# PARAMETER ESTIMATION IN SOME MULTIVARIATE COMPOUND DISTRIBUTIONS

by

GEORGE E. J. SMITH

B.Sc., University of British Columbia, 1962

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ARTS

in the Department

of

Mathematics

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August, 1965

# ABSTRACT

During the past three decades or so there has been much work done concerning contagious probability distributions in an attempt to explain the behavior of certain types of biological populations. The distributions most widely discussed have been the Poisson-binomial, the Poisson Pascal or Poisson-negative binomial, and the Poisson-Poisson or Neyman Type A. Many generalizations of the above distributions have also been discussed.

The purpose of this work is to discuss the multivariate analogues of the above three distributions, i.e. the Poisson-multinomial, Poisson-negative multinomial, and Poisson-multivariate Poisson, respectively.

In chapter one the first of these distributions is discussed. Initially a biological model is suggested which leads us to a probability generating funtion. From this a recursion formula for the probabilities is found. Parameter estimation by the methods of moments and maximum likelihood is discussed in some detail and an approximation for the asymptotic efficiency of the former method is found. The latter method is asymptotically efficient. Finally sample zero and unit sample frequency estimators are briefly discussed.

In chapter two, exactly the same procedure is followed for the Poisson-negative multinomial distribution. Many close similarities are obvious between the two distributions.

The last chapter is devoted to a particular common limiting case of the first two distributions. This is the Poisson-multivariate Poisson. In this case the desired results are obtained by carefully considering appropriate limits in either of the previous two cases.

# TABLE OF CONTENTS

page

INTRODUCTION 1

CHAPTER I The Poisson Multinomial Distribution

1-1. A Biological Model . . . . . . 2

1-2. Probability Generating Function and Recursion Formula for Probabilities . . 4

1-3. Estimation of Parameters by the Method of Moments . . . . . 6

1-4. Maximum Likelihood Estimators . . 10

1-5. Covariance Matrix of Maximum Likelihood Estimators

A. Method of Calculation . . . . 16

B. Calculation of the elements of J . . 22

1-6. Efficiency of Method of Moments

A. Method of Calculation . . . . 26

B. Calculation of 'det M' in Terms of the Parameters . . . . . . 28

C. Determination of the Jacobian 'det $\phi$' 32

1-7. Sample Zero Frequency and Unit Sample Frequency Estimators .

A. Sample Zero Frequency and First Moments 34

B. Unit Sample Frequency Estimation 36

page

CHAPTER II The Poisson Negative-Multinomial Distribution

2-1. A Biological Model . . . . . . 38

2-2. Probability Generating Function and Recursion Formula for Probabilities . . 39

2-3. Estimation of Parameters by the Method of Moments . . . . . . 43

2-4. Maximum Likelihood Estimators . . . 46

2-5. Covariance Matrix of Maximum Likelihood Estimators

A. Method of Calculation . . . . 52

B. Calculation of the Elements of J* . 57

2-6. Efficiency of Method of Moments

A. Method of Calculation . . . . 61

B. Calculation of 'det M' in Terms of the Parameters . . . . . . 62

C. Determination of the Jacobian 'det $\phi^*$' 64

1-7. Sample Zero Frequency and Unit Sample Frequency Estimators

A. Sample Zero Frequency and First Moments 66

B. Unit Sample Frequency Estimators 68

page

CHAPTER III Limiting Distributions of the Poisson-Multinomial and Poisson-Negative Multinomial Distributions

3-1. Introduction . . . . . . . . 70

3-2. The Poisson-Poisson Distribution . . 70

3-3. The Information Matrix . . . . . 73

3-4. Efficiency of Method of Moments . . 75

3-5. Sample Zero Frequency and First Moments 78

3-6. Unit Sample Frequency Estimation . . 79

APPENDIX 1A Obtaining an Explicit Expression for the Probability Function . . 82

APPENDIX 1B Calculation of Moments from Factorial Cumulant Generating Function . . 86

APPENDIX 1C Calculation of the Entries of the Information Matrix "J" . . . 92

APPENDIX 1D Calculation of the Inverse of the Matrix $J/\beta$ . . . . . . 95

APPENDIX 1E A Lemma Showing $\Omega = \phi^T M \phi$ . . . 100

APPENDIX 1F Calculation of (1-6.12) and (2-6.5) 102

APPENDIX 1G Calculation of (1-6.16) and (2-6.6) 103

page

APPENDIX 1H Calculation of "det M" . . . . . 105

APPENDIX 2A Obtaining the Probability Generating Function $g^*(\vec{s})$ . . . . . 106

APPENDIX 2B Calculation of the Entries of the Information Matrix "J*" . . . 107

APPENDIX 2C Calculation of det $\phi^*$ . . . . . 110

APPENDIX 3A Calculation of Efficiency of Method of Moments for the Poisson-Multivariate Poisson Distribution . . 112

BIBLIOGRAPHY . . . . . . . . . . 115

## ACKNOWLEDGMENT

The author would like to express his sincere thanks to Dr. Stanley W. Nash whose many hours of guidance and constructive criticism made this thesis possible. He would also like to thank Dr. R.A. Restrepo for reading the final form, and Miss Carol Lambert for typing it. Finally, he wishes to thank the National Research Council of Canada and the University of British Columbia for their financial assistance.

# INTRODUCTION

In recent years there have been many attempts to investigate statistically the behavior of various insect and plant populations. It has been found that models using a simple normal, Poisson, or binomial distribution are generally inadequate. The negative binomial distribution has been used somewhat successfully by Fisher [1941], Anscombe [1950], Bliss [1953], and others.

More recently, compound or 'contagious' distributions have been applied to biological models with somewhat greater success. The three which are most commonly used are (1), the Poisson-binomial - McGuire, Brindley, and Bancroft [1957], Sprott [1958], Shumway and Gurland [1960]; (2) the Poisson-Pascal (or Poisson-negative binomial) - Katti and Gurland [1961]; and (3) the Poisson-Poisson (or Neyman Type A) - Neyman [1938], Douglas [1955]. Models based on these distributions, however, must assume homogeneity in the characteristics of the experimental plot. These might include soil type, amount of moisture present, type of vegetation present, etc. Attempts to relax this assumption or to generalize in other ways have been made by Neyman [1938], Feller [1943], Thomas [1949], Beall and Rescia [1953], and Gurland [1958].

So far, only the univariate case has been considered for the above compound distributions and their generalizations. The object of this treatise is to extend some of the results of the three compound distributions mentioned above to the multivariate case.

# CHAPTER I

# THE POISSON-MULTINOMIAL DISTRIBUTION

## 1-1. A Biological Model

In this model we assume there is a large field $A$ of area $S_A$, and homogeneous throughout, where batches of insect eggs are laid. Homogeneity implies that the probability density of the batches follows a uniform distribution over the field. We will also assume the position of a particular batch is independent of the positions of the others. This seems to be a reasonable assumption as long as the average distance between the batches is much greater than their size. Next, let us choose a region $B$ of $A$ which is far enough from the boundary of $A$ so that boundary effects will be negligable. Let us divide $B$ into many small plots or quadrats, $B_1$, $B_2$, ..., all having the same shape and area $S_{B_o}$, which is much smaller than $S_A$.

Let $Z$ be a random variable denoting the number of batches laid in a particular quadrat. If $M$ is the total number of eggs laid in $A$, then

$$P(Z = z) = \binom{M}{z} \left(\frac{S_{B_o}}{S_A}\right)^z \left(1 - \frac{S_{B_o}}{S_A}\right)^{Mz} .$$

If $M$ is assumed to be large, since $S_{B_o} << S_A$, this is approximately the Poisson distribution

$$\left.\begin{aligned} P(Z = z) &= e^{-\lambda} \cdot \frac{\lambda^z}{z!} \\ \text{where} \quad \lambda &= M \cdot \frac{S_{B_o}}{S_A} \end{aligned}\right\} \qquad (1\text{-}1.1)$$

Having outlined the breeding ground, let us consider the insects themselves. We suppose the insects coming from those eggs that hatch can be divided into $n-1$ classes on the basis of some distinguishing characteristic (e.g. colour, size, type of insect, etc.). For each integer $i$, $1 \leq i \leq n-1$, let $X_i$ be a random variable denoting the number of insects in a quadrat that are born into the $i^{th}$ class. We assume the probability of an insect being born into the $i^{th}$ class is $p_i$ and is independent of what happens to any other insect. The probability $p_n$ that an insect does not hatch is therefore

$$p_n = 1 - \sum_{i=1}^{n-1} p_i \qquad 1\text{-}1.2)$$

We now make the arbitrary assumption that exactly $N$ eggs are laid in each batch. Assume all the eggs hatch about the same time, and sometime later we count the number of insects in a quadrat, noting how many belong to each class. If we assume the effect of insects migrating into and out of the quadrat is negligable, we have that the conditional joint density of $X_1$, $X_2$, ..., $X_{n-1}$ is

$$P_{\vec{X}}(\vec{x}|Z=z) = P_{\vec{X}}(\vec{X} = \vec{x}|Z = z) = \binom{Nz}{\vec{x}} \prod_{i=1}^{n} p_i^{x_i} \qquad (1\text{-}1.3)$$

where we define

$$\left.\begin{aligned}
&\vec{X} = (X_1, X_2, \ldots, X_{n-1}) \\
&\binom{Nz}{x} = \begin{cases} \dfrac{(Nz)!}{x_1!\; x_2! \;\ldots\; x_n!} & \text{if } \sum_{i=1}^{n-1} x_i \le Nz,\; x_i \ge 0 \\ 0 & \text{otherwise} \end{cases} \\
&\quad \text{where} \quad x_n = Nz - \sum_{i=1}^{n-1} x_i
\end{aligned}\right\} \qquad (1\text{-}1.4)$$

Thus we have a multinomial distribution. We can combine (1-1.1) and (1-1.3) to get the joint density

$$P(\vec{X}=\vec{x}) = P_{\vec{X}}(\vec{x}) = \sum_{z=o}^{\infty} P(\vec{X}=\vec{x} \mid Z=z)\, P(Z=z)$$

$$= e^{-\lambda} \sum_{z=o}^{\infty} \frac{\lambda^z}{z!} \binom{Nz}{\vec{x}} \prod_{i=1}^{n} p_i^{\,x_i} \qquad (1\text{-}1.5)$$

## 1-2. Probability Generating Function and Recursion Formula for Probabilities

Generally it is much easier to calculate individual probabilities using a recursion formula rather than the density function. The first step in this direction is to find the probability generating function $g(\vec{s})$ where $\vec{s} = (s_1, s_2, \ldots, s_{n-1})$

$$g(\vec{s}) = E(s_1^{\,x_1}\, s_2^{\,x_2} \ldots s_{n-1}^{\,x_{n-1}})$$

$$= e^{-\lambda} \sum_{z=o}^{\infty} (\lambda^z/z!) \sum_{x_1=o}^{\infty} \ldots \sum_{x_{n-1}=o}^{\infty} \binom{Nz}{\vec{x}} p_n^{\,x_n} \prod_{i=1}^{n-1} (s_i p_i)^{x_i}$$

We may assume the upper limit of the sums to be $\infty$ because of the definition of $\binom{Nz}{\vec{x}}$ given in (1-1.4).

To simplify the above expression, we simply note that the n-1 summations on the right are the multinomial expansion of

$$\left(\sum_{i=1}^{n-1} s_i p_i + p_n\right)^{Nz}$$

Then

$$g(\vec{s}) = e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} \left[\sum_{i=1}^{n-1} s_i p_i + p_n\right]^{Nz}$$

and the result of summing this is

$$g(\vec{s}) = \exp\left\{\lambda\left[\sum_{i=1}^{n-1} s_i p_i + p_n\right]^N - \lambda\right\} \tag{1-2.1}$$

From (1-2.1) and the definition of a probability generating function, we can calculate the individual probabilities by means of

$$P_{\vec{x}}\ (\vec{x}+\vec{e}_k) = \frac{1}{x_k+1} \prod_{i=1}^{n-1} \frac{1}{x_i!}\ D_k D_1^{x_1} D_2^{x_2} \ldots D_{n-1}^{x_{n-1}} g(\vec{s})\Big|_{\vec{s}=0} \tag{1-2.2}$$

where $\vec{e}_k$ is the (n-1)-vector with 1 in the $k^{th}$ position and zeros elsewhere, and $D_i$ means the partial derivative with respect to $s_i$. Using the Leibnitz rule for multiple differentiation, we obtain the following results which are explicitely calculated in Appendix 1A.

$$\left.\begin{aligned} P_{\vec{x}}(\vec{0}) &= g(\vec{0}) = \exp\,[\lambda(p_n^{\,N}-1)] \\ P_{\vec{x}}(\vec{x}+\vec{e}_k) &= \frac{\lambda p_k p_n^{\,N-1}}{(x_k+1)} \sum_{y_1=0}^{x_1} \cdots \sum_{y_{n-1}=0}^{x_{n-1}} \frac{N!}{[N-\sum\limits_{i=1}^{n-1}(x_i-y_i)-1]!} \\ &\qquad [\prod_{i=1}^{n-1} \left(\frac{p_i}{p_n}\right)^{x_i-y_i} \frac{1}{(x_i-y_i)!}]\, P_{\vec{x}}(\vec{y}) \end{aligned}\right\} \qquad (1\text{-}2.3)$$

## 1-3. Estimation of Parameters by the Method of Moments

The first step in this method is necessarily to find the moments of the distribution. This can probably be best accomplished if we realize that $g(\vec{s})$ is the factorial moment generating function if we set $\vec{s} = (1, 1, \ldots, 1) \equiv \vec{1}$ instead of $\vec{s}=\vec{0}$. Now set

$$c(\vec{s}) = \log g(\vec{s}) = \lambda \left\{ [\sum_{i=1}^{n-1} s_i p_i + p_n]^N - 1 \right\} \qquad (1\text{-}3.1)$$

This is the factorial cumulant generating function from which we can easily calculate the factorial cumulants. Then, using tables relating moments and cumulants such as the one in David and Barton [1962], pages 142-3, we may find the factorial moments and finally the moments about the origin.

Since the above mentioned table relates cumulants and moments about the origin, it also relates factorial moments and cumulants since both have the same relationships. (David and Barton [1962], page 51).

Using the above procedure, we first define

$$\left.\begin{aligned}
G_1 &= N(\lambda+1) - 1 \\
G_2 &= N^2(\lambda^2+3\lambda+1) - 3N(\lambda+1) + 2 \\
G_3 &= N^3(\lambda^3+6\lambda+7\lambda+1) - 6N^2(\lambda^2+3\lambda+1) \\
&\qquad + 11N(\lambda+1) - 6
\end{aligned}\right\} \qquad (1\text{-}3.2)$$

Then the moments are, according to Appendix 1B,

$$\left.\begin{aligned}
E(X_i) &= N\lambda p_i \\
E(X_i^{\,2}) &= N\lambda p_i(p_i G_1+1) \\
&= N\lambda p_i[p_i N(\lambda+1) - p_i+1] \\
E(X_i X_j) &= N\lambda p_i p_j G_1 \\
&= N\lambda p_i p_j[N(\lambda+1) - 1]
\end{aligned}\right\} \qquad (1\text{-}3.3)$$

$$\left.\begin{aligned}
E(X_i^{\,3}) &= N\lambda p_i(p_i^{\,2}G_2+3p_i G_1+1) \\
E(X_i^{\,2}X_j) &= N\lambda p_i p_j(p_i G_2+G_1) \\
E(X_i X_j X_k) &= N\lambda p_i p_j p_k G_2 \\
E(X_i^{\,4}) &= N\lambda p_i(p_i^{\,3}G_3+6p_i^{\,2}G_2+7p_i G_1+1) \\
E(X_i^{\,3}X_j) &= N\lambda p_i p_j(p_i^{\,2}G_3+3p_i G_2+G_1) \\
E(X_i^{\,2}X_j^{\,2}) &= N\lambda p_i p_j[p_i p_j G_3+(p_i+p_j)G_2+G_1] \\
E(X_i X_j X_k X_m) &= N\lambda p_i p_j p_k p_m G_3 \\
E(X_i^{\,2}X_j X_k) &= N\lambda p_i p_j p_k(p_i G_3+G_2)
\end{aligned}\right\} \qquad (1\text{-}3.4)$$

Only the moments given in (1-3.3) are needed to estimate the parameters. The remaining ones, however, will be needed later to calculate the efficiency.

Let us define a new random variable W,

$$W = \sum_{i=1}^{n-1} X_i \tag{1-3.5}$$

Then $$E(W) = \sum_{i=1}^{n-1} E(X_i) = N\lambda(1-p_n) \tag{1-3.6}$$

and $$E(W^2) = E\left(\sum_{i=1}^{n-1} \sum_{j=1}^{n-1} X_i X_j\right)$$

$$= \sum_{i=1}^{n-1} E(X_i^2) + \sum_{i=1}^{n-1} \sum_{j \neq 1} E(X_i X_j)$$

If we substitute for the expectations from (1-3.3) and sum,

$$E(W^2) = N\lambda(1-p_n)\left\{1+[N(\lambda+1)-1](1-p_n)\right\} \tag{1-3.7}$$

Substituting (1-3.6) into (1-3.7), we obtain

$$E(W^2) = E(W)\ [E(W) + (N-1)(1-p_n)+1] \tag{1-3.8}$$

Now we can solve (1-3.6) for $1-p_n$, substitute into (1-3.8), and solve the resulting equation for $\lambda$. We obtain

$$\lambda = \frac{N-1}{N} \frac{E^2(W)}{E(W^2)-E^2(W)-E(W)} \tag{1-3.9}$$

From (1-3.3)

$$p_i = \frac{E(X_i)}{N\lambda} = \frac{E(X_i)[E(W^2)-E^2(W)-E(W)]}{(N-1)\,E^2(w)}$$
$$p_n = 1 - \sum_{i=1}^{n-1} p_i = 1 - \frac{E(W)}{N\lambda} \qquad (1\text{-}3.10)$$

(1-3.9) and (1-3.10) give the moment estimators $\lambda^*$ and $p_i^*$ for the paramters $\lambda$ and $p_i$ respectively, if the population moments are replaced by the corresponding sample moments.

Before proceeding, let us define some notation. Let $\beta$ be the number of samples we observe. Next define

$$\vec{x}_\alpha = (x_{1\alpha}, x_{2\alpha}, \ldots, x_{n-1,\alpha})$$

where $x_{j\alpha}$ is the observation of the $j^{th}$ characteristic from the $\alpha^{th}$ sample. Finally, define

$$x_{j\cdot} = \frac{1}{\beta}\sum_{\alpha=1}^{\beta} x_{j\alpha}$$
$$x_{\cdot\cdot} = \sum_{j=1}^{n-1} x_{j\cdot} = \frac{1}{\beta}\sum_{j=1}^{n-1}\sum_{\alpha=1}^{\beta} x_{j\alpha} \qquad (1\text{-}3.11)$$
$$w_\alpha = \sum_{j=1}^{n-1} x_{j\alpha}$$

So, $x_{i\cdot}$ is the mean of the number of insects with the $i^{th}$ characteristic per sample and $x_{\cdot\cdot}$ is the mean number of insects per sample.

Thus, from (1-3.9), we may write

$$\lambda^* = \frac{N-1}{N}\ \frac{x_{\cdot\cdot}^2}{(1/\beta)\sum_{\alpha=1}^{\beta} w_\alpha^2 - x_{\cdot\cdot}^2 - x_{\cdot\cdot}}$$

Also, $p_i$ is estimated by

$$\frac{\text{observed no. of insects with } i^{th} \text{ property}}{\text{no. of insects observed + estimated no. of unhatched eggs}}$$

i.e. $p_i^* = x_{i.}/N\lambda^*$

and $p_n$ is estimated by

$$\frac{\text{estimated no. of unhatched eggs}}{\text{no. of insects observed + estimated no. of unhatched eggs}}$$

$$p_n^* = 1 - \sum_{i=1}^{n-1} p_i^* = (N\lambda^* - x_{..})/N\lambda^*$$

## 1-4. Maximum Likelihood Estimators

The likelihood function, L, is given by

$$L = \prod_{\alpha=1}^{\beta} P_{\vec{x}}(\vec{x}_\alpha) \qquad (1\text{-}4.1)$$

To find the maximum likelihood estimators $\hat{\lambda}$ and $\hat{p}_i$ of $\lambda$ and $p_i$, we must solve the following system of equations.

$$\left.\begin{array}{l} \partial L/\partial\lambda = 0 \\ \partial L/\partial p_i = 0 \qquad i = 1, 2, \ldots, n-1 \end{array}\right\} \qquad (1\text{-}4.2)$$

Equation (1-4.1) can be written as follows

$$\log L = \sum_{\alpha=1}^{\beta} \log P_{\vec{x}}\,(\vec{x}_\alpha)$$

Because "log L" is a monotone increasing function of L, (1-4.2) is equivalent to solving the system

$$\left.\begin{aligned} \frac{\partial}{\partial\lambda} \log L &= \sum_{\alpha=1}^{\beta} \frac{1}{P_{\vec{x}}(\vec{x}_\alpha)} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}_\alpha) = 0 \\ \frac{\partial}{\partial p_j} \log L &= \sum_{\alpha=1}^{\beta} \frac{1}{P_{\vec{x}}(\vec{x}_\alpha)} \frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x}_\alpha) = 0 \end{aligned}\right\} \qquad (1\text{-}4.3)$$

Using (1-1.5) we can see that

$$\frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x}) = e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} \binom{Nz}{\vec{x}} \prod_{i=1}^{n-1} p_i^{x_i} \left[\frac{x_j}{p_j} - \frac{x_n}{p_n}\right]$$

Let us expand this expression into two terms. In the second term replace $\binom{Nz}{\vec{x}} x_n$ by $\binom{Nz}{\vec{x}+\vec{e}_j}(x_j+1)$ since these two expressions are equal. Then it is clear that

$$\begin{aligned} \frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x}) &= \frac{x_j}{p_j} P_{\vec{x}}(\vec{x}) - e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} \binom{Nz}{\vec{x}+\vec{e}_j} \frac{p_j}{p_n} \left(\prod_{i=1}^{n} p_i^{x_i}\right) \frac{x_j+1}{p_j} \\ &= \frac{x_j}{p_j} P_{\vec{x}}(\vec{x}) - \frac{x_j+1}{p_j} P_{\vec{x}}(\vec{x}+\vec{e}_j) \end{aligned} \qquad (1\text{-}4.4)$$

Equation (1-4.4) must hold for each observation, i.e. it holds for $\vec{x}=\vec{x}_\alpha$ and $x_j=x_{j\alpha}$. Substituting this into (1-4.3) we obtain

$$\frac{\partial}{\partial p_j} \log L = \sum_{\alpha=1}^{\beta} \frac{1}{P_{\vec{x}}(\vec{x}_\alpha)} \left[\frac{x_{j\alpha}}{p_j} P_{\vec{x}}(\vec{x}_\alpha) - \frac{x_{j\alpha}+1}{p_j} P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_j)\right] = 0$$

Multiplying by $p_j$ and using (1-3.11), this becomes

$$\hat{p}_j \frac{\partial}{\partial p_j} \log L = \beta x_{j.} - \sum_{\alpha=1}^{\beta} \frac{(x_{j\alpha}+1) P_{\vec{x}}(\vec{x}_\alpha + \vec{e}_j)}{P_{\vec{x}}(\vec{x}_\alpha)} = 0 \qquad (1\text{-}4.5)$$

Considering again the probability function (1-1.5), we can differentiate with respect to $\lambda$ and obtain

$$\frac{\partial}{\partial \lambda} P_{\vec{x}}(\vec{x}) = e^{-\lambda} \sum_{z=0}^{\infty} \frac{z\lambda^{z-1}}{z!} \binom{Nz}{\vec{x}} \prod_{i=1}^{n} p_i^{x_i} - P_{\vec{x}}(\vec{x}) \qquad (1\text{-}4.6)$$

Equation (1-1.4) implies

$$z = x_n/N + (1/N) \sum_{k=1}^{n-1} x_k .$$

If we substitute for the first $z$ in (1-4.6) using this expression, then

$$\frac{\partial}{\partial \lambda} P_{\vec{x}}(\vec{x}) = e^{-\lambda} \sum_{z=0}^{\infty} \frac{x_n}{N} \frac{\lambda^{z-1}}{z!} \binom{Nz}{\vec{x}} \prod_{i=1}^{n} p_i^{x_i}$$

$$+ e^{-\lambda} \sum_{z=0}^{\infty} \frac{1}{N} \left( \sum_{k=1}^{n-1} x_k \right) \frac{\lambda^{z-1}}{z!} \binom{Nz}{\vec{x}} \prod_{i=1}^{n} p_i^{x_i} - P_{\vec{x}}(\vec{x})$$

$$= \frac{p_n(x_m+1)}{p_m N\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} \binom{Nz}{\vec{x}+\vec{e}_m} \frac{p_m}{p_n} \prod_{i=1}^{n} p_i + \left(\frac{1}{N\lambda} \sum_{k=1}^{n-1} x_k - 1\right) P_{\vec{x}}(\vec{x})$$

where $m$ may be any integer between 1 and $n-1$ inclusive. Thus

$$\frac{\partial}{\partial \lambda} P_{\vec{x}}(\vec{x}) = \frac{p_n(x_m+1)}{p_m N\lambda} P_{\vec{x}}(\vec{x}+\vec{e}_m) + \left(\frac{1}{N\lambda} \sum_{k=1}^{n-1} x_k - 1\right) P_{\vec{x}}(\vec{x}) \qquad (1\text{-}4.7)$$

We should notice at this point that we actually have $n-1$ expressions for $\partial/\partial\lambda(P_{\vec{x}}(\vec{x}))$. These are all equal, however, and it does not matter which one we choose.

Equation (1-4.7) holds for each observation, i.e. when $\vec{x} = \vec{x}_\alpha$, $x_j = x_{j\alpha}$. Then, substituting (1-4.7) with this modification into the second equation of (1-4.3), we obtain

$$\sum_{\alpha=1}^{\beta} \left[-\frac{\hat{p}_n(x_{m\alpha}+1)}{\hat{p}_m N\hat{\lambda}} \frac{P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m)}{P_{\vec{x}}(\vec{x}_\alpha)} + \frac{1}{N\hat{\lambda}} \sum_{k=1}^{n-1} x_{k\alpha} - 1\right] = 0 \qquad (1\text{-}4.8)$$

Multiplying this by $N\hat{\lambda}\hat{p}_m$, substituting from (1-4.5) for the first term, and using the simplifying notation defined in (1-3.11),

$$\hat{p}_n x_{m.} + \hat{p}_m x_{..} - N\lambda\hat{p}_m = 0 \qquad (1\text{-}4.9)$$

If we sum over $m$, this yields

$$x_{..} - N\hat{\lambda}(1-\hat{p}_n) = 0$$

Hence $$\hat{p}_n = (N\hat{\lambda} - x_{..}) / N\hat{\lambda} \qquad (1\text{-}4.10)$$

Finally, let us substitute this into (1-4.9)

$$\hat{p}_m = x_{m.} / N\hat{\lambda} \quad m = 1, \ldots, n-1 \qquad (1\text{-}4.11)$$

Equations (1-4.10) and (1-4.11) give the maximum likelihood estimators for the $p_i$ if we know the corresponding estimator for $\lambda$. We could conceivably solve for $\hat{\lambda}$ by substituting (1-4.10) and (1-4.11) into (1-4.5) or (1-4.7) with the expression set equal to zero. However, it is not hard to see that it would be impossible to solve this directly for $\lambda$. Thus it is necessary to use a numerical procedure.

