Approximating the negative moments of the Poisson distribution

C. Matthew Jones[1,2 †] **and Anatoly A. Zhigljavsky**[1 ††]

[1] *School of Mathematics, Cardiff University, CF24 4YH, UK.*

[2] *Cardiff Research Consortium, Cardiff Medicentre, CF14 4UJ, UK.*

Abstract

Through the use of the Stirling numbers of the first kind, a general class of approximations is derived for the negative moments of the Poisson distribution.

Keywords: Negative moments; Stirling numbers of the first kind; Poisson distribution

1 Introduction

The approximation of negative moments of the Poisson distribution is an important practical problem and has attracted significant attention in literature, see for example Chao and Strawderman (1972), Gupta (1979), Stancu (1968), Stephan (1945) and Tiku (1964). The main applications of the first negative moment are related to the theory of mixed Poisson distributions: indeed, if η_i $(i = 1, 2, \ldots n)$ are i.i.d.r.v. with variance σ^2 and the sample size n has the positive Poisson distribution, then the variance of the sample mean $\frac{1}{n}\sum_{i=1}^{n}\eta_i$ is $\sigma^2 \mathrm{E}\,(1/n)$, see e.g. Grab and Savage (1954). The negative moments of the Poisson distribution are known to be useful in life testing problems, as shown in Bartholomew (1957), David and Johnson (1956) and Epstein and Sobel (1954).

A recent application has occurred in the field of multi-centre clinical trials, which involves the simultaneous recruitment of patients to numerous centres. The approximation of negative moments are required in the subsequent calculations of MSE's of the mean treatment effect differences. As demonstrated in Dragalin et al. (2002), three least squared estimators can be defined for the mean treatment difference based on fixed effect models of increasing

† *E-mail address:* matthew.jones@cardiff-research-consortium.co.uk

†† *E-mail address:* ZhigljavskyAA@cardiff.ac.uk

complexity. In particular, the MSE of the Type III estimator is

$$\frac{\sigma^2}{N^2}\sum_{i=1}^{N}\left(\frac{1}{n_{i2}}+\frac{1}{n_{i1}}\right), \tag{1}$$

where N is the number of centres, σ^2 is the variance of the response observed at each centre, and n_{ij} is the number of patients of treatment j at centre i. If n_{ij} are assumed to be positive Poisson random variables then (1) becomes a random variable, with moments that are expressed through the negative moments of the Poisson distribution. Similar phenomena occurs with the Type I and II estimators, however, with a more complicated form. See Fedorov et al. (2003) for further details.

Let $Poisson(\lambda)$ denote the Poisson distribution with parameter λ, and let ξ be a random variable, $\xi \sim Poisson(\lambda)$. Additionally, let ξ_+ be the so-called positive Poisson r.v. with parameter λ; that is,

$$\Pr(\xi_+ = x) = \frac{1}{1-e^{-\lambda}}\frac{\lambda^x}{x!}e^{-\lambda}, \quad \text{for} \quad x = 1, 2\ldots$$

The negative moments of the Poisson distribution are defined as the negative moments of ξ_+:

$$\mu_{-\alpha} = E\left(\frac{1}{\xi_+^{\,\alpha}}\right) = \frac{e^{-\lambda}}{1-e^{-\lambda}}\sum_{x=1}^{\infty}\frac{\lambda^x}{x^\alpha\, x!} \quad \text{for any } \alpha > 0\,. \tag{2}$$

We develop the following k^{th} order approximations for $\mu_{-\alpha}$, with integer α:

$$\mu_{-\alpha} \cong \mu_{-\alpha}^{(k)} = \sum_{u=\alpha}^{k}\frac{s(u,\alpha)}{\lambda^u}\,, \tag{3}$$

where $s(u, \alpha)$ are the so-called signless Stirling numbers of the first kind, see Section 3.1. The case of non-integer α is considered in Section 3.6.

Approximations of the negative moments $\mu_{-\alpha}$ introduced in Tiku (1964) have been cited in a number of reference books; see, for instance, Haight (1966), Johnson et al. (1992). We shall use these approximations as a benchmark for comparison in subsequent calculations, and we will demonstrate that our approximations are more precise. Tiku's approximations for the first negative moment (see Equation (11) in Tiku (1964)) are defined as follows:

$$\mu_{-1} \cong T_j^{(1)} = \frac{\left(1+\sum_{r=3}^{j}\beta_r\right)}{(\lambda-1)(1-e^{-\lambda})}\ ;\ \text{with } T_0^{(1)} = \frac{1}{(\lambda-1)(1-e^{-\lambda})}\,, \tag{4}$$

where $\beta_r = a^{(r)}/\lambda(\lambda+1)(\lambda+2)\ldots(\lambda+r-1),\ r = 3,4,5,\ldots$ with $a^{(3)} = 1,\ a^{(4)} = 7,\ a^{(5)} = 43,\ a^{(6)} = 271$, etc.

