# When Is the Conway-Maxwell-Poisson Distribution Infinitely Divisible?

Xi Geng[*1] and Aihua Xia[†1]

[1]School of Mathematics and Statistics, the University of Melbourne, Parkville VIC 3010, Australia

**Abstract**

An essential character for a distribution to play a central role in the limit theory is infinite divisibility. In this note, we prove that the Conway-Maxwell-Poisson (CMP) distribution is infinitely divisible iff it is the Poisson or geometric distribution. This explains that, despite its applications in a wide range of fields, there is no theoretical foundation for the CMP distribution to be a natural candidate for the law of small numbers.

## 1 Introduction and the main result

The fundamental driving force for the success of modelling count data is the law of small numbers and the Poisson distribution is undoubtedly the cornerstone of the approximation theory in capturing the distribution of counts of rare events [Barbour, Holst and Janson (1992)]. The main disadvantage of the Poisson distribution

---

[*]email: xi.geng@unimelb.edu.au.

is that its mean and variance are identical, so when the count data is over-dispersed (resp. under-dispersed), i.e. its variance is bigger (resp. less) than its mean, the Poisson distribution is inadequate in fitting the data. There are many alternatives aiming to overcome the deficiency, e.g., compound Poisson (including negative binomial), translated Poisson [Barbour and Xia (1999), Röllin (2005)] and convolutions of infinitely divisible distributions [Presman (1983), Kruopis (1986)]. By introducing an extra parameter to the Poisson distribution for looking after the dispersion behaviour of the count data, the *Conway-Maxwell-Poisson distribution* CMP$(\lambda, \nu)$ [Conway and Maxwell (1962)] assumes the probability mass function

$$\mathbb{P}(X = k) = \frac{1}{Z(\lambda, \nu)} \cdot \frac{\lambda^k}{(k!)^\nu}, \ \ k = 0, 1, 2, \cdots,$$

where the parameters $(\lambda, \nu)$ satisfy $\lambda, \nu > 0$ or $\nu = 0, \lambda \in (0, 1)$, and $Z(\lambda, \nu)$ denotes the normalising constant. The CMP distribution is over-dispersed if $\nu < 1$ and under-dispersed if $\nu > 1$. Due to its smooth transition between over-dispersion and under-dispersion, it plays a significant role in modelling count data [Hilbe (2014)] and has an extraordinarily diverse range of applications, see [Sellers, Borle and Shmueli (2012)] for a brief survey. However, despite some initiatives [Daly and Gaunt (2016), Li, Zhang and He (2020)], there is disproportionately little advance in its analytical properties and an approximation theory based on this popular distribution. This note aims to explain the fundamental reason behind the unbalanced development.

An essential character for a distribution to play a central role in the limit theory is infinite divisibility. For the CMP distribution, we have the following result.

**Theorem 1.** *The distribution* CMP$(\lambda, \nu)$ *is infinitely divisible if and only if* $\nu = 0$ *or* $\nu = 1$.

In other words, the CMP distribution is infinitely divisible iff it is the Poisson or geometric distribution. This explains that, despite its applications in a wide range of fields, the only member in the CMP family that can capture the distribution of the count of weakly dependent rare events is the Poisson distribution and there is no theoretical foundation for the CMP distribution to be a natural candidate for the law of small numbers. In particular, Theorem 1 implies that the CMP process possessing independent stationary increments with $\nu > 0$ in [Zhu *et al.* (2017)] does not exist except it is a Poisson process.

# 2 The proof

The sufficiency part requires no comment. In addition, the distribution on non-negative integers is infinitely divisible if and only if it is a compound Poisson [Feller (1968), p. 290], and as such compound Poisson cannot be under-dispersed (see subsection 2.1), we know that $\mathrm{CMP}(\lambda, \nu)$ is not infinitely divisible when $\nu > 1$. By comparing with the probability mass function of a compound Poisson distribution, [Mao (2020)] concludes that $\mathrm{CMP}(\lambda, \nu)$ is not infinitely divisible when $\nu \in (0.33, 1)$. However, as $\nu$ edges towards 0, the elementary approach of comparing the probability mass functions offers little hope in determining infinite divisibility of the CMP distribution, we need an entirely different approach to tackle the problem. Our strategy is to use complex analysis.