The following calculation is based on Newton's method, which says that if $f(\hat{\lambda}) = 0$, then

$$\hat{\lambda}_{n+1} = \hat{\lambda}_n - [f(\hat{\lambda}_n)/D_{\hat{\lambda}} f(\hat{\lambda}_n)] \tag{1-4.12}$$

where $\hat{\lambda}_n$ is the $n^{th}$ iteration of $\hat{\lambda}$ and $\hat{\lambda}_o$ is the initial estimate, which might be the moment estimator of $\lambda$. To find a suitable $f(\hat{\lambda})$ in our case, note from (1-4.9) that

$$\hat{p}_n/\hat{p}_m = (N\hat{\lambda} - x..)/x_{m.} \tag{1-4.13}$$

If we substitute for $\hat{p}_n/\hat{p}_m$ and $\sum_{\alpha=1}^{\beta} \sum_{k=1}^{n-1} x_{k\alpha}$ in (1-4. 8), using (1-4.13) and (1-3.11) respectively, and multiply the resulting equation by $N\hat{\lambda}/(N\hat{\lambda}-x..)$, we obtain an expression which is zero. We may use this as our $f(\hat{\lambda})$. Hence

$$f(\hat{\lambda}) = \sum_{\alpha=1}^{\beta} \frac{(x_{m\alpha}+1)}{x_{m.}} \frac{P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m)}{P_{\vec{x}}(\vec{x}_\alpha)} - \beta = 0 \tag{1-4.14}$$

It remains to find $D_{\hat{\lambda}} f(\hat{\lambda})$ for substitution into (1-4.12). From (1-4.14),

$$D_{\hat{\lambda}} f(\hat{\lambda}) = \sum_{\alpha=1}^{\beta} \frac{x_{m\alpha}+1}{x_{m.}} \Big\{ P_{\vec{x}}(\vec{x}_\alpha)\, D_{\hat{\lambda}} P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m) - P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m)\, D_{\hat{\lambda}} P_{\vec{x}}(\vec{x}_\alpha) \Big\} \frac{1}{P_{\vec{x}}^{2}(\vec{x}_\alpha)} \tag{1-4.15}$$

Now $$D_{\hat{\lambda}} P_{\vec{x}}(\vec{x}_\alpha) = \sum_{k=1}^{n-1} \frac{\partial}{\partial \hat{p}_k} P_{\vec{x}}(\vec{x}_\alpha) \cdot D_{\hat{\lambda}} \hat{p}_k + \frac{\partial}{\partial \hat{\lambda}} P_{\vec{x}}(\vec{x}_\alpha) \tag{1-4.16}$$

and from (1-4.11), $$D_{\hat{\lambda}} \hat{p}_m = -x_{m.}/N\hat{\lambda}^2 = -\hat{p}_m/\hat{\lambda} \tag{1-4.17}$$

Let us substitute expressions for the derivatives in (1-4.16).

Use (1-4.4) to replace $\partial/\partial\hat{p}_k[P_{\vec{x}}(\vec{x}_\alpha)]$, (1-4.7) to replace $\partial/\partial\lambda[P_{\vec{x}}(\vec{x}_\alpha)]$ and (1-4.17) to replace $D_{\hat{\lambda}}\hat{p}_m$. Then we may use (1-4.13) to replace $\hat{p}_n/\hat{p}_m$ and (1-3.11) to replace the resulting expression.

$$D_{\hat{\lambda}}\, P_{\vec{x}}(\vec{x}_\alpha) = (1/\hat{\lambda}) \sum_{k=1}^{n-1} (x_{k\alpha}+1)\, P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_k) - [\langle (N-1)w_\alpha/N\hat{\lambda}\rangle +1]$$

$$\cdot P_{\vec{x}}(\vec{x}_\alpha) + \left(\frac{N\hat{\lambda}-x_{..}}{x_{m.}}\right) \frac{(x_{m\alpha}+1)}{N\hat{\lambda}}\, P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m) \tag{1-4.18}$$

If we replace $\vec{x}_\alpha$ by $\vec{x}_\alpha+\vec{e}_m$, and hence $x_{m\alpha}$ by $x_{m\alpha}+1$, and use (1-3.11) to replace $\sum_{k=1}^{n-1} x_{k\alpha}$

$$D_{\hat{\lambda}}P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m) = (1/\hat{\lambda}) \sum_{k=1}^{n-1} (x_{k\alpha}+1) P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_k+\vec{e}_m)$$

$$- \left[1+ \frac{N-1}{N\hat{\lambda}}\, (w_\alpha+1)\right] P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m) + \left[1/\hat{\lambda} + \frac{N\hat{\lambda}-x_{..}-1}{x_{m.}+1} \cdot \frac{x_{m\alpha}+2}{N\hat{\lambda}}\right]$$

$$\cdot P_{\vec{x}}(\vec{x}_\alpha+2\vec{e}_m) \tag{1-4.19}$$

The $(r+1)^{st}$ iterated values of $\lambda$, $p_1$, ..., $p_{n-1}$ can be calculated from the $r^{th}$ iterated values by the following procedure. First substitute the $r^{th}$ iterated values into (1-4.18) and (1-4.19) to obtain $D_{\hat{\lambda}}P_{\vec{x}}(\vec{x}_\alpha)$ and $D_{\hat{\lambda}}P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m)$. Then substitute these into (1-4.15) to find $D_{\hat{\lambda}}f(\hat{\lambda})$ which is finally substituted into (1-4.12) for $\hat{\lambda}_{r+1}$. Then the $(r+1)^{st}$ iterated values of the $\hat{p}_m$ can be found from (1-4.10) and (1-4.11).

## 1-5. Coveriance Matrix of Maximum Likelihood Estimators

### A. Method of Calculation

Direct calculation of the covariance matrix, $\hat{\Omega}$, of the maximum likelihood estimators is practically impossible. Under certain conditions, however, we may find the asymptotic $\hat{\Omega}$ as $\beta \to \infty$ by first calculating the information matrix

$$J = \begin{pmatrix} I_{\lambda\lambda} & I_{\lambda p_1} \cdots & I_{\lambda p_{n-1}} \\ I_{p_1\lambda} & I_{p_1 p_1} \cdots & I_{p_1 p_{n-1}} \\ \vdots & \vdots & \vdots \\ I_{p_{n-1}\lambda} & I_{p_{n-1}\lambda} & I_{p_{n-1}p_{n-1}} \end{pmatrix} \tag{1-5.1}$$

where $I_{st} = E(\frac{\partial}{\partial s} \log L \cdot \frac{\partial}{\partial t} \log L) = -E(\frac{\partial^2}{\partial s \partial t} \log L)$ (1-5.2)

where $L$ is the likelihood function. The second equality is true by the argument presented in Kendall and Stuart [1961] pp. 52-53. From the remarks at the beginning of § 1-3 we conclude that the factorial moment generating function is given by (1-2.1) where $\vec{s}$ is set equal to $\vec{1}$ instead of $\vec{0}$. Since $g(\vec{s})$ is clearly infinitely differentiable, all the factorial moments are finite. Because each moment about the origin is a finite linear combination of factorial moments, these moments are also finite.

Lemma 1-1.

Let $X, Y,$ and $Z$ be non-negative, identically distributed and mutually independent random variables with $E(X^3) < \infty$. Then

$$0 \leq E(XYZ) \leq E(X^2Y) \leq E(X^3).$$

Proof:

Consider $0 \leq E(XY^{\frac{1}{2}} - Y^{\frac{1}{2}}Z)^2 = E(X^2Y) - 2E(XYZ) + E(XY^2)$. Because of the mutual independence and identical distribution of the random variables, $E(X^2Y) = E(X^2)\,E(Y) = E(Z^2)\,E(Y) = E(YZ^2)$. Thus $0 \leq E(X^2Y) - E(XYZ)$. Now, noting that $XYZ \geq 0$ always, we have

$$0 \leq E(XYZ) \leq E(X^2Y)$$

Now consider $0 \leq E(X^{3/2} - X^{1/2}Y)^2 = E(X^3) - 2E(X^2Y) + E(XY^2)$. But $E(X^2Y) = E(XY^2)$ since distributions are identical. Therefore $0 \leq E(X^3) - E(X^2Y)$ and so $E(X^2Y) \leq E(X^3)$. Combining this with the previous result we have $0 \leq E(XYZ) \leq E(X^2Y) \leq E(X^3)$.

Q.E.D.

Lemma 1-2.

For the Poisson-multinomial distribution

$$P_{\vec{X}}(\vec{x}+\vec{e}_m) \leq (p_m/p_n)\left\{ \sum_{i=1}^{n-1} x_i + N + \lambda N \left( \sum_{i=1}^{n-1} x_i + N\right)^N \right\} P_{\vec{X}}(\vec{x})$$

Proof:

From (1-1.5),

$$P_{\vec{X}}(\vec{x}+\vec{e}_m) = e^{-\lambda} \sum_{z=0}^{\infty} (\lambda^z/z!) \binom{Nz}{\vec{x}+\vec{e}_m} \left(\prod_{i=1}^{n} p_i^{x_i}\right)(p_m/p_n) \quad (1\text{-}5.3)$$

Let $T$ be an integer such that

$$N(T-1) \le \sum_{i=1}^{n-1} x_i < NT \tag{1-5.4}$$

T depends on $\vec{x}$. From the definition of the multinomial coefficient, we see that the first T terms of the sum in (1-5.3) are zero. Thus, if we write the multinomial coefficient in a slightly different way,

$$P_{\vec{x}}(\vec{x}+\vec{e}_m) = \sum_{z=T}^{\infty} \frac{\lambda^z}{z!} \binom{Nz}{\vec{x}} \left(\prod_{i=1}^{n} p_i^{x_i}\right) \frac{Nz-x_1-\ldots-x_{n-1}}{x_m+1} (p_m/p_n)$$

$$\le (p_m/p_n) \sum_{z=T}^{\infty} (\lambda^z/z!) \binom{Nz}{\vec{x}} \left(\prod_{i=1}^{n} p_i^{x_i}\right) Nz$$

$$= (p_m/p_n)\left\{ NT(\lambda^T/T!)\binom{NT}{\vec{x}} \left(\prod_{i=1}^{n} p_i^{x_i}\right) + N \sum_{z=T+1}^{\infty} \frac{\lambda^z}{(z-1)!} \binom{Nz}{\vec{x}} \left(\prod_{i=1}^{n} p_i^{x_i}\right)\right\} \tag{1-5.5}$$

Note that the first term in the brackets is NT times the "z=T" term in the expansion of $P_{\vec{x}}(\vec{x})$ and, since each term is positive,

$$NT(\lambda^T/T!)\binom{NT}{\vec{x}} \left(\prod_{i=1}^{n} p_i^{x_i}\right) \le NTP_{\vec{x}}(\vec{x}) \le \left(\sum_{i=1}^{n-1} x_i+N\right) P_{\vec{x}}(\vec{x}) \tag{1-5.6}$$

Now consider the second term in (1-5.5). Because $\sum_{i=1}^{n-1} x_i < NT$ and $z \ge T+1$,

$$\binom{Nz}{\vec{x}} = \frac{(Nz-N)!}{\left(\prod_{i=1}^{n-1} x_i!\right)\left(Nz-N-\sum_{i=1}^{n-1} x_i\right)!} \prod_{k=o}^{N-1} \frac{(Nz-k)}{\left(Nz-k-\sum_{i=1}^{n-1} x_i\right)} \tag{1-5.7}$$

Now since $k<N$ and $z-1 \ge T$ we have

$$\frac{Nz - k}{Nz-k-\sum_{i=1}^{n-1} x_i} \leq \frac{Nz - N}{Nz-N-\sum_{i=1}^{n-1} x_i} \leq \frac{NT}{NT-\sum_{i=1}^{n-1} x_i} \leq \sum_{i=1}^{n-1} x_i + N$$

Thus (1-5.7) becomes

$$\binom{Nz}{\vec{x}} \leq \binom{Nz - N}{\vec{x}} \left( \sum_{i=1}^{n-1} x_i + N\right)^N$$

Hence the second term on the right side of (1-5.5) is less than or equal to

$$\lambda N \sum_{z=T+1}^{\infty} \frac{\lambda^{z-1}}{(z-1)!} \binom{Nz-N}{\vec{x}} \left(\prod_{i=1}^{n} p_i^{\,x_i}\right) \left(\sum_{i=1}^{n-1} x_i+N\right)^N$$

This expression is $\lambda N(\sum_{i=1}^{n-1} x_i+N)^N$ times part of the expansion of $P_{\vec{x}}(\vec{x})$ and thus is less than $\lambda N(\sum_{i=1}^{n-1} x_i+N)^N$.

Using this fact along with (1-5.6) in (1-5.5) we have

$$P_{\vec{x}}(\vec{x}+\vec{e}_m) \leq (p_m/p_n)\left\{ \sum_{i=1}^{n-1} x_i + N + \lambda N \left(\sum_{i=1}^{n-1} x_i + N\right)^N \right\} P_{\vec{x}}(\vec{x})$$

Q.E.D.

<u>Lemma 1-3</u>.

For the Poisson-multinomial distribution, $E[\,|\partial/\partial\lambda(\log L)|^3] < \infty$

<u>Proof</u>:

Using (1-4.1) and differentiating the logarithm,

$$E\left(\left|\frac{\partial}{\partial\lambda} \log L\right|^3\right) = E\left(\left|\sum_{\alpha=1}^{\beta} \frac{1}{P_{\vec{x}}(\vec{x}_\alpha)} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}_\alpha)\right|^3\right)$$

$$\leq \sum_{\alpha=1}^{\beta} \sum_{\gamma=1}^{\beta} \sum_{\delta=1}^{\beta} E\left(\left|\frac{\partial/\partial\lambda[P_{\vec{x}}(\vec{x}_\alpha)]}{P_{\vec{x}}(\vec{x}_\alpha)} \frac{\partial/\partial\lambda[P_{\vec{x}}(\vec{x}_\gamma)]}{P_{\vec{x}}(\vec{x}_\gamma)} \frac{\partial/\partial\lambda[P_{\vec{x}}(\vec{x}_\delta)]}{P_{\vec{x}}(\vec{x}_\delta)}\right|\right)$$

Because the observations are mutually independent, the expressions $\frac{\partial/\partial\lambda[P_{\vec{X}}(\vec{x}_\alpha)]}{P_{\vec{X}}(\vec{x}_\alpha)}$ are independent for $\alpha=1, 2, \ldots, \beta$. Hence we may apply lemma 1-1 to the above inequality.

$$E\left(\left|\frac{\partial}{\partial\lambda}\log L\right|^3\right) \leq \beta^2 \sum_{\alpha=1}^{\beta} E\left(\left|\frac{\partial/\partial\lambda[P_{\vec{X}}(\vec{x}_\alpha)]}{P_{\vec{X}}(\vec{x})}\right|^3\right)$$

Since the observations are identically distributed, the above expression is independent of $\alpha$, hence

$$E\left(\left|\frac{\partial}{\partial\lambda}\log L\right|^3\right) \leq \beta^3 E\left(\left|\frac{\partial/\partial\lambda[P_{\vec{X}}(\vec{x})]}{P_{\vec{X}}(\vec{x})}\right|^3\right) \tag{1-5.8}$$

and substituting for $\partial/\partial\lambda[P_{\vec{X}}(\vec{x})]$ from (1-4.7),

$$E\left(\left|\frac{\partial}{\partial\lambda}\log L\right|^3\right) \leq \beta^3 E\left(\left|\frac{p_n(x_m+1)}{p_m N\lambda}\cdot\frac{P_{\vec{X}}(\vec{x}+\vec{e}_m)}{P_{\vec{X}}(\vec{x})} + (1/N\lambda)\sum_{i=1}^{n-1} x_i - 1\right|^3\right)$$

Replacing the absolute value of the sum by the sum of the absolute values and using lemma 1-2,

$$\leq \beta^3 E\left\{\frac{x_m+2}{N\lambda}\sum_{i=1}^{n-1} x_i + (x_m+1)[1/\lambda + (\sum_{i=1}^{n-1} x_i+N)^N] + 1\right\}$$

When the above expression is expanded, it will yield a finite sum of terms of the following type -

$$\left.\begin{array}{l} \text{constant}\cdot E\left(\prod_{j=1}^{n-1} x_j^{n_j}\right) \\ \text{where the } n_j \text{ are non-negative integers} \end{array}\right\} \tag{1-5.9}$$

These terms are all finite since we know all the moments are finite. Hence the result follows.

Q.E.D.

Lemma 1-4.

For the Poisson-multinomial distribution

$E(|\partial/\partial\, p_i(\log L)|^3) < \infty \quad , \quad i = 1, 2, \ldots, n-1.$

Proof:

By the same argument as in lemma 1-3 but using $p_i$ instead of $\lambda$ , we will obtain (1-5.3) with $p_i$ replacing $\lambda$ , i.e.

$$E(|\frac{\partial}{\partial p_i} \log L|^3) \leq \beta^3 E(|\frac{\partial/\partial p_i[P_{\vec{x}}(\vec{x})]}{P_{\vec{x}}(\vec{x})}|^3)$$

Substituting for the derivative using (1-4.4), we find

$$E(|\frac{\partial}{\partial p_i} \log L|^3) \leq \beta^3 E(|\frac{x_j}{p_j} - \frac{x_j+1}{p_j} \frac{P_{\vec{x}}(\vec{x}+\vec{e}_j)}{P_{\vec{x}}(\vec{x})}|^3)$$

Replacing the absolute value of the sum by the sum of the absolute values and using lemma 1-2.

$$\leq E\left\{ \frac{x_j}{p_j} - \frac{x_j+1}{p_n} [\sum_{i=1}^{n-1} x_i + N + \lambda N (\sum_{i=1}^{n-1} x_i + N)^N] \right\}$$

Upon expansion, the above expression becomes a finite sum of terms of the type described in (1-5.9), and by the same reasoning as was used there, $E[|\partial/\partial p_i\ (\log L)|^3]$ is finite.

Q.E.D.

Let us appeal to theorem 2, page 282 in Rao [1947]. This theorem says the following -

Let $\hat{\Omega}$ be the covariance matrix of the maximum likelihood estimators $\theta_1, \theta_2, \ldots, \theta_{n-1}$, and $L$ be the likelihood function.

Then, if $E[|\partial/\partial\theta_i\ (\log L)|^{2+\eta}] < \infty$, $i = 1, 2, \ldots, n$ for some $\eta > 0$,

$$\lim_{\beta\to\infty} \beta\, J^{-1} = \lim_{\beta\to\infty} \beta\hat{\Omega} \tag{1-5.10}$$

where $J$ is the information matrix and $\beta$ is the number of samples observed.

Lemmas (1-3) and (1-4) show that the Poisson-multinomial distribution satisfies the conditions of Rao's theorem if we choose $\eta=1$. Hence (1-5.10) holds, and thus for samples of reasonable size we can make the approximation

$$J^{-1} = \hat{\Omega} \tag{1-5.11}$$

## B. Calculation of the Elements of J

Before proceeding let us first prove the following result.

Lemma 1-5.

Define $$A_{ij} = -1 + \sum_{x_1=o}^{\infty} \cdots \sum_{x_{n-1}=o}^{\infty} \frac{(x_i+1)(x_j+1)P_{\vec{x}}(\vec{x}+\vec{e}_i)P_{\vec{x}}(\vec{x}+\vec{e}_j)}{N^2\lambda^2\, p_i p_j\ P_{\vec{x}}(\vec{x})} \tag{1-5.12}$$

Then $A_{ij} = A_{mk} = A$, say, for $m,i,j,k = 1, 2, \ldots, n-1$.

Proof:

$$A_{ij} - A_{mk} = A_{ij} - A_{im} + A_{im} - A_{mk}$$

$$= \sum_{x_1=o}^{\infty} \cdots \sum_{x_{n-1}=o}^{\infty} \frac{(x_i+1)P_{\vec{x}}(\vec{x}+\vec{e}_i)}{N^2\lambda^2 p_i P_{\vec{x}}(\vec{x})} \left[\frac{(x_j+1)P_{\vec{x}}(\vec{x}+\vec{e}_j)}{p_j} - \frac{(x_m+1)P_{\vec{x}}(\vec{x}+\vec{e}_m)}{p_m}\right]$$

$$+\sum_{x_1=o}^{\infty}\cdots\sum_{x_{n-1}=o}^{\infty}\frac{(x_m+1)P_{\vec{x}}(\vec{x}+\vec{e}_m)}{N^2\lambda^2 p_m P_{\vec{x}}(\vec{x})}\left[\frac{(x_i+1)P_{\vec{x}}(\vec{x}+\vec{e}_i)}{p_i}-\frac{(x_k+1)P_{\vec{x}}(\vec{x}+e_k)}{p_k}\right.$$

Recall now that (1-4.7) is true for all values of $m$ from 1 to n-1. This is possible only if

$$\frac{(x_m+1)}{p_m N\lambda}\ P_{\vec{x}}(\vec{x}+\vec{e}_m)=\frac{(x_k+1)}{p_k N\lambda}\ P_{\vec{x}}(\vec{x}+\vec{e}_k).$$

Using this fact, the terms in the brackets of the above equation are zero and hence the whole expression is zero.

Q.E.D.

Let us consider $I_{\lambda\lambda}$. From (1-5.2) and (1-4.1)

$$I_{\lambda\lambda}=E[(\frac{\partial}{\partial\lambda}\log L)^2]=E([\sum_{\alpha=1}^{\beta}\frac{\partial}{\partial\lambda}\log P_{\vec{x}}(\vec{x}_\alpha)]^2)$$

$$=\sum_{\alpha=1}^{\beta}\sum_{\gamma=1}^{\beta}E[\frac{\partial}{\partial\lambda}\log P_{\vec{x}}(\vec{x}_\alpha)\cdot\frac{\partial}{\partial\lambda}\log P_{\vec{x}}(\vec{x}_\gamma)]$$

Because the observations are independent of each other and also have identical distributions,

$$I_{\lambda\lambda}=\beta\ E([\frac{\partial}{\partial\lambda}\log P_{\vec{x}}(\vec{x})]^2)+\beta(\beta-1)\ E^2[\frac{\partial}{\partial\lambda}\log P_{\vec{x}}(\vec{x})] \qquad (1\text{-}5.13)$$

At this point let us observe that

$$E[\frac{\partial}{\partial\lambda}\log P_{\vec{x}}(\vec{x})]=E[\frac{1}{P_{\vec{x}}(\vec{x})}\ \frac{\partial}{\partial\lambda}P_{\vec{x}}(\vec{x})]$$

$$=\sum_{x_1=o}^{\infty}\cdots\sum_{x_{n-1}=o}^{\infty}\frac{\partial}{\partial\lambda}\ P_{\vec{x}}(\vec{x})=\frac{\partial}{\partial\lambda}\sum_{x_1=o}^{\infty}\cdots\sum_{x_{n-1}=o}^{\infty}P_{\vec{x}}(\vec{x})$$

$$=\frac{\partial}{\partial\lambda}(1)=0 \qquad (1\text{-}5.14)$$

Thus (1-5.13) becomes, after differentiating the logarithm,

$$I_{\lambda\lambda} = \beta\ E\ ([\frac{1}{P_{\vec{x}}(\vec{x})} \cdot \frac{\partial}{\partial\lambda}\ P_{\vec{x}}(\vec{x})]^2) \tag{1-5.15}$$

Substituting for the derivative from (1-4.7) and using the definition of expectation,

$$(1/\beta)\ I_{\lambda\lambda} = \sum_{x_1=0}^{\infty} \cdots \sum_{x_{n-1}=0}^{\infty} \frac{1}{P_{\vec{x}}(\vec{x})} \left\{ \frac{p_n(x_m+1)}{p_m N\lambda} \quad P_{\vec{x}}(\vec{x}+\vec{e}_m) \right.$$

$$\left. + [(1/N\lambda) \sum_{k=1}^{n-1} x_k - 1]\ P_{\vec{x}}(\vec{x}) \right\}^2$$

If we now square, replace the resulting sums by the moments they represent and use (1-5.12) along with lemma 1-5,

$$\frac{1}{\beta}\ I_{\lambda\lambda} = p_n^{\ 2}\ (A+1) + (1/N^2\lambda^2)\ E[(\sum_{k=1}^{n-1} X_k)^2] + 1$$

$$- (2p_n/p_m N\lambda)\ E[X_m(\sum_{k=1}^{n-1} X_k - 1)] + (2p_n/p_m N^2\lambda^2)\ E(X_m)$$

$$- (2/N\lambda)\ E(\sum_{k=1}^{n-1} X_k)$$

We can evaluate the various moments by expansion and the use of (1-3.3). After simplifying the resulting expression, we obtain

$$(1/\beta)\ I_{\lambda\lambda} = p_n^{\ 2}\ A + (1-p_n)(N\lambda + Np_n - p_n)/N\lambda \tag{1-5.16}$$

By a similar procedure the other entries of the information matrix may be calculated. The results are

$$\left.\begin{aligned}
(1/\beta)\, I_{\lambda p_j} &= (1/\beta)\, I_{p_j\lambda} = -p_n N\lambda A + Np_n + 1 - p_n \\
(1/\beta)\, I_{p_i p_i} &= N^2\lambda^2 A + N\lambda(1/p_i + 1 - N) \\
(1/\beta)\, I_{p_i p_j} &= N^2\lambda^2 A + N\lambda(1-N) \qquad \text{for} \quad i \neq j
\end{aligned}\right\} \tag{1-5.17}$$

The calculations of the above results are done in Appendix 1C. Let us define

$$\left.\begin{aligned}
B_{\lambda\lambda} &= (1/\beta)\, I_{\lambda\lambda} \\
B_{\lambda p} &= (1/\beta)\, I_{\lambda p_i} \\
B_{pp} &= (1/\beta)\, I_{p_i p_j} \qquad j \neq i
\end{aligned}\right\} \tag{1-5.18}$$

Thus $\quad B_{pp} + N\lambda/p_i = (1/\beta)\, I_{p_i p_i}$ (1-5.19)

If we substitute (1-5.18) and (1-5.19) into (1-5.1),

$$J = \beta \begin{bmatrix}
B_{\lambda\lambda} & B_{\lambda p} & \cdots & B_{\lambda p} \\
B_{\lambda p} & B_{pp}+N\lambda/p_1 & \cdots & B_{pp} \\
\vdots & \vdots & & \vdots \\
B_{\lambda p} & B_{pp} & \cdots & B_{pp}+N\lambda/p_{n-1}
\end{bmatrix} \tag{1-5.20}$$

By (1-5.11) the inverse of this matrix is $\hat{\Omega}$. In Appendix 1D the inverse of such a matrix is calculated in detail. By making the appropriate association of variables we have

$$\det J = \beta^n \left\{ B_{\lambda\lambda} + (B_{\lambda\lambda}B_{pp} - B_{\lambda p}{}^2)(1-p_n)/N\lambda \right\} \prod_{i=1}^{n-1} (N\lambda/p_i) \tag{1-5.21}$$

$$\left.\begin{aligned}
\operatorname{var}\hat{\lambda} &= \frac{1}{\beta}\,\frac{N\lambda + B_{pp}(1-p_n)}{B_{\lambda\lambda}N\lambda + (B_{\lambda\lambda}B_{pp}-B_{\lambda p}^{2})(1-p_n)} \\
\operatorname{cov}(\hat{\lambda},\hat{p}_i) &= -\frac{1}{\beta}\,\frac{B_{\lambda p}\,p_i}{B_{\lambda\lambda}N\lambda+(B_{\lambda\lambda}B_{pp} - B_{\lambda p}^{2})(1-p_n)} \\
\operatorname{cov}(\hat{p}_i,\hat{p}_j) &= -\frac{1}{\beta}\,\frac{(B_{\lambda\lambda}B_{pp} - B_{\lambda p}^{2})\,p_ip_j}{[B_{\lambda\lambda}N\lambda + (B_{\lambda\lambda}B_{pp} - B_{\lambda p}^{2})(1-p_n)]N\lambda} \\
\operatorname{var}\hat{p}_i &= \frac{1}{\beta}\left[\frac{p_i}{N\lambda} - \frac{(B_{\lambda\lambda}B_{pp}-B_{\lambda p}^{2})\,p_i^{2}}{[B_{\lambda\lambda}N\lambda + (B_{\lambda\lambda}B_{pp}-B_{\lambda p}^{2})(1-p_n)]N\lambda}\right]
\end{aligned}\right\} \quad (1\text{-}5.22)$$

Corollaries 2.1 and 2.2 in Rao [1947] state that if the distribution satisfies lemmas 1-3 and 1-4, then the maximum likelihood estimators are minimum variance estimators for large samples and in terms of the generalized variance, det $\hat{\Omega}$, are asymptotically efficient.

