get

$$Z(\lambda, \nu) = {}_0F_{\nu-1}(;1,\ldots,1;\lambda) \tag{4}$$

Using the pgf, the expected value and variance of X can be obtained as

$$E(X) = \lambda \frac{\partial}{\partial \lambda} \log(Z(\lambda, \nu)) \tag{5}$$

and

$$V(X) = \lambda \frac{\partial}{\partial \lambda} \left[\lambda \frac{\partial}{\partial \lambda} \log(Z(\lambda, \nu)) \right]$$

For further properties and characterizations of CMP distribution, one may refer to Nadarajah (2009), Daly and Gaunt (2016) and Li *et al.* (2019). In the sequel, two new results on CMP distribution are presented.

2.1. Recurrence relationship of probabilities

Boswell and Patil (1973) have shown that any discrete distribution can be characterized in terms of differential equations involving its parameters. For example, Poisson distribution with mean λ satisfy the following recurrence relationship, namely,

$$\frac{dp_x}{d\lambda} = p_{x-1} - p_x$$

where p_x is the pmf of the Poisson distribution. A similar recurrence relationship for CMP distribution is obtained below.

Result 1: Let p_x denote the pmf of CMP distribution. Then

$$\frac{\partial p_x}{\partial \lambda} = \frac{1}{x^{\nu-1}} p_{x-1} - \frac{E(X)}{\lambda} p_x$$

Proof: Partially differentiating the pmf in equation (1) with respect to λ, we get,

$$\frac{\partial p_x}{\partial \lambda} = \frac{x\lambda^{x-1}Z(\lambda,\nu) - \lambda^x \frac{\partial}{\partial \lambda}Z(\lambda,\nu)}{(x!)^\nu Z^2(\lambda,\nu)}$$

Note that $\frac{\partial}{\partial \lambda}\log(Z(\lambda,\nu)) = \frac{1}{Z(\lambda,\nu)}\frac{\partial}{\partial \lambda}Z(\lambda,\nu)$. Thus

$$\begin{aligned}\frac{\partial p_x}{\partial \lambda} &= \frac{x\lambda^{x-1}}{(x!)^\nu Z(\lambda,\nu)} - \frac{\lambda^x \frac{\partial}{\partial \lambda}\log(Z(\lambda,\nu))}{(x!)^\nu Z(\lambda,\nu)} \\ &= \frac{1}{x^{\nu-1}}\frac{\lambda^{x-1}}{((x-1)!)^\nu Z(\lambda,\nu)} - \frac{\lambda^x}{(x!)^\nu Z(\lambda,\nu)}\frac{\partial}{\partial \lambda}\log(Z(\lambda,\nu)) \end{aligned} \tag{6}$$

Using equations (1) and (5) in equation (6), we get,

$$\frac{\partial p_x}{\partial \lambda} = \frac{1}{x^{\nu-1}}p_{x-1} - \frac{E(X)}{\lambda}p_x$$

□

An illustration of the computation of the probabilities using the recurrence relation for the parameter choice $(\lambda,\nu) = (1.2, 0.5)$ is shown below. To carry out the recursive process, the values of $P(X = 0)$ and $E(X)$ need to be computed. $P(X = 0)$ is computed by substituting $x = 0$ in equation (1) and $E(X)$ is computed from equation (5) using `com.mean()` function available in `compoisson` package in R. The values are found to be 0.2096 and 1.992285, respectively. The successive probabilities are computed using the recurrence relation

$$p_x = \frac{\lambda}{E(X)}\left[\frac{1}{x^{\nu-1}}p_{x-1} - \frac{\partial p_x}{\partial \lambda}\right], \quad x = 1, 2, \ldots$$

and are tabulated below for $x = 1, 2, 3, 4$.

Table 1: Recurrence probabilities for $(\lambda,\nu) = (1.2, 0.5)$

x	$\frac{\partial p_x}{\partial \lambda}$	p_x
0	-	0.2096
1	0.0962	0.2511
2	0.0020	0.2127
3	0.1246	0.1468
4	0.1487	0.0872

2.2. CMP as Generalized Hypergeometric distribution

A discrete random variable X with pmf

$$P(X = k) = C\frac{\lambda^k}{k!}\gamma_k[(\mathbf{a});(\mathbf{c})], \quad k = 0, 1, 2, \ldots \tag{7}$$

is said to belong to generalized hypergeometric family of distributions, provided its pgf can be expressed in terms of generalized hypergeometric series as (See Dacey, 1972)

$$P_X(z) = C_pF_q[(\mathbf{a});(\mathbf{c});\lambda z] \tag{8}$$

Here C denotes the normalizing constant and

$$\gamma_k[(\mathbf{a});(\mathbf{c})] = \frac{\Gamma[(\mathbf{a}+k);(\mathbf{c}+k)]}{\Gamma[(\mathbf{a});(\mathbf{c})]}$$

with

$$\Gamma[(\mathbf{a});(\mathbf{c})] = \frac{\Gamma(a_1)\Gamma(a_2)\ldots\Gamma(a_p)}{\Gamma(c_1)\Gamma(c_2)\ldots\Gamma(c_q)}$$

Result 2: CMP distribution belongs to the generalized hypergeometric family of distributions.