Throughout the rest, let $\nu \in (0, 1)$ and $\lambda > 0$ be given. We assume on the contrary that $X \sim \mathrm{CMP}(\lambda, \nu)$ is infinitely divisible. The main steps towards reaching a contradiction are summarised as follows.

1. Under the assumption, $X$ must be a compound Poisson random variable, say

$$X = \sum_{n=1}^{N} Y_n,$$

   where $N$ is Poisson distributed and $\{Y_1, Y_2, \cdots\}$ are i.i.d. positive integer-valued random variables and independent of $N$.

2. The probability generating function $G(z)$ of $Y_1$ must be an entire function, i.e. holomorphic on the entire complex plane.

3. The exponential-type growth property of the probability generating function $F(z)$ of $X$ forces $G(z)$ to have no more than polynomial growth.

4. $G(z)$ must be a polynomial as a consequence of complex analysis.

5. The precise growth estimate for $F(z)$ obtained in Step 3 further forces $G(z)$ to be a monomial, which then leads to a contradiction trivially.

In the rest of this note, we develop the above steps carefully.

## 2.1 The compound Poisson decomposition and related probability generating functions

Under the assumption of infinite divisibility, we know from [Feller (1968), p. 290] that $X$ is compound Poisson, i.e.

$$X \stackrel{d}{=} \sum_{n=1}^{N} Y_n,$$

where $N \sim \mathrm{Pn}(\mu)$ with some $\mu > 0$, $\{Y_1, Y_2, \cdots\}$ is an i.i.d. sequence of $\mathbb{N} \triangleq \{0, 1, 2, \dots\}$-valued random variables that is independent of $N$. By absorbing the mass of $Y_1$ at the origin to the parameter of $N$ if necessary, we may assume that $Y_1$ takes values in the positive integers $\mathbb{N}_+ \triangleq \{1, 2, \dots\}$. Note that in this case we have

$$\mathbb{V}[X] - \mathbb{E}[X] = \mathbb{E}[N]\big(\mathbb{E}[Y_1^2] - \mathbb{E}[Y_1]\big) \geqslant 0.$$

This already implies that $\nu$ cannot be greater than 1 if $\mathrm{CMP}(\lambda, \nu)$ were infinitely divisible.

We now write

$$q_k \triangleq \mathbb{P}(Y_1 = k), \quad k \in \mathbb{N}_+,$$

and let $G(z)$ denote the probability generating function of $Y_1$, i.e.

$$G(z) = \sum_{r=1}^{\infty} q_r z^r. \tag{2.1}$$

We treat $z$ as a complex variable and note that $G(z)$ is holomorphic at least in the unit disk. From elementary probability theory, $e^{\mu(G(z)-1)}$ is the probability generating function of $X$. By using the CMP-distribution of $X$, we have

$$e^{\mu G(z)} = \frac{e^{\mu}}{Z(\lambda, \nu)} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{(k!)^{\nu}} \tag{2.2}$$

for those $z$'s within the radius of convergence of $G(z)$. By taking $z = 0$, we get $e^{\mu} = Z(\lambda, \nu)$ and thus

$$e^{\mu G(z)} = \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{(k!)^{\nu}}.$$

Let us set

$$F(z) \triangleq \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{(k!)^{\nu}}.$$

Note that $F(z)$ defines an entire function.

