Far East J.Theo.Stat.,23(1),103-117.

BINOMIAL TEST STATISTICS USING PSI FUNCTIONS

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.

August 3, 2007

Abstract

For the negative binomial model (probability generating function $(p+1-pt)^{-k}$) a logarithmic derivative is the Psi function difference $\psi(k+x)-\psi(k)$; this and its derivatives lead to a test statistic to decide on the validity of a specified model. The test statistic uses a data base so there exists a comparison available between theory and application. Note that the test function is not dominated by outliers.

Applications to (i) Fisher's tick data, (ii) accidents data, (iii) Weldon's dice data are included.

2000 Mathematics Subject Classification: 62E20.
Key words and phrases: accident data, logarithmic derivatives, maximum likelihood estimator, moment estimator, Weldon's dice data.

1 Introduction

In a recent paper (Bowman and Shenton, 2007a, 2007b) on the skewness of maximum likelihood estimators of the parameters of the negative binomial distribution (nbd), we have introduced the function of Psi functions

$$S_1(x,k) = \psi(k+x) - \psi(k), \qquad (k>0, x=0,1,\cdots)$$

which is equivalent to

$$\begin{aligned} S_1(x,k) &= \frac{1}{k} + \frac{1}{k+1} + \frac{1}{k+2} + \cdots + \frac{1}{k+x-1}, \qquad (x=1,2,\cdots) \\ &= 0 \ \text{ when } \ x=0. \end{aligned} \tag{1}$$

Now from a previous study (Bowman and Shenton, 1965, p30) we have the expectations

$$\begin{aligned} E[S_1(x,k)] &= L = \ln(p+1), \\ E[xS_1(x,k)] &= p + pkL, \\ E[S_1^2(x,k)] &= L^2 + E[S_2(x,k)] \end{aligned} \tag{2}$$

where

$$\begin{aligned} S_2(x,k) &= \frac{1}{k^2} + \frac{1}{(k+1)^2} + \cdots + \frac{1}{(k+x-1)^2}, \qquad (x=1,2,\cdots) \\ &= 0 \ \text{ when } \ x=0 \end{aligned}$$

the probability generating function of the negative binomial distribution being $(p+1-pt)^{-k}$.

2 A generalization of $S_1(x,k)$

We consider

$$\begin{aligned} S_j(x,k) &= \frac{1}{k^j} + \frac{1}{(k+1)^j} + \cdots + \frac{1}{(k+x-1)^j} \qquad (j=1,2,\cdots; k>0, x=0,1,\cdots) \\ &= 0 \ \text{ when } \ x=0, \end{aligned}$$

noting that it takes the value zero when $x=0$, it being related to the Psi functions

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

We have

$$E[S_j(x,k)] = E\left\{\frac{1}{k^j} + \frac{1}{(k+1)^j} + \cdots + \frac{1}{(k+x-1)^j}\right\}$$

$$= \frac{1}{(j-1)!} E\left\{\int_0^\infty (e^{-k\omega}\omega^{j-1} + e^{-k\omega}\omega^{j-1}e^{-\omega} + e^{-k\omega}\omega^{j-1}e^{-2\omega} + \cdots + e^{-k\omega}\omega^{j-1}e^{-(x-1)\omega}) d\omega\right\}$$
$$= \frac{1}{(j-1)!} E\left\{\int_0^\infty e^{-k\omega}\omega^{j-1}(1 + e^{-\omega} + \cdots + e^{-\omega(x-1)}) d\omega\right\}$$
$$= \frac{1}{(j-1)!} E \int_0^\infty \frac{e^{-k\omega}\omega^{j-1}(1 - e^{-\omega x}) d\omega}{1 - e^{\omega}}$$

$$= \frac{1}{(j-1)!} \int_0^1 \frac{t^{k-1}(\ln \frac{1}{t})^{j-1}}{1-t} \left\{1 - \frac{1}{(p+1-pt)^k}\right\} dt, \quad (k > 0, p > 0) \tag{3}$$

after the transformation $e^{-\omega} = t$.