Proof: Let ${}_pF_q[(\mathbf{a});(\mathbf{c});t]$ denote the generalized hypergeometric series where $\mathbf{a} = (a_1, a_2, \ldots, a_p)$ and $\mathbf{c} = (c_1, c_2, \ldots, c_q)$. The pgf of CMP distribution given in equation (3) is obtained by taking $p = 0$, $q = \nu - 1$ and $\mathbf{c} = (1, 1, \ldots, 1)$. Comparing equation (3) with equation (8), we get,

$$\begin{aligned}
P(X=k) &= \frac{1}{{}_0F_{\nu-1}(;1,\ldots,1;\lambda)}\frac{\lambda^k}{k!}\gamma_k[;(1,\ldots,1)] \\
&= \frac{1}{Z(\lambda,\nu)}\frac{\lambda^k}{k!}\frac{\Gamma[;(1+k)]}{\Gamma[;(1)]} \qquad \text{(using equation (4))} \\
&= \frac{1}{Z(\lambda,\nu)}\frac{\lambda^k}{k!}\left[\frac{1}{\underbrace{\Gamma(1+k)\ldots\Gamma(1+k)}_{(\nu-1)\text{ terms}}}\right]\left[\frac{1}{\underbrace{\Gamma(1)\ldots\Gamma(1)}_{(\nu-1)\text{ terms}}}\right]^{-1} \\
&= \frac{1}{Z(\lambda,\nu)}\frac{\lambda^k}{k!}\frac{1}{(k!)^{\nu-1}} \\
&= \frac{1}{Z(\lambda,\nu)}\frac{\lambda^k}{(k!)^{\nu}}
\end{aligned}$$

which is the pmf of the CMP distribution. Hence the result. □

3. Minimum variance unbiased estimation

In this section, we propose a minimum variance unbiased estimator for the location parameter λ of the CMP distribution when ν is known. For fixed ν, CMP distribution belongs to the one-parameter power series family of distributions.
From Roy and Mitra (1957), the pmf of the complete sufficient statistic T of one-parameter power series family of distributions with parameter θ is given by

$$P(T=t) = \frac{A(t,n)\theta^t}{[c(\theta)]^n} \tag{9}$$

Accordingly, the MVU estimator of θ^r, $r = 1, 2, \ldots$, denoted by $\delta(t, r)$ is

$$\delta(t,r) = \begin{cases} 0, & \text{if } t < r \\ \dfrac{A(t-r,n)}{A(t,n)}, & \text{if } t \geq r \end{cases} \tag{10}$$

Since CMP distribution belongs to the exponential family of distributions, for fixed ν, $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic for λ. The pmf of $T = \sum_{i=1}^{n} X_i$ is given by (Sellers *et al.*, 2017)

$$P(T=t) = P(t) = \frac{\lambda^t}{(t!)^\nu [Z(\lambda,\nu)]^n} \sum_{\substack{x_1,x_2,\ldots,x_n \\ x_1+x_2+\ldots+x_n=t}} \binom{t}{x_1 \ldots x_n}^\nu, t = 0, 1, \ldots \tag{11}$$

Comparing equation (11) with equation (9), it can be seen that

$$A(t,n) = \frac{1}{(t!)^\nu} \sum_{\substack{x_1,x_2,\ldots,x_n \\ x_1+x_2+\ldots+x_n=t}} \binom{t}{x_1 \ldots x_n}^\nu \tag{12}$$

Using equation (12) in equation (10) with $r = 1$, the MVU estimator of λ, namely, $\delta(t, 1) = \delta(t)$ (say) is obtained as

$$\delta(t) = \begin{cases} 0, & \text{if } t < 1 \\ \dfrac{\sum\limits_{\substack{x_1,x_2,\ldots,x_n \\ x_1+x_2+\ldots+x_n=t-1}} \binom{t-1}{x_1 \ldots x_n}^\nu}{\sum\limits_{\substack{x_1,x_2,\ldots,x_n \\ x_1+x_2+\ldots+x_n=t}} \binom{t}{x_1 \ldots x_n}^\nu} \, t^\nu, & \text{if } t \geq 1. \end{cases} \tag{13}$$

To verify that $\delta(t)$ is indeed an unbiased estimator of λ, we proceed as follows.

Consider the ratio of consecutive probabilities of T, namely,

$$\frac{P(t-1)}{P(t)} = \frac{t^\nu}{\lambda} \frac{\sum\limits_{\substack{x_1,x_2,\ldots,x_n \\ x_1+x_2+\ldots+x_n=t-1}} \binom{t-1}{x_1 \ldots x_n}^\nu}{\sum\limits_{\substack{x_1,x_2,\ldots,x_n \\ x_1+x_2+\ldots+x_n=t}} \binom{t}{x_1 \ldots x_n}^\nu} \tag{14}$$

Using equation (14) in equation (13) and taking expectation, we get,

$$\begin{aligned} E[\delta(t)] &= E\left[\frac{P(t-1)}{P(t)}\lambda\right] \\ &= \lambda \sum_{t=1}^{\infty} \frac{P(t-1)}{P(t)} P(t) \\ &= \lambda \sum_{t=1}^{\infty} P(t-1) \\ &= \lambda \end{aligned}$$

Thus, $\delta(t)$ is an unbiased estimator of λ.

