EXISTENCE AND UNIQUENESS OF THE MAXIMUM
LIKELIHOOD ESTIMATOR FOR THE TWO-PARAMETER
NEGATIVE BINOMIAL DISTRIBUTION

by

J. Aragón, D. Eberly and S. Eberly

Technical Report No. 282
August 1991

Prepared Under the Auspices
of
National Science Foundation Grant
DMS 90-02502
Ingram Olkin, Project Director

Department of Statistics
Stanford University
Stanford, California
EXISTENCE AND UNIQUENESS OF THE MAXIMUM LIKELIHOOD ESTIMATOR FOR THE TWO-PARAMETER NEGATIVE BINOMIAL DISTRIBUTION

by

J. Aragón, D. Eberly and S. Eberly
Stanford University and University of Texas at San Antonio

Technical Report No. 282

Prepared Under the Auspices of
National Science Foundation Grant
DMS 90-02502
Ingram Olkin, Project Director

Department of Statistics
Stanford University
Stanford, California
Existence and Uniqueness of the Maximum Likelihood Estimator
for the Two-Parameter Negative Binomial Distribution

J. Aragón, D. Eberly, and S. Eberly

Stanford University and University of Texas at San Antonio

SUMMARY

It is known that there exists at least one solution to the maximum likelihood equations for the two-parameter negative binomial distribution if the second sample moment $s^2$ is greater than the sample mean $\bar{x}$. Here we show that such condition is necessary and sufficient and that the solution must be unique if it exists.

1 Introduction

The negative binomial distribution (NBD) has been applied widely in Biology, Psychology, Communications, Insurance, Economics, Medicine, Military, etc. The following parametrization is used here:

$$f(x) = \binom{x+k-1}{k} p^k (1-p)^x, \quad x = 0, 1, 2, \ldots$$

We treat $k$ as a continuous parameter with $k \in (0, \infty)$. We refer to the distribution as NB$(k, p)$.

As a result of the frequent application of the NBD an increasing number of papers on estimation have appeared in the literature (Fisher, 1941 and 1953; Haldane, 1941; Wise, 1946; Anscombe 1949 and 1950; Bliss and Fisher, 1953; Bliss and Owen, 1958; Shah, 1961; Katti and Gurland, 1962; O’Carroll, 1962. Shenton and Wallington, 1962; Shenton, 1963; Martin and Katti, 1965; Shenton and Myers, 1965; Johnson and Kotz, 1970).

*Key words:* Maximum likelihood estimator; Negative binomial distribution; Newton’s method
Anscombe (1950) conjectured that the MLE exists and is unique when the second sample moment $s^2 = \sum_{i=1}^{n} x_i^2/n - \bar{x}^2$ is greater than the sample mean $\bar{x} = \sum_{i=1}^{n} x_i/n$ and no MLE exists when $s^2 \leq \bar{x}$. Proofs for existence of the MLE when $s^2 > \bar{x}$ were given by Johnson and Kotz (1969), although the book contains some misprints, and by Willson et al. (1986). Among others, several issues have been raised and addressed: (i) What to do when $s^2 < \bar{x}$? (ii) Treatment of negative estimates of $k$ (iii) Comparison of different estimation methods for small samples. However, the questions of uniqueness when $s^2 > \bar{x}$ and of existence when $s^2 \leq \bar{x}$ have gone unanswered. We answer these remaining questions by proving the following:

**Theorem.** Let $x_i, i = 1, \ldots, n$ be a random sample from NB($k, p$). The MLE of ($k, p$) exists if and only if $s^2 > \bar{x}$. Moreover, if the MLE exists, then it must be unique.

It appears that the difficulty in solving the likelihood equations resides on the fact that the loglikelihood function is very flat with respect to the parameter $k$ (Willson, et. al., 1986). This result is consistent with the property of the NBD which states that the variance is larger than the mean. It suggests that the NBD should not be used for large data sets with $s^2 \leq \bar{x}$.

### 2 The Results

Let $M = \max x_i$, so $0 \leq x_i \leq M$. If $f_j$ is the proportion of the sample values equal to $j$, then $\bar{x} = \sum_{j=1}^{M} jf_j$ and $s^2 = \sum_{j=1}^{M} j^2f_j - (\sum_{j=1}^{M} jf_j)^2$. We wish to compute the maximum likelihood estimators $(\hat{k}, \hat{p})$ for the
sample. The MLE for \( p \) is given by \( \hat{p} = \frac{\hat{k}}{\hat{\bar{x}} + \hat{k}} \) and \( \hat{k} \) is a solution to

\[
g(k) = \sum_{j=1}^{M} \frac{F_j}{k+j-1} - \log\left(1 + \frac{\bar{x}}{k}\right) = 0
\]

(1)

where \( F_j = \sum_{i=j}^{M} f_i \) is the proportion of the sample values greater than or equal to \( j \).

