THE DIFFERENCE CALCULUS AND THE NEGATIVE BINOMIAL DISTRIBUTION

K.O. BOWMAN[1] and L.R. SHENTON[2]

[1]Computational Sciences and Engineering Division
Oak Ridge National Laboratory, P.O.Box 2008, 4500N, MS-6191
Oak Ridge, TN 37831-6191, U.S.A.
e-mail: bowmanko@ornl.gov

[2]Department of Statistics, University of Georgia
Athens, Georgia 30602, U.S.A.

May 11, 2007

Abstract

In a previous paper we state the dominant term in the third central moment of the maximum likelihood estimator $\hat{k}$ of the parameter k in the negative binomial probability function where the probability generating function is $(p+1-pt)^{-k}$. A partial sum of the series $\sum 1/(k+x)^3$ is involved, where x is a negative binomial random variate. In expectation this sum can only be found numerically using the computer. Here we give a simple definite integral in (0,1) for the generalized case. This means that now we do have a valid expression for $\sqrt{\beta_{11}(\hat{k})}$ and $\sqrt{\beta_{11}(\hat{p})}$. In addition we use the finite difference operator Δ, and $E = 1 + \Delta$ to set up formulas for low order moments. Other examples of the operators are quoted relating to the orthogonal set of polynomials associated with the negative binomial probability function used as a weight function.

2000 Mathematics Subject Classification: 62E20.
Key words and phrases: asymptotic skewness, maximum likelihood, Maple code, moments, operators E and Δ, Vandermond.

1 Introduction

In a previous paper (Bowman and Shenton, 2007) we have given a formulas for the basic skewness of the maximum likelihood estimators $\hat{k}$, $\hat{p}$ of k, p in sampling from the probability function whose probability generating function is $(p+1-pt)^{-k}$, $k > 0$, $p > 0$. The skewness of $\hat{k}$, for example, is $\sqrt{\beta_{11}(\hat{k})}$ which is the $1/\sqrt{N}$ coefficient of the skewness itself, N the sample size. In fact from Bowman and Shenton (Section (4.4) and (5), 2007), we have

$$\sqrt{\beta_{11}(\hat{k})} = [Var_1(\hat{k})]^{3/2}\left\{-2\mathcal{E}_3(x)] - 3\frac{\partial F_1(k,p)}{\partial k} + \frac{3r}{kp}\frac{\partial F_1(k,p)}{\partial r} - \frac{4r(1+r)}{k^2}\right\}$$

where

$$r\frac{\partial F_1(k,p)}{\partial k} = -\sum_{x=1}^{\infty}\frac{r^x}{x}\frac{(x-1)!\Gamma(k)}{\Gamma(k+x)}[\psi(k+x)-\psi(k)],$$

and

$$r\frac{\partial F_1(k,p)}{\partial r} = \sum_{x=1}^{\infty}\frac{r^x(x-1)!\Gamma(k)}{\Gamma(k+x)} \quad (0 < r < 1.k > 0)$$

and $\psi(k+x) - \psi(x) = \frac{1}{k} + \frac{1}{k+1} + \cdots + \frac{1}{k+x-1}$, $x = 1, 2, \ldots$.

For the skewness of $\sqrt{\beta_{11}(\hat{p})}$, the triplet $T_r(\alpha\beta\gamma) = \{-2\mathcal{E}[S_3(x)] - \cdots\}$ terms idendical to $\hat{k}$ displayed in above equation. Only the multiplier triplet of L^{ij} will be different.

$$B(kkk) = -Var_1(\hat{k})^3\frac{p^3}{k^3}, \quad B(kkp) = Var_1(\hat{k})^2\frac{p^2}{k^2}\left(\frac{pq}{k} + \frac{p^2}{k^2}Var_1(\hat{k})\right),$$

$$B(kpp) = -Var_1(\hat{k})\frac{p}{k}\left(\frac{pq}{k} + \frac{p^2}{k^2}Var_1(\hat{k})\right)^2, \quad B(ppp) = \left(\frac{pq}{k} + \frac{p^2}{k^2}Var_1(\hat{k})\right)^3.$$

The term $-2\mathcal{E}S_3 = -2\mathcal{E}S_3(x,k)$ in algebraic form has remained a problem, although numerically it is easily evaluated using the Maple implementation. In this study we have discovered a simple definite integral form exploiting the fact that

$$\frac{1}{k^j} = \frac{1}{(j-1)!}\int_0^{\infty} e^{-k\omega}\omega^{j-1}d\omega. \quad (k > 0, j = 1, 2, \ldots).$$