An alternate expression for $\delta(t)$ can be obtained by using the approximation for the sums of powers of multinomial coefficients given below.

$$\sum_{\substack{x_1,x_2,\ldots,x_n \\ x_1+x_2+\ldots+x_n=t}} \binom{t}{x_1 \ldots x_n}^{\nu} \simeq n^{\nu t}\sqrt{\frac{K_{n\nu}}{(\pi t)^{(n-1)(\nu-1)}}}, \tag{15}$$

where

$$K_{n\nu} = \frac{n^{n(\nu-1)}}{\nu^{(n-1)}2^{(n-1)(\nu-1)}}.$$

From equation (13), an approximate expression for $\delta(t)$ is

$$\delta(t) \simeq \frac{t^{\nu}}{n^{\nu}}\sqrt{\left(\frac{t}{t-1}\right)^{(\nu-1)(n-1)}}, \quad t > 1. \tag{16}$$

An approximate expression for the variance of $\delta(t)$ in equation (16) is obtained as follows. Consider

$$\begin{aligned}
V[\delta(t)] &= E[\delta^2(t)] - (E[\delta(t)])^2 \\
&\simeq E\left[\frac{t^{2\nu}}{n^{2\nu}}\left(\frac{t}{t-1}\right)^{(\nu-1)(n-1)}\right] - \lambda^2 \\
&= \left[\sum_{t=2}^{\infty}\frac{t^{2\nu}}{n^{2\nu}}\left(\frac{t}{t-1}\right)^{(\nu-1)(n-1)} P(t)\right] - \lambda^2 \\
&= \left[\sum_{t=2}^{\infty}\frac{t^{2\nu}}{n^{2\nu}}\left(\frac{t}{t-1}\right)^{(\nu-1)(n-1)} \frac{\lambda^t}{(t!)^{\nu}(Z(\lambda,\nu))^n} \sum_{\substack{x_1,x_2,\ldots,x_n \\ x_1+x_2+\ldots+x_n=t}} \binom{t}{x_1 \ldots x_n}^{\nu}\right] - \lambda^2 \\
&= \frac{1}{n^{2\nu}(Z(\lambda,\nu))^n}\sum_{t=2}^{\infty} t^{2\nu}\left(\frac{t}{t-1}\right)^{(\nu-1)(n-1)} \frac{\lambda^t}{(t!)^{\nu}}\left[n^{\nu t}\left(\frac{n^{n(\nu-1)}}{(2\pi t)^{(n-1)(\nu-1)}\nu^{n-1}}\right)^{1/2}\right] - \lambda^2 \\
&\qquad \text{(using equation (15))} \\
&= \frac{1}{n^{2\nu}(Z(\lambda,\nu))^n}\left(\frac{n^{n(\nu-1)}}{2\pi^{(n-1)(\nu-1)}\nu^{n-1}}\right)^{1/2} \sum_{t=2}^{\infty}\frac{n^{\nu t}\lambda^t t^{2\nu}}{(t!)^{\nu}} \left(\frac{1}{t-1}\right)^{(n-1)(\nu-1)} - \lambda^2.
\end{aligned}$$

4. Comparison of MVU and likelihood estimation

In this section, the performance of the MVU estimator of λ is compared with that of the maximum likelihood (ML) estimator through MAB and RE using simulated and real-life datasets. The likelihood function based on n iid observations, namely, $\vec{x} = (x_1, x_2, \ldots, x_n)$ on X having CMP distribution is given by

$$L(\lambda;\nu,\vec{x}) = \lambda^{\sum_{i=1}^{n} x_i}\prod_{i=1}^{n}(x_i!)^{-\nu}\left[Z(\lambda,\nu)\right]^{-n}$$

Since $Z(\lambda,\nu)$ involve an infinite sum, a closed form expression for the likelihood estimator of λ, namely, $\hat{\lambda}_{ML}$ cannot be obtained. However, an estimate of $\hat{\lambda}_{ML}$, can be obtained using

Newton-Raphson method. `COMPoissonReg` package in R contains functions to compute the ML estimates. The RE of $\delta(t)$ with respect to $\hat{\lambda}_{ML}$ is defined as

$$RE(\delta(t), \hat{\lambda}_{ML}) = \frac{V[\hat{\lambda}_{ML}]}{V[\delta(t)]}$$

A value of RE more than one imply that $V[\delta(t)]$ is less than $V[\hat{\lambda}_{ML}]$, suggesting that $\delta(t)$ is efficient than $\hat{\lambda}_{ML}$. However, $V[\hat{\lambda}_{ML}]$ does not have a closed-form expression. Therefore, we make use of bootstrap approach to compute the RE values. The method to compute RE using bootstrap samples is given in the following steps.

1. Generate a random sample n^* of size n from CMP distribution for fixed λ and ν.

2. Draw B bootstrap samples each of size n with replacement from n^*.

3. For each bootstrap sample, compute $\delta(t)$ and $\hat{\lambda}_{ML}$. Denote these values as $\delta^{[b]}(t)$, $\hat{\lambda}_{ML}^{[b]}$, $b = 1, 2, \ldots, B$.