Note that a number of alternative approximations to μ_{-1} can be constructed through using the fact that

$$\mu_{-1} = \frac{e^{-\lambda}}{1-e^{-\lambda}} \left(\mathrm{E}i(\lambda) - \log\lambda - \gamma\right) \ , \tag{5}$$

where $\gamma \simeq 0.5772\ldots$ is Euler's constant and $\mathrm{E}_i(\lambda) = \int_{-\infty}^{\lambda}(e^u/u)\,du$ is the exponential integral, see for example Grab and Savage (1954). Note that the relation (5) is often used to approximate values of the exponential integral $\mathrm{E}i(\lambda)$; consequently, this broadens the scope of our approximation into several other fields.

The approximations of Tiku (1964) for the higher order integer negative moments ($\alpha \geq 2$) are defined as

$$\mu_{-\alpha} \cong T^{(\alpha)} = \frac{1}{(\lambda-1)(\lambda-2)\ldots(\lambda-\alpha)} \ . \tag{6}$$

The suggested approximations are far more precise than Tiku's approximations. The criterion by which we determine the accuracy of the approximations is through studying their relative error:

$$Relative\,Error = \frac{Exact\,Value \ - \ Approximate\,Value}{Exact\,Value} \ .$$

Note that positive values for the relative error indicate that the respective approximation is underestimating the true value.

In this paper we are only interested in moderate and large values of λ. For small λ the moments can be easily computed using the definition (2). In particular, the relative error of the following simple approximation

$$\mu_{-\alpha} = \frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{\infty} \frac{\lambda^k}{k^\alpha\,k!} \cong \frac{e^{-\lambda}}{1-e^{-\lambda}} \sum_{k=1}^{3\lambda+10} \frac{\lambda^k}{k^\alpha\,k!} \ ;$$

is smaller than 10^{-10} in absolute value for all $\lambda > 0$ and $\alpha \geq 1$.

2 Deriving recurrence relations and approximations

2.1 General Recurrences

Denote

$$A_{l,m}(\lambda) = \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{x^l (x+m-1)!} \,. \tag{7}$$

In particular,

$$\mu_{-\alpha} = \frac{1}{1-e^{-\lambda}} A_{\alpha,1}(\lambda) \,. \tag{8}$$

For each l, $m > 0$, we obviously have

$$\frac{1}{x^l} = \frac{1}{x^{l-1}(x+m)} + \frac{m}{x^l (x+m)} \,. \tag{9}$$

Through applying Equation (9) to (7) we obtain the basic recurrence:

$$A_{l,\,m}(\lambda) = \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^x}{x^{l-1}(x+m)!} + \sum_{x=1}^{\infty} \frac{m e^{-\lambda}\lambda^x}{x^l (x+m)!} = A_{l-1,\,m+1}(\lambda) + m A_{l,\,m+1}(\lambda) \,, \tag{10}$$

and from (7) we have

$$A_{0,m}(\lambda) \;=\; \frac{1}{\lambda^{m-1}} \left(1 - \sum_{x=0}^{m-1} \frac{e^{-\lambda}\lambda^x}{x!} \right) \,. \tag{11}$$

Note that for any fixed $m \geq 1$ and $\lambda \geq 1$ we have

$$\frac{1}{\lambda^{m-1}} \left(1 - m\,\lambda^{m-1} e^{-\lambda} \right) \leq A_{0,m}\,(\lambda) \leq \frac{1}{\lambda^{m-1}} \tag{12}$$

so that the difference $A_{0,m}(\lambda) - 1/\lambda^{m-1}$ is exponentially small as $\lambda \to \infty$.

Through performing several iterations of the basic recurrence (10), we obtain

$$\begin{aligned} A_{l,\,m}(\lambda) &= A_{l-1,m+1}(\lambda) + m A_{l-1,m+2}(\lambda) + m(m+1) A_{l,m+2}(\lambda) = \dots \\ &= \sum_{r=0}^{N} \frac{(m+r-1)!}{(m-1)!} A_{l-1,m+r+1}(\lambda) + \frac{(m+N)!}{(m-1)!} A_{l,m+N+1}(\lambda) \,. \end{aligned} \tag{13}$$