4

## 2.2  Holomorphicity of $G(z)$ on $\mathbb{C}$

Our next step is to show that $G(z)$ must be an entire function. Before doing so, we first recall the construction of the logarithm of a holomorphic function (cf. [Lang (1999), p. 123]). Suppose that $f(z)$ is a holomorphic function on a simply connected domain $\Omega$ and is everywhere non-zero, then there exists a holomorphic function $L(z)$ on $\Omega$ such that

$$e^{L(z)} = f(z). \tag{2.3}$$

Indeed, the equation (2.3) suggests that $L'(z) = \frac{f'(z)}{f(z)}$, which leads us to defining

$$L(z) \triangleq \int_{z_0}^{z} \frac{f'(w)}{f(w)} dw, \quad z \in \Omega,$$

where $z_0$ is a fixed based point in $\Omega$ and the integral is performed along an arbitrary path joining $z_0$ to $z$. The well-definedness of $L(z)$ is a simple consequence of the simply connectedness of $\Omega$ and Cauchy's theorem.

**Lemma 1.** *The function $G(z)$ defined by the power series (2.1) is an entire function.*

*Proof.* Let $R$ be the radius of convergence of $G(z)$. Suppose on the contrary that $R < \infty$. We know that

$$e^{\mu G(z)} = F(z), \quad z \in B_R \triangleq \{z : |z| < R\}.$$

In particular,

$$e^{\mu G(\rho)} = F(\rho), \quad \rho \in (0, R).$$

Since the coefficients of $G(z)$ are non-negative, by the monotone convergence theorem we have

$$e^{\mu G(R)} = F(R) < \infty.$$

It follows that $G(z)$ is convergent on the entire boundary of $B_R$. The dominated convergence theorem further implies that $G(z)$ is continuous on the closed ball $\overline{B_R}$ and we thus have

$$e^{\mu G(z)} = F(z), \quad z \in \overline{B_R}. \tag{2.4}$$

Now let $U \triangleq \{z \in \mathbb{C} : F(z) \neq 0\}$. Then $U$ is an open subset of $\mathbb{C}$. Since the exponential function is everywhere non-vanishing, it follows from (2.4) that $\overline{B_R} \subseteq U$, and hence $B_{R+\varepsilon} \subseteq U$ for some $\varepsilon > 0$. As $B_{R+\varepsilon}$ is simply connected, there

is a well-defined logarithm of $F(z)$ on $B_{R+\varepsilon}$, namley a holomorphic function $L(z)$ such that

$$e^{L(z)} = F(z), \quad z \in B_{R+\varepsilon}.$$

Combining with (2.4), we obtain

$$L(z) = \mu G(z) + 2\pi i k(z), \quad z \in B_R$$

with some function $k : B_R \to \mathbb{Z} \triangleq \{0, \pm 1, \pm 2, \dots\}$. Since both $L(z)$ and $G(z)$ are continuous, the function $k(z)$ must be constant (say $k(z) \equiv k^*$) and we arrive at

$$L(z) = \mu G(z) + 2\pi k^* i = 2\pi k^* i + \sum_{r=1}^{\infty} \mu q_r z^r.$$

The power series on the right hand side gives the Taylor expansion of $L(z)$. Since $L(z)$ is holomorphic on $B_{R+\varepsilon}$, its radius of convergence must be at least $R + \varepsilon$. This contradicts the assumption that $R$ is the radius of convergence for $G(z)$. Therefore, the series (2.1) is convergent on the entire complex plane and $G(z)$ is thus an entire function. $\qquad\square$

## 2.3 Precise growth-type estimate of $F(z)$

In this part, we investigate the precise growth of $F(z)$. This is the core step of the argument.

We first prepare a simple analytical lemma.