4. Using the B bootstrap estimates of $\delta(t)$ and $\hat{\lambda}_{ML}$, calculate $v(\delta(t))$ and $v(\hat{\lambda}_{ML})$ defined respectively as
$$v(\delta(t)) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\delta^{[b]}(t) - \delta^*(t)\right)^2$$
where
$$\delta^*(t) = \frac{1}{B}\sum_{b=1}^{B}\delta^{[b]}(t)$$
and
$$v(\hat{\lambda}_{ML}) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\lambda}_{ML}^{[b]} - \hat{\lambda}_{ML}^*\right)^2$$
where
$$\hat{\lambda}_{ML}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{\lambda}_{ML}^{[b]}$$

5. RE based on bootstrap samples is computed as the ratio of $v(\hat{\lambda}_{ML})$ to $v(\delta(t))$

4.1. Simulation study

A simulation study is carried out to examine the behaviour of the ML and MVU estimates by computing the MAB and RE. Random samples of sizes $n = 25, 50$ are generated from the CMP distribution by fixing the parameters λ and ν as below.

- Case 1: $\nu = 0.2$, $\lambda \in \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0\}$
- Case 2: $\nu = 2.0$, $\lambda \in \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0\}$

The `COMPoissonReg` package in R (Kimberly *et al.*, 2019) is used to generate the sample observations. Case 1 corresponds to over-dispersed counts and Case 2 to under-dispersed counts. Based on the simulated observations, $\delta^{[b]}(t)$, $\hat{\lambda}^{[b]}_{ML}$ and RE are computed using the bootstrap procedure given in steps 1 to 5 of the previous section taking $B = 200$. To find $\hat{\lambda}^{[b]}_{ML}$, the in-built function `glm.cmp()` available in `COMPoissonReg` package is used. To understand the fluctuations in the above values when the sample observations change, the procedure is repeated for 100 runs. One run of the bootstrap procedure will yield a value for the RE, $v(\hat{\lambda}_{ML})$ and $v(\delta(t))$. Based on the B bootstrap estimates in each run, MAB of the estimators are computed. MAB of the bootstrap estimator L of the parameter θ is defined as $MAB = \dfrac{1}{B}\sum_{b=1}^{B}|L^{[b]} - \theta|$. The summary statistics of the MAB values under case 1 and 2 for $n = 25$ and 50 are presented in Table 2.

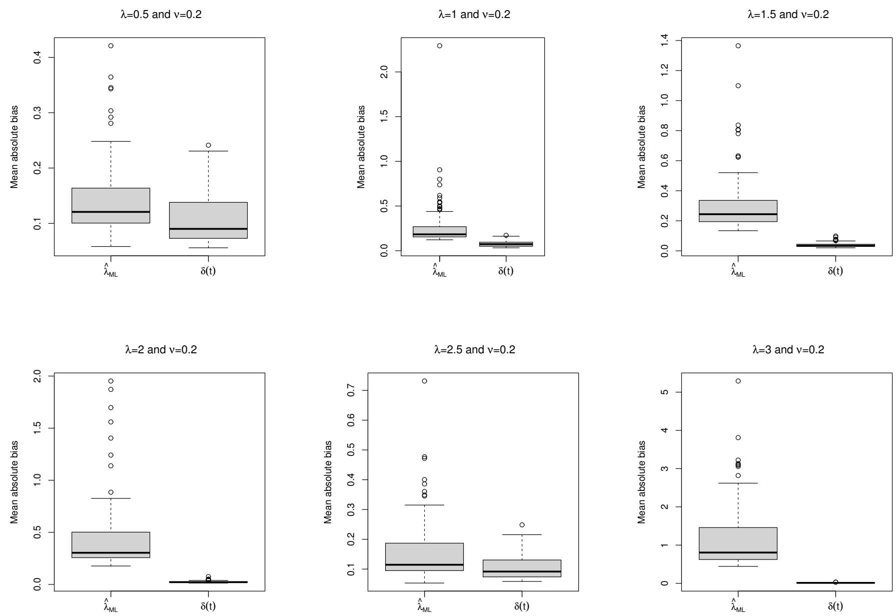
Table 2: Summary statistics (min, 25^{th} quantile, median, mean, 75^{th} quantile, max) of MAB values of $\hat{\lambda}_{ML}$ and $\delta(t)$ under case 1 and case 2 for $n = 25$ and 50

