THE UNIVERSITY OF CALGARY

ESTIMATION PROBLEMS FOR SOME GENERALIZED DISCRETE PROBABILITY DISTRIBUTIONS

by

MOHAMED SHOUKRI

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE

STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

JANUARY, 1980

Revised March, 1980

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled

Estimation Problems for Some Generalized Discrete Probability Distributions

submitted by Mohamed Mohamed Shoukri in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

P. C. Consul

P. C. Consul, Chairman
Department of Mathematics & Statistics

Junjiro Ogawa

J. Ogawa
Department of Mathematics & Statistics

John R Collins

J. R. Collins
Department of Mathematics & Statistics

Nigel Michael Waters

N. M. Waters
Department of Geography

Bruce Johnston

Bruce Johnston, External Examiner
University of Manitoba

April 21, 1980

ABSTRACT

During the last three decades, the statistical literature was enriched by many of the generalized families and classes of discrete distributions, which proved to have many important applications in a wide variety of disciplines, such as biological and medical sciences, social sciences, physical sciences, engineering, operations research, and so on. The problem of estimation and the study of the structural properties of many of the generalized forms of discrete distributions, attracted the attention of many statisticians and research workers. In this present thesis, we discuss the problem of simultaneous estimation of the parameters of two of the *recently* generalized families that have many applications specially in the theory of queues, and we introduce a new class of bivariate generalized discrete distributions.

A presentation for some of the important classes of discrete probability distributions will be given in Chapter I, where we exhibit the scope of work constituting the subsequent chapters of this thesis. Chapter II deals with the simultaneous estimation of the two parameters of the generalized Poisson distribution. We study the asymptotic properties of the moment estimators as far as terms of order n^{-1} and n^{-2} in the biases, variances, and the covariance. Also, we give expressions for the first order terms in the moments of the maximum likelihood estimators, and compare the

performances of the two types of estimators with reference to the first order terms in the biases.

In Chapter III, we study the maximum likelihood and the moment estimators for two parameters of the generalized negative binomial distribution, which has three parameters. We derive the biases and variances of these estimators. Higher order terms in the biases, variances, and the covariance of the moment estimators will be given, and we determine the sample size required to ensure some stability in the behaviour of the moments of these estimators.

In Chapter IV, we introduce the bivariate generalization of the class of modified power series distributions, discussed by Gupta (21), under the title of "Bivariate Modified Power Series Distribution". In this chapter we give the recurrence relations among the moments, some examples of the generalized bivariate distributions, which belong in this class, and discuss some of their properties, applications, marginal and conditional distributions.

The problem of estimation of the parameters in the Bivariate Modified Power Series Distribution is taken up in Chapter V, where we discuss the maximum likelihood estimation. Moreover, we develop a theorem proving the necessary and sufficient conditions for the existence of a minimum variance unbiased estimator for a real valued parametric function of the parameters of the mentioned class.

ACKNOWLEDGEMENTS

I wish to express my deep and sincere gratitude to my supervisor, professor P.C. Consul, for introducing me, in the first place, to the subject of Discrete Distributions and for his close supervision of my graduate program. In particular, I am greatly indebted to him for his perpetual guidance and the valuable suggestions in the course of preparation of this thesis.

I would also like to thank the examiners for their detailed suggestions which have substantially improved the thesis.

I would like to extend my thanks to Mrs. Betty Teare, for her careful typing of the final manuscript and for making the revisions and corrections.

TABLE OF CONTENTS

CHAPTER I

SYSTEMS OF DISCRETE PROBABILITY DISTRIBUTIONS

1.0 Introduction.

The statistical distributions arose initially in connection with some specific situations, and once their relevance was established, there was little further interest in the theoretical analysis of these distributions, as they were mainly used for descriptive purposes. During the last quarter of the 19th century and the first quarter of the 20th century, the determination of sampling distributions of statistics based on random variables, and the study of various systems of distributions, with special reference to their use in model construction, had received a great deal of attention from theoretical as well as applied statisticians. This area deals with a large number of discrete probability distributions which may be classified as compound, mixed, modified, contagious, and generalized distributions. Trying to keep track of such a wide-ranging and rapidly expanding literature is rather a difficult task. Some account of these distributions can be found in the works of Gurland (23), Haight (26), Patil and Joshi (61), Krishnamoorthy (43), Neyman (53), Consul and Shenton (12), and the books by Johnson and Kotz (35), Mardia (49), and Ord (56).

There are many ways according to which statisticians classified systems of non-negative lattice distributions, and in the following sections, we shall view two such classifications. The first classification is based upon the existence of a recursive scheme among the successive probabilities $f_j, f_{j+1}, \ldots$, where $f_j = P(X=j)$, and in this respect, we present systems defined by difference equations. The other classification is based directly on the form of f_j as a function of j, and in this respect, we introduce the power series distributions.

1.1 Some Systems Defined by Difference Equations.

Pearson (63) noted that the successive probabilities for a hypergeometric probability distribution, denoted by f_j and f_{j+1}, satisfy the ratio $(f_{j+1}-f_j)/(f_{j+1}+f_j)$ = linear function of j/quadratic function of j. He used this as a base for obtaining (by a limiting process) the differential equation defining the Pearson system of continuous distribution functions. It may also be used as a basis for defining a system of discrete distributions. Moreover, it was realized by Guldberg (19) that if

$$f_{j+1} = \frac{K(j-a_1)\ \ldots\ (j-a_r)}{(j-b_1)\ \ldots\ (j-b_s)} f_j \quad , \tag{1.1}$$

then one can establish a recurrence relation among the moments. For the special case $r = 1$, $s = 2$, Ord (57) employed the difference equation

(1.2) $$\Delta f_{j-1} = \frac{(a-j)f_{j-1}}{b_0+b_1 j+b_2 j(j-1)}$$

to define a class of discrete distributions, based on a lattice of unit width. He noticed that the form of the density function will depend on the roots of the denominator in equation (1.2), and he used the values of the constant $k = (b_1-b_2-1)^2/(4b_2(b_2+2))$ as a criterion to distinguish between the distributions of this class. He also provided tables to summarize the distributions devised by using this technique and graphs for the corresponding regions of the (β_1,β_2) plane.

Katz (38) devoted his dissertation entirely to the investigation of the properties and sampling characteristics of the class of discrete probability distributions defined by the difference equation,

(1.3) $$\frac{f_{j+1}}{f_j} = \frac{\alpha+\beta j}{j+1} \quad , \qquad j = 0,1,2,\ldots \ .$$

The probability generating function (p.g.f.) of this class is given by

$$g(t) = \left(\frac{1-\beta t}{1-\beta}\right)^{-\alpha/\beta} \quad ,$$

and hence we have the special cases

(i) $\lim_{\beta\to 0} g(t) = e^{\alpha(t-1)}$, which is the p.g.f. of a Poisson distribution with mean α.

(ii) If $\beta < 0$, $\frac{-\alpha}{\beta} = n$ (n is a positive integer) and $\frac{-\beta}{1-\beta} = P$, then $g(t) = (1-P+Pt)^n$, which is the p.g.f. of a binomial distribution.

(iii) If $0 < \beta < 1$, $\frac{\alpha}{\beta} = n$, and $\frac{\beta}{1-\beta} = P$, then $g(t) = (1+P-Pt)^{-n}$, which is the p.g.f. of a negative binomial distribution.

1.2 Systems Associated with Series Expansions.

The fact that any power series of the form

$$\sum_{n=0}^{\infty} a_n t^n \quad , \tag{1.4}$$

is the probability generating function of a discrete probability distribution, provided that $a_n \geq 0$, $n = 0,1,2,\ldots$, and that they sum to unity, attracted the attention of many authors to the usefulness of series representation of classes of probability distributions. In this section we present some of the most important classes of discrete probability distributions associated with series expansions.

1.2.1 The General Dirichlet Series Distribution.

Using the general Dirichlet series defined as

$$f(\theta) = \sum_{x=1}^{\infty} a_x \exp(-\lambda_x \theta) \tag{1.5}$$

where $a_x \geq 0$, $\theta > 0$ and $\{\lambda_x\}$ is a sequence of real positive increasing numbers whose limit is infinity, Siromoney (35) defined the general Dirichlet series distribution with parameter θ by the probability function

$$P(X=x) = \frac{a_x \exp(-\lambda_x \theta)}{f(\theta)} , \qquad x = 1,2,\ldots \quad . \tag{1.6}$$

1.2.2 Power Series Distributions (PSD's).

The first form of a PSD was introduced by Tweedie (82) in 1947, where he was concerned about the Laplace transform of a more general form of a probability distribution defined as

$$(1.7) \qquad P(X=x) = a(x)\, e^{-x\theta}/f(\theta)$$

where

$$(1.8) \qquad f(\theta) = \sum_x a(x)\, e^{-x\theta} \quad .$$

In an attempt to characterize the Poisson distribution by the equality of the mean and the variance, Kosambi (42) visualized the PSD's as a class of discrete probability distributions.

The PSD was formally introduced by Noack (54) in 1950, as a mathematical model. A random variable X is said to have a PSD if it takes non-negative integral values with probability density function

$$(1.9) \qquad P(X=x) = a(x)\, \theta^x/f(\theta) \quad , \qquad x \in T$$

and zero otherwise, where T is the entire set of non-negative integers. Patil (59) defined the generalized power series distribution (GPSD) in the form as in (1.9), except that the random variable X is permitted to take values on a subset of the set T. Roy and Mitra (68) derived the minimum variance unbiased (MVU) estimator for the parametric function θ^r of the PSD from a complete sample or a left truncated sample, with known truncation points. The MVU estimators for the parameters of the PSD were studied extensively by Charalambides (7), when the sample is truncated on the left with unknown truncation points. The problem of deriving MVU estimators for doubly truncated samples have been considered by Joshi and Park (36).

1.2.3 Modified Power Series Distribution (MPSD).

A more general form of PSD's was given by Gupta (1974; (21)), by the name of MPSD, defined by a discrete random variable X with the probability distribution

$$(1.10) \qquad P(X=x) = a(x)(g(\theta))^{x}/f(\theta) \qquad x \in S \subseteq T$$

where $g(\theta)$ and $f(\theta)$ are positive, finite, and differentiable functions of θ and such that $f(\theta) = \sum_{x \in S} a(x)(g(\theta))^{x}$. Obviously, when $g(\theta) = \phi$ gives a unique value for $f(\theta)$ as a function of ϕ, the MPSD is reduced to a GPSD. The maximum likelihood (M.L.) estimator for θ and the first order terms in its bias and variance were given by Gupta (22). Kumar and Consul (44), evaluated the recurrence relations among the inverse moments of an MPSD, and used their results to find the exact bias of the M.L. estimator for θ, in some particular families of this class. Kumar (45) further extended his study and developed a theorem proving the necessary and sufficient conditions for the existence of an MVU estimator for a real valued parametric function of θ.

1.2.4 Lagrange Series Distributions.

Consul and Shenton (12) have introduced a new class of univariate discrete distributions under the title of Lagrangian distributions, on account of their relationship with the Lagrange's expansion of an inverse function. If $g(t)$ and $f(t)$ are two probability generating functions, then under the transformation

$$t = u \cdot g(t) \tag{1.11}$$

and within the circle of convergence, $f(t)$ can be expanded in powers of u, by Lagrange's expansion as

$$f(t) = f(0) + \sum_{j=1}^{\infty} \frac{u^j}{j!} \frac{\partial^{j-1}}{\partial t^{j-1}} \left[(g(t))^j \frac{\partial f(t)}{\partial t} \right] \bigg|_{t=0} . \tag{1.12}$$

Since, for $t = 1$, $u = 1$, then a probability distribution defined on a subset of non-negative integers will be given by

$$P(X=k) = \frac{1}{k!} \frac{\partial^{k-1}}{\partial t^{k-1}} \left[(g(t))^k \frac{\partial f(t)}{\partial t} \right] \bigg|_{t=0} . \tag{1.13}$$

The authors have further studied some interesting properties of the Lagrange distributions. The univariate families of Lagrange distributions include a large number of important discrete distributions such as the Borel-Tanner distribution (BTD), the generalized negative binomial distribution (GNBD) and the generalized Poisson distribution (GPD). Many of the Lagrange distributions belong to the MPSD class.

After the development of satisfactory systems, for use in the univariate case, it was only natural to extend them to bivariate and multivariate systems. The multivariate generalizations of important discrete distributions, their applications and some of their properties have been discussed by a number of scientists which include

Bhat and Kulkarni (3), Olkin and Sobel (55), Consul and Shenton (14), and many others.

1.2.5 Multivariate Power Series Distribution.

The multivariate generalization of the class of PSD was first introduced by Khatri ((1959; (40)). The n-dimensional random vector $\underline{x} = (x_1, x_2, \ldots, x_n)$ is said to have a multivariate PSD if its probability distribution function is given by

$$(1.14) \quad P(X_1=x_1, \ldots, X_n=x_n) = a(x_1, \ldots, x_n) \prod_{i=1}^{n} \theta_i^{x_i} / f(\theta_1, \ldots, \theta_n), \qquad \underline{x} \in T_n$$

and zero otherwise, where, T_n is the n-dimensional subspace of non-negative integers. He discussed the moments relations, and some characterization problems for this class. The estimation by the M.L. method and the problem of existence of an MVU was discussed by Patil (60).

