CHAPTER 6

ESTIMATION OF THE PARAMETERS OF THE GNBD

6.1 INTRODUCTION

Negative binomial is, perhaps, a distribution which is widely used in fitting the biological data (Anscombe [14], Bliss [6] etc.). 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 [49], Johnson and Kotz [36] etc.). Recently Jain and Consul [26] have obtained a such distribution by compounding the negative binomial distribution with an additional parameter, which takes into account the variations in the mean and the variance. This distribution is a GNBD defined in (1.5). For $\beta < 1$, there are only finite number of non-zero terms in the series $\sum p_x = 1$ and hence, except for a few lucky choices for $\beta < 1$ and $\theta, k$, the series $\sum p_x$ will not converge to
unity (Helson | 45 | ). For β = 1, (1.5) reduces to the negative binomial and for β = 0 it reduces to the binomial. For large value of k, the GMBD gives a Poisson-type approximation and as k → 0, the decapitated GMBD tends to a GLSD.

Jain and Consul [26] obtained moment estimators of the parameters \( \theta, \beta \) and \( k \) and used these in fitting the distribution. But, moment estimators may not always be efficient. Moreover, estimators without some knowledge of their asymptotic distribution are, in general, of little use. In the present chapter we discuss three methods viz. (i) The M.L. method (ii) The method of moments (Method A) and (iii) The method using the zero-cell frequency and the first two sample moments.

We derive the asymptotic variances and covariances of the estimators obtained by all the three methods and the asymptotic efficiencies of Methods A and B relative to the M.L. method. We present a table of efficiencies which have been computed for different values of the parameters. An illustrative example of Jain and Consul [26] has been given. A few properties which have been not described by Jain and Consul [26] are discussed here.
6.2 PROPERTIES OF THE GIBB

In the following we shall discuss a few properties of the GIBB which are not described by Jain and Consul [26].

(i) While defining the GIBB (1.5), Jain and Consul [26] have stated a condition that \( p_x = 0 \) whenever \((k + x\beta) < 0\). Now, if \((k + x\beta) < 0\) then \((k + x\beta - x + 1) < 0\) but \((k + x\beta - x + 1)\) can be negative even if \((k + x\beta) > 0\). e.g. for \(k = 1\), \(\beta = 0.5\) and \(x = 6\), \(k + x\beta = 4 > 0\) but \((k + x\beta - x + 1) = -1 < 0\). So, it would be more proper that the condition is stated as \( p_x = 0 \), whenever \((k + x\beta - x + 1) < 0\).

(ii) Nelson [45] remarked that for \(\beta < 0\), there will be finite number of terms in the series \(\sum p_x\), because of condition stated in (i). But, even for \(\beta > 0\) there are finite number of terms in the series \(\sum p_x\). Infact, for \(\beta < 1\), there will be finite number of terms. Table 6.2.1 shows the values of \(\sum_{x=0}^{n} p_x\) and \(n\), for some values of \(\beta\), \(\theta\) and \(k\).

(iii) For \(\beta < 1\), except for a few lucky choices, \(\sum_{x=0}^{n} p_x\) will not converge to unity. \(\sum p_x\) will be greater
than 1. Table 6.2.1 confirms this. Hence, for 
\[ \beta < 1, \] only those values are admissible for which 
\[ \sum p_x = 1. \]

(iv) For small values of \( \beta \) (say \( < 0.5 \)) and large values of 0 (say \( > 0.5 \)) the generalized negative binomial probabilities are the binomial type probabilities.

(v) For \( \beta > 1, \) as \( \beta \) increases the distribution will have longer and longer tails. e.g. for \( \beta = 1.5, \) 
\[ \theta = 0.5 \text{ and } k = 3, \quad \sum_{x=0}^{\infty} p_x = 0.999999 \text{ when } n = 123. \]

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
\( n \) & \( \sum p_x \) & \( n \) & \( \sum p_x \) & \( n \) & \( \sum p_x \) \\
\hline
\( k = 0.5 \) & & \( k = 1.0 \) & & & \\
\hline
\( \beta = 0.1 \) & \( \beta = 0.5 \) & \( \beta = 0.1 \) & \( \beta = 0.5 \) & \\
\( n = 2 \) & \( n = 3 \) & \( n = 3 \) & \( n = 4 \) & \\
\hline
0.1 & 1.00084 & 1.00000 & 1.00004 & 1.00000 \\
0.5 & 1.03698 & 1.00130 & 1.01004 & 1.00068 \\
0.9 & 1.44638 & 1.08640 & 1.32597 & 1.07777 \\
\hline
\end{tabular}
\end{table}
6.3 THE M.L. METHOD

