

On Size-Biased Negative Binomial Distribution and its Use in Zero-Truncated Cases

Khurshid Ahmad Mir

Department of Statistics, Govt. Degree College (Boys), Baramulla, Kashmir India.

E-mail: khrshdmir@yahoo.com

A size-biased negative binomial distribution, a particular case of the weighted negative binomial distribution, taking the weights as the variate values has been defined. A Bayes' estimator of size-biased negative binomial distribution (SBNBD) has been obtained by using non-informative and gamma prior distributions. Also comparison has been made of this estimator with the corresponding maximum likelihood estimator (MLE) with the help of R- Software.

Keywords: size-biased negative binomial distribution, Bayes' estimator, R- Software.

1. INTRODUCTION

NEGATIVE MOMENTS have been under study for quite some time. Many authors have worked on the negative moments of discrete distributions, mainly binomial, Poisson, geometric and negative binomial distributions (see [1]-[13]). The first negative moment of the hyper-Poisson distribution and the condition under which this moment is identical to that of the Poisson distribution was studied, for details see [14]. An expression for $E(X + A)^{-1}$, where X has a negative binomial distribution was derived and later, a relation between inverse ascending factorial moments of the geometric distribution and the negative moments of the negative binomial distribution was studied, for details see ([12], [13]). Pascal defined the negative binomial distribution as

$$P_1(X=x) = \binom{m+x-1}{x} \alpha^x (1-\alpha)^m; x=0,1,2,\dots \quad (1)$$

where $0 < \alpha < 1, m > 0$

The mean and variance of the distribution are given as

$$\mu'_1 = \frac{m\alpha}{1-\alpha} \quad (2)$$

$$\mu_2 = \frac{m\alpha}{(1-\alpha)^2} \quad (3)$$

It can be easily seen that variance is always greater than the mean.

The zero-truncated negative binomial distribution is given by

$$P_2(X=x) = \frac{\binom{m+x-1}{x} \alpha^x}{(1-\alpha)^{-m} - 1}; x=1,2,\dots \quad (4)$$

In this paper, a size-biased negative binomial distribution is defined. Negative moments and negative factorial moments of size-biased negative binomial distribution in terms of hypergeometric series function are derived. Recurrence relation of negative ascending factorial moments is also derived so that higher moments are easily calculated. The distribution is comparatively easy to be estimated and hence it can be used as a simple alternative to the zero-truncated negative binomial

distribution. Comparison among different estimation methods is also studied.

2. SIZE-BIASED NEGATIVE BINOMIAL DISTRIBUTION

A size-biased negative binomial distribution is defined by the probability function as

$$P_3(X=x) = \binom{m+x-1}{x-1} \alpha^{x-1} (1-\alpha)^{m+1}; x=1, 2 \quad (5)$$

where $0 < \alpha < 1, m > 0$

The moment generating function of the distribution (5) is given as

$$M_x(t) = e^t (1-\alpha)^{m+1} (1-\alpha e^t)^{-(m+1)} \quad (6)$$

The mean and variance of the distribution (5) can be obtained by using the relation (6) and are given as

$$\mu'_1 = \frac{m\alpha + 1}{1-\alpha} \quad (7)$$

$$\mu_2 = \frac{m\alpha + \alpha}{(1-\alpha)^2} \quad (8)$$

3. ESTIMATION OF PARAMETERS

In this section, we discuss the various estimation methods for size-biased negative binomial distribution and verify their efficiencies.