In another direction we consider the moments of the random variate $1/(k+x)$ using the finite difference operators Δ and $E = 1+\Delta$. Simple illustrations of the application of these operators are given for examples $\mathcal{E}(x^r)$, the rth noncentral moment, $\mathcal{E}(x-kp)^r$, the rth central moment. Cumulants of the negative binomial variate are also mentioned. A brief mention is given for the work of A.C. Aitken and H.T. Gronin, relating to Vandermond and Gregory-Newton.

2 Basic formulas

2.1 The probability function

The probability generating function for the negative binomial distribution we are studying is

$$B(t; k, p) = (p + 1 - pt)^{-k} \quad (p > 0, k > 0)$$

and

$$Pr(X = x) = (q)^{-k} \left(\frac{p}{q}\right)^x \frac{\Gamma(k+x)}{\Gamma(k)x!}. \quad (q = p + 1, x = 0, 1, \ldots) \tag{1}$$

We may sometimes set $r = p/q = p/(p+1)$, so that $0 < r < 1$ and

$$Pr(X = x) = (1 - r)^k r^x \frac{\Gamma(k+x)}{\Gamma(k)x!}.$$

The probability generating function (p.g.f.) is

$$(p + 1 - pt)^{-k}. \tag{2}$$

2.2 Finite difference operators E and Δ and moments

Johnson and Kotz (1969) quote a formula for μ'_j (the $j$the non-central moment) in the form

$$\mu'_j = (1 - p\Delta)^{-k} 0^j.$$

They prove this starting with

$$\mu'_j = \sum_{s=1}^{\infty} s^j \binom{k+s-1}{k-1} \left(\frac{p}{q}\right)^s \left(1 - \frac{p}{q}\right)^k$$

allowing for our notation. They then say "Formally" (2) follows. Why formally?

A simpler demonstration is

$$\mu'_j = \mathcal{E}x^j - \mathcal{E}E^x(\underline{0}^j) = (1 - p\Delta)^{-k}(\underline{0}^j),$$

fundamentally two lines. $\mathcal{E}$ represents expectation.

For central moments we have

$$\mu_j = \mathcal{E}(x - kp)^j = \mathcal{E}E^x(\underline{0} - kp)^j = (1 - p\Delta)^{-k}(\underline{0} - kp)^j.$$

The underlined elements are those on which the E operator works. Factorial moments $\mu_{[j]}$ are found from

$$\mathcal{E}(1+\alpha)^x = [p+1-p(1+\alpha)]^{-k} = (1-p\alpha)^{-k}.$$

Then $\mu_{[1]} = kp$, $\mu_{[2]} = k(k+1)p^2$ and so on.

Cumulants can be set up using the equation (moments about the mean)

$$e^{\frac{\kappa_2\alpha^2}{2!}+\frac{\kappa_3\alpha^3}{3!}+\frac{\kappa_4\alpha^4}{4!}+\cdots} = \frac{e^{-kp\alpha}}{(p+1-pe^{\alpha})^k},$$

and taking logarithmic derivatives with respect to α, leads to

$$\frac{1}{k}\left(\frac{\kappa_2\alpha^2}{1!}+\frac{\kappa_3\alpha^2}{2!}+\frac{\kappa_4\alpha^3}{3!}+\cdots\right) = -p+\frac{pe^{\alpha}}{q-pe^{\alpha}} = -q+\frac{q}{q-pe^{\alpha}}.$$

Hence

$$\left\{1+\frac{1}{kq}\left(\frac{\kappa_2\alpha^2}{1!}+\frac{\kappa_3\alpha^2}{2!}+\frac{\kappa_4\alpha^3}{3!}+\cdots\right)\right\}H(\alpha,p)=1$$

where

$$H(\alpha,p) = 1-p\alpha-\frac{p^2\alpha^2}{2!}-\cdots-\frac{p^j\alpha^j}{j!}-\cdots.$$

Equating coefficients of $\frac{\alpha^j}{j!}$ to zero leads to the recurrence for cumulants, namely

$$\kappa_{j-1} = p\left\{\binom{j}{1}\kappa_j+\binom{j}{2}\kappa_{j-1}+\cdots+\binom{j}{j-1}\kappa_2+kq\right\}$$

for $j=1,2,\ldots$, κ_1 being zero.