Rather than using the definition (7), we can define $A_{l,m}(\lambda)$ through the recurrences (13) and the initial conditions (11). Let the quantities $B_{l,m}(\lambda)$, which can be considered as a simplified version of $A_{l,m}(\lambda)$, be defined through the recurrences (13), that is,

$$B_{l,m}(\lambda) = \sum_{r=0}^{N} \frac{(m+r-1)!}{(m-1)!} B_{l-1,m+r+1}(\lambda) + \frac{(m+N)!}{(m-1)!} B_{l,m+N+1}(\lambda) \quad \text{for all } N > 0, \tag{14}$$

and the initial conditions that are similar to (11) but exclude the exponential terms:

$$B_{0,m}(\lambda) \quad = \quad \frac{1}{\lambda^{m-1}} \,. \tag{15}$$

Deriving approximations using $B_{l,m}(\lambda)$ as opposed to $A_{l,m}(\lambda)$ implies that we ignore all the terms that are exponentially decreasing when $\lambda \to \infty$. Since $B_{0,m}(\lambda) = \frac{1}{\lambda^{m-1}} \geq A_{0,m}(\lambda)$ for all m, see (12), we have $B_{l,m}(\lambda) \geq A_{l,m}(\lambda)$ for all $l \geq 0$, $m \geq 1$.

2.2 Deriving approximations and determining their precision

Let m, l and k be fixed. Consider the expansions of $A_{l,m}(\lambda)$ and $B_{l,m}(\lambda)$ of the form:

$$A_{l,m}(\lambda) = \sum_{r=0}^{k} \frac{a_{l,m}(r)}{\lambda^r} + o\left(\frac{1}{\lambda^k}\right), \ \lambda \to \infty \,; \tag{16}$$

$$B_{l,m}(\lambda) = \sum_{r=0}^{k} \frac{b_{l,m}(r)}{\lambda^r} + o\left(\frac{1}{\lambda^k}\right), \ \lambda \to \infty \,, \tag{17}$$

where $a_{l,m}(r)$ and $b_{l,m}(r)$ are some coefficients. The properties of these coefficients are summarized in the following statement.

Theorem 1. *Let $l, k \geq 0$, $m \geq 1$ be some integers and let $A_{l,m}(\lambda)$ and $B_{l,m}(\lambda)$ be defined through* (16) *and* (17), *respectively. Then we have*

(i) $a_{l,m}(r) = b_{l,m}(r)$ for all l, m, r;

(ii) $a_{l,m}(r) = 0$ for all $r < l + m - 1$;

(iii) $a_{l,m}(r) = 1$ for $r = l + m - 1$;

(iv) $a_{l,m}(r) \geq 0$ for all r.

Proof. The statement (i) follows from the fact that for fixed l and m the difference between $A_{l,m}(\lambda)$ and $B_{l,m}(\lambda)$ is exponentially small as $\lambda \to \infty$.

In order to prove the other three statements, we use induction in l. For $l = 0$ the statements obviously follow from either (12) or (15). Assume that (ii) - (iv) hold for $l - 1$. For all fixed $m \geq 1$, $l \geq 1$ we have for all $x \geq 1$:

$$\frac{1}{(x+m)\dots(x+m+l-1)} \leq \frac{1}{x^l} \leq \frac{1}{(x+m)\dots(x+m+l-1)}\left(1 + \frac{C_{m,l}}{x+m+l}\right)$$

for some positive constants $C_{m,l}$; these constants are $C_{m,l} = (m+l+1)\,((m+l)!/m! - 1)$. Through substituting this into (7) we obtain

$$A_{0,\,m+l}(\lambda) \le A_{l,\,m}(\lambda) \le A_{0,\,m+l}(\lambda) + C_{m,\,l} A_{0,\,m+l+1}(\lambda)\,.$$

In view of (12) we obtain as $\lambda \to \infty$

$$\frac{1}{\lambda^{m+l-1}} + o\left(\frac{1}{\lambda^{m+l}}\right) \le A_{l,\,m}(\lambda) \le \frac{1}{\lambda^{m+l-1}} + \frac{C_{m,\,l}}{\lambda^{m+l}}\,;$$

this yields (ii) and (iii).

Assume that (iv) is true for $l-1$, through the recurrence relation in (13) it follows that (iv) is true for l; as it implies that each $a_{l,\,m}(r)$ is a weighted sum of non-negative coefficients $a_{l-1,\,u}(v)$ with non-negative weights. □

In view of (8) and (16), we define the k^{th} order approximations to μ_{-l} as

$$\mu_{-l}^{(k)} = \sum_{r=0}^{k} \frac{a_{l,\,1}(r)}{\lambda^r}\,. \tag{18}$$

The statements (ii) and (iii) of Theorem 1 imply the summation in (18) actually starts at $r = l$, they also imply that as $\lambda \to \infty$ the order of decrease of μ_{-l} is $\frac{1}{\lambda^l}$; more precisely:

$$\mu_{-l} - \frac{1}{\lambda^l} = o\left(\frac{1}{\lambda^l}\right) \text{ as } \lambda \to \infty\,. \tag{19}$$

Statement (iv) yields that for each l, μ_{-l} is monotonously increasing as k grows, that is $\mu_{-l}^{(k)} \le \mu_{-l}^{(k+1)}$ for all k.