Let us write the probability function of the GNBDD (1.5) as

\[ (6.3.1) \quad P(X=x) = p_x(k, \beta, \theta) \]

\[ = k(k+x\beta-1)(k+x\beta-2) \cdots (k+x\beta-x+1) \]

\[ \cdot (1-\theta)^k (\theta(1-\theta)^{\beta-1})^x / x! , \]

\[ x \in I, \quad 0 < \theta < 1, \quad |\beta| < 1, \quad k > 0 \quad \text{and} \quad p_x = 0, \]

whenever \( k + x\beta - x + 1 < 0. \)

Now, consider a random sample of size \( N \) from the population (6.3.1) and let \( N_x \) be the observed frequency in the sample corresponding to \( X = x. \) Then, the likelihood function \( L \) is given by

\[ L = \prod_{x=0}^{n} p_x^{N_x} \]

Taking the natural logarithm of \( L \) and differentiating w.r.t. \( \theta, \beta \) and \( k \) and writing \( \sum_{x=0}^{n} xN_x / N = M \) we get

\[ (6.3.2) \quad \frac{\partial \ln L}{\partial \theta} = N (1-\theta)^{-1} \left[ \frac{M}{1} (\theta^{-1} - \beta) - k \right] \]
Equating (6.3.2) to (6.3.4) to zero we get the M.L. equations for estimating \( \theta \), \( \beta \) and \( k \) as

\[
(6.3.5) \quad (N.M')^{-1} \sum_{x=2}^{n} \sum_{j=1}^{x-1} \frac{x.N}{x} = -\ln(1 - \hat{\theta})
\]

\[
(6.3.6) \quad N^{-1} \sum_{x=2}^{n} \sum_{j=1}^{x-1} \frac{N}{x} + \frac{N_{-N}}{N.k} = -\ln(1 - \hat{\theta})
\]

\[
(6.3.7) \quad \frac{1}{\hat{\theta}} - \frac{\hat{k}}{M'} = \hat{\beta}
\]

These equations can be solved for \( \hat{\theta} \), \( \hat{\beta} \) and \( \hat{k} \) by using the method of iterations such as the Newton-Raphson method or the method of scoring (Rao [55], p. 370), although both methods may fail because of non-convergence.

The Fisher information matrix \( U \) of the M.L. estimators \( \hat{\theta} \), \( \hat{\beta} \), \( \hat{k} \) can be found to be
(6.3.8) \( u = N \sum u_{ij} \),

where the elements

\[
u_{ij} = E \left[ -N^{-1} \frac{\partial^2 \ln L}{\partial r_i \partial r_j} \right],
\]

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

(6.3.9) \( u_{11} = k / (6\theta(1 - \theta)) \)

(6.3.10) \( u_{12} = u_{21} = \theta^2 u_{11} \)

(6.3.11) \( u_{13} = u_{31} = 1 / (1 - \theta) \)

(6.3.12) \( u_{22} = \sum_{x=2}^\infty \sum_{j=1}^x (k + x\beta - j)^{-2} p_x \)

(6.3.13) \( u_{23} = u_{32} = \sum_{x=2}^\infty \sum_{j=1}^x (k + x\beta - j)^{-2} p_x \)

(6.3.14) \( u_{33} = k^2(1 - (1 - \theta)k) + \sum_{x=2}^\infty \sum_{j=1}^x (k + x\beta - j)^{-2} p_x \)
\[(6.3.15) \quad \delta = 1 - \beta 0\]

and \(p_x = p_x(k, \beta, \theta)\) as given in (1.5). The asymptotic variance-covariance matrix \(U^{-1}\) of the M.L. estimators \(\hat{\delta}, \hat{\beta}, \hat{\theta}\) can easily be obtained by inverting \(U\).