3.1. Method of maximum likelihood estimation

The log-likelihood function of the size-biased negative binomial distribution is given as

$$L = \prod_{i=1}^n \binom{m+x_i-1}{x_i-1} \alpha^{y-n} (1-\alpha)^{mn+n} \quad (9)$$

$$\log L = \sum_{i=1}^n \log \binom{m+x_i-1}{x_i-1} + (y-n) \log \alpha + (mn+n) \log(1-\alpha) \quad (10)$$

Where $y = \sum_{i=1}^n x_i$

And the likelihood equation is given as

$$\frac{\partial}{\partial \alpha} \log L = \frac{(y-n)}{\alpha} - \frac{(mn+n)}{(1-\alpha)} = 0$$

$$\hat{\alpha} = \frac{y-n}{mn+y} \quad (11)$$

3.2. Method of moments

Replacing μ'_1 from (7) by \bar{x} , we get

$$\hat{\alpha} = \frac{\bar{x}-1}{\bar{x}+m} \quad (12)$$

3.3. Bayes' estimation

Consider the prior distribution of α to be beta with known hyper parameters a, b>0 having density function

$$f(\alpha/a,b) = \frac{\alpha^{a-1}(1-\alpha)^{b-1}}{B(a,b)}; \quad 0 < \alpha < 1 \quad (13)$$

The posterior distribution from (9) and (13) is obtained as

$$\prod\left(\frac{\alpha}{y}\right) = \frac{\alpha^{a+y-n-1}(1-\alpha)^{mn+n+b-1}}{\int_0^1 \alpha^{a+y-n-1}(1-\alpha)^{mn+n+b-1} d\alpha}$$

The Bayes' estimator of α is given as

$$\begin{aligned} \tilde{\alpha} &= \int_0^1 \alpha \prod\left(\frac{\alpha}{y}\right) d\alpha \\ &= \frac{B(a+y-n+1, mn+n+b)}{B(a+y-n, mn+n+b)} \\ &= \frac{a+y-n}{a+y+mn+b} \end{aligned} \quad (14)$$

where $y = \sum_{i=1}^n x_i$

4. NEGATIVE MOMENTS AND INVERSE ASCENDING FACTORIAL MOMENTS OF SIZE-BIASED NEGATIVE BINOMIAL DISTRIBUTION

Theorem I:

Suppose X is size-biased negative binomially distributed. Then

$$E(X+A)^{-1} = \frac{(1-\alpha)^{m+1}}{(A+1)} {}_2F_1[A+1, m+1; A+2; \alpha] \quad (15)$$

where $A > 0$ and

$$\begin{aligned} {}_2F_1[A+1, m+1; A+2; \alpha] &= 1 + \frac{(A+1)(m+1)}{(A+2)}\alpha \\ &\quad + \frac{(A+1)(A+2)(m+1)(m+2)}{(A+2)(A+3)2!}\alpha^2 + \dots \end{aligned}$$

Proof: $E(X+A)^{-1} = \sum_{x=1}^{\infty} \binom{m+x-1}{x-1} \alpha^{x-1} (1-\alpha)^{m+1} \frac{1}{X+A}$

$$\begin{aligned} &= \frac{(1-\alpha)^{m+1}}{(A+1)} \left\{ 1 + \frac{(A+1)(m+1)}{(A+2)}\alpha \right. \\ &\quad \left. + \frac{(A+1)(m+1)(m+2)}{(A+3)2!}\alpha^2 + \dots \right\} \\ &= \frac{(1-\alpha)^{m+1}}{(A+1)} \left\{ 1 + \frac{(A+1)(m+1)}{(A+2)}\alpha \right. \\ &\quad \left. + \frac{(A+1)(A+2)(m+1)(m+2)}{(A+2)(A+3)2!}\alpha^2 + \dots \right\} \\ &= \frac{(1-\alpha)^{m+1}}{(A+1)} {}_2F_1[A+1, m+1; A+2; \alpha] \end{aligned}$$

Theorem 2:

Suppose X is a size-biased negative binomial random variable, then the inverse ascending factorial moment of X is:

$$\mu'_{[k]} = E\left[\prod_{i=1}^k (X+i)\right]^{-1} = (1-\alpha)^{m+1} [(k+1)!]^{-1} {}_2F_1[2, m+1; k+2; \alpha] \quad (16)$$

k = 1, 2 ...