The approach we took to finding the MLE was motivated by Eberly (1991). In this thesis, data sets were generated from a length-biased truncated negative binomial distribution \( 1 + \text{NB}(k + 1, p) \) where \( k > 0 \). To construct an MLE, we used equation (1) with \( k \) replaced by \( k + 1 \). In applying Newton's method to (1), we had problems with the zero at infinity for \( g(k) \). With an inappropriate initial guess, the iterates tended towards infinity. To avoid this problem, we defined \( z = 1/k \) and \( G(z) = g(k) \), so

\[
G(z) = \sum_{j=1}^{M} \frac{z F_j}{(j-1)z + 1} - \log(1 + \bar{x} z), \quad z \in (0, \infty).
\]

(2)

Lam et al. (1988) and Clark and Perry (1983) have used this reparametrization. For \( k \geq 1 \), we need only consider (1) for \( z \in (0, 1] \). One has much more control on the behavior of Newton's method on this finite interval. The function \( G \) has the properties \( G(0) = G'(0) = 0 \) and \( \lim_{z \to \infty} G(z) = F_1 > 0 \). We observed that the graphs of \( G \) were of two types:

![Figure 1. Graphs for \( G(z) \)](image)

The convexity/concavity at \( z = 0 \) determines the shape of the graph. Computing \( G''(0) \) and using the
definition for \( F_j \), we have

\[
G''(0) = \bar{x}^2 - 2 \sum_{j=1}^{M} (j-1)F_j = \bar{x} - \bar{s}^2.
\]

If \( G''(0) < 0 \), then eventually the graph of \( G \) must intersect the \( z \)-axis, thereby providing a root \( \bar{z} \) to \( G(z) = 0 \) and a MLE \( \hat{k} = 1/\bar{z} \). This reproduces the results in Johnson and Kotz (1969) and Willson et al. (1986). We now give a more detailed analysis to show uniqueness of the zero when \( G''(0) < 0 \) and the nonexistence of zeros when \( G''(0) \geq 0 \).

**Case 1:** \( x_i \in \{0, 1\} \).

Let \( F_1 > 0 \), \( F_j = 0 \) for \( j \geq 2 \). Then \( G(z) = F_1z - \log(1 + F_1z) \), \( G'(z) = F_1 - F_1/(1 + F_1z) \), and \( G''(z) = [F_1/(1 + F_1z)]^2 \). Since \( G''(z) > 0 \) and \( G(0) = G'(0) = 0 \), the graph of \( G \) never intersects the \( z \)-axis.

Thus, there is no root \( \bar{z} \). Note that \( x_i \in \{0, 1\} \) implies \( \bar{x} = \bar{x}^2 \leq \bar{x} \).

**Case 2:** \( x_i \in \{0, 1, 2\} \).

The essential geometric ideas for the general case are illustrated by this case. Let \( F_j = 0 \) for \( j \geq 3 \) and \( 1 \geq F_1 \geq F_2 > 0 \). The equation \( G(z) = 0 \) yields \( F_1z + F_2z/z + 1 = \log(1 + \bar{z}z) \). Rename \( F_1 = u \) and use \( \bar{z} = F_1 + F_2 \in (0, 2] \) to solve for \( F_1 \):

\[
u = u(z, \bar{z}) = \left(\frac{z + 1}{z^2}\right) \log(1 + \bar{z}z) - \bar{z}z, \quad z > 0.
\]

(3)

The parameters \( (u, \bar{z}) \) lie in the convex polygonal solid \( D = \{ (u, \bar{z}) : 1 \geq u \geq \bar{z} - u \geq 0 \} \). Using L'Hopital's rule, \( u(0, \bar{z}) = \bar{z} - \bar{z}^2/2 \), which in terms of \( F_1, F_2, \) and \( \bar{z} \) is equivalent to \( G''(0) = 0 \). Define the subregion \( E \subset D \) by \( E = D \cap \{ (w, \bar{z}) : w < u(0, \bar{z}) \} \). Figure 2 illustrates the sets \( D \) and \( E \), and the graphs of \( u(0, \bar{z}) \) and \( u(1, \bar{z}) \) in \( D \).
Figure 2. Region of existence

We can view (3) in the following way. For each value of $z > 0$ the graph of (3) in $D$ represents all pairs $(u, \tilde{z})$ for which $G$ has a zero at $z$. We prove that $E$ is a disjoint union of these graphs. A consequence is that $G$ has a zero if and only if $F_1 < u(0, \tilde{z})$, or equivalently, if and only if $s^2 > \tilde{z}$. Moreover, since the graphs are disjoint in $D$, a given pair $(u, \tilde{z})$ has exactly one graph containing it and the corresponding zero $z$ for $G$ is unique. Figure 2 shows some of the graphs that fill $E$.

