ESTIMATING THE GENERALIZED NEGATIVE BINOMIAL DISTRIBUTION

By

P.N. JANI AND S.M. SHAH

Sardar Patel University, Vallabh Vidyanagar, Gujarat

(Received : September, 1978)

SUMMARY

The different methods for estimating the parameters of the generalized negative binomial distribution have been discussed. The asymptotic variances and covariances of the estimators obtained by all the methods have been derived and the asymptotic efficiencies of the method of moments and the method using the zero-cell frequency, relative to the method of maximum likelihood, have been derived and computed for a wide ranging set of the values of the parameters.

INTRODUCTION

The negative binomial was, perhaps, a distribution which widely used in fitting the biological data. (Anscombe [1], Bliss [2]). At present the large number of mixture and compound distributions obtained by different authors, by compounding the negative binomial distribution in different ways, are available in the literature (Patil [8], Johnson and Kotz [5]). Recently Jain and consul [4] have obtained a such distribution, called the generalized negative binomial distribution (GNBD), compounding the negative binomial distribution with an additional parameter, which takes into account the variations in the mean and the variance.

For a random variable X, the GNBD is defined by the probability function

$$P(X=i)=p_i=\frac{k\,\Gamma(k+i\beta)}{i\,!\,\Gamma(k+i\beta-i+1)}\,\theta^i(1-\theta)^{i\beta-i+k}, \qquad \ldots(1.1)$$

$$i=0,1,2,\ldots,\ k>0,\ 0<\theta<1,\ |\beta\theta|<1 \text{ and}$$

$$p_i=0 \text{ for } i\leqslant n, \text{ if } K+n\beta<0. \qquad \ldots(1.2)$$

For $\beta < 1$, there are only a finite number of non-zero terms in the series $\sum_{i=0}^{\infty} p_i = 1$, because of (1.2). Hence, except for a few lucky choices for $\beta < 1$ and θ, k, the series $\sum_{i=0}^{\infty} p_i$ will not converge to unity (Nelson, [7]. For $\beta = 1$, (1.1) reduces to the classical negative binomial distribution and for $\beta = 0$ it reduces to the binomial distribution. For large k, the GNBD gives a Poisson-type approximation and as $k \to 0$, the zero-truncated GNBD tends to a generalized logarithmic series distribution.

Jain and Consul obtained moment estimators of parameters θ, β and k of the GNBD (1.1) and used these in fitting the distribution. But, moment estimators are not always found efficient. In the present paper we suggest an efficient method, using the zero-cell frequency and the first two sample moments, and study this method along with the method of maximum likelihood and the method of moments. The asymptotic variances and covariances of the estimators obtained by all the three methods and the asymptotic efficiencies of the method of moments and the method using the zero-cell frequency, relative to the method of maximum likelihood, have been deiived. The efficiencies have been computed for a wide ranging set of values of the parameters θ, β and k.

METHOD OF MAXIMUM LIKELIHOOD

Applying the usual technique of the maximum likelihood method we get the following equations for estimating θ, β and k of the GNBD (1.1) as

$$(N\bar{X})^{-1} \sum_{i=2}^{\infty} \sum_{j=1}^{i-1} i.N_i\,(\hat{k}+i\hat{\beta}-j)^{-1} = -\ln(1-\hat{\theta}) \qquad \ldots(2.1)$$

$$N^{-1} \sum_{i=2}^{\infty} \sum_{j=1}^{i-1} N_i\,(\hat{k}+i\hat{\beta}-j)^{-1} + (N\hat{k})^{-1}(N-N_0) = -\ln(1-\hat{\theta}) \qquad \ldots(2.2)$$

$$(\hat{\theta})^{-1} - \hat{k}\,(\bar{X})^{-1} = \hat{\beta} \qquad \ldots(2.3)$$

where Ni=observed frequency in the sample corresponding to $X=i$, $\Sigma\, Ni=N$; the sample size and $\Sigma\, Ni/N=\bar{X}$; the sample mean. These equations can be solved for θ, β and k by using the method of iterations such as the Newton-Raphson method or the method of scoring (Rao, [9], although both methods may fail because of non-convergence.