In particular when $j = 1$

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

and from (2)

$$\begin{aligned} ES_1(x,k) &= E\{\psi(k+x) - \psi(k)\} \\ &= \int_0^1 \frac{t^{k-1}}{1-t} \left\{1 - \frac{1}{(p+1-pt)^k}\right\} dt \\ &= \ln(p+1) \quad (k > 0, p > 0). \end{aligned}$$

The result may be proved by differentiating with respect to p and setting

$$\omega = t/(p+1-pt). \quad (\omega = 1, t = 1; \omega = 0, t = 0)$$

Then

$$t = \omega(p+1)/(1+\omega p),$$

and

$$\frac{dt}{dp} = \frac{p+1}{(1+\omega p)^2}.$$

Finally

$$\frac{d}{dp}[ES_1(x,k)] = \frac{1}{(p+1)^2}(1+p)k \int_0^1 \omega^{k-1} d\omega = \frac{1}{p+1}.$$

So

$$E[S_1(x,k)] = \ln(p+1) + \text{constant},$$

and letting $p \to 0$ shows the constant is zero.

In a sense the integral is independent of k for $k > 0$. For several pairs of values (k, p) the result has been checked numerically by computer. It is easily checked algebraically for $k = 1$ and $k = 2$. Expanding $[1 + p(1-t)]^{-k}$ in powers of $(1-t)$, we find

$$\begin{aligned} E[S_1(x,k)] &= \frac{p}{k} - \frac{p^2}{k+1} + \frac{p^3}{k+1} - \frac{p^4}{k+3} + \cdots \quad (k > 0, 0 < p \leq 1) \\ &= \ln(p+1) \text{ when } k = 1. \end{aligned}$$

(A reader has pointed out that we have assumed the sum of integrands equals the integrand of sums)

3 A peripheral remark on skewness

In Bowman and Shenton (2007a, 2007b) we state formulas for the skewness of the maximum likelihood estimators $\hat{k}$, $\hat{p}$; see §4.4 in the paper. The skewness for $\hat{k}$ appears as the third standard central moment (μ_s/σ^3) and its $1/\sqrt{N}$ coefficient, where N the sample size. All four terms in $\sqrt{\beta_{11}(\hat{k})}$ are defined algebraically except for the term $-2E[S_3(x,k)]$. Expression (3) of the present paper provides an answer. The skewness of the estimator $\hat{k}$ can now be set up numerically by computer and also in algebraic form.

4 Test statistics for binomial, Poisson, and negative binomial distributions

4.1 Negative binomial distribution

The probability generating function for the negative binomial distribution is

$$(p+1-pt)^{-k} \qquad (k>0, p>0)$$

and the probability function contains the factor $\Gamma(k+x)/\Gamma(k)$ with logarithmic derivative in Psi functions $S_1(x,k)$.

Suppose the statistical model for a data base is the negative binomial distribution. Then we have the test function

$$T_1(\underline{n};\bar{k}) = \frac{n_1}{N}\left(\frac{1}{\bar{k}}\right) + \frac{n_2}{N}\left(\frac{1}{\bar{k}} + \frac{1}{\bar{k}+1}\right) + \frac{n_3}{N}\left(\frac{1}{\bar{k}} + \frac{1}{\bar{k}+1} + \frac{1}{\bar{k}+2}\right) + \cdots \tag{4}$$

for a sample of size N with frequencies $n_1, n_2, \cdots$; $\bar{k}$ may refer to either k or p.

Now from §1 or ORNL Report 1643

$$E[S_1(x,k)] = \ln(p+1)$$

where p is taken to be a consistent estimator of p. From (2)

$$Var[S_1(x,k)] = E[S_2(x,k)],$$

and from Fisher (1941),

$$Var[S_1(x,k)] = i_{kk} = \sum_{x=1}^{\infty} \frac{r^x}{x} \frac{(x-1)!\Gamma(k)}{\Gamma(k+x)}, \qquad (k>0, r=p/q=p/(p+1)).$$