Another sub-class of the multivariate PSD was discussed by Patil (62), with the name of sum-symmetric PSD's. He realized that some of the most important multivariate discrete models such as the multinomial, the negative multinomial, the multiple Poisson, and the multivariate logarithmic series distributions, have a common mathematical property in that they can be written as multivariate PSD's with sum-symmetric series functions. Patil defined the multivariate PSD with series function $f(\theta_1, \theta_2, \ldots, \theta_n)$ to be symmetric if $f(\theta_{i_1}, \theta_{i_2}, \ldots, \theta_{i_n}) = f(\theta_1, \theta_2, \ldots, \theta_n)$ for every permutation $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$. Further, he called the n-dimensional distribution (1.14) sum-symmetric if the series function is sum-symmetric, in that $f(\theta_1,\ldots,\theta_n) = u(\theta_1+\theta_2+\ldots+\theta_n)$ for some $u(\cdot)$. Thus, for the sum-symmetric PSD, we have the series function

$$f(\theta_1,\theta_2,\ldots,\theta_n) = u(\theta_1+\theta_2+\ldots+\theta_n) = \sum_{z=0}^{\infty} a(z)(\theta_1+\theta_2+\ldots+\theta_n)^z$$

$$= \sum_{z=0}^{\infty} a(z) \sum_{x_1+x_2+\ldots+x_n=z} z!\left(\prod_{i=1}^{n} x_i!\right)^{-1}\theta_1^{x_1}\theta_2^{x_1}\ldots\theta_n^{x_n} \tag{1.15}$$

$$= \sum \frac{(x_1+\ldots+x_n)!}{x_1!x_2!\ldots x_n!}\, a(x_1+x_2+\ldots+x_n)\,\theta_1^{x_1}\theta_2^{x_2}\ldots\theta_n^{x_n} \quad ,$$

and therefore the probability function of the sum-symmetric PSD is of the form

$$P(x_1,\ldots,x_n) = \frac{(x_1+x_2+\ldots+x_n)!}{x_1!x_2!\ldots x_n!}\, a(x_1+x_2+\ldots+x_n)\frac{\theta_1^{x_1}\theta_2^{x_2}\ \ldots\ \theta_n^{x_n}}{u(\theta_1+\theta_2+\ldots+\theta_n)} \; . \tag{1.16}$$

For this class of models, Patil discussed the moment recursions, partial sums and relative conditionals, marginals, conditional and regression properties, characterization by regression and other characteristic properties, multiple correlation coefficient, M.L. estimation and MVU estimation.

Our investigation in the present thesis is chiefly in two directions:

(i) We would like to extend the work done by different authors, in estimating one parameter of the GNBD and the GPD, to a more practical situation, where we consider the problem of simultaneous estimation for the two parameters of these families.

(ii) We introduce a bivariate generalization for the class of univariate MPSD defined in (1.10), under the name of a bivariate modified power series distribution (BMPSD). We study a number of properties of these distributions and obtain the MVU estimators of some functions of the parameters. This new class includes many of the well-known classical bivariate discrete distributions as well as some of the newly introduced bivariate statistical probability distributions. Many of the BMPSD families are related to the bivariate Lagrange distributions discussed by Shenton and Consul (74).

Before proceeding with the problem of estimation for the GPD and the GNBD, and to avoid repetition, we state a number of well known results as they are required for the derivation of many results.

1.3 Taylor's Expansion.

Widder (82) defines C^k as the class of those bivariate functions $f(x,y)$ such that all the partial derivatives of order k are continuous and gives the Taylor expansion for bivariate functions as follows:

Theorem (1.1): *If $f(x,y) \in C^{k+1}$ and (a,b) is any interior point in the domain of x and y, then for any non-negative integer k*

$$(1.17) \qquad f(x,y) = \sum_{j=0}^{k} \frac{1}{j!} \left\{ (x-a) \frac{\partial}{\partial a} + (y-b) \frac{\partial}{\partial b} \right\}^{j} f(a,b) + R_k \ ,$$

where

$$(1.18)\qquad R_k = \frac{1}{(k+1)!}\left\{(x-a)\frac{\partial}{\partial r} + (y-b)\frac{\partial}{\partial s}\right\}^{k+1} f(r,s)\ ,$$

r,s being replaced by $r = a+\theta(x-a)$, $s = b+\theta(y-b)$, and $0<\theta<1$ after differentiation.

As particular cases of the above theorem one gets the famous law of the mean for functions of two variables when $k=0$ and the following result when $k=1$

$$f(x,y) = f(a,b) + \alpha f_1(a,b) + \beta f_2\ (a,b) + \frac{1}{2}\{\alpha^2 f_{11}(a+\theta\alpha,\ b+\theta\beta) +$$

$$+ \alpha\beta f_{12}(a+\theta\alpha\ ,\ b+\theta\beta) + \beta^2 f_{22}(a+\theta\alpha\ ,\ b+\theta\beta)\}$$

where $\alpha = x-a$, $\beta = y-b$ and $0<\theta<1$.

The expansion (1.17) is extensively used for approximations of different orders by showing that the contribution made by the remainder term is of that order.

1.4 Regularity Conditions for Some Asymptotic Results.

Shenton and Bowman (75) have assumed the following four regularity conditions for the validity of the expressions for the asymptotic variances, covariances, and asymptotic biases of the maximum likelihood estimators derived by them for large samples.

1. the population consists of a denumerable set of classes,
2. the log likelihood function, $\log L \in C^5$, where C^5 is as defined in the previous section,
3. the partial derivatives of the first and second order of $\log L$ are bounded,
4. the random variable X has a range independent of the parameters to be estimated.

Other regularity conditions that are required to prove the efficiency, consistency and asymptotic normality of those estimators can be found in Cramer (16), and Rao (67). These regularity conditions are not listed as we are not seeking such optimum properties for our estimators.

1.5 Newton-Raphson Method of Iteration.

Let $x = f(x)$ be an equation where $f(x)$ satisfies the conditions

(i) $f(x)$ is continuous on $I = [a,b]$

(ii) $f(x) \in I$ for all $x \in I$

(iii) $|f(x_1) - f(x_2)| \leq L|x_1 - x_2|$ where L is a constant < 1 .

The condition (iii) is called the Lipschitz condition. It is proved in Henrici (32) that if the above conditions are satisfied the equation $x = f(x)$ has a unique solution which can be reached by an iterative algorithm. The proof of the following theorem for a function $F(x)$ of the form $F(x) = x - f(x)$ is outlined in (32).

Theorem (1.2): *Let the function $F(x)$, defined as above on a closed interval $[a,b]$, be twice differentiable and satisfy the following conditions:*

(i) $F(a)F(b) < 0$

(ii) $F'(x) \neq 0, \quad x \in [a,b]$

(iii) $F''(x) \geq 0 \quad$ *or* $\leq 0 \quad$ *for all* $x \in [a,b]$

(iv) $F'''(x)$ *exists and is continuous,*

then for any choice of $x_0 \in [a,b]$, *the sequence* $\{x_n\}$ *determined from the recurrence relation*

$$x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)} \quad , \qquad n = 0,1,2,\ldots$$

converges to the unique solution s *of* $F(x) = 0$.

CHAPTER II

ESTIMATION OF PARAMETERS FOR THE GENERALIZED POISSON DISTRIBUTION

2.1 Introduction.

The generalized Poisson distribution (GPD), is defined by the probability function

$$(2.1)\quad P_x(\lambda_1,\lambda_2) = P(X=x) = \begin{cases} \lambda_1(\lambda_1+\lambda_2 x)^{x-1}(x!)^{-1}\exp -(\lambda_1+\lambda_2 x) & , x=0,1,2,\ldots \\ 0 & \text{elsewhere} \end{cases}$$

where, $\lambda_1 > 0$ and $0 \le \lambda_2 < 1$.

Consul and Shenton (13) obtained the distribution by expanding the p.g.f. $f(t) = e^{\lambda_1(t-1)}$, using (1.12), under the transformation $t = ue^{\lambda_2(t-1)}$. Thus one can show that

$$e^{\lambda_1(t-1)} = \sum_{x=0}^{\infty} \frac{u^x}{x!}\lambda_1(\lambda_1+\lambda_2 x)^{x-1}\, e^{-(\lambda_1+\lambda_2 x)}$$

since $u = 1$, whenever $t = 1$ we get

$$1 = \sum_{x=0}^{\infty} \frac{\lambda_1}{x!} (\lambda_1+\lambda_2 x)^{x-1} e^{-(\lambda_1+\lambda_2 x)} \quad ,$$

which proves that the probabilities given by (2.1) sum to unity.

The distribution was formally introduced by Consul and Jain (10), with the conditions $|\lambda_2| < 1$ and $P_x(\lambda_1,\lambda_2) = 0$ for all $x \geq m$ when $\lambda_1+\lambda_2 m < 0$, and they had shown that the GPD provided a very close fit to many different types of data. Consul and Shenton (13) changed the condition $|\lambda_2| < 1$ to $0 < \lambda_2 < 1$, and it was pointed out by Nelson (52) that the conditions, given by Consul and Jain, would not make the probabilities (2.1), in general, sum up to unity as $P_x(\lambda_1,\lambda_2) = 0$ for all $x \geq m$ when $\lambda_1+\lambda_2 m < 0$.

In their paper about the properties and applications of the generalized Lagrange distribution, Consul and Shenton (13), have proved that in a single server queue, with constant service time, if the number of customers initiating the queue is a Poisson variate with probability generating function $f(t) = \exp[\lambda_1(t-1)]$, and if the input is Poissonian with probability generating function $g(t) = \exp[\lambda_2(t-1)]$, where the customers are served in the order of their arrival, then the probability distribution of the number of customers served before the queue first vanishes is given by (2.1). They have also shown that for the GPD, the cumulants satisfy the recurrence relation

(2.1.a) $$(1-\lambda_2)L_{r+1} = \lambda_2\partial_2(L_r) + \lambda_1\partial_1(L_r), \qquad r = 1,2,\ldots$$

where L_r is the rth cumulant, $\partial_s = \dfrac{\partial}{\partial\lambda_s}$, $(s = 1,2)$ and $L_1 = \lambda_1(1-\lambda_2)^{-1}$.

The first six central moments of the GPD are

$$\mu_1' = \lambda_1(1-\lambda_2)^{-1}$$

$$\mu_2 = \lambda_1(1-\lambda_2)^{-3}$$

$$\mu_3 = \lambda_1(1+2\lambda_2)(1-\lambda_2)^{-5}$$

(2.1.b) $$\mu_4 = 3\lambda_1^2(1-\lambda_2)^{-6} + \lambda_1(1+8\lambda_2+6\lambda_2^2)(1-\lambda_2)^{-7}$$

$$\mu_5 = 10\lambda_1^2(1+2\lambda_2)(1-\lambda_2)^{-8} + \lambda_1(1+22\lambda_2+58\lambda_2^2+24\lambda_2^3)(1-\lambda_2)^{-9}$$

$$\mu_6 = 15\lambda_1^3(1-\lambda_2)^{-9} + 5\lambda_1^2(5+32\lambda_2+26\lambda_2^2)(1-\lambda_2)^{-10}$$

$$+ \lambda_1(1+52\lambda_2+328\lambda_2^2+444\lambda_2^3+120\lambda_2^4)(1-\lambda_2)^{-11} \quad .$$

The problem of estimation of the GPD has been studied by many researchers, by considering the following two different forms:

(i) By writing $\lambda_1 = \alpha_1\theta$, $\lambda_2 = \alpha_2\theta$, the two parameters λ_1 and λ_2 become linear functions of a common parameter θ and the probability function (2.1) becomes

$$(2.2) \qquad P(X=x) = \begin{cases} \alpha_1(\alpha_1+\alpha_2 x)^{x-1} \dfrac{(\theta e^{-\alpha_2\theta})^x}{x!\ e^{\alpha_1\theta}}, & x = 0,1,2,\ldots \\ \\ 0 & \text{otherwise.} \end{cases}$$

Charalambides (8) considered θ as the only unknown parameter to be estimated, and he derived the distribution of the sufficient statistic for θ, and hence its minimum variance unbiased estimator. Gupta (22), derived the M.L. estimator for θ, and found the first order terms in the bias and the variance of that estimator.

(ii) Kumar and Consul (44), took $\lambda_1 = \theta$ and $\lambda_2 = \alpha\theta$, so that the number of parameters did not increase, but the parameter λ_2 became a linear function of the other parameter. They assumed α to be a known constant, and derived the exact amount of the bias and the variance of the M.L. estimator of θ, by using the negative moments of the GPD. Kumar (45), further derived the first order terms in the biases and covariances of the moment estimators of the parameters α and θ, and measured the asymptotic efficiency of the minimum chi-square method of estimation relative to the method of moments.