The Fisher information matrix U of the maximum likelihood estimators (m,l.e's) $\hat{\theta}$, $\hat{\beta}$, $\hat{k}$ can be found to be

$$U=N\,[(u_{ij})] \qquad \ldots(2.4)$$

where the elements $u_{ij}=E\left[-N^{-1}\cdot\dfrac{\partial^2 \ln L}{\partial r_i\,.\,\partial r_j}\right]$,

$i, j=1, 2, 3$ and $(r_1, r_2, r_3)=(\theta, \beta, k)$, are given by

$$u_{11}=k\,(\delta\theta)^{-1}.\ u_{13}$$

$$u_{12}=u_{21}=\theta^2.\ u_{11}$$

$$u_{13}=u_{31}=(1-\theta)^{-1} \qquad \ldots(2.5)$$

$$u_{22}=\sum_{i=1}^{\infty}\sum_{j=1}^{i-1} i^2\,(k+i\beta-j)^{-2}.\ p_i$$

$$u_{23}=u_{32}=\sum_{i=2}^{\infty}\sum_{j=1}^{i-1} i\,(k+i\beta-j)^{-2}.\ p_i$$

$$u_{33}=k^{-2}\,(1-(1-\theta)^k)+\sum_{i=2}^{\infty}\sum_{j=1}^{i-1}(k+i\beta-j)^{-2}\,p_i,$$

and

$$\delta=1-\beta\theta \qquad \ldots(2.6)$$

The asymptotic variance-covariance matrix U^{-1} of the m.l.e's $\hat{\theta}$, $\hat{\beta}$ and $\hat{k}$ can easily be obtained by inverting U given in (2.4).

METHOD OF MOMENTS

Using (2.1) and (2.4) of Gupta [3], we can obtain the moments of the GNBD (1.1). Considering first three moments and solving for

θ, β, k and replacing μ, μ_2, μ_3 by their respective sample estimators $\bar{X}$, $M_2=(NT_2-T_1^2)\,(N(N-1))^{-1}$ and $M_3=N^{-1}\,(T_3-3\bar{X}\,T_2+2N\,\bar{X}^3)$, $T_r=\sum_{i=0}^{\infty} i^r.\,N_i$, we get the moment estimators θ^*, β^* and k^* as

$$\theta^*=1-A/2+(A^2/4-1)^{1/2} \quad ...(3.1)$$

$$\beta^*=(\theta^*)^{-1}\,[1-\{(1-\theta^*)\,\bar{X}M_2^{-1}\}^{1/2}] \quad ...(3.2)$$

$$k^*=\bar{X}\,((\theta^*)^{-1}-\beta^*) \quad ...(3.3)$$

where

$$A=-2+(\bar{X}\,M_3-3M_2^2)(\bar{X}\,M_2^3)^{-1} \quad ...(3.4)$$

Using the differential method (Kendall and Stuart [6], we obtain the asymptotic variance-covariance matrix V of θ^*, β^* and k^* to the order N^{-1} as

$$V=N^{-1}\,[v_{ij}], \quad ...(3.5)$$

where the elements v_{ij}, $i, j=1, 2, 3$ are given by

$$\begin{aligned} v_{11}=&(1-\theta)\,(k\theta^3\delta^3)^{-1}\,[6k\theta\,(1-\theta)\,\delta\,\{4k\theta\,(1-\theta)\,\delta \\ &+48\,(1-\theta)^2-36\delta\,(1-\theta)\,(2-\theta) \\ &+3\delta^2\,(8-8\theta+\theta^2)\,\} \\ &+864\,(1-\theta)^4-1248\,\delta\,(1-\theta)^3\,(2-\theta) \\ &+18\,\delta^2\,(1-\theta)^2\,(140-140\,\theta+27\,\theta^2) \\ &-\delta^4\,(120-240\theta+138\theta^2-18\theta^3+\theta^4)] \end{aligned}$$

$$v_{12}=v_{21}=-\theta^{-1}\,(A_1.v_{11}+A_2)$$

$$v_{22}=\theta^{-2}\,(A_1^2\,.\,v_{11}+2A_1\,.\,A_2+A_3)$$

$$v_{13}=v_{31}=-k(\theta^2.v_{12}+v_{11})\,(\delta\theta)^{-1}+(1-\theta)$$

$$v_{23}=v_{32}=-k\,(\theta^2.v_{22}+v_{12})\,(\delta\theta)^{-1}$$

$$v_{33}=-k\,(\theta^2.v_{23}+v_{13})\,(\delta\theta)^{-1}$$

where

$$A_1=\frac{2\,(1-\theta)-\delta\,(2-\theta)}{2\theta\,(1-\theta)} \quad ...(3.7)$$