In the case of a geometric distribution, ($k=1$), the reader is referred to is Shenton and Bowman (2001).

2.3 The shape of the probability function

Johnson and Kotz (1969) sketch the probability function in four cases (p.128); $(k,p) = (1,2)$, (2,1), and (5,0.4), the fourth case being $kp=2$.

Some properties

(i) If $kp > q$ there is a mode at the least integer not less than kpq; two equal modes if $kp=q$.

(ii) If $kp < q$ the mode is at $k=0$.

These brief descriptions seem sufficient to demonstrate that shape characteristics of the probability function may be complicated.

2.4 Further formulas involving E and Δ

For the binomial random variate, Aiken and Gonen (1934-1935) find the expression

$$P_j(x) = (1+p\Delta)^{-(n-j+1)}x^{(j)} \quad (j=1,2,\ldots;0<p<1)$$

for a member of the orthogonal set $\{P_j(x)\}$ with

$$\Delta P_j(x) = jP_{j-1}(x:X,n-1),$$

and the inverse,

$$x^{(j)} = (1+p\Delta)^{n-j+1}P_j(x).$$

For a Poisson variate,

$$\kappa_j(x) = e^{-m\Delta}x^{(j)}$$

with $\Delta\kappa_j = j\kappa_{j-1}$ and inverse $x^{(j)} = e^{m\Delta}\kappa_j(x)$. These results have been, to some extent, overlooked in the literature partly because of the unfortunate titles referring to "fourfold sampling".

A year earlier than the above paper (vol. LIM, 1932-1933), Aiken published a paper "On the Graduation of Data by the Orthogonal Polynomials of Least Squares". In the first few pages he set out the fundamental properties of the operators Δ and E (note that Euler in the eighteenth century mentions Δ). Aiken defines 'summation by parts' and the half interval difference $\delta u_x = u_{x+\frac{1}{2}} - u_{x-\frac{1}{2}}$, and also states several identities including some due to Vandermond. Examples:

Identities in Factorial Polynomial.

(xxi) $(x+m)^{(r)} = x^{(r)} + rx^{(r-1)}m + r_{[1]}x^{(r-2)}m^{(2)} + \cdots + m^{(r)}$.

(xxii) $(x+m)_{(r)} = x^{(r)} + x_{(r-1)}m + x_{(r-2)}(m+1)_{(2)} + \cdots + m_{(r)}$.

(xxiii) $(x-m)^{(r)} = x^{(r)} - rx^{(r-1)}m + r_{[1]}x^{(r-2)}(m+1)^{(2)} - \cdots + (-)^r(m+r-1)^{(r)}$.

(xxiv) $(x-m)_{(r)} = x^{(r)} - x_{(r-1)}m + x_{(r-2)}(m+1)_{(2)} - \cdots + (-)^r(m+r-1)_{(r)}$.

The above give various forms of Vandermonds's familiar algebraic forms. or these it is generally assured that r is a positive integer, x and m reals in symbollic forms we may work

$$(x+m)^{(r)} = E^m(x+0)^{(m)} = (1+\Delta)^m x^{(r)},$$
$$(x-m)^{(r)} = E^{-m}(x-0)^{(m)} = (1+\Delta)^{-m} x^{(r)}.$$

Aiken seems to regard these identities as examples of the Gregory-Nelson interpolation formulas.

In the sequel we shall consider the psi function difference $\psi(k+x) - \psi(k) = \frac{1}{k} + \frac{1}{k+1} + \cdots + \frac{1}{k+x-1}$, $(k > 0, x = 1, 2, \ldots)$ and its derivatives using the finite dirrerence calculus.

2.5 Skewness, $\sqrt{\beta_{11}(\hat{k})}$, $\sqrt{\beta_{11}(\hat{p})}$

In a recent paper, Bowman and Shenton (2007), we have studied the skewness (μ_3/σ^3) of maximum likelihood estimators $\hat{k}$, $\hat{p}$ for k, p respectively. The final approximation for the skewness of $\hat{k}$ is

$$\sqrt{\beta_{11}(\hat{k})} = [Var_1(\hat{k})]^{3/2} \left\{ -2E[S_3(x)] - 3\frac{\partial F_1(k,p)}{\partial k} + \frac{3r}{kp}\frac{\partial F_1(k,p)}{\partial r} - \frac{4r(1+r)}{k^2} \right\}$$