For the probability distribution given in (2.1) we discuss the simultaneous estimation of λ_1 and λ_2 by two methods.

We exclude the case $\lambda_2 = 0$ from our study in (2.1) because this boundary gives the Poisson distribution for which the results are very simple. Thus our subsequent study will be for $0 < \lambda_2 < 1$.

In section 2, we shall give the moment estimators of λ_1 and λ_2, their biases, variances and covariances, correct to terms of order n^{-1} and n^{-2}.

In section 3, we consider the M.L. estimators for the parameters λ_1 and λ_2, and shall give the expressions for the biases and covariances of these estimators up to terms of order n^{-1}. In section 4 we study the accuracy of the M.L. estimators from a decapitated and double truncated sample of size n, compared to a complete sample of the same size.

In section 5 we shall measure the asymptotic efficiency of the moment estimators relative to the M.L. estimators, and then give a numerical comparison between the performance of the two sets of estimators for some selected values of the parameters λ_1 and λ_2.

2.2 Moment Estimators for the Parameters of the GPD.

Based upon a random sample of size n taken from the probability distribution given by (2.1), the moment estimators of λ_1 and λ_2 are defined as the estimators which satisfy the equations obtained by equating the sample mean and sample variance with the population mean and variance respectively. The two equations are

$$\lambda_1(1-\lambda_2)^{-1} = m_1'$$

and

$$\lambda_1(1-\lambda_2)^{-3} = m_2 = n^{-1}\sum_{i=1}^{n}(x_i-m_1')^2 ,$$

where m_1' is the sample mean, and m_2 is the sample variance. On solving for λ_1 and λ_2, the moment estimators of λ_1 and λ_2 were given by Consul and Jain (10), as

$$\lambda_1^* = m_1'^{3/2} \cdot m_2^{-1/2}$$

(2.3)

$$\lambda_2^* = 1 - (m_1'/m_2)^{1/2}$$

2.2.1 <u>Lemma on Taylor's expansion of moment estimators</u>.

Since the moment estimators can be considered as functions $f(m_1',m_2)$ of the first two sample moments m_1' and m_2 and it can be easily shown that $f(m_1',m_2)$ belongs to the class C^5 as defined in 1.3, the bivariate Taylor's expansion of $f(m_1',m_2)$ becomes

(2.4) $$f(\mu_1'+h+\mu_2+k) = f(\mu_1',\mu_2) + \sum_{j=1}^{4}\frac{1}{j!}[h(\overline{\frac{\partial}{m_1'}} + k\,\overline{\frac{\partial}{m_2}})^j f(m_1',m_2)] + R_4$$

where $R_4 = \frac{1}{5!}\{h\frac{\partial}{\partial r} + k\frac{\partial}{\partial\beta}\}^5 f(r,s)$,

h and k represent the increments $(m_1'-\mu_1')$ and $(m_2-\mu_2)$ respectively, and the bar over the partial derivatives means that the values of these derivatives are to be evaluated at $m_1' = \mu_1'$ and $m_2 = \mu_2$ and r, s being replaced by $r = \mu_1'+\theta h$, $s = \mu_2+\theta k$, $0 < \theta < 1$ after differentiation. The first five terms of the expansion (2.4) are

$$
\begin{aligned}
f(m_1',m_2) &= f(\mu_1',\mu_2) + (m_1'-\mu_1')A + (m_2-\mu_2)B \\
&+ \frac{1}{2!}\left[(m_1'-\mu_1')^2C + 2(m_1'-\mu_1')(m_2-\mu_2)D + (m_2-\mu_2)^2E\right] \\
&+ \frac{1}{3!}\left[(m_1'-\mu_1')^3F + 3(m_1'-\mu_1')^2(m_2-\mu_2)G + 3(m_1'-\mu_1')(m_2-\mu_2)^2H\right. \\
&\left.+ (m_2-\mu_2)^3I\right] + \frac{1}{4!}\left[(m_1'-\mu_1')^4J + 4(m_1'-\mu_1')^3(m_2-\mu_2)K\right. \\
&+ 6(m_1'-\mu_1')^2(m_2-\mu_2)^2L + 4(m_1'-\mu_1')(m_2-\mu_2)^3M + \\
&\left.+ (m_2-\mu_2)^4Z\right] + \ldots
\end{aligned}
\tag{2.5}
$$

where

$$
\begin{aligned}
&\frac{\overline{\partial} f}{\partial m_1'} = A\ , &&\frac{\overline{\partial} f}{\partial m_2} = B\ , &&\frac{\overline{\partial}^2 f}{\partial m_1'\partial m_2} = D\ , \\
&\frac{\overline{\partial}^2 f}{\partial m_1'^2} = C\ , &&\frac{\overline{\partial}^2 f}{\partial m_2^2} = E\ , &&\frac{\overline{\partial}^3 f}{\partial m_1'^2\partial m_2} = G\ , \\
&\frac{\overline{\partial}^3 f}{\partial m_1'\partial m_2^2} = H\ , &&\frac{\overline{\partial}^3 f}{\partial m_1'^3} = F\ , &&\frac{\overline{\partial}^3 f}{\partial m_2^3} = I\ , \\
&\frac{\overline{\partial}^4 f}{\partial m_1'^4} = J\ , &&\frac{\overline{\partial}^4 f}{\partial m_1'^3\partial m_2} = K\ , &&\frac{\overline{\partial}^4 f}{\partial m_1'^2\partial m_2^2} = L\ , \\
&\frac{\overline{\partial}^4 f}{\partial m_1'\partial m_2^3} = M\ , &&\frac{\overline{\partial}^4 f}{\partial m_2^4} = Z\ ,
\end{aligned}
\tag{2.6.a}
$$

Thus, to obtain the particular terms for the expected value which give an accuracy of the order n^{-2}, one has to calculate expectations of the forms

(2.6.b) $\mu_{rs} = E[(m_1'-\mu_1')^r \ (m_2-\mu_2)^s] \qquad r,s = 0,1,2,3,4$ and $1 \le r+s \le 4$.

Their values are as follows:

$$\mu_{10} = E[(m_1'-\mu_1')] = 0$$

$$\mu_{01} = E[(m_2-\mu_2)] = \frac{-\mu_2}{n}$$

$$\mu_{20} = E[(m_1'-\mu_1')^2] = \frac{\mu_2}{n}$$

$$\mu_{02} = E[(m_2-\mu_2)^2] = \frac{1}{n}\ (\mu_4-\mu_2^2) + \frac{1}{n^2}\ (5\mu_2^2-2\mu_4) + o(\frac{1}{n^3})$$

$$\mu_{11} = E[(m_1'-\mu_1')(m_2-\mu_2)] = \frac{\mu_3}{n} - \frac{\mu_3}{n^2} +$$

(2.7.a) $$\mu_{30} = E[(m_1'-\mu_1')^3] = \frac{\mu_3}{n^2} + o(\frac{1}{n^3})$$

$$\mu_{21} = E[(m_1'-\mu_1')^2(m_2-\mu_2)] = \frac{1}{n^2}\ (\mu_4-4\mu_2^2) + o(\frac{1}{n^3})$$

$$\mu_{12} = E[(m_1'-\mu_1')(m_2-\mu_2)^2] = \frac{1}{n^2}\ (\mu_5-8\mu_2\mu_3) + o(\frac{1}{n^3})$$

$$\mu_{03} = E[(m_2-\mu_2)^3] = \frac{1}{n^2}\ (\mu_6-6\mu_2\mu_4-6\mu_3^2+5\mu_2^3) + o(\frac{1}{n^3})$$

$$\mu_{40} = E[(m_1'-\mu_1')^4] = \frac{3\mu_2^2}{n^2} + o(\frac{1}{n^3})$$

$$\mu_{31} = E[(m_1'-\mu_1')^3(m_2-\mu_2)] = \frac{3\mu_2\mu_3}{n^2} + o(\frac{1}{n^3})$$

$$\mu_{22} = E[(m_1'-\mu_1')^2(m_2-\mu_2)^2] = \frac{1}{n^2}\ (\mu_2\mu_4-\mu_2^3+2\mu_3^2) + o(\frac{1}{n^3})$$

$$\mu_{13} = E[(m_1'-\mu_1')(m_2-\mu_2)^3] = \frac{1}{n^2}\ (3\mu_3\mu_4-3\mu_2^2\mu_3) + o(\frac{1}{n^3})$$

$$\mu_{04} = E[(m_2-\mu_2)^4] = \frac{1}{n^2}\ (3\mu_4^2-6\mu_4\mu_2^2+3\mu_2^4) + o(\frac{1}{n^3}) \quad .$$

The accuracy of the above expressions for expectations was checked carefully either by the Shenton-Myer's (71) "orthogonal" statistics technique or by the method of symmetric functions given by Kendall and Stuart (39) as shown below.

It can be easily proved that μ_{rs} as defined in (2.6.b) is independent of the location parameter μ_1'. Thus, without any loss of generality, we can assume $\mu_1' = 0$. The expectations μ_{r0}', up to order n^{-2} are given by Cramer (16) as

$$E(m_1'^2) = \mu_2/n$$

$$E(m_1'^3) = \mu_3/n^2$$

$$E(m_1'^4) = 3\mu_2^2/n^2 + o(\frac{1}{n^3})$$

and

$$E(m_1'^r) = o(\frac{1}{n^3}) \qquad r \geq 5 \ .$$

To derive μ_{31}, μ_{13}, μ_{22}, and μ_{04}, we shall make use of the following Shenton-Myer's method (71) of orthogonal statistics. The authors have defined the rth orthogonal statistic Q_r in a general form and have used their moment generating function to provide tables of expectations of powers and products of these statistics correct up to order n^{-5}. Since we need the first two orthogonal statistics Q_1 and Q_2, we quote the corresponding results

$$Q_1 = \frac{1}{n} \sum_{\alpha=1}^{n} (x_\alpha - \mu_1') = m_1' - \mu_1'$$

$$Q_2 = \frac{1}{n} \sum_{\alpha=1}^{n} (x_\alpha - \mu_1')^2 - \frac{\mu_3}{\mu_2} \frac{1}{n} \sum_{\alpha=1}^{n} (x_\alpha - \mu_1') - \mu_2$$

$$= m_2 + Q_1^2 - Q_1 \frac{\mu_3}{\mu_2} - \mu_2$$

and

$$E(Q_1 Q_2) = 0 \ ,$$

$$E(Q_1^2 Q_2) = n^{-2}(\mu_4 - \frac{\mu_3^2}{\mu_2} - \mu_2^2) + o(\frac{1}{n^3}) \ ,$$

$$E(Q_1 Q_2^2) = n^{-2}(\mu_5 + \frac{\mu_3^3}{\mu_2^2} - 2\,\frac{\mu_3\mu_4}{\mu_2}) + o(\frac{1}{n^3}) \ ,$$

$$E(Q_2^2) = n^{-1}(\mu_4 - \frac{\mu_3^2}{\mu_2} - \mu_2^2) + o(\frac{1}{n^3}) \ ,$$

$$E(Q_1^2 Q_2^2) = \frac{\mu_2}{n^2}\,(\mu_4 - \frac{\mu_3^2}{\mu_2} - \mu_2^2) + o(\frac{1}{n^3}) \ ,$$

$$E(Q_2^3) = n^{-2}(\mu_6 - \frac{\mu_3^4}{\mu_2^3} + 2\mu_2^3 - \frac{3\mu_3\mu_5}{\mu_2} + \frac{3\mu_3^2\mu_4}{\mu_2^2} - 3\mu_2\mu_4 + 3\mu_3^2) + o(\frac{1}{n^3}) \ ,$$

$$E(Q_2^4) = \frac{3}{n^2}\,(\mu_4 - \frac{\mu_3^2}{\mu_2} - \mu_2^2)^2 + o(\frac{1}{n^3}) \ ,$$

$$E(Q_1^i Q_2^j) = o(\frac{1}{n^3}), \qquad (i,j > 2).$$

By using the relations (2.7.a) and the above expectations, we have

(2.7.b)
$$\begin{aligned}\mu_{31} &= E[m_1'^3(m_2-\mu_2)] \\ &= E[Q_1^3(Q_2+Q_1\frac{\mu_3}{\mu_2}-Q_1^2)] = 3\mu_2\mu_3/n^2 + o(\frac{1}{n^3})\end{aligned}$$

(2.7.c)
$$\begin{aligned}\mu_{22} &= E[m_1'^2(m_2-\mu_2)^2] \\ &= E[Q_1^2(Q_2+Q_1\frac{\mu_3}{\mu_2}-Q_1^2)^2] \\ &= \frac{1}{n^2}(\mu_2\mu_4-\mu_2^3+2\mu_3^2) + o(\frac{1}{n^3}) .\end{aligned}$$