6.4 METHOD A

Using (1.8), (1.9) and (1.11), the first six central moments are derived as:

\[(6.4.1) \quad \mu = k \theta \delta^{-1}\]

\[(6.4.2) \quad \mu_2 = k \theta (1 - \theta) \delta^{-3}\]

\[(6.4.3) \quad \mu_3 = k \theta (1 - \theta) \delta^{-5} \left[3\theta (1 - \theta) - 5(2 - \theta)\right]\]

\[(6.4.4) \quad \mu_4 = k \theta (1 - \theta) \delta^{-7} \left[3k\theta (1 - \theta) \delta + 15(1 - \theta)^2 - 10\delta (1 - \theta)(2 - \theta) + \delta^2 (6 - 6\theta + \theta^2)\right]\]
(6.4.5) \[ \mu_5 = k_0 (1-\theta) \delta^{-9} \left[ 10k_0 \delta (1-\theta) (3(1-\theta) - \delta(2-\theta)) \right. \\
+ 105(1-\theta)^3 - 105\delta (1-\theta)^2(2-\theta) \\
+ 5\delta^2 (1-\theta)(26-26\theta+50^2) - \delta^3 (2-\theta)(12-12\theta+6^2) \right] \\
(6.4.6) \mu_6 = k_0 (1-\theta) \delta^{-11} \left[ 5k_0 \delta (1-\theta) (3k_0 \delta (1-\theta) + 63(1-\theta)^2 \\
- 42\delta(1-\theta)(2-\theta) + \delta^2 (26-26\theta+50^2) \right) \\
+ 345(1-\theta)^4 - 1260 \delta (1-\theta)^3(2-\theta) \\
+ 708^2 (1-\theta)^2 (34 - 34\theta + 70^2) \\
- 146^3 (1-\theta)(2-\theta)(33 - 33\theta + 46^2) \\
+ \delta^4 (129 - 240\theta + 1506^2 - 306^3 + \theta^4) \right] \\

Solving (6.4.1), (6.4.2) and (6.4.3) for \theta, \beta, k and replacing \mu, \mu_2, \mu_3 by their respective estimators \hat{M}_1, \hat{M}_2 = (N.T_2 - T_1^2) / (N(N-1)) and \\
\hat{M}_3 = (T_3 - 3M_1^2.T_2 + 2N.M_1^3) / N, where T_r = \sum x^r.N_x, \\
r = 1, 2, 3, we get the moment estimators \hat{\theta}*, \hat{\beta}*, \\
and \hat{k} as
(6.4.7) \( \theta^* = 1 - A_0 / 2 + (A_0^2 / 4 - 1)^{1/2} \)

(6.4.8) \( \beta^* = [1 - (1 - \theta^*) M_1 / M_2]^{1/2} / \theta^* \)

(6.4.9) \( k^* = M_1^*(1 / \theta^* - \beta^*) \)

where

(6.4.10) \( A_0 = -2 + (M_1^* M_3 - 3M_2^2)^2 / (M_1^* M_2) \)

Using the differential formulae (Kendall and Stuart [42], § 10.6), used in the previous chapters for the GLSD and the GPD, we obtain the elements \( v_{ij} \) of the asymptotic variance-covariance matrix

(6.4.11) \( V = N^{-1} \sum v_{ij} I \), \( i, j = 1, 2, 3 \)

of \( \theta^*, \beta^*, k^* \), to the order \( N^{-1} \), as
\[(6.4.12) \quad v_{11} = N.V(0^*) \]

\[= (1-\theta)(k^0,\delta^0)^{-1} \left[ 6k^0(1-\theta)\delta \left( 4k^0(1-\theta)\delta 
+ 48(1-\theta)^2 - 36\delta(1-\theta)(2-\theta) + 36^2(3-30+\theta^2) \right) 
+ 864(1-\theta)^4 - 1248\delta(1-\theta)^3(2-\theta) 
+ 18\delta^2(1-\theta)^2(140 - 140\theta + 27\theta^2) 
- 356^3(1-\theta)(2-\theta)(14 - 14\theta + \theta^2) 
+ \delta^4(120 - 240\theta + 138\theta^2 - 18\theta^3 + \theta^4) \right] \]

\[(6.4.13) \quad v_{12} = v_{21} = N.Cov(\theta^*,\beta^*) = - \left( A_1 \cdot v_{11} + A_2 \right) / \theta \]

\[(6.4.14) \quad v_{22} = N.V(\beta^*) = \left( A_1^2 \cdot v_{11} + 2A_1 \cdot A_2 + A_3 \right) / \theta^2 \]

\[(6.4.15) \quad v_{13} = v_{31} = N.Cov(\theta^*,\kappa^*) = - k(\theta^2 \cdot v_{12} + v_{11})/(\delta \theta) 
+ (1 - \theta) \]

\[(6.4.16) \quad v_{23} = v_{32} = N.Cov(\beta^*,\kappa^*) = - k(\theta^2 \cdot v_{22} + v_{12})/(\delta \theta) \]

\[(6.4.17) \quad v_{33} = N.V(\kappa^*) = - k(\theta^2 \cdot v_{23} + v_{13})/(\delta \theta) \]
where