where

$$\frac{\partial F_1(k,p)}{\partial k} = -\sum_{x=1}^{\infty} \frac{r^x}{x} \frac{(x-1)!\Gamma(k)}{\Gamma(k+x)} [\psi(k+x) - \psi(k)],$$

and

$$r\frac{\partial F_1(k,p)}{\partial r} = \sum_{x=1}^{\infty} \frac{r^x (x-1)!\Gamma(k)}{\Gamma(k+x)} \qquad (0 < r < 1, k > 0)$$

and $\psi(k+x) - \psi(x) = \frac{1}{k} + \frac{1}{k+1} + \cdot + \frac{1}{k+x-1}$, $x = 1, 2, \ldots$.

In this the term $-2\mathcal{E}S_3(x)$ occurs. Now

$$S_3(x) \equiv S_3(x,k) = \frac{1}{k^3} + \frac{1}{(k+1)^3} + \cdots + \frac{1}{(k+x_1)^3} \qquad (k > 0, x = 1, 2, \ldots).$$

We could not find an exact value for its expectation, namely

$$(1+p)^{-k} \sum_{x=1}^{\infty} \left(\frac{p}{q}\right)^x \frac{\Gamma(k+x)}{\Gamma(k)x!} \left(\frac{1}{k^3} + \frac{1}{(k+1)^3} + \cdots + \frac{1}{(k+x-1)^3}\right)$$

but noted it is simpler to use

$$(1-r)^k \sum_{x=1}^{\infty} \frac{r^x \Gamma(k+x)}{x!\Gamma(k)} \left(\frac{1}{k^3} + \frac{1}{(k+1)^3} + \cdots + \frac{1}{(k+x-1)^3}\right)$$

where $r = p/(1+p)$, so $0 < r < 1$; clearly convergence now seems possible.

Using the Maple code implementation we found the first few terms for this and similar expressions, in powers of r. Here are examples,

$$\mathcal{E}(S_1 S_2) \ \text{coeff.}r = \frac{1}{k^2}$$

$$\text{coeff.}r^2 = \frac{2k^3 + 2k^2 + 2k + 1}{2k^2(k+1)^2}$$

$$\text{coeff.}r^3 = \frac{3k^5 + 21k^4 + 42k^3 + 36k^2 + 24k + 8}{6k^2(k+1)^2(k+1)^2}$$

$$\text{coeff.}r^4 = \frac{4k^7 + 31k^6 + 273k^5 + 723k^4 + 971k^3 + 726k^2 + 396k + 108}{12k^2(k+1)^2(k+2)^2(k+3)^2}$$

$$\text{coeff.}r^5 = \frac{A}{60k^2(k+1)^2(k+2)^2(k+3)^2(k+4)^2}$$

$$A = 15k^9 + 310k^8 + 2760k^7 + 13940k^6 + 42985k^5 + 80430k^4 + 88760k^3 + 60000k^2 + 28800k + 6912$$

$$\mathcal{E}(S_3) \;\; \text{coeff.}r = \frac{1}{k^2}$$

$$\text{coeff.}r^2 = \frac{k^2 - k - 1}{2k^2(k+1)^2}$$

$$\text{coeff.}r^3 = \frac{3k^3 + 3k^2 - 6k - 4}{3k^2(k+1)^2(k+2)^2}$$

$$\text{coeff.}r^4 = \frac{11k^4 + 42k^3 + 13k^2 - 66k - 36}{4k^2(k+1)^2(k+2)^2(k+3)^2}$$

$$\text{coeff.}r^5 = \frac{2(25k^5 + 190k^4 + 395k^3 - 10k^2 - 600k - 288)}{5k^2(k+1)^2(k+2)^2(k+3)^2(k+4)^2}$$

It is difficulat to derive a pattern so further terms are not considered.

