Filomat 30:4 (2016), 953–960
DOI 10.2298/FIL1604953S

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Approximating the Conway-Maxwell-Poisson normalizing constant

Burçin Şimşek and Satish Iyengar

[a]*Department of Statistics, University of Pittsburgh, USA*

Abstract. The Conway-Maxwell-Poisson is a two-parameter family of distributions on the nonnegative integers. Its parameters λ and ν model the intensity and the dispersion, respectively. Its normalizing constant is not always easy to compute, so good approximations are needed along with an assessment of their error. Shmueli, et al. [11] derived an approximation assuming that ν is an integer, and gave an estimate of the relative error. Their numerical work showed that their approximation performs well in some parameter ranges but not in others. Our aims are to show that this approximation applies to all real $\nu > 0$; to provide correction terms to this approximation; and to give different approximations for ν very small and very large. We then investigate the error terms numerically to assess our approximations. In parameter ranges for which Shmueli's approximation does poorly we show that our correction terms or alternative approximations give considerable improvement.

1. Introduction

The Poisson distribution is widely used to model count data. However, in many applications the equality of the mean and variance of the Poisson is often too restrictive. Overdispersion (underdispersion), where the variance is greater (less) than the mean, are commonly encountered in count data [3, 9]; in such cases, the Poisson model can be a poor fit. The Conway-Maxwell-Poisson family is a useful alternative in such cases. Conway and Maxwell [5] introduced it for studying queuing systems with state-dependent service rates. It has two parameters, making it flexible enough to fit a wider range of count data than the Poisson.

There were no systematic studies of the probabilistic or statistical properties of the Conway-Maxwell-Poisson family until the series of papers by Shmueli and her colleagues [10, 11]. They noted that the Conway-Maxwell-Poisson is an exponential family, can fit overdispersed and underdispersed data, and can be modified to account for zero-inflated data. They also provide the details of standard approaches to maximum likelihood and Bayes estimation of the parameters. Imoto [6] proposed a three-parameter generalized Conway-Maxwell-Poisson distribution which includes the negative binomial distribution as a special case.

2010 *Mathematics Subject Classification*. 60E05, 62E17

Keywords. asymptotic expansion, Gaussian approximation, generalized hypergeometric function, geometric distribution, modified Bessel function, relative error.

Received: 30 June 2015; Accepted: 30 July 2015

Communicated by Gradimir Milovanović

We thank Professor Tibor Pogány for pointing out the connection to the generalized hypergeometric functions. Burçin Şimşek's work was supported by a Andrew Mellon Predoctoral Fellowship at the University of Pittsburgh.

Email address: bus5@pitt.edu and ssi@pitt.edu (Burçin Şimşek and Satish Iyengar)

For integer $k \geq 0$, the Conway-Maxwell-Poisson variate Y has the probabilities

$$P(Y = k|\lambda, \nu) = C(\lambda, \nu)^{-1} \frac{\lambda^k}{(k!)^\nu}, \quad \text{where} \quad C(\lambda, \nu) = \sum_{k=0}^{\infty} \frac{\lambda^k}{(k!)^\nu}.$$

The parameter space is

$$\Theta = \{(\lambda, \nu) : \lambda > 0, 0 < \nu \leq \infty\} \cup \{(\lambda, \nu) : 0 < \lambda < 1, \nu = 0\}.$$

The Conway-Maxwell-Poisson includes three well-known distributions as special cases: geometric when $\nu = 0$ and $0 < \lambda < 1$; Poisson when $\nu = 1$; and Bernoulli with parameter $p = \lambda/(\lambda + 1)$ when $\nu = \infty$. Thus, it acts as a bridge between these distributions.

