



Some Approximation Results for Bayesian Posteriors that Involve the Hurwitz–Lerch Zeta Distribution

Hongxiang Li¹ · Tsung Fei Khang^{1,2}

Received: 4 April 2022 / Revised: 28 November 2022 / Accepted: 11 January 2023 /
Published online: 31 January 2023

© The Author(s), under exclusive licence to Malaysian Mathematical Sciences Society and Penerbit Universiti Sains Malaysia 2023

Abstract

Consider the generalized Poisson and the negative binomial model with mean parameter equal to kb , where $k \geq 0$ is a count parameter and $0 < b < 1$ is a hyperparameter. We show that conditioning on counts from both models and assuming a uniform prior for k lead to the following Bayesian posterior distributions: (i) geometric for conditioning value of 0; (ii) extended negative binomial for conditioning value of 1; (iii) approximately extended Hurwitz–Lerch zeta distribution for conditioning value of 2 or more. Kullback–Leibler divergence for measuring the quality of the approximating distributions for some combinations of b and the mean–variance ratio is given.

Keywords Approximation · Generalized Poisson distribution · Hurwitz–Lerch zeta distribution · Negative binomial distribution · Posterior distribution

Mathematics Subject Classification 62E17 · 62F15

1 Introduction

The family of generalized Poisson (GP) distribution [1] has been used for more than 40 years to model count data that may be overdispersed or underdispersed. Some of

Communicated by Rosihan M. Ali.

✉ Tsung Fei Khang
tfkhang@um.edu.my

Hongxiang Li
chelsea.divo@hotmail.com ; 17221472@siswa.um.edu.my

¹ Institute of Mathematical Sciences, Faculty of Science, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

² Universiti Malaya Centre for Data Analytics, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

its interesting theoretical properties include a Poisson mixture interpretation [2], and a heavier tail compared to the negative binomial distribution [2, 3]. Various chance mechanisms have been found to generate the GP distribution [4]. Numerous applications are given in [5]. In bioinformatics, the GP distribution has been used for modeling RNA-Seq count data [6–10].

The negative binomial (NB) distribution has a long history of use in the analysis of biological count data [11–13]. In bioinformatics, it is the primary model for modelling RNA-Seq count data and forms the basis of statistical tests of differential gene expression [14–17].

Low et al. [10] pointed out that the observed gene counts in RNA-Seq experiments are the consequence of stochastic variation acting on true gene counts. They proposed to view the true gene count as a count parameter k . Under the GP model, the expectation of the observed gene counts given k is equal to the product of k and a hyperparameter b . Thus, the modeling focus shifts to finding the posterior distribution of k , assuming a specific prior distribution. Since posterior distributions often have complicated forms that make further analysis difficult and computationally expensive, finding appropriate approximations to them is important to improve their applied value in statistics.

We first introduce the GP model. Let X be a random variable following a GP(λ_1, λ_2) distribution. Its probability mass function (pmf) is given by

$$P(X = x|\lambda_1, \lambda_2) = \frac{\lambda_1(\lambda_1 + x\lambda_2)^{x-1} e^{-(\lambda_1+x\lambda_2)}}{x!}, \quad (1)$$

where $x = 0, 1, 2, \dots$, $\lambda_1 > 0$ and $\max\{-1, -\frac{\lambda_1}{4}\} < \lambda_2 < 1$. Negative values of λ_2 correspond to overdispersion, positive values to underdispersion, and $\lambda_2 = 0$ reduces Eq. (1) to the Poisson distribution with mean λ_1 . Consider the following parametrisation: $\lambda_2 = 1 - \sqrt{m}$, $\lambda_1 = kb\sqrt{m}$, where $k = 0, 1, 2, \dots$ is the count parameter, $0 < b < 1$ is a hyperparameter, and $m > 0$ is the mean–variance ratio. Under this parametrisation, the mean and the variance of the GP model are given by $E(X|k) = \lambda_1/(1 - \lambda_2) = kb$, and $\text{Var}(X|k) = E(X|k)/(1 - \lambda_2)^2 = kb/m$, respectively.

Now, consider a random variable Y that has an NB distribution with parameters r (the number of failures until y successes) and p (the probability of success). Its pmf is given by

$$P(Y = y|r, p) = \frac{\Gamma(y+r)}{\Gamma(y+1)\Gamma(r)} p^y (1-p)^r, \quad (2)$$

for $y = 0, 1, 2, \dots$. The mean and the variance of Y are given by $E(Y|k) = pr/(1-p) = kb$, and $\text{Var}(Y|k) = pr/(1-p)^2 = kb/m$, respectively, where $k = 0, 1, 2, \dots$, $0 < b < 1$, $0 < m < 1$, and m is the mean–variance ratio. Thus, $r = kbm/(1-m) = k\tau$, where $\tau = bm/(1-m)$.

We are interested in the posterior distribution of k conditioned on observations from these two models, using an improper uniform prior $P(k) = 1, k = 0, 1, 2, \dots$. The

posterior distribution of k under a GP model (Eq. (1)) is given by

$$\begin{aligned}
 P(k|X = x) &= \frac{k(bk\sqrt{m} + x(1 - \sqrt{m}))^{x-1} e^{-bk\sqrt{m}}}{\sum_{j=x}^{\infty} j(bj\sqrt{m} + x(1 - \sqrt{m}))^{x-1} e^{-bj\sqrt{m}}} \\
 &= \frac{k(k + g(x))^{x-1} e^{-bk\sqrt{m}}}{\sum_{j=x}^{\infty} j(j + g(x))^{x-1} e^{-bj\sqrt{m}}},
 \end{aligned}
 \tag{3}$$

for $k \geq x$, where $0 < m < \min\{(1 - b)^{-2}, 4\}$ and $g(x) = \{(1 - \sqrt{m})/(b\sqrt{m})\}x$. Then, the posterior distribution of k under an NB model (Eq. (2)) is given by

$$\begin{aligned}
 P(k|Y = y) &= \frac{\frac{\Gamma(y+k\tau)}{\Gamma(y+1)\Gamma(k\tau)} (1 - m)^y m^{k\tau}}{\sum_{j=y}^{\infty} \frac{\Gamma(y+j\tau)}{\Gamma(y+1)\Gamma(j\tau)} (1 - m)^y m^{j\tau}} \\
 &= \frac{\frac{\Gamma(y+k\tau)}{\Gamma(k\tau)} m^{k\tau}}{\sum_{j=y}^{\infty} \frac{\Gamma(y+j\tau)}{\Gamma(j\tau)} m^{j\tau}},
 \end{aligned}
 \tag{4}$$

for $k \geq y$.

The posterior distributions (Eq. (3) and (4)) are proper even though an improper uniform prior distribution is used. The aim of this paper is to find their exact distributions, and where this is not possible, approximating distributions that are mathematically tractable. By doing so, their mean and variance can be determined directly from the theoretical properties of the approximating distribution.

2 Results

We first show that when the GP and the NB models have count of 0 or 1, the posterior distribution of k is geometric, and extended NB, respectively.