3 The psi function difference, its derivatives and associated moments

Using the operators Δ and E of the finite difference calculus, we have

$$\begin{aligned}\mathcal{E}\frac{1}{(k+x)^j} &= \mathcal{E}E^x\frac{1}{k^j} = (1-p\Delta)^{-k}\frac{1}{k^j} \qquad (j = 1, 2, \ldots)\\ &= \frac{1}{k^j} - \frac{kp}{1!}\left(\frac{1}{k^j} - \frac{1}{(k+1)^j}\right) + \frac{k(k+1)p^2}{2!}\left(\frac{1}{k^j} - \frac{1}{(k+1)^j} + \frac{1}{(k+2)^j}\right) - \cdots\\ &= \frac{1}{(j-1)!}\int_0^1 \frac{t^{k-1}(\ln\frac{1}{t})^{j-1}dt}{(p+1-pt)^k},\end{aligned}$$

valid for $k > 0$, $p > 0$, $j = 1, 2, \ldots$. Using moments for $j = 1, 2, 3, 4$ we can set up the skewness (μ_3/σ^3) and kurtosis (μ_4/σ^4) for the negative binomial random variate $1/(k+x)$. Figures showing the structure and form of these measures are given in Figure 1.

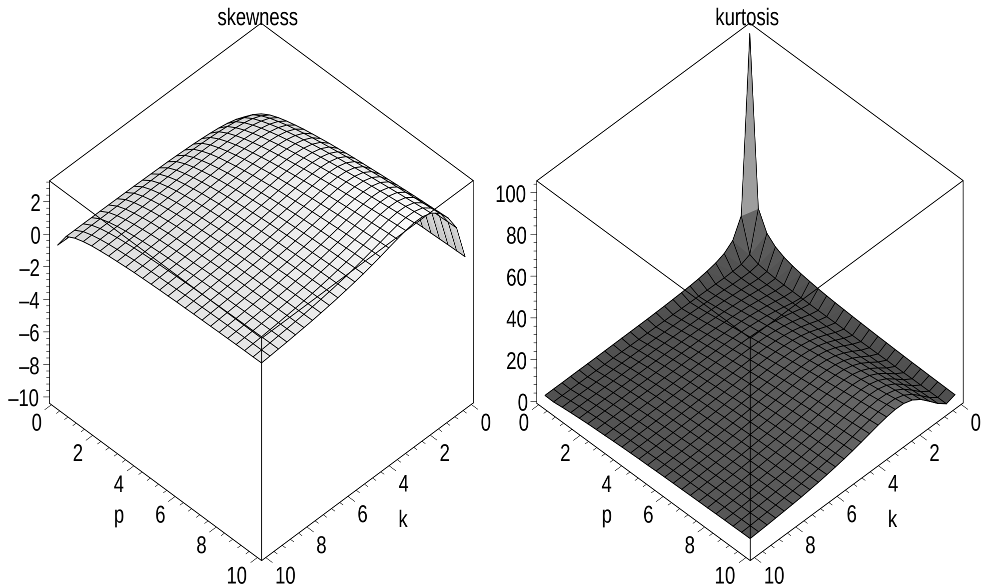

Figure 1: $\sqrt{\beta}_1 \mathcal{E}\left(\frac{1}{k+x}\right)$ and $\beta_2 \mathcal{E}\left(\frac{1}{k+x}\right)$

Tabulation of the moments are given in Table 1.

Table 1. Moments of the variate $\frac{1}{k+x}$ of the negative binomial distribution

k	p	0.1	0.5	1.0	5.0	10.0
0.1	μ_1'	9.9136	9.6351	9.3809	8.4515	7.9693
	σ	0.8834	1.7943	2.3116	3.4980	3.9040
	$\sqrt{\beta_1}$	-10.1227	-4.7161	-3.4691	-1.8200	-1.4061
	β_2	103.4818	23.2540	13.0468	4.3222	2.9859
0.5	μ_1'	1.9371	1.7408	1.5708	1.0288	0.7998
	σ	0.2853	0.5498	0.6733	0.8240	0.8095
	$\sqrt{\beta_1}$	-4.3280	-1.6835	-0.9754	0.2522	0.7123
	β_2	19.8195	3.9288	2.0532	1.2042	1.6809
1.0	μ_1'	0.9531	0.8109	0.6931	0.3584	0.2398
	σ	0.1491	0.2735	0.3190	0.3169	0.2736
	$\sqrt{\beta_1}$	-2.9115	-0.8618	-0.2301	1.1700	1.8887
	β_2	9.6847	1.9890	1.3453	2.9941	5.5811
5.0	μ_1'	0.1846	0.1416	0.1098	0.0396	0.0221
	σ	0.0212	0.0334	0.0335	0.0183	0.0112
	$\sqrt{\beta_1}$	-1.0487	0.1558	0.6252	1.7541	2.2195
	β_2	3.0766	2.3058	3.0432	8.2987	12.5965
10.0	μ_1'	0.0917	0.0688	0.0525	0.0181	0.0100
	σ	0.0081	0.0121	0.0116	0.0056	0.0033
	$\sqrt{\beta_1}$	-0.6732	0.2299	0.5679	1.2067	1.3774
	β_2	2.8220	2.7234	3.3027	5.6950	6.7217

Comments: The coefficient of variation ($\sigma/mean$) is in general small, especially if k is small. The general structure is difficult to describe and the main interest is to give evidence that the Δ and E approach to moment description works.