A formula for $\mu_3[S_1(x,k)]$ is given in the appendix. In general numerical values of the moments may be calculated directly from

$$\mu'_s[S_1(x,k)] = \frac{1}{(p+1)^k}\sum_{x=1}^{\infty}\left(\frac{p}{q}\right)\frac{\Gamma(k+x)}{x!\Gamma(k)}[S_1(x,k)]^s,$$

and central moments using the usual correction formulas. For the statistic in (4), the moments will be

$$\begin{aligned}\mu'_1(T_1) &= \mu'_1[S_1(x,k)],\\ \mu_2(T_1) &= \mu_2[S_1(x,k)/N] = \sigma^2(T_1),\\ \sqrt{\beta_1(T_1)} &= \sqrt{\beta_1[S_1(x,k)]}/\sqrt{N},\\ \beta_2(T_1) &= 3 + \frac{\beta_2[S_1(x,k)] - 3}{N}.\end{aligned}$$

Thus a 4-moment approximating distribution to T_1 may be set up (Pearson (see Bowman and Shenton, 1979a, 1979b), S_U or S_B of the Johnson (1949) system). The significance of the standardized difference

$$\bar{T_1} = \frac{T_1 - E(T_1)}{\sigma(T_1)}$$

is considered..

Example 1

Two sets of data considered by Fisher (1941) correcting ticks on sheep, samples of 60 and 82.

Maximum likelihood estimators $(\hat{k},\hat{p})$ and moment estimators (k^*,p^*) are considered.

Table 1 Fisher's data $N = 60$, $\hat{k} = 3.7513$, $\hat{p} = 0.8651$

x	0	1	2	3	4	5	6	7	8	9	10
n_x	7	9	8	13	8	5	4	3	0	1	2
$\frac{1}{k+x-1}$	0	0.27	0.21	0.17	0.15	0.13	0.11	0.10	0.09	0.09	0.08
$S_1(x,k)$	0	0.27	0.48	0.65	0.80	0.93	1.04	1.14	1.24	1.32	1.40
$S_1(x,k)n_x/N$	0	0.04	0.06	0.14	0.11	0.08	0.07	0.06	0	0.02	0.05

$\ln(\hat{p}+1) = 0.6233$ and $T_1(\underline{n};\hat{k},\hat{p}) = 0.6240$. If we use moment estimators $k^* = 3.9567$ and $p^* = 0.8213$ that $\ln(p^*+1) = 0.5996$ and $T_1(\underline{n};k^*,p^*) = 0.5984$. Moments of the $E[S_1(x,k)]$ are shown in the Table 2.

Table 2 Moments of the $E[S_1(x,k)]$

	μ'_1	μ_2	σ	$\sqrt{\beta_1}$	β_2
$S_1(x,\hat{k})$	0.6233	0.1305	0.3612	0.1567	2.5481
$S_1(x,k^*)$	0.5996	0.1198	0.3461	0.1682	2.5662

Comment: There is good agreement between the test function $T_1(\underline{n}; k, p)$ and its expected value $\ln(p+1)$, using maximum likelihood, or moment estimators. Fisher (1941) was mainly interested in estimator's efficiency.

For the case of Fishers second data with sample size $N = 82$,

Table 3 Fisher's data $N = 82$.

	Estimator of k	Estimator of p	$\ln(p+1)$	$T_1(\underline{n}; k, p)$
Maximum likelihood	1.7775	3.6918	1.5458	1.5457
Moment	1.5261	4.2992	1.6676	1.6988

For $N = 82$, this sample has a "long tail", there being sheep with as many as 25 ticks. Outliers for moment procedures may play a significant role. The $\bar{T}_1$ test reduces the influences of outliers. We have $T_1(\underline{n}; \hat{k}, \hat{p}) = 1.5457$ and $T_1(\underline{n}; k^*, p^*) = 1.6988$ supporting the conjecture that a negative binomial distribution model is appropriate.

4.2 The Poisson distribution

Example 2

Kendall (1977) considered an accident data set using the Poisson distribution.

$$Pr(X = x) = \frac{e^{-\theta}\theta^x}{x!}, \qquad (\theta > 0, x = 0, 1, \cdots)$$

with $\hat{\theta} = \theta^* = \sum(n_x/N) = 0.4652$.

$$\begin{aligned} E[S_1(x,\theta)] &= E\int_0^1 \frac{t^{\theta-1}(1-t^x)dt}{1-t} \\ &= \int_0^1 \frac{t^{\theta-1}(1-e^{\theta(t-1)})dt}{1-t} \end{aligned}$$

and $S_1(x,\theta) = \frac{1}{\theta} + \frac{1}{\theta+1} + \cdots + \frac{1}{x+\theta-1}$.