Theorem 1 *The posterior distribution of k is (i) geometric with mean $e^{-b\sqrt{m}}(1 - e^{-b\sqrt{m}})^{-1}$ and variance $e^{-b\sqrt{m}}(1 - e^{-b\sqrt{m}})^{-2}$ for $k \geq 0$, when $x = 0$ for the GP model; and with mean $m^\tau(1 - m^\tau)^{-1}$ and variance $m^\tau(1 - m^\tau)^{-2}$ for $k \geq 0$, when $y = 0$ for the NB model; (ii) extended NB with mean $(1 + e^{-b\sqrt{m}})(1 - e^{-b\sqrt{m}})^{-1}$ and variance $2e^{-b\sqrt{m}}(1 - e^{-b\sqrt{m}})^{-2}$ for $k \geq 1$, when $x = 1$ for the GP model; and with mean $(1 + m^\tau)(1 - m^\tau)^{-1}$ and variance $2m^\tau(1 - m^\tau)^{-2}$ for $k \geq 1$, when $y = 1$ for the NB model.*

Proof We only show the proof for the GP model with $x = 0, 1$ since the proof for the NB model with $y = 0, 1$ is similar with $m^\tau = e^{-b\sqrt{m}}$.

(i) When $x = 0$:

$$P(k|X = 0) = \frac{e^{-bk\sqrt{m}}}{\sum_{j=0}^{\infty} e^{-bj\sqrt{m}}} = (e^{-b\sqrt{m}})^k (1 - e^{-b\sqrt{m}}),$$

where $k = 0, 1, \dots$. Hence, the posterior distribution of k is geometric with success probability $p = 1 - e^{-b\sqrt{m}}$. The mean and the variance follow from standard results.

(ii) When $x = 1$:

$$P(k|X = 1) = \frac{ke^{-bk\sqrt{m}}}{\sum_{j=1}^{\infty} je^{-bj\sqrt{m}}} = \binom{k}{1} (1 - e^{-b\sqrt{m}})^2 (e^{-b\sqrt{m}})^{k-1},$$

where $k = 1, 2, \dots$. Therefore, the posterior distribution of k is extended NB with parameters $p = e^{-b\sqrt{m}}$ and $r = 2$. The mean and the variance are

$$E(k|X = 1) = \sum_{k=1}^{\infty} kP(k|X = 1) = \frac{1 + e^{-b\sqrt{m}}}{1 - e^{-b\sqrt{m}}},$$

$$\text{Var}(k|X = 1) = \sum_{k=1}^{\infty} k^2P(k|X = 1) - [E(k|X = 1)]^2 = \frac{2e^{-b\sqrt{m}}}{(1 - e^{-b\sqrt{m}})^2},$$

respectively. □

Corollary 1 *The cumulative distribution function (cdf) of the posterior distribution of k is*

- (i) $F_{k|X=0}(k) = 1 - e^{-b\sqrt{m}(k+1)}$ for $k \geq 0$;
- (ii) $F_{k|X=1}(k) = 1 - [k(1 - e^{-b\sqrt{m}}) + 1]e^{-b\sqrt{m}k}$ for $k \geq 1$;
- (iii) $F_{k|Y=0}(k) = 1 - m^{\tau(k+1)}$ for $k \geq 0$;
- (iv) $F_{k|Y=1}(k) = 1 - [k(1 - m^{\tau}) + 1]m^{\tau k}$ for $k \geq 1$.

We now show that the extended Hurwitz–Lerch zeta distribution [18] is an appropriate approximation for the posterior distribution of k under the GP model with $x \geq 2$.

Theorem 2 *The posterior distribution of k , given $X = x$ has a GP distribution with mean kb and variance kb/m , can be approximated by an extended Hurwitz–Lerch zeta distribution with mean $x/b + 1/(b\sqrt{m})$ and variance $(x + 1)/(b^2m)$ for some $k \geq l$ where $x \geq 2$ is the conditioning value, and $l \geq x$.*

Proof First, we note that the denominator in Eq. (3) can be written as

$$\sum_{j=x}^{\infty} (j + g(x))^x e^{-b\sqrt{m}j} - g(x) \sum_{j=x}^{\infty} (j + g(x))^{x-1} e^{-b\sqrt{m}j}.$$

The Lerch transcendent $\Phi(u, s, v)$ (see [19]) is given by

$$\Phi(u, s, v) = \sum_{k=0}^{\infty} \frac{u^k}{(v + k)^s},$$

where u is complex and $|u| < 1$, $v \neq 0, -1, -2, \dots$, and $s \neq 1, 2, \dots$. Representing the denominator using the Lerch transcendent, we get

$$e^{-b\sqrt{m}x} \Phi(e^{-b\sqrt{m}}, -x, wx) - (w - 1)x e^{-b\sqrt{m}x} \Phi(e^{-b\sqrt{m}}, -(x - 1), wx), \tag{5}$$

where $w = 1 + (1 - \sqrt{m})/(b\sqrt{m})$.

The following identity (Eq.1.11(11) in [19]) relates the Lerch transcendent to the Bernoulli polynomials for $s = -h$:

$$\Phi(u, -h, v) = \frac{h!}{u^v} \left(\log \frac{1}{u} \right)^{-(h+1)} - \frac{1}{u^v} \sum_{r=0}^{\infty} \frac{B_{h+r+1}(v)(\log u)^r}{r!(h+r+1)}, \tag{6}$$

where $|\log u| < 2\pi$, $v \neq 0, -1, -2, \dots$, $h \neq -1, -2, \dots$, and $B_n(v)$ is the n th Bernoulli polynomial with argument v . The Bernoulli polynomial is defined as

$$B_n(v) = \sum_{j=0}^n \binom{n}{j} B_{n-j}(0)v^j,$$

where $B_i(0)$ is the i th Bernoulli number.

If we substitute $u = e^{-b\sqrt{m}}$, $h = x$, $v = wx$ into Eq. (6), and then multiply both sides by $e^{-bw\sqrt{m}x}$, we obtain

$$e^{-bw\sqrt{m}x} \Phi(e^{-b\sqrt{m}}, -x, wx) = \frac{\Gamma(x+1)}{(b\sqrt{m})^{x+1}} - \sum_{r=0}^{\infty} \frac{B_{x+r+1}(wx)(-b\sqrt{m})^r}{r!(x+r+1)}. \tag{7}$$

We can approximate Eq. (7) as

$$e^{-bw\sqrt{m}x} \Phi(e^{-b\sqrt{m}}, -x, wx) \approx \frac{\Gamma(x+1)}{(b\sqrt{m})^{x+1}}, \tag{8}$$

provided that

$$\left| \sum_{r=0}^{\infty} \frac{B_{x+r+1}(wx)(-b\sqrt{m})^r}{r!(x+r+1)} \right| = o\left(\frac{\Gamma(x+1)}{(b\sqrt{m})^{x+1}} \right),$$

as $x \rightarrow \infty$, o being the little o notation. To show this, we use an identity involving the Bernoulli polynomials and the sum of x th powers (Eq. 1.13(10) in [19]):

$$\frac{B_{x+1}(z) - B_{x+1}(0)}{x+1} = \sum_{t=0}^{z-1} t^x,$$

for $x = 2, 3, \dots$ and $z \in \mathbb{Z}^+$. For sufficiently large z , $B_{x+1}(z)$ is positive and dominates $B_{x+1}(0)$. Suppose w_x is a positive integer that is sufficiently large, then

$$\frac{B_{x+r+1}(wx)}{x+r+1} \approx \sum_{t=0}^{wx-1} t^{x+r},$$

if $B_{x+r+1}(0) > 0$ and for some b and m such that $B_{x+r+1}(wx) \gg B_{x+r+1}(0)$; if $B_{x+r+1}(0) \leq 0$, we have