Assume that we able to compute the coefficients $a_{l,\,1}(r)$ (which we can, unlike the coefficients $a_{l,\,m}(r)$ for $m > 1$). Thus, we can compute the k^{th} order approximations $\mu_{-l}^{(k)}$ to μ_{-l} with $k > l$, with Theorem 1 implying that the precision is:

$$|\mu_{-l}^{(k)} - \mu_{-l}| = O\left(\frac{1}{\lambda^{k+1}}\right) \text{ as } \lambda \to \infty\,. \tag{20}$$

Formulae (19) and (20) yield that the relative error of $\mu_{-l}^{(k)}$ with $k > l$ is:

$$\left|\frac{\mu_{-l}^{(k)} - \mu_{-l}}{\mu_{-l}}\right| = \frac{O\left(\frac{1}{\lambda^{k+1}}\right)}{\frac{1}{\lambda^l} + o\left(\frac{1}{\lambda^l}\right)} = \frac{O\left(\frac{1}{\lambda^{k-l+1}}\right)}{1 + o(1)} = O\left(\frac{1}{\lambda^{k-l+1}}\right) \text{ as } \lambda \to \infty\,. \tag{21}$$

Despite the rate of decrease (as $\lambda \to \infty$) of the relative error, $\mu_{-l}^{(k)}$ increases as k grows, we cannot take k arbitrarily large. This is attributable to the constant in $O\left(\frac{1}{\lambda^{k-l+1}}\right)$ in (21) rapidly increasing as k grows. This constant, as follows from (2) has the order of at least $s(k+1,\, l)$, which is roughly $k!(\log k)$. Numerical results (shown below) demonstrate that one can guarantee extremely precise approximations by taking k up to $\lceil \lambda \rceil + l$.

Note also that Theorem 1 implies that to construct a k^{th} order approximation to μ_{-l}, we are free to ignore the last terms in (14) and use

$$B_{l,m}(\lambda) \cong \sum_{r=0}^{N} \frac{(m+r-1)!}{(m-1)!} B_{l-1,m+r+1}(\lambda)\,, \tag{22}$$

for any $N \geq k-(m+l)$.

3 Construction of the approximations

In the approximations that follow we will need the so-called *signless Stirling numbers of the first kind.*

3.1 Signless Stirling numbers of the first kind

Denote the signless Stirling numbers of the first kind as $s(n,\, j)$, see Pólya et al. (1983) and Comtet (1974). These numbers satisfy the recurrence

$$s(n,\, j) = (n-1)s(n-1,\, j) + s(n-1,\, j-1), \quad n,\, j \geq 1\,,$$

with the initial conditions $s(n,\, j) = 0$ if $n \leq 0$ or $j \leq 0$, except $s(0,\, 0) = 1$. Values for these numbers can be obtained from numerous combinatorial books, for example Pólya et al. (1983).

We shall need the following properties of these Stirling numbers, $s(n, j)$. Let $H_n = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$ be the harmonic numbers, then (see Comtet (1974), p. 217) $s(n, j) = 0$ for $n < j$ and for all $n \geq j-1$

$$s(n+1, 1) \quad = \quad n!\,; \tag{23}$$

$$s(n+1,2) = n!\left(1+\frac{1}{2}+\ldots+\frac{1}{n}\right) = n!H_n\ ; \tag{24}$$

$$s(n+1,3) = \frac{n!}{2}\left[H_n^2 - \left(1+\frac{1}{2^2}+\ldots+\frac{1}{n^2}\right)\right]\ . \tag{25}$$

Additionally, for n fixed and variable k, $s(n+1, k+1)$ are the elementary symmetric functions of the first n integers; that is, for $p = 1, 2, \ldots n$ we have

$$s(n+1, n+1-p) = \sum_{1\le i_1<i_2<\ldots<i_p\le n} i_1 i_2 \ldots i_p\ . \tag{26}$$

3.2 Approximating the first negative moment

Through using (14) and (15) we have

$$B_{1,1}(\lambda) = \frac{1}{\lambda} + \frac{1}{\lambda^2} + \frac{2}{\lambda^3} + \ldots = \sum_{r=0}^{k-1} \frac{r!}{\lambda^{r+1}} + k!B_{1,k+1}(\lambda)\ . \tag{27}$$

To construct k^{th} order approximations for μ_{-1}, we keep the first k terms in Equation (27), ignoring $B_{1,k+1}(\lambda)$, respectively. In this way we obtain

$$\mu_{-1} \cong B_{1,1}(\lambda) \cong \mu_{-1}^{(k)} = \sum_{r=0}^{k-1} \frac{r!}{\lambda^{r+1}} = \frac{1}{\lambda} + \frac{1}{\lambda^2} + \frac{2}{\lambda^3} + \ldots + \frac{(k-1)!}{\lambda^k}\ . \tag{28}$$

Through using (23) it is evident that we can write $\mu_{-1}^{(k)}$ in the form of (3).