Similarly,

(2.7.d)
$$\begin{aligned}\mu_{13} &= E[Q_1(Q_2+Q_1\frac{\mu_3}{\mu_2}-Q_1^2)^3] \\ &= \frac{\mu_3}{\mu_2}E(Q_1^2Q_2^2) + (\frac{\mu_3}{\mu_2})^3 E(Q_1^4) + \frac{2\mu_3}{\mu_2}E(Q_1^2Q_2^2) \\ &= \frac{1}{n^2}(3\mu_3\mu_4-3\mu_2^2\mu_3) + o(\frac{1}{n^3}) .\end{aligned}$$

Finally,

(2.7.e)
$$\begin{aligned}\mu_{04} &= E[(Q_2+Q_1\frac{\mu_3}{\mu_2}-Q_1^2)^4] \\ &= \frac{1}{n^2}[3(\mu_4-\frac{\mu_3^2}{\mu_2}-\mu_2^2)^2 - 3\frac{\mu_3^4}{\mu_2^2}+\frac{6\mu_3^2\mu_4}{\mu_2}-6\mu_2\mu_3^2] + o(\frac{1}{n^3}) \\ &= \frac{1}{n^2}(3\mu_4^2-6\mu_2^2\mu_4+3\mu_2^4) + o(\frac{1}{n^3}) .\end{aligned}$$

On taking the expectations on both sides of (2.5), using the notations in (2.6.a) and the expectations (2.7.a), we have for arbitrary function $f(m_1',m_2)$, of the sample moments m_1' and m_2,

$$E(f(m_1',m_2)) = f(\mu_1',\mu_2) - \frac{\mu_2 B}{n} + \frac{1}{2}\,[\frac{\mu_2}{n}\,C + 2(\frac{\mu_3}{n} - \frac{\mu_3}{n^2})D$$

$$+ (\frac{1}{n}\,(\mu_4-\mu_2^2) + \frac{1}{n^2}\,(5\mu_2^2-2\mu_4))E] + \frac{1}{6}\,[\frac{\mu_3}{n^2}\,F + \frac{3}{n^2}\,(\mu_4-4\mu_2^2)G$$

$$(2.8.a) \qquad + \frac{3}{n^2}\,(\mu_5-8\mu_2\mu_3)H + \frac{1}{n^2}\,(\mu_6-6\mu_2\mu_4-6\mu_3^2+5\mu_2^3)I]$$

$$+ \frac{1}{24}\,[\frac{3\mu_2^2}{n^2}\,J + \frac{4}{n^2}\,(3\mu_2\mu_3)K + \frac{6}{n^2}\,(\mu_2\mu_4-\mu_2^3+2\mu_3^2)L$$

$$+ \frac{4}{n^2}\,(3\mu_3\mu_4-3\mu_2^2\mu_3)M + \frac{1}{n^2}\,(3\mu_4^2-6\mu_4\mu_2^2+3\mu_2^4)Z] \qquad ,$$

correct to terms up to order n^{-2}.

Thus, in general one can write

$$(2.8.b) \qquad E(f(m_1',m_2)) - f(\mu_1',\mu_2) = \frac{k_1}{n} + \frac{1}{n^2}\sum_{i=1}^{11} \pi_i + o(\frac{1}{n^3})$$

where

$$(2.8.c) \qquad k_1 = -\mu_2 B + \frac{1}{2}\mu_2 C + \mu_3 D + \frac{1}{2}(\mu_4-\mu_2^2)E$$

and

$$(2.8.d) \qquad \pi_1 = -\mu_3 D, \qquad \pi_2 = \frac{1}{2}(5\mu_2^2-2\mu_4)E, \qquad \pi_3 = \frac{\mu_3 F}{6}\,,$$

$$(2.8.e) \qquad \pi_4 = \frac{1}{2}(\mu_4-4\mu_2^2)G, \qquad \pi_5 = \frac{1}{2}(\mu_5-8\mu_2\mu_3)H,$$

$$\pi_6 = \frac{1}{6}(\mu_6-6\mu_2\mu_4-6\mu_3^2+5\mu_2^3)I, \qquad \pi_7 = \frac{\mu_2^2}{8}\,J,$$

$$\pi_8 = \frac{\mu_2\mu_3}{2}\,K, \qquad \pi_9 = \frac{1}{4}(\mu_2\mu_4-\mu_2^3+2\mu_3^2)L,$$

$$\pi_{10} = \frac{1}{6}(3\mu_3\mu_4 - 3\mu_2^2\mu_3)M, \qquad \pi_{11} = \frac{1}{24}(3\mu_4^2 - 6\mu_4\mu_2^2 + 3\mu_2^4)Z \quad .$$

2.2.2 Biases of Moment Estimators.

Considering the expression for the moment estimator λ_1^*, of the parameter λ_1, which is given in (2.3), by

$$\lambda_1^* = (m_1')^{3/2} \cdot (m_2)^{-1/2} \quad ,$$

as the function $f(m_1',m_2)$, taking the derivatives partially w.r.t. m_1' and m_2 and putting $m_1' = \mu_1' = \lambda_1(1-\lambda_2)^{-1}$, and $m_2 = \mu_2 = \lambda_1(1-\lambda_2)^{-3}$, we get the values of the constants in (2.6.a) as

$$A = (3/2)(1-\lambda_2), \qquad B = -(1-\lambda_2)^3/2, \qquad D = -3(1-\lambda_2)^4/4\lambda_1,$$

$$C = 3(1-\lambda_2)^2/4\lambda_1, \qquad E = 3(1-\lambda_2)^6/4\lambda_1, \qquad G = -3(1-\lambda_2)^5/8\lambda_1^2,$$

$$(2.9) \quad H = 9(1-\lambda_2)^7/8\lambda_1^2, \qquad F = -3(1-\lambda_2)^3/8\lambda_1^2, \quad I = -15(1-\lambda_2)^9/8\lambda_1^2,$$

$$J = 9(1-\lambda_2)^4/16\lambda_1^3, \qquad K = 3(1-\lambda_2)^6/16\lambda_1^3, \quad L = 9(1-\lambda_2)^8/16\lambda_1^3,$$

$$M = -45(1-\lambda_2)^{10}/16\lambda_1^3, \quad \text{and} \quad Z = 105(1-\lambda_2)^{12}/16\lambda_1^3 \quad .$$

Now by using the expansion (2.5) for λ_1^*, taking the expectations on both sides and by using the values in (2.9) and the relations (2.7.a - e), and by substituting the values of μ_r, ($r = 2,3,4,5,6$) given in (2.1.b), we get the bias $b_1(\lambda_1^*) = E[\lambda_1^*-\lambda_1]$ as

$$(2.10)\qquad b_1(\lambda_1^*) = \frac{1}{n}\left[\frac{5}{4}\lambda_1 + \frac{3}{4}\lambda_2(2+3\lambda_2)(1-\lambda_2)^{-1}\right] + \frac{1}{n^2}\sum_{i=1}^{11}\pi_i + o\left(\frac{1}{n^3}\right)$$

where

$$\pi_1 = \frac{3(1+2\lambda_2)}{4(1-\lambda_2)}, \qquad \pi_2 = -\left(\frac{3\lambda_1}{8} + \frac{3(1+8\lambda_2+6\lambda_2^2)}{4(1-\lambda_2)}\right),$$

$$\pi_3 = \frac{-(1+2\lambda_2)}{16\lambda_1(1-\lambda_2)^2}, \qquad \pi_4 = \frac{3}{16(1-\lambda_2)} - \frac{3(1+8\lambda_2+6\lambda_2^2)}{16\lambda_1(1-\lambda_2)^2},$$

$$\pi_5 = \frac{9(1+2\lambda_2)}{8(1-\lambda_2)} + \frac{9(1+22\lambda_2+58\lambda_2^2+24\lambda_2^3)}{16\lambda_1(1-\lambda_2)^2},$$

$$\pi_6 = -\left(\frac{5\lambda_1}{8} + \frac{5(13+88\lambda_2+70\lambda_2^2)}{16(1-\lambda_2)} + \frac{5(1+52\lambda_2+328\lambda_2^2+444\lambda_2^3+120\lambda_2^4)}{16\lambda_1(1-\lambda_2)^2}\right),$$

$$\pi_7 = \frac{9}{128\lambda_1(1-\lambda_2)^2}, \qquad \pi_8 = \frac{3(1+2\lambda_2)}{32\lambda_1(1-\lambda_2)^2},$$

$$\pi_9 = \frac{9}{32(1-\lambda_2)} + \frac{9(3+16\lambda_2+14\lambda_2^2)}{64\lambda_1(1-\lambda_2)^2},$$

$$\pi_{10} = -\left(\frac{45(1+2\lambda_2)}{16(1-\lambda_2)} + \frac{45(1+2\lambda_2)(1+8\lambda_2+6\lambda_2^2)}{32\lambda_1(1-\lambda_2)^2}\right),$$

$$\pi_{11} = \frac{105\lambda_1}{32} + \frac{105(1+8\lambda_2+6\lambda_2^2)}{32(1-\lambda_2)} + \frac{105(1+8\lambda_2+6\lambda_2^2)^2}{128\lambda_1(1-\lambda_2)^2}\ .$$

To give an indication of the work involved in evaluating the above expressions, we give below the derivation of π_6.

$$\pi_6 = \frac{1}{6}[\mu_6-6\mu_2\mu_4-6\mu_3^2+5\mu_2^3]I$$

$$= \frac{1}{6}[15\lambda_1^3(1-\lambda_2)^{-9}+5\lambda_1^2(5+32\lambda_2+26\lambda_2^2)(1-\lambda_2)^{-10}+\lambda_1(1+52\lambda_2+328\lambda_2^2+444\lambda_2^3+$$

$$+120\lambda_2^4)(1-\lambda_2)^{-11}-18\lambda_1^3(1-\lambda_2)^{-9}-6\lambda_1^2(1+8\lambda_2+6\lambda_2^2)(1-\lambda_2)^{-10}-6\lambda_1^2(1+4\lambda_2$$

$$+4\lambda_2^2)(1-\lambda_2)^{-10}+5\lambda_1^3(1-\lambda_2)^{-9}]\left[\frac{-15(1-\lambda_2)^9}{8\lambda_1^2}\right]$$

$$= \frac{1}{6}[2\lambda_1^3(1-\lambda_2)^{-9}+\lambda_1^2(1-\lambda_2)^{-10}(25+160\lambda_2+130\lambda_2^2-6-24\lambda_2-24\lambda_2^2-6-48\lambda_2-36\lambda_2^2)$$

$$+ \lambda_1(1+52\lambda_2+328\lambda_2^2+444\lambda_2^3+120\lambda_2^4)(1-\lambda_2)^{-11}]\left[\frac{-15(1-\lambda_2)^9}{8\lambda_1^2}\right] \quad .$$

Thus

$$\pi_6 = -\left[\frac{5\lambda_1}{8} + \frac{5(13+88\lambda_2+70\lambda_2^2)}{16(1-\lambda_2)} + \frac{5(1+52\lambda_2+328\lambda_2^2+444\lambda_2^3+120\lambda_2^4)}{16\lambda_1(1-\lambda_2)^2}\right] \quad .$$

Denoting the moment estimator λ_2^* by $f^*(m_1',m_2)$ and its partial derivatives with respect to m_1', m_2 calculated at $m_1' = \mu_1'$ and $m_2 = \mu_2$, by $A^*,B^*,\ldots$ etc. as given in (2.6.a), the expectation

$$E(f^*(m_1',m_2)) - f^*(\mu_1',\mu_2) \quad ,$$

will be the same as in (2.8.b), but with the difference that the values of the constants k_1, $A,B,C,\ldots$ etc. will be different. Let

$$E(f^*(m_1',m_2)) - f^*(\mu_1',\mu_2) = k_1^*/n + n^{-2}\sum_{i=1}^{11}\pi_i^* + o(\frac{1}{n^3})$$

where

(2.11.a) $$k_1^* = -\mu_2 B^* + \tfrac{1}{2}\mu_2 C^* + \mu_3 D^* + \tfrac{1}{2}(\mu_4 - \mu_2^2)E^* \quad ,$$

π_i^* ($i = 1,\ldots,11$) are those expressions for π_i given in (2.8.e), when $A,B,C,\ldots$ are replaced by $A^*,B^*,C^*,\ldots$. Thus, for $f^*(m_1',m_2) = \lambda_2^* = 1 - \left(\frac{m_1'}{m_2}\right)^{\frac{1}{2}}$, the values of the required partial derivatives are given as

$$A^* = -(1-\lambda_2)^2/2\lambda_1, \quad B^* = (1-\lambda_2)^4/2\lambda_1, \quad D^* = (1-\lambda_2)^5/4\lambda_1^2,$$

$$C^* = (1-\lambda_2)^3/4\lambda_1^2, \quad E^* = -3(1-\lambda_2)^7/4\lambda_1^2, \; G^* = -(1-\lambda_2)^6/8\lambda_1^3,$$