		Case 1	
		ν=0.2	
λ	Estimator	$n = 25$	$n = 50$
0.5	$\hat{\lambda}_{ML}$	(0.0750, 0.1404,0.1829,0.2504,0.2831,1.1164)	(0.0582,0.1006,0.1207,0.1460,0.1636,0.4210)
	$\delta(t)$	(0.0680,0.0960,0.1196,0.1319,0.1578,0.2877)	(0.0558,0.0730,0.0900,0.1097,0.1379,0.2413)
1.0	$\hat{\lambda}_{ML}$	(0.1680, 0.2274,0.2870,0.4468,0.5143,1.7551)	(0.1221,0.1546,0.1831,0.2734,0.2631,2.2928)
	$\delta(t)$	(0.0421,0.0679,0.0802,0.0900,0.1166,0.1743)	(0.0327,0.0506,0.0734,0.0771,0.0947,0.1727)
1.5	$\hat{\lambda}_{ML}$	(0.0421,0.0679,0.0802,0.0900,0.1166,0.1743)	(0.1343,0.1952,0.2439,0.3188,0.3358,1.3648)
	$\delta(t)$	(0.0288,0.0381,0.0467,0.0519,0.0599,0.1357)	(0.0204,0.0290,0.0340,0.0390,0.0439,0.0986)
2.0	$\hat{\lambda}_{ML}$	(0.2569,0.4035,0.5755,1.1046,1.0828,10.7535)	(0.1773,0.2574,0.3040,0.4520,0.5003,1.9523)
	$\delta(t)$	(0.0187,0.0256,0.0293,0.0335,0.0348,0.0987)	(0.0135,0.0190,0.0220,0.0247,0.0269,0.0753)
2.5	$\hat{\lambda}_{ML}$	(0.4173,0.6163,0.8842,1.9223,2.0632,24.4187)	(0.0530,0.0950,0.1145,0.1581,0.1869,0.7314)
	$\delta(t)$	(0.0141,0.0186,0.0225,0.0252,0.0278,0.0625)	(0.0585,0.0739,0.0916,0.1069,0.1297,0.2484)
3.0	$\hat{\lambda}_{ML}$	(0.5448,1.0515,1.5072,3.8200,3.2055,67.8926)	(0.4440,0.6244,0.8051,1.1763,1.4477,5.2907)
	$\delta(t)$	(0.0092,0.0141,0.0176,0.0205,0.0245,0.0563)	(0.0078,0.0101,0.0126,0.0144,0.0175,0.0341)
		Case 2	
		ν=2.0	
λ	Estimator	$n = 25$	$n = 50$
0.5	$\hat{\lambda}_{ML}$	(0.0750,0.1404,0.1829,0.2504,0.2831,1.1164)	(0.0582,0.1006,0.1207,0.1460,0.1637,0.4210)
	$\delta(t)$	(0.0680,0.0960,0.1196,0.1319,0.1578,0.2877)	(0.0558,0.0730,0.0901,0.1097,0.1379,0.2413)
1.0	$\hat{\lambda}_{ML}$	(0.0000,0.0000,0.0000,9.7e+05,1.0000,9.7e+07)	(0.2016,0.2394,0.2754,0.3733,0.4210,1.3105)
	$\delta(t)$	(0.1378,0.1953,0.2286,0.2703,0.3053,1.2498)	(0.1218,0.1505,0.1869,0.2168,0.2527,0.7250)
1.5	$\hat{\lambda}_{ML}$	(0.0000,1.0000,1.0000,1.3e+06,1.0000,7.3e+07)	(0.3160,0.3915,0.4694,0.5979,0.6645,2.0552)
	$\delta(t)$	(0.1686,0.2887,0.3404,0.3877,0.4182,1.0897)	(0.1683,0.2167,0.2513,0.2860,0.3255,0.6965)
2.0	$\hat{\lambda}_{ML}$	(1.0000,1.0000,1.0000,3.3e+06,3.0000, 1.1e+08)	(0.4564,0.5459,0.6823,1.0315,1.1313,7.3437)
	$\delta(t)$	(0.2765,0.3927,0.4420,0.5051,0.5440,1.4926)	(0.1950,0.2934,0.3349,0.3841,0.4470,0.9988)
2.5	$\hat{\lambda}_{ML}$	(1.0000,1.0000,2.0000,3.8e+06,2.3e+05,9.5e+07)	(0.5956,0.7449,0.8666,1.6464,1.7517,10.5577)
	$\delta(t)$	(0.3069,0.4858,0.5898,0.6579,0.7665,1.3116)	(0.2547,0.3283,0.3813,0.4420,0.5051,1.1666)
3.0	$\hat{\lambda}_{ML}$	(1.0000,2.0000,4.0000,6.8e+06,6.8e+05,2.1e+08)	(1.0000, 1.0000,1.0000,4.2e+05,3.0000,4.2e+07)
	$\delta(t)$	(0.4241,0.5499,0.6534,0.7568,0.8361,2.5188)	(0.2375,0.3989,0.4527,0.5337,0.6380,1.4019)

The boxplots of MAB values under both the cases for $n = 50$ are given in Table 3. It is observed from the plots and the summary statistics that the MAB values corresponding to MVU estimator are comparatively small and less dispersed than that of ML estimator under both the cases. Also, the presence of extreme values in the plots corresponding to the ML estimator indicate that the likelihood approach at times over or under estimates the parameter. The line plots displayed in Table 4 correspond to the variances of the ML and MVU estimates based on bootstrap samples for 100 runs. The x-axis in the plots denote the runs and the y-axis denote the variances of the estimates.