Table I: Comparison of the relative errors of the approximations for the first negative moment, against different values of λ.

λ	Relative Error					
	$\mu_{-1}^{(5)}$	$\mu_{-1}^{(\lceil 0.5\lambda\rceil)}$	$\mu_{-1}^{(\lceil\lambda\rceil)}$	$T_0^{(1)}$	$T_3^{(1)}$	$T_6^{(1)}$
10	1.605 10^{-3}	1.605 10^{-3}	-1.217 10^{-3}	1.686 10^{-2}	1.611 10^{-2}	1.546 10^{-2}
15	2.606 10^{-4}	5.539 10^{-5}	-1.459 10^{-5}	6.117 10^{-3}	5.934 10^{-3}	5.799 10^{-3}
25	1.580 10^{-5}	1.718 10^{-8}	-1.305 10^{-9}	1.912 10^{-3}	1.855 10^{-3}	1.837 10^{-3}
50	4.289 10^{-7}	1.124 10^{-17}	-4.333 10^{-20}	4.350 10^{-4}	4.275 10^{-4}	4.264 10^{-4}
100	1.265 10^{-8}	6.292 10^{-36}	-1.935 10^{-41}	1.042 10^{-4}	1.032 10^{-4}	1.031 10^{-4}
200	3.847 10^{-10}	1.496 10^{-72}	-1.633 10^{-79}	2.551 10^{-5}	2.539 10^{-5}	2.538 10^{-5}

The accuracy of the approximations used are shown in Table I, which gives the relative error against different values of λ; all of our approximations compare favorably to the more complex Tiku approximations $T_j^{(1)}$ defined in Equation (4). Already the simple approximation $\mu_{-1}^{(5)} = \frac{1}{\lambda} + \frac{1}{\lambda^2} + \frac{2}{\lambda^3} + \frac{6}{\lambda^4} + \frac{24}{\lambda^5}$, provides very good precision. The approximations $\mu_{-1}^{(3)}$ and $\mu_{-1}^{(4)}$ are also accurate. Table I shows that the approximation $\mu^{(\lceil 0.5\lambda \rceil)}$ is very accurate for $\lambda \geq 10$, and always underestimates μ_{-1}. On the other hand, $\mu_{-1}^{(k)}$ with $k = \lceil \lambda \rceil$ always overestimates μ_{-1}, and it is even more precise than $\mu_{-1}^{(k)}$ with $k = \lceil \lambda/2 \rceil$. This yields that any approximation $\mu_{-1}^{(k)}$, with $\lceil \frac{\lambda}{2} \rceil \leq k \leq \lceil \lambda \rceil$ provides an extremely accurate approximation.

3.3 Approximating the second negative moment

Using Equation (22) with $l = 2$ and $m = 1$, we obtain

$$\mu_{-2} \cong \sum_{r=0}^{N} r! B_{1,r+2}(\lambda) \,. \tag{29}$$

We will be constructing approximations of the form

$$\mu_{-2} \cong \mu_{-2}^{(k)} = \sum_{r=1}^{k} \frac{a_{2,1}(r)}{\lambda^r} \,,$$

which we shall obtain from (29) with N large enough (any $N > k$ would suffice), by the repeated application of the approximations in (22), and keeping all the terms that contribute to $a_{2,1}(r)$'s with $r \leq k - 1$.

Table II: Comparison of the relative errors for the approximations of the second negative moment, against different values of λ.

λ	Relative Error			
	$\mu_{-2}^{(6)}$	$\mu_{-2}^{(\lceil 0.5\lambda \rceil+1)}$	$\mu_{-2}^{(\lceil \lambda \rceil+1)}$	$T^{(2)}$
10	2.925 10^{-2}	2.925 10^{-2}	-1.158 10^{-2}	9.355 10^{-2}
15	4.148 10^{-3}	1.425 10^{-3}	-2.587 10^{-4}	2.993 10^{-1}
25	2.316 10^{-4}	7.840 10^{-7}	-4.671 10^{-8}	8.338 10^{-3}
50	6.250 10^{-6}	1.157 10^{-15}	-3.824 10^{-18}	1.808 10^{-3}
100	1.849 10^{-7}	1.460 10^{-33}	-4.077 10^{-39}	4.244 10^{-4}
200	5.639 10^{-9}	7.871 10^{-70}	-8.023 10^{-82}	1.030 10^{-4}