## 1-6. Efficiency of the Method of Moments

### A. Method of Calculation

The efficiency of a method of parameter estimation for a multivariate distribution is defined to be

$$\text{Eff} = \frac{\det C_M}{\det C} \qquad (1\text{-}6.1)$$

where $C$ is the covariance matrix of the estimators for the method in question and $C_M$ is the covariance matrix of the minimum variance estimators.

Because the Poisson-multinomial distribution satisfies the conditions of lemmas 1-3 and 1-4, corollary 2-2 in Rao [1947] states that the maximum likelihood estimators have minimum variance. Thus, in our case, $C_M = \hat{\Omega}$ and $C = \bar{\Omega}$, the covariance matrix of the moment estimators. Hence Eff = (det $\hat{\Omega}$)/(det $\bar{\Omega}$) and by (1-5.11),

$$\text{Eff} = \frac{1}{\det \bar{\Omega} \cdot \det J} \tag{1-6.2}$$

To calculate $\bar{\Omega}$, we will first find the covariance matrix, M, of the moment estimators. Let us define $\overline{W^2}$ and $\bar{X}_i$ as follows

$$\left.\begin{aligned} \overline{W^2} &= (1/\beta) \sum_{\alpha=1}^{\beta} \Big( \sum_{k=1}^{n-1} X_{k\alpha} \Big)^2 \\ \bar{X}_i &= (1/\beta) \sum_{\alpha=1}^{\beta} X_{i\alpha} \end{aligned}\right\} \tag{1-6.3}$$

where $X_{k\alpha}$ is the random variable denoting the number of insects observed with the $k^{th}$ characteristic on the $\alpha^{th}$ observation. $\overline{W^2}$ estimates $E(W^2)$ and $\bar{X}_i$ estimates $E(X_i)$. By definition

$$M = \begin{pmatrix} \text{var } \overline{W^2} & \text{cov}(\overline{W^2},\bar{X}_1) & \text{cov}(\overline{W^2},\bar{X}_2) & \cdots & \text{cov}(\overline{W^2},\bar{X}_{n-1}) \\ \text{cov}(\overline{W^2},\bar{X}_1) & \text{var } \bar{X}_1 & \text{cov}(\bar{X}_1,\bar{X}_2) & \cdots & \text{cov}(\bar{X}_1,\bar{X}_{n-1}) \\ \text{cov}(\overline{W^2},\bar{X}_2) & \text{cov}(\bar{X}_1,\bar{X}_2) & \text{var } \bar{X}_2 & \cdots & \text{cov}(\bar{X}_2,\bar{X}_{n-1}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{cov}(\overline{W^2},\bar{X}_{n-1}) & \text{cov}(\bar{X}_1,\bar{X}_{n-1}) & \text{cov}(\bar{X}_2,\bar{X}_{n-1}) & \cdots & \text{var } \bar{X}_{n-1} \end{pmatrix} \tag{1-6.4}$$

By Appendix 1E we can approximate $\det \bar{\Omega}$ by

$$\det \bar{\Omega} = (\det \phi)^2 \det M \tag{1-6.5}$$

where $\det \phi$ is the Jacobian

$$\det \phi = \frac{\partial[\lambda, p_1, \ldots, p_{n-1}]}{\partial[E(\overline{W^2}), E(\bar{X}_1), \ldots, E(\bar{X}_{n-1})]} \tag{1-6.6}$$

## B. Calculation of det M in Terms of the Parameters

First we must express the elements of $M$ in terms of $p_1, p_2, \ldots, p_{n-1}, \lambda$. Consider

$$\operatorname{var} \bar{X}_i = \operatorname{var} [(1/\beta) \sum_{\alpha=1}^{\beta} X_{i\alpha}]$$

Because the observations are independent and identically distributed for each observation,

$$\begin{aligned} \operatorname{var} \bar{X}_i &= (1/\beta^2) \sum_{\alpha=1}^{\beta} \operatorname{var} X_{i\alpha} = (1/\beta) \operatorname{var} X_i \\ &= (1/\beta) [E(X_i^{\,2}) - E^2(X_i)] \end{aligned} \tag{1-6.7}$$

If we now replace the expectations by the expressions given in (1-3.3), we will have

$$\operatorname{var} \bar{X}_i = (1/\beta) N\lambda\, p_i \, [p_i(N-1) + 1] \tag{1-6.8}$$

Now consider $\operatorname{cov} (\bar{X}_i, \bar{X}_j)$ $i \neq j$.

$$\operatorname{cov} (\bar{X}_i, \bar{X}_j) = \operatorname{cov} [(1/\beta) \sum_{\alpha=1}^{\beta} X_{i\alpha}, (1/\beta) \sum_{\gamma=1}^{\beta} X_{j\gamma}]$$

$$= (1/\beta^2)\, E\left\{ [\sum_{\alpha=1}^{\beta} X_{i\alpha} - \beta E(X_i)][\sum_{\gamma=1}^{\beta} X_{j\alpha} - \beta E(X_j)]\right\}$$

The last equality is true since $E(X_{i\alpha}) = E(X_i)$ for all $\alpha$. Also, since the observations are independent,

$$\operatorname{cov}(\overline{X}_i,\overline{X}_j) = (1/\beta^2)\sum_{\alpha=1}^{\beta} E\left\{[X_{i\alpha} - E(X_i)][X_{j\alpha} - E(X_j)]\right\}$$

$$+ (1/\beta^2)\sum_{\alpha=1}^{\beta}\sum_{\gamma\neq\alpha} E[X_{i\alpha} - E(X_i)]\, E[X_{j\alpha} - E(X_j)]$$

But $E[X_{i\alpha} - E(X_i)] = E(X_{i\alpha}) - E(X_i) = 0$. Hence, since the $X_{i\alpha}$ are identically distributed in $\alpha$, we can write

$$\begin{aligned}\operatorname{cov}(\overline{X}_i,\overline{X}_j) &= (1/\beta)\, E\left\{[X_i-E(X_i)][X_j-E(X_j)]\right\} \\ &= (1/\beta)\,[E(X_iX_j) - E(X_i)\, E(X_j)]\end{aligned} \tag{1-6.9}$$

Substituting (1-3.3) into the above equation we obtain

$$\operatorname{cov}(\overline{X}_i,\overline{X}_j) = (1/\beta)\, N(N-1)\lambda\, p_i p_j \tag{1-6.10}$$

Now consider $\operatorname{cov}(\overline{W^2},\overline{X}_k)$. By the same reasoning as before

$$\operatorname{cov}(\overline{W^2},\overline{X}_k) = \operatorname{cov}\left\{(1/\beta)\sum_{\alpha=1}^{\beta}\left(\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} X_{i\alpha}X_{j\alpha}\right),\ (1/\beta)\sum_{k=1}^{\beta} X_{k\alpha}\right\}$$

$$= (1/\beta)\left\{\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} E(X_kX_iX_j) - E(X_k)\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} E(X_iX_j)\right\}$$

$$= (1/\beta)\left\{E(X_k^{\,3}) + 2\sum_{j\neq k} E(X_k^{\,2}X_j) + \sum_{i\neq k}\ \sum_{j\neq k,i} E(X_kX_iX_j)\right.$$

$$\left. + \sum_{i\neq k} E(X_kX_i^{\,2}) - E(X_k)\left[\sum_{i=1}^{n-1} E(X_i^{\,2}) + \sum_{i=1}^{n-1}\ \sum_{j\neq i} E(X_iX_j)\right]\right\}$$

(1-6.11)

By substituting (1-3.3) and (1-3.4) into the above expression, we obtain an equation in terms of the parameters which ultimately reduces to

$$\operatorname{cov}(\overline{W^2}, \overline{X}_k) = (1/\beta)N\lambda p_k \Big\{ G_2(1-p_n)^2 + 3G_1(1-p_n) - N\lambda(1-p_n)[G_1(1-p_n) + 1] + 1 \Big\} \tag{1-6.12}$$

The steps between (1-6.11) and (1-6.12) are outlined in Appendix 1F. For simplicity, let us define

$$H_1 = G_2(1-p_n)^2 + 3G_1(1-p_n) - N\lambda(1-p_n)[G_1(1-p_n) + 1] + 1 \tag{1-6.13}$$

Then $$\operatorname{cov}(\overline{W^2}, \overline{X}_k) = (1/\beta)\, N\lambda p_k H_1 \tag{1-6.14}$$

Finally consider $\operatorname{var} \overline{W^2}$. Using again the arguments of identity and independence, we obtain

$$\operatorname{var} \overline{W^2} = \operatorname{var}\Big\{ (1/\beta) \sum_{\alpha=1}^{\beta} \Big( \sum_{i=1}^{n-1} X_{i\alpha} \Big)^2 \Big\}$$

$$= (1/\beta) \Big\{ E\Big( \sum_{i=1}^{n-1} X_i \Big)^4 - E^2\Big( \sum_{i=1}^{n-1} X_i \Big)^2 \Big\}$$

$$= (1/\beta) \Big\{ \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sum_{k=1}^{n-1} \sum_{m=1}^{n-1} E(X_i X_j X_k X_m) - \Big[ \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} E(X_i X_j) \Big]^2 \Big\}$$

$$= (1/\beta) \Big\{ \sum_{i=1}^{n-1} \sum_{j\neq i} \sum_{k\neq i,j} \sum_{m\neq i,j,k} E(X_i X_j X_k X_m)$$

$$+ 6 \sum_{i=1}^{n-1} \sum_{j\neq i} \sum_{k\neq i,j} E(X_i^{\,2} X_j X_k) + 3 \sum_{i=1}^{n-1} \sum_{j\neq i} E(X_i^{\,2} X_j^{\,2})$$

$$+ 4\sum_{i=1}^{n-1}\sum_{j\neq i} E(X_i^{3}X_j) + \sum_{i=1}^{n-1} E(X_i^{4}) - [\sum_{i=1}^{n-1} E(X_i^{2})$$

$$\left. + \sum_{j=1}^{n-1}\sum_{k\neq j} E(X_i X_k)]^2 \right\} \qquad (1\text{-}6.15)$$

Now substitute (1-3.3) and (1-3.4) into the above equation and simplify as is done in Appendix 1G. Then

$$\text{Var}\ \overline{W^2} = (1/\beta)N\lambda(1-p_n)\Big\{G_3(1-p_n)^3 + 6G_2(1-p_n)^2 + 7G_1(1-p_n)$$

$$\left. + 1 - N\lambda(1-p_n)[G_1(1-p_n) + 1]^2\right\} \qquad (1\text{-}6.16)$$

For simplicity we may define

$$H_2 = G_3(1-p_n)^3 + 6G_2(1-p_n)^2 + 7G_1(1-p_n) + 1$$

$$- N\lambda(1-p_n)[G_1(1-p_n) + 1]^2 \qquad (1\text{-}6.17)$$

Then $\quad \text{var}\ \overline{W^2} = (1/\beta)N\lambda(1-p_n)H_2 \qquad (1\text{-}6.18)$

Thus, substituting (1-6.8), (1-6.10), (1-6.14), and (1-6.18) into (1-6.4),

$$M = \frac{1}{\beta}\begin{pmatrix} N\lambda(1-p_n)H_2 & N\lambda p_1 H_1 & \cdots & N\lambda p_{n-1}H_1 \\ N\lambda p_1 H_1 & N\lambda p_1[p_1(N-1)+1] & \cdots & N(N-1)\lambda p_1 p_{n-1} \\ N_\lambda p_2 H_1 & N(N-1)\lambda p_1 p_2 & & N(N-1)\lambda p_2 p_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ N\lambda p_{n-1}H_1 & N(N-1)\lambda p_1 p_{n-1} & \cdots & N\lambda p_{n-1}[p_{n-1}(N-1)+1] \end{pmatrix}$$

(1-6.19)

The determinant of this matrix is calculated explicitly in Appendix 1H. We need only set $N' = N-1$ and $R = N\lambda$ in the matrix in the appendix and we have (1-6.19). Then

$$\det M = (N\lambda/\beta)^n(1-p_n)\left(\prod_{i=1}^{n-1} p_i\right)[H_2(N-Np_n+p_n)-H_1{}^2] \qquad (1\text{-}6.20)$$

## C. Determination of the Jacobian, det $\phi$

To evaluate $\det\ \phi$ as defined in (1-6.6) we must evaluate the determinant of the following matrix -

$$\phi = \begin{pmatrix} \frac{\partial\lambda}{\partial E(W^2)} & \frac{\partial\lambda}{\partial E(X_1)} & & \cdots & \frac{\partial\lambda}{\partial E(X_{n-1})} \\ \frac{\partial p_1}{\partial E(W^2)} & \frac{\partial p_1}{\partial E(X_1)} & \frac{\partial p_1}{\partial E(X_2)} & \cdots & \frac{\partial p_1}{\partial E(X_{n-1})} \\ \vdots & \vdots & & \ddots & \vdots \\ \frac{\partial p_{n-1}}{\partial E(W^2)} & \frac{\partial p_{n-1}}{\partial E(X_1)} & & \cdots & \frac{\partial p_{n-1}}{\partial E(X_{n-1})} \end{pmatrix} \qquad (1\text{-}6.21)$$

To find the above partial derivatives we appeal to (1-3.9) and (1-3.10), noting that

$$E(W) = \sum_{i=1}^{n-1} E(X_i).$$

After differentiating we obtain

$$\frac{\partial\lambda}{\partial E(W^2)} = -\frac{N-1}{N}\,\frac{E^2(W)}{[E(W^2)-E^2(W)-E(W)]^2}$$

$$\frac{\partial\lambda}{\partial E(X_i)} = \frac{N-1}{N} \; \frac{E(W)[2E(W^2)-E(W)]}{[E(W^2)-E^2(W)-E(W)]^2}$$

$$\frac{\partial p_i}{\partial E(W^2)} = \frac{E(X_i)}{(N-1)E^2(W)}$$

$$\frac{\partial p_i}{\partial E(X_j)} = \frac{E(X_i)}{N-1} \cdot \frac{E(W)-2E(W^2)}{E^3(W)} \qquad \text{if} \quad i \neq j$$

$$\frac{\partial p_i}{\partial E(X_i)} = \frac{E(X_i)}{N-1} \cdot \frac{E(W)-2E(W^2)}{E^3(W)} + \frac{E(W^2)-E^2(W)-E(W)}{(N-1)E^2(W)}$$

It will be more convenient if we are able to express the entries of $\phi$ in terms of the parameters. To this end we may substitute for the expectations in the above set of equations using (1-3.3), (1-3.6), and (1-3.7). After simplification we will find

$$\left.\begin{aligned}
\frac{\partial\lambda}{\partial E(W^2)} &= -\frac{1}{N(N-1)(1-p_n)^2} \\
\frac{\partial\lambda}{\partial E(X_i)} &= \frac{2[p_n+N(\lambda+1)(1-p_n)]-1}{N(N-1)(1-p_n)^2} \\
\frac{\partial p_i}{\partial E(W^2)} &= \frac{p_i}{N(N-1)\lambda(1-p_n)^2} \\
\frac{\partial p_i}{\partial E(X_j)} &= \frac{p_i\left\{1-2[p_n+N(\lambda+1)(1-p_n)]\right\}}{N(N-1)\lambda(1-p_n)^2} \quad \text{if} \quad i \neq j \\
\frac{\partial p_i}{\partial E(X_i)} &= \frac{p_i\left\{1-2[p_n+N(\lambda+1)(1-p_n)]\right\}}{N(N-1)\lambda(1-p_n)^2} + \frac{1}{N\lambda}
\end{aligned}\right\} \qquad (1\text{-}6.22)$$

The substitution of (1-6.22) into (1-6.21) gives an explicit expression for $\phi$. Its determinant may be found by multiplying the first row by $p_i/\lambda$ and adding it to the $(i+1)^{st}$ row for $i = 1, 2, \ldots, n-1$. This will give an upper triangular matrix which may be expanded by the first column to give

$$\det \phi = - \frac{1}{(N\lambda)^{n-1} N(N-1)(1-p_n)^2} \tag{1-6.23}$$

If we now substitute (1-6.20) and (1-6.23) into (1-6.5),

$$\det \bar{\Omega} = \frac{[H_2(N-Np_n+p_n)-H_1{}^2]}{N^2(N-1)^2(1-p_n)^3(N\lambda)^{n-2}\beta^n} \prod_{i=1}^{n-1} p_i$$

At this point we can find the efficiency by substituting the above equation along with (1-5.21) into (1-6.2). Hence

$$\text{Eff} = \frac{N^2(N-1)^2(1-p_n)^3}{[B_{\lambda\lambda}N\lambda+(1-p_n)(B_{\lambda\lambda}B_{pp}-B_{\lambda p}{}^2)][H_2(N-Np_n+p_n)-H_1{}^2]} \tag{1-6.24}$$

## 1-7. Sample Zero Frequency and Unit Sample Frequency Estimators

### A. Sample Zero Frequency and First Moments

Sample zero frequency estimation is useful if the zero sample, (i.e. $\vec{X} = \vec{0}$), occurs quite frequently. From (1-2.3), if we set $\vec{X}=\vec{0}$

$$P_{\vec{X}}(\vec{0}) = \exp[\lambda(p_n{}^N-1)] \tag{1-7.1}$$

Let us define $F(\vec{a})$ to be the frequency with which

$\vec{X} = \vec{a} = (a_1, a_2, \ldots, a_{n-1})$ occurs in $\beta$ observations. Consider the estimator $(1/\beta)\, F(\vec{0})$ for $P_{\vec{X}}(\vec{0})$

$$E[\frac{1}{\beta} F(\vec{0})] = P_{\vec{X}}(\vec{0}) = \exp[\lambda(p_n{}^N - 1)] \qquad (1\text{-}7.2)$$

Hence $(1/\beta)\, F(\vec{0})$ is an unbiased estimator for $P_{\vec{X}}(\vec{0})$.

We may obtain the sample zero frequency estimators $\tilde{\lambda}$ and $\tilde{p}_i$ for $\lambda$ and $p_i$ by using the moment estimators for the first moments given by (1-3.3) and (1-6.3) ff. together with the estimator just defined in (1-7.2) to obtain the equations

$$\left.\begin{aligned} \bar{X}_i &= N\tilde{\lambda}\tilde{p}_i \\ (1/\beta)\, F(\vec{0}) &= \exp[\tilde{\lambda}(\tilde{p}_n{}^N - 1)] \end{aligned}\right\} \qquad (1\text{-}7.3)$$

To solve for $\tilde{\lambda}$, $\tilde{p}_i$, let us first add the top equation of (1-7.3) for $i=1, 2, \ldots, n-1$.

$$N\tilde{\lambda}(1-\tilde{p}_n) = \sum_{i=1}^{n-1} \bar{X}_i$$

Solving for $\tilde{p}_n$ and substituting into the bottom equation of (1-7.3),

$$\log \frac{F(\vec{0})}{\beta} = \tilde{\lambda}[(1 - \frac{1}{N\tilde{\lambda}} \sum_{i=1}^{n-1} \bar{X}_i)^N - 1] \qquad (1\text{-}7.4)$$

We can use a numerical method to find $\tilde{\lambda}$, and then from (1-7.2),

$$\tilde{p}_i = \bar{X}_i / N\tilde{\lambda} \qquad (1\text{-}7.5)$$

## B. Unit Sample Frequency Estimation

If the unit samples (i.e. $\vec{X} = \vec{e}_k$, k=1, 2, ..., n-1). occur fairly frequently it may be advantageous to use this estimator. From (1-2.1) we can see that

$$P_X(\vec{e}_k) = \left.\frac{\partial g(\vec{s})}{\partial s_k}\right|_{\vec{s}=\vec{0}} = \{\exp\,[\lambda(p_n{}^N-1)]\}N\; p_n{}^{N-1} p_k \lambda \qquad (1\text{-}7.6)$$

Consider the estimator $(1/\beta)\,F(\vec{e}_k)$ for $P_X(\vec{e}_k)$. We notice that $E[\frac{1}{\beta}\,F(\vec{e}_k)] = P_X(\vec{e}_k)$. Hence the estimator is unbiased. Thus we may solve the equations

$$(1/\beta)\,F(\vec{e}_k) = \{\exp[\lambda(p_n{}^N-1)]\}Np_n{}^{N-1}p_k\lambda \quad k=1,\ 2,\ \dots,\ n-1 \qquad (1\text{-}7.7)$$

along with (1-7.3) for $p_1$ and $\lambda$ to obtain their unit sample estimators $\check{p}_1$ and $\check{\lambda}$.

To solve these equations let us divide (1-7.7) by (1-7.7) with k=1. Then

$$\frac{F(\vec{e}_k)}{F(\vec{e}_1)} = \frac{\check{p}_k}{\check{p}_1} \qquad (1\text{-}7.8)$$

Dividing (1-7.7) with k=1 by (1-7.3)

$$F(\vec{e}_1)/F(\vec{0}) = N(1 - \sum_{k=1}^{n-1} \check{p}_k)^{N-1}\; \check{p}_1\check{\lambda}$$

If we substitute for $\check{p}_k$ from (1-7.8) and solve for $\check{\lambda}$,

$$1/\check{\lambda} = \frac{\check{p}_1 NF(\vec{0})}{F(\vec{e}_1)}\,[1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^{N-1} \qquad (1\text{-}7.9)$$

From (1-7.3) and the fact that the sum of the $p_i$ is one, we obtain

$$\log \frac{F(\vec{0})}{\beta} = \check{\lambda}[(1 - \sum_{k=1}^{n-1} \check{p}_k)^N - 1]$$

Upon substitution for $\check{p}_k$ from (1-7.8) and division by $\lambda$, we find

$$\frac{1}{\check{\lambda}} \log \frac{F(\vec{0})}{\beta} + 1 = [1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^N \qquad (1\text{-}7.10)$$

If we now substitute for $1/\check{\lambda}$ from (1-7.9), equation (1-7.10) becomes

$$\frac{\check{p}_1 NF(\vec{0})}{F(\vec{e}_1)} \; [1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^{N-1} \log \frac{F(\vec{0})}{\beta} + 1$$

$$= [1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^N$$

This may be rewritten as

$$\left[1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)\right]^{N-1} \left\{ \frac{\check{p}_1}{F(\vec{e}_1)} [NF(\vec{0}) \log \frac{F(\vec{0})}{\beta} + \sum_{k=1}^{n-1} F(\vec{e}_k)] - 1 \right\}$$

$$+ 1 = 0 \qquad (1\text{-}7.11)$$

We can now use a numerical method to find $\check{p}_1$. We can then calculate $\check{\lambda}$ from (1-7.9) and finally $\check{p}_i$, $i=2, 3, \ldots, n-1$ from (1-7.8).

# CHAPTER II

# THE POISSON-NEGATIVE MULTINOMIAL DISTRIBUTION

## 2-1. A Biological Model

This distribution arises from a model very similar to the one given in § 1.1 for the Poisson-multinomial distribution. The only variations are the following.

Let N represent, instead of the total number of eggs laid in each batch, the mean number of eggs that do not hatch in each batch. Let Z be a random variable denoting the number of batches of eggs laid in a particular quadrat and assume the egg laying stops as soon as the $(Nz)^{th}$ egg is laid that will not hatch. Hence if we define $p_1, p_2, \ldots, p_{n-1}$ the same as in § 1.1 and $p_n$ by (1-1.2), then

$$\left.\begin{aligned} P_{\vec{x}}(\vec{x}\,|Z = z) &= \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!x_2!\ldots x_{n-1}!(Nz-1)!}\, p_n^{Nz} \prod_{i=1}^{n-1} p_i^{x_i} \\ &\qquad\qquad\qquad\qquad \text{if} \quad z>0 \\ P_{\vec{x}}(\vec{x}\,|Z = 0) &= \begin{cases} 1 & \text{if} \quad x=0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}\right\} \qquad (2.1.1)$$

If Z again has a Poisson $(\lambda)$ distribution as in § 1.1, then

$$\begin{aligned} P_{\vec{x}}(\vec{x}) &= \sum_{z=0}^{\infty} P_{\vec{x}}(\vec{x}\,|Z = z)\, P(Z=z) \\ &= e^{-\lambda} \sum_{z=0}^{\infty} (\lambda^z/z!)\, \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!}\; p_n^{Nz} \prod_{i=1}^{n-1} p_i^{x_i} \end{aligned} \qquad (2\text{-}1.2)$$

where we adopt the convention that

$$\frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!} = \begin{cases} 1 & \text{if } \vec{x} = \vec{0},\ z = 0 \\ 0 & \text{if } \vec{x} \neq \vec{0},\ z = 0 \end{cases} \tag{2-1.3}$$

## 2-2. Probability Generating Function and Recursion Formula for Probabilities

The probability generating function is defined by

$$g^*(\vec{s}) = E(s_1^{x_1} s_2^{x_2} \ldots s_{n-1}^{x_{n-1}}) \tag{2-2.1}$$

Thus from (2-1.2), (2-1.3), and (2-2.1)

$$g^*(\vec{s}) = \sum_{x_1=0}^{\infty} \ldots \sum_{x_{n-1}=0}^{\infty} e^{-\lambda}\left\{ \sum_{z=1}^{\infty} (\lambda^z/z!) \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!} p_n^{Nz} \prod_{i=1}^{n-1} (s_i p_i)^{x_i} + \delta_{\vec{x},\vec{0}} \right\}$$

where $\delta_{\vec{x},\vec{0}} = \begin{cases} 1 & \text{if } \vec{x} = \vec{0} \\ 0 & \text{if } \vec{x} \neq \vec{0} \end{cases}$

Let us sum the terms in $\delta_{\vec{x},\vec{0}}$ separately and rearrange the order of the sums of the other terms. Then

$$g^*(\vec{s}) = e^{-\lambda} \sum_{z=1}^{\infty} (\lambda^z/z!) p_n^{Nz} \sum_{x_1=0}^{\infty} \ldots \sum_{x_{n-1}=0}^{\infty} \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!} \prod_{i=1}^{n-1} (s_i p_i)^{x_i} + e^{-\lambda} \tag{2-2.2}$$

To evaulate the above expression we use the identity

$$\frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!} = \prod_{j=1}^{n-1} \binom{\sum_{k=j}^{n-1} x_k+Nz-1}{x_j} \qquad (2\text{-}2.3)$$

From Feller [1950], page 61, (12.4), we have the identity

$$\binom{-a}{x} = \binom{x+a-1}{x}(-1)^x \qquad (2\text{-}2.4)$$

Hence $\sum_{x=0}^{\infty} \binom{-a}{x}(-p)^x = \sum_{x=0}^{\infty} \binom{x+a-1}{x}(-1)^x(-p)^x = (1-p)^{-a}$ (2-2.5)

The last equality is true because the middle term is simply the binomial expansion of $(1-p)^{-a}$. We may use (2-2.4) to replace the combinatorial expression on the right side of (2-2.3) and then use the result to replace the factorials in (2-2.2). We obtain an expression involving negative binomial coefficients

$$g^*(\vec{s}) = e^{-\lambda} \sum_{z=1}^{\infty} (\lambda^z/z!)p_n^{Nz} \sum_{x_{n-1}=0}^{\infty} \cdots \sum_{x_1=0}^{\infty} \prod_{j=1}^{n-1} \binom{-\sum_{k=j+1}^{n-1} x_k-Nz}{x_j} \qquad (2\text{-}2.5\tfrac{1}{2})$$

$$\cdot \prod_{i=1}^{n-1} (-s_i p_i)^{x_i} + e^{-\lambda}$$

After carrying out the indicated summations using (2-2.5) as is done in Appendix 2A,

$$g^*(\vec{s}) = e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} p_n^{Nz} \left(1 - \sum_{i=1}^{n-1} s_i p_i\right)^{-Nz}$$

and hence

$$g^*(\vec{s}) = \exp \left\{ \lambda \left( \frac{p_n}{1 - \sum_{i=1}^{n-1} s_i p_i} \right)^N - \lambda \right\} \tag{2-2.6}$$

Let us consider the following change of variables

$$\left.\begin{aligned} p_n &= 1/b_n \\ p_i &= -b_i/b_n \quad i = 1, \ldots, n-1 \\ N &= -V > 0 \end{aligned}\right\} \tag{2-2.7}$$

Then (2-2.6) becomes

$$g^*(\vec{s}) = \exp \left\{ \lambda [b_n + \sum_{i=1}^{n-1} s_i b_i ]^V - \lambda \right\} \tag{2-2.8}$$

Notice this formula is exactly the same as (1-2.1) where $V$ corresponds to $N$, and $b$ to $p$. Hence whatever we say about $\lambda$ and $p_i$ in the Poisson-multinomial distribution, we may say the same thing about $\lambda$ and $b_i$ respectively in the Poisson-negative multinomial distribution by virtue of (2-2.8). In many results obtained in this chapter this fact will greatly reduce the length of calculations, while in others, especially those involving derivatives with respect to $p_i$, it is better to calculate directly.