Table 3: Boxplots of MABs of $\hat{\lambda}_{ML}$ and $\delta(t)$ for $n = 50$

Case 1



Case 2

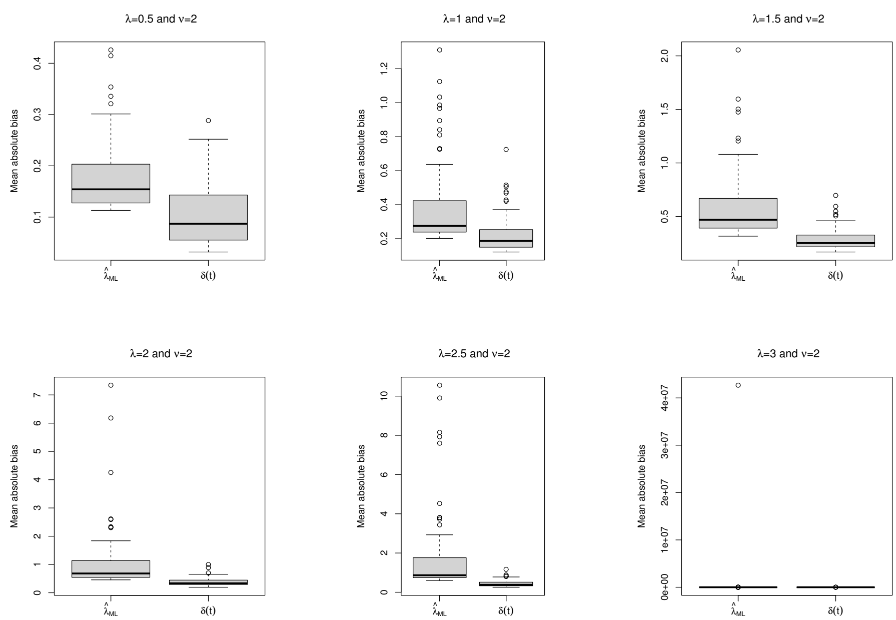
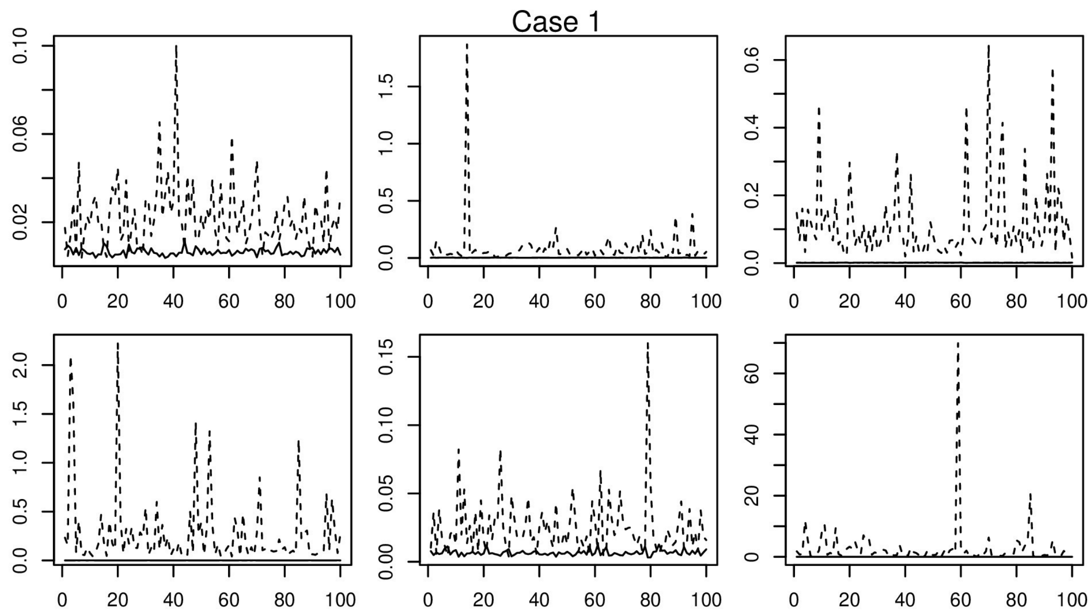
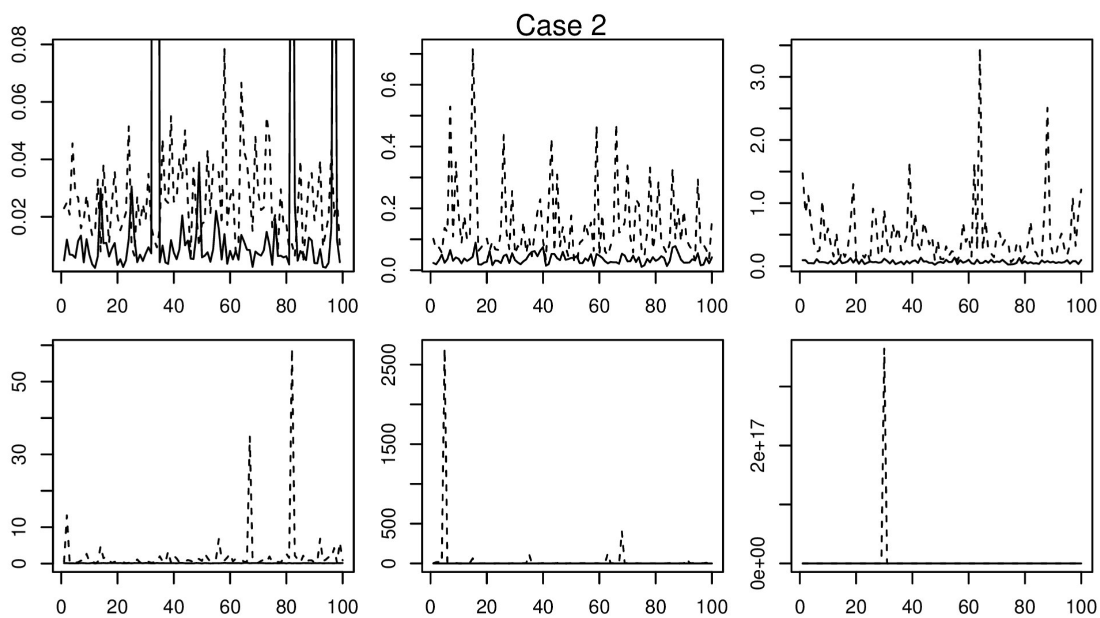