(2.12) $$H^* = -3(1-\lambda_2)^8/8\lambda_1^3, \quad F^* = -3(1-\lambda_2)^4/8\lambda_1^3, \; I^* = 15(1-\lambda_2)^{10}/8\lambda_1^3,$$

$$J^* = 15(1-\lambda_2)^5/16\lambda_1^4, \; K^* = 3(1-\lambda_2)^7/16\lambda_1^4, \; L^* = 3(1-\lambda_2)^9/16\lambda_1^4,$$

$$M^* = 15(1-\lambda_2)^{11}/16\lambda_1^4, \; Z^* = -105(1-\lambda_2)^{13}/16\lambda_1^4,$$

Now by using the expansion (2.5) for λ_2^*, and taking the expectations on both sides and using the values in (2.12) with the relations (2.7.a) and by substituting the values of μ_r, ($r = 2,3,4,5,6$) given in (2.1.b) we get the bias $b_2(\lambda_2^*) = E[\lambda_2^* - \lambda_2]$ as

(2.13) $$b_2(\lambda_2^*) = \frac{1}{n}\left[\frac{-5}{4}(1-\lambda_2) - \frac{\lambda_2}{4\lambda_1}(10+9\lambda_2^2)\right] + n^{-2}\sum_{i=1}^{11}\pi_i^* + o\left(\frac{1}{n^3}\right)$$

and the corresponding values of π_i^* become

$$\pi_1^* = \frac{-(1+2\lambda_2)}{4\lambda_1}, \qquad \pi_2^* = \frac{3(1-\lambda_2)}{8} + \frac{3(1+8\lambda_2+6\lambda_2^2)}{4\lambda_1},$$

$$\pi_3^* = \frac{-(1+2\lambda_2)}{16\lambda_1^2(1-\lambda_2)}, \qquad \pi_4^* = \frac{1}{16\lambda_1} - \frac{(1+8\lambda_2+6\lambda_2^2)}{16\lambda_1^2(1-\lambda_2)},$$

$$\pi_5^* = -\left[\frac{3(1+2\lambda_2)}{8\lambda_1} + \frac{3(1+22\lambda_2+58\lambda_2^2+24\lambda_2^3)}{16\lambda_1^2(1-\lambda_2)}\right], \tag{2.14}$$

$$\pi_6^* = \frac{5(1-\lambda_2)}{8} + \frac{5(13+88\lambda_2+70\lambda_2^2)}{16\lambda_1} + \frac{5(1+52\lambda_2+328\lambda_2^2+444\lambda_2^3+120\lambda_2^4)}{16\lambda_1^2(1-\lambda_2)},$$

$$\pi_7^* = \frac{15}{128\lambda_1^2(1-\lambda_2)}, \qquad \pi_8^* = \frac{3(1+2\lambda_2)}{32\lambda_1^2(1-\lambda_2)},$$

$$\pi_9^* = \frac{3}{32\lambda_1} + \frac{3(3+16\lambda_2+14\lambda_2^2)}{64\lambda_1^2(1-\lambda_2)}, \qquad \pi_{10}^* = \frac{15(1+2\lambda_2)}{16\lambda_1} + \frac{15(1+2\lambda_2)(1+8\lambda_2+6\lambda_2^2)}{32\lambda_1^2(1-\lambda_2)},$$

and

$$\pi_{11}^* = -\left[\frac{105(1-\lambda_2)}{32} + \frac{105(1+8\lambda_2+6\lambda_2^2)}{32\lambda_1} + \frac{105(1+8\lambda_2+6\lambda_2^2)^2}{128\lambda_1^2(1-\lambda_2)}\right].$$

2.2.3 Asymptotic Variances and Covariance of Moment Estimators.

Now

$$\begin{aligned}\mathrm{Var}(f(m_1',m_2)) &= \mathrm{Var}(f(m_1',m_2) - f(\mu_1',\mu_2)) \\ &= E[(f(m_1',m_2)-f(\mu_1',\mu_2))^2] - [E(f(m_1',m_2)-f(\mu_1',\mu_2))]^2.\end{aligned}$$

Subtracting $f(\mu_1',\mu_2)$ from both sides of (2.5), squaring and taking expectations, one can show that

$E[(f(m_1',m_2)-f(\mu_1',\mu_2))^2]$

$$= \frac{\mu_2}{n} A^2 + 2\left(\frac{\mu_3}{n} - \frac{\mu_3}{n^2}\right)AB + \left(\frac{1}{n}(\mu_4-\mu_2^2) + \frac{1}{n^2}(5\mu_2^2-2\mu_4)\right)B^2$$

$$+ \frac{\mu_3}{n^2} AC + \frac{1}{n^2}(\mu_4-4\mu_2^2)(2AD+BC) + \frac{1}{n^2}(\mu_5-8\mu_2\mu_3)(2BD+AE)$$

$$\text{(2.15)} \quad + \frac{1}{n^2}(\mu_6-6\mu_2\mu_4-6\mu_3^2+5\mu_2^3)BE + \frac{3\mu_2^2}{n^2}\left(\frac{C^2}{4} + \frac{AF}{3}\right)$$

$$+ \frac{1}{n^2}(3\mu_2\mu_3)(AG + \frac{BF}{3} + DC) + \frac{1}{n^2}(\mu_2\mu_4-\mu_2^3+2\mu_3^2)(D^2+BG+AH+ \frac{CE}{2})$$

$$+ \frac{1}{n^2}(3\mu_3\mu_4-3\mu_2^2\mu_3)(DE + \frac{AI}{3} + BH) + \frac{1}{n^2}(3\mu_4^2-6\mu_4\mu_2^2+3\mu_2^4)\left(\frac{E^2}{4} + \frac{BI}{3}\right) + o(\frac{1}{n^3}).$$

Collecting terms in (2.15) for the coefficients of n^{-1} and n^{-2}, one can show that, the asymptotic expansion of the variance of $f(m_1',m_2)$ can be written in the form

$$\text{(2.16.a)} \qquad \operatorname{Var}(f(m_1',m_2)) = v_1/n + n^{-2}\sum_{i=1}^{12} \alpha_i + o(\frac{1}{n^3}) ,$$

where, in general

$$\text{(2.16.b)} \qquad v_1 = \mu_2 A^2 + 2\mu_3 AB + (\mu_4-\mu_2^2)B^2$$

and

$$\alpha_1 = -2\mu_3 AB, \qquad \alpha_2 = (5\mu_2^2-2\mu_4)B^2, \qquad \alpha_3 = \mu_3 AC,$$

$$\alpha_4 = (\mu_4-4\mu_2^2)(2AD+BC), \qquad \alpha_5 = (\mu_5-8\mu_2\mu_3)(2BD+AE),$$

$$\text{(2.16.c)} \quad \alpha_6 = (\mu_6-6\mu_2\mu_4-6\mu_3^2+5\mu_2^3)BE, \qquad \alpha_7 = 3\mu_2^2\left(\frac{C^2}{4}+\frac{AF}{3}\right),$$

$$\alpha_8 = 3\mu_2\mu_3(AG+\frac{BF}{3}+DC), \qquad \alpha_9 = (\mu_2\mu_4-\mu_2^3+2\mu_3^2)(D^2+BG+AH+\frac{CE}{2}),$$

$$\alpha_{10} = (3\mu_3\mu_4-3\mu_2^2\mu_3)(DE+\frac{AI}{3}+BH), \qquad \alpha_{11} = (3\mu_4^2-6\mu_4\mu_2^2-3\mu_2^4)\left(\frac{E^2}{4}+\frac{BI}{3}\right),$$

$$\alpha_{12} = -[-\mu_2 B+\frac{\mu_2 C}{2}+\frac{1}{2}(\mu_4-\mu_2^2)E+\mu_3 D]^2 \quad .$$

Thus the asymptotic expansion of the variance of $\lambda_1^* = f(m_1', m_2)$ given in (2.16.a) will be

$$\text{(2.17)} \qquad \mathrm{Var}(\lambda_1^*) \doteq \frac{1}{n}\left[\frac{\lambda_1^2}{2}+\frac{\lambda_1}{2(1-\lambda_2)}(2-2\lambda_2+3\lambda_2^2)\right]+\frac{1}{n^2}\sum_{i=1}^{12}\alpha_i + o(\frac{1}{n^3})$$

where α_i $(i = 1,2,\ldots,12)$ can be evaluated by substituting the values of μ_r $(r = 2,\ldots,6)$ and $A,B,\ldots$ etc., in (2.16.c).

Denoting the moment estimator λ_2^* of the parameter λ_2 by $f^*(m_1', m_2)$, the asymptotic variance of λ_2^* is similarly given by

$$\mathrm{Var}(\lambda_2^*) = \frac{1}{n}[\mu_2 A^{*2} + 2\mu_3 A^* B^* + (\mu_4-\mu_2^2)B^{*2}] + o(\frac{1}{n^3})$$

$$\text{(2.18)} \qquad = \frac{1}{n}\left[\frac{(1-\lambda_2)}{2\lambda_1}(\lambda_1-\lambda_1\lambda_2+2\lambda_2+3\lambda_1^2)\right]+\frac{1}{n^2}\sum_{i=1}^{12}\alpha_i^* + o(\frac{1}{n^3})$$

where α_i^* have the same expression (2.16.c) when $A,B,C,\ldots$ are replaced by $A^*, B^*, C^*, \ldots$ given in (2.12).

To find the covariance of $f^*(m_1', m_2)$ and $f(m_1', m_2)$ we shall use the bivariate-Taylor expansion of both functions.

Now

$$\begin{aligned} \mathrm{Cov}(f,f^*) &= E[(f(m_1',m_2)-f(\mu_1',\mu_2))(f^*(m_1',m_2)-f^*(\mu_1',\mu_2))] \\ &\quad -E[f(m_1',m_2)-f(\mu_1',\mu_2)]\cdot E[f^*(m_1',m_2)-f^*(\mu_1',\mu_2)]. \end{aligned} \tag{2.19.a}$$

But

$$[f(m_1',m_2)-f(\mu_1',\mu_2)][f^*(m_1',m_2)-f^*(\mu_1',\mu_2)]$$

$$\begin{aligned} &= (m_1'-\mu_1')^2 AA^* + (m_1'-\mu_1')(m_2-\mu_2)(A^*B+B^*A) + (m_2-\mu_2)^2 BB^* \\ &+ \tfrac{1}{2}[(m_1'-\mu_1')^3(AC^*+CA^*) + (m_1'-\mu_1')^2(m_2-\mu_2)(2AD^*+BC^*+B^*C+2A^*D) \\ &+ (m_1'-\mu_1')(m_2-\mu_2)^2(AE^*+2BD^*+2B^*D+A^*E) + (m_2-\mu_2)^3(BE^*+B^*E)] \\ &+ \tfrac{1}{12}[(m_1'-\mu_1')^4(2A^*F+2AF^*+3CC^*) + (m_1'-\mu_1')^3(m_2-\mu_2)(6AG^*+2BF^* \\ &+ 6CD^*+2B^*F+6C^*D+6A^*G) + (m_1'-\mu_1')^2(m_2-\mu_2)^2(6AH^*+6BG^*+6A^*H \\ &+ 6B^*G+3CE^*+12DD^*+3C^*E) + (m_1'-\mu_1')(m_2-\mu_2)^3(6BH^*+2AI^*+6B^*H \\ &+ 2A^*I+6D^*E+6DE^*) + (m_2-\mu_2)^4(2BI^*+2B^*I+3EE^*)] \quad . \end{aligned} \tag{2.19.b}$$

Taking the expectation on both sides of (2.19.b), and substituting together with (2.7.a) in (2.19.a) we have

$$(2.19.c)\quad \mathrm{Cov}(f,f^*) \doteq \frac{1}{n}[A\mu_2A^*+\mu_3(A^*B+B^*A)+B(\mu_4-\mu_2^2)B^*] + \frac{1}{n^2}\sum_{j=1}^{11}\omega_j$$

where

$$\omega_1 = (5\mu_2^2-2\mu_4)BB^* - \mu_3(A^*B+B^*A), \qquad \omega_2 = \frac{\mu_3}{2}(AC^*+CA^*),$$

$$\omega_3 = \frac{1}{2}(\mu_4-4\mu_2^2)(2AD^*+BC^*+B^*C+2A^*D), \qquad \omega_4 = \frac{1}{2}(\mu_5-8\mu_2\mu_3)(AE^*+2BD^*+A^*E+2B^*D),$$

$$\omega_5 = \frac{1}{2}(\mu_6-6\mu_2\mu_4-6\mu_3^2+5\mu_2^3)(BE^*+B^*E), \qquad \omega_6 = \frac{\mu_2^2}{4}(2A^*F+3CC^*+2AF^*),$$

$$(2.19.d)\quad \omega_7 = \frac{1}{12}(3\mu_2\mu_3)(6AG^*+2BF^*+6CD^*+6C^*D+2B^*F+6A^*G),$$

$$\omega_8 = \frac{1}{12}(\mu_2\mu_4-\mu_2^3+2\mu_3^2)(6AH^*+6BG^*+3CE^*+12DD^*+6A^*H+6B^*G+3C^*E),$$

$$\omega_9 = \frac{1}{12}(3\mu_3\mu_4-3\mu_2^2\mu_3)(6BH^*+2AI^*+6B^*H+2A^*I+6D^*E+6DE^*),$$

$$\omega_{10} = \frac{1}{12}(3\mu_4^2-6\mu_4\mu_2^2+3\mu_2^4)(2BI^*+2B^*I+3EE^*),$$

$$\omega_{11} = -[-\mu_2B+\frac{\mu_2C}{2}+\frac{(\mu_4-\mu_2^2)E}{2}+\mu_3D][-\mu_2B^*+\frac{\mu_2C^*}{2}+\frac{(\mu_4-\mu_2^2)E^*}{2}+\mu_3D^*] .$$

Thus, the covariance between the moment estimators of $f = f(m_1',m_2) = \lambda_1^*$ and $f^* = f^*(m_1',m_2) = \lambda_2^*$ of the parameters λ_1 and λ_2, is obtained by using (2.2.b), (2.12) and (2.9) in (2.19.c) with the result

$$(2.20)\quad \mathrm{Cov}(\lambda_1^*,\lambda_2^*) = \frac{1}{2n}[-\lambda_1+\lambda_1\lambda_2-3\lambda_2^2] + \frac{1}{n^2}\sum_{j=1}^{11}\omega_j + o(\frac{1}{n^3})$$

and ω_j $(j = 1,2,\ldots,11)$ are obtained from (2.19.d).