To obtain an expression for the probabilities we must differentiate $g^*(\vec{s})$ an appropriate number of times. Starting with (2-2.8) we can exactly follow the procedure in Appendix 1A with the obvious change in symbols and obtain (1A-4) which will be written as

$$P_{\vec{x}}(\vec{x}+\vec{e}_k) = \frac{\lambda b_k b_n^{V-1}}{x_k+1} \sum_{y_1=0}^{x_1} \cdots \sum_{y_{n-1}=0}^{x_{n-1}} V(V-1) \ldots [V-\sum_{i=1}^{n-1}(x_i-y_i)]$$
$$\left[\prod_{i=1}^{n-1} \left(\frac{b_i}{b_n}\right)^{x_i-y_i} \cdot \frac{1}{(x_i-y_i)!}\right] P_{\vec{x}}(\vec{y})$$

From (2-2.7) it is clear that

$$b_i \;=\; -\,p_i/p_n \tag{2-2.9}$$

Upon substitution for the b's and $V$ in the above equation from (2-2.7) and (2-2.9), we obtain

$$P_{\vec{x}}(\vec{x}+\vec{e}_k) = -\,\frac{\lambda(p_k/p_n)(1/p_n)^{-N-1}}{x_k+1} \sum_{y_1=0}^{x_1} \cdots \sum_{y_{n-1}=0}^{x_{n-1}} (-N)(-N-1) \ldots$$

$$[-N-\sum_{i=1}^{n-1}(x_i-y_i)][\prod_{i=1}^{n-1}(-p_i)^{x_i-y_i} \frac{1}{(x_i-y_i)!}] P_{\vec{x}}(\vec{y})$$

If we now factor the "minus one" out of each term immediately following the multiple sum and each $p_i$, and note that $N$ is an integer, we obtain the second equation of (2-2.10). The first comes from (2-2.6).

$$\left.\begin{aligned}
&P_{\vec{x}}(\vec{0}) = g^*(\vec{0}) = e^{\lambda(p_n{}^N-1)} \\
&P_{\vec{x}}(\vec{x}+\vec{e}_k) = \frac{\lambda p_k p_n{}^N}{x_k+1} \sum_{y_1=o}^{x_1} \cdots \sum_{y_{n-1}=o}^{x_{n-1}} \frac{[N+\sum_{i=1}^{n-1}(x_i-y_i)]!}{(N-1)!} \\
&\prod_{i=1}^{n-1} [p_i{}^{x_i-y_i}/(x_i-y_i)!]\ P_{\vec{x}}(\vec{y})
\end{aligned}\right\} \qquad (2\text{-}2.10)$$

## 2-3. Estimation of Parameters by the Method of Moments

To obtain the moments of this distribution, we use the same method as in § 1-3. From (2-2.8) we may form the cumulant generating function

$$c(\vec{s}) = \lambda\left\{[b_n + \sum_{i=1}^{n-1} s_i b_i]^V - 1\right\} \qquad (2\text{-}3.1)$$

By following the calculations of Appendix 1B, but replacing $N$ with $V$, and $p_i$ with $b_i$, $i = 1, 2, \ldots, n$, we will get (1-3.2), (1-3.3), and (1-3.4) with the above replacement. Let us call these modified equations (1-3.2)', (1-3.3)', and (1-3.4)'. If we apply the transformation given by (2-2.8) and (2-2.9) so as to express (1-3.2)', (1-3.3)', and (1-3.4)' in terms of $N$ and $p_i$ and then define

$$\left.\begin{aligned}
G_1^* &= N(\lambda+1) + 1 \\
G_2^* &= N^2(\lambda^2+3\lambda+1) + 3N(\lambda+1) + 2 \\
G_3^* &= N^3(\lambda^3+6\lambda^2+7\lambda+1) + 6N^2(\lambda^2+3\lambda+1) \\
&\qquad + 11N(\lambda+1) + 6
\end{aligned}\right\} \qquad (2\text{-}3.2)$$

the moments will be

$$
\left.\begin{aligned}
E(X_i) &= N\lambda p_i/p_n \\
E(X_i^2) &= \frac{N\lambda p_i}{p_n}\left(\frac{p_i}{p_n} G_1^* + 1\right) \\
&= \frac{N\lambda p_i}{p_n}\left[\frac{p_i N}{p_n}(\lambda+1) + \frac{p_i}{p_n} + 1\right] \\
E(X_i X_j) &= \frac{N\lambda p_i p_j}{p_n^2} G_1^* \\
&= \frac{N\lambda p_i p_j}{p_n^2}[N(\lambda+1) + 1]
\end{aligned}\right\} \qquad (2\text{-}3.3)
$$

$$
\left.\begin{aligned}
E(X_i^3) &= \frac{N\lambda p_i}{p_n}\left(\frac{p_i^2}{p_n^2} G_2^* + 3\frac{p_i}{p_n} G_1^* + 1\right) \\
E(X_i^2 X_j) &= \frac{N\lambda p_i p_j}{p_n^2}\left(\frac{p_i}{p_n} G_2^* + G_1^*\right) \\
E(X_i X_j X_k) &= \frac{N\lambda p_i p_j p_k}{p_n^3} G_2^* \\
E(X_i^4) &= \frac{N\lambda p_i}{p_n}\left(\frac{p_i^3}{p_n^3} G_3^* + 6\frac{p_i^2}{p_n^2} G_2^* + 7\frac{p_i}{p_n} G_1^* + 1\right) \\
E(X_i^3 X_j) &= \frac{N\lambda p_i p_j}{p_n^2}\left(\frac{p_i^2}{p_n^2} G_3^* + 3\frac{p_i}{p_n} G_2^* + G_1^*\right) \\
E(X_i^2 X_j^2) &= \frac{N\lambda p_i p_j}{p_n^2}\left[\frac{p_i p_j}{p_n^2} G_3^* + \frac{(p_i+p_j)}{p_n} G_2^* + G_1^*\right] \\
E(X_i^2 X_j X_k) &= \frac{N\lambda p_i p_j p_k}{p_n^3}\left(\frac{p_i}{p_n} G_3^* + G_2^*\right) \\
E(X_i X_j X_k X_m) &= \frac{N\lambda p_i p_j p_k p_m}{p_n^4} G_3^*
\end{aligned}\right\} \qquad (2\text{-}3.4)
$$

Let us now define the random variable, $W$, by

$$W = \sum_{i=1}^{n-1} X_i \tag{2-3.5}$$

Then $$E(W) = \sum_{i=1}^{n-1} E(X_i) = \frac{N\lambda(1-p_n)}{p_n} \tag{2-3.6}$$

and $$E(W^2) = E\left(\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} X_i X_j\right)$$

$$= \sum_{i=1}^{n-1} E(X_i^{\,2}) + \sum_{i=1}^{n-1}\sum_{j\neq i} E(X_i X_j)$$

If we substitute for the expectations from (2-3.2) and sum,

$$E(W^2) = \frac{N\lambda}{p_n}(1-p_n)\left\{1 + [N(\lambda+1)+1]\frac{(1-p_n)}{p_n}\right\} \tag{2-3.7}$$

Then, the substitution of (2-3.6) into this yields

$$E(W^2) = E(W)\left\{E(W) + (N+1)\frac{(1-p_n)}{p_n}+1\right\} \tag{2-3.8}$$

Consequently we can solve (2-3.6) for $p_n$, substitute into (2-3.8), and solve for $\lambda$ to obtain

$$\lambda = \frac{N+1}{N}\;\frac{E^2(W)}{E(W^2)-E^2(W)-E(W)} \tag{2-3.9}$$

From (2-3.3), $p_i = p_n E(X_i)/N\lambda$ and from (2-3.6) $p_n = N\lambda/[E(W)+N\lambda]$. Thus we may eliminate $p_n$ from these two equations and substitute for $\lambda$ from (2-3.9). After simplifying

$$\left.\begin{aligned} p_i &= \frac{E(X_i)}{E(W)+N\lambda} = \frac{E(X_i)[E(W^2)-E^2(W)-E(W)]}{E(W)[E(W^2)-E^2(W)+NE(W)]} \\ p_n &= N\lambda/[E(W) + N\lambda] \end{aligned}\right\} \qquad (2\text{-}3.10)$$

To obtain the moment estimators $\lambda^*$ and $p_i^*$ of $\lambda$ and $p_i$ respectively, we simply replace the population moments by their corresponding sample moments. Using the notation defined in (1-3.11), (2-3.9) and (2-3.10) yield

$$\lambda^* = \frac{x..^2}{(1/\beta)\sum_{\alpha=1}^{\beta} w_\alpha^2 - x..^2 - x..} \cdot \frac{N+1}{N}$$

$$p_i^* = x_{i.}/[x.. + N\lambda^*]$$

$$p_n^* = 1 - \sum_{i=1}^{n-1} p_i^* = N\lambda^*/[x.. + N\lambda^*]$$

## 2-4 Maximum Likelihood Estimators

In this section we will see that the derivation of the maximum likelihood estimators closely parallels that for the Poisson-multinomial distribution. Let us define $\beta$, $\vec{x}_\alpha$, and $x_{i\alpha}$ as at the end of § 1.3. Then we may define the likelihood function, L, as in (1-4.1) and obtain (1-4.3). For convenience we will record this set of equations again

$$\frac{\partial}{\partial\lambda} \log L = \sum_{\alpha=1}^{\beta} \frac{1}{P_{\vec{x}}(\vec{x}_\alpha)} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}_\alpha) = 0$$

$$\frac{\partial}{\partial p_i} \log L = \sum_{\alpha=1}^{\beta} \frac{1}{P_{\vec{x}}(\vec{x}_\alpha)} \frac{\partial}{\partial p_i} P_{\vec{x}}(\vec{x}_\alpha) = 0 \qquad (2\text{-}4.1)$$

Here, of course, $P_{\vec{x}}$ denotes the Poisson-negative multinomial rather than the Poisson-multinomial density.

It is possible to find the derivatives of $P_{\vec{x}}(\vec{x})$ by differentiating (2-1.2)

$$\frac{\partial}{\partial p_i} P_{\vec{x}}(\vec{x}) = e^{-\lambda} \sum_{z=0}^{\infty} \frac{\lambda^z}{z!} \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!} p_n^{Nz} \cdot \prod_{j=1}^{n-1} p_j^{x_j} (x_i/p_i - Nz/p_n)$$

Consider the following identity

$$z = (1/N)[(Nz + \sum_{j=1}^{n-1} x_j) - \sum_{j=1}^{n-1} x_j] \qquad (2\text{-}4.2)$$

If we use this identity to substitute for the last $z$ in the above equation, then

$$\frac{\partial}{\partial p_i} P_{\vec{x}}(\vec{x}) = (x_i/p_i) P_{\vec{x}}(\vec{x}) - e^{-\lambda} \sum_{z=0}^{\infty} \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!} \cdot p_n^{Nz-1} \prod_{j=1}^{n-1} p_j^{x_j} (Nz + \sum_{k=1}^{n-1} x_k) + (1/p_n) \sum_{k=1}^{n-1} x_k P_{\vec{x}}(\vec{x})$$

$$= [x_i/p_i + (1/p_n) \sum_{k=1}^{n-1} x_k] P_{\vec{x}}(\vec{x})$$

$$- \frac{(x_i+1)e^{-\lambda}}{p_i p_n} \sum_{z=o}^{\infty} \frac{\lambda^z}{z!} \frac{(x_1+\ldots+x_{n-1}+Nz)!}{x_1!\ldots x_{n-1}!(x_i+1)(Nz-1)!} p_n^{Nz} p_i \prod_{j=1}^{n-1} p_j^{x_j}$$

Clearly this reduces to

$$\frac{\partial}{\partial p_i} P_{\vec{x}}(\vec{x}) = [x_i/p_i + (1/p_n)\sum_{k=1}^{n-1} x_k]P_{\vec{x}}(\vec{x}) - \frac{x_i+1}{p_n p_i} P_{\vec{x}}(\vec{x}+\vec{e}_i) \tag{2-4.3}$$

Now let us differentiate (2-1.2) with respect to $\lambda$.

$$\frac{\partial}{\partial \lambda} P_{\vec{x}}(\vec{x}) = e^{-\lambda} \sum_{z=o}^{\infty} \frac{z\lambda^{z-1}}{z!} \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!} p_n^{Nz} \prod_{j=1}^{n-1} p_j^{x_j}$$

$$- P_x(x)$$

The use of (2-4.2) to substitute for the first $z$ in the numerator of the above expression results in

$$= (e^{-\lambda}/N\lambda) \sum_{z=o}^{\infty} \frac{\lambda^z}{z!} [(Nz + \sum_{i=1}^{n-1} x_i) - \sum_{i=1}^{n-1} x_i] \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!}$$

$$p_n^{Nz} \prod_{j=1}^{n-1} p_j^{x_j} - P_{\vec{x}}(\vec{x})$$

Therefore we simplify to

$$\frac{\partial}{\partial \lambda} P_{\vec{x}}(\vec{x}) = \frac{x_m+1}{N\lambda p_m} P_{\vec{x}}(\vec{x}+\vec{e}_m) - [(1/N\lambda)\sum_{j=1}^{n-1} x_j + 1] P_{\vec{x}}(\vec{x})$$

$$m=1,\ 2,\ \ldots,\ n-1 \tag{2-4.4}$$

Equations (2-4.3) and (2-4.4) will hold for each observation, i.e. when $\vec{x} = \vec{x}_\alpha$, $x_m = x_{m\alpha}$, $\alpha = 1, 2, \ldots, \beta$. With this in mind we may substitute (2-4.4) into the top equation of (2-4.1) to obtain

$$\sum_{\alpha=1}^{\beta} \frac{x_{m\alpha}+1}{N\hat{\lambda}\hat{p}_m} \frac{P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m)}{P_{\vec{x}}(\vec{x}_\alpha)} - \sum_{\alpha=1}^{\beta} [1 + (1/N\hat{\lambda}) \sum_{k=1}^{n-1} x_{k\alpha}] = 0 \qquad (2\text{-}4.5)$$

Using the same idea we substitute (2-4.3) into the bottom equation of (2-4.1)

$$\sum_{\alpha=1}^{\beta} [x_{i\alpha}/\hat{p}_i + (1/\hat{p}_n) \sum_{k=1}^{n-1} x_{k\alpha}] - (N\hat{\lambda}/\hat{p}_n) \sum_{\alpha=1}^{\beta} \frac{x_{i\alpha}+1}{N\hat{\lambda}\hat{p}_i} \frac{P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_i)}{P_{\vec{x}}(\vec{x}_\alpha)} = 0$$

It is possible to replace the last sum by substituting from (2-4.5) and then simplifying the notation by means of (1-3.11) to get $x_{i.}/\hat{p}_i + x_{..}/\hat{p}_n - (N\hat{\lambda}/\hat{p}_n)(\beta+x_{..}/N\hat{\lambda}) = 0$. Solving for $\hat{p}_i$, this becomes

$$\hat{p}_i = \hat{p}_n x_{i.}/N\hat{\lambda} \qquad i = 1, 2, \ldots, n-1 \qquad (2\text{-}4.6)$$

If we now multiply this equation by $N\hat{\lambda}$ and add for $i=1, 2, \ldots, n-1$, we will obtain

$$N\hat{\lambda} \sum_{i=1}^{n-1} \hat{p}_i = N\hat{\lambda}(1-\hat{p}_n) = \hat{p}_n \sum_{i=1}^{n-1} x_{i.} = \hat{p}_n x_{..}$$

Hence $\qquad \hat{p}_n = N\hat{\lambda}/(N\hat{\lambda} + x_{..}) \qquad (2\text{-}4.7)$

and upon substitution of this into (2-4.6) we have the estimator for $p_i$.

$$\hat{p}_i = x_{i.}/(N\hat{\lambda} + x_{..}) \qquad i = 1, 2, \ldots, n-1 \qquad (2\text{-}4.8)$$

It still remains to find the estimator for $\lambda$, that is, $\hat{\lambda}$. As is the case with the Poisson-multinomial distribution, it is almost impossible to solve for $\hat{\lambda}$ directly, and hence we must use a numerical method. The following calculation is based on Newton's formula which is given by (1-4.12). Writing it again,

$$\hat{\lambda}_{n+1} = \hat{\lambda}_n - [f(\hat{\lambda}_n) \;/\; D_\lambda f(\hat{\lambda}_n)] \tag{2-4.9}$$

where $\quad f(\hat{\lambda}) = 0. \quad$ (2-4.10)

If we use (2-4.8) to substitute for $\hat{p}_m$ in (2-4.5), the latter equation reduces to an expression which is a candidate for $f(\hat{\lambda})$ since it satisfies (1-4.13), i.e.

$$f(\hat{\lambda}) = \sum_{\alpha=1}^{\beta} \frac{(x_{m\alpha}+1)}{x_{m.}} \; \frac{P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m)}{P_{\vec{x}}(\vec{x}_\alpha)} - \beta = 0 \tag{2-4.11}$$

The final step in our procedure is to find $D_{\hat{\lambda}} f(\hat{\lambda})$ for substitution into (2-4.9). Differentiating (2-4.10) with respect to $\hat{\lambda}$ gives

$$D_{\hat{\lambda}} f(\hat{\lambda}) = \sum_{\alpha=1}^{n-1} \frac{x_{m\alpha}+1}{P_{\vec{x}}^{2}(\vec{x}_\alpha)} \; [P_{\vec{x}}(\vec{x}_\alpha) \, D_{\hat{\lambda}} P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m)$$

$$- P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m) \, D_{\hat{\lambda}} \, P_{\vec{x}}(\vec{x}_\alpha)] \tag{2-4.12}$$

But we know

$$D_{\hat{\lambda}} \, P_{\vec{x}}(\vec{x}_\alpha) = \sum_{m=1}^{n-1} \frac{\partial}{\partial \hat{p}_m} \, P_{\vec{x}}(\vec{x}_\alpha) D_{\hat{\lambda}} \hat{p}_m + \frac{\partial}{\partial \hat{\lambda}} \, P_{\vec{x}}(\vec{x}_\alpha) \tag{2-4.13}$$

and from (2-4.7), $\quad D_{\hat{\lambda}} \hat{p}_i = - x_{i.}/(N\hat{\lambda} + x_{..})^2 \quad$ (2-4.14)

Now let us substitute for the derivatives in (2-4.12). Using (2-4.13) to substitute for the partials with respect to $\hat{p}_i$ and $\hat{\lambda}$ respectively

$$D_{\hat{\lambda}}\, P_{\vec{x}}(\vec{x}_\alpha) = \sum_{j=1}^{n-1} \left\{ [x_{j\alpha}/\hat{p}_j + (1/\hat{p}_n) \sum_{i=1}^{n-1} x_{i\alpha}]\, P_{\vec{x}}(\vec{x}_\alpha) \right.$$

$$\left. - \frac{x_{j\alpha}+1}{\hat{p}_n \hat{p}_j}\, P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_j) \right\} \left( - \frac{Nx_{j.}}{(N\hat{\lambda}+x_{..})^2} \right) - [1 + (1/N\hat{\lambda}) \sum_{i=1}^{n-1} x_{i\alpha}]\, P_{\vec{x}}(\vec{x}_\alpha)$$

$$+ \frac{x_{m\alpha}+1}{N\hat{\lambda}\hat{p}_m}\, P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m)$$

Next, substituting for $\hat{p}_n$ and $\hat{p}_i$ using (2-4.7) and (2-4.8) respectively, and using (1-3.11) to replace the expression $\sum_{j=1}^{n-1} x_{j\alpha}$, we finally have

$$D_{\hat{\lambda}}\, P_{\vec{x}}(\vec{x}_\alpha) = (1/\hat{\lambda}) \sum_{j=1}^{n-1} (x_{j\alpha}+1)\, P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_j) - [(w_\alpha/\hat{\lambda})(1+1/N\hat{\lambda})$$

$$+1]\, P_{\vec{x}}(\vec{x}_\alpha) + \left(\frac{N\hat{\lambda}+x_{..}}{x_{m.}}\right)\left(\frac{x_{m\alpha}+1}{N\hat{\lambda}}\right) P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m) \qquad (2\text{-}4.15)$$

If we now replace $\vec{x}_\alpha$ by $\vec{x}_\alpha+\vec{e}_m$ and hence $x_{m\alpha}$ by $x_{m\alpha}+1$, we obtain

$$D_{\hat{\lambda}}\, P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m) = (1/\hat{\lambda}) \sum_{j=1}^{n-1} (x_{j\alpha}+1)\, P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_j+\vec{e}_m)$$

$$-[\frac{w_\alpha+1}{\hat{\lambda}}(1+1/N\hat{\lambda}) + 1]\, P_{\vec{x}}(\vec{x}_\alpha+\vec{e}_m) + [1/\hat{\lambda} + \frac{N\hat{\lambda}+x_{..}+1}{x_{m.}+1} \cdot \frac{x_{m\alpha}+2}{N\hat{\lambda}}]$$

$$\cdot\, P_{\vec{x}}(\vec{x}_\alpha+2\vec{e}_m) \qquad (2\text{-}4.16)$$

If we know the $r^{th}$ iterated values of $\lambda, p_1, \ldots, p_{n-1}$, the substitution of (2-4.15) and (2-4.16) into (2-4.12) and the result into (2-4.9) yields $\hat{\lambda}_{r+1}$. Then the $(r+1)^{st}$ iterated values of $p_1, \ldots, p_n$ may be found from (2-4.7) and (2-4.8). One suggestion for initial estimates of $\lambda, p_1, \ldots, p_{n-1}$ is the moment estimators.

## 2-5 Covariance Matrix of the Maximum Likelihood Estimators

### A. Method of Calculation

As with the Poisson-multinomial distribution, direct calculation of $\hat{\Omega}$ is nearly impossible. We wish to show however, that it may be calculated indirectly by the same method as in § 1-5A, i.e. using Rao's theorem which is stated in that section.

To prove this distribution satisfies Rao's theorem, the same procedure as in § 1-5A is followed. From the remarks at the beginning of § 1-3 we conclude that the factorial moment generating function is given by (2-2.6) where $\vec{s}$ is set equal to $\vec{1}$ instead of $\vec{0}$. Since $1 - \sum_{i=1}^{n-1} p_i = p_n > 0$, it is clear that $g^*(\vec{s})$ is infinitely differentiable. Hence all the factorial moments are finite. Because each moment about the origin is a finite linear combination of factorial moments, these moments are also finite.

**Lemma 2-1.**

For the Poisson-negative multinomial distribution,

$$P_{\vec{X}}(\vec{x}+\vec{e}_m) \leq \left\{ \sum_{i=1}^{n-1} x_i + N\lambda[1 + (\sum_{i=1}^{n-1} x_i + 1)^N] \right\} P_{\vec{X}}(\vec{x})$$

Proof:

Using the convention defined in (2-1.3),

$$P_{\vec{x}}(\vec{x}+\vec{e}_m) = e^{-\lambda} \sum_{z=1}^{\infty} \frac{\lambda^z}{z!} \frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!} \frac{(x_1+\ldots+x_{n-1}+Nz)}{x_m+1}$$

$$\cdot\, p_n^{Nz}\Big(\prod_{i=1}^{n-1} p_i^{x_i}\Big)\, p_m$$

$$= \frac{p_m}{x_m+1}\left\{\sum_{i=1}^{n-1} x_i\, P_{\vec{x}}(\vec{x}) + Ne^{-\lambda}\sum_{z=1}^{\infty}(\lambda^z/z!)\,\frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!}\right.$$

$$\left.\cdot\, p_n^{Nz}\Big(\prod_{i=1}^{n-1} p_i^{x_i}\Big)\, z\right\} \qquad (2\text{-}5.1)$$

Let us observe that for $z \geq 2$

$$\frac{(x_1+\ldots+x_{n-1}+Nz-1)!}{x_1!\ldots x_{n-1}!(Nz-1)!}$$

$$= \frac{(x_1+\ldots+x_{n-1}+Nz-N-1)!}{x_1!\ldots x_{n-1}!(Nz-N-1)!} \prod_{k=1}^{N} \frac{(x_1+\ldots+x_{n-1}+Nz-k)}{(Nz-k)}$$

$$\leq \frac{(x_1+\ldots+x_{n-1}+Nz-N-1)!}{x_1!\ldots x_{n-1}!(Nz-N-1)!} \Big(\sum_{i=1}^{n-1} x_i + 1\Big)^N$$

We may use the above inequality to substitute for the factorials in (2-5.1) for $z=2, 3, \ldots$ . Thus

$$P_{\vec{x}}(\vec{x}+\vec{e}_m) \leq \frac{p_m}{x_m+1}\left\{\sum_{i=1}^{n-1} x_i\, P_{\vec{x}}(\vec{x}) + N\lambda e^{-\lambda}\frac{(x_1+\ldots+x_{n-1}+N-1)!}{x_1!\ldots x_{n-1}!(N-1)!}\right.$$

$$\cdot p_n^N \prod_{i=1}^{n-1} p_i^{x_i} + p_n^N N\lambda e^{-\lambda}(\sum_{i=1}^{n-1} x_i+1)^N \sum_{z=2}^{\infty} \frac{\lambda^{z-1}}{(z-1)!} \frac{(x_1+\ldots+x_{n-1}+Nz-N-1)!}{x_1!\ldots x_{n-1}!(Nz-N-1)!}$$

$$\cdot p_n^{N(z-1)} \prod_{i=1}^{n-1} p_i^{x_i}\Big\}$$

The second term in the braces is the "z=1" term in the expansion of $P_{\vec{X}}(\vec{x})$. The third term in the braces is equal to $P_{\vec{X}}(\vec{x})$ minus the "z=0" term in its expansion, a fact which is easily seen if we replace $z$ by $z-1$. Since each term in the expansion of $P_{\vec{X}}(\vec{x})$ is non-negative, we can conclude

$$P_{\vec{X}}(\vec{x}+\vec{e}_m) \leq \frac{p_m}{x_m+1}\Big\{ \sum_{i=1}^{n-1} x_i\ P_{\vec{X}}(\vec{x}) + N\lambda\ P_{\vec{X}}(\vec{x})$$

$$+ p_n^N (\sum_{i=1}^{n-1} x_i+1)^N N\lambda\ P_x(x)\Big\}$$

If we now note that $e^{-\lambda}$, $p_m$, $p_n$ are positive and less than one and $x_i \geq 0$ for all $i$, we can write

$$P_{\vec{X}}(\vec{x}+\vec{e}_m) \leq \Big\{ \sum_{i=1}^{n-1} x_i+N\lambda\ [1 + (\sum_{i=1}^{n-1} x_i+1)^N] \Big\}\ P_{\vec{X}}(\vec{x})$$

Q.E.D.