In this paper, we reparametrize this family with $\alpha = \lambda^{1/\nu}$ in place of λ. Thus,

$$P(Y = k|\alpha, \nu) = C(\alpha, \nu)^{-1} \left(\frac{\alpha^k}{k!}\right)^\nu, \quad \text{where} \quad C(\alpha, \nu) = \sum_{k=0}^{\infty} \left(\frac{\alpha^k}{k!}\right)^\nu.$$

The normalizing constant, $C(\alpha, \nu)$, plays an important role in the computations commonly used for fitting models to data: for computing probabilities, moments, and maximum likelihood estimates and their standard errors. Thus, it is of interest to get a good approximation for the normalizing constant. For the case of integer ν, Shmueli, et al. [11] expressed $C(\lambda, \nu)$ as a $(\nu - 1)$-dimensional integral involving complex exponentials and used Laplace's method to derive an approximation. Their numerical calculations showed that this approximation was good when $\lambda > 10^\nu$, or $\alpha > 10$.

Our aims in this paper are first to show that this approximation is in principle valid for all $\nu > 0$, not just integers; second, to use correction terms based on certain expansions to improve upon this approximation; and third, for ν near 0, to propose the use of the geometric approximation with correction terms. In Section 2, we sketch the approach of Shmueli, et al. [11] and discuss its limitations. In Section 3, we give a new derivation of the same approximation of $C(\alpha, \nu)$ by using well known statistical methods: in particular, we express $C(\alpha, \nu)$ as an expectation of a function of a Poisson variate, and use a Gaussian approximation to the square root of a Poisson for large α and all $\nu > 0$. In Section 4, we connect the normalizing constant to two special functions: the generalized hypergeometric function ${}_0F_{\nu-1}$ for integer ν, and the modified Bessel function of the first kind, I_0 for $\nu = 2$. The well-known asymptotic expansion for $I_0(z)$ for $|z| \to \infty$ points to the need for correction terms to the leading term. Thus, in Section 5, we compute the first two correction terms and assess them numerically. We show that we have an improvement over the approximation due to Shmueli, et al. [11], considerably so where the latter does poorly. In Section 6, we end with a discussion and propose future directions.

2. Laplace Approximation for Integer ν

Shmueli and her colleagues [11] derived an asymptotic approximation and upper bound for integer ν using Laplace's method on a $(\nu - 1)$-dimensional integral that represents $C(\lambda, \nu)$. A brief account of their steps and approximation: First, since $\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{e^{ix}} e^{-ixk} \, dx = \frac{1}{k!}$, we have

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{e^{ix}} C(\lambda e^{-ix}, \nu) \, dx = \sum_{k=0}^{\infty} \frac{\lambda^k}{(k!)^\nu} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{e^{ix}} e^{-ixk} \, dx = C(\lambda, \nu + 1).$$

This expresses $C(\lambda, \nu + 1)$ as an integral of $C(\lambda, \nu)$. Iterating this process starting from $C(\lambda, 1) = e^\lambda$ leads to a representation of $C(\lambda, \nu)$ for integer $\nu > 0$, as a multiple integral:

$$C(\lambda, \nu) = \frac{1}{(2\pi)^{\nu-1}} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \exp\left(\sum_{j=1}^{\nu-1} e^{ix_j} + \lambda \, e^{-\sum_{j=1}^{\nu-1} ix_j}\right) dx_1 \ldots dx_{\nu-1}.$$

Then, Laplace's method [2] applied to the multiple integral yields

$$C(\lambda,\nu)=\sum_{k=0}^{\infty}\frac{\lambda^k}{(k!)^\nu}=\frac{\exp\left(\nu\lambda^{1/\nu}\right)}{\lambda^{(\nu-1)/2\nu}(2\pi)^{(\nu-1)/2}\sqrt{\nu}}\left[1+O(\lambda^{-1/\nu})\right], \tag{1}$$

which in our parametrization is

$$C(\alpha,\nu)=\frac{e^{\nu\alpha}}{(2\pi\alpha)^{\frac{\nu-1}{2}}\sqrt{\nu}}\left[1+O(\alpha^{-1})\right] \quad \text{as} \quad \alpha=\lambda^{1/\nu}\to\infty. \tag{2}$$