2.3 Maximum Likelihood Estimation for the GPD from a Complete Sample.

The following recurrence relations for the probabilities of the GPD given in (2.1) are of chief importance in deriving the M.L. estimators and their asymptotic biases and covariances.

Consul and Jain (11), have shown that the probabilities (2.1) satisfy the recurrence relations

$$\text{(2.21.a)} \quad \partial_1 P_x(\lambda_1,\lambda_2) = P_{x-1}(\lambda_1+\lambda_2,\lambda_2) - P_x(\lambda_1,\lambda_2)$$

$$\text{(2.21.b)} \quad \partial_2 P_x(\lambda_1,\lambda_2) = \frac{1-\lambda_2}{\lambda_2} \cdot x \cdot P_x(\lambda_1,\lambda_2) - \frac{\lambda_1}{\lambda_2} P_{x-1}(\lambda_1+\lambda_2,\lambda_2)$$

$$\text{(2.21.c)} \quad xP_x(\lambda_1,\lambda_2) = \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}(x-1)P_{x-1}(\lambda_1+\lambda_2,\lambda_2) - \lambda_1 P_{x-1}(\lambda_1+\lambda_2,\lambda_2)$$

where $P_x(\lambda_1,\lambda_2)$ is defined in (2.1).

For the sake of convenience and brevity we shall denote the probabilities $P_x(\lambda_1,\lambda_2)$, $P_{x-1}(\lambda_1+\lambda_2,\lambda_2)$, and $P_{x-2}(\lambda_1+2\lambda_2,\lambda_2)$ by P_x, $\overset{-}{P}_{x-1}$, $\overset{+}{P}_{x-2}$ respectively.

Let a sample of size n be taken from the GPD (2.1) which gives the frequencies n_x, $x = 0,1,2,\ldots,k$, for the different classes, and let $\sum_{x=0}^{k} n_x = n$. The log-likelihood function

$$\begin{aligned}\log L &= \sum_{x=0}^{k} n_x \log P_x \\ &= \sum_{x=0}^{k} n_x[-\lambda_1 - x\lambda_2 + (x-1)\log(\lambda_1+x\lambda_2) + \log\lambda_1 - \log x!]\end{aligned}$$

where P_x is the probability of observation x, n_x is the frequency of the observation x and k is the largest of the observations.

On differentiating log L partially with respect to λ_1 we have

$$\frac{\partial \log L}{\partial \lambda_1} = -n + \frac{n}{\lambda_1} + \sum_{x=0}^{k} n_x \frac{(x-1)}{\lambda_1+\lambda_2 x}$$

$$\frac{\partial^2 \log L}{\partial \lambda_1^2} = -\left[\frac{n}{\lambda_1^2} + \sum_{x=0}^{k} n_x \frac{(x-1)}{(\lambda_1+\lambda_2 x)^2}\right] .$$

Now, the ML equation $\frac{\partial \log L}{\partial \lambda_1} = 0$ for λ_1 can be written in the form

$$\hat{\lambda}_1 = \sum_{x=0}^{k} \frac{n_x}{n} \frac{x(\hat{\lambda}_1+\hat{\lambda}_2)}{\hat{\lambda}_1+x\hat{\lambda}_2} \tag{2.22}$$

where the function of λ_1 on the right, is continuous and is more than 1 and less than m_1' if $\lambda_2 > 0$ and satisfies the Lipschitz condition given in (1.5). Thus the ML equation in λ_1 must have a unique solution, corresponding to each value of λ_2, which must maximize the likelihood function because the second partial derivative is always negative.

It should be noted that in the case the probability mass is concentrated at Zero, where the probability of the sample becomes

$$P\left(\bigcap_{i=1}^{n} X_i = 0\right) = e^{-n\lambda_1} ,$$

the M.L. estimator is $\hat{\lambda}_1 = 0$. For that particular case equation (2.22) will not hold.

Similarly, on differentiating log L partially with respect to λ_2, we get

$$\frac{\partial \log L}{\partial \lambda_2} = \sum_{x=0}^{k} n_x \left[-x + \frac{x(x-1)}{\lambda_1 + x\lambda_2} \right]$$

$$\frac{\partial^2 \log L}{\partial \lambda_2^2} = - \sum_{x=0}^{k} \left[n_x \frac{x^2(x-1)}{(\lambda_1 + x\lambda_2)^2} \right] .$$

Again, the ML equation $\frac{\partial \log L}{\partial \lambda_2} = 0$ for λ_2 can be written in the form

(2.23) $$\lambda_2 = 1 - \frac{1}{m_1'} \sum_{x=0}^{k} \frac{n_x}{n} \frac{x(\lambda_1 + \lambda_2)}{\lambda_1 + x\lambda_2} .$$

For simultaneous ML estimation of the two parameters we eliminate λ_1 between (2.22) and (2.23) and get

(2.24) $$\lambda_2 = 1 - \frac{1}{m_1'} \sum_{x=0}^{k} \frac{n_x}{n} \frac{\{m_1'(1-\lambda_2) + \lambda_2\}x}{m_1'(1-\lambda_2) + x\lambda_2} ,$$

as the ML estimating equation in one variable λ_2. The equation (2.24) is of the same form as (2.23), with the difference that λ_1 is replaced by $m_1'(1-\lambda_2)$.

It has been shown in Appendix I that the matrix of the second partial derivatives of the log likelihood function is negative definite. Accordingly, if the ML equation (2.24) has a solution it must be the unique M.L. estimator.

We now define

$$(2.25) \qquad H(\lambda_2) = 1 - \sum_{x=0}^{k} \frac{n_x}{n} \frac{\{m_1'(1-\lambda_2)+\lambda_2\}x}{\{m'(1-\lambda_2)+x\lambda_2\}m_1'} - \lambda_2 .$$

The function $H(\lambda_2)$ is the sum of a finite number of expressions in λ_2. Obviously, $H(0) = 0$, and $H(1) = -\frac{1}{m_1'}(1 - \frac{n_0}{n})$. Thus to apply Newton-Raphson method of iteration one should find an interval $[a,b] \subset (0,1)$, where neither a, nor b depend upon the sample values, and for some value of $\lambda_2 \in [a,b]$ the function $H(\lambda_2)$ is positive and its first order partial derivative w.r.t. λ_2 does not vanish. We could not find such an interval, and the graph of the function $H(\lambda_2)$ does not seem to satisfy the condition $H(a) \cdot H(b) < 0$. This means the Newton-Raphson method of iteration fails as a tool to detect the unique root of $H(\lambda_2) = 0$. Hence other methods of iteration should be tried to find the unique root of the equation $H(\lambda_2) = 0$.

In the following section we derive the asymptotic biases, variances and covariance of the M.L. estimators, based upon the work by Shenton and Bowman (1975).

2.3.1 Asymptotic Variances and Covariances of the ML Estimators.

Assuming the four regularity conditions given in (1.4), Shenton and Bowman (72), (75) have proved that the first order terms of the variances and covariance of the ML estimators for two parameters, are given by

$$\begin{aligned} \operatorname{Var}(\hat{\lambda}_1) &= P_{22}/n|I| + o(1/n) \\ \operatorname{Cov}(\hat{\lambda}_1, \hat{\lambda}_2) &= -P_{12}/n|I| + o(1/n) \\ \operatorname{Var}(\hat{\lambda}_2) &= P_{11}/n|I| + o(1/n) \end{aligned} \tag{2.26.a}$$

where

$$\begin{aligned} P_{11} &= E(P^{-2}(\partial_1 P)^2) \\ P_{12} &= E(P^{-2}(\partial_1 P)(\partial_2 P)) \\ P_{22} &= E(P^{-2}(\partial_2 P)^2) \end{aligned} \tag{2.26.b}$$

where p is the probability defined in (2.1), and

$$|I| = P_{11}P_{22} - P_{12}^2 \quad . \tag{2.26.c}$$

Since the four regularity conditions given by (1.4) are satisfied by the GPD defined in (2.1), the above results will hold for the ML estimators $\hat{\lambda}_1$ and $\hat{\lambda}_2$.

By using (2.21.a), (2.21.b), and (2.21.c), we shall now evaluate the expressions for P_{ij} $(i, j = 1,2)$.

$$P_{11} = \sum_{x=1}^{\infty} \frac{\bar{P}_{x-1}^2}{P_x} - 1$$

$$\begin{aligned} &= \frac{1}{\lambda_1} \sum_{x=1}^{\infty} \left\{1 + \frac{\lambda_1(x-1)}{\lambda_1+\lambda_2 x}\right\} \bar{P}_{x-1} - 1 = \frac{1}{\lambda_1} + \sum_{x=2}^{\infty} \frac{\lambda_1+\lambda_2}{\lambda_1+2\lambda_2} \overset{+}{P}_{x-2} - 1 \\ &= \frac{1}{\lambda_1} - \frac{\lambda_2}{\lambda_1+2\lambda_2} = \frac{\lambda_1+2\lambda_2-\lambda_1\lambda_2}{\lambda_1(\lambda_1+2\lambda_2)} \quad . \end{aligned} \tag{2.27}$$

Similarly,

$$P_{22} = E\left(\frac{1}{P^2}\left(\frac{\partial P}{\partial \lambda_2}\right)^2\right) = E\left[\left(\frac{1-\lambda_2}{\lambda_2}\,x\right)^2 - \frac{2\lambda_1(1-\lambda_2)}{\lambda_2^2}\,x\,\frac{\bar{P}_{x-1}}{P_x} + \frac{\lambda_1^2}{\lambda_2^2}\,\frac{\bar{P}^2_{x-1}}{P_x^2}\right]$$

(2.28)

$$= \frac{\lambda_1(2+\lambda_1)}{(1-\lambda_2)(\lambda_1+2\lambda_2)}$$

and

$$P_{12} = E\left(\frac{1}{P^2}\,\frac{\partial P}{\partial \lambda_1}\,\frac{\partial P}{\partial \lambda_2}\right) = E\left[\frac{1-\lambda_2}{\lambda_2}\,\frac{\bar{P}_{x-1}}{P_x}\,x - \frac{(1-\lambda_2)}{\lambda_2}\,x + \frac{\lambda_1}{\lambda_2}\,\frac{\bar{P}_{x-1}}{P_x} - \frac{\lambda_1}{\lambda_2}\,\frac{\bar{P}^2_{x-1}}{P_x^2}\right]$$

(2.29)

$$= \frac{1}{\lambda_2} - \frac{\lambda_1}{\lambda_2}\,P_{11} = \frac{\lambda_1}{\lambda_1+2\lambda_2} \quad .$$

Fisher's information determinant $|I|$, given by (2.26.c) becomes

(2.30) $$|I| = 2[(1-\lambda_2)(\lambda_1+2\lambda_2)]^{-1} \ .$$

Thus

(2.31) $$\mathrm{Var}(\hat{\lambda}_1) = \frac{\lambda_1(2+\lambda_1)}{2n} + o\left(\frac{1}{n}\right)$$

(2.32) $$\mathrm{Var}(\hat{\lambda}_2) = \frac{(\lambda_1+2\lambda_2-\lambda_1\lambda_2)(1-\lambda_2)}{2\lambda_1 n} + o\left(\frac{1}{n}\right)$$

(2.33) $$\mathrm{Cov}(\hat{\lambda}_1,\hat{\lambda}_2) = -\frac{\lambda_1(1-\lambda_2)}{2n} + o\left(\frac{1}{n}\right) \ .$$