<u>Lemma 2-2</u>.

For the Poisson-negative multinomial distribution, $E(|\partial/\partial\lambda\ (\log L)|^3) < \infty$.

Proof:

By the same argument as in lemma 1-3, we are able to obtain (1-5.3) i.e.

$$E\left(\left|\frac{\partial}{\partial\lambda}\log L\right|^3\right) \leq \beta^3 E\left(\left|\frac{\partial/\partial\lambda[P_{\vec{x}}(\vec{x})]}{P_{\vec{x}}(\vec{x})}\right|^3\right)$$

Substituting from (2-4.3) for the derivative

$$= \beta^3 E\left\{\left|\frac{x_m+1}{N\lambda p_m}\;\frac{P_{\vec{x}}(\vec{x}+\vec{e}_m)}{P_{\vec{x}}(\vec{x})} - [(1/N\lambda)\sum_{j=1}^{n-1} x_j+1\right|^3\right\}$$

After replacing the absolute value of the sum by the sum of the absolute values and using the result of lemma 2-1, we have

$$E\left(\left|\frac{\partial}{\partial\lambda}\log L\right|^3\right) \leq \beta^3 E\left\langle \frac{x_m+1}{N\lambda p_m}\left\{\sum_{i=1}^{n-1} x_i+N\lambda[1+(\sum_{i=1}^{n-1} x_i+1)^N]\right\}\right.$$

$$\left. + (1/N\lambda)\sum_{j=1}^{n-1} x_j+1\right\rangle^3$$

Upon expansion this will be a finite sum of terms of the following type -

$$\left.\begin{array}{l} \text{constant} \;.\; E\left(\prod_{j=1}^{n-1} x_j^{\,n_j}\right) \\ \text{where the } n_j \text{ are non-negative integers} \end{array}\right\} \qquad (2\text{-}5.2)$$

These terms are all finite since we know all the moments are finite. Hence the result follows.

Q.E.D.

Lemma 2-3.

For the Poisson-negative multinomial distribution, $E(|\partial/\partial p_i\ (\log L)|^3) < \infty$ for $i = 1, 2, \ldots, n-1$.

Proof:

By the same argument as in lemma 1-3, but replacing $\lambda$ by $p_i$ we get (1-5.3) with $p_i$ replacing $\lambda$. Thus

$$E\left(\left|\frac{\partial}{\partial p_i} \log L\right|^3\right) \leq \beta^3 E\left(\left|\frac{\partial/\partial p_i [P_{\vec{x}}(\vec{x})]}{P_{\vec{x}}(\vec{x})}\right|^3\right)$$

Let us substitute from (2-4.2) for the derivative.

$$\leq \beta^3 E\left(\left|x_i/p_i + (1/p_n) \sum_{k=1}^{n-1} x_k - \frac{x_i+1}{p_n p_i} \frac{P_{\vec{x}}(\vec{x}+\vec{e}_i)}{P_{\vec{x}}(\vec{x})}\right|^3\right)$$

By the manipulation of absolute value sign and use of lemma 2-1, we arrive at

$$E\left(\left|\frac{\partial}{\partial p_i} \log L\right|^3\right) \leq \beta^3 E\left\langle x_i/p_i + (1/p_n) \sum_{k=1}^{n-1} x_k + \frac{x_i+1}{p_n p_i}\left\{\sum_{j=1}^{n-1} x_j + N\lambda\ [1+(\sum_{j=1}^{n-1} x_j+1)^N]\right\}\right\rangle$$

If this expression is expanded, a sum of terms like those in (2-5.2) is obtained and by the same reasoning as there, the result is obtained.

Q.E.D.

If we consider Rao's theorem which is stated near the end of §1-5A, we notice that lemmas 2-2 and 2-3 show that the Poisson-negative multinomial distribution satisfies its conditions by choosing $\eta = 1$. Hence (1-5.5) will hold, and for samples of reasonable size we may use the approximation

$$J^{*-1} = \hat{\Omega}^* \tag{2-5.3}$$

where
$$J^* = \begin{pmatrix} I_{\lambda\lambda}^* & I_{\lambda p_1}^* & \cdots & I_{\lambda p_{n-1}}^* \\ I_{p_1\lambda}^* & I_{p_1 p_1}^* & & I_{p_1 p_{n-1}}^* \\ \vdots & & & \vdots \\ I_{p_{n-1}\lambda}^* & I_{p_{n-1}p_1}^* & & I_{p_{n-1}p_{n-1}}^* \end{pmatrix} \tag{2-5.4}$$

is the information matrix, i.e.

$$I_{st}^* = E\left(\frac{\partial}{\partial s} \log L \cdot \frac{\partial}{\partial t} \log L\right) \tag{2-5.5}$$

where $L$ is the likelihood function for the Poisson-negative multinomial distribution.

## B. Calculation of the Elements of $J^*$.

Before proceeding with the calculation, let us first prove the following lemma which will be of use to simplify the notation.

Lemma 2-4.

$$\text{Define}\quad A_{ij} = -1 + \sum_{x_1=0}^{\infty} \cdots \sum_{x_{n-1}=0}^{\infty} \frac{(x_i+1)(x_j+1)P_{\vec{x}}(\vec{x}+\vec{e}_i)P_{\vec{x}}(\vec{x}+\vec{e}_j)}{N^2\lambda^2 p_i p_j\, P_{\vec{x}}(\vec{x})} \tag{2-5.6}$$

Then $A_{ij} = A_{mk} = A$, say $i,j,k,m = 1, 2, \ldots, n-1$

Proof:

The proof is identical to that of lemma 1-5 except that in place of equation (1-4.7) we refer to (2-4.4). We must also note that $P_{\vec{x}}$ now refers to the probability function of the Poisson-negative multinomial whereas in lemma 1-5 it referred to that of the Poisson-multinomial.

Q.E.D.

Consider $I_{\lambda\lambda}^*$. If we substitute (1-4.1) into (2-5.5) and use the same argument as was used for obtaining (1-5.13), we get a formula for $I_{\lambda\lambda}^*$ which is exactly the same as that for $I_{\lambda\lambda}$ in the last mentioned equation. Now substitute for the derivative using (2-4.4). Then

$$(1/\beta)I_{\lambda\lambda}^* = \sum_{x_1=0}^{\infty} \cdots \sum_{x_{n-1}=0}^{\infty} \frac{1}{P_{\vec{x}}(\vec{x})} \left\{ \frac{x_m+1}{N\lambda p_m} P_{\vec{x}}(\vec{x}+\vec{e}_m) - [1 + (1/N\lambda) \sum_{i=1}^{n-1} x_i]\, P_{\vec{x}}(\vec{x}) \right\}^2$$

Expansion of this expression using the definition of expectation and lemma 2-4 yields

$$(1/\beta)I_{\lambda\lambda}^* = A + 2 + (1/N^2\lambda^2)\, E[(\sum_{i=1}^{n-1} X_i)^2]$$

$$- (2/N^2\lambda^2 p_m)\, E[(\sum_{i=1}^{n-1} X_i - 1)\, X_m] - (2/N\lambda p_m)\, E(X_m)$$

$$+ (2/N\lambda) \sum_{i=1}^{n-1} E(X_i)$$

Let us make use of equations (2-3.3) to substitute for the expectations. After simplification, we find

$$(1/\beta)I_{\lambda\lambda}^* = A - (1-p_n)[N(\lambda+1)(1+p_n) + 1] / N\lambda p_n^{\;2} \qquad (2\text{-}5.7)$$

Similarly we may obtain the other entries of the information matrix

$$\left.\begin{aligned}
(1/\beta)I_{\lambda p_j}^* &= (1/\beta)I_{p_j\lambda}^* \\
&= N\lambda A/p_n + [(N\lambda+N+1)p_n^{\;2} - 1/p_n - N\lambda]/p_n \\
(1/\beta)I_{p_i p_j}^* &= N^2\lambda^2 A/p_n^{\;2} + (N\lambda/p_n^{\;2})[-(N\lambda+N+1)/p_n^{\;2} \\
&\quad + 1/p_n + N\lambda + 1] \quad \text{for} \quad i \neq j \\
(1/\beta)I_{p_i p_j}^* &= N^2\lambda^2 A/p_n^{\;2} + (N\lambda/p_n^{\;2})[-(N\lambda+N+1)/p_n^{\;2} \\
&\quad + 1/p_n + N\lambda + 1] + N\lambda/p_n p_i
\end{aligned}\right\} \qquad (2\text{-}5.8)$$

The calculations of the above results are outline in Appendix 2B. Now let us define

$$\left.\begin{aligned} B_{\lambda\lambda}^* &= (1/\beta)\, I_{\lambda\lambda}^* \\ B_{\lambda p}^* &= (1/\beta)\, I_{\lambda p_i}^* \\ B_{pp}^* &= (1/\beta)\, I_{p_i p_j}^* \quad , \quad j \neq i \end{aligned}\right\} \tag{2-5.9}$$

Hence $B_{pp}^* + N\lambda/p_n p_i = (1/\beta)\, I_{p_i p_i}^*$ (2-5.10)

$$J^* = \beta \begin{bmatrix} B_{\lambda\lambda}^* & B_{\lambda p}^* & \cdots & B_{\lambda p}^* \\ B_{\lambda p}^* & B_{pp}^* + N\lambda/p_n p_1 & & B_{pp}^* \\ \vdots & \vdots & & \vdots \\ B_{\lambda p}^* & B_{pp}^* & & B_{pp}^* + N\lambda/p_n p_{n-1} \end{bmatrix}$$

Now, from (2-5.3) we have that $\hat{\Omega} = J^{-1}$, and using Appendix 1D, formulas (1D-1) and (1D-6) with the appropriate association of variables,

$$\det J^* = \beta^n \left\{ B_{\lambda\lambda}^* + (B_{\lambda\lambda}^* B_{pp}^* - B_{\lambda p}^{*2}) p_n (1-p_n)/N\lambda \right\} \prod_{i=1}^{n-1} \left(\frac{N\lambda}{p_n p_i}\right)$$

(2-5.11)

$$\left.\begin{aligned}
\operatorname{var}\hat{\lambda} &= \frac{1}{\beta}\;\frac{N\lambda + B_{pp}{}^{*}\,p_n(1-p_n)}{B_{\lambda\lambda}{}^{*}N\lambda+(B_{\lambda\lambda}{}^{*}B_{pp}{}^{*}-B_{\lambda p}{}^{*2})p_n(1-p_n)}\\
\operatorname{cov}(\hat{\lambda},\hat{p}_i) &= -\frac{1}{\beta}\;\frac{B_{\lambda p}{}^{*}\,p_np_i}{B_{\lambda\lambda}{}^{*}N\lambda+(B_{\lambda\lambda}{}^{*}B_{pp}{}^{*}-B_{\lambda p}{}^{*2})p_n(1-p_n)}\\
\operatorname{cov}(\hat{p}_i,\hat{p}_j) &= -\frac{1}{\beta}\;\frac{(B_{\lambda\lambda}{}^{*}B_{pp}{}^{*}-B_{\lambda p}{}^{*2})p_n{}^{2}p_ip_j}{[B_{\lambda\lambda}{}^{*}N\lambda+(B_{\lambda\lambda}{}^{*}B_{pp}{}^{*}-B_{\lambda p}{}^{*2})p_n(1-p_n)]N\lambda}\\
\operatorname{var}\hat{p}_i &= \frac{1}{\beta}\left\{\frac{p_np_i}{N\lambda} - \frac{(B_{\lambda\lambda}{}^{*}B_{pp}{}^{*}-B_{\lambda p}{}^{*2})p_n{}^{2}p_i{}^{2}}{[B_{\lambda\lambda}{}^{*}N\lambda+(B_{\lambda\lambda}{}^{*}B_{pp}{}^{*}-B_{\lambda p}{}^{*2})p_n(1-p_n)]N\lambda}\right\}
\end{aligned}\right\}$$

(2-5.12)

Corollaries 2.1 and 2.2 in Rao [1947] state that if the distribution satisfies lemmas 2-2 and 2-3, then the maximum likelihood estimators are minimum variance estimators for large samples and in terms of the generalized variance, det $\hat{\Omega}^*$, are asymtotically efficient.

## 2-6 Efficiency of Method of Moments

### A. Method of Calculation

The method used is identical to the one described in 1-6A. To distinguish certain quantities such as the information matrix, covariance matrix, etc. for the distribution now under consideration from those for the distribution described in chapter 1, superscript stars will be written after the symbols (e.g. $\hat{\Omega}^*$, $J^*$, $M^*$, etc.). Thus the efficiency is given by

$$\text{Eff} = \frac{1}{\det \bar{\Omega}^* \det J^*} \qquad (2\text{-}6.1)$$

where $\bar{\Omega}^*$ is the covariance matrix of the moment estimators. As is shown in Appendix 1E,

$$\det \bar{\Omega}^* = (\det \phi^*)^2 \det M^* \qquad (2\text{-}6.2)$$

where $\det \phi^*$ is given by (1-6.6) and $M^*$ is given by (1-6.4) except that the quantities now refer to the Poisson-negative multinomial distribution.

## B. Calculation of $\det M^*$ in Terms of the Parameters

We may repeat the argument that led to (1-6.7) for the present distribution and obtain an identical result, namely

$$\text{var } \bar{X}_i = (1/\beta)[E(X_i^2) - E^2(X_i)]$$

Substituting for the expectations using (2-3.3), we obtain

$$\text{var } \bar{X}_i = (1/\beta)(N\lambda p_i/p_n)[(p_i/p_n)(N+1)+1] \qquad (2\text{-}6.3)$$

If we repeat the argument that led to (1-6.9) we will obtain $\text{cov}(\bar{X}_i,\bar{X}_j) = (1/\beta)[E(X_iX_j) - E(X_i)\, E(X_j)]$ and substitution from (2-3.3) yields

$$\text{cov}(\bar{X}_i,\bar{X}_j) = (1/\beta)N(N+1)\lambda\ p_ip_j/p_n^2 \qquad (2\text{-}6.4)$$

Similarly, repetition of the arguments leading to (1-6.11) and (1-6.15) lead to identical equations for the Poisson-negative multinomial distribution. By substituting (2-3.3) into (1-6.11) and (1-6.15) and using Appendices 1F and 1G with $G_i^*$ replacing $G_i$ and $p_k/p_n$ replacing $p_k$, we have respectively

$$\left.\begin{aligned} &\operatorname{cov}(\overline{W}^2,\overline{X}_k) = (1/\beta)(N\lambda\, p_i/p_n)H_1^* \\ &\text{where}\quad H_1^* = \frac{1-p_n}{p_n}\left\{G_2^*\left(\frac{1-p_n}{p_n}\right) + 3G_1^*\right\} \\ &\qquad\qquad - N\lambda\left[G_1^*\left(\frac{1-p_n}{p_n}\right)+1\right] + 1 \end{aligned}\right\} \qquad (2\text{-}6.5)$$

and

$$\left.\begin{aligned} &\operatorname{var}\overline{W}^2 = (1/\beta)[N\lambda(1-p_n)/p_n]H_2^* \\ &\text{where}\quad H_2^* = G_3^*\left(\frac{1-p_n}{p_n}\right)^3 + 6G_2^*\left(\frac{1-p_n}{p_n}\right)^2 + 7G_1^*\left(\frac{1-p_n}{p_n}\right) \\ &\qquad\qquad + 1 - N\lambda\left(\frac{1-p_n}{p_n}\right)\left[G_1^*\left(\frac{1-p_n}{p_n}\right)+1\right]^2 \end{aligned}\right\} \qquad (2\text{-}6.6)$$

Now let us stubstitute (2-6.3) through (2-6.6) into the expression for $M^*$ which is given by (1-6.4). Then

$$M^* = \frac{1}{\beta}\begin{pmatrix} \frac{N\lambda(1-p_n)}{p_n}H_2^* & \frac{N\lambda p_1}{p_n}H_1^* & \cdots & \frac{N\lambda p_{n-1}}{p_n}H_1^* \\ \frac{N\lambda p_1}{p_n}H_1^* & \frac{N\lambda p_1}{p_n}\left[\frac{p_1}{p_n}(N+1)+1\right] & \cdots & N(N+1)\lambda\frac{p_1 p_{n-1}}{p_n^2} \\ \vdots & N(N+1)\lambda\frac{p_1\,p_2}{p_n^2} & \ddots & \vdots \\ & \vdots & & \\ \frac{N\lambda p_{n-1}}{p_n}H_1^* & N(N+1)\lambda\frac{p_1 p_{n-1}}{p_n^2} & & \frac{N\lambda p_{n-1}}{p_n}\left[\frac{p_{n-1}}{p_n}(N+1)+1\right] \end{pmatrix}$$

(2-6.7)

The determinant of the above matrix is found in Appendix 1H if we replace $H_i$, $N'$, and $R$ by $H_i^*$, $(N+1)/p_n$, and $N\lambda/p_n$ respectively in the result given in the appendix. Thus

$$\det M = (N\lambda/\beta p_n)^n (1-p_n)\left(\prod_{i=1}^{n-1} p_i\right)\left\{H_2^*[1+(1-p_n)(N+1)/p_n] - H_1^{*2}\right\} \tag{2-6.8}$$

## C. Determination of the Jacobian, det $\phi^*$

The expression for $\phi^*$ is the same as the one given for $\phi$ in (1-6.21) with the exception that the quantities refer now to the Poisson-negative multinomial distribution. The partial derivatives in (1-6.21) can be obtained in a straightforward manner from (2-3.9) and (2-3.10). Hence, by differentiating,

$$\frac{\partial\lambda}{\partial E(W^2)} = -\frac{N+1}{N}\,\frac{E^2(W)}{[E(W^2)-E^2(W)-E(W)]^2}$$

$$\frac{\partial\lambda}{\partial E(X_i)} = \frac{N+1}{N}\cdot\frac{E(W)[2E(W^2)-E(W)]}{[E(W^2)-E^2(W)-E(W)]^2}$$

$$\frac{\partial p_i}{\partial E(W^2)} = \frac{(N+1)E(X_i)}{[E(W^2)-E^2(W)+NE(W)]^2}$$

$$\frac{\partial p_i}{\partial E(X_j)} = \frac{E(X_i)}{E^2(W)[E(W^2)-E^2(W)+NE(W)]^2}\left\{-E(W^2)[E(W^2)-E^2(W)+NE(W)] + E(W)[E(W^2)-E(W)][E(W)-N] - E^3(W)\right\} \quad \text{if } i \neq j$$

$$\frac{\partial p_i}{\partial E(X_i)} = \frac{E(X_i)}{E^2(W)[E(W^2)-E^2(W)+NE(W)]^2}\left\{-E(W^2)[E(W^2)-E^2(W)+NE(W)]\right.$$

$$+ E(W)[E(W^2)-E(W)][E(W) - N] - E^3(W)\Big\}$$

$$+ \frac{E(W^2)-E^2(W)-E(W)}{E(W)[E(W^2)-E^2(W)+NE(W)]}$$

Let us use (2-3.3), (2-3.6), and (2-3.7) to substitute for the expectations and then simplify the results to obtain

$$\left.\begin{aligned}
\frac{\partial\lambda}{\partial E(W^2)} &= -\frac{p_n{}^2}{N(N+1)(1-p_n)^2} \\
\frac{\partial\lambda}{\partial E(X_i)} &= \frac{Fp_n}{N(N+1)(1-p_n)^2} \\
\frac{\partial p_i}{\partial E(W^2)} &= \frac{p_i p_n{}^3}{N(N+1)\lambda(1-p_n)^2} \\
\frac{\partial p_i}{\partial E(X_j)} &= \frac{Dp_i p_n}{N(N+1)\lambda(1-p_n)^2} \quad \text{for} \quad i \neq j \\
\frac{\partial p_i}{\partial E(X_i)} &= \frac{Dp_i p_n}{N(N+1)\lambda(1-p_n)^2} + \frac{p_n}{N\lambda}
\end{aligned}\right\} \qquad (2\text{-}6.9)$$

where $F$ and $D$ are defined by

$$\left.\begin{aligned}
F &= 2(1-p_n)[N(\lambda+1)+1] + p_n \\
D &= -N(1-p_n{}^2) - 1 - \frac{1}{N+1}[2N^2\lambda p_n(1-p_n) \\
&\quad - N^2\lambda^2(1-p_n)^2 + N\lambda(1-p_n{}^2)]
\end{aligned}\right\} \qquad (2\text{-}6.10)$$

We may now substitute these values into (1-6.21) and obtain an explicit expression for $\phi^*$. The determinant of $\phi^*$ is calculated in Appendix 2C if we replace $Q$ in the appendix by $N(N+1)(1-p_n)^2/p_n$. The result is

$$\det \phi^* = -\frac{p_n^{n+1}[(N+1)(1-p_n)+D+Fp_n]}{(N+1)^2 N(N\lambda)^{n-1}(1-p_n)^3} \tag{2-6.11}$$

Now we are able to substitute (2-6.8) and (2-6.11) into (2-6.2) to get

$$\det \bar{\Omega}^* = \frac{p_n^{n+2}\lambda[(N+1)(1-p_n)+D+Fp_n]^2}{(N+1)^4 N(N\lambda)^{n-1}(1-p_n)^5\beta^n}\Big\{H_2^* - H_1^{*2} + (1-p_n)H_2^*\,(N+1)/p_n\Big\}\prod_{i=1}^{n-1} p_i \tag{2-6.12}$$

By the same argument that led to (1-6.2), it is easy to see that

$$\mathrm{Eff} = \frac{1}{\det \Omega^* \det J^*}$$

If we replace the determinants by their explicit expressions given in (2-5.11) and (2-6.12), we find that

$$\mathrm{Eff} = \frac{N^2(N+1)^4(1-p_n)^5}{p_n^{\,2}[B_{\lambda\lambda}^* N\lambda + p_n(1-p_n)(B_{\lambda\lambda}^* B_{pp}^* - B_{\lambda p}^{*2})]}\cdot \frac{1}{[(N+1)(1-p_n)+D+Fp_n]^2}\cdot\frac{1}{[(H_2^*-H_1^{*2})p_n+(1-p_n)(N+1)H_2^*]} \tag{2-6.13}$$

## 2-7 Sample Zero Frequency and Unit Sample Frequency Estimators

### A. Sample Zero Frequency and First Moments

This type of estimation is useful under the same conditions as outlined in § 1-7A, i.e. the zero sample occurs fairly frequently. Setting $\vec{X} = \vec{0}$ in (2-2.10), we obtain

$$P_{\vec{X}}(\vec{0}) = \exp\left[\lambda(p_n{}^N - 1)\right] \tag{2-7.1}$$

If we define $F(\vec{a})$ in the same manner as in §1-7A, then $E[\frac{1}{\beta} F(\vec{0})] = P_{\vec{X}}(\vec{0}) = \exp[\lambda(p_n{}^N-1)]$ and hence $(1/\beta)F(\vec{0})$ is an unbiased estimator for $P_{\vec{X}}(\vec{0})$.

Thus, to obtain the sample zero frequency estimators $\tilde{\lambda}$ and $\tilde{p}_i$ of $\lambda$ and $p_i$, we solve the equation $(1/\beta)F(\vec{0}) = \exp[\tilde{\lambda}(\tilde{p}_n{}^N-1)]$ along with the equations of (2-3.3) which involve the first moments. Hence we solve the set

$$\left.\begin{aligned} (1/\beta)F(\vec{0}) &= \exp[\tilde{\lambda}(\tilde{p}_n{}^N-1)] \\ X_i &= N\tilde{\lambda}\,\tilde{p}_i/\tilde{p}_n \qquad i = 1, 2, \ldots, n-1 \end{aligned}\right\} \tag{2-7.2}$$

for $\tilde{\lambda}$ and $\tilde{p}_i$ .

To do this we add the second equation of (2-7.2) for $i = 1, 2, \ldots, n-1$. Then we have

$$\sum_{i=1}^{n-1} X_i = N\tilde{\lambda}\,(1-\tilde{p}_n)/\tilde{p}_n$$

and hence
$$\tilde{p}_n = \frac{N\tilde{\lambda}}{\sum_{i=1}^{n-1} X_i + N\tilde{\lambda}} \tag{2-7.3}$$

We may now substitute this quantity into the logarithm of the first equation of (2-7.2) and find

$$\log \frac{F(\vec{0})}{\beta} = \tilde{\lambda}\left[\left(\frac{N\tilde{\lambda}}{\sum_{i=1}^{n-1} X_i+N\tilde{\lambda}}\right)^N - 1\right]$$

We must now solve for $\tilde{\lambda}$. It is best to use some numerical procedure. Once having done this we may find $\tilde{p}_n$ from (2-7.3) and finally $\tilde{p}_i$ from (2-7.2).

B. Unit Sample Frequency Estimation

For cases where the unit samples occur frequently, the unit sample estimators are sometimes useful. From (2-2.7)

$$P_{\vec{x}}(\vec{e}_k) = \left.\frac{\partial g^*(\vec{s})}{\partial s_k}\right|_{s=0} = \exp\left[\lambda(p_n{}^N-1)\right] N p_n{}^N p_k \lambda \tag{2-7.4}$$

Now $E[(1/\beta)F(\vec{e}_k)] = P_{\vec{x}}(\vec{e}_k)$ and thus the unit frequency estimator is unbiased for $P_{\vec{x}}(\vec{e}_k)$. To find the unit sample estimators $\check{\lambda}$ and $\check{p}_i$ for $\lambda$ and $p_i$ respectively, we must solve

$$(1/\beta)F(\vec{e}_k) = \exp[\check{\lambda}(\check{p}_n{}^N-1)] N\check{p}_n{}^N \check{p}_k \check{\lambda} \tag{2-7.5}$$

together with the zero sample estimator

$$(1/\beta)F(\vec{0}) = \exp\ [\check{\lambda}(\check{p}_n{}^N-1)] \tag{2-7.6}$$

After dividing (2-7.5) by (2-7.5) with $k=1$, we get

$$\frac{F(\vec{e}_k)}{F(\vec{e}_1)} = \frac{\check{p}_k}{\check{p}_1} \tag{2-7.7}$$

Also, after dividing (2-7.5) with $k=1$ by the first equation of (2-7.2) and noting the definition of $p_n$

$$F(\vec{e}_1)/F(\vec{0}) = N\left(1 - \sum_{k=1}^{n-1} \check{p}_k\right)^N \check{p}_1\ \check{\lambda} \tag{2-7.8}$$

If we now take the logarithm of (2-7.6) and substitute for the $\check{p}_k$ from (2-7.7)

$$\frac{1}{\check{\lambda}} \log \frac{F(\vec{0})}{\beta} + 1 = [1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^N \tag{2-7.9}$$

Let us divide (2-7.8) by $\check{\lambda}$ and substitute for $\check{p}_k$, k = 2, 3, ..., n-1 from (2-7.7)

$$1/\check{\lambda} = \frac{\check{p}_1 NF(\vec{0})}{F(\vec{e}_1)} [1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^N \tag{2-7.10}$$

We may now substitute this into (2-7.9) and obtain

$$\frac{\check{p}_1 NF(\vec{0})}{F(\vec{e}_1)} [1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^N \log \frac{F(\vec{0})}{\beta} + 1 = [1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^N$$

This may be rewritten as

$$[1 - \frac{\check{p}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k)]^N [\frac{\check{p}_1 NF(\vec{0})}{F(\vec{e}_1)} \log \frac{F(\vec{0})}{\beta} - 1] + 1 = 0 \tag{2-7.11}$$

Probably the best way to solve this for $\check{p}_1$ is to use a suitable numerical procedure. After doing this, $\check{\lambda}$ may be obtained from (2-7.10) and $\check{p}_k$, k=2, 3, ..., n-1 from (2-7.7).