\[ (6.4.18) \quad A_1 = \frac{2(1 - \theta) - \delta(2 - \theta)}{2\theta(1 - \theta)} \]

\[ (6.4.19) \quad A_2 = (25k^3)^{-1} \left[ 12 A_1 k^3(1 - \theta)^2 + 45(1 - \theta)^3 \right. \] 
\[ - 546(1 - \theta)^2(2 - \theta) + 26^2(1 - \theta)(36 - 36\theta + 36^2) \]
\[ - 3^2(2 - \theta)(6 - 36\theta^2) \right] \]

\[ (6.4.20) \quad A_3 = \delta(4k^2)^{-1}(1 - \theta)^{-1} \left[ 2k^3(1 - \theta) + 10(1 - \theta)^2 \right. \]
\[ - 36(1 - \theta)(2 - \theta) + \delta^2(6 - 36\theta^2) \right] \] .

The asymptotic efficiency \( E_A \) of the method A, relative to the M.L. method, is the joint asymptotic efficiency of the moment estimators \((\theta^*, \beta^*, k^*)\) relative to the M.L. estimators \((\hat{\theta}, \hat{\beta}, \hat{k})\) and is given by

\[ (6.4.21) \quad E_A = ( | U | . | V | )^{-1}, \]

where \(| X |\) is the determinant of a matrix \(X\).

6.5 METHOD B

Now we discuss a method which uses the zero-cell frequency and the first two sample moments.
From (1.5) we have

$$p_0 = P(X = 0) = (1 - \theta)^k$$

which after taking the natural logarithm yields

(6.5.1) $$k = \ln p_0 - (\ln(1-\theta))^{-1}.$$  

Using (6.5.1), (6.4.1) and (6.4.2) and replacing $p_0$, $n$ and $\mu_2$ by their respective estimators $F_0 = N_0 / N$, $M_1$ and $M_2$ we get estimators $\bar{\theta}$, $\bar{\beta}$, $\bar{k}$ as

(6.5.2) $$\bar{\theta} = 1 - (\bar{\alpha} \bar{\theta} / \bar{G})^2$$

(6.5.3) $$\bar{\beta} = \left[1 - (1-\bar{\theta}) M_1 / M_2 \right]^{1/2} / \bar{\theta}$$

(6.5.4) $$\bar{k} = M_1^{1/2} (1 / \bar{\theta} - \bar{\beta})$$

where

(6.5.5) $$\bar{\alpha} = \left[1 - \ln(1-\theta) \right]^{-1}$$

and

(6.5.6) $$\bar{G} = (M_1^{3/2} / M_2)^{1/2} / \ln(p_0).$$
can easily be obtained by solving (6.5.2) iteratively.
The table of (u₉) given in the appendix may be found useful for this purpose.

Using the differential formulae (Kendall and Stuart § 10.6), we derive the elements wᵢⱼ
(i, j = 1, 2, 3) of the asymptotic variance-covariance matrix

\[(6.5.7) \quad w = N^{-1} \sum wᵢⱼ, \quad i, j = 1, 2, 3.\]

of Ý, Þ, Ý, to the order \(N^{-1}\), as

\[(6.5.8) \quad w₁₁ = N \cdot V(Ý) = \theta(1-\theta) A₂⁻¹ \sum \delta(1-\theta)\]

\[\cdot \left( (1-\theta)^{-k} - 1 \right) - 4k\theta \alpha⁻¹(2(1-\theta)\]

\[+ k\delta(2k\delta(1-\theta) + 6(1-\theta)^2\]

\[-4\delta(1-\theta)(2-\delta) + \delta^2(6-6\theta+\theta^2)\right]\]

\[(6.5.9) \quad w₁₂ = w₂₁ = N \cdot Cov(Ý, Þ) = -\left( A₁·w₁₁ + A₄ \right) / \theta\]

\[(6.5.10) \quad w₂₂ = N \cdot V(Þ) = (A₁²·w₁₁ + 2A₁·A₄ + A₃)\]
\[(6.5.11) \quad w_{13} = w_{31} = N.\text{Cov}(\tilde{\theta}, \tilde{k}) = -k(\theta^2 w_{12} + w_{11}) / (\delta \theta) + (1-\theta)\]

\[(6.5.12) \quad w_{23} = w_{32} = N.\text{Cov}(\tilde{\beta}, \tilde{k}) = -k(\theta^2 w_{22} + w_{12}) / (\delta \theta)\]

\[(6.5.13) \quad w_{33} = N.V(\tilde{k}) = -k(\theta^2 w_{23} + w_{13}) / (\delta \theta)\]

where

\[(6.5.14) \quad A_4 = A_5 \theta (k \delta)^{-1} \left[ k \delta \delta^3 + 2\alpha (1-\theta) (k A_3 - 3(1-\theta)A_1) \right]\]

\[(6.5.15) \quad A_5 = (2\theta - (2-\theta) / \alpha)^{-1}\]

and \(A_1, A_3, \alpha\) are as given in (6.4.18), (6.4.20) and (6.5.5) respectively.