2.3.2 Asymptotic Biases of the ML Estimators.

As we have already shown, explicit forms for the ML estimators cannot be obtained, and consequently, terms in the biases cannot be easily expressed as functions of the parameters λ_1 and λ_2. Under the same regularity conditions given in section (1.4), Shenton and Bowman (75), Shenton and Wallington (70) have proved that the first order approximations to the biases of the ML estimators of two parameters satisfy the following two equations

$$\text{(2.34.a)}\qquad \begin{aligned} b_1(\hat{\lambda}_1)P_{11} + b_2(\hat{\lambda}_2)P_{12} &= -L_1/2|I| \\ b_1(\hat{\lambda}_1)P_{21} + b_2(\hat{\lambda}_2)P_{22} &= -L_2/2|I| \end{aligned}$$

where

$$\text{(2.34.b)}\qquad \begin{aligned} L_1 &= P_{22}P_{1,11} - 2P_{12}P_{1,12} + P_{11}P_{1,22} \\ L_2 &= P_{22}P_{2,11} - 2P_{12}P_{2,12} + P_{11}P_{2,22} \end{aligned}$$

and

$$P_{1,11} = E\left(P^{-2}\,\frac{\partial P}{\partial\lambda_1}\,\frac{\partial^2 P}{\partial\lambda_1^2}\right), \qquad P_{1,12} = E\left(P^{-2}\,\frac{\partial P}{\partial\lambda_1}\,\frac{\partial^2 P}{\partial\lambda_1\partial\lambda_2}\right),$$

$$\text{(2.35)}\quad P_{1,22} = E\left(P^{-2}\,\frac{\partial P}{\partial\lambda_1}\,\frac{\partial^2 P}{\partial\lambda_2^2}\right), \qquad P_{2,11} = E\left(P^{-2}\,\frac{\partial P}{\partial\lambda_2}\,\frac{\partial^2 P}{\partial\lambda_1^2}\right),$$

$$P_{2,12} = E\left(P^{-2}\,\frac{\partial P}{\partial\lambda_2}\,\frac{\partial^2 P}{\partial\lambda_1\partial\lambda_2}\right), \qquad P_{2,22} = E\left(P^{-2}\,\frac{\partial P}{\partial\lambda_2}\,\frac{\partial^2 P}{\partial\lambda_2^2}\right).$$

The formal derivation of $b_1(\hat{\lambda}_1)$ and $b_2(\hat{\lambda}_2)$ will be given in Appendix (II). We shall now calculate the values of $P_{\alpha,ij}$ $(\alpha,i,j = 1,2)$ given in (2.35).

Using the recurrence relations given in (2.21.a), (2.21.b) and (2.21.c), we have

$$P_{1,11} = \sum_{x=0}^{\infty} \frac{\bar{P}_{x-1}\overset{+}{P}_{x-2}}{P_x} - 2\sum_{x=0}^{\infty} \frac{\bar{P}_{x-1}}{P_x} + 1 \; . \tag{2.36.a}$$

Let

$$C_{12} = \sum_{x=0}^{\infty} \frac{\bar{P}_{x-1}\overset{+}{P}_{x-2}}{P_x} \; .$$

Since

$$\bar{P}_{x-1} = \frac{(\lambda_1+\lambda_2)x}{\lambda_1(\lambda_1+\lambda_2 x)} P_x \; , \tag{2.36.b}$$

and

$$\frac{x}{\lambda_1+\lambda_2 x} = \frac{2}{\lambda_1+2\lambda_2} + \frac{\lambda_1(x-2)}{(\lambda_1+2\lambda_2)(\lambda_1+\lambda_2 x)} \; , \tag{2.36.c}$$

then

$$\begin{aligned}\sum_{x=2}^{\infty} \frac{x}{\lambda_1+\lambda_2 x}\overset{+}{P}_{x-2} &= \frac{2}{\lambda_1+2\lambda_2} + \frac{\lambda_1}{\lambda_1+2\lambda_2}\sum_{x=2}^{\infty} \frac{(x-2)}{\lambda_1+\lambda_2 x} \\ &\quad \times (\lambda_1+2\lambda_2)(\lambda_1+\lambda_2 x)^{x-3}((x-2)!)^{-1}\exp[-(\lambda_1+\lambda_2 x)] \; .\end{aligned} \tag{2.36.d}$$

Thus, from (2.36.b), (2.36.c) and (2.36.d),

$$C_{12} = \frac{2(\lambda_1+\lambda_2)}{\lambda_1(\lambda_1+2\lambda_2)} + \frac{(\lambda_1+\lambda_2)}{\lambda_1+3\lambda_2} \; . \tag{2.36.e}$$

Hence

$$P_{1,11} = \frac{2\lambda_1\lambda_2^2-2\lambda_1\lambda_2-6\lambda_2^2}{\lambda_1(\lambda_1+2\lambda_2)(\lambda_1+3\lambda_2)} \; . \tag{2.36.f}$$

Since

$$\frac{\partial P_x}{\partial \lambda_2}\,\frac{P_x}{\partial \lambda_1 \partial \lambda_2} = \left[\frac{1-\lambda_2}{\lambda_2}\,xP_x - \frac{\lambda_1}{\lambda_2}\,\bar{P}_{x-1}\right]\left[\frac{1-\lambda_2}{\lambda_2}\,x\bar{P}_{x-1} - \frac{1-\lambda_2}{\lambda_2}\,xP_x - \frac{\lambda_1}{\lambda_2}\,P_{x-2}\right.$$

(2.37.a)

$$\left. - \frac{(1-\lambda_1)}{\lambda_2}\,\bar{P}_{x-1}\right] \quad ,$$

then

$$P_{2,12} = \left(\frac{1-\lambda_2}{\lambda_2}\right)^2 \sum x^2\bar{P}_{x-1} - \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\sum x\,\frac{\bar{P}^2_{x-1}}{P_x} - \left(\frac{1-\lambda_2}{\lambda_2}\right)^2 \sum x^2 P_x$$

(2.37.b)
$$+ \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\sum x\bar{P}_{x-1} - \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\sum x\overset{+}{P}_{x-2} + \frac{\lambda_1^2}{\lambda_2^2}\,C_{12}$$

$$- \frac{(1-\lambda_1)(1-\lambda_2)}{\lambda_2^2}\sum x\bar{P}_{x-1} + \frac{\lambda_1(1-\lambda_1)}{\lambda_2^2}\sum \frac{\bar{P}^2_{x-1}}{P_x} \quad .$$

Moreover

(2.37.c)
$$\sum x^2\bar{P}_{x-1} = \frac{\lambda_1+\lambda_2}{(1-\lambda_2)^3} + \frac{(\lambda_1+\lambda_2)^2}{(1-\lambda_2)^2} + \frac{2(\lambda_1+\lambda_2)}{1-\lambda_2} + 1$$

and

(2.37.d)
$$\sum x\,\frac{\bar{P}^2_{x-1}}{P_x} = \sum (x-1)\,\frac{P^2_{x-1}}{P_x} + (P_{11}+1).$$

On substituting in (2.37.b), and after simplification, one can show that

(2.37.e)
$$P_{2,12} = \frac{5\lambda_1^2\lambda_2+13\lambda_1\lambda_2^2-2\lambda_1^2\lambda_2^2+12\lambda_2^3-6\lambda_2^2-8\lambda_1\lambda_2-2\lambda_1^2}{\lambda_2(1-\lambda_2)(\lambda_1+2\lambda_2)(\lambda_1+3\lambda_2)} \quad .$$

Since

$$(2.38.a) \qquad \frac{\partial P_x}{\partial \lambda_2}\frac{\partial^2 P_x}{\partial \lambda_1^2} = \left[\frac{1-\lambda_2}{\lambda_2} xP_x - \frac{\lambda_1}{\lambda_2}\bar{P}_{x-1}\right]\left[\frac{\partial \bar{P}_{x-1}}{\partial \lambda_1} - \frac{\partial P_x}{\partial \lambda_1}\right] .$$

then

$$(2.38.b) \qquad \begin{aligned} P_{2,11} &= \frac{1-\lambda_2}{\lambda_2}\sum x \frac{\partial \bar{P}_{x-1}}{\partial \lambda_1} - \frac{\lambda_1}{\lambda_2}\sum \frac{\bar{P}_{x-1}}{P_x}\frac{\partial \bar{P}_{x-1}}{\partial \lambda_1} \\ &- \frac{1-\lambda_2}{\lambda_2}\sum x \frac{\partial P_x}{\partial \lambda_1} + \frac{\lambda_1}{\lambda_2}\sum \frac{\bar{P}_{x-1}}{P_x}(\bar{P}_{x-1} - P_x) . \end{aligned}$$

Since the P function, given by (2.1), satisfies the regularity conditions with respect to λ_1 and λ_2 and the operations of summation and differentiation can be interchanged, accordingly,

$$(2.38.c) \qquad \begin{aligned} P_{2,11} &= \frac{(1-\lambda_2)}{\lambda_2}\left(\frac{1}{1-\lambda_2}\right) - \frac{\lambda_1}{\lambda_2}C_{12} + 2\frac{\lambda_1}{\lambda_2}(P_{11} + 1) \\ &- \frac{1-\lambda_2}{\lambda_2}\left(\frac{1}{1-\lambda_2}\right) - \frac{\lambda_1}{\lambda_2} . \end{aligned}$$

Substituting the values of P_{11} and C_{12} given respectively in (2.27) and (2.36.e), we get

$$(2.38.d) \qquad P_{2,11} = \frac{-2(\lambda_1+\lambda_2)}{\lambda_2(\lambda_1+2\lambda_2)} - \frac{\lambda_1(\lambda_1+\lambda_2)}{\lambda_2(\lambda_1+3\lambda_2)} + \frac{\lambda_1}{\lambda_2} + \frac{2}{\lambda_2} - \frac{2\lambda_1}{\lambda_2(\lambda_1+2\lambda_2)}$$

which, upon simplification, can be written as

$$(2.38.e) \qquad P_{2,11} = \frac{(2\lambda_1^2\lambda_2 - 2\lambda_1^2 - 4\lambda_1\lambda_2 + 4\lambda_1\lambda_2^2 + 6\lambda_2^2)}{\lambda_2(\lambda_1+2\lambda_2)(\lambda_1+3\lambda_2)} .$$

$$\frac{\partial P_x}{\partial\lambda_1}\frac{\partial^2 P_x}{\partial\lambda_2^2} = (\bar{P}_{x-1}-P_x)\left(-\frac{xP_x}{\lambda_2^2}+\left(\frac{1-\lambda_2}{\lambda_2}\right)^2 x^2P_x - \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}x\bar{P}_{x-1}\right.$$

(2.39.a)

$$\left.+\frac{\lambda_1}{\lambda_2^2}\bar{P}_{x-1}-\frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}(x-1)\bar{P}_{x-1}+\frac{\lambda_1^2}{\lambda_2^2}\overset{+}{P}_{x-2}\right) .$$

Multiplying both sides of (2.39.a) by P^{-2}, and taking expectations, we have

$$P_{1,22} = -\frac{1}{\lambda_2^2}\sum x\bar{P}_{x-1}+\frac{1}{\lambda_2^2}\sum xP_x+\frac{(1-\lambda_2)^2}{\lambda_2^2}\sum x^2\bar{P}_{x-1}-\frac{(1-\lambda_2)^2}{\lambda_2^2}\sum x^2P_x$$

$$-\frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\sum x\frac{\bar{P}^2_{x-1}}{P_x}+\frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\sum x\bar{P}_{x-1}+\frac{\lambda_1}{\lambda_2^2}\sum\frac{\bar{P}^2_{x-1}}{P_x}$$

(2.39.b)

$$-\frac{\lambda_1}{\lambda_2^2}\sum\bar{P}_{x-1}-\frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\sum(x-1)\frac{\bar{P}^2_{x-1}}{P_x}+\frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\sum(x-1)\bar{P}_{x-1}$$

$$+\frac{\lambda_1^2}{\lambda_2^2}C_{12}-\frac{\lambda_1^2}{\lambda_2^2}\sum\overset{+}{P}_{x-2}$$

or

$$P_{1,22} = \frac{-(1+\lambda_1)}{\lambda_2^2(1-\lambda_2)}+\frac{\lambda_1}{\lambda_2^2(1-\lambda_2)}+\frac{(1-\lambda_2)^2}{\lambda_2^2}\left[\frac{\lambda_1+\lambda_2}{(1-\lambda_2)^3}+\frac{(\lambda_1+\lambda_2)^2}{(1-\lambda_2)^2}+\frac{2(\lambda_1+\lambda_2)}{(1-\lambda_2)}+1\right]$$