# CHAPTER III

# LIMITING DISTRIBUTIONS OF THE POISSON-MULTINOMIAL AND POISSON-NEGATIVE MULTINOMIAL DISTRIBUTIONS

## 3-1 Introduction

In many applications certain parameters may be known already to be very large, to be a particular value, or to be almost negligable. Usually, if circumstances permit, it is much easier to consider the limiting distributions as the parameters approach their respective limits.

If a particular $p_i$ is allowed to approach zero in either of the above distributions, ~~the form of the distributions~~, the form of the distribution will remain unchanged except that the $i^{th}$ variate will be completely ignored.

## 3-2 The Poisson-Poisson Distribution

The most interesting limiting distribution is the one in which $N \to \infty$ and $p_i \to 0$ for $i=1, 2, \ldots, n-1$ in such a way that $Np_i = \alpha_i$ = constant. Then, by using the fact that

$$p_n = 1 - (1/N) \sum_{i=1}^{n-1} \alpha_i , \quad \text{we have from (1-2.1)}$$

$$g_L(\vec{s}) = \lim_{N\to\infty} g(\vec{s}) = \exp\left\{ \lambda \lim_{N\to\infty} [(1/N) \sum_{i=1}^{n-1} \alpha_i(s_i - 1) + 1]^N - \lambda \right\}$$

If we now apply the mathematical identity

$$e^Q = \lim_{N\to\infty} (1+Q/N)^N = \lim_{N\to\infty} (1+1/N)^{QN} \qquad (3\text{-}2.1)$$

we will have

$$g_L(\vec{s}) = \exp\left\{ \lambda \exp\left[ \sum_{i=1}^{n-1} \alpha_i(s_i-1) \right] - \lambda \right\} \tag{3-2.2}$$

If, instead, we start from (2-2.6), we get

$$g_L^*(\vec{s}) = \lim_{N\to\infty} g^*(\vec{s}) = \exp\left\{ \lambda \lim_{N\to\infty} \left[ \frac{p_n}{1 - \sum_{i=1}^{n-1} s_i p_i} \right]^N - \lambda \right\}$$

By using the definition of the $\alpha_i$ and a familiar theorem about limits,

$$= \exp\left\{ \lambda \lim_{N\to\infty} [1-(1/N) \sum_{i=1}^{n-1} \alpha_i]^N \lim_{N\to\infty} [1-(1/N) \sum_{i=1}^{n-1} \alpha_i s_i]^{-N} - \lambda \right\}$$

Finally, let us apply identity (3-2.1) to the above limits. After slight simplification we get

$$g_L^*(\vec{s}) = \exp\left\{ \lambda \exp\left[ \sum_{i=1}^{n-1} \alpha_i(s_i-1) \right] - \lambda \right\} \tag{3-2.3}$$

Equations (3-2.2) and (3-2.3) show that $g_L(\vec{s}) = g_L^*(\vec{s})$. Hence both the Poissson-multinomial and Poisson-negative multinomial have the same limiting distribution. From the form of the probability generating function we see that this is a Poisson-multivariate Poisson, or the multivariate analogue of the Neyman Type A distribution.

From what has been done in chapters I and II it is, for the most part, an easy matter to obtain the same results for the limiting distribution as we obtained for the two previous distributions. We simply allow the parameters to approach their limits in the formulas which give the quantities we wish to find.

The moments of the distribution may be found from (1-3.3),

$E(X_i) = \lim_{N\to\infty} N\lambda \; p_i = \lambda\,\alpha_i$ , or from (2-3.3),

$$\left.\begin{aligned} E(X_i) &= \lim_{N\to\infty} N\lambda p_i/p_n = \lambda\,\alpha_i \\ \text{Similarly,}\quad E(X_i^{\,2}) &= \lambda(\lambda+1)\alpha_i^{\,2} + \lambda\alpha_i \\ E(X_iX_j) &= \lambda(\lambda+1)\alpha_i\alpha_j \end{aligned}\right\} \qquad (3\text{-}2.4)$$

Here we may consider estimating the parameters. We must be aware that these are now $\lambda$, $\alpha_1$, ..., $\alpha_{n-1}$. From either (1-3.9) and (1-3.10), or (2-3.9) and (2-3.10), we can derive the moment estimators

$$\left.\begin{aligned} \lambda &= \frac{E(W^2)}{E(W^2)-E^2(W)-E(W)} \\ \alpha_i = \lim_{N\to\infty} Np_i &= \frac{E(X_i)[E(W^2)-E^2(W)-E(W)]}{E^2(W)} \end{aligned}\right\} \qquad (3\text{-}2.5)$$

Also, the maximum likelihood estimators for the $\alpha_i$ may be found from (1-4.11) or (2-4.8) by taking the limit of $Np_i$ as $N\to\infty$. Thus we have

$$\hat{\alpha}_m = x_{m.}/\hat{\lambda}$$

The maximum likelihood estimator for $\lambda$ is given by either (1-4.14) or (2-4.11) since both of these remain unchanged as $N\to\infty$. We must realize, however, that now $P_{\vec{X}}$ refers to the density of the limiting distribution. We shall henceforth denote the limiting density by $\underline{P}_{\vec{X}}$.

## 3-3 The Information Matrix

From either (1-5.12) or (2-5.6) it is clear that in the limiting case

$$A = -1 + \sum_{x_1=0}^{\infty} \cdots \sum_{x_{n-1}=0}^{\infty} \frac{(x_i+1)(x_j+1)\underline{P}_{\vec{x}}(\vec{x}+\vec{e}_i)\underline{P}_{\vec{x}}(\vec{x}+\vec{e}_j)}{\lambda^2 \alpha_i \alpha_j \, \underline{P}_{\vec{x}}(\vec{x})}$$

Let $\underline{J}$ denote the information matrix and $\underline{I}_{st}$ denote its entries. Then

$$\underline{I}_{\lambda\lambda} = \lim_{N\to\infty} I_{\lambda\lambda} = \lim_{N\to\infty} I_{\lambda\lambda}{}^*$$

From (1-5.16) and (1-5.18), or (2-5.7) and (2-5.9)

$$(1/\beta)\, \underline{I}_{\lambda\lambda} = \lim_{n\to\infty} B_{\lambda\lambda} = \lim_{N\to\infty} B_{\lambda\lambda}{}^* = A \qquad (3\text{-}3.1)$$

In calculating the other entries we must be careful since $\alpha_i$ corresponds to $Np_i$ rather than $p_i$. Thus from the definition of the information matrix, (1-5.2)

$$\underline{I}_{\lambda\alpha_i} = E[\frac{1}{\underline{P}_{\vec{x}}{}^2(\vec{x})} \frac{\partial}{\partial\lambda} \underline{P}_{\vec{x}}(\vec{x}) \frac{\partial}{\partial\alpha_i} \underline{P}_{\vec{x}}(\vec{x})]$$

$$= \lim_{N\to\infty} E[\frac{1}{P_{\vec{x}}{}^2(\vec{x})} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}) \frac{\partial}{\partial(Np_i)} P_{\vec{x}}(\vec{x})]$$

$$= \lim_{N\to\infty} (1/N)\, E[\frac{1}{P_{\vec{x}}{}^2(\vec{x})} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}) \frac{\partial}{\partial p_i} P_{\vec{x}}(\vec{x})]$$

$$= \lim_{N\to\infty} (I_{\lambda p_i}/N) = \lim_{N\to\infty} (I_{\lambda p_i}{}^*/N)$$

Now from (1-5.17) and (1-5.18), or (2-5.8) and (2-5.9)

$$(1/\beta)\underline{I}_{\lambda\alpha_i} = \lim_{N\to\infty} (B_{\lambda p}/N) = \lim_{N\to\infty} (B_{\lambda p}{}^*/N) = -\lambda A+1 \qquad (3\text{-}3.2)$$

Similarly, using the same equations,

$$(1/\beta)\underline{I}_{\alpha_i\alpha_j} = (1/\beta) \lim_{N\to\infty} (I_{p_i p_j}/N^2) = (1/\beta) \lim_{N\to\infty} (I_{p_i p_j}{}^*/N^2)$$

$$= \lim_{N\to\infty} (B_{pp}/N^2) = \lim_{N\to\infty} (B_{pp}{}^*/N^2)$$

$$= \lambda(\lambda A - 1) \qquad (3\text{-}3.3)$$

Also, $$(1/\beta)\underline{I}_{\alpha_i\alpha_i} = \lambda(\lambda A-1+1/\alpha_i) \qquad (3\text{-}3.4)$$

Hence we may summarize $\underline{J}$ as

$$\underline{J} = \begin{pmatrix} \underline{I}_{\lambda\lambda} & \underline{I}_{\lambda p_1} & \cdots & \underline{I}_{\lambda p_{n-1}} \\ \underline{I}_{\lambda p_1} & \underline{I}_{p_1 p_1} & \cdots & \underline{I}_{p_1 p_{n-1}} \\ \vdots & \vdots & & \vdots \\ \underline{I}_{\lambda p_{n-1}} & \underline{I}_{p_1 p_{n-1}} & \cdots & \underline{I}_{p_{n-1} p_{n-1}} \end{pmatrix}$$

where

$$\left.\begin{aligned} (1/\beta)\underline{I}_{\lambda\lambda} &= \underline{B}_{\lambda\lambda} = A \\ (1/\beta)\underline{I}_{\lambda p_i} &= \underline{B}_{\lambda p} = -\lambda A+1 \\ (1/\beta)\underline{I}_{p_i p_j} &= \underline{B}_{pp} = \lambda(\lambda A-1) \quad \text{if} \quad i\neq j \\ (1/\beta)\underline{I}_{p_i p_i} &= \underline{B}_{pp} + \lambda/\alpha_i = \lambda(\lambda A-1+1/\alpha_i) \end{aligned}\right\} \qquad (3\text{-}3.5)$$

## 3-4 Efficiency of Method of Moments

By using the same arguments as are used to obtain (1-6.2) and (1-6.5), we can show that

$$\text{Eff} = \frac{1}{(\det \underline{\phi})^2 \det \underline{M} \det \underline{J}} \tag{3-4.1}$$

where $\underline{M}$ is the covariance matrix of the moment estimators of $\lambda$, $\alpha_1$, ..., $\alpha_{n-1}$ and $\det \underline{\phi}$ is the Jacobian

$$\frac{\partial[\lambda,\alpha_1,\ldots,\alpha_{n-1}]}{\partial[E(W^2),E(X_1),\ldots,E(X_{n-1})]}$$

First let us consider $\underline{M}$. $\underline{M}$ may be found by taking limits in either (1-6.19) or (2-6.7). This results in

$$\left.\begin{aligned}
\text{cov}\,(\overline{X}_i,\overline{X}_j) &= (1/\beta)\,\lambda\alpha_i\alpha_j \quad \text{for} \quad i \neq j \\
\text{var}\;\; \overline{X}_i &= (1/\beta)\,\lambda\alpha_i(\alpha_i+1) \\
\text{cov}\,(\overline{W^2},\overline{X}_i) &= (1/\beta)\,\lambda\alpha_i\underline{H}_1 \\
\text{var}\;\; \overline{W^2} &= (1/\beta)\,\lambda\alpha\,\underline{H}_2
\end{aligned}\right\} \tag{3-4.2}$$

where

$$\alpha = \sum_{i=1}^{n-1} \alpha_i$$

$$\left.\begin{aligned}
\underline{H}_1 &= \lim_{N\to\infty} H_1 = \lim_{N\to\infty} H_1^* \\
&= (2\lambda+1)\alpha^2 + (4\lambda+3)\alpha + 1 \\
\underline{H}_2 &= \lim_{N\to\infty} H_2 = \lim_{N\to\infty} H_2^* \\
&= \alpha^3(4\lambda^2+6\lambda+1)+2\alpha^2(2\lambda^2+8\lambda+3)+\alpha(6\lambda+7)+1
\end{aligned}\right\} \tag{3-4.3}$$

Because each entry of $M$ and $M^*$ converges to the corresponding entry of $\underline{M}$ as $N\to\infty$,

$$\det \underline{M} = \lim_{N\to\infty} \det M = \lim_{N\to\infty} \det M^* \tag{3-4.4}$$

Now consider $\det \underline{\phi}$

$$\det \underline{\phi} = \lim_{N\to\infty} \frac{\partial[\lambda,\alpha_1,\ldots,\alpha_{n-1}]}{\partial[E(W^2),E(X_1),\ldots,E(X_{n-1})]}$$

$$= \lim_{N\to\infty} \begin{vmatrix}
\frac{\partial\lambda}{\partial E(W^2)} & \frac{\partial\lambda}{\partial E(X_1)} & \cdots & & \frac{\partial\lambda}{\partial E(X_{n-1})} \\
\frac{\partial\alpha_1}{\partial E(W^2)} & \frac{\partial\alpha_1}{\partial E(X_1)} & \frac{\partial\alpha_1}{\partial E(X_2)} & \cdots & \frac{\partial\alpha_1}{\partial E(X_{n-1})} \\
\vdots & & & & \\
\frac{\partial\alpha_{n-1}}{\partial E(W^2)} & \frac{\partial\alpha_{n-1}}{\partial E(X_1)} & \cdots & & \frac{\partial\alpha_{n-1}}{\partial E(X_{n-1})}
\end{vmatrix}$$

By virtue of (1-6.21) and the fact that $\alpha_i = Np_i$,

$$\det \underline{\phi} = \lim_{N\to\infty} N^{n-1} \det \underline{\phi} = \lim_{N\to\infty} N^{n-1} \det \phi^* \qquad (3\text{-}4.5)$$

From § 3-3 we see that

$$\det \underline{J} = \lim_{N\to\infty} \begin{vmatrix} I_{\lambda\lambda} & I_{\lambda p_1}/N & & I_{\lambda p_{n-1}}/N \\ I_{\lambda p_1}/N & I_{p_1 p_1}/N^2 & . & I_{p_1 p_{n-1}}/N^2 \\ \vdots & \vdots & \ddots & \vdots \\ I_{\lambda p_{n-1}}/N & I_{p_{n-1} p_1}/N^2 & \cdots & I_{p_{n-1} p_{n-1}}/N^2 \end{vmatrix}$$

$$= \lim_{N\to\infty} N^{-2(n-1)} \det J \qquad (3\text{-}4.6)$$

The same relation holds if we consider $J^*$. Now if we substitute (3-4.4), (3-4.5) and (3-4.6) into (3-4.1) we have

$$\begin{aligned} \underline{\mathrm{Eff}} &= \Big\{ (\lim_{N\to\infty} N^{n-1} \det \phi)^2 (\lim_{N\to\infty} \det M) \\ &\qquad \cdot (\lim_{N\to\infty} N^{-2(n-1)} \det J) \Big\}^{-1} \\ &= \lim_{N\to\infty} \Big\{ (\det \phi)^2 \det M \det J \Big\}^{-1} \\ &= \lim_{N\to\infty} (\mathrm{Eff}) \end{aligned} \qquad (3\text{-}4.7)$$

Exactly the same result will hold if we consider $\phi^*$, $M^*$, and $J^*$ instead of $\phi$, $M$, and $J$. By taking limits either in (1-6.24) or in (2-6.13) and then substituting for the $\underline{B}$'s from (3-3.5),

$$\underline{\text{Eff}} = \frac{\alpha^3}{[\lambda A(\alpha+1)-\alpha][\underline{H}_2(\alpha+1)-\underline{H}_1^{\,2}]} \tag{3-4.8}$$

where $\underline{H}_1$ and $\underline{H}_2$ are given by (3-4.3). This calculation is outlined in Appendix 3A.

## 3-5 Sample Zero Frequency and First Moment Estimators

From either (1-7.1) or (2-7.1) we can deduce that the probability of the zero sample is

$$\begin{aligned}\underline{P}_{\vec{x}}(\vec{0}) &= \lim_{N\to\infty} \exp\,[\lambda(p_n{}^N-1)] \\ &= \exp\,[\lambda(\lim_{N\to\infty} \langle p_n{}^N \rangle -1)].\end{aligned}$$

Because $\alpha = N(1-p_n)$, we have

$$\lim_{N\to\infty} p_n{}^N = \lim_{N\to\infty} (1-\alpha/N)^N = e^{-\alpha} \tag{3-5.1}$$

Thus $\underline{P}_{\vec{x}}(\vec{0}) = \exp\,[\lambda(e^{-\alpha}-1)]$

Defining $F(\vec{a})$ to be the frequency with which $\vec{X} = \vec{a} = (a_1, \ldots, a_{n-1})$ occurs in $\beta$ observations, we have

$$E(\tfrac{1}{\beta}\,F(0)) = \underline{P}_x(0) = \exp\,[\lambda(e^{-\alpha}-1)] \tag{3-5.2}$$

Thus $(1/\beta)F(0)$ is an unbiased estimator for $\underline{P}_x(0)$. The sample zero frequency estimators $\tilde{\lambda}$ and $\tilde{p}_i$ for $\lambda$ and $p_i$ respectively may be found by solving the above equations simultaneously with the first moment equations of (3-2.4), i.e.

$$(1/\beta)F(\vec{0}) = \exp[\tilde{\lambda}(e^{-\tilde{\alpha}}-1)]$$

$$X_i = \tilde{\lambda}\,\tilde{\alpha}_i \qquad (3\text{-}5.3)$$

Adding the second equation of (3-5.3) for $i=1, 2, \ldots, n-1$,

$$\sum_{i=1}^{n-1} X_i = \tilde{\lambda}\sum_{i=1}^{n-1}\tilde{\alpha}_i = \tilde{\lambda}\,\tilde{\alpha}$$

This equation may be solved for $\tilde{\alpha}$ and used to replace $\tilde{\alpha}$ in the first equation of (3-5.3). Hence

$$(1/\beta)F(\vec{0}) = \exp\left\{\tilde{\lambda}\exp[(-1/\tilde{\lambda})\sum_{i=1}^{n-1} X_i - \tilde{\lambda}\right\}$$

This may be solved for $\tilde{\lambda}$ using a suitable numerical method. Once having done this we can solve for the $\tilde{\alpha}_i$ using the second equation of (3-5.2), i.e. $\tilde{\alpha}_i = X_i/\tilde{\lambda}$.

## 3-6 Unit Sample Frequency Estimation

Taking limits in either (1-7.7) or (2-7.4) with the help of (3-5.1) yields

$$\underline{P}_{\vec{x}}(\vec{e}_k) = \exp[\lambda(e^{-\alpha}-1)]\,\alpha_k\lambda e^{-\alpha} \quad k=1, 2, \ldots, n-1 \qquad (3\text{-}6.1)$$

Let us note that $E[(1/\beta)F(\vec{x})] = \underline{P}_{\vec{x}}(\vec{x})$, hence $(1/\beta)F(\vec{e}_k)$ is an unbiased estimator for $\underline{P}_{\vec{x}}(\vec{e}_k)$. Thus, if we let $\check{\lambda}$ and $\check{\alpha}_k$ be the unit sample estimators for $\lambda$ and $\alpha_k$ respectively, then we may solve (3-6.1) with $\underline{P}_{\vec{x}}(\vec{e}_k)$ replaced by its estimator, and (3-5.2) with $\underline{P}_{\vec{x}}(\vec{0})$ replaced by its estimator for $\check{\lambda}$ and

$\alpha_k$, k=1, 2, ..., n-1, i.e. we must solve the system

$$\left.\begin{aligned} (1/\beta)F(\vec{0}) &= \exp\,[\check{\lambda}(e^{-\check{\alpha}}-1)] \\ (1/\beta)F(\vec{e}_k) &= \exp\,[\check{\lambda}(e^{-\check{\alpha}}-1)]\check{\alpha}_k\check{\lambda}e^{-\check{\alpha}} \\ &\qquad k=1,\ 2,\ldots,\ n-1 \end{aligned}\right\} \tag{3-6.2}$$

Upon division of the second equation by the second equation with k=1, we obtain

$$\frac{F(\vec{e}_k)}{F(\vec{e}_1)} = \frac{\check{\alpha}_k}{\check{\alpha}_1} \tag{3-6.3}$$

and upon division of the second equation by the first and setting k=1,

$$\frac{F(\vec{e}_1)}{F(\vec{0})} = \check{\alpha}_1\ \check{\lambda}e^{-\check{\alpha}} \tag{3-6.4}$$

By taking the logarithm of the first equation of (3-6.2) and substituting for the $\alpha_k$ in $\alpha = \alpha_1 + \ldots + \alpha_k$ from (3-6.3) we obtain

$$(1/\check{\lambda})\log\frac{F(\vec{0})}{\beta} = \exp\left\{-\frac{\check{\alpha}_1}{F(\vec{e}_1)}\sum_{k=1}^{n-1} F(\vec{e}_k)\right\} -1 \tag{3-6.5}$$

We may solve (3-6.4) for $1/\check{\lambda}$ and then replace the $\check{\alpha}_k$ in $\check{\alpha}$ using (3-6.3) to obtain

$$1/\check{\lambda} = \frac{\check{\alpha}_1 F(\vec{0})}{F(\vec{e}_1)}\ \exp\left\{-\frac{\check{\alpha}_1}{F(\vec{e}_1)}\sum_{k=1}^{n-1} F(\vec{e}_k)\right\}$$

Let us now substitute this equation for $1/\check{\lambda}$ in (3-6.5). Then

$$\left\{ \frac{\check{\alpha}_1 F(\vec{0})}{F(\vec{e}_1)} \log \frac{F(\vec{0})}{\beta} - 1 \right\} \exp \left\{ - \frac{\check{\alpha}_1}{F(\vec{e}_1)} \sum_{k=1}^{n-1} F(\vec{e}_k) \right\} + 1 = 0$$

Because this equation does not lend itself easily to exact solution, it is best to try a numerical procedure to find $\check{\alpha}_1$. Then $\check{\alpha}_2, \ldots, \check{\alpha}_{n-1}$ may be found from (3-6.3) and $\lambda$ from (3-6.4).

# APPENDIX 1A

## OBTAINING AN EXPLICIT EXPRESSION FOR THE PROBABILITY FUNCTION

We will start with the probability generating function given by (1-2.1).

$$g(\vec{s}) = e^{\lambda(T^N-1)}$$

where $$T = \sum_{i=1}^{n-1} s_i p_i + p_n$$

Then $$D_k g(\vec{s}) = \lambda g(\vec{s})\, D_k(T^N) \qquad (1A\text{-}1)$$

For simplicity let $\lambda g(\vec{s}) = A$, $D_k(T^N) = B$. Then (1A-1) becomes $D_k g(\vec{s}) = AB$.

To find the higher order derivatives, let us resort to Leibnitz Rule which states that

$$D_{n-1}^{\,x_{n-1}}(AB) = \sum_{y_{n-1}=0}^{x_{n-1}} \binom{x_{n-1}}{y_{n-1}} D_{n-1}^{\,y_{n-1}}A \cdot D_{n-1}^{\,x_{n-1}-y_{n-1}}B$$

Now, we may differentiate termwise with respect to $s_{n-2}$ and obtain

$$D_{n-2}^{\,x_{n-2}} D_{n-1}^{\,x_{n-1}}(AB) = \sum_{y_{n-1}=0}^{x_{n-1}} \binom{x_{n-1}}{y_{n-1}} D_{n-2}^{\,x_{n-2}}[D_{n-1}^{\,y_{n-1}}A \cdot D_{n-1}^{\,x_{n-1}-y_{n-1}}B]$$

Use Leibnitz Rule on each term.

$$= \sum_{y_{n-1}=0}^{x_{n-1}} \binom{x_{n-1}}{y_{n-1}} \sum_{y_{n-2}=0}^{x_{n-2}} \binom{x_{n-2}}{y_{n-2}} [D_{n-2}^{\,y_{n-2}} D_{n-1}^{\,y_{n-1}}A] \cdot [D_{n-2}^{\,x_{n-2}-y_{n-2}} D_{n-1}^{\,x_{n-1}-y_{n-1}}B]$$

The sums are independent of each other and consequently we may reverse their order

$$= \sum_{y_{n-2}=o}^{x_{n-2}} \sum_{y_{n-1}=o}^{x_{n-1}} \binom{x_{n-2}}{y_{n-2}}\binom{x_{n-1}}{y_{n-1}} D_{n-2}^{\;y_{n-2}} D_{n-1}^{\;y_{n-1}} A$$

$$\cdot\ D_{n-2}^{\;x_{n-2}-y_{n-2}} D_{n-1}^{\;x_{n-1}-y_{n-1}} B$$

Continuing this process we find that, after replacing A and B by the expressions they represent,

$$D_1^{\;x_1} \ldots D_{n-1}^{\;x_{n-1}} D_k g(\vec{s}) = \sum_{y_1=o}^{x_1} \ldots \sum_{y_{n-1}=o}^{x_{n-1}} \binom{x_{n-1}}{y_{n-1}} \ldots \binom{x_1}{y_1}$$

$$\cdot\ D_{n-1}^{\;y_{n-1}} \ldots D_1^{\;y_1} [\lambda g(\vec{s})] D_{n-1}^{\;x_{n-1}-y_{n-1}} \ldots D_1^{\;x_1-y_1} D_k(T^N) \tag{1A-2}$$

The remaining problem is to find the above multiple derivative of $T^N$. Notice

$$D_k(T^N) = NT^{N-1}\, D_kT = NT^{N-1}\, p_k$$

Continuing,

$$D_1^{\;x_1-y_1}(D_kT^N) = Np_k\, D_1^{\;x_1-y_1}(T^{N-1})$$

$$= p_k\, N(N-1)(N-2)\,\ldots\,[N-(x_1-y_1)]T^{N-(x_1-y_1)-1}\, p_1^{\;x_1-y_1}$$

We may proceed in the same manner through the derivatives with respect to all the $s_i$. The resulting expression is

$$D_{n-1}^{x_{n-1}-y_{n-1}} \ldots D_1^{x_1-y_1} D_k(T^N)$$
$$= p_k\left(\prod_{i=1}^{n-1} p_i^{x_i-y_i}\right) N(N-1) \ldots [N-\sum_{i=1}^{n-1}(x_i-y_i)]T^{N-\sum_{i=1}^{n-1}(x_i-y_i)-1} \tag{1A-3}$$

From the definition of a probability generating function,

$$P_{\vec{x}}(\vec{x}+\vec{e}_k) = \left.\frac{D_{n-1}^{x_{n-1}} \ldots D_1^{x_1} D_k g(\vec{s})}{(x_k+1)\prod_{i=1}^{n-1} x_i!}\right|_{\vec{s}=\vec{0}}$$

and, substituting from (1A-2),

$$= \frac{\lambda}{(x_k+1)\prod_{i=1}^{n-1} x_i!} \sum_{y_1=0}^{x_{n-1}} \cdots \sum_{y_{n-1}=0}^{x_{n-1}} \prod_{i=1}^{n-1} \binom{x_i}{y_i} [D_{n-1}^{y_{n-1}} \ldots D_1^{y_1} g(\vec{s})]$$
$$\left. \cdot [D_{n-1}^{x_{n-1}-y_{n-1}} \ldots D_1^{x_1-y_1} D_k(T^N)]\right|_{\vec{s}=\vec{0}}$$

If we use (1A-3) to evaluate the second square bracket

$$= \frac{\lambda}{(x_k+1)\prod_{i=1}^{n-1} x_i!} \sum_{y_1=0}^{x_1} \cdots \sum_{y_{n-1}=0}^{x_{n-1}} \prod_{i=1}^{n-1} \binom{x_i}{y_i} [\prod_{i=1}^{n-1} y_i!\ P_{\vec{x}}(\vec{y})]$$
$$p_k[\prod_{i=1}^{n-1} p_i^{x_i-y_i}]\ N(N-1) \ldots [N-\sum_{i=1}^{n-1}(x_i-y_i)]p_n^{N-\sum_{i=1}^{n-1}(x_i-y_i)-1}$$

Finally, we can express the combinations as factorials and, after rearranging the terms, we arrive at

$$P_{\vec{x}}(\vec{x}+\vec{e}_k) = \frac{\lambda p_k p_n^{N-1}}{x_k+1} \sum_{y_1=0}^{x_1} \cdots \sum_{y_{n-1}=0}^{x_{n-1}} N(N-1)\ldots[N-\sum_{i=1}^{n-1}(x_i-y_i)]$$

$$[\prod_{i=1}^{n-1} (\frac{p_i}{p_n})^{x_i-y_i} \cdot \frac{1}{(x_i-y_i)!}] P_{\vec{x}}(\vec{y}) \tag{1A-4}$$

If, as in the case of the Poisson-multinomial distribution, $N$ is a positive integer, then

$$P_{\vec{x}}(\vec{x}+\vec{e}_k) = \begin{cases} \dfrac{\lambda p_k p_n^{N-1}}{x_k+1} \displaystyle\sum_{y_1=0}^{x_1} \cdots \sum_{y_{n-1}=0}^{x_{n-1}} \dfrac{N!}{[N-\sum_{i=1}^{n-1}(x_i-y_i)-1]!} \cdot \\ [\displaystyle\prod_{i=1}^{n-1} (\frac{p_i}{p_n})^{x_i-y_i} \cdot \frac{1}{(x_i-y_i)!}] P_{\vec{x}}(\vec{y}) \quad \text{if} \quad N > \sum_{i=1}^{n-1}(x_i-y_i) \\ 0 \qquad \text{otherwise} \end{cases}$$

# APPENDIX 1B

## CALCULATION OF MOMENTS FROM FACTORIAL CUMULANT GENERATING FUNCTION

From (1-3.1), the cumulant generating function for the Poisson-multinomial distribution is

$$c(\vec{s}) = \lambda \left\{ \left[ \sum_{i=1}^{n-1} s_i p_i + p_n \right]^N - 1 \right\}$$

For simplicity of notation, in the following calculations the expectations of products of $X_1$, $X_2$, $X_3$, $X_4$ will be calculated explicitely. It is clear, however that these results may be generalized to products of any $X_i$'s.