The asymptotic efficiency \(E_B\) of Method B relative to the M.L. method is the joint asymptotic efficiency of the estimators \((\tilde{\theta}, \tilde{\beta}, \tilde{k})\) relative to the M.L. estimators \((\hat{\theta}, \hat{\beta}, \hat{k})\) and is given by

\[(6.5.16) \quad E_B = (|U| \cdot |W|)^{-1} \]
Remark 6.5.1. As discussed in the case of the GLSD, expanding \( \alpha = (-\ln(1 - \theta))^{-1} \), occurring in (6.5.2), in power series expansion we can get the approximate estimator of \( \theta \) in explicit form. But, since these estimators are found less efficient, they are not discussed here.

6.6 COMPARISON OF ASYMPTOTIC EFFICIENCIES

Using (6.4.21) and (6.5.16) the asymptotic efficiencies \( E_A \) and \( E_B \) have been computed for different sets of values of \( k, \theta \) and \( \beta \), keeping the restrictions discussed in Section 6.2 in mind, and tabulated in Table 6.6.1. The table shows that for \( \beta > 1 \), the method of moments is not fairly efficient.
The relative asymptotic efficiencies $E_A$ and $E_B$ of Methods A and B

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$k = 0.5$</th>
<th>$k = 1.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta$: 0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>$E_A$</td>
<td>0.874</td>
<td>0.866</td>
</tr>
<tr>
<td>$E_B$</td>
<td>0.991</td>
<td>0.937</td>
</tr>
<tr>
<td>$E_A$</td>
<td>0.679</td>
<td>0.483</td>
</tr>
<tr>
<td>$E_B$</td>
<td>0.994</td>
<td>0.984</td>
</tr>
<tr>
<td>$E_A$</td>
<td>0.371</td>
<td>0.270</td>
</tr>
<tr>
<td>$E_B$</td>
<td>0.998</td>
<td>0.929</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$k = 2.0$</th>
<th>$k = 3.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\theta$: 0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>$E_A$</td>
<td>0.917</td>
<td>0.847</td>
</tr>
<tr>
<td>$E_B$</td>
<td>0.973</td>
<td>0.837</td>
</tr>
<tr>
<td>$E_A$</td>
<td>0.795</td>
<td>0.628</td>
</tr>
<tr>
<td>$E_B$</td>
<td>0.999</td>
<td>0.914</td>
</tr>
<tr>
<td>$E_A$</td>
<td>0.630</td>
<td>0.426</td>
</tr>
<tr>
<td>$E_B$</td>
<td>0.970</td>
<td>0.941</td>
</tr>
</tbody>
</table>
6.7 EXAMPLE

In Table 6.7.1 we consider a numerical data from the set of four examples considered by Jain and Consul [26] who fitted the GMBD (1.5) by the method of moments. We fit the GMBD by Method B. Since the convergence is not achieved by the M.L. equations, the fitting by the M.L. method is not shown here. The Table shows that, though not much difference in \( \chi^2 \) values, the expected frequencies obtained by Method B are much closer than obtained by Method A. The residual sum of squares (RSS) is sufficiently lower in the case of Method B.
Counts of the number of European red mites on apple leaves: data of P. Garman

<table>
<thead>
<tr>
<th>No. of mites per leaf ( x )</th>
<th>Leaves observed</th>
<th>Expected frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Method A</td>
</tr>
<tr>
<td>0</td>
<td>70</td>
<td>71.43</td>
</tr>
<tr>
<td>1</td>
<td>38</td>
<td>33.98</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>19.80</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>11.59</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>6.57</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>3.55</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>1.30</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.64</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Total \( 150 \) \( 150.00 \) \( 150.00 \)

Mean \( 1.14667 \)

Variance \( 2.27365 \)

\( \chi^2 \) \( 35.14 \) \( 36.14 \)

Estimates \( \theta : \)

\( \beta : \)

\( k : \)