Notice that this approximation requires $\nu > 0$; we will see at the end of Section 5 below that when $\nu < 1$ and $\lambda < 1$ an approximation based on the geometric distribution with correction terms is a better approximation. Shmueli and her colleagues [11] also present results of numerical examples to check the relative error of their approximation. Their numerical studies show that the leading term in the asymptotic expression in (1) is a good approximation for certain parameter ranges; however for $\nu > 1$ it consistently underestimates the true value of the $C(\alpha,\nu)$. In addition, their results for non-integer values of ν suggests that this approximation also applies to all real ν. An investigation into other aspects of this family by Nadarajah [8] also assumes that ν is an integer, and connects this series to the generalized hypergeometric functions, which we discuss in Section 4. In the next section, we use a statistical approach to derive the asymptotic approximation with the same leading term: our argument works for all $\nu > 0$; it also provides explicit expressions for the first two terms of an asymptotic expansion which can be used to attempt to improve underestimation by the leading term.

3. A Statistical Approach for all $\nu > 0$

We first describe the steps of our approach. We write the normalizing constant as the expectation of a function of X, a Poisson(α) variate, and use Stirling's formula to approximate that function. Next, we approximate $2\sqrt{X}$ by a normal distribution for large values of α; we then use elementary expansions for the log gamma and log functions, and end by taking the expectation of the resulting expression.

STEP 1. EXPRESS NORMALIZING CONSTANT AS AN EXPECTATION

Notice that

$$C(\alpha,\nu)=\sum_{k=0}^{\infty}\left(\frac{\alpha^k}{k!}\right)^{\nu}=e^{\alpha}\sum_{k=0}^{\infty}e^{-\alpha}\frac{\alpha^k}{k!}\left(\frac{\alpha^k}{k!}\right)^{\nu-1}=e^{\alpha}E_\alpha\left(\frac{\alpha^X}{X!}\right)^{\nu-1}. \tag{3}$$

When α is large, X will be large with high probability, with the most likely value near α. Stirling's approximation [1] says that

$$\Gamma(\alpha+1)\sim\alpha^{\alpha}e^{-\alpha}\sqrt{2\pi\alpha} \quad \text{as } \alpha\to\infty.$$

Using Stirling's approximation in (3), we get

$$C(\alpha,\nu)=e^{\alpha}E_\alpha\left(\frac{\alpha^X}{\alpha^\alpha e^{-\alpha}\sqrt{2\pi\alpha}}\frac{\alpha^\alpha e^{-\alpha}\sqrt{2\pi\alpha}}{X!}\right)^{\nu-1}=\frac{e^{\nu\alpha}}{(2\pi\alpha)^{\frac{\nu-1}{2}}}E_\alpha\left(\frac{\alpha^X e^{-\alpha}\sqrt{2\pi\alpha}}{\Gamma(X+1)}\right)^{\nu-1}=\frac{e^{\nu\alpha}}{(2\pi\alpha)^{\frac{\nu-1}{2}}}E_\alpha[U(\alpha,X)]. \tag{4}$$

Note that the constant term here is the same as the asymptotic expression in (1) above, except for the $\sqrt{\nu}$ term in the denominator. To show that $E_\alpha[U(\alpha,X)]$ is approximately $\nu^{-1/2}$ and assess the order of magnitude of the error, we study $U(\alpha,X)$ or equivalently

$$\ln[U(\alpha,X)]=(\nu-1)\left[\left(X+\frac{1}{2}\right)\ln\alpha-\alpha+\frac{\ln(2\pi)}{2}-\ln\Gamma(X+1)\right]. \tag{5}$$

STEP 2. USE NORMAL APPROXIMATION TO $2\sqrt{X}$ AND EXPANSIONS

Recall that for large α, the distribution of $2\sqrt{X}$ is approximately normal with mean $2\sqrt{\alpha}$ and variance 1. There are better variance stabilizing transformations such as Anscombe's $2\sqrt{X+3/8}$ (see [7]) but we use the simpler expression here. Writing Z for a standard normal variate, we have

$$2\sqrt{X} \sim (Z + 2\sqrt{\alpha}) \quad \text{so that} \quad X \sim \frac{(Z+2\sqrt{\alpha})^2}{4} = \alpha + Z\sqrt{\alpha} + \frac{Z^2}{4}.$$