$$0 < \frac{B_{x+r+1}(wx)}{x+r+1} \leq \sum_{t=0}^{wx-1} t^{x+r}.$$

Thus,

$$\begin{aligned} & \frac{(b\sqrt{m})^{x+1}}{\Gamma(x+1)} \left| \sum_{r=0}^{\infty} \frac{B_{x+r+1}(wx)(-b\sqrt{m})^r}{r!(x+r+1)} \right| \\ & < \frac{1}{x!} (b\sqrt{m})^{x+1} \sum_{r=0}^{\infty} \left[\frac{(b\sqrt{m})^r}{r!} \left| \frac{B_{x+r+1}(wx)}{x+r+1} \right| \right] \\ & \lesssim \frac{1}{x!} (b\sqrt{m})^{x+1} \sum_{r=0}^{\infty} \left[\frac{(b\sqrt{m})^r}{r!} \sum_{t=0}^{wx-1} t^{x+r} \right] \\ & < \frac{1}{x!} (b\sqrt{m})^{x+1} \sum_{r=0}^{\infty} \left[\frac{(b\sqrt{m})^r}{r!} \int_0^{wx} t^{x+r} dt \right] \\ & = \frac{1}{x!} (b\sqrt{m})^{x+1} \sum_{r=0}^{\infty} \left[\frac{(b\sqrt{m})^r}{r!} \frac{(wx)^{x+r+1}}{x+1} \right] \\ & = \frac{(bw\sqrt{m}x)^{x+1}}{(x+1)!} e^{bw\sqrt{m}x}. \end{aligned}$$

Applying Stirling’s approximation for $x!$, the right-hand-side simplifies to $U = (c/\sqrt{2\pi x})(ce^{c+1})^x$, where $c = bw\sqrt{m}$. For $0 < ce^{c+1} \leq 1$, U converges to 0 as $x \rightarrow \infty$. Therefore, $0 < c \leq W(e^{-1}) \approx 0.2785$, where $W(\cdot)$ is the Lambert W function. Thus, for $\min\{[1 - W(e^{-1})]^2(1 - b)^{-2}, 4\} \leq m < \min\{(1 - b)^{-2}, 4\}$, Eq. (8) should give reasonably good approximation.

Subsequently, we can approximate Eq. (5) using Eq. (8):

$$\begin{aligned} & e^{-b\sqrt{m}x} \Phi(e^{-b\sqrt{m}}, -x, wx) - (w-1)xe^{-b\sqrt{m}x} \Phi(e^{-b\sqrt{m}}, -(x-1), wx) \\ & \approx e^{-b\sqrt{m}x} e^{bw\sqrt{m}x} \frac{\Gamma(x+1)}{(b\sqrt{m})^{x+1}} - (w-1)xe^{-b\sqrt{m}x} e^{bw\sqrt{m}x} \frac{\Gamma(x)}{(b\sqrt{m})^x} \end{aligned}$$

$$\begin{aligned}
 &= e^{(1-\sqrt{m})x} \frac{\sqrt{m} \Gamma(x+1)}{(b\sqrt{m})^{x+1}} \\
 &\approx e^{(1-\sqrt{m})x} \sqrt{m} e^{-bw\sqrt{m}x} \Phi(e^{-b\sqrt{m}}, -x, wx) \\
 &= \sqrt{m} e^{-b\sqrt{m}x} \Phi(e^{-b\sqrt{m}}, -x, wx).
 \end{aligned} \tag{9}$$

Substituting Eq. (9) for the denominator in Eq. (3), we obtain

$$\begin{aligned}
 P(k|X = x) &\approx \frac{k(k+g(x))^{x-1} e^{-b\sqrt{m}k}}{\sqrt{m} \Phi(e^{-b\sqrt{m}}, -x, wx) e^{-b\sqrt{m}x}} \\
 &= \underbrace{\frac{(e^{-b\sqrt{m}})^{k-x}}{\Phi(e^{-b\sqrt{m}}, -x, wx)(k-x+wx)^{-x}}}_A \times \underbrace{\frac{k}{\sqrt{m}(k+g(x))}}_B.
 \end{aligned} \tag{10}$$

To see that A is the pmf of an extended Hurwitz–Lerch zeta distribution with extended parameter space [20], we start with the pmf of the Hurwitz–Lerch zeta distribution:

$$q_k = \frac{1}{\theta \Phi(\theta, s+1, a+1)} \frac{\theta^k}{(k+a)^{s+1}}, \tag{11}$$

for $k = 1, 2, \dots$, where $a > -1$ and $s \in \mathbb{R}$ if $0 < \theta < 1$ or $s > 0$ if $\theta = 1$. Shifting the support to $k = x, x+1, \dots$, Eq. (11) becomes

$$q_k = \frac{\theta^{k-x}}{\Phi(\theta, s+1, a+1)(k-x+1+a)^{s+1}}, \tag{12}$$

for $k = x, x+1, \dots$, where $x \geq 2$. Taking $\theta = e^{-b\sqrt{m}}$, $s+1 = -x$, $a+1 = wx$ and substituting them into Eq. (12) leads to A in Eq. (10).

For some combinations of b and m , there exists an $l \geq x$ such that $k \geq l$ results in $\sqrt{m}(k+g(x)) = k + o(k)$. In this case, B of Eq. (10) becomes approximately 1, thus

$$P(k|X = x) \approx \frac{(e^{-b\sqrt{m}})^{k-x}}{\Phi(e^{-b\sqrt{m}}, -x, wx)(k-x+wx)^{-x}}, \tag{13}$$

which is the pmf of the extended Hurwitz–Lerch zeta distribution with parameters $\theta = e^{-b\sqrt{m}}$, $s+1 = -x$, $a+1 = wx$.

To derive the mean and the variance of the extended Hurwitz–Lerch distribution (Eq. (13)), we first note that the latter is a special case of the modified power series distribution [21], which has pmf

$$P(Z = z) = \frac{A(z)[g(\theta)]^z}{f(\theta)}, \quad z \in \mathbb{B},$$

where $\mathbb{B} \subset \mathbb{Z}^+$, $A(z) > 0$, and $f(\theta)$, $g(\theta)$ are positive, finite and differentiable functions of θ . In this case, we have

$$g(\theta) = \theta, \quad f(\theta) = \theta^x \Phi(\theta, s + 1, a + 1), \quad A(z) = \frac{1}{(z - x + 1 + a)^{s+1}}.$$

The expectation of Z is

$$E(Z) = \frac{g(\theta)f'(\theta)}{f(\theta)g'(\theta)} = \frac{\theta}{\theta^x \Phi(\theta, s + 1, a + 1)} \frac{\partial}{\partial \theta} \theta^x \Phi(\theta, s + 1, a + 1), \quad (14)$$

and the variance is

$$\text{Var}(Z) = \frac{g(\theta)}{g'(\theta)} \frac{\partial}{\partial \theta} E(Z). \quad (15)$$

Note that

$$\frac{\partial}{\partial \theta} \Phi(\theta, s, a) = \frac{1}{\theta} \Phi(\theta, s - 1, a) - \frac{a}{\theta} \Phi(\theta, s, a). \quad (16)$$

Therefore,

$$\frac{\partial}{\partial \theta} \theta^x \Phi(\theta, s + 1, a + 1) = \theta^{x-1} [(x - a - 1) \Phi(\theta, s + 1, a + 1) + \Phi(\theta, s, a + 1)]. \quad (17)$$