Rewrite (22) as

$$B_{l,m}(\lambda) \cong \sum_{r \geq 0} \frac{(m + r - 1)!}{(m - 1)!} B_{l-1,m+r+1}(\lambda) \,, \tag{30}$$

and, respectively, (29) as

$$\mu_{-2} \cong \sum_{r\geq 0} r! B_{1,r+2}(\lambda)\,. \tag{31}$$

Hence, from (31) and (30) we obtain

$$\mu_{-2} \quad\cong\quad \sum_{r\geq 0} r! B_{1,r+2}(\lambda) \cong \sum_{r\geq 0} r! \sum_{s\geq 0} \frac{1}{\lambda^{s+r+2}} \frac{(s+r+1)!}{(r+1)!} = \sum_{r\geq 0} \frac{1}{r+1} \sum_{s\geq 0} \frac{1}{\lambda^{s+r+2}} (s+r+1)!\,.$$

Let $t = r+s+1$; using Equation (24) we obtain

$$\mu_{-2} \quad\cong\quad \sum_{r\geq 0} \sum_{t\geq r+1} \frac{t!}{(r+1)\lambda^{t+1}} = \sum_{t\geq 1} \frac{t!}{\lambda^{t+1}} \sum_{r=0}^{t-1} \frac{1}{r+1} = \sum_{t\geq 1} \frac{t! H_t}{\lambda^{t+1}} = \sum_{t\geq 2} \frac{s(t,2)}{\lambda^t}\,. \tag{32}$$

The k^{th} order approximation to μ_{-2} is obtained when we keep the first k terms in the right-hand side of (32).

In Table II we discover that $\mu_{-2}^{(\lceil 0.5\lambda \rceil + 1)}$, $\mu_{-2}^{(\lceil \lambda \rceil + 1)}$ and even the very simple approximation $\mu_{-2}^{(6)}$ compare very favorable against Tiku's approximation $T^{(2)}$. We find that $\mu_{-2}^{\lceil \lambda \rceil + 1}$ over-estimates μ_{-2}, whereas $\mu_{-2}^{(\lceil 0.5\lambda \rceil + 1)}$ understimates, this suggests any approximation $\mu_{-2}^{(k)}$ with $\lceil \frac{\lambda}{2} \rceil + 1 \leq k \leq \lceil \lambda \rceil + 1$ would provide an extremely accurate approximation.

3.4 Approximating the third negative moment

To approximate the third negative moment, μ_{-3}, we express $B_{3,1}(\lambda)$ in a similar way:

$$\begin{aligned}
\mu_{-3} \quad&\cong\quad B_{3,1}(\lambda) \cong \sum_{r\geq 0} r! B_{2,r+2}(\lambda) \cong \sum_{r\geq 0} r! \sum_{s\geq 0} B_{1,r+s+3}(\lambda) \frac{(r+s+1)!}{(r+1)!} \\
&\cong\quad \sum_{r\geq 0} \frac{1}{r+1} \sum_{s\geq 0} (r+s+1)! \sum_{t\geq 0} \frac{1}{\lambda^{t+r+s+3}} \frac{(r+s+t+2)!}{(r+s+2)!} \\
&=\quad \sum_{r\geq 0} \frac{1}{r+1} \sum_{s\geq 0} \frac{1}{r+s+2} \sum_{t\geq 0} \frac{1}{\lambda^{t+r+s+3}} (r+s+t+2)!\,.
\end{aligned}$$

Let $u = r+s+t+2$ and $s' = r+s+1$, so that

$$\mu_{-3} \quad\cong\quad \sum_{u\geq 2} \frac{u!}{\lambda^{u+1}} \sum_{r=0}^{u-2} \sum_{s'=r+1}^{u-1} \frac{1}{(r+1)(s'+1)}\,. \tag{33}$$

Since

$$H_u^2 = \sum_{r=0}^{u-1}\frac{1}{r+1}\sum_{s=0}^{u-1}\frac{1}{s+1} = \sum_{r=0}^{u-2}\frac{1}{r+1}\sum_{s=r+1}^{u-1}\frac{1}{s+1} + \sum_{r=0}^{u-1}\left[\frac{1}{r+1}\right]^2 + \sum_{r=1}^{u-1}\frac{1}{r+1}\sum_{s=0}^{r-1}\frac{1}{s+1}$$
$$= 2\sum_{r=0}^{u-2}\frac{1}{r+1}\sum_{s=r+1}^{u-1}\frac{1}{s+1} + \left(1+\frac{1}{2^2}+\frac{1}{3^2}+\ldots+\frac{1}{u^2}\right), \quad (34)$$

we have from (33) and (25)

$$\mu_{-3} \cong \sum_{u\geq 2}\frac{u!}{2\lambda^{u+1}}\left[H_u^2 - \left(1+\frac{1}{2^2}+\ldots+\frac{1}{u^2}\right)\right] = \sum_{u\geq 2}\frac{s(u+1,3)}{\lambda^{u+1}} = \sum_{u\geq 3}\frac{s(u,3)}{\lambda^u}. \quad (35)$$

The k^{th} order approximation $\mu_{-3}^{(k)}$ is obtained by keeping the first k terms in (35).