$$\begin{aligned} A_2=&(2\delta k\theta^2)^{-1}\,[12A_1\,k\delta\theta^2(1-\theta)^2+48(1-\theta)^3-54\delta(1-\theta)^2 \\ &(2-\theta)+2\delta^2(1-\theta)\,(36-36\theta-5\theta^2)-\delta^3(2-\theta) \\ &(6-6\theta-\theta^2)] \end{aligned} \quad ...(3.8)$$

$$\begin{aligned} A_3=&\delta(4k\theta)^{-1}(1-\theta)^{-1}\,[2k\theta\delta(1-\theta)+10(1-\theta)^2 \\ &-8\delta(1-\theta)\,(2-\theta)+\delta^2(6-6\theta+\theta^2)] \end{aligned} \quad ...(3.9)$$

The joint asymptotic efficiency E_M of the moment estimators (θ^*, β^*, k^*) relative to the m.l.e's $(\hat{\theta}, \hat{\beta}, \hat{k})$ is given by

$$E_M=(|U| \cdot |V|)^{-1} \quad \ldots(3\text{-}10)$$

where $|X|$ is the determinant of a matrix X.

Method Using the Zero-Cell Frequency

The Method of moments is not always efficient. In the case of GNBD (1.1) for certain estimates of the parameters this method is found inefficient. Since the maximum likelihood equations are too-complicated to solve for the m.l.e.'s some other efficient methods should be search out. We, in the following suggest a method for estimating the parameters θ, β, k of the GNBD (1.1) which makes use of the observed zero-cell frequency and the first two sample moments.

From (1.1) we have

$$p_o=P(X=0)=(1-\theta)^k$$

which after taking the natural logarithm yields

$$k=\ln p_o\,(\ln(1-\theta))^{-1} \quad \ldots(4.1)$$

Using (4.1) and the first two moments and replacing p_o, μ, μ_2 by their respective estimators $N^{-1}N_o$, $\bar{X}$ and M_2 we get estimators for θ, β and k as

$$\tilde{\theta}=C(1-\tilde{\theta})^{1/2}\ln(1-\tilde{\theta}) \quad \ldots(4.2)$$

$$\tilde{\beta}=(\tilde{\theta})^{-1}\,[1-\{(1-\tilde{\theta})\bar{X}M_2^{-1}\}^{1/2}] \quad \ldots(4.3)$$

$$\tilde{k}=\bar{X}((\tilde{\theta})^{-1}-\tilde{\beta}) \quad \ldots(4.4)$$

where

$$C=(\bar{X}^3M_2^{-1})^{1/2}\,(\ln(N^{-1}N_o))^{-1} \quad \ldots(4.5)$$

$\tilde{\theta}$ can easily be obtained by solving equation (4.2) iteratively.

Following the method given in Section 3, we derive the asymptotic variance-covariance matrix W of estimators $\tilde{\theta}$, $\tilde{\beta}$, $\tilde{k}$ to the order N^{-1} as

$$W=N^{-1}\,[w_{ij}], \quad \ldots(4.6)$$

where the elements w_{ij}, $i, j=1, 2, 3$ are given by

$$w_{11}=\theta(1-\theta)\phi^2(\delta k2)^{-1}[4\theta\delta(1-\theta)\{(1-\theta)^{-k}-1\}+$$
$$+4k\delta\theta\alpha\{2(1-\theta)+k\delta\theta\}$$
$$+k\alpha^2\{2k\theta\delta(1-\theta)+6(1-\theta)^2$$
$$-4\delta(1-\theta)(2-\theta)+\delta^2(6-6\theta+\theta^2)\}]$$