Table 4: Plots of variances of $\hat{\lambda}_{ML}$ and $\delta(t)$ for $n = 50$



The plots in clockwise direction correspond to (λ,ν) in the order (0.5,0.2), (1.0,0.2), (1.5,0.2), (3.0,0.2), (2.5,0.2) and (2.0,0.2)



The plots in clockwise direction correspond to (λ,ν) in the order (0.5,2.0), (1.0,2.0), (1.5,2.0), (3.0,2.0), (2.5,2.0) and (2.0,2.0)

-- $\hat{\lambda}_{ML}$ — δ(t)

It can be observed from the plots that the estimates of $\delta(t)$ are less dispersed compared to $\hat{\lambda}_{ML}$. Also, it is observed that the variances of the ML estimates are large when compared to that of MVU estimates. In particular, for $\nu = 2$ and $\lambda = 3$, the variance is found to be much larger than 2e+17 in some runs. However, the corresponding variances of the MVU estimates are very close to zero for all the runs.

For each of the cases, RE value is computed and the proportion of RE values greater than one in the 100 runs are obtained for $n = 25$ and 50 respectively. The proportions are given in Tables 5 and 6.

Table 5: Proportion of RE values greater than one for Case 1

λ	0.5	1.0	1.5	2.0	2.5	3.0
$n = 25$	0.86	0.99	1.00	1.00	1.00	1.00
$n = 50$	0.88	0.99	0.99	1.00	0.86	1.00

Table 6: Proportion of RE values greater than one for Case 2

λ	0.5	1.0	1.5	2.0	2.5	3.0
$n = 25$	0.79	0.94	0.98	1.00	0.98	0.97
$n = 50$	0.88	0.96	0.98	0.99	0.99	0.99

As seen from Tables 5 and 6, the proportion of times RE values greater than one is more than 0.8, at times closer to 1, for both the cases indicating $\delta(t)$ yields estimates having smaller variance than $\hat{\lambda}_{ML}$. Thus, the simulation results indicate that the proposed MVU estimator is better than the ML in terms of MAB and variance for both over- and under-dispersed data.

4.2. Real-life illustration

As an application of the proposed estimation method to real-life data, we consider the article publishing dataset given in Long (1997). The data relates to the number of articles (X) published by Ph.D. biochemists (B). The dataset is as given in Table 7.

Table 7: Article publication data

X	0	1	2	3	4	5	6	7	8	9	10	11	12	16	19
B	275	246	178	84	67	27	17	12	1	2	1	1	2	1	1

The data is tested for equi-dispersion using `dispersiontest()` in `AER` package in R and the results indicate that the data is over-dispersed (p-value is 1.44e-06, dispersion index is 2.1889). Hence, Poisson distribution is not a suitable choice to model the data and therefore it can be modelled using CMP distribution. The dispersion parameter ν is estimated using the method of moments and is found to be $\hat{\nu} = 0.1249$. The estimate of the location parameter of λ is obtained using the proposed MVU estimator $\delta(t)$ and the ML estimator by fixing $\nu = 0.1249$. The MVU and the ML estimates are found to be 0.8248 and 0.7809, respectively. To compute the sample variances of the estimates, bootstrap samples each of size $n = 915$ are replicated for $B = 200$ times from the data set. The corresponding sample variances are found to be 0.0001304 and 0.0021546. The plot of the observed and the expected frequencies from CMP distribution using $\delta(t)$, $\hat{\lambda}_{ML}$ and $\nu = 0.1249$ is shown in Figure 1. The corresponding residual (difference of the observed and expected frequency) plots are also presented. From the plots, it can be observed that both the estimators provide similar fits. However, the variance of the MVU estimator of λ is smaller than that of the ML estimator suggesting that MVU estimation is efficient.

