arXiv:2304.02624v1 [math.PR] 5 Apr 2023

关于康威–麦克斯韦–泊松分布无限可分性的评论

Greg Markowsky, Preet Patel

gmarkowsky@gmail.com ppat0019@student.monash.edu澳大利亚莫纳什大学数学系

2023年4月6日

摘要

在一篇近期发表的优雅论文中[2], Geng 和 Xia 解决了康威–麦克斯韦–泊松分布的无限可除性问题，主要运用了复分析中的若干结果。在本文中，我们展示了如何避开这些复分析方法，从而以完全初等的方式证明他们的结果。

康威–麦克斯韦–泊松分布是定义在非负整数上的离散概率分布，其概率质量函数为

$$P(X=k)=\frac{1}{Z(\lambda,\nu)}\frac{\lambda^k}{(k!)^\nu}. \tag{1}$$

其中，λ、$\nu>0$，或 $\nu=0$和 $0<\lambda<1$，而 $Z(\lambda,\nu)$ 是使得$\sum_{k=0}^{\infty}P(X=k)=1$成立的正确常数。该分布已广泛用于计数数据建模及其他应用；参见[1,3–7]。最近，[2], 证明了以下结果。

定理 0.1。由（*1*）定义的分布具有无限可分性，当且仅当 $\nu=0$ 或$\nu=1$。

若要了解这一结果对于极限定理和小数定律的重要性，请读者参阅 [2]中的讨论。该文的证明在多个地方用到了复分析方法。在这篇简短的注记中，我们将说明如何在需要时避开这些方法。

我们先假设 X服从 $CMP(\lambda,\mu)$所给的分布；也就是说，其概率质量函数由（1）给出。若假设该分布无限可分，那么正如[2], 的作者所指出的，它必为复合泊松分布，并且存在一个独立同分布序列 $Y_1,Y_2,\ldots$，其随机变量满足

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

其中 $N \sim Pois(\mu)$ $(\mu > 0)$ 与 Y 各项相互独立。若以 $G(x)$ 表示 Y_1 的概率生成函数，即，

$$G(x) = \sum_{r=1}^{\infty} P(Y_1 = r)x^r, \tag{2}$$

then

$$e^{\mu G(x)} = \sum_{k=0}^{\infty} \frac{(\lambda x)^k}{(k!)^\nu}. \tag{3}$$

现在，式（3）的一个有趣特征是，右侧的收敛半径显然为无穷大，而左侧只能保证在 $|x| < 1$ 时存在。文献 [2] 的作者在此步骤要求知晓级数（2）具有无穷收敛半径，因此他们转而使用复分析，将 x视为复变量。由于对于一般的幂级数，各项之间可能存在复杂的抵消，因此该论证需要相当精细；然而，注意到这里所涉及的所有级数都具有非负系数，因而排除了由抵消导致的困难，事情便可以简化，并且可以用下面的引理替代复数论证。

引理 0.2。Su　假设 $(q_n)_n$ 是一个满足 $\sum_{n=0}^{\infty} q_n < \infty$ 的非负实数序列。定义

$$G(x) = \sum_{n=0}^{\infty} q_n x^n \tag{4}$$

并令 $F(x) = e^{\mu G(x)}$ 对某个 $\mu > 0$成立，该函数在级数收敛的所有地方都有定义。则 F 是解析的，并且在零附近具有幂级数展开式 $F(x) = \sum_{n=0}^{\infty} t_n x^n$。假设这个关于 F 的级数具有无限收敛半径。那么定义 G 的级数也具有无限收敛半径。

证明。 该 μ 可吸收到 G中，且系数仍然非负，因此我们忽略μ。假设（4）的收敛半径有限，R （注意到 $R \geq 1$ 因为这些 q_n 的总和有限）。对于 $0 < x < R$，我们有

$$\sum_{n=0}^{\infty} q_n x^n = G(x) < e^{G(x)} = \sum_{k=0}^{\infty} \frac{\left(\sum_{n=0}^{\infty} q_n x^n\right)^k}{k!} \tag{5}$$

由于这个二重级数中的所有项均为正，且该级数在0 附近绝对且一致收敛，因此我们可以直接展开并重新排列，而无需谨慎处理；我们还记得，右侧的级数必须重新排列为 $\sum_{k=0}^{\infty} t_k x^k$。我们得出，每个 t_m 都可以表示为各个 q_n的线性组合，且这个线性组合必须包含来自 q_m 中 $k = 1$ 项的 5。我们得出 $q_m \leq t_m$，这足以说明 $R = \infty$。

□

[2] 的作者随后通过一系列概率论论证，旨在证明对于某个 C、$d > 0$ （d 为整数）以及充分大的 $|x|$ ， $|G(x)| \leq C|x|^d$ 。从复平面的角度来看，现在已知 G 是整函数；在复分析课程中，这是一个标准练习（李乌维尔定理的推广）：证明这意味着 G是

一个次数为 d的多项式。这个结果可以由柯西积分公式或其他一些密切相关的结果推导出来。然而，同样在我们的情形中，所有系数均为非负数这一事实使得可以用以下引理替代这一论证。

引理 0.3。假设 $G(x) = \sum_{n=0}^{\infty} q_n x^n$ 是一个幂级数，且对于所有 n，都有 $q_n \geq 0$ 。此外，假设该级数的收敛半径为无穷大，并且当 C、$d > 0$ 和 $|x|$ 充分大时，满足 $|G(x)| \leq C|x|^d$ 。那么，G 是次数至多为 d 的多项式。

证明。假设存在某个非零的 q_m ，使得 $m > d$。对于 $x \geq 0$ ，我们有 $G(x) \geq q_m x^m$；然而，由于 $m > d$ ，当 $C > 0$、 $n \leq d$ 和 x 充分大时，该量将超过任意常数 Cx^n 。引理由此得证。

□

致谢

作者谨向 Xi Geng 和 Aihua Xia 致谢，感谢他们的宝贵交流。

参考文献

[1] F. Daly 和 R. Gaunt。康威-麦克斯韦-泊松分布：分布理论与近似。《ALEA》，13：635–658，2016年。[2] X. Geng 和 A. Xia。康威–麦克斯韦–泊松分布何时具有无限可分性?《统计与概率快报》，181：109264，2022年。[3] J. Hilbe。《计数数据建模》。剑桥大学出版社，2014年。[4] B. Li、H. Zhang 和 Jiao He。康威-麦克斯韦-泊松随机变量的一些特征与性质。《统计学通讯：理论与方法》，49(6)：1311–1329，2020年。[5] T. Pog´any。康威-麦克斯韦-泊松归一化常数的积分形式。《统计与概率快报》，119：144–145，2016年。[6] K. Sellers、D. Borle 和 G. Shmueli。用于计数数据的康威-麦克斯韦-泊松模型：方法与应用综述。《商业与工业应用随机模型》，28(2)：104–116，2012年。[7] S. Yip、Y. Zou、R. Hung 和 K. Yiu。使用过度离散分布预测足球比赛中的角球数。arXiv 预印本 arXiv:2112.13001，2021年。