**Lemma 2.** *There exist positive constants $C_1, C_2$ and $K$ depending only on $\nu$, such that*

$$C_1 k^{\frac{1-\nu}{2}} \nu^{\nu k} \leqslant \frac{(k\nu)!}{(k!)^{\nu}} \leqslant C_2 k^{\frac{1-\nu}{2}} \nu^{\nu k} \tag{2.5}$$

*for all $k \geqslant K$.*

*Proof.* We recall the following Stirling's approximation for the Gamma function for all positive real $x$:

$$x! \triangleq \Gamma(x+1) \sim \sqrt{2\pi x} \left(\frac{x}{e}\right)^x \quad \text{as } x \to \infty. \tag{2.6}$$

As a result, we have

$$\frac{(k\nu)!}{(k!)^{\nu}} \sim \sqrt{\nu}(2\pi k)^{\frac{1-\nu}{2}} \nu^{\nu k} \quad \text{as } k \to \infty,$$

and the claim thus follows. $\qquad\square$

6

The main result for this part is stated below. It quantifies the precise growth rate of $F(z)$ as $z$ approaches infinity along the positive axis.

**Lemma 3.** *Let $M \triangleq \lambda \nu^\nu$. There exists a constant $C > 0$ as well as two polynomials $p_1, p_2$ with positive coefficients such that*

$$\frac{C}{R^{1/\nu}} e^{(MR)^{1/\nu}} - p_1(R) \leqslant F(R) \leqslant p_2(R) e^{(MR)^{1/\nu}} \qquad (2.7)$$

*for all $R$ with $MR > 1$.*

*Proof.* We first establish the upper bound. Let $K$ be as in Lemma 2. Enlarging $K$ if necessary, we assume $K\nu \geq 2$. For $k \geq K$, the right hand side of (2.5) ensures

$$\frac{(\lambda R)^k}{(k!)^\nu} \leq C_2 \frac{(MR)^k}{(k\nu)!} k^{\frac{1-\nu}{2}} \leq C_3 \frac{(MR)^k}{(k\nu - 1)!},$$

where $C_3 \triangleq C_2/\nu$. Hence,

$$F(R) = \sum_{k=0}^{K-1} \frac{(\lambda R)^k}{(k!)^\nu} + \sum_{k=K}^{\infty} \frac{(\lambda R)^k}{(k!)^\nu} \leqslant \sum_{k=0}^{K-1} \frac{(\lambda R)^k}{(k!)^\nu} + C_3 \sum_{k=K}^{\infty} \frac{(MR)^k}{(k\nu - 1)!}. \qquad (2.8)$$

We now examine the last summation in the above inequality by introducing the division $k = mp + r$ where $p \triangleq 1/\nu$, $m \in \mathbb{N}$ and $r \in [0, p)$. To be more precise, for each $m \in \mathbb{N}$ we set

$$R_m \triangleq \{r \in [0, p) : mp + r \in \mathbb{Z}\}.$$

Note that $R_m \neq \emptyset$ (since $p > 1$) and contains at most $[p] + 1$ elements, where $[p]$ denotes the integer part of $p$. It is clear that each $k \in \mathbb{N}$ can be written as $k = mp + r$ with some $m \in \mathbb{N}$ and $r \in R_m$. To see the uniqueness of such decomposition, suppose that

$$mp + r = m'p + r'.$$

Then

$$(m - m')p = r' - r \in (-p, p).$$

As a result, $m = m'$ and $r = r'$. It follows that $k \leftrightarrow (m, r)$ is a one-to-one correspondence. By using this decomposition, we have

$$\sum_{k=K}^{\infty} \frac{(MR)^k}{(k\nu - 1)!} \leq \sum_{m=1}^{\infty} \sum_{r \in R_m} \frac{(MR)^{mp+r}}{(m + r\nu - 1)!}$$

$$\leqslant \sum_{m=1}^{\infty} \frac{(MR)^{mp}}{(m-1)!} \sum_{r \in R_m} (MR)^r$$

$$\leqslant ([p] + 1)(MR)^{2p} e^{(MR)^p},$$

7

provided that $MR > 1$. Since the second last summation in (2.8) is polynomial, the desired upper bound follows.