4 Integrals for $\mathcal{E}S_j(x,k)$

By definition

$$S_j(x,k) = \frac{1}{k^j} + \frac{1}{(k+1)^j} + \cdots + \frac{1}{(k+x-1)^j} \qquad (x = 1, 2, \ldots)$$

and is related to the psi function

$$\psi_{j-1}(x) - \psi_{j-1}(x,k)$$

$\psi(x) = \frac{d}{dx} \ln \Gamma(x)$, $\psi_1(x) = \frac{d}{dx}\psi(x)$ and so on

We have

$$\begin{aligned}
\mathcal{E}[S_j(x,k)] &= \mathcal{E}\left\{\frac{1}{k^j} + \frac{1}{(k+1)^j} + \cdots + \frac{1}{(k+x-1)^j}\right\} \\
&= \frac{1}{(j-1)!}\mathcal{E}\left\{\int_0^\infty (e^{-k\omega}\omega^{j-1} + e^{-k\omega}\omega^{j-1}e^{-\omega} + e^{-k\omega}\omega^{j-1}e^{-2\omega} + \cdots + e^{-k\omega}\omega^{j-1}e^{-(x-1)\omega})d\omega\right\} \\
&= \frac{1}{(j-1)!}\mathcal{E}\left\{\int_0^\infty (e^{-k\omega}\omega^{j-1}(1 + e^{-\omega} + \cdots + e^{-\omega(x-1)})d\omega\right\} \\
&= \frac{1}{(j-1)!}\mathcal{E}\int_0^\infty \frac{e^{-k\omega}\omega^{j-1}(1-e^{-\omega x})d\omega}{1-e^{-\omega x}}
\end{aligned}$$

and from (2)

$$\begin{aligned}
\mathcal{E}S_j(x,k) &= \frac{1}{(j-1)!}\int_0^\infty \frac{e^{-k\omega}\omega^{j-1}(1-(p+1-pe^{-\omega})^{-k}d\omega}{1-e^{-\omega}} \\
&= \frac{1}{(j-1)!}\int_0^1 \frac{t^{k-1}(\ln\frac{1}{t})^{j-1}[1-(p+1-pt)^{-k}]dt}{1-t} \qquad (3)
\end{aligned}$$

where $(j = 1, 2, \cdots, k > 0, p > 0)$. This new form is readily evaluated, the range being $0 < t < 1$, and checks against the tabulation of $S_1(x,k)$, $S_2(x,k)$, $S_3(x,k)$ given in section 3.2.3 of Bowman and Shenton (2007); it also provides an exact expression for the term $-2\mathcal{E}S_3(x,k)$ appearing in section 4.4 for the skewness of $\hat{k}$, the maximum likelihood estimator of k. The skewness $\sqrt{\beta_{11}(\hat{k})}$ refer to the normed value of the second order term in the third central moment. A similar term appears in $\sqrt{\beta_{11}(\hat{p})}$.

There is interest in expression (3) when $j = 1$ and for $j = 2$. Thus for $j = 1$

$$\mathcal{E}S_1(x,k) = \int_0^1 \frac{t^{k-1}[1-(p+1-pt)^{-k}]dt}{1-t} \quad (k > 0, p > 0) \qquad (4)$$

But it is readily shown from the probability generating function (3) that $\mathcal{E}S_1(x,k) = \ln(p+1)$, for $p > 0$. Thus (4) is independent of k provided $k > 0$.

From the generating function (2) we have when $j = 2$

$$(1-r)^k \sum_{x=1}^{\infty} \frac{r^x\Gamma(k+x)}{x!\Gamma(k)} S_1(x,k) = \ln(p+1)$$

leading to

$$\mathcal{E}[S_1(x,k)]^2 - \mathcal{E}S_2(x,k) = \ln^2(p+1)$$

so the left side of this expression does not depend on k. Note that (4) defines $\mathcal{E}S_2(x,k)$ as an integral and is equal to Fisher's i_{kk} (Fisher, 1941).