We now prove that $E$ is a disjoint union of graphs. Note that $u(z, 0) = 0$ for all $z > 0$, so it is sufficient to show that the slopes $u_{zz}(z, \tilde{z})$ are strictly decreasing in $z$, or $u_{zz}(z, \tilde{z}) < 0$. Differentiate $u$ to obtain $u_{zz} = -(1 - \tilde{z})\tilde{z}/(1 + \tilde{z})^2 < 0$ for $\tilde{z} \in (0, 1)$. Integrate over $(z_1, z_2)$ with $z_1 < z_2$ to obtain $u_{zz}(z_2, \tilde{z}) - u_{zz}(z_1, \tilde{z}) < 0$ where the strict inequality holds since the integral of a continuous negative function over a set of positive measure is negative. Integrate over $(0, \tilde{z})$ and use $u(z, 0) = 0$ to obtain $u(z_2, \tilde{z}) < u(z_1, \tilde{z})$. Therefore, the graphs are ordered as claimed. As $z \to \infty$, the graphs of $u(z, \tilde{z})$ in $D$ approach the graph of $u(\infty, \tilde{z}) \equiv 0$ in $D$, which is the single point $(0, 0)$, and so $E$ is a disjoint union of the graphs.

Case 3: $z_i \in \{0, \ldots, M\}$, $M \geq 3$.  

5
The idea of constructing the region of existence as a union of disjoint graphs applies here as it did in case 2.

Let $F_j = 0$ for $j \geq M + 1$ and $1 \geq F_1 \geq \cdots \geq F_M > 0$. The equation $G(z) = 0$ yields

$$\sum_{j=1}^{M} \frac{zF_j}{(j-1)z + 1} = \log(1 + \bar{z}z).$$

Rename $F_1 = u$ and $\bar{v} = (F_2, \ldots, F_{M-1})$, and use $\bar{x} = \sum_{j=1}^{M} F_j$ to solve for $F_1$:

$$u = u(z, \bar{x}, \bar{v}) = \frac{[(M-1)z + 1] \log(1 + \bar{z}z) - \bar{z}z}{(M-1)z^2} - \frac{1}{M-1} \sum_{j=2}^{M-1} \frac{(M-j)F_j}{(j-1)z + 1}, \quad z > 0. \quad (4)$$

The parameters $(u, \bar{x}, \bar{v})$ lie in the $M$-dimensional convex polygonal solid:

$$D = \{(u, \bar{x}, \bar{v}) : 1 \geq u \geq v_1 \geq \cdots \geq v_{M-2} \geq \bar{x} - (u + v_1 + \cdots + v_{M-2})\}.$$

Using L'Hopital's rule,

$$u(0, \bar{x}, \bar{v}) = \bar{x} - \frac{\bar{x}^2}{2(M-1)} - \frac{1}{M-1} \sum_{j=2}^{M-1} (M-j)F_j,$$

which is equivalent to $G''(0) = 0$.

As in case 2, for each $z > 0$ the graph of (4) in the $(u, \bar{x}, \bar{v})$ domain represents all triples $(u, \bar{x}, \bar{v})$ for which $G$ has a zero at $z$. We show that $E = D \cap \{(w, \bar{x}, \bar{v}) : w < u(0, \bar{x}, \bar{v})\}$ is a disjoint union of these graphs. Existence of a zero for $G$ is guaranteed if and only if $F_1 < u(0, \bar{x}, \bar{v})$, or equivalently, if and only if $s^2 > \bar{x}$. The disjointness of the union implies that for a given triple $(u, \bar{x}, \bar{v}) \in E$ there is exactly one graph which passes through it, and so $G$ has exactly one zero $z$.

Differentiate $u$ with to obtain $u_{\bar{x}x} = -(M-1 - \bar{x})\bar{x}/[(M-1)(1 + \bar{x}z)]^2 < 0$ for $\bar{x} \in (0, M-1)$. Note that the graph of $u(0, \bar{x}, \bar{v})$ exists $D$ when $\bar{x} = M - 1$. Integrate over $(z_1, z_2)$ with $z_1 < z_2$ to obtain $u_{\bar{x}}(z_2, \bar{x}, \bar{v}) - u_{\bar{x}}(z_1, \bar{x}, \bar{v}) < 0$. Finally, integrate over $(0, \bar{x})$ to obtain

$$u(z_2, \bar{x}, \bar{v}) - u(z_2, 0, \bar{v}) < u(z_1, \bar{x}, \bar{v}) - u(z_1, 0, \bar{v}).$$

Unlike case 2, we have two extra terms (where $\bar{x} = 0$) which may affect the ordering of the graphs. However, restricting our attention to the domain $D$, when $\bar{x} = 0$, all the parameters must be zero: $u = 0$ and $\bar{v} = 0$. 

6
so on D we have \( u(z_2, \bar{z}, \bar{v}) < u(z_1, \bar{z}, \bar{v}) \) and the graphs are ordered as claimed. As \( z \to \infty \), the graphs of \( u(z, \bar{z}, \bar{v}) \) in \( D \) approach the graph of \( u(\infty, \bar{z}, \bar{v}) \equiv 0 \) in \( D \), which is the single point \( \bar{0} \), so \( E \) is the disjoint union of the graphs.

Acknowledgements. We thank Professor Ingram Olkin for his valuable comments.

3 References