Let us now calculate all the partial derivatives of $c(\vec{s})$ with respect to the $s_i$ of order less than or equal to four. These are

1. $$\frac{\partial c}{\partial s_1} = \lambda N \left( \sum_{i=1}^{n-1} s_i p_i + p_n \right)^{N-1} p_1$$

2. $$\frac{\partial^2 c}{\partial s_1^2} = \lambda N(N-1) \left( \sum_{i=1}^{n-1} s_i p_i + p_n \right)^{N-2} p_1^2$$

3. $$\frac{\partial^3 c}{\partial s_1^3} = \lambda N(N-1)(N-2) \left( \sum_{i=1}^{n-1} s_i p_i + p_n \right)^{N-3} p_1^3$$

4. $$\frac{\partial^4 c}{\partial s_1^4} = \lambda N(N-1)(N-2)(N-3) \left( \sum_{i=1}^{n-1} s_i p_i + p_n \right)^{N-4} p_1^4$$

5. $$\frac{\partial^2 c}{\partial s_2 \partial s_1} = \lambda N(N-1) \left( \sum_{i=1}^{n-1} s_i p_i + p_n \right)^{N-2} p_1 p_2$$

6. $$\frac{\partial^3 c}{\partial s_2 \partial s_1^2} = \lambda N(N-1)(N-2)\left(\sum_{i=1}^{n-1} s_i p_i + p_n\right)^{N-3} p_1^2 p_2$$

7. $$\frac{\partial^4 c}{\partial s_2 \partial s_1^3} = \lambda N(N-1)(N-2)(N-3)\left(\sum_{i=1}^{n-1} s_i p_i + p_n\right)^{N-4} p_1^3 p_2$$

8. $$\frac{\partial^3 c}{\partial s_3 \partial s_2 \partial s_1} = \lambda N(N-1)(N-2)\left(\sum_{i=1}^{n-1} s_i p_i + p_n\right)^{N-3} p_1 p_2 p_3$$

9. $$\frac{\partial^4 c}{\partial s_3 \partial s_2 \partial s_1^2} = \lambda N(N-1)(N-2)(N-3)\left(\sum_{i=1}^{n-1} s_i p_i + p_n\right)^{N-4} p_1^2 p_2 p_3$$

10. $$\frac{\partial^4 c}{\partial s_2^2 \partial s_1^2} = \lambda N(N-1)(N-2)(N-3)\left(\sum_{i=1}^{n-1} s_i p_i + p_n\right)^{N-4} p_1^2 p_2^2$$

11. $$\frac{\partial^4 c}{\partial s_4 \partial s_3 \partial s_2 \partial s_1} = \lambda N(N-1)(N-2)(N-3)\left(\sum_{i=1}^{n-1} s_i p_i + p_n\right)^{N-4} p_1 p_2 p_3 p_4$$

To obtain the factorial cumulants corresponding to the above derivatives we set $\vec{s}=\vec{1}$. We denote the factorial cumulants by $K_{ijkm}$ where this symbol represents the $i^{th}$ cumulant with respect to $X_1$, $j^{th}$ with respect $X_2$, $k^{th}$ with respect to $X_3$, $m^{th}$ with respect to $X_4$. If the last subscripts are zero, they may be omitted (e.g. $K_{1100} = K_{110} = K_{11}$). Thus formulas 1-11 become respectively

1. $K_1 = N\lambda p_1$
2. $K_2 = N(N-1)\,\lambda p_1^2$
3. $K_3 = N(N-1)(N-2)\,\lambda p_1^3$

4. $K_4 = N(N-1)(N-2)(N-3)\ \lambda p_1^{\ 4}$

5. $K_{11} = N(N-1)\ \lambda p_1 p_2$

6. $K_{21} = N(N-1)(N-2)\ \lambda p_1^{\ 2} p_2$

7. $K_{31} = N(N-1)(N-2)(N-3)\ \lambda p_1^{\ 3} p_2$

8. $K_{111} = N(N-1)(N-2)\ \lambda p_1 p_2 p_3$

9. $K_{211} = N(N-1)(N-2)(N-3)\ \lambda p_1^{\ 2} p_2 p_3$

10. $K_{22} = N(N-1)(N-2)(N-3)\ \lambda p_1^{\ 2} p_2^{\ 2}$

11. $K_{1111} = N(N-1)(N-2)(N-3)\ \lambda p_1 p_2 p_3 p_4$

Now define $E(X_1^{\ (r_1)} \ldots X_t^{\ (r_t)})$ to be the $r_1^{\ th}$ factorial moment with respect to $X_1$, ..., and the $r_t^{\ th}$ factorial moment with respect to $X_t$. Then, using tables converting factorial cumulants to factorial moments, we have if we define the $G_i$ by (1-3.2)

1. $E(X_1) \quad = \quad K_1 \quad = \quad N\lambda p_1$

2. $E(X_1^{\ (2)}) = K_2 + K_1^{\ 2} = N\lambda p_1^{\ 2}[N(\lambda+1)-1]$

$\qquad\qquad = N\lambda p_1^{\ 2} G_1$

3. $E(X_1 X_2) = K_{11} + K_{10}K_{01} = N\lambda p_1 p_2 [N(\lambda+1)-1]$

$\qquad\qquad = N\lambda p_1 p_2 G_1$

4. $E(X_1^{\ (3)}) = K_3 + 3K_2K_1 + K_1^{\ 3}$

$\qquad\qquad = N\lambda p_1^{\ 3}[N^2(\lambda^2+3\lambda+1)-3N(\lambda+1)+2]$

$\qquad\qquad = N\lambda p_1^{\ 3} G_2$

$$5.\quad E(X_1^{(2)}X_2) = K_{21}+K_{20}K_{01}+2K_{11}K_{10}+K_{10}^{\,2}K_{01}$$

$$= N\lambda p_1^{\,2}p_2G_2$$

$$6.\quad E(X_1X_2X_3) = K_{111}+K_{110}K_{001}+K_{101}K_{010}+K_{011}K_{100} + K_{100}K_{010}K_{001}$$

$$= N\lambda p_1p_2p_3G_2$$

$$7.\quad E(X_1^{(4)}) = K_4 + 4K_3K_1 + 3K_2^{\,2} + 6K_2K_1^{\,2} + K_1^{\,4}$$

$$= N\lambda p_1^{\,4}[N^3(\lambda^3+6\lambda^2+7\lambda+1)-6N^2(\lambda^2+3\lambda+1) + 11N(\lambda+1)-6]$$

$$= N\lambda p_1^{\,4}\,G_3$$

$$8.\quad E(X_1^{(3)}X_2) = K_{31} + 3K_{21}K_{10} + K_{30}K_{01} + 3K_{11}K_{20} + 3K_{20}K_{10}K_{01} + 3K_{11}K_{10}^{\,2} + K_{10}^{\,3}K_{01}$$

$$= N\lambda p_1^{\,3}p_2\,G_3$$

$$9.\quad E(X_1^{(2)}X_2^{(2)}) = K_{22} + 2K_{12}K_{10} + 2K_{21}K_{01} + 2K_{11}^{\,2} + K_{02}K_{20} + K_{02}K_{10}^{\,2} + 4K_{11}K_{10}K_{01} + K_{20}K_{01}^{\,2} + K_{10}^{\,2}K_{01}^{\,2}$$

$$= N\lambda p_1^{\,2}p_2^{\,2}\,G_3$$

$$10.\quad E(X_1^{(2)}X_2X_3) = K_{211} + 2K_{111}K_{100} + K_{210}K_{001} + K_{201}K_{010} + K_{011}K_{200} + 2K_{110}K_{101} + K_{011}K_{100}^{\,2} + 2K_{110}K_{100}K_{001} + 2K_{101}K_{100}K_{010} + K_{200}K_{010}K_{001} + K_{100}^{\,2}K_{010}K_{001}$$

$$= N\lambda p_1^{\,2}p_2p_3\,G_3$$

$$11. \quad E(X_1X_2X_3X_4) = K_{1111} + K_{1110}K_{0001} + K_{1101}K_{0010} + K_{1011}K_{0100}$$
$$+ K_{0111}K_{1000} + K_{1100}K_{0011} + K_{1010}K_{0101}$$
$$+ K_{1001}K_{0110} + K_{1100}K_{0010}K_{0001}$$
$$+ K_{1010}K_{0100}K_{0001} + K_{1001}K_{0100}K_{0010}$$
$$+ K_{0110}K_{1000}K_{0001} + K_{0101}K_{1000}K_{0010}$$
$$+ K_{0011}K_{1000}K_{0100} + K_{1000}K_{0100}K_{0010}K_{0001}$$
$$= N\lambda p_1p_2p_3p_4\ G_3$$

Now it is an easy matter to find the moments about the origin. Equations 1, 3, 6, and 11 need no change. The others need slight modification which is done as follows -

$$2'. \quad E(X_1^2) = E(X_1^{(2)}) + E(X_1) = N\lambda p_1 \left\{ [N(\lambda+1)-1]p_1 + 1 \right\}$$
$$= N\lambda p_1 ( p_1G_1 + 1)$$

$$4'. \quad E(X_1^3) = E(X_1^{(3)}) + 3E(X_1^{(2)}) + E(X_1)$$
$$= N\lambda p_1(p_1^2G_2 + 3p_1G_1 + 1)$$

$$5'. \quad E(X_1^2X_2) = E(X_1^{(2)}X_2) + E(X_1X_2)$$
$$= N\lambda p_1p_2(p_1G_2 + G_1)$$

$$7'. \quad E(X_1^4) = E(X_1^{(4)}) + 6E(X_1^{(3)}) + 7E(X_1^{(2)}) + E(X_1)$$
$$= N\lambda p_1(p_1^3G_3 + 6p_1^2G_2 + 7p_1G_1 + 1)$$

$$8'. \quad E(X_1^3X_2) = E(X_1^{(3)}X_2) + 3E(X_1^{(2)}X_2) + E(X_1X_2)$$
$$= N\lambda p_1p_2\ (p_1^2G_3 + 3p_1G_2 + G_1)$$

9'. $E(X_1^{\,2}X_2^{\,2}) \;= E(X_1^{(2)}X_2^{(2)}) + E(X_1^{(2)}X_2) + E(X_2^{(2)}X_1) + E(X_1X_2)$

$$= N\lambda p_1 p_2\,[p_1 p_2 G_3 + (p_1+p_2)G_2 + G_1]$$

10'. $E(X_1^{\,2}X_2X_3) \;= E(X_1^{(2)}X_2X_3) + E(X_1X_2X_3)$

$$= N\lambda p_1 p_2 p_3 (p_1 G_3 + G_2)$$

Formulas 1, 3, 6 and 11 together with the primed formulas give expressions for the moments of order four or less. These, in a slightly generalized form are summarized in (1-3.4).

# APPENDIX 1C

## CALCULATION OF THE ENTRIES OF THE INFORMATION MATRIX "J"

The calculation of $I_{\lambda\lambda}$ is outlined in § 1.5B.

Consider $I_{\lambda p_j}$ and $I_{p_j\lambda}$

$$I_{\lambda p_j} = I_{p_j\lambda} = E\left(\frac{\partial}{\partial\lambda}\log L \ \frac{\partial}{\partial p_j}\log L\right) \quad \text{by definition.}$$

Differentiating after substituting from (1-4.1) yields

$$I_{\lambda p_j} = E\left\{ \left[\sum_{\alpha=1}^{\beta} \frac{1}{P_{\vec{x}}(\vec{x}_\alpha)} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}_\alpha)\right]\left[\sum_{\gamma=1}^{\beta} \frac{1}{P_{\vec{x}}(\vec{x}_\gamma)} \frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x}_\gamma)\right]\right\}$$

$$= \sum_{\alpha=1}^{\beta}\sum_{\gamma=1}^{\beta} E\left[\frac{1}{P_{\vec{x}}(\vec{x}_\alpha)} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}_\alpha) \cdot \frac{1}{P_{\vec{x}}(\vec{x}_\gamma)} \frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x}_\gamma)\right]$$

Because the observations are independent and identically distributed

$$I_{\lambda p_j} = \beta\ E\left\{\frac{1}{[P_{\vec{x}}(\vec{x})]^2} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}) \cdot \frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x})\right\}$$

$$+ \beta(\beta-1)\ E\left[\frac{1}{P_{\vec{x}}(\vec{x})} \frac{\partial}{\partial\lambda} P_{\vec{x}}(\vec{x}_\alpha)\right] E\left[\frac{1}{P_{\vec{x}}(\vec{x})} \frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x}_\alpha)\right] \qquad (1C\text{-}1)$$

The second term is zero by the same reasoning as in (1-5.8). From the definition of expectation, and substituting from (1-4.4) and (1-4.7) for the derivatives,

$$(1/\beta)\, I_{\lambda p_j} = \sum_{x_1=0}^{\infty} \cdots \sum_{x_{n-1}=0}^{\infty} \frac{1}{P_{\vec{x}}(\vec{x})} \left\{ \frac{p_n(x_m+1)}{p_m N\lambda} P_{\vec{x}}(\vec{x}+\vec{e}_m) \right.$$

$$\left. + (1/N\lambda) \sum_{k=1}^{n-1} x_k\, P_{\vec{x}}(\vec{x}) - P_{\vec{x}}(\vec{x}) \right\} \left\{ (x_j/p_j) P_{\vec{x}}(\vec{x}) - \frac{x_j+1}{p_j} P_{\vec{x}}(\vec{x}+\vec{e}_j) \right\}$$

As is seen from (1-4.7) this holds for all $m$. If we choose $m=j$, multiply the expression out and replace the sums by corresponding expectations and use lemma 1-5 where necessary,

$$(1/\beta)\, I_{\lambda p_j} = \frac{p_n}{p_j^{\,2} N\lambda} E[X_j(X_j-1)] - p_n N\lambda(A+1) + (1/p_j N\lambda) E(X_j)$$

If we choose $m \neq j$ and follow the same procedure,

$$(1/\beta)\, I_{\lambda p_j} = \frac{p_n}{p_j p_m N\lambda} E(X_j X_n) - p_n N\lambda(A+1) + (1/p_j N\lambda) E(X_j)$$

In either case, we may substitute for the expectations using (1-3.3). Both will lead to the same result which is

$$(1/\beta)\, I_{\lambda p_j} = -\, p_n N\lambda A + p_n N + 1 - p_n \tag{1C-2}$$

Now consider $I_{p_i p_j}$. By definition $I_{p_i p_j} =$ $E[\partial/\partial p_i\ (\log L)\ .\ \partial/\partial p_j\ (\log L)]$. Let us now substitute from (1-4.1) and realize that the observations are independent and identically distributed. Then

$$I_{p_i p_j} = \beta\, E\Big[\frac{1}{[P_{\vec{x}}(\vec{x})]^2} \frac{\partial}{\partial p_i} P_{\vec{x}}(\vec{x})\ .\ \frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x})\Big]$$

$$+\ \beta(\beta-1)\, E\Big[\frac{1}{P_{\vec{x}}(\vec{x})} \frac{\partial}{\partial p_i} P_{\vec{x}}(\vec{x})\Big]\, E\Big[\frac{1}{P_{\vec{x}}(\vec{x})}\ .\ \frac{\partial}{\partial p_j} P_{\vec{x}}(\vec{x})\Big]$$

Reasoning as in (1-5.8), we see that the second term is zero. The substitution for the derivatives from (1-4.4) yields

$$(1/\beta)\, I_{p_i p_j} = \sum_{x_1=o}^{\infty} \cdots \sum_{x_{n-1}=o}^{\infty} \frac{1}{P_{\vec{X}}(\vec{x})} [(x_j/p_j) P_{\vec{X}}(\vec{x}) + \frac{x_j+1}{p_j} P_{\vec{X}}(\vec{x}+\vec{e}_j)]$$

$$[(x_i/p_i)\, P_{\vec{X}}(\vec{x}) + \frac{x_i+1}{p_i}\ P_{\vec{X}}(\vec{x}+\vec{e}_i)]$$

If $i=j$, we use the definition of expectation and expand the expression. Then

$$(1/\beta)\, I_{p_i p_i} = (1/p_i^{\ 2})[E(X_i^{\ 2}) - 2E(X_i(X_i-1)) + N^2\lambda^2 p_i^{\ 2}(A+1)]$$

and the use of (1-3.3) and simplification gives

$$(1/\beta)\, I_{p_i p_i} = N^2\lambda^2 A + N\lambda\ (1/p_i + 1-N) \tag{1C-3}$$

If $i \neq j$, we may use the same procedure to get

$$(1/\beta)\, I_{p_i p_j} = (1/p_i p_j)[-E(X_i X_j) + N^2\lambda^2 p_i p_j (A+1)]$$

$$= N^2\lambda^2 A + N\lambda(1-N) \tag{1C-4}$$

Equations 1C-1, 2, 3, and 4 give the remaining entries of the information matrix.

# APPENDIX 1D

## CALCULATION OF THE INVERSE OF THE MATRIX

$$J/\beta = \begin{pmatrix} Q & R & R & \cdots & R \\ R & S+W/p_1 & S & \cdots & S \\ R & S & S+W/p_2 & \cdots & S \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R & S & S & \cdots & S+W/p_{n-1} \end{pmatrix}$$

**Step 1. Find det $(J/\beta)$**

Let us perform the following elementary row operations on $J/\beta$.

1. Subtract column 2 from columns 3, 4, ..., n-1.
2. Subtract $(S/R)$ (column 1) from column 2.
3. Add $(p_k/p_1)$ [row $(k+1)$] to row 2 for $k=2, 3, \ldots, n-1$.
4. Add $(p_1/W)(QS/R-R)$ (row 2) to row 1.

These operations leave the determinant unchanged. Hence we will have the result

$$\det\left(\frac{J}{\beta}\right) = \begin{vmatrix} Q + \frac{p_1}{W}\left(\frac{QS}{R} - R\right)\left(R + R\sum_{k=2}^{n-1}\frac{p_k}{p_1}\right) & & & \bigcirc \\ * & \frac{W}{p_1} & & \\ \vdots & \bigcirc & \ddots & \\ * & & & W/p_{n-1} \end{vmatrix}$$

$$\det(J/\beta) = [Q+(QS-R^2)(1-p_n)/W]\prod_{i=1}^{n-1}(W/p_i) \qquad \text{(1D-1)}$$

Step 2. Find minors of matrix elements

Let $K_{\alpha\gamma}$ be the minor of the $(\alpha,\gamma)$ entry of $J/\beta$.

$$K_{\lambda\lambda} = \begin{vmatrix} S + W/p_1 & S & \cdots & S \\ S & S + W/p_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & S \\ S & \cdots & S & S + W/p_{n-1} \end{vmatrix}$$

Carry out the following elementary row operations.

1. Subtract column 1 from columns 2, 3, ..., n-1.
2. Add $(p_k/p_1)$ (row k) to row 1 for $k=2, 3, \ldots, n-1$.

The result is a lower triangular matrix whose determinant is

$$K_{\lambda\lambda} = [1 + S(1-p_n)/W] \prod_{i=1}^{n-1} (W/p_i) \tag{1D-2}$$

$$K_{\lambda p_i} = \begin{vmatrix} R & S + \frac{W}{p_1} & & & & & \\ & S & \ddots & & & S & \\ & & \ddots & S + \frac{W}{p_{i-1}} & & & \\ & & & S & S & & \\ & S & & & S + \frac{W}{p_{i+1}} & \ddots & \\ & & & & \ddots & \ddots & S \\ & & & & & & S + \frac{W}{p_{n-1}} \end{vmatrix}$$

Do the following operations.

1. Subtract row $i$ from all other rows.
2. Expand by column 1.

The result is a diagonal matrix whose determinant is

$$K_{\lambda p_i} = (-1)^{i+1}\, R \prod_{k \neq i} (W/p_k) \tag{1D-3}$$

$$\begin{array}{l l}
 & K_{p_i p_j} = \begin{vmatrix}
Q & R & R & \cdot & \cdot & \cdot & \cdot & \cdot & R \\
R & S+\frac{W}{p_1} & S & & & & & & \\
\cdot & & \ddots & S & & & S & & \\
\cdot & & & S+\frac{W}{p_{i-1}} & S & & & & \\
\cdot & & & S & S+\frac{W}{p_{i+1}} & & & & \\
\cdot & & & & S & \ddots & & & \\
\cdot & & S & & & \ddots & S+\frac{W}{p_{j-1}} & & \\
\text{row } j & \cdot & & & & & S & S & \ddots \\
\cdot & & & & & & & S+\frac{W}{p_{j+1}} & S \\
R & & & & & & & \ddots & S+\frac{W}{p_{n-1}}
\end{vmatrix}
\end{array}$$

column $i+1$ $\qquad$ $i \neq j$

Do the following operations

1. Subtract row $j$ from all other rows except row 1.
2. Subtract $(R/S)$ (row $j$) from row 1.
3. Expand by column $(i+1)$ and then by row 1.

$$K_{p_i p_j} = (-1)^{i+j+1}(QS-R^2)\prod_{k\neq i,j}(W/p_k) \tag{1D-4}$$

$$K_{p_i p_i} = \begin{vmatrix} Q & R & \cdot & \cdot & & R \\ R & S+\frac{W}{p_1} & & & & \\ & & \ddots & & \text{\Huge S} & \\ \cdot & & & S+\frac{W}{p_{i-1}} & & \\ \cdot & & \text{\Huge S} & & S+\frac{W}{p_{i+1}} & \\ & & & & & \ddots \\ R & & & & & S+\frac{W}{p_{n-1}} \end{vmatrix}$$

This will be the same as $\det(J/\beta)$ but with the expressions in $p_i$ missing.

Thus $$K_{p_i p_i} = [Q+(QS-R^2)(1-p_n-p_i)/W]\prod_{k\neq i}(W/p_k) \tag{1D-5}$$

From either (1-5.6) or (2-5.3) we know that $\hat{\Omega} = J^{-1}$, i.e. $\hat{\Omega} = (1/\beta)(J/\beta)^{-1}$. By a well known theorem in matrix theory, the elements of $\hat{\Omega}$ can be expressed as the cofactors of corresponding elements of $J^T$ divided by $\det J$. Hence from (1D-1), ..., (1D-5),

$$\left.\begin{aligned}
\text{var } \hat{\lambda} &= \frac{1}{\beta} \cdot \frac{K_{\lambda\lambda}}{\det(J/\beta)} = \frac{1}{\beta} \cdot \frac{W+S(1-p_n)}{WQ+(QS-R^2)(1-p_n)} \\
\text{cov } (\hat{\lambda},\hat{p}_i) &= \frac{1}{\beta} \; \frac{K_{p_i\lambda}}{\det(J/\beta)} = -\frac{1}{\beta} \; \frac{Rp_i}{WQ+(QS-R^2)(1-p_n)} \\
\text{cov } (\hat{p}_i,\hat{p}_j) &= \frac{1}{\beta} \cdot \frac{K_{p_jp_i}}{\det(J/\beta)} = -\frac{1}{\beta} \; \frac{(QS-R^2)p_ip_j}{[WQ+(QS-R^2)(1-p_n)]W} \\
\text{var } \hat{p}_i &= \frac{1}{\beta} \cdot \frac{K_{p_ip_i}}{\det(J/\beta)} = -\frac{1}{\beta} \left[\frac{p_i}{W} - \frac{(QS-R^2)p_i{}^2}{[WQ+(QS-R^2)(1-p_n)]W}\right]
\end{aligned}\right\} \quad \text{(1D-6)}$$

# APPENDIX 1E

**Lemma:**

Let $\vec{\omega} = (\omega_1, \ldots, \omega_n)$ and $\vec{\mu} = (\mu_1, \ldots, \mu_n)$ be two sets of random variables related by the equations $\omega_i = \omega_i(\vec{\mu})$, $i = 1, 2, \ldots, n$. Let $\overline{\omega}_i$ and $\overline{\mu}_i$ denote the expected values of $\omega_i$ and $\mu_i$ respectively, and $\Omega = (\Omega_{ij})$ and $M = (M_{ij})$ denote the covariance matrices of $\vec{\omega}$ and $\vec{\mu}$ respectively. Also, denote by $\phi$, the Jacobian

$$\frac{\partial(\vec{\omega})}{\partial(\vec{\mu})} = \left|\left(\frac{\partial \omega_i}{\partial \mu_j}\bigg|_{\vec{\omega}=\vec{\overline{\omega}}}\right)\right|$$

Then if higher order derivatives of $\omega_i$ with respect to $\mu_j$ are negligable compared to the first order, we have $\Omega = \phi^T M \phi$.