Using this normal approximation, we obtain

$$\frac{\ln[U(\alpha,X)]}{(\nu-1)} \simeq \left[\left(\alpha + Z\sqrt{\alpha} + \frac{Z^2}{4} + \frac{1}{2}\right)\ln(\alpha) - \alpha + \frac{\ln(2\pi)}{2}\right] - \ln\Gamma\left(\alpha + Z\sqrt{\alpha} + \frac{Z^2}{4} + 1\right) = A - B. \tag{6}$$

In (6), we leave A as is; for B, we use the first three terms of the following expansions

$$\ln\Gamma(x+1) = \left(x + \frac{1}{2}\right)\ln(x+1) - (x+1) + \frac{\ln(2\pi)}{2} + O(x^{-1}) \quad \text{for} \quad x \to \infty,$$

$$\ln(1+t) = t - \frac{t^2}{2} + \frac{t^3}{3} + \dots \quad \text{for} \quad |t| < 1$$

to obtain

$$B = \left(\alpha + Z\sqrt{\alpha} + \frac{Z^2}{4} + \frac{1}{2}\right)\ln\left(\alpha + Z\sqrt{\alpha} + \frac{Z^2}{4} + 1\right) - \left(\alpha + Z\sqrt{\alpha} + \frac{Z^2}{4} + 1\right) + \frac{\ln(2\pi)}{2} \tag{7}$$

$$= \left(\alpha + Z\sqrt{\alpha} + \frac{Z^2}{4} + \frac{1}{2}\right)\ln(\alpha) + \left(\alpha + Z\sqrt{\alpha} + \frac{Z^2}{4} + \frac{1}{2}\right)\ln\left(1 + \frac{Z}{\sqrt{\alpha}} + \frac{Z^2}{4\alpha} + \frac{1}{\alpha}\right) - \left(\alpha + Z\sqrt{\alpha} + \frac{Z^2}{4} + 1\right) + \frac{\ln(2\pi)}{2}.$$

Combining (6) and (7), we get

$$\frac{\ln[U(\alpha,X)]}{\nu-1} = A - B \simeq -\frac{Z^2}{2} - \frac{6Z + Z^3}{12\sqrt{\alpha}} + O(\alpha^{-1}).$$

Then, using $e^x \simeq 1 + x$ for small x, we have

$$\begin{aligned} U(\alpha,X) &\simeq \exp\left[-(\nu-1)\frac{Z^2}{2}\right]\exp\left[-(\nu-1)\frac{6Z+Z^3}{12\sqrt{\alpha}}\right]\exp\left[O(\alpha^{-1})\right] \\ &= \exp\left[-(\nu-1)\frac{Z^2}{2}\right]\left[1 - (\nu-1)\frac{6Z+Z^3}{12\sqrt{\alpha}}\right][1 + O(\alpha^{-1})]. \end{aligned} \tag{8}$$

STEP 3. TAKE EXPECTATIONS

Finally, subsitute (8) in (4), and notice that $6Z + Z^3$ is an odd function, so its expectation is 0, to get

$$E_\alpha[U(\alpha,X)] \simeq \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-(\nu-1)z^2/2}\left[1 - (\nu-1)\frac{6z+z^3}{12\sqrt{\alpha}}\right]e^{-z^2/2}[1 + O(\alpha^{-1})]\,dz = \frac{1}{\sqrt{\nu}}[1 + O(\alpha^{-1})],$$

so that

$$C(\alpha,\nu) = e^{\alpha}E_\alpha\left(\frac{\alpha^X}{X!}\right)^{\nu-1} = \frac{e^{\nu\alpha}}{(2\pi\alpha)^{\frac{\nu-1}{2}}\sqrt{\nu}}[1 + O(\alpha^{-1})] \text{ as } \alpha \to \infty, \tag{9}$$

which is the same asymptotic expression as in (2). In summary, we have used a standard statistical approach and elementary expansions to show that the expression is valid for all real $\nu > 0$.