Table III: Comparison of the relative errors for approximations of the third negative moment, against different values of λ.

λ	Relative Error			
	$\mu_{-3}^{(7)}$	$\mu_{-3}^{(\lceil 0.5\lambda\rceil+2)}$	$\mu_{-3}^{(\lceil\lambda\rceil+2)}$	$T^{(3)}$
10	1.941 10^{-1}	1.941 10^{-1}	-2.834 10^{-2}	3.159 10^{-1}
15	3.337 10^{-2}	1.720 10^{-2}	-1.440 10^{-3}	9.704 10^{-2}
25	1.669 10^{-3}	1.719 10^{-5}	-5.923 10^{-7}	2.287 10^{-2}
50	4.534 10^{-5}	5.567 10^{-14}	-1.360 10^{-16}	4.700 10^{-3}
100	1.358 10^{-6}	1.590 10^{-31}	-3.855 10^{-37}	1.081 10^{-3}
200	4.169 10^{-8}	1.963 10^{-67}	-1.924 10^{-79}	2.597 10^{-4}

In Table III we discover that all approximations compare favorably against Tiku's approximation $T^{(3)}$, and if we had and approximation of the form $\mu_{-3}^{(k)}$ with $\lceil 0.5\lambda\rceil + 2 \leq k \leq \lceil\lambda\rceil + 2$, the approximation would be extremely accurate.

3.5 General Case

For the l^{th} order negative moment we approximate $B_{l,1}(\lambda)$. This can be obtained by applying the same methodology:

$$\mu_{-l} \cong \tilde{\mu}_{-l} = \sum_{r_1\geq 0}\frac{1}{r_1+1}\sum_{r_2\geq 0}\frac{1}{r_1+r_2+2}\cdots\sum_{r_{l-1}\geq 0}\frac{1}{r_1+r_2+\ldots+r_{l-1}+l-1}\sum_{r_l\geq 0}\frac{(r_1+r_2+\ldots+r_l+l-1)!}{\lambda^{r_1+r_2+\ldots+r_l+l}}.$$

Let $r'_m = \sum_{i=1}^{m} r_i + (m-1)$, for $2 \leq r \leq l$.

Thus, we have

$$\tilde{\mu}_{-l} = \sum_{r'_l \geq l-1} \frac{r'_l!}{\lambda^{r'_l+1}} \sum_{r_1=0}^{r'_l-(l-1)} \sum_{r'_2=r_1+1}^{r'_l-(l-2)} \sum_{r'_3=r'_2+1}^{r'_l-(l-3)} \cdots \sum_{r'_{l-1}=r'_{l-2}+1}^{r'_l-1} \frac{1}{(r_1+1)(r'_2+1)\ldots(r'_{l-1}+1)} .$$

Let $t_1 = r_1 + 1$ and $t_m = r'_m + 1$, $2 \leq m \leq l$. Hence, using Equation (26)

$$\begin{aligned} \tilde{\mu}_{-l} &= \sum_{t_l \geq l-1} \frac{t_l!}{\lambda^{t_l+1}} \sum_{t_1=1}^{t_l-(l-2)} \sum_{t_2=t_1}^{t_l-(l-1)} \sum_{t_3=t_2}^{t_l-(l-2)} \cdots \sum_{t_{l-1}=t_{l-2}}^{t_l-2} \frac{1}{t_1\, t_2\, \ldots\, t_{l-1}} \\ &= \sum_{t_l \geq l-1} \frac{s(t_l+1, l)}{\lambda^{t_l+1}} = \sum_{r \geq l} \frac{s(r,l)}{\lambda^r} . \end{aligned} \tag{36}$$

Similarly to the above, the k^{th} order approximation to μ_{-l} is obtained when we keep the first k terms in (36).