Proof:

$$\begin{aligned} \mu'_{[k]} &= E\left[\prod_{i=1}^k (X+i)\right]^{-1} = \sum_{x=1}^k \binom{m+x-1}{x-1} \alpha^{x-1} (1-\alpha)^{m+1} \prod_{i=1}^k \left[\frac{1}{(X+i)}\right] \\ &= \sum_{x=1}^k \binom{m+x-1}{x-1} \alpha^{x-1} (1-\alpha)^{m+1} \frac{X!}{(k+X)!} \\ &= \frac{(1-\alpha)^{m+1}}{(k+1)!} \left[1 + \frac{(m+1)2}{(k+2)}\alpha + \frac{(m+1)(m+2)3!}{(k+2)(k+3)2!}\alpha^2 + \dots \right] \\ &= (1-\alpha)^{m+1} [(k+1)!]^{-1} {}_2F_1[2, m+1; k+2; \alpha] \end{aligned}$$

5. RECURRENCE RELATION FOR NEGATIVE MOMENTS AND INVERSE ASCENDING FACTORIAL MOMENTS OF SIZE-BIASED NEGATIVE BINOMIAL DISTRIBUTION

Theorem 3:

Suppose X is size-biased negative binomially distributed. Then

$$E(X+A)^{-1} = \frac{A}{(A-m)\alpha} E(X+A-1)^{-1} - \frac{(1-\alpha)}{\alpha(A-m)} \quad (17)$$

Proof: From (15), we have:

$$E(X+A)^{-1} = \frac{(1-\alpha)^{m+1}}{(A+1)} {}_2F_1[A+1; m+1; A+2; \alpha]$$

Using the identity (see [15], page 71)

$$\begin{aligned} (1-z) {}_2F_1[a, b; c; z] &= {}_2F_1[a-1, b; c; z] \\ &\quad - \left(\frac{c-b}{c}\right) z {}_2F_1[a, b; c+1; z] \end{aligned} \quad (18)$$

When $a_1=A+1, b_1=m+1, c_1=A+1$ and $z=\alpha$, we have

$$(1-\alpha)_2 F_1[A+1, m+1; A+1; \alpha]$$

$$= {}_2 F_1[A, m+1; A+1; \alpha] - \frac{(A-m)\alpha}{(A+1)} {}_2 F_1[A+1, m+1; A+2; \alpha]$$

Also $E(X + A-1)^{-1} = \frac{(1-\alpha)^{m+1}}{A} {}_2 F_1[A, m+1; A+1; \alpha]$

$$(1-\alpha)_2 F_1[A+1, m+1; A+1; \alpha]$$

$$= \frac{A E(X+A-1)^{-1}}{(1-\alpha)^{m+1}} - \frac{(A-m)\alpha}{(A+1)} {}_2 F_1[A+1, m+1; A+2; \alpha]$$

$$(1-\alpha)(1-\alpha)^{-(m+1)}$$

$$= \frac{A E(X+A-1)^{-1}}{(1-\alpha)^{m+1}} - \frac{(A-m)\alpha(A+1) E(X+A)^{-1}}{(A+1)(1-\alpha)^{m+1}}$$

$$(1-\alpha) = A E(X+A-1)^{-1} - (A-m)\alpha E(X+A)^{-1}$$

$$E(X + A)^{-1} = \frac{A}{(A-m)\alpha} E(X + A-1)^{-1} - \frac{(1-\alpha)}{\alpha(A-m)}$$

6. FITTING OF SBNBD MODEL

Two different varieties of Mulberry Ichinose and Kokuso-20 having different leaf spot disease intensity were chosen for the study in the local Kashmir Sericulture division. Three trees of each variety were selected randomly, from each tree three branches were selected at random and then from each branch the spots were recorded from all the leaves. The leaves with no spot were referred as disease free and named as grade zero (0 grade). The leaves having 1 to 5, 6 to 10, 11 to 15, 16 to 20 and more than 20 spots were graded as 1,2,3,4 and 5 grades respectively. In our study the leaves of zero grades were not found. The data is listed in table-1 for each disease. Regarding the choice of values of (a, b) in the Bayes' estimator, there was no information about their values except that they are real and positive numbers. Therefore 25 combinations of values of (a, b) were considered for a, b=1, 2,3,4,5 and those values of a, b were selected for which the Bayes' estimator $\tilde{\alpha}$ has minimum variance. It was found that for a=b=5, the Bayes' estimator has minimum variance. Also, from table 1, it is obvious that the Bayes' estimator provides us the better fit in comparison to maximum likelihood estimator.