$$w_{12}=w_{21}=-\theta^{-1}(A_1.w_{11}+A_4)$$

$$w_{22}=\theta^{-2}(A_1^2\cdot w_{11}+2A_1\cdot A_4+A_3) \qquad \ldots(4.7)$$

$$w_{33}=w_{31}=-k(\theta^2.w_{12}+w_{11})(\delta\theta)^{-1}+(1-\theta)$$

$$w_{23}=w_{32}=-k(\theta^2.w_{22}+w_{12})(\delta\theta)^{-1}$$

$$w_{33}=-k(\theta^2.w_{23}+w_{13})(\delta\theta)^{-1}$$

where

$$A_4=\theta\phi(k\delta)^{-1}[k\theta\delta^3+2\alpha(1-\theta)\{kA_3-\delta(1-\theta)A_1\}] \qquad \ldots(4.8)$$

$$\phi=[2\theta+(2-\theta)\alpha]^{-1} \qquad \ldots(4.9)$$

$$\alpha=\ln(1-\theta) \qquad \ldots(4.10)$$

and A_1 and A_3 are as given in (3.14) and (3.16).

The joint asymptotic efficiency Ez of $(\tilde{\theta}, \tilde{\beta}, \tilde{k})$ relative to the $m.1.e's$ $(\hat{\theta}, \hat{\beta}, \hat{k})$ is given by

$$Ez=(\,|U|\,.\,|W|\,)^{-1}. \qquad \ldots(4.11)$$

COMPARISION OF ASYMPTOTIC EFFICIENCIES

For comparision, the asymptotic efficiencies E_M and E_Z have been computed for different values of k, β and θ and tabulated in Table—1. Since the distribution has very long tail for example for $k=3$; $\beta=1.1$. and $\theta=0.7$ the series $\sum_{i=1}^{n} p_i$ has value 0.99999907 when $n=93$, we have computed the efficiencies upto the accuracy $\Sigma\, p_i=0.999999$. If we increase the accuracy, the results may slightly change. However, the table provides a good study of comparision between the efficiencies E_M and E_Z. The table shows that for $\beta>1$, the method of moment is not fairly efficient.

TABLE 1

The values of efficiencies E_M and E_Z

		$k=0.5$			$k=1.0$			$k=2.0$			$k=3.0$		
		0.3	0.5	0.7	0.3	0.5	0.7	0.3	0.5	0.7	0.3	0.5	0.7
0.9	E_M	0.874	0.866	0.998	0.866	0.837	0.968	0.917	0.847	0.894	0.921	0.874	0.892
	E_Z	0.991	0.937	0.908	0.980	0.929	0.737	0.973	0.837	0.602	0.894	0.709	0.404
1.0	E_M	0.679	0.488	0.315	0.715	0.540	0.376	0.795	0.628	0.478	0.827	0.689	0.556
	E_Z	0.994	0.984	0.965	0.996	0.981	0.940	0.999	0.914	0.757	0.935	0.783	0.514
1.1	E_M	0.571	0.270	0.071	0.599	0.327	0.105	0.680	0.426	0.175	0.771	0.501	0.244
	E_Z	0.998	0.929	0.778	0.973	0.966	0.840	0.970	0.941	0.764	0.963	0.829	0.556

REFERENCES

[1] Anscombe, F.J. (1949) : The statistical analysis of insect counts based on the negative binomial distribution, *Biometrics* 5, 165-173.

[2] Bliss, C.I. (1953) : Fitting the negative binomial distribution to biological data. *Biometrics*, **9**, 176-196.

[3] Gupta, R.C. (1974) : Modified power series distribution and some of its applications. *Sankhya*, B, **36**, 288-298.

[4] Jain, J.C. and Consul, P.C. (1971) : A generalized negative binomial distribution. *SIAM J. of Appl. Math*. **21**, 501-513.

[5] Johnson, N.L. and Kotz's S. (1969) : 'Distributions in Statistics : Discrete Distributions', *Houghton Mifflin Co., Boston.*

[6] Kendall, M.G. and Stuart, A. (1969) : 'The Advanced Theory of Statistics', Vol. **1**, Charles Griffin and Co., London.

[7] Nelson, D.L. (1975) : Some remarks on generalization of the negative binomial and Poisson distributions. *Technometrics*, **17**, 135-136.

[8] Patil, G.P. (1965) : 'Classical and Contagious Discrete Distribution', Statistical Publishing Society, *Calcutta.*

[9] Rao, C.R. (1974) : 'Linear Statistical Inference and its Applications', Wiley Eastern (Pvt.) Ltd., New Delhi.