3.6 Approximating non-integer negative moments

Assume that $\alpha > 0$ is not an integer. Set $l = \lfloor \alpha \rfloor$ and let $\theta = \alpha - l$ with $0 \leq \theta \leq 1$. Accurate approximations for non-integer negative moments can be obtained through using the following approximation:

$$\mu_{-\alpha}^{(t+\alpha)} = \sum_{r=0}^{t} \frac{s(r+1,\, l) + \theta\, [s(r+l+1,\, l+1) - s(r+1,\, l)]}{\lambda^{\alpha+r}} . \tag{37}$$

Table IV clearly demonstrates that all of the approximations for the non-integer negative moments provide very accurate approximations for $\lambda \geq 50$, with a relative error less than 0.003. However, it can also be seen that for larger values of α (e.g. $\alpha = 5.1$ and 5.5) the approximations are nor very accurate for small λ. This can be explained by the fact that the approximations for the larger integer negative moments are only accurate for fairly large values of λ. Additionally, we find that the approximations are less accurate as θ departs further away from zero or one.

Table IV: Comparison of relative errors of the non-integer negative moments for different values of α versus different values of λ, for $t = 5$.

λ	Relative Error					
	$\mu_{-0.1}$	$\mu_{-0.5}$	$\mu_{-2.1}$	$\mu_{-2.5}$	$\mu_{-5.1}$	$\mu_{-5.5}$
10	-6.872 10^{-3}	-1.959 10^{-2}	8.463 10^{-4}	2.122 10^{-3}	8.303 10^{-1}	8.916 10^{-1}
15	-3.852 10^{-3}	-1.095 10^{-2}	-5.835 10^{-3}	-1.633 10^{-2}	5.832 10^{-1}	7.441 10^{-1}
25	-2.072 10^{-3}	-5.834 10^{-2}	-3.299 10^{-3}	-9.216 10^{-3}	1.288 10^{-2}	1.762 10^{-1}
50	-9.629 10^{-3}	-2.692 10^{-2}	-1.225 10^{-3}	-3.416 10^{-3}	-1.853 10^{-3}	-5.351 10^{-3}
100	-4.652 10^{-3}	-1.296 10^{-2}	-5.245 10^{-4}	-1.461 10^{-3}	-7.002 10^{-4}	-1.946 10^{-3}
200	-2.287 10^{-4}	-6.363 10^{-3}	-2.429 10^{-4}	-6.756 10^{-4}	-2.835 10^{-4}	-7.878 10^{-4}

Acknowledments

We kindly acknowledge the support of this work by GlaxoSmithKline. We are especially grateful to Valerii Fedorov and Byron Jones. We are also very grateful to our Cardiff University colleague Nikolai Leonenko for useful advice. Finally, we thank the referee for useful comments.

References

Chao M.T. and Strawderman W.E., (1972). Negative moments of positive random variables, *Journal of the American Statistical Association* **67**, 429–431.

Comtet, L., (1974). *Advanced combinatorics: the art of finite and infinite expansions*, (D. Reidel Publishing Company, Dordrecht.).

David F. N. and Johnson N.L., (1956). Reciprocal Bernoulli and Poisson varaibles, *Metron* **18** (No.1-2), 77–81.

Dragalin F., Fedorov V., Jones B. and Rockhold F, (2002). Estimation of the combined response to treatment in multicentre trials, *Journal of Biopharmaceutical Statistics* **11**, 275–295.

Epstein B. and Sobel M., (1954). Some theorems relevant to life testing, *Annals of Mathematical Statistics* **25**, 373–381.

Fedorov V., Jones B., Jones C.M. and Zhigljavsky, (2003). Multicentre clinical trials with random enrolment: theoretical approximations to the moments of the mean squared errors of estimators, *GlaxoSmithkline Technical Report*, GSK BDS Technical Report 2003.

Grab E.L. and Savage I.R., (1954). Tables of the expected value of 1/X for positive Bernoulli and Poisson variables, *Journal of the American Statistical Association* **49**, 169–177.

Gupta R.C., (1979). On negative moments of generalized Poisson distribution, *Math. Operationsforsch. Statist.* **10**, No. 1, 169–172.

Haight F. A., (1966). *Handbook of the Poisson distribution*, (Wiley, N.Y.).

Johnson N., Kotz S. and Kemp A., (1992). *Univariate discrete distributions*, (Wiley, N.Y.).

Pólya G., Tarjan R. E., Woods. D. R., (1983). *Notes on introductory combinatorics*, (Birkḧuser, Boston).

Stancu D. D. ,(1968). On the moments of negative order of the positive Bernoulli and Poisson variables, *Studia Universitatis Babes Bolyai Series, Math. Phys.* **1**, 29–31.

Stephan F. F. ,(1945). The expected value and variance of the reciprocal and other negative powers of a positive Bernoullian variate, *Annals of Mathematical Statistics* **16**, 50–61.

Tiku M. L., (1964). A note on the negative moments of a truncated poissson variate, *Journal of the American Statistical Association*, **59**, 1220–1224.