The idea of establishing the lower bound is similar. By using Lemma 2, we have

$$F(R) \geqslant C_1 \sum_{k=K}^{\infty} \frac{k^{\frac{1-\nu}{2}}(MR)^k}{(k\nu)!} \geqslant C_1 \sum_{k=K}^{\infty} \frac{(MR)^k}{(k\nu)!}$$

$$= C_1 \left( \sum_{k=0}^{\infty} \frac{(MR)^k}{(k\nu)!} - \sum_{k=0}^{K-1} \frac{(MR)^k}{(k\nu)!} \right)$$

$$= C_1 \sum_{m=0}^{\infty} \sum_{r \in R_m} \frac{(MR)^{mp+r}}{(m+r\nu)!} - p_1(R),$$

where

$$p_1(R) \triangleq C_1 \sum_{k=0}^{K-1} \frac{(MR)^k}{(k\nu)!}.$$

To estimate the double summation, we observe that (assuming $MR > 1$)

$$\sum_{r \in R_m} \frac{(MR)^{mp+r}}{(m+r\nu)!} \geqslant \frac{(MR)^{mp}}{(m+1)!} \sum_{r \in R_m} (MR)^r \geqslant \frac{(MR)^{mp}}{(m+1)!}$$

where the last inequality follows from the fact that $R_m \neq \emptyset$. Therefore,

$$\sum_{m=0}^{\infty} \sum_{r \in R_m} \frac{(MR)^{mp+r}}{(m+r\nu)!} \geqslant \sum_{m=0}^{\infty} \frac{(MR)^{mp}}{(m+1)!} = (MR)^{-p} \left( e^{(MR)^p} - 1 \right).$$

The desired lower bound thus follows.

$\square$

## 2.4 The function $G(z)$ is a polynomial

Let $d \geqslant 2$ be the unique integer such that $1/\nu \in (d-1, d]$. It follows from Lemma 3 that

$$F(R) \leqslant e^{M'R^d} \quad \text{for } R \text{ sufficiently large}$$

with some $M' > 0$. On the other hand, since $G(z)$ has non-negative coefficients, we have

$$e^{\mu|G(z)|} \leqslant e^{\mu G(|z|)} = F(|z|).$$

8

As a result, we have

$$|G(z)| \leqslant \frac{1}{\mu} \log F(|z|) \leqslant \frac{M'}{\mu}|z|^d \quad \text{for } z \text{ large.}$$

Since $G(z)$ is an entire function, the following complex analysis lemma implies that $G(z)$ has to be a polynomial.

**Lemma 4.** *Let $g(z)$ be an entire function such that*

$$|g(z)| \leqslant C|z|^d$$

*for all large $z$. Then $g(z)$ is a polynomial of degree at most $d$.*

*Proof.* We write

$$g(z) = \sum_{k=0}^{\infty} a_k z^k, \quad z \in \mathbb{C}.$$

From Cauchy's integral formula, the Taylor coefficients are given by

$$a_k = \frac{1}{2\pi i} \int_{\partial B_R} \frac{g(z)}{z^{k+1}} dz = \frac{1}{2\pi} \int_0^{2\pi} \frac{g(Re^{i\theta})}{(Re^{i\theta})^k} d\theta, \quad k = 0, 1, 2, \cdots.$$

Note that the above formula is true for all $R > 0$. By the assumption, we have

$$|a_k| \leqslant \frac{C}{2\pi} \int_0^{2\pi} \frac{R^d}{R^k} d\theta = CR^{d-k}$$

for all large $R$. For each $k > d$, by taking $R \to \infty$ we conclude that $a_k = 0$. $\quad\square$

**Corollary 1.** *The function $G(z)$ is a polynomial of degree at most $d$, i.e. $G(z) = q_1 z + \cdots + q_d z^d$. In particular, $Y_1$ is supported on $\{1, \cdots, d\}$.*

## 2.5 Reaching the contradiction

We are now in a position to complete the proof of Theorem 1. The main point is that the growth estimate (2.7) of $F(R)$ forces the polynomial $G(z)$ to consist of the single monomial $q_d z^d$ only, which then trivially leads to a contradiction since $X$ achieves all possible values in $\mathbb{N}$. Recall that $d \geqslant 2$ is the unique integer such that $1/\nu \in (d-1, d]$.