4. Generalized Hypergeometric Function Approach

Nadarajah [8] has connected the Conway-Maxwell-Poisson family with the generalized hypergeometric (Fox-Wright) functions and studied the moments and cumulative distribution functions; however, she did not deal specifically with the problem of approximating the normalizing constant. Even though ν must be an integer for this connection to hold, we pursue it for illustration because the special case $\nu = 2$ motivates the use of correction terms. The generalized hypergeometric functions are

$$ {}_pF_q\left[\begin{matrix} (a_1A_1) & (a_2A_2) & \dots & (a_pA_p) \\ (b_1B_1) & (b_2B_2) & \dots & (b_qB_q) \end{matrix}; z\right] = \sum_{n=0}^{\infty} \frac{\prod_{i=1}^{p}\Gamma(a_i + A_i n)}{\prod_{j=1}^{q}\Gamma(b_j + B_j n)} \frac{z^n}{n!}, $$

where an empty product is defined as 1. When $p = 0$, $q = \nu - 1$, and $a_i = A_i = b_j = B_j = 1$ and set $z = \lambda$, we get

$$ C(\lambda,\nu) = {}_0F_{\nu-1}[-;1,\cdots,1;\lambda] \quad \text{or} \quad C(\alpha,\nu) = {}_0F_{\nu-1}[-;1,\cdots,1;\alpha^\nu]. $$

When $\nu = 2$, the generalized hypergeometric function reduces to the modified Bessel function of the first kind: $C(\alpha, 2) = I_0(2\alpha)$. Using the known asymptotic expansion [1]

$$ I_\eta(z) \sim \frac{e^z}{\sqrt{2\pi z}}\left(1 - \frac{4\eta^2-1}{8z} + \frac{(4\eta^2-1)(4\eta^2-9)}{2!(8z)^2} + O(z^{-3})\right) \quad \text{as } |z| \to \infty, $$

we have

$$ C(\alpha,2) = I_0(2\alpha) \sim \frac{e^{2\alpha}}{\sqrt{4\pi\alpha}}\left(1 + \frac{1}{16\alpha} + \frac{9}{512\alpha^2} + O(\alpha^{-3})\right) \tag{10} $$

as $\alpha \to \infty$. We see from this expansion that the estimate in (1) is most likely an underestimate of the normalizing constant because the first two correction terms of order α^{-1} and α^{-2} are both positive.

All relative errors in this paper are expressed in percents. Table 1 presents the relative errors of the two of the approximations in (10): the leading term only, and the leading term with first order correction. We note that Shmueli et al. [11] presented the relative errors for the leading term only for λ and ν values from 0.1 to 1.9. Based on their table, they stated that the approximation is not good when $\alpha = \lambda^{1/\nu} < 10$. Our illustration in Table 1 (using the λ parametrization) demonstrates that even in this range including a correction term improves the approximation considerably. Including the second correction term will in this case lessen the underestimate; however, that correction is typically much smaller than the first-order correction, so we omit it.

Table 1: Relative errors in (10) for leading term and for first-order correction; $\nu = 2$.

values of λ:	0.1	0.3	0.5	0.7	0.9	1.1	1.3	1.5	1.7	1.9
leading term	-14.4	-13.9	-11.9	-10.3	-9.1	-8.1	-7.4	-6.7	-6.2	-5.8
first correction	2.6	-4.0	-4.1	-3.6	-3.1	-2.6	-2.3	-2.0	-1.7	-1.5

This illustration shows the usefulness of correction terms. The literature [4] on the asymptotics of the generalized hypergeometric functions ${}_pF_q$ require integer values of p and q , which in our case restricts ν to be an integer. Therefore, in the next section we extend our statistical approach by including more terms in the expansions for $\ln(1 + t)$ and $\ln\Gamma(x + 1)$ to get correction terms of order α^{-1} and α^{-2} for all $\nu > 0$.