Tab.1

| Leaf spot grade | Ichinose | | | Kokuso-20 | | |
|-----------------|--------------------|--------------------|-------------------------|--------------------|--------------------|-------------------------|
| | Observed Frequency | Expected Frequency | | Observed Frequency | Expected Frequency | |
| | | MLE $\hat{\alpha}$ | Bayes' $\tilde{\alpha}$ | | MLE $\hat{\alpha}$ | Bayes' $\tilde{\alpha}$ |
| 1 | 18 | 17.1 | 17.5 | 37 | 36.12 | 36.95 |
| 2 | 15 | 14.3 | 14.9 | 16 | 15.90 | 15.97 |
| 3 | 10 | 9.76 | 9.95 | 15 | 14.86 | 14.92 |
| 4 | 14 | 13.91 | 13.92 | 8 | 7.45 | 7.95 |
| 5 | 13 | 14.93 | 13.73 | 8 | 9.67 | 8.21 |
| Total | 70 | 70 | 70 | 84 | 84 | 84 |
| χ^2 | | 0.34 | 0.054 | | 0.35 | 0.006 |

ACKNOWLEDGEMENT

The author is highly thankful to the editor and the two reviewers for their valuable suggestions.

REFERENCES

- [1] Grab, E.L., Savage, I.R. (1954). Tables of the expected value of 1/X for positive Bernoulli and Poisson variables. *JASA*, 49, 169-177.
- [2] Mendenhall, W., Lehmann, E.H. Jr. (1960). An approximation to the negative moments of the positive binomial useful in life testing. *Technometrics*, 2, 227-41.
- [3] Govindarajulu, Z. (1962). The reciprocal of the decapitated negative binomial variable. *JASA*, 57, 906-913.
- [4] Govindarajulu, Z. (1963). Recurrence relations for the inverse moments of the positive binomial variable. *JASA*, 58, 467-473.
- [5] Tiku, M.L. (1964). A note on the negative moments of a truncated Poisson variate. *JASA*, 59, 1220-4.
- [6] Stancu, D.D. (1968). On the moments of negative order of the positive Bernoulli and Poisson variables. *Studia Universitatis Babeş Bolyai Series, Mathematics and Physics*, 1, 29-31.
- [7] Chao, M.T., Strawderman, W.E. (1972). Negative moments of positive random variables. *JASA*, 67, 429-431.
- [8] Lepage, Y. (1978). Negative factorial moments of positive random variables. *Industrial Mathematics*, 28, 95-100.
- [9] Cressie, N., Davids, A.S., Folks, J.L., Policello, G.E. (1981). The moment generating function and negative integer moments. *American Statistician*, 35 (3), 148-150.
- [10] Cressie, N., Borkent, M. (1986). The moment generating function has its moments. *Journal of Statistical Planning and Inference*, 13, 337-344.
- [11] Jones, M.C. (1986). Inverse moments of negative binomial. *Journal of Statistical Computation and Simulation*, 23, 241-243.
- [12] Jones, M.C. (1987). Inverse factorial moments. *Statistics and Probability Letters*, 6, 37-42.
- [13] Roohi, A. (2002). *Negative Moments of Non-Negative Random Variables*. Unpublished M. Phil Dissertation. Lahore: University of Lahore, 135.
- [14] Ahmad, M., Sheikh, A.K. (1983). Some estimation problem of a class of inverted distributions. In *Bulletin of the International Statistical Institute: Proceedings of the 44th Session*. International Statistical Institute, 1-6.
- [15] Rainville, E.D. (1960). *Special Functions*. New York: Chelsea Pub. Co.