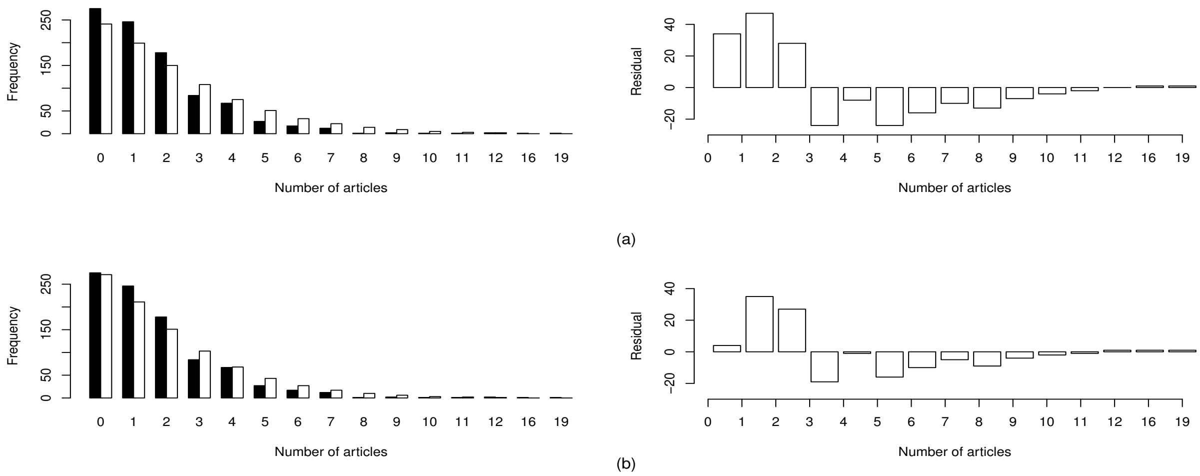


Figure 1: Observed (■) and expected (□) frequencies of CMP distribution with (a) MVU estimate of λ and (b) ML estimate of λ with the corresponding residual plots

5. Concluding remarks

The method of MVU estimation of the location parameter of CMP distribution proposed in this paper is simple and easy to compute. Unlike the existing estimators available in the literature, the proposed MVU estimator has a closed-form expression and does not require iterative procedures for computation. The estimator is based on the distribution of the complete sufficient statistic of the parameter. Application of the proposed estimator to simulated and real-life data reveals that the resulting estimates are less biased and efficient. Unlike the ML estimator, the proposed MVU estimator does not over or under estimate the parameter. However, to implement the proposed method, the value of the dispersion parameter ν should be known. In case it is not available, the same can be estimated using the ratio of the sample mean to the sample variance or by the method of moments.

Acknowledgements

The authors are grateful to the anonymous referees for their suggestions and comments, which helped in improvising the paper.

References

Boswell, M. and Patil, G. (1973). Characterization of certain discrete distributions by differential equations with respect to their parameters. *Australian Journal of Statistics*, **15**, 128–131.

Chakraborty, S. and Imoto, T. (2016). Extended Conway-Maxwell-Poisson distribution and its properties and applications. *Journal of Statistical Distributions and Applications*, **3**, 1–19.

Conway, R. and Maxwell, W. L. (1961). A queuing model with state-dependent service rate. *Journal of Industrial Engineering*, **12**, 132–136.

Cordeiro, G. M., Rodrigues, J. and de Castro, M. (2012). The exponential COM-Poisson distribution. *Statistical Papers*, **53**, 653-664.

Dacey, M. F. (1972). A family of discrete probability distributions defined by the generalized hypergeometric series. *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, **34**, 243–250.

Daly, F. and Gaunt, R. (2016). The Conway-Maxwell-Poisson distribution: Distributional theory and approximation. *Latin American Journal of Probability and Mathematical Statistics*, **13**, 635–658.

Kimberly, S., Thomas, L. and Andrew, R. (2019). COMPoissonReg: Conway-Maxwell Poisson (COM-Poisson) Regression. R package version 0.7.0. URL https://CRAN.R-project.org/package=COMPoissonReg.

Li, B., Zhang, H. and He, J. (2019). Some characterizations and properties of COM-Poisson random variables. *Communications in Statistics - Theory and Methods*, **49**, 1311–1329.

Long, J.S. (1997). *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA, Sage Publications.

Nadarajah, S. (2009). Useful moment and cdf formulations for the COM–Poisson distribution. *Statistical Papers*, **50**, 617–622.

Roy, J. and Mitra, S. K. (1957). Unbiased minimum variance estimation in a class of discrete distributions. *Sankhyā: The Indian Journal of Statistics(1933-1960)*, **18**, 371–378.

Roy, S., Tripathi, R. C. and Balakrishnan, N. (2020). A Conway Maxwell Poisson type generalization of the negative hypergeometric distribution. *Communications in Statistics-Theory and Methods*, **49**, 2410–2428.

Sellers, K. F., Borle, S. and Shmueli, G. (2012). The COM-Poisson model for count data: a survey of methods and applications. *Applied Stochastic Models in Business and Industry*, **28**, 104–116.

Sellers, K. F., Swift, A. W. and Weems, K. S. (2017). A flexible distribution class for count data. *Journal of Statistical Distributions and Applications*, **4**, 1–21.

Shmueli, G., Minka, T. P., Kadane, J. B., Borle, S. and Boatwright, P. (2005). A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **54**, 127–142.