Let $\theta = e^{-b\sqrt{m}}$, $s + 1 = -x$ and $a + 1 = wx$. Substituting Eq. (17) into Eq. (14) yields

$$E(k|X = x) \approx \frac{\Phi(e^{-b\sqrt{m}}, -x - 1, wx)}{\Phi(e^{-b\sqrt{m}}, -x, wx)} - \frac{(1 - \sqrt{m})x}{b\sqrt{m}}. \quad (18)$$

Then, substituting Eq. (18) into Eq. (15) and using Eq. (16) yields

$$\text{Var}(k|X = x) \approx \frac{\Phi(e^{-b\sqrt{m}}, -x - 2, wx)}{\Phi(e^{-b\sqrt{m}}, -x, wx)} - \left[\frac{\Phi(e^{-b\sqrt{m}}, -x - 1, wx)}{\Phi(e^{-b\sqrt{m}}, -x, wx)} \right]^2. \quad (19)$$

Eq. (18) can be further approximated by applying Eq. (8):

$$E(k|X = x) \approx \frac{\frac{\Gamma(x+2)}{(b\sqrt{m})^{x+2}}}{\frac{\Gamma(x+1)}{(b\sqrt{m})^{x+1}}} - \frac{1 - \sqrt{m}}{b\sqrt{m}} x = \frac{x + 1}{b\sqrt{m}} - \frac{(1 - \sqrt{m})x}{b\sqrt{m}} = \frac{x}{b} + \frac{1}{b\sqrt{m}}.$$

Similarly, Eq. (19) can be further approximated as

$$\text{Var}(k|X = x) \approx \frac{\frac{\Gamma(x+3)}{(b\sqrt{m})^{x+3}}}{\frac{\Gamma(x+1)}{(b\sqrt{m})^{x+1}}} - \left[\frac{\frac{\Gamma(x+2)}{(b\sqrt{m})^{x+2}}}{\frac{\Gamma(x+1)}{(b\sqrt{m})^{x+1}}} \right]^2 = \frac{(x + 2)(x + 1)}{(b\sqrt{m})^2} - \frac{(x + 1)^2}{(b\sqrt{m})^2} = \frac{x + 1}{b^2m}.$$

□

For the NB model with $y \geq 2$, we again find that the posterior distribution of k is approximately given by the extended Hurwitz–Lerch zeta distribution.

Theorem 3 *Let Y have an NB distribution with pmf given by Eq. (2). The posterior distribution of k given $Y = y$ approximately follows the extended Hurwitz–Lerch zeta distribution with mean $-(y + 1)/(\tau \log m) - (y - 1)/(2\tau)$ and variance $(y + 1)/(\tau \log m)^2$, for $k \geq y$, where $y \geq 2$.*

Proof For non-negative real α, β such that $\alpha \neq \beta$, Laforgia & Natalini [22] give the following approximation for the quotient of gamma functions:

$$\frac{\Gamma(y + \alpha)}{\Gamma(y + \beta)} \approx \frac{1}{(y + c)^{\beta - \alpha}}, \tag{20}$$

when $y \rightarrow \infty$, with $c = (\alpha + \beta - 1)/2$. Applying Eq. (20) in Eq. (4) yields

$$\begin{aligned} P(k|Y = y) &\approx \frac{(k\tau + \frac{y-1}{2})^y m^{k\tau}}{\sum_{j=y}^{\infty} (j\tau + \frac{y-1}{2})^y m^{j\tau}} \\ &= \frac{(k + \frac{y-1}{2\tau})^y m^{k\tau}}{\sum_{j=y}^{\infty} (j + \frac{y-1}{2\tau})^y m^{j\tau}}, \end{aligned} \tag{21}$$

as $k\tau, j\tau \rightarrow \infty$, for $k \geq y$. Then, representing the denominator of Eq. (21) using the Lerch transcendent, we obtain

$$\sum_{j=y}^{\infty} \left(j + \frac{y-1}{2\tau} \right)^y m^{j\tau} = m^{\tau y} \Phi \left(m^{\tau}, -y, \left(\frac{1}{2\tau} + 1 \right) y - \frac{1}{2\tau} \right). \tag{22}$$

Hence, Eq. (21) can be expressed as

$$P(k|Y = y) \approx \frac{\left(k + \frac{y-1}{2\tau} \right)^y m^{\tau(k-y)}}{\Phi \left(m^{\tau}, -y, \left(\frac{1}{2\tau} + 1 \right) y - \frac{1}{2\tau} \right)}, \tag{23}$$

where $k = y, y + 1, \dots$. Eq. (23) is just Eq. (12) with $\theta = m^{\tau}, s + 1 = -y, a + 1 = \left(\frac{1}{2\tau} + 1 \right) y - \frac{1}{2\tau}$. Therefore, we conclude that the pmf of the extended Hurwitz–Lerch zeta distribution approximates the posterior distribution of k under NB model for $y \geq 2$.

By the same approach used in Theorem 2 to derive the mean and the variance of the posterior distribution of k , we obtain

$$E(k|Y = y) \approx \frac{\Phi(m^\tau, -y - 1, (\frac{1}{2\tau} + 1)y - \frac{1}{2\tau})}{\Phi(m^\tau, -y, (\frac{1}{2\tau} + 1)y - \frac{1}{2\tau})} - \frac{y - 1}{2\tau}, \tag{24}$$

$$\begin{aligned} \text{Var}(k|Y = y) \approx & \frac{\Phi(m^\tau, -y - 2, (\frac{1}{2\tau} + 1)y - \frac{1}{2\tau})}{\Phi(m^\tau, -y, (\frac{1}{2\tau} + 1)y - \frac{1}{2\tau})} \\ & - \left[\frac{\Phi(m^\tau, -y - 1, (\frac{1}{2\tau} + 1)y - \frac{1}{2\tau})}{\Phi(m^\tau, -y, (\frac{1}{2\tau} + 1)y - \frac{1}{2\tau})} \right]^2, \end{aligned} \tag{25}$$

for $k \geq y$, where $y \geq 2$. By similar argument leading to Eq. (8), we obtain

$$m^{\tau g(y)} \Phi(m^\tau, -y, g(y)) \approx \frac{\Gamma(y + 1)}{(-\tau \log m)^{y+1}}, \tag{26}$$

where $g(y) = (\frac{1}{2\tau} + 1)y - \frac{1}{2\tau}$. Using Eq. (26), Eq. (24) can be approximated as

$$E(k|Y = y) \approx \frac{\frac{\Gamma(y+2)}{(-\tau \log m)^{y+2}}}{\frac{\Gamma(y+1)}{(-\tau \log m)^{y+1}}} - \frac{y - 1}{2\tau} = \frac{y + 1}{\tau \log m} - \frac{y - 1}{2\tau}.$$

Similarly, we can use Eq. (26) to approximate Eq. (25) as

$$\begin{aligned} \text{Var}(k|Y = y) \approx & \frac{\frac{\Gamma(y+3)}{(-\tau \log m)^{y+3}}}{\frac{\Gamma(y+1)}{(-\tau \log m)^{y+1}}} - \left(\frac{y + 1}{-\tau \log m} \right)^2 \\ = & \frac{(y + 2)(y + 1)}{(\tau \log m)^2} - \frac{(y + 1)^2}{(\tau \log m)^2} \\ = & \frac{y + 1}{(\tau \log m)^2}. \end{aligned}$$

□

The results of Theorem 2 and Theorem 3 lead us to the following corollary.