5. Correction Terms and Numerical Examples

DERIVING THE CORRECTION TERMS

To derive the correction terms, we use more terms in the expansions below to further study B in (6):

$$\ln\Gamma(x+1) = \left(x+\frac{1}{2}\right)\ln(x+1)-(x+1)+\frac{\ln(2\pi)}{2}+\frac{1}{12(x+1)}+O(x^{-3}) \quad \text{for} \quad x\to\infty$$

and

$$\ln(1+t) = \sum_{n=1}^{6}(-1)^{n+1}\frac{t^n}{n}+O(t^7) \quad \text{for} \quad |t|\to 0.$$

Following the same steps as in Section 3, we get

$$\frac{\ln U(\alpha,\nu)}{(\nu-1)} \simeq -\frac{Z^2}{2}-\frac{Z^3+6Z}{12\sqrt{\alpha}}+\frac{Z^4+12Z^2+8}{96\alpha}-\frac{Z^5+2Z^3+4Z}{48\alpha\sqrt{\alpha}}+\frac{Z^6+30Z^4+120Z^2}{1920\alpha^2}+O(\alpha^{-5/2}).$$

As in (8), we first exponentiate this expression, and expand all but the $e^{-(\nu-1)z^2/2}$ term in a two-term Taylor series; we note that the terms with odd powers of Z integrate to 0, so we omit them as before. The resulting expansion is

$$\begin{aligned} EU(\alpha,\nu) &= E\left\{e^{-(\nu-1)Z^2/2}\left[1+\frac{Z^4+12Z^2+8}{96\alpha}+\frac{Z^6+30Z^4+120Z^2}{1920\alpha^2}+O(\alpha^{-3})\right]\right\} \\ &= \frac{1}{\sqrt{\nu}}\left[1+\frac{\nu-1}{12\alpha}\left(\frac{3}{8\nu^2}+\frac{3}{2\nu}+1\right)+\frac{\nu-1}{16\alpha^2}\left(\frac{1}{8\nu^3}+\frac{3}{4\nu^2}+\frac{1}{\nu}\right)+O(\alpha^{-3})\right]. \end{aligned}$$

Because odd powers of Z integrate to zero, the error term here is $O(\alpha^{-3})$ rather than $O(\alpha^{-5/2})$. Thus, the normalizing constant with correction terms of order up to α^{-2} is

$$C(\alpha,\nu) = \frac{e^{\nu\alpha}}{(2\pi\alpha)^{\frac{\nu-1}{2}}\sqrt{\nu}}\left[1+(\nu-1)\left(\frac{8\nu^2+12\nu+3}{96\alpha\nu^2}+\frac{1+6\nu}{144\alpha^2\nu^3}\right)+O(\alpha^{-3})\right] \quad \text{as} \quad \alpha\to\infty. \tag{11}$$

Before turning to the details of the numerical illustration, we return to the special case $\nu = 2$: note that the leading term is the same as in (10); however the two correction terms in (11) are

$$\frac{1}{\alpha}\frac{59}{16(24)} \simeq \frac{0.1536}{\alpha} \quad \text{and} \quad \frac{1}{\alpha^2}\frac{13}{2(512)} \simeq \frac{0.0127}{\alpha^2}$$

instead of $1/16\alpha = 0.0625/\alpha$ and $9/512\alpha^2 = 0.0177/\alpha^2$ in (10). We attribute this difference to the fact that our statistical approach to the case for all $\nu > 0$ is different from the Laplace method approach for the modified Bessel function $I_\eta(z)$.

NUMERICAL RESULTS

In order to assess the performance of the approximations, we need to know the exact values of the normalizing constant. When α is small, we simply sum a finite number of terms in the series because it converges quickly. When α is very large, we sum the series around the modal value after factoring it out. In particular, it is easy to see that for $a_k = (\alpha^k/k!)^\nu$, then $a_{k+1}/a_k < 1$ if and only if $\alpha < k+1$, so that the modal probability is k_α is the integer nearest to $\alpha - 1$. Thus, we write

$$C(\alpha,\nu) = \sum_{k=0}^{\infty}\left(\frac{\alpha^k}{k!}\right)^\nu = \left(\frac{\alpha^{k_\alpha}}{k_\alpha!}\right)^\nu\sum_{k=0}^{\infty}\left(\frac{\alpha^{k-k_\alpha}k_\alpha!}{k!}\right)^\nu \simeq \left(\frac{\alpha^{k_\alpha}}{k_\alpha!}\right)^\nu\sum_{k_\alpha-l}^{k_\alpha+h}\left(\frac{\alpha^{k-k_\alpha}k_\alpha!}{k!}\right)^\nu,$$

where l and h can be determined by examining the successive differences, and stopping when those differences are sufficiently small.