Table 4 Accident data $N = 647$ using Poisson distribution.

x	0	1	2	3	4	5 and over
n_x	447	132	42	21	3	2
$\frac{1}{\theta+x-1}$	0	2.1495	0.6825	0.4056	0.2886	0.2240
$S_1(x,\theta)$	0	2.1495	2.8320	3.2376	3.5262	3.7502
$S_1(x,\theta)n_x/N$	0	0.4385	0.1838	0.1051	0.0164	0.0116
$S_2(x,\theta)n_x/N$	0	4.6204	5.0861	5.2507	5.3339	5.3841

so that $ES_1(x,\theta) = 0.8594$ (computed by integral) and $T_1(\theta) = 0.7554$ s discrepancy is not surprising since the Poisson fit is known to be a poor fit. For $E[S_2(x,\theta)] = \sum_1^5 S_2(x,\theta)n_x/N = \mu_2'(S_1(x,\theta))$.The standard deviation was computed with the usual formula and $\mu_2 = 0.7461$; for sample size of $N = 647$, $\sigma = 0.03396$.

4.3 Test statistic for a binomial distribution ($pgf = (pt+q)^n$)

4.3.1 Test statistics

The random variables and statistics are:

$$\begin{cases} S_1(x, n^+) = \frac{1}{n} + \frac{1}{n+1} + \cdots + \frac{1}{n+x-1}, \\ T_1(\underline{n}; n^+) = \frac{n_1}{N}\left(\frac{1}{n}\right) + \frac{n_2}{N}\left(\frac{1}{n} + \frac{1}{n+1}\right) + \cdots; \end{cases}$$

$$\begin{cases} S_1(x; p) = \frac{1}{p} + \frac{1}{p+1} + \cdots + \frac{1}{p+x-1}, \\ T_1(\underline{n}; p) = \frac{n_1}{N}\left(\frac{1}{p}\right) + \frac{n_2}{N}\left(\frac{1}{p} + \frac{1}{p+1}\right) + \cdots; \end{cases}$$

$$\begin{cases} S_1(x, n^-) = \frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{n-x+1}, \\ T_1(\underline{n}; n^-) = \frac{n_1}{N}\left(\frac{1}{n}\right) + \frac{n_2}{N}\left(\frac{1}{n} + \frac{1}{n-1}\right) + \cdots; \end{cases}$$

there being a binomial data base $(n_0, n_1, \cdots)$ for a sample of size N. Similar expressions involving jth powers occur. For example

$$S_j(x, n^+) = \frac{1}{n^j} + \frac{1}{(n+1)^j} + \cdots + \frac{1}{(n+x-1)^j}. \quad (j = 1, 2, \cdots)$$

4.3.2 Integrals and series

We have, in expectation

$$\begin{aligned} E[S_1(x; n^+)] &= E\int_0^1 (t^{n-1} + t^n + \cdots + t^{n+x-1})dt \\ &= E\int_0^1 \frac{t^{n-1}(1-t^x)dt}{1-t} \\ &= \int_0^1 \frac{t^{n-1}\{1-(pt+q)^n\}dt}{1-t}; \end{aligned}$$

similarly

$$ES_j(x; n^+) = \frac{1}{(j-1)!}\int_0^1 \frac{t^{n-1}(\ln\frac{1}{t})^{j-1}}{1-t}\left\{1-(pt+q)^n\right\}dt.$$

Omitting details we also have

$$ES_j(x; p) = \frac{1}{(j-1)!}\int_0^1 \frac{t^{p-1}(\ln\frac{1}{t})^{j-1}}{1-t}\left\{1-(pt+q)^n\right\}dt, \quad (j = 1, 2, \cdots)$$

$$ES_j(x; n^-) = \frac{1}{(j-1)!}\int_0^1 \frac{(p+qt)^n - t^n}{1-t}(\ln\frac{1}{t})^{j-1}dt, \quad (j = 1, 2, \cdots)$$

and in particular

$$ES_1(x, n^-) = \int_0^1 \frac{(p+qt)^n - t^n}{1-t}dt = p + \frac{p^2}{2} + \cdots + \frac{p^n}{n}. \quad (0 < p < 1)$$

The last case is found by differentiating the integral with respect to p.