Corollary 2 *The cdf of the posterior distribution of k is approximately*

- (i) $F_{k|X=x}(k) \approx 1 - e^{-b\sqrt{m}(k-x+1)} \Phi(e^{-b\sqrt{m}}, -x, k+wx-x+1) / \Phi(e^{-b\sqrt{m}}, -x, wx)$, for $k \geq x$, where $x \geq 2$, for the GP model;
- (ii) $F_{k|Y=y}(k) \approx 1 - m^{\tau(k-y+1)} \Phi(m^\tau, -y, k + \frac{y-1}{2\tau} + 1) / \Phi(m^\tau, -y, \frac{y-1}{2\tau} + y)$, for $k \geq y$, where $y \geq 2$, for the NB model.

Table 1 Kullback–Leibler divergence for Eq. (3) versus Eq. (13)

<i>b</i>	<i>m</i>	<i>X</i> = 2	<i>X</i> = 3	<i>X</i> = 10	<i>X</i> = 20	<i>X</i> = 50
0.2	0.2	0.0552	0.0546	0.0440	0.0324	0.0160
	0.4	0.0278	0.0258	0.0157	0.0090	0.0035
	0.6	0.0117	0.0101	0.0046	0.0023	0.0009
	0.8	0.0028	0.0022	0.0008	0.0004	0.0001
	1.2	0.0026	0.0017	0.0004	0.0002	0.0001
	1.5	0.0131	0.0073	0.0018	0.0009	0.0003
0.4	0.2	0.0411	0.0391	0.0299	0.0231	0.0138
	0.4	0.0185	0.0168	0.0110	0.0074	0.0035
	0.6	0.0072	0.0062	0.0035	0.0021	0.0009
	0.8	0.0016	0.0014	0.0007	0.0004	0.0001
	1.2	0.0015	0.0011	0.0004	0.0002	0.0001
	2.0	0.0304	0.0178	0.0046	0.0022	0.0009
0.6	0.2	0.0313	0.0286	0.0200	0.0152	0.0098
	0.4	0.0128	0.0112	0.0070	0.0050	0.0029
	0.6	0.0046	0.0039	0.0023	0.0016	0.0008
	0.8	0.0010	0.0008	0.0004	0.0003	0.0001
	1.2	0.0008	0.0006	0.0003	0.0002	0.0001
	2.5	0.0311	0.0215	0.0070	0.0035	0.0014
0.8	0.2	0.0242	0.0214	0.0134	0.0096	0.0058
	0.4	0.0091	0.0076	0.0043	0.0029	0.0017
	0.6	0.0031	0.0025	0.0013	0.0009	0.0005
	0.8	0.0006	0.0005	0.0003	0.0002	0.0001
	1.2	0.0005	0.0004	0.0002	0.0001	0.0001
	3.0	0.0207	0.0155	0.0064	0.0038	0.0018

3 Computational Validation

Table 1 shows how well the extended Hurwitz–Lerch zeta approximates the posterior distribution of *k* under GP for different combinations of *b* and *m*. For a fixed *b*, the approximation is best for *m* in the neighborhood of 1. For a fixed *m*, the approximation improves as *b* becomes closer to 1. Finally, for fixed *m* and *b*, the approximation improves as *x* increases.

Table 2 shows that for a given *y*, the larger the values of *m* and *b*, the better the extended Hurwitz–Lerch zeta distribution approximates the posterior distribution of *k* under the NB model. For fixed *m* and *b*, the approximation deteriorates as *y* increases up to 50. However, the Kullback–Leibler divergence remains well below 0.02 when $0.6 \leq m < 1$ for the *b* values considered.

Table 3 and 4 show results of approximating the mean and the standard deviation of the posterior distribution of *k* given *X* = *x* has a GP distribution. Similar results for the posterior distribution of *k* given *Y* = *y* has a NB distribution are given in Table 5 and 6. In general, the approximations have relative error that stays within 10% of the

Table 2 Kullback–Leibler divergence for Eq. (4) versus Eq. (23)

m	b	$Y = 2$	$Y = 3$	$Y = 10$	$Y = 20$	$Y = 50$
0.2	0.2	0.0165	0.0287	0.1214	0.2928	0.8883
	0.4	0.0105	0.0184	0.0802	0.2014	0.6834
	0.6	0.0070	0.0122	0.0524	0.1292	0.4401
	0.8	0.0048	0.0084	0.0343	0.0802	0.2494
0.4	0.2	0.0036	0.0064	0.0272	0.0557	0.1265
	0.4	0.0017	0.0033	0.0172	0.0432	0.1225
	0.6	0.0009	0.0018	0.0098	0.0267	0.0951
	0.8	0.0005	0.0010	0.0054	0.0141	0.0503
0.6	0.2	0.0006	0.0010	0.0036	0.0061	0.0129
	0.4	0.0002	0.0004	0.0024	0.0055	0.0129
	0.6	0.0001	0.0002	0.0013	0.0036	0.0117
	0.8	0.0000	0.0001	0.0006	0.0018	0.0066
0.8	0.2	0.0000	0.0001	0.0002	0.0002	0.0005
	0.4	0.0000	0.0000	0.0001	0.0002	0.0005
	0.6	0.0000	0.0000	0.0001	0.0002	0.0005
	0.8	0.0000	0.0000	0.0000	0.0001	0.0003

true value when $m \geq 0.6$, for the b and X, Y values considered. Approximations that use the Lerch transcendent function (e.g., Eq. (18), Eq. (19), Eq. (24) and Eq. (25) have smaller relative error compared to the simpler equations in Theorem 2 and Theorem 3. However, for large values of X and Y , both approximations generally have similar relative error for $m \geq 0.6$. Since the Lerch transcendent cannot be evaluated for some combinations of b and m for large X, Y values, the use of the simpler equations in these two theorems suffices.

4 Concluding Remarks

In this paper, we have clarified several theoretical properties of the posterior distribution of a count parameter k arising in the GP and the NB model. Thus, for conditioning values of 0 and 1 from these two models, the posterior distribution of k is found to be geometric and extended NB, respectively. For conditioning values of 2 or more, the posterior distribution of k under either a GP or an NB model is approximated by the extended Hurwitz–Lerch zeta distribution. To our knowledge, this is the first instance where a connection between the Hurwitz–Lerch zeta distribution and a Bayesian posterior distribution is demonstrated. The present results open up the possibility of using the posterior mean to correct for observed gene counts in RNA-Seq data analysis.

Table 3 Posterior mean of k when X has a GP distribution. For the array in each cell, the first value is the exact posterior mean computed using Eq. (3) (first 10, 000 terms); the second and the third give the deviation from the first value using Eq. (18) and Theorem 2, respectively