Lastly, we may remined readers that Fisher's i_{kk} parameter is given by Fisher as

$$\begin{aligned} i_{kk} &= \frac{r}{k} + \frac{r^2}{2k(k+1)} + \frac{4r^3}{6k(k+1)(k+2)} + \cdots \qquad (r = \frac{p}{p+1}) \\ &= \sum_{x=1}^{\infty} \frac{r^x (x-1)!}{x} \frac{\Gamma(k)}{\Gamma(k+x)} \qquad (p > 0, k > 0, 0 < r < 1) \end{aligned}$$

and for the most part its validity depends on pattern recognition, actually the first three terms in the series. Our result in (3) depends on manipulation of

$$i_{kk} = \sum_{x=0}^{\infty} \frac{1}{q^k} \left(\frac{p}{q}\right)^x \frac{\gamma(k+x)}{\Gamma(k)x!} \left[\psi_1(k) - \psi_1(k+x)\right),$$

where the psi function difference is

$$\frac{1}{k^2} + \frac{1}{(k+1)^2} + \cdots + \frac{1}{(k+x-1)^2}. \qquad (k > 0, x = 1, 2, \ldots)$$

which relates to (3) with $j = 2$.

5 Conclusion

The integral in (4) above for $\mathcal{E}S_j(x,k)$, namely

$$\mathcal{E}S_j(x,k) = (1+p)^{-k} \sum_{x=1}^{\infty} \left(\frac{p}{q}\right) \frac{\Gamma(k+x)}{x!\Gamma(k)} \left(\frac{1}{k^j} + \frac{1}{(k+1)^j} + \cdots + \frac{1}{(k+x-1)^j}\right) \qquad (5)$$

is simple in form and probably new, we have not found it in Nielsen's (1965) *Handbook for the Gamma Functions* (first published in 1906).

Now it is exactly proved from the generating function in §1.1 that

$$\mathcal{E}S_1(x,k) = \int_0^{\infty} t^{k-1} \frac{1 - [1 + p(1-t)]^{-k} dt}{1-t} = \ln(p+1) \qquad (6)$$

for $k > 0$, $p > 0$. The integral involves k yet is independent of $\Re k$ for all $k > 0$,

For completness we must add that there are many identities involving $S_1(x,k)$, $S_2(x,k)$ and so on; see Bowman and Shenton (1965, p.30). For example

$$2\mathcal{E}S_3(x,k) = 3\mathcal{E}S_1(x,k)S_2(x,k) - \mathcal{E}[S_1(x,k)]^3 + \ln^3(p+1).$$

Note that a portion of this will be independent of k, $k > 0$. Howeveer we have been unable to find integral forms like expression (6) for expressions such as products $\mathcal{E}[S_i(x,k)S_j(x,k)]$, i, j being positive integers.

Reference

Aitken, A.C. (1932-1933). On the graduaton of data by the orthogonal polynomial of least squres, *Proc.Royal.Soc.Edinbugh*, Vol. LII, 54-78.

Aitken, A.C. and Gonin, H.T. (1934-1935). On fourfold sampling with and withoutreplacement, *Proc.Roy.Soc.Edinbugh*, Vol. LV, 114-125.

Bowman, K.O. and Shenton, L.R. (1965). *Asymptotic Covariances for the Maximum likelihood estimators of the Parameters of a Negative Binomial Distribution*, K-1643, Union Carbide Nuclear Division, Oak Ridge, TN

Bowman, K.O. and Shenton, L.R. (2007). Skewness for maximum likelihood estimators of the negative binomial distribution, *Far East Journal of Theoretical Statistics*, submitted.

Fisher, R.A. (1941), The negative binomial distribution, *Annals of Eugenics*, XI(II), 182-187.

Johnson, N.L. and Kotz, S. (1969). *discrete Distributions*, Houghton Mifflin Co., Boston.

Nielsen, N. (1965). *Handbook for the Gamma Functions* (printed in English), Chelsea Publishing Co, New York.

Shenton, L.R. and Bowman, K.O. (2001), The geometric distribution, cumulanats, Eulerian numbers, and the logarithmic function, *Far East Journal of Theoretical Statistics*, 5(1), 113-142.