**Proof:**

Using Taylor's theorem,

$$\omega_i - \overline{\omega}_i = \sum_{k=1}^{n} \frac{\partial \omega_i}{\partial \mu_k}\bigg|_{\vec{\omega}=\vec{\overline{\omega}}} (\mu_k - \overline{\mu}_k) + \text{second order terms}$$

Thus

$$E[(\omega_i - \overline{\omega}_i)(\omega_j - \overline{\omega}_j)] = \sum_{k=1}^{n} \sum_{m=1}^{n} \frac{\partial \omega_i}{\partial \mu_k} \frac{\partial \omega_j}{\partial \mu_m} E[(\mu_k - \overline{\mu}_k)(\mu_m - \overline{\mu}_m)]$$

where the derivatives are evaluated at $\vec{\omega} = \vec{\overline{\omega}}$. From the definitions of $\Omega$ and $M$,

$$\Omega_{ij} = \sum_{k=1}^{n} \sum_{m=1}^{n} \frac{\partial \omega_i}{\partial \mu_k} \frac{\partial \omega_j}{\partial \mu_m} M_{km}$$

$$= (\phi^T M \phi)_{ij}$$

Thus each entry of $\Omega$ agrees with each entry of $\phi^T M\phi$.

Q.E.D.

By a well known theorem from matrix theory, the lemma implies $\det \Omega = \det \phi^T . \det M . \det \phi$, or

$$\det \bar{\Omega} = (\det \phi)^2 \det M.$$

# APPENDIX 1F

## CALCULATION OF (1-6.12) AND (2-6.5)

Let us substitute (1-3.3) and (1-3.4) into (1-6.11).

$$\mathrm{cov}\,(\overline{W^2}, X_k) = (1/\beta)\Big\{ p_k N\lambda (p_k{}^2 G_2 + 3p_k G_1 + 1)$$

$$+ 2\sum_{j\neq k} N\lambda p_k p_j (p_k G_2 + G_1) + \sum_{i\neq k}\ \sum_{j\neq k,i} N\lambda\, G_2 p_k p_i p_j$$

$$+ \sum_{i\neq k} N\lambda p_i p_k\, (p_i G_2 + G_1) - N\lambda p_k \sum_{i=1}^{n-1} N\lambda p_i (p_i G_1 + 1)$$

$$- N\lambda p_k \sum_{i=1}^{n-1}\ \sum_{j\neq i} N\lambda G_1 p_i p_j \Big\}$$

After carrying out the summations we obtain

$$= (1/\beta)\Big\{ p_k N\lambda (p_k{}^2 G_2 + 3p_k G_1 + 1) + 2N\lambda p_k (p_k G_2 + G_1)(1-p_k-p_n)$$

$$+ N\lambda p_k G_2 [(1-p_k-p_n)^2 - \sum_{i\neq k} p_i{}^2] + N\lambda p_k [G_2 \sum_{i\neq k} p_i{}^2 + G_1 (1-p_k-p_n)$$

$$- N^2\lambda^2 p_k (G_1 \sum_{i=1}^{n-1} p_i{}^2 + 1 - p_n) - N^2\lambda^2 G_1 p_k [(1-p_n)^2 - \sum_{i=1}^{n-1} p_i{}^2]$$

Upon simplification, this yields

$$\mathrm{cov}\,(\overline{W^2}, X_k) = (1/\beta) N\lambda p_k\ \Big\{ (1-p_n)[G_2(1-p_n) + 3G_1$$

$$- N\lambda \left\langle G_1(1-p_n) + 1 \right\rangle ] + 1 \Big\}$$

This equation is (1-6.12).

# APPENDIX 1G

## CALCULATION OF (1-6.16) AND (2-6.6)

Let us substitute (1-3.3) and (1-3.4) into (1-6.15)

$$\mathrm{var}\ \overline{W}^2 = (1/\beta)\Big\{\sum_{i=1}^{n-1}\sum_{j\neq i}\sum_{k\neq i,j}\sum_{m\neq i,j,k} N\lambda G_3\, p_i p_j p_k p_m$$

$$+ 6\sum_{i=1}^{n-1}\sum_{j\neq i}\sum_{k\neq i,j} N\lambda(G_3 p_i^{\,2} p_j p_k + G_2 p_i p_j p_k) + 3\sum_{i=1}^{n-1}\sum_{j\neq i} N\lambda$$

$$[G_3 p_i^{\,2} p_j^{\,2} + G_2(p_i^{\,2} p_j + p_i p_j^{\,2}) + G_1 p_i p_j] + 4\sum_{i=1}^{n-1}\sum_{j\neq i} N\lambda$$

$$(G_3 p_i^{\,3} p_j + 3G_2 p_i^{\,2} p_j + G_1 p_i p_j) + \sum_{i=1}^{n-1} N\lambda(G_3 p_i^{\,4} + 6G_2 p_i^{\,2} + 7G_1 p_i + p_i)$$

$$-\Big[\sum_{i=1}^{n-1} N\lambda p_i(p_i G_1 + 1) + \sum_{i=1}^{n-1}\sum_{j\neq i} N\lambda G_1 p_i p_j\Big]^2\Big\}$$

After carrying out the summations we obtain

$$=(1/\beta)\Big\{N\lambda G_3[(1-p_n)^4 - 6(1-p_n)^2\sum_{k=1}^{n-1} p_k^{\,2} + 3(\sum_{k=1}^{n-1} p_k^{\,2})^2 + 8(1-p_n)\sum_{i=1}^{n-1} p_i^{\,3}$$

$$-6\sum_{i=1}^{n-1} p_i^{\,4}] + 6N\lambda[G_3\langle(1-p_n)^2\sum_{i=1}^{n-1} p_i^{\,2} - 2(1-p_n)\sum_{i=1}^{n-1} p_i^{\,3} - (\sum_{j=1}^{n-1} p_j^{\,2})^2$$

$$+2\sum_{i=1}^{n-1} p_i^{\,4}\rangle + G_2\langle(1-p_n)^3 - 3(1-p_n)\sum_{j=1}^{n-1} p_j^{\,2} + 2\sum_{i=1}^{n-1} p_i^{\,3}\rangle]$$

$$+3N\lambda[G_3\langle(\sum_{i=1}^{n-1} p_i^{\,2})^2 - \sum_{i=1}^{n-1} p_i^{\,4}\rangle + G_2\langle 2(1-p_n)\sum_{j=1}^{n-1} p_j^{\,2} - 2\sum_{i=1}^{n-1} p_i^{\,3}\rangle$$

$$+G_1\langle(1-p_n)^2 - \sum_{i=1}^{n-1} p_i^{\,2}\rangle] + 4N\lambda[G_3\langle(1-p_n)\sum_{i=1}^{n-1} p_i^{\,3} - \sum_{i=1}^{n-1} p_i^{\,4}\rangle$$

$$+3G_2\langle(1-p_n)\sum_{i=1}^{n-1}p_i{}^2-\sum_{i=1}^{n-1}p_i{}^3\rangle+G_1\langle(1-p_n)^2-\sum_{i=1}^{n-1}p_i{}^2\rangle]$$

$$+N\lambda(G_3\sum_{i=1}^{n-1}p_i{}^4+6G_2\sum_{i=1}^{n-1}p_i^3+7G_1\sum_{i=1}^{n-1}p_i{}^2+1-p_n)$$

$$-[N\lambda\langle G_1\sum_{i=1}^{n-1}p_i{}^2+1-p_n\rangle+N\lambda G_1\langle(1-p_n)^2-\sum_{i=1}^{n-1}p_i{}^2\rangle]^2\}$$

Upon simplification, this yields

$$\operatorname{var}\overline{W}^2=(1/\beta)N\lambda(1-p_n)\left\{G_3(1-p_n)^3+6G_2(1-p_n)^2+7G_1(1-p_n)\right.$$

$$\left.+1-N\lambda(1-p_n)[G_1(1-p_n)+1]^2\right\}$$

This equation is (1-6.16)

# APPENDIX 1H

## CALCULATION OF det M WHERE M IS GIVEN BY

$$
M=\frac{1}{\beta}\begin{pmatrix}
R(1-p_n)H_2 & Rp_1H_1 & & Rp_{n-1}H_1 \\
Rp_1H_1 & Rp_1(p_1N'+1) & RN'p_1p_2 \cdots & RN'p_1p_{n-1} \\
Rp_2H_1 & RN'p_1p_2 & \ddots & \vdots \\
\vdots & & & RN'p_{n-2}p_{n-1} \\
Rp_{n-1}H_1 & RN'p_1p_{n-1} & \cdots & Rp_{n-1}(p_{n-1}N'+1)
\end{pmatrix}
$$

Let us perform the following elementary operations.

1. Take the common factor $R$ out of each row, the common factor $H_1$ out of row 1 and column 1, and the common factor $p_i$ out of row $i+1$, $i=1, 2, \ldots, n-1$.
2. Multiply row 1 by $-N'$ and add to rows $2, 3, \ldots, n$.
3. Multiply row $i+1$ by $-p_i$ and add to row 1, $i=1, 2, \ldots, n-1$.

We now have

$$
\det M = \left(\frac{R}{\beta}\right)^n H_1^{\,2}\left(\prod_{i=1}^{n-1} p_i\right)\begin{vmatrix}
\frac{(1-p_n)H_2}{H_1^{\,2}}-\left[1-\frac{(1-p_n)H_2N'}{H_1^{\,2}}\right](1-p_n) & & 0 & \\
* & 1 & & \\
\vdots & & \ddots & \\
* & 0 & & 1
\end{vmatrix}
$$

$$
=(R/\beta)^n\left(\prod_{i=1}^{n-1} p_i\right)(1-p_n)\left\{H_2-H_1^{\,2}(1-p_n)H_2N'\right\}
$$

# APPENDIX 2A

## OBTAINING THE PROBABILITY GENERATING FUNCTION $g^*(\vec{s})$

Let us start with (2-2.5$\frac{1}{2}$). This expression is equal to

$$g^*(\vec{s}) = e^{-\lambda} \sum_{z=1}^{\infty} (\lambda^z/z!) p_n^{Nz} \sum_{x_{n-1}=0}^{\infty} \binom{-Nz}{x_{n-1}} (-s_{n-1}p_{n-1})^{x_{n-1}} \cdots$$

$$\sum_{x_2=0}^{\infty} \binom{-x_3-\ldots-Nz}{x_2} (-s_2p_2)^{x_2} \sum_{x_1=0}^{\infty} \binom{-x_2-\ldots-Nz}{x_1} (-s_1p_1)^{x_1} + e^{-\lambda}$$

We can now apply (2-2.5) to the last sum and obtain

$$g^*(\vec{s}) = \sum_{z=1}^{\infty} (\lambda^z/z!) p_n^{Nz} \sum_{x_{n-1}=0}^{\infty} \binom{-Nz}{x_{n-1}} (-s_{n-1}p_{n-1})^{x_{n-1}} \cdots$$

$$\sum_{x_2=0}^{\infty} \binom{-x_3-\ldots-Nz}{x_2} (-s_2p_2)^{x_2} (1-s_1p_1)^{-x_2-\ldots-Nz} + e^{-\lambda}$$

$$= \cdots \sum_{x_2=0}^{\infty} \binom{-x_3-\ldots-Nz}{x_2} \left(\frac{-s_2p_2}{1-s_1p_1}\right)^{x_2} (1-s_1p_1)^{-x_2-\ldots-Nz} + e^{-\lambda}$$

Applying (2-2.5) again and continuing in this manner

$$= e^{-\lambda} \sum_{z=1}^{\infty} (\lambda^z/z!) p_n^{Nz} (1-s_1p_1-\ldots-s_{n-1}p_{n-1})^{-Nz} + e^{-\lambda}$$

$$= e^{-\lambda} \sum_{z=0}^{\infty} (\lambda^z/z!) p_n^{Nz} (1-s_1p_1-\ldots-s_{n-1}p_{n-1})^{-Nz}$$

The sum is simply the power series expansion of

$$\exp\left\{ \lambda \left( \frac{p_n}{1-\sum_{i=1}^{n-1} s_i p_i} \right)^N - \lambda \right\}$$

from which (2-2.6) follows.

## APPENDIX 2B

## CALCULATION OF THE ENTRIES OF THE INFORMATION MATRIX J*

The calculation of $I_{\lambda\lambda}{}^*$ is done in §2.5B.

Consider $I_{\lambda p_j}{}^*$ and $I_{p_j\lambda}{}^*$. By the same argument as in Appendix 1C, we may obtain the same equation as (1C-1), and as in (1-5.8), second term will be zero, i.e.

$$I_{\lambda p_j}{}^* = I_{p_j\lambda}{}^* = \beta E\left\{\frac{1}{[P_{\vec{X}}(\vec{x})]^2}\ \frac{\partial}{\partial\lambda}P_{\vec{X}}(\vec{x}).\frac{\partial}{\partial p_j}\ P_{\vec{X}}(\vec{x})\right\}$$

Substitution of (2-4.2) and (2-4.3) into the above equation results in

$$(1/\beta)I_{\lambda p_i}{}^*=\sum_{x_1=o}^{\infty}\cdots\sum_{x_{n-1}=o}^{\infty}\ \frac{1}{P_{\vec{X}}(\vec{x})}\left\{\frac{x_m+1}{N\lambda p_m}\ P_{\vec{X}}(\vec{x}+\vec{e}_m)-[(1/N\lambda)\sum_{j=1}^{n-1}x_j\right.$$
$$\left.+1]P_{\vec{X}}(\vec{x})\right\}\cdot\left\{[x_i/p_i+(1/p_n)\sum_{k=1}^{n-1}x_k]P_{\vec{X}}(\vec{x})-\frac{x_i+1}{p_np_i}\ P_{\vec{X}}(\vec{x}+\vec{e}_i)\right\}$$

(2B-1)

From (2-4.3), we see this must hold for any m, m=1, 2, ..., n-1. If we choose m=i, expand this expression, and make use of lemma 2-4, we can write the above equation in expectation notation as

$$(1/\beta)I_{\lambda p_i}{}^*=(1/p_i)[(-1+1/p_n)E(X_i)-(1/N\lambda)E(X_i\sum_{k=1}^{n-1}X_k)$$
$$+(1/N\lambda p_n)\left\{(2/p_i)E[X_i(\sum_{k=1}^{n-1}X_k-1)]-E[(\sum_{k=1}^{n-1}X_k)^2]\right\}$$

$$-(1/p_n)\sum_{k=1}^{n-1} E(X_k)-(N\lambda/p_n)(A+1)+(1/N\lambda)[E(X_i^2)-E(X_i)]$$

Similarly, if we choose $m\neq i$, we obtain

$$(1/\beta)I_{\lambda p_i}{}^*=(1/p_i)[(-1+1/p_n)E(X_i)-(1/N\lambda)E(X_i\sum_{k=1}^{n-1} X_k)$$
$$+(1/N\lambda p_m)\left\{(2/p_m)E[X_m(\sum_{k=1}^{n-1} X_k-1)]-E(\sum_{k=1}^{n-1} X_k)^2\right\}$$
$$-(1/p_n)\sum_{k=1}^{n-1} E(X_k)-(N\lambda/p_n)(A+1)+(1/N\lambda p_i p_m)E(X_iX_m)$$

Upon simplification by means of substitution from (2-3.3) of either of the above expressions, we find that

$$(1/\beta)I_{\lambda p_i}{}^*=N\lambda A/p_n+[(N\lambda+N+1)p_n{}^2-1/p_n-N\lambda]/p_n \tag{2B-2}$$

Next consider $I_{p_ip_j}$. From (2-5.5) and same procedure as above,

$$I_{p_ip_j}{}^* = \beta E[\frac{1}{[P_{\vec{X}}(\vec{x})]^2}\frac{\partial}{\partial p_i}P_{\vec{X}}(\vec{x})\cdot\frac{\partial}{\partial\lambda}P_{\vec{X}}(\vec{x})]$$

Let us substitute for the derivatives from (2-4.2) and expand

$$(1/\beta)I_{p_ip_j}{}^* = \sum_{x_1=o}^{\infty}\cdots\sum_{x_{n-1}=o}^{\infty}\frac{1}{P_{\vec{X}}(\vec{x})}\left\{[x_i/p_i+(1/p_n)\sum_{k=1}^{n-1} x_k]\,P_{\vec{X}}(\vec{x})\right.$$
$$\left.-\frac{x_i+1}{p_np_i}P_{\vec{X}}(\vec{x}+\vec{e}_i)\right\}\left\{[x_j/p_j+(1/p_n)\sum_{k=1}^{n-1} x_k]\,P_{\vec{X}}(\vec{x})\right.$$
$$\left.-\frac{x_j+1}{p_np_j}P_{\vec{X}}(\vec{x}+\vec{e}_j)\right\} \tag{2B-3}$$

If $i=j$, expansion and the definition of expectation yield

$$(1/\beta)I_{p_ip_i}{}^* = (1/p_i{}^2)(1-2/p_n)E(X_i{}^2)-(2/p_np_i)[(1/p_i$$

$$+1/p_n)E(X_i)+(1-1/p_n)E(X_i\sum_{k=1}^{n-1} X_k)+(1/p_n{}^2)E(\sum_{k=1}^{n-1} X_k)^2$$

$$+ (N^2\lambda^2/p_n{}^2)(A+1)$$

Substitution for the expectations using (2-3.3) will give us

$$(1/\beta)I_{p_ip_i}{}^* = N^2\lambda^2A/p_n{}^2 + (N\lambda/p_n{}^2)[-(N\lambda+N+1)/p_n{}^2 + 1/p_n$$

$$+N\lambda+1] + N\lambda/p_np_i \qquad (2B-4)$$

If $i\neq j$, (2B-4) yields

$$(1/\beta)I_{p_ip_j}{}^* = (1-2/p_n)E(X_iX_j)+(1/p_n)[(1/p_i)E(X_i\sum_{k=1}^{n-1} X_k)$$

$$+ (1/p_j)E(X_j\sum_{k=1}^{n-1} X_k)] + (1/p_n{}^2)\left\{ E(\sum_{k=1}^{n-1} X_k)^2\right.$$

$$- (1/p_j)E[X_j(\sum_{k=1}^{n-1} X_k-1)] - (1/p_i)E[X_i(\sum_{k=1}^{n-1} X_k-1)]$$

$$+ (N^2\lambda^2/p_n{}^2)(A+1)$$

$$= N^2\lambda^2A/p_n{}^2+(N\lambda/p_n{}^2)[-(N\lambda+N+1)/p_n{}^2+1/p_n+N\lambda+1] \qquad (2B-5)$$

Equations 2B-1, 2, 4 and 5 give the remaining entries of the information matrix.

# APPENDIX 2C

## CALCULATION OF det $\phi^*$ WHERE

$$\det \phi^* = \begin{vmatrix} \frac{-p_n}{Q} & \frac{F}{Q} & \frac{F}{Q} & \cdots & \frac{F}{Q} \\ \frac{p_n^2 p_1}{\lambda Q} & \frac{p_1 D}{\lambda Q} + \frac{p_n}{N\lambda} & \frac{p_1 D}{\lambda Q} & \cdots & \frac{p_1 D}{\lambda Q} \\ \frac{p_n^2 p_2}{\lambda Q} & \frac{p_2 D}{\lambda Q} & \frac{p_2 D}{\lambda Q} + \frac{p_n}{N\lambda} & \cdots & \frac{p_2 D}{\lambda Q} \\ \vdots & \vdots & & \ddots & \vdots \\ \frac{p_n^2 p_{n-1}}{\lambda Q} & \frac{p_{n-1} D}{\lambda Q} & \frac{p_{n-1} D}{\lambda Q} & \cdots & \frac{p_{n-1} D}{\lambda Q} + \frac{p_n}{N\lambda} \end{vmatrix}$$

To find the determinant we must perform the following elementary operations.

1. Factor $Q$ out of every row of the matrix, $p_n$ out of column 1, and finally $p_i/\lambda$ out of row $i+1$ for $i=1, 2, \ldots, n-1$.
2. Multiply column 1 by $F$ and add to the other columns.
3. Subtract row 2 from rows 3, 4, ..., n.
4. Multiply column $i+1$ by $p_i/p_1$ and add to column 2 for $i = 2, 3, \ldots, n-1$.
5. Expand by row 1 and then by column 1.

The result is a constant times the determinant of a diagonal matrix, and so

$$\det \phi^* = \frac{-Np_n^{\,n-1}}{Q^2(N\lambda)^{n-1}} \left[Qp_n/N + (D+Fp_n)(1-p_n)\right]$$

# APPENDIX 3A

## CALCULATION OF EFFICIENCY OF METHOD OF MOMENTS FOR THE POISSON-MULTIVARIATE POISSON DISTRIBUTION

Equation (3-4.7) ff. suggests two ways to find the efficiency. We may take the limit of the efficiency of either the Poisson Multinomial or the Poisson-Negative Multinomial distribution as $N \to \infty$, i.e. we may take limits in either (1-6.24) or (2-6.13).

Method 1. Take limiting value of (1-6.24).

After dividing numerator and denominator by $N$ and rearranging the expression slightly,

$$\underline{\text{Eff}} = \lim_{N \to \infty} \left\{ \frac{(1-1/N)^2 N^3 (1-p_n)^3}{[\lambda B_{\lambda\lambda} + N(1-p_n) \langle B_{\lambda\lambda}(B_{pp}/N^2) - (B_{\lambda p}/N)^2 \rangle ]} \right.$$

$$\left. \cdot \frac{1}{[H_2 \langle N(1-p_n) + p_n \rangle - H_1{}^2]} \right\}$$

Use (3-3.1), (3-3.2), and (3-3.3) to find the limit of the first factor in the denominator and (3-4.3) to find the limit of the second. Also use the fact that $\alpha = N(1-p_n)$. Then

$$\underline{\text{Eff}} = \alpha^3 / \left\{ [\lambda A + \alpha \langle A\lambda(\lambda A - 1) - (-\lambda A + 1)^2 \rangle ][\underline{H}_2 \langle \alpha + 1 \rangle - H_1{}^2] \right\}$$

This simplifies to

$$\underline{\text{Eff}} = \alpha^3 / \left\{ [\lambda A(\alpha+1) - \alpha][\underline{H}_2(\alpha+1) - \underline{H}_1{}^2] \right\} \qquad (3A-1)$$

Method 2. Take limiting value of (2-6.13).

After dividing numerator and denominator by N and rearranging the terms slightly,

$$\underline{\text{Eff}} = \lim_{N\to\infty} \left\{ \frac{(1+1/N)^4 N^5 (1-p_n)^5}{p_n^{\,2}[\lambda B_{\lambda\lambda}{}^* + \langle B_{\lambda\lambda}{}^*(B_{pp}{}^*/N^2) - (B_{\lambda p}{}^*/N)^2 \rangle p_n N(1-p_n)]} \right.$$

$$\left. \cdot \frac{1}{[(1+1/N)N(1-p_n)+D+Fp_n]^2[(H_2{}^* - H_1{}^{*2})p_n + (1-p_n)N(1+1/N)H_2{}^*]} \right\}$$

Let us note that from (2-6.10),

$$\lim_{N\to\infty} Fp_n = \lim_{N\to\infty} \left\{ 2(1-p_n)[N(\lambda+1)+1]p_n + p_n^{\,2} \right\} = 2(\lambda+1)\alpha + 1$$

$$\lim_{N\to\infty} D = \lim_{N\to\infty} \left\{ -N(1-p_n)(1+p_n) - 1 - 2\left(\frac{N}{N+1}\right) \lambda p_n N(1-p_n) \right.$$

$$\left. + \left(\frac{1}{N+1}\right) \lambda^2 N^2 (1-p_n)^2 + \left(\frac{N}{N+1}\right) \lambda (1-p_n)^2 \right\} = -2(\lambda+1)\alpha - 1$$

Hence we see that $\lim_{N\to\infty} (D+Fp_n) = 0$

If we use this fact and equations (3-3.1), (3-3.2), and (3-3.3) to find the limit of the second factor in the denominator and (3-4.3) to find the limit of the last factor

$$\underline{\text{Eff}} = \alpha^3 / \left\{ 1 \cdot [\lambda A + \langle A\lambda\ (\lambda A - 1) - (-\lambda A + 1)^2 \rangle \alpha] . \alpha^2 \right.$$

$$\left. \cdot [\underline{H}_2 - \underline{H}_1{}^2 + \alpha \underline{H}_2] \right\}$$

This simplifies to

$$\underline{\text{Eff}} = \alpha^3 / \left\{ [\lambda A(\alpha+1) - \alpha][\underline{H}_2(\alpha+1) - \underline{H}_1{}^2] \right\}$$

This is exactly the same as (3A-1).

# BIBLIOGRAPHY

[1] Anscombe, F.J., (1950), "Sampling Theory of the Negative Binomial and Logarithmic Series Distributions". Biometrika, Vol. 37, pp. 358-382.

[2] Barrett, M.E., (1962), "Stochastic Processes in Population Studies." An unpublished M.A. thesis at the University of British Columbia.

[3] Beall, G. and Rescia, R.R., (1953), "A Generalization of Neyman's Contagious Distributions." Biometrics, Vol. 9, pp. 354-386.

[4] Bliss, C.I., (1953), "Fitting the Negative Binomial Distribution to Biological Data." Biometrics, Vol. 9, pp. 176-200.

[5] David, F.N. and Barton, D.E., (1962), Combinatorial Chance. Hafner Publishing Company, New York.

[6] Douglas, J.B., (1955), "Fitting the Neyman Type A (two parameter) Contagious Distribution." Biometrics, Vol. 11, pp. 149-173.

[7] Feller, W., (1943), "On a General Class of Contagious Distributions." Ann. Math. Stat., Vol. 14, pp. 389-400.

[8] Feller, W., (1950), An Introduction to Probability Theory and its Applications, Vol. I, John Wiley and Sons, Inc., New York, London.

[9] Fisher, R.A., (1941-2), "The Negative Binomial Distribution." Ann. Eug., Vol. 11, pp. 182-187.

[10] Gurland, J., (1958), "A Generalized Class of Contagious Distributions." Biometrics, Vol. 14, pp. 229-249.

[11] Katti, S.K. and Gurland, J., (1961), "The Poisson Pascal Distribution." Biometrics Vol. 17, pp. 527-538.

[12] Kendall, M.G. and Stuart, A., (1958), The Advanced Theory of Statistics, Vol I. Hafner Publishing Company, New York.

[13] Kendall, M.G. and Stuart, A., (1961), The Advanced Theory of Statistics, Vol II. Hafner Publishing Company, New York.

[14] McGuire, J.U., Brindley, T.A., and Bancroft T.A., (1957), "The Distribution of European Corn Borer Larvae Pyrausta Nubilalis (HBN) in Field Corn." Biometrics, Vol. 13, pp. 65-78.

[15] Neyman, J., (1938), "On a New Class of 'Contagious' Distributions." Ann. Math. Stat., Vol. 9, pp. 35-57.

[16] Shumway, R., and Gurland, J., (1960), "Fitting the Poisson Binomial Distributions." Biometrics, Vol. 16, pp. 522-533.

[17] Skellam, J.G., (1952), "Studies in Statistical Ecology I. Spatial Pattern." Biometrika, Vol. 39, pp. 346-362.

[18] Sprott, D.A., (1958), "Method of Maximum Likelihood Applied to the Poisson Binomial Distribution." Biometrics, Vol. 14, pp. 97-106.

[19] Thomas, M., (1949), "A Generalization of Poisson's Binomial Limit for Use in Ecology." Biometrika, Vol. 36, pp. 18-25.

[20] Rao, C.R., (1945), "Information and the Accuracy Attainable in the Estimation of Statistical Parameters." Calcutta Math. Soc. Bull., Vol. 37, pp. 81-91.

[21] Rao, C.R., (1947) "Minimum Variance and the Estimation of Several Parameters." Proc. Camb. Phil. Soc., Vol. 43, pp. 280-283.