b	m	$X = 2$	$X = 3$	$X = 10$	$X = 20$	$X = 50$
0.2	0.2	(29.6, -4.9, -8.5)	(36.1, -5.8, -9.9)	(74.5, -9.4, -13.3)	(125.2, -11.7, -14.0)	(275.1, -13.6, -13.9)
	0.4	(21.7, -2.5, -3.8)	(27.1, -2.8, -4.2)	(62.6, -4.1, -4.7)	(112.5, -4.5, -4.6)	(262.5, -4.6, -4.6)
	0.6	(18.3, -1.3, -1.8)	(23.4, -1.4, -1.9)	(58.4, -1.8, -1.9)	(108.3, -1.9, -1.9)	(258.3, -1.9, -1.9)
	0.8	(16.3, -0.5, -0.7)	(21.3, -0.6, -0.7)	(56.3, -0.7, -0.7)	(106.3, -0.7, -0.7)	(256.3, -0.7, -0.7)
	1.2	(14.2, 0.4, 0.4)	(19.2, 0.4, 0.4)	(54.2, 0.4, 0.4)	(104.2, 0.4, 0.4)	(254.2, 0.4, 0.4)
0.4	1.5	(13.3, 0.7, 0.7)	(18.3, 0.7, 0.7)	(53.3, 0.7, 0.7)	(103.3, 0.7, 0.7)	(253.3, 0.7, 0.7)
	0.2	(14.9, -2.2, -4.4)	(18.3, -2.5, -5.2)	(38.0, -3.8, -7.4)	(63.5, -4.8, -7.9)	(138.1, -6.2, -7.5)
	0.4	(11.0, -1.0, -2.0)	(13.8, -1.2, -2.3)	(31.8, -1.7, -2.8)	(56.6, -2.0, -2.6)	(131.3, -2.3, -2.3)
	0.6	(9.3, -0.5, -1.1)	(11.9, -0.6, -1.2)	(29.4, -0.8, -1.2)	(54.2, -0.9, -1.0)	(129.2, -0.9, -0.9)
	0.8	(8.3, -0.2, -0.5)	(10.8, -0.2, -0.5)	(28.2, -0.3, -0.4)	(53.1, -0.3, -0.3)	(128.1, -0.3, -0.3)
0.6	1.2	(7.2, 0.2, 0.1)	(9.7, 0.2, 0.1)	(27.1, 0.2, 0.2)	(52.1, 0.2, 0.2)	(127.1, 0.2, 0.2)
	2.0	(6.2, 0.5, 0.5)	(8.7, 0.5, 0.5)	(26.3, 0.5, 0.5)	(51.3, 0.5, 0.5)	(126.3, 0.5, 0.5)
	0.2	(10.1, -1.3, -3.0)	(12.4, -1.4, -3.7)	(26.2, -2.0, -5.8)	(43.8, -2.5, -6.8)	(94.0, -3.3, -6.9)
	0.4	(7.5, -0.6, -1.5)	(9.5, -0.6, -1.8)	(21.9, -0.9, -2.6)	(38.6, -1.1, -2.7)	(88.1, -1.4, -2.1)
	0.6	(6.4, -0.3, -0.9)	(8.2, -0.3, -1.0)	(20.2, -0.4, -1.4)	(36.7, -0.5, -1.2)	(86.3, -0.6, -0.8)
0.8	(5.7, -0.1, -0.5)	(7.5, -0.1, -0.6)	(19.2, -0.2, -0.7)	(35.7, -0.2, -0.5)	(85.5, -0.2, -0.3)	

Table 3 continued

b	m	$X = 2$	$X = 3$	$X = 10$	$X = 20$	$X = 50$
0.8	1.2	(4.9, 0.1, -0.1)	(6.7, 0.1, -0.1)	(18.3, 0.1, -0.1)	(34.8, 0.1, 0.1)	(84.7, 0.1, 0.1)
	2.5	(4.0, 0.4, 0.4)	(5.7, 0.4, 0.4)	(17.3, 0.4, 0.4)	(34.0, 0.4, 0.4)	(84.0, 0.4, 0.4)
	0.2	(7.7, -0.9, -2.4)	(9.5, -0.9, -3.0)	(20.6, -1.2, -5.3)	(34.7, -1.4, -6.9)	(74.3, -1.7, -9.0)
	0.4	(5.8, -0.4, -1.3)	(7.4, -0.4, -1.6)	(17.4, -0.5, -2.9)	(30.6, -0.6, -3.7)	(68.7, -0.7, -4.2)
	0.6	(4.9, -0.2, -0.8)	(6.4, -0.2, -1.1)	(16.0, -0.2, -1.9)	(28.9, -0.3, -2.3)	(66.5, -0.3, -2.4)
	0.8	(4.4, -0.1, -0.5)	(5.9, -0.1, -0.7)	(15.2, -0.1, -1.3)	(28.0, -0.1, -1.6)	(65.3, -0.1, -1.4)
	1.2	(3.9, 0.0, -0.2)	(5.3, 0.1, -0.4)	(14.4, 0.1, -0.7)	(26.9, 0.1, -0.8)	(64.1, 0.1, -0.5)
	3.0	(3.0, 0.2, 0.2)	(4.3, 0.2, 0.1)	(13.1, 0.2, 0.1)	(25.6, 0.3, 0.2)	(62.9, 0.3, 0.3)

Table 4 Posterior standard deviation of k when X has a GP distribution. For the array in each cell, the first value is the exact posterior standard deviation computed using Eq. (3) (first 10, 000 terms); the second and the third give the deviation from the first value using Eq. (19) and Theorem 2, respectively. NA indicates an instance of failure to compute the Lerch transcendent using the VGAM package for $X = 50$ and $m = b = 0.8$

b	m	$X = 2$	$X = 3$	$X = 10$	$X = 20$	$X = 50$
0.2	0.2	(18.9, -0.7, 0.5)	(21.5, -0.7, 0.8)	(35.1, -0.5, 2.0)	(49.2, 0.1, 2.0)	(78.6, 0.9, 1.3)
	0.4	(13.5, -0.3, 0.2)	(15.5, -0.3, 0.4)	(25.7, 0.0, 0.5)	(35.9, 0.2, 0.3)	(56.3, 0.2, 0.2)
	0.6	(11.1, -0.1, 0.1)	(12.8, -0.1, 0.2)	(21.3, 0.0, 0.1)	(29.5, 0.1, 0.1)	(46.1, 0.0, 0.0)
	0.8	(9.6, 0.0, 0.1)	(11.1, 0.0, 0.1)	(18.5, 0.0, 0.0)	(25.6, 0.0, 0.0)	(39.9, 0.0, 0.0)
	1.2	(7.9, 0.0, 0.0)	(9.1, 0.0, 0.0)	(15.1, 0.0, 0.0)	(20.9, 0.0, 0.0)	(32.6, 0.0, 0.0)
0.4	1.5	(7.0, 0.0, 0.0)	(8.1, 0.0, 0.0)	(13.5, 0.0, 0.0)	(18.7, 0.0, 0.0)	(29.1, 0.0, 0.0)
	0.2	(9.4, -0.4, 0.3)	(10.7, -0.4, 0.5)	(17.1, -0.5, 1.4)	(23.9, -0.4, 1.7)	(38.7, 0.0, 1.2)
	0.4	(6.7, -0.2, 0.2)	(7.6, -0.2, 0.3)	(12.6, -0.2, 0.6)	(17.7, -0.1, 0.4)	(28.1, 0.1, 0.1)
	0.6	(5.5, -0.1, 0.1)	(6.3, -0.1, 0.2)	(10.5, 0.0, 0.2)	(14.7, 0.0, 0.1)	(23.0, 0.0, 0.0)
	0.8	(4.8, 0.0, 0.1)	(5.5, 0.0, 0.1)	(9.2, 0.0, 0.1)	(12.8, 0.0, 0.0)	(20.0, 0.0, 0.0)
0.6	1.2	(3.9, 0.0, 0.0)	(4.5, 0.0, 0.1)	(7.6, 0.0, 0.0)	(10.5, 0.0, 0.0)	(16.3, 0.0, 0.0)
	2.0	(3.0, 0.0, 0.0)	(3.5, 0.0, 0.0)	(5.8, 0.0, 0.0)	(8.1, 0.0, 0.0)	(12.6, 0.0, 0.0)
	0.2	(6.2, -0.3, 0.2)	(7.0, -0.3, 0.4)	(11.0, -0.5, 1.3)	(15.1, -0.5, 1.9)	(24.4, -0.4, 2.2)
	0.4	(4.4, -0.1, 0.2)	(5.0, -0.1, 0.3)	(8.0, -0.2, 0.8)	(11.2, -0.2, 0.9)	(18.2, -0.1, 0.6)
	0.6	(3.6, -0.1, 0.1)	(4.1, -0.1, 0.2)	(6.7, -0.1, 0.5)	(9.4, -0.1, 0.5)	(15.2, 0.0, 0.2)
0.8	0.8	(3.1, 0.0, 0.1)	(3.5, 0.0, 0.2)	(5.9, 0.0, 0.3)	(8.3, 0.0, 0.3)	(13.3, 0.0, 0.1)
	1.2	(2.5, 0.0, 0.1)	(2.9, 0.0, 0.1)	(4.9, 0.0, 0.2)	(6.9, 0.0, 0.1)	(10.9, 0.0, 0.0)
	2.5	(1.8, 0.1, 0.0)	(2.1, 0.0, 0.1)	(3.5, 0.0, 0.0)	(4.8, 0.0, 0.0)	(7.5, 0.0, 0.0)