9

*Proof of Theorem 1.* According to Lemma 3 and Corollary 1, we have

$$\frac{C}{R^{1/\nu}}e^{(MR)^{1/\nu}} - p_1(R) \leqslant e^{\mu(q_1 R + \cdots + q_d R^d)} \leqslant p_2(R)e^{(MR)^{1/\nu}} \qquad (2.9)$$

for all sufficiently large $R$. The lower bound forces $q_d$ to be non-zero. In the case when $\nu \neq 1/d$, the upper bound in (2.9) cannot hold true when $R \to \infty$, giving a contradiction. It now remains to consider the case when $\nu = 1/d$. In this case, the upper bound implies $\mu q_d \leqslant M^{1/\nu}$ while the lower bound implies that $\mu q_d \geqslant M^{1/\nu}$. Therefore, $\mu q_d = M^{1/\nu}$. If $q_1, \cdots, q_{d-1}$ were not all zero, the upper bound cannot hold true. As a result, we have $q_1 = \cdots = q_{d-1} = 0$. In particular, $q_d = 1$ and $G(z) = z^d$. Since $d \geqslant 2$, this clearly contradicts the fact that $\mathbb{P}(X = 1) > 0$.

Now the proof of Theorem 1 is complete.

$\square$

# References

[Barbour, Holst and Janson (1992)] A. D. Barbour, L. Holst and S. Janson, *Poisson approximation*, Oxford Univ. Press, 1992.

[Barbour and Xia (1999)] A. D. Barbour and A. Xia, *Poisson perturbations*, European Series in Applied and Industrial Mathematics: Probability and Statistics **3** (1999), 131–150.

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

[Daly and Gaunt (2016)] F. Daly and R. E. Gaunt, *The Conway-Maxwell-Poisson distribution: distributional theory and approximation*, ALEA Latin American Journal of Probabability and Mathematical Statistics **13** (2016), 635–658.

[Feller (1968)] W. Feller, *An introduction to probability theory and its applications I*, Third Edition, John Wiley & Sons, 1968.

[Hilbe (2014)] J. M. Hilbe, *Modeling Count Data*, Cambridge University Press, 2014.

[Kruopis (1986)] J. Kruopis, *Precision of approximations of the generalized Binomial distribution by convolutions of Poisson measures*, Lithuanian Math. J. **26** (1986), 37–49.

[Lang (1999)] S. Lang, *Complex analysis*, Graduate Texts in Mathematics, Springer-Verlag, 1999.

[Li, Zhang and He (2020)] B. Li, H. Zhang and J. He, *Some characterizations and properties of COM-Poisson random variables*, Communications in Statistics - Theory and Methods **49** (2020), 1311–1329.

[Mao (2020)] C. Mao, *The Conway-Maxwell Poisson distribution – from infinite divisibility to its applications*, MSc Thesis, University of Melbourne, 2020.

[Presman (1983)] E. L. Presman, *Approximation of binomial distributions by infinitely divisible ones*, Theory. Probab. Appl. 28 (1983), 393–403.

[Röllin (2005)] A. Röllin, *Approximation of sums of conditionally independent variables by the translated Poisson distribution*, Bernoulli **11** (2005), 1115–1128.

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

[Zhang (2015)] H. Zhang, *Characterizations and Infinite Divisibility of Extended COM-Poisson Distribution*, International Journal of Statistical Distributions and Applications **1** (2015), 1–5.

[Zhu *et al.* (2017)] L. Zhu, K. F. Sellers, D. S. Morris and G. Shmueli, *Bridging the Gap: A Generalized Stochastic Process for Count Data*, Amer. Statist. **71** (2017), 71–80.