From the table in Shmueli, et al. [11], we see that for $\nu > 1$ the leading term underestimates the true value. Note that for $\nu > 1$ our correction terms are all positive; thus, we focus on $\nu > 1$. From their work, we know that the leading term is a good approximation for $\alpha > 10$; hence we consider $\alpha \leq 10$, and assess the accuracy of the leading term and the first-order correction in Tables 2 and 3, respectively; we did not include the correction term of order α^{-2} because it is much smaller, hence less consequential. We see considerable improvement when we use the first-order correction, unless λ is very small, in which case a direct summation of the series is easy and accurate.

Table 2: Relative error using the leading term in (1)

	values of ν							
λ	1.1	1.3	1.5	1.7	1.9	2.1	2.3	2.5
0.1	0.1	-1.9	-5.2	-8.9	-12.6	-16.0	-19.1	-21.9
0.3	-1.4	-4.4	-7.4	-10.2	-12.7	-14.9	-16.9	-18.6
0.5	-1.5	-4.2	-6.8	-9.0	-11.0	-12.7	-14.3	-15.6
0.7	-1.3	-3.8	-6.0	-7.9	-9.5	-11.0	-12.3	-13.4
0.9	-1.2	-3.3	-5.2	-6.9	-8.4	-9.7	-10.9	-11.9
1.1	-1.0	-2.9	-4.6	-6.2	-7.5	-8.7	-9.8	-10.7
1.3	-0.9	-2.6	-4.1	-5.5	-6.8	-7.9	-8.9	-9.8
1.5	-0.8	-2.3	-3.7	-5.0	-6.2	-7.3	-8.2	-9.1
1.7	-0.7	-2.1	-3.4	-4.6	-5.7	-6.7	-7.7	-8.5
1.9	-0.6	-1.9	-3.1	-4.3	-5.3	-6.3	-7.2	-8.1
2.1	-0.6	-1.7	-2.9	-3.9	-4.9	-5.9	-6.8	-7.6
2.3	-0.5	-1.6	-2.7	-3.7	-4.7	-5.6	-6.5	-7.3
2.5	-0.5	-1.5	-2.5	-3.5	-4.4	-5.3	-6.2	-7.0

Table 3: Relative errors using first correction term in (11)

	values of ν							
λ	1.1	1.3	1.5	1.7	1.9	2.1	2.3	2.5
0.1	9.1	15.2	14.6	11.7	8.2	4.7	1.4	-1.6
0.3	1.9	2.8	1.9	0.5	-1.0	-2.5	-3.8	-4.9
0.5	0.6	0.6	-0.1	-1.0	-1.9	-2.7	-3.4	-4.0
0.7	0.19	-0.02	-0.6	-1.2	-1.8	-2.3	-2.8	-3.1
0.9	0.03	-0.2	-0.6	-1.1	-1.5	-1.9	-2.2	-2.4
1.1	-0.03	-0.3	-0.6	-0.9	-1.3	-1.5	-1.7	-1.8
1.3	-0.04	-0.3	-0.5	-0.8	-1.0	-1.2	-1.3	-1.4
1.5	-0.04	-0.2	-0.4	-0.6	-0.8	-0.9	-1.1	-1.1
1.7	-0.04	-0.2	-0.3	-0.5	-0.6	-0.8	-0.8	-0.9
1.9	-0.03	-0.1	-0.3	-0.4	-0.5	-0.6	-0.7	-0.7
2.1	-0.01	-0.1	-0.2	-0.3	-0.4	-0.5	-0.5	-0.5
2.3	-0.0	-0.1	-0.1	-0.2	-0.3	-0.4	-0.4	-0.4
2.5	0.0	-0.03	-0.1	-0.2	-0.2	-0.3	-0.3	-0.3