Table 4 continued

b	m	$X = 2$	$X = 3$	$X = 10$	$X = 20$	$X = 50$
0.8	0.2	(4.6, -0.2, 0.2)	(5.2, -0.3, 0.4)	(7.9, -0.4, 1.4)	(10.6, -0.4, 2.2)	(16.3, -0.5, 3.6)
	0.4	(3.3, -0.1, 0.2)	(3.7, -0.1, 0.3)	(5.6, -0.2, 0.9)	(7.6, -0.2, 1.4)	(12.1, -0.2, 2.1)
	0.6	(2.6, -0.1, 0.1)	(3.0, -0.1, 0.3)	(4.6, -0.1, 0.7)	(6.3, -0.1, 1.1)	(10.2, -0.1, 1.3)
	0.8	(2.3, 0.0, 0.1)	(2.6, 0.0, 0.2)	(4.0, 0.0, 0.6)	(5.6, 0.0, 0.8)	(9.0, NA, 0.9)
	1.2	(1.9, 0.0, 0.1)	(2.1, 0.0, 0.2)	(3.3, 0.0, 0.4)	(4.7, 0.0, 0.6)	(7.7, 0.0, 0.5)
	3.0	(1.2, 0.1, 0.1)	(1.3, 0.1, 0.1)	(2.2, 0.1, 0.2)	(3.2, 0.0, 0.1)	(5.1, 0.0, 0.0)

Table 5 Posterior mean of k when Y has an NB distribution. For the array in each cell, the first value is the exact posterior mean computed using Eq. (4) (first 10, 000 terms); the second and the third give the deviation from the first value using Eq. (24) and Theorem 3, respectively

m	b	$Y = 2$	$Y = 3$	$Y = 10$	$Y = 20$	$Y = 50$
0.2	0.2	(31.8, -2.5, -4.6)	(37.9, -4.1, -8.2)	(75.2, -16.3, -28.5)	(125.3, -37.0, -54.4)	(275.0, -109.3, -131.3)
	0.4	(16.0, -1.0, -2.4)	(19.2, -1.7, -4.4)	(38.4, -6.5, -15.1)	(63.6, -14.5, -28.2)	(138.0, -43.4, -66.1)
	0.6	(10.8, -0.6, -1.7)	(13.1, -1.0, -3.2)	(26.6, -3.4, -11.0)	(44.0, -7.2, -20.4)	(93.9, -20.5, -46.0)
	0.8	(8.3, -0.4, -1.4)	(10.1, -0.6, -2.6)	(21.0, -2.0, -9.3)	(35.0, -4.0, -17.2)	(74.4, -10.0, -38.4)
0.4	0.2	(22.1, -0.7, -1.3)	(27.4, -1.2, -2.2)	(62.6, -5.3, -6.3)	(112.5, -11.4, -11.9)	(262.5, -28.8, -28.8)
	0.4	(11.2, -0.3, -0.8)	(14.0, -0.5, -1.3)	(31.8, -2.1, -3.6)	(56.5, -4.9, -6.2)	(131.3, -14.0, -14.4)
	0.6	(7.6, -0.1, -0.7)	(9.6, -0.2, -1.2)	(22.0, -1.0, -3.2)	(38.6, -2.4, -5.1)	(88.1, -7.7, -10.2)
	0.8	(5.9, -0.1, -0.7)	(7.5, -0.1, -1.1)	(17.4, -0.5, -3.4)	(30.7, -1.2, -5.5)	(68.7, -3.6, -10.3)
0.6	0.2	(18.3, -0.2, -0.4)	(23.4, -0.4, -0.7)	(58.4, -1.5, -1.6)	(108.3, -3.0, -3.0)	(258.3, -7.2, -7.2)
	0.4	(9.3, -0.1, -0.4)	(11.9, -0.1, -0.6)	(29.4, -0.6, -1.0)	(54.2, -1.4, -1.6)	(129.2, -3.6, -3.6)
	0.6	(6.4, 0.0, -0.4)	(8.2, -0.1, -0.6)	(20.2, -0.3, -1.3)	(36.7, -0.7, -1.6)	(86.3, -2.3, -2.5)
	0.8	(5.0, 0.0, -0.5)	(6.4, 0.0, -0.8)	(16.0, -0.2, -1.8)	(28.9, -0.4, -2.6)	(66.5, -1.1, -3.7)
0.8	0.2	(16.3, 0.0, -0.1)	(21.3, -0.1, -0.2)	(56.3, -0.3, -0.3)	(106.3, -0.5, -0.5)	(256.3, -1.2, -1.2)
	0.4	(8.3, 0.0, -0.2)	(10.8, 0.0, -0.3)	(28.2, -0.1, -0.2)	(53.1, -0.2, -0.3)	(128.1, -0.6, -0.6)
	0.6	(5.7, 0.0, -0.3)	(7.5, 0.0, -0.4)	(19.2, -0.1, -0.6)	(35.7, -0.1, -0.5)	(85.5, -0.4, -0.4)
	0.8	(4.4, 0.0, -0.4)	(5.9, 0.0, -0.6)	(15.2, 0.0, -1.2)	(28.0, -0.1, -1.5)	(65.3, -0.2, -1.5)