4.3.3 Moments of random variable

For the three forms of the test statistics, low order moments can be found using computers.

For $S_1(x; n^+)$

$$\mu_r'[S_1(x, n^+)] = \sum_{x=1}^{n} \binom{n}{x} p^x q^{n-x} \left(\frac{1}{n} + \frac{1}{n+1} + \cdots + \frac{1}{n+x-1} \right)^r,$$

and from $r = 1, 2, 3, 4$ central moments are set up. For example $\mu_2 = \mu_2' - (\mu_1')^2$. For the corresponding test statistic

$$\sigma^2(T_1) = \mu_2[S_1(x, n^+)]/N,$$
$$\sqrt{\beta_1(T_1)} = \sqrt{\beta_1[S_1(x, n^+)]}/\sqrt{N},$$

and $\beta_2 = 3 + \frac{\beta_2[S_1(x,n^+)]-3}{N}$; $\sqrt{\beta_1}$ and β_2 are sometimes defined as α_3 and α_4. In this case the standardized test statistic is

$$\bar{T}_1 = \frac{T_1(\underline{n}; n^+) - E[T_1(\underline{n}; n^+)]}{\sigma[T_1(\underline{n}; n^+)]}.$$

4.3.4 Application to dice data

Example 3 W.F. Weldon's dice data (Kendall and Stuart(1977))

Table 5 Frequency-distribution of 26,306 throws of 12 dice,
the occurrence of a 5 or 6 being counted a success

No.of successes	Observed frequency	Theoretical frequency from the binomial 26,306 $(0.6623 + 0.3377)^{12}$	No.of successes	Observed frequency	Theoretical frequency from the binomial 26,306 $(0.6623 + 0.3377)^{12}$
0	185	187	6	3067	3043
1	1149	1146	7	1331	1330
2	3265	3215	8	403	424
3	5475	5465	9	105	96
4	6114	6269	10 and over	18	15
5	5194	5115	11	0	1
			12	0	0
			Totals	26306	26306

If the dice were perfect (a condition rarely realized in practice) the proportion p of successes would be $\frac{1}{3}$; and the appropriate binomial would be in the form of binomial distribution given in Kendal and Stuart (1977), $(\frac{2}{3} + \frac{1}{3})^{12}$. In this particular case the dice were not quite perfect, the proportion of cases exhibiting a 5 or 6 being 0.3377. Taking this as the value of p, we get the frequency function $(0.6623 + 0.3377)^{12}$, which when multiplied by the total frequency 26.306 gives the theoretical frequencies shown in the third column Table 5. The agreement with observation is evidently fairly good.

Test statistics moments for the three cases are:

Table 6. The moments of three test statistics

	$E(\frac{1}{n}+\frac{1}{n+1}+\cdots+\frac{1}{n+x-1})$	$E(\frac{1}{p}+\frac{1}{p+1}+\cdots+\frac{1}{p+x-1})$	$E(\frac{1}{n}+\frac{1}{n-1}+\cdots+\frac{1}{n-x+1})$
T_1	0.2962	4.3265	0.4120
$\sigma(T_1)$	0.0024	0.0411	0.0032
$\sqrt{\beta_1(T_1)}$	-0.0090	-0.0083	-0.0076
$\beta_2(T_1)$	3.0000	2.9999	3.0000
$\bar{T}_1$	-0.0154	-0.0083	0.0209

The last row indicates that the discrepancies are not significant.

5 A second test for the negative binomial distribution model

We define the test statistic T_2 as

$$T_2(\underline{n},k)=\frac{n_0}{N}\left(\frac{1}{k}\right)+\frac{n_1}{N}\left(\frac{1}{k+1}\right)+\frac{n_2}{N}\left(\frac{1}{k+2}\right)+\cdots$$

and is associated with random variable

$$s(x)=\frac{1}{k+x}\qquad(x=0,1,\cdots,k>0)$$

where x refers to the negative binomial (it also may refer to other binomial distributions defined on $x=0,1,2,\cdots$). Non-central moment are

$$\mu_s'(s(x))\equiv\frac{1}{(j-1)!}\int_0^1\frac{t^{k-1}(\ln\frac{1}{t})^{j-1}dt}{(p+1-pt)^k},\qquad(k>0,p>0)$$

and in particular the mean is

$$\mu_1'(s(x))=\int_0^1\frac{t^{k-1}dt}{(p+1-pt)^k}.$$

We can set up a 4-moment approximating distribution and consider the significant if the standardized statistic

$$\bar{T}_2=\frac{T_2-E(\bar{T}_2)}{\sigma(\bar{T}_2)}.$$

Since the procedure follows closely that of the first test statistic T_1, our description here has been accordingly attenuated.