Next, we turn to the case where both ν and λ are small because the leading term in (1) can be very poor in that region. For small ν it is more natural to consider an expansion around the geometric distribution than the Poisson ($\nu = 1$) because the Conway-Maxwell-Poisson distribution reduces to the geometric when $\nu = 0$. In this case, the expansion is the much simpler Taylor expansion of $C(\lambda, \nu)$ about $\nu = 0$:

$$C(\lambda, \nu) = \frac{1}{1-\lambda} - \nu \sum_{k=0}^{\infty} \ln(k!)\lambda^k + O(\nu^2). \tag{12}$$

In Table 4, we compare leading term in (1) with the geometric (G0), and its linear (G1) correction term. The linear correction term already gives very accurate results, so we do not consider a quadratic correction term in (12).

Table 4: Comparison of relative errors for (1) and (12)

(λ, ν)	(.1,.1)	(.3,.1)	(.5,.1)	(.7,.1)	(.1,.3)	(.3,.3)	(.1,.5)	(.3,.5)	(.1,.7)
Eq (1)	-100.0	-97.7	-83.3	-49.1	-78.7	-38.4	-35.8	-6.8	-10.8
G0	0.1	1.0	4.4	16.4	0.2	2.6	0.3	3.8	0.4
G1	-0.0	-0.1	-0.9	-8.4	-4.0	-0.8	-0.1	-1.9	-0.2

Finally, when ν is very big, the denominator $(k!)^\nu$ grows rapidly, so the defining series for $C(\lambda, \nu)$ converges rapidly, and will approximate the limiting Bernoulli distribution.

6. Discussion

The primary approximation (1) to the normalizing constant for the Conway-Maxwell-Poisson family can perform quite poorly in certain cases. Thus, in this paper, we have presented a more comprehensive study of the normalizing constant for this family using standard statistical methods, which also provide correction terms to the widely used approximation. The statistical approach above does have one loose end that must be tied up: we must rigorously justify the taking of the expectation after using the normal approximation to the square root of a Poisson variate. Our numerical results are encouraging, because they clearly demonstrate that our approach gives reasonably accurate results. Our contributions are the following: our method applies for all $\nu > 0$, not just integer ν. Next, we provide correction terms and give numerical examples that help guide us in determining when to use them. We also give an alternate approximation for very small ν, when the primary approximation (1) is poor. The fact that $C(\lambda, \nu) = {}_0F_{\nu-1}[-; 1, \cdots, 1; \lambda]$ for integer ν, and the special case of $\nu = 2$ for which we have an asymptotic expansion for $I_0(z)$ suggested to us that the asymptotics of generalized hypergeometric functions could be useful for calculating $C(\lambda, \nu)$.

References

[1] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions, Dover, New York, 1965.
[2] N. Bleistein, R. Handelsman, Asymptotic Expansion of Integrals, Dover, New York, 1986.
[3] N. Breslow, Tests of hypotheses in overdispersed Poisson regression and other quasi-likelihood models, Journal of American Statistical Association 85 (1990) 565–571.
[4] B. L. J. Braaksma, Asymptotic expansions and analytic continuations for a class of Barnes' integrals, Compositio Mathematica 15 (1964) 239–341.
[5] R. W. Conway, W. L. Maxwell A queuing model with state dependent service rates, Journal of Industrial Engineering, 12 (1962) 132–136.
[6] K. Imoto, A generalized Conway-Maxwell-Poisson distribution which includes the negative binomial distribution, Applied Mathematics and Computation, 247 (2014) 824–834.
[7] E. L. Lehmann, Testing Statistical Hypotheses, (2nd edition), Wiley, 1986.
[8] S. Nadarajah, Useful moment and CDF formulations for the COM-Poisson distribution, Statistics Papers, 50 (2009) 617–622.
[9] M. A. Ridout, P. Besbeas, An empirical model for underdispersed count data, Statistical Modelling, 4 (2004) 77–89.
[10] K. F. Sellers, S. Borle, 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.
[11] G. Shmueli, T. Minka, J. B. Kadane, S. Borle, P. Boatwright, A useful distribution for fitting discrete data: revival of the Conway-Maxwell-Poisson distribution, Journal of Royal Statistical Society C, 54 (2005) 127–142.