Table 6 Posterior standard deviation of k when Y has an NB distribution. For the array in each cell, the first value is the exact posterior standard deviation computed using Eq. (4) (first 10, 000 terms); the second and the third give the deviation from the first value using Eq. (25) and Theorem 3, respectively. NA indicates an instance of failure to compute the Lerch transcendent using the VGAM package for $Y = 50$ and $m = b = 0.8$

m	b	$Y = 2$	$Y = 3$	$Y = 10$	$Y = 20$	$Y = 50$
0.2	0.2	(20.8, 0.0, 0.8)	(23.4, -0.2, 1.4)	(37.0, -1.6, 4.2)	(51.0, -3.4, 6.0)	(79.8, -5.0, 8.9)
	0.4	(10.3, 0.0, 0.4)	(11.6, -0.2, 0.8)	(18.1, -1.2, 2.5)	(24.8, -2.7, 3.7)	(39.4, -6.5, 4.9)
	0.6	(6.8, -0.1, 0.3)	(7.6, -0.1, 0.6)	(11.6, -0.9, 2.1)	(15.7, -2.0, 3.3)	(24.9, -5.7, 4.7)
0.4	0.8	(5.1, 0.0, 0.3)	(5.7, -0.1, 0.6)	(8.4, -0.6, 2.0)	(11.0, -1.5, 3.2)	(16.7, -4.1, 5.5)
	0.2	(13.9, 0.0, 0.3)	(15.9, 0.0, 0.5)	(26.1, 0.3, 1.0)	(36.2, 0.9, 1.3)	(56.5, 2.0, 2.0)
	0.4	(6.9, 0.0, 0.2)	(7.8, 0.0, 0.4)	(12.8, -0.1, 0.8)	(17.8, -0.1, 0.9)	(28.2, 0.6, 1.0)
0.6	0.6	(4.5, 0.0, 0.2)	(5.1, 0.0, 0.3)	(8.1, -0.2, 0.9)	(11.3, -0.4, 1.2)	(18.3, -0.6, 1.2)
	0.8	(3.4, 0.0, 0.2)	(3.8, 0.0, 0.3)	(5.7, -0.1, 1.1)	(7.7, -0.4, 1.7)	(12.1, -1.0, 2.5)
	0.2	(11.2, 0.0, 0.1)	(12.9, 0.0, 0.2)	(21.4, 0.2, 0.3)	(29.6, 0.3, 0.3)	(46.1, 0.5, 0.5)
0.8	0.4	(5.5, 0.0, 0.1)	(6.3, 0.0, 0.2)	(10.5, 0.0, 0.3)	(14.7, 0.1, 0.2)	(23.0, 0.2, 0.3)
	0.6	(3.6, 0.0, 0.1)	(4.1, 0.0, 0.2)	(6.7, 0.0, 0.5)	(9.4, -0.1, 0.6)	(15.2, 0.0, 0.3)
	0.8	(2.7, 0.0, 0.2)	(3.0, 0.0, 0.3)	(4.6, 0.0, 0.8)	(6.3, -0.1, 1.1)	(10.2, -0.2, 1.5)
0.8	0.2	(9.6, 0.0, 0.1)	(11.1, 0.0, 0.1)	(18.5, 0.0, 0.0)	(25.6, 0.1, 0.1)	(39.9, 0.1, 0.1)
	0.4	(4.8, 0.0, 0.1)	(5.5, 0.0, 0.1)	(9.2, 0.0, 0.1)	(12.8, 0.0, 0.0)	(20.0, 0.0, 0.0)
	0.6	(3.1, 0.0, 0.1)	(3.6, 0.0, 0.2)	(5.9, 0.0, 0.3)	(8.3, 0.0, 0.3)	(13.3, 0.0, 0.1)
0.8	(2.3, 0.0, 0.1)	(2.6, 0.0, 0.2)	(4.0, 0.0, 0.6)	(5.6, 0.0, 0.8)	(9.0, NA, 1.0)	

Acknowledgements We thank two anonymous reviewers for their constructive comments which helped improve the clarity of the present work.

Author Contributions TFK was involved in the conceptualisation; all authors contributed to the methodology, formal analysis, writing—original draft preparation and writing—review and editing.

Funding None

Data Availability None

Code Availability R package used: VGAM [23]. R codes for reproducing the analyses are available at https://github.com/Divo-Lee/R-Code/blob/main/Post_dist.R

Declarations

Conflict of interest: None

Ethics approval: Not applicable

Consent to participate: Not applicable

Consent for publication: Both authors agree to publish

References

1. Consul, P.C., Jain, G.C.: A generalization of the Poisson distribution. *Technometrics* **15**(4), 791–799 (1973)
2. Joe, H., Zhu, R.: Generalized Poisson distribution: the property of mixture of Poisson and comparison with negative binomial distribution. *Biom. J.* **47**(2), 219–229 (2005)
3. Nikoloulopoulos, A.K., Karlis, D.: On modeling count data: a comparison of some well-known discrete distributions. *J. Stat. Comput. Simul.* **78**(3), 437–457 (2008)
4. Shoukri, M., Consul, P.: Some chance mechanisms generating the generalized Poisson probability models. In: *Biostatistics*, pp. 259–268. Springer, Dordrecht (1987)
5. Consul, P.C.: *Generalized Poisson Distribution: Properties and Applications*. Marcel Dekker, Marcel Dekker, New York (1989)
6. Srivastava, S., Chen, L.: A two-parameter generalized Poisson model to improve the analysis of RNA-seq data. *Nucleic Acids Res.* **38**(17), 170 (2010)
7. Li, W., Jiang, T.: Transcriptome assembly and isoform expression level estimation from biased RNA-Seq reads. *Bioinformatics* **28**(22), 2914–2921 (2012)
8. Zhang, J., Kuo, C.-C.J., Chen, L.: WemIQ: an accurate and robust isoform quantification method for RNA-seq data. *Bioinformatics* **31**(6), 878–885 (2015)
9. Wang, Z., Wang, J., Wu, C., Deng, M.: Estimation of isoform expression in RNA-seq data using a hierarchical Bayesian model. *J. Bioinform. Comput. Biol.* **13**(06), 1542001 (2015)
10. Low, J.Z.-B., Khang, T.F., Tammi, M.T.: CORNAS: coverage-dependent RNA-Seq analysis of gene expression data without biological replicates. *BMC Bioinformatics* **17**, 575 (2017)
11. Anscombe, F.: The statistical analysis of insect counts based on the negative binomial distribution. *Biometrics* **5**(2), 165–173 (1949)
12. Bliss, C.I., Fisher, R.A.: Fitting the negative binomial distribution to biological data. *Biometrics* **9**(2), 176–200 (1953)
13. White, G.C., Bennetts, R.E.: Analysis of frequency count data using the negative binomial distribution. *Ecology* **77**(8), 2549–2557 (1996)
14. Anders, S., Huber, W.: Differential expression analysis for sequence count data. *Genome Biol.* **11**, 106 (2010)

15. Hardcastle, T.J., Kelly, K.A.: baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* **11**(1), 422 (2010)
16. Wu, H., Wang, C., Wu, Z.: A new shrinkage estimator for dispersion improves differential expression detection in RNA-seq data. *Biostatistics* **14**(2), 232–243 (2013)
17. Love, M.I., Huber, W., Anders, S.: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**(12), 550 (2014)
18. Gupta, P.L., Gupta, R.C., Ong, S.H., Srivastava, H.: A class of Hurwitz-Lerch Zeta distributions and their applications in reliability. *Appl. Math. Comput.* **196**(2), 521–531 (2008)
19. Bateman, H.: *Higher Transcendental Functions*, vol. 1. McGraw-Hill Book Company, New York (1953)
20. Liew, K.W., Ong, S.H., Toh, K.K.: The Poisson-stopped Hurwitz-Lerch zeta distribution. *Communications in Statistics - Theory and Methods* **51**(16), 5638–5652 (2022)
21. Gupta, R.C.: Modified power series distribution and some of its applications. *Sankhya, Series B* **36**(3), 288–298 (1974)
22. Laforgia, A., Natalini, P.: On the asymptotic expansion of a ratio of gamma functions. *J. Math. Anal. Appl.* **389**(2), 833–837 (2012)
23. Yee, T.W.: VGAM: Vector Generalized Linear and Additive Models. (2022). R package version 1.1-7. <https://CRAN.R-project.org/package=VGAM>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.