6 Remarks in conclusion

The random variable

$$S_1(x,k) = \frac{1}{k} + \frac{1}{k+1} + \cdots + \frac{1}{k+x-1} \qquad (k>0)$$

has a distribution when x has a probability with support from the set of non-negative integers $(0,1,2,\cdots)$. The distributions considered are the binomial, Poisson, negative binomial with probability generating functions $(pt+q)^n$, $\exp(\theta(t-1))$, $(p+1-pt)^{-k}$.

There is a statistical version of (1), using a data base leading to the test statistic

$$T_1(\underline{n},k) = \frac{n_1}{N}\left(\frac{1}{k}\right) + \frac{n_2}{N}\left(\frac{1}{k} + \frac{1}{k+1}\right) + \cdots + .$$

Notice that the zero frequency n_0 does not appear.

The moments of the test statistic T_1, assuming existence, are readily set up using a computer; from time to time algebraic or integral forms may be available as checks. Although the from of the test is peculiar to the negative binomial distribution, there are many applications relating to discrete distributions. Some examples are:

(i) Contagious distributions such as a Neyman forms.

(ii) Compound distributions, such as Poisson-binomial with probability generating function $\exp\{\theta[(pt+q)^n - 1]\}$.

(iii) Mixture distribution as published by Everett and Hands (1981).

The binomial itself has two forms involving expressions such as

$$\int_0^1 \frac{t^{n-1}(1-(pt+q)^n)dt}{1-t}$$

and

$$\int_0^1 \frac{(p+qt)^n - t^n}{1-t} dt.$$

Appendix

Moments of $S_1(x,k)$, $S_2(x,k)$ and similar terms and the negative binomial distribution

A.1 The third central moment of $S_1(x,k)$ from ORNL-1643, we have

$$E(S_1^3) = 3E(S_1S_2) - 2E(S_3) + L^3 \qquad (L = \ln(1+p))$$

and similarly

$$E(S_1S_2) = 2E(S_3) + \frac{\partial}{\partial k}E(S_2) + LE(S_2)$$

so

$$\mu_3(S_1) = 3\left\{2E(S_2) + \frac{\partial}{\partial k}E(S_2) + LE(S_2)\right\}$$
$$-2E(S_3) + L^3 - 3LE(S_1)^2 + 2L^3$$
$$= 4E(S_3) + 3\frac{\partial}{\partial k}E(S_2) - L^3$$

A.2 Other cases. We have not been able to discover an algebraic form for the fourth central moment; the problem relates to $E(S_2^2)$; similar forms such as $E(S_3^2)$ appear intractable algebraically.

Reference

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. (1979a). Approximate percentage points for Pearson distributions, *Biometrika*, 66, 147-151.

Bowman, K.O. and Shenton, L.R. (1979b). Further approximate Pearson percentage points and Cornish-Fisher, *Comm. Stat. Part B. Simulation and Computation*, B(8), 231-244.

Bowman, K.O. and Shenton, L.R. (2007a). Skewness for maximum likelihood estimators of the negative binomial distribution, *Far East Journal of Theoretical Statistics*, 22(1),103-129.

Bowman, K.O. and Shenton, L.R. (2007b). *The Difference Calculus and the Negative Binomial Distribution*, Technical Report Number 2007-1, Department of Statistics, University of Georgia.

Everitt, B.S. and Hand, D.J. (1981). *Finite Mixture Distributions*, Chapman and Hall, London.

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

Johnson, N.L. (1949). System of frequency curves generated by methods of translation, *Biometrika*, 36, 149-176.

Kendall, M.G. and Stuart, A. (1977). *The Advanced Theory of Statistics, Volume 1, Distribution Theory, Fourth edition*, MacMillan Publishing Co., Inc., New York.

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