(2.39.c)

$$-\frac{(1-\lambda_2)^2}{\lambda_2^2}\left[\frac{\lambda_1}{(1-\lambda_2)^3}+\frac{\lambda_1^2}{(1-\lambda_2)^2}\right]-\frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\left[\frac{(\lambda_1+\lambda_2)^2}{\lambda_1}\left(\frac{2}{\lambda_1+2\lambda_2}+\right.\right.$$

$$+ \frac{1}{1-\lambda_2}\Big) + (P_{11}+1)\Big] + \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\left[\frac{1+\lambda_1}{1-\lambda_2}\right] + \frac{\lambda_1}{\lambda_2^2}(P_{11}+1) - \frac{\lambda_1}{\lambda_2^2}$$

$$(2.39.c) \quad - \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\left[\frac{(\lambda_1+\lambda_2)^2}{\lambda_1}\left(\frac{2}{\lambda_1+2\lambda_2} + \frac{1}{1-\lambda_2}\right) + \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}\left(\frac{\lambda_1+\lambda_2}{1-\lambda_2}\right)\right.$$

$$+ \frac{\lambda_1^2}{\lambda_2^2}\left[\frac{2(\lambda_1+\lambda_2)}{\lambda_1(\lambda_1+2\lambda_2)} + \frac{\lambda_1+\lambda_2}{\lambda_1+3\lambda_2}\right] - \frac{\lambda_1^2}{\lambda_2^2} \quad .$$

The last expression can be shown to be equal to

$$(2.39.d) \qquad P_{1,22} = \frac{3\lambda_1^2\lambda_2^2-\lambda_1^2\lambda_2-5\lambda_1\lambda_2^2+3\lambda_1\lambda_2^3-6\lambda_2^3}{\lambda_2^2(\lambda_1+2\lambda_2)(\lambda_1+3\lambda_2)} \quad .$$

$$\frac{\partial P_x}{\partial\lambda_1}\frac{\partial^2 P_x}{\partial\lambda_1\partial\lambda_2} = (\bar{P}_{x-1}-P_x)\left(\frac{1-\lambda_2}{\lambda_2}x\bar{P}_{x-1} - \frac{1-\lambda_2}{\lambda_2}xP_x - \frac{\lambda_1}{\lambda_2}\overset{+}{P}_{x-2} - \frac{1-\lambda_1}{\lambda_2}\bar{P}_{x-1}\right)$$

$$(2.40.a) \qquad = \frac{1-\lambda_2}{\lambda_2}x\bar{P}^2_{x-1} - \frac{1-\lambda_2}{\lambda_2}xP_x\bar{P}_{x-1} - \frac{1-\lambda_2}{\lambda_2}xP_x\bar{P}_{x-1} + \frac{1-\lambda_2}{\lambda_2}xP_x^2$$

$$- \frac{\lambda_1}{\lambda_2}\bar{P}_{x-1}\overset{+}{P}_{x-2} + \frac{\lambda_1}{\lambda_2}P_x\overset{+}{P}_{x-2} - \frac{1-\lambda_1}{\lambda_2}\bar{P}^2_{x-1} + \frac{1-\lambda_1}{\lambda_2}P_x\bar{P}_{x-1} \quad .$$

Thus, dividing both sides by P^2 and taking the expectations, we get

$$P_{1,12} = \frac{1-\lambda_2}{\lambda_2}\sum x\frac{\bar{P}^2_{x-1}}{P_x} - 2\,\frac{1-\lambda_2}{\lambda_2}\sum x\bar{P}_{x-1} + \frac{1-\lambda_2}{\lambda_2}\sum xP_x - \frac{\lambda_1}{\lambda_2}\sum\frac{\bar{P}_{x-1}\overset{+}{P}_{x-2}}{P_x}$$

(2.40.b)

$$+ \frac{\lambda_1}{\lambda_2}\sum\overset{+}{P}_{x-2} - \frac{1-\lambda_1}{\lambda_2}\sum\frac{\bar{P}^2_{x-1}}{P_x} + \frac{1-\lambda_1}{\lambda_2}\sum\bar{P}_{x-1} \quad .$$

To derive a closed form for $\sum xP^2_{x-1}/P_x$ we proceed as follows:

On multiplying both sides of (2.21.c) by P_{x-1}/λ_2, we get

$$\frac{1}{\lambda_2}\, xP_x\bar{P}_{x-1} = \frac{\lambda_1}{\lambda_1+\lambda_2}\,(x-1)\bar{P}^2_{x-1} + \frac{\lambda_1}{\lambda_2}\,\bar{P}^2_{x-1} \quad ,$$

from which

$$\frac{\lambda_1}{\lambda_1+\lambda_2}\sum (x-1)\,\frac{\bar{P}^2_{x-1}}{P_x} = \frac{1+\lambda_1}{\lambda_2(1-\lambda_2)} - \frac{\lambda_1}{\lambda_2}\,(P_{11}+1)$$

or

$$\sum (x-1)\,\frac{\bar{P}^2_{x-1}}{P_x} = \frac{(\lambda_1+\lambda_2)(1+\lambda_1)}{\lambda_1\lambda_2(1-\lambda_2)} - \frac{\lambda_1+\lambda_2}{\lambda_2}\,(P_{11}+1)$$

$$= \frac{(\lambda_1+\lambda_2)^2}{\lambda_1}\left(\frac{2}{\lambda_1+2\lambda_2} + \frac{1}{1-\lambda_2}\right)$$

Hence

$$(2.40.c) \qquad \sum x\,\frac{\bar{P}^2_{x-1}}{P_x} = \frac{(\lambda_1+\lambda_2)^2}{\lambda_1}\left(\frac{2}{\lambda_1+2\lambda_2} + \frac{1}{1-\lambda_2}\right) + (P_{11}+1) \quad .$$

Substituting (2.40.c) and (2.36.e) in (2.40.b), we have

$$P_{1,12} = \frac{1-\lambda_2}{\lambda_2}\left[\frac{(\lambda_1+\lambda_2)^2}{\lambda_1}\left(\frac{2}{\lambda_1+2\lambda_2} + \frac{1}{1-\lambda_2}\right) + (P_{11}+1)\right] - \frac{2(1-\lambda_2)}{\lambda_2}\left(\frac{1+\lambda_1}{1-\lambda_2}\right)$$

$$+ \frac{1-\lambda_2}{\lambda_2}\left(\frac{\lambda_1}{1-\lambda_2}\right) - \frac{\lambda_1}{\lambda_2}\left[\frac{2(\lambda_1+\lambda_2)}{\lambda_1(\lambda_1+2\lambda_2)} + \frac{\lambda_1+\lambda_2}{\lambda_1+3\lambda_2}\right] + \frac{\lambda_1}{\lambda_2} - \frac{1-\lambda_1}{\lambda_2}\,(P_{11}+1) + \frac{1-\lambda_1}{\lambda_2} \; ,$$

which, after some simplifications, will give

$$P_{1,12} = \frac{(\lambda_2^2+\lambda_1^3+3\lambda_1^2\lambda_2-\lambda_1^2-2\lambda_1\lambda_2)}{\lambda_2(\lambda_1+2\lambda_2)(\lambda_1+3\lambda_2)} \quad . \tag{2.40.d}$$

Finally,

$$\frac{\partial P_x}{\partial\lambda_2}\frac{\partial^2 P_x}{\partial\lambda_2^2} = \left(\frac{1-\lambda_2}{\lambda_2} xP_x - \frac{\lambda_1}{\lambda_2}\bar{P}_{x-1}\right)\left(\frac{-xP_x}{\lambda_2^2} + \frac{(1-\lambda_2)^2}{\lambda_2^2} x^2P_x\right.$$

$$- \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2} x\bar{P}_{x-1} + \frac{\lambda_1}{\lambda_2^2}\bar{P}_{x-1} - \frac{\lambda_1(1-\lambda_2)}{\lambda_2^2}(x-1)\bar{P}_{x-1} \tag{2.41.a}$$

$$\left. + \frac{\lambda_1^2}{\lambda_2^2}\overset{+}{P}_{x-2}\right) \quad .$$

Multiplying the brackets in (2.41.a) and dividing both sides by P^2, and then taking expectations, one can write

$$P_{2,22} = \frac{-(1-\lambda_2)}{\lambda_2^3}\sum x^2P_x + \frac{\lambda_1}{\lambda_2^3}\sum xP_{x-1} + \left(\frac{1-\lambda_2}{\lambda_2}\right)^3 \sum x^2P_x$$

$$- \frac{2\lambda_1(1-\lambda_2)^2}{\lambda_2^3}\sum x^2\bar{P}_{x-1} + \frac{\lambda_1^2(1-\lambda_2)}{\lambda_2^3}\sum x\frac{\bar{P}^2_{x-1}}{P_x}$$

(2.41.b)

$$+ \frac{\lambda_1(1-\lambda_2)}{\lambda_2^3}\sum x\bar{P}_{x-1} - \frac{\lambda_1^2}{\lambda_2^3}\sum\frac{P^2_{x-1}}{P_x} - \frac{\lambda_1(1-\lambda_2)^2}{\lambda_2^3}\sum x(x-1)\bar{P}_{x-1}$$

$$+ \frac{\lambda_1^2(1-\lambda_2)}{\lambda_2^3}\sum(x-1)\frac{\bar{P}^2_{x-1}}{P_x} + \frac{\lambda_1^2(1-\lambda_2)}{\lambda_2^3}\sum x\overset{+}{P}_{x-2} - \frac{\lambda_1^3}{\lambda_2^3}\sum\frac{\bar{P}_{x-1}\overset{+}{P}_{x-2}}{P_x} \quad .$$

On substituting the computed values for

$$\sum \frac{\bar{P}_{x-1}\overset{+}{P}_{x-2}}{P_x}, \qquad \sum x\,\frac{\bar{P}^2_{x-1}}{P_x}, \qquad \text{and} \qquad \sum \frac{\bar{P}^2_{x-1}}{P_x},$$

in (2.41.b), we get

$$\begin{aligned}
P_{2,22} = {} & \frac{-(1-\lambda_2)}{\lambda_2^3}\left[\frac{\lambda_1}{(1-\lambda_2)^3}+\frac{\lambda_1^2}{(1-\lambda_2)^2}\right]+\frac{\lambda_1}{\lambda_2^3}\left[\frac{1+\lambda_1}{1-\lambda_2}\right]+\left(\frac{1-\lambda_2}{\lambda_2}\right)^3\left[\frac{\lambda_1(1+2\lambda_2)}{(1-\lambda_2)^5}\right. \\
& \left. +\frac{3\lambda_1}{1-\lambda_2}\left(\frac{\lambda_1}{(1-\lambda_2)^3}+\frac{\lambda_1^2}{(1-\lambda_2)^2}\right)-\frac{2\lambda_1^3}{(1-\lambda_2)^3}\right]-\frac{2\lambda_1(1-\lambda_2)^2}{\lambda_2^3}\left[\frac{\lambda_1+\lambda_2}{(1-\lambda_2)^3}\right. \\
& \left. +\frac{(\lambda_1+\lambda_2)^2}{(1-\lambda_2)^2}+\frac{2(\lambda_1+\lambda_2)}{1-\lambda_2}+1\right]+\frac{\lambda_1^2(1-\lambda_2)}{\lambda_2^3}\left[(P_{11}+1)+\frac{(\lambda_1+\lambda_2)^2}{\lambda_1}\left(\frac{2}{\lambda_1+2\lambda_2}\right.\right. \\
\text{(2.41.c)} \quad & \left.\left. +\frac{1}{1-\lambda_2}\right)\right]+\frac{\lambda_1(1-\lambda_2)}{\lambda_2^3}\left[\frac{1+\lambda_1}{1-\lambda_2}\right]-\frac{\lambda_1^2}{\lambda_2^3}\left[P_{11}+1\right]-\frac{\lambda_1(1-\lambda_2)^2}{\lambda_2^3}\left[\frac{\lambda_1+\lambda_2}{1-\lambda_2}\right. \\
& \left. +\frac{\lambda_1+\lambda_2}{(1-\lambda_2)^3}+\frac{(\lambda_1+\lambda_2)^2}{(1-\lambda_2)^2}\right]+\frac{\lambda_1^2(1-\lambda_2)}{\lambda_2^3}\left[\frac{(\lambda_1+\lambda_2)^2}{\lambda_1}\left(\frac{2}{\lambda_1+2\lambda_2}+\frac{1}{1-\lambda_2}\right)\right] \\
& +\frac{\lambda_1^2(1-\lambda_2)}{\lambda_2^3}\left[\frac{\lambda_1+2\lambda_2}{1-\lambda_2}+2\right]-\frac{\lambda_1^3}{\lambda_2^3}\left[\frac{2(\lambda_1+\lambda_2)}{\lambda_1(\lambda_1+2\lambda_2)}+\frac{(\lambda_1+\lambda_2)}{\lambda_1+3\lambda_2}\right] .
\end{aligned}$$

The last expression can be simplified and written as

$$\begin{aligned}
P_{2,22} = {} & [\lambda_2^3(1-\lambda_2)^2(\lambda_1+2\lambda_2)(\lambda_1+3\lambda_2)]^{-1}[3\lambda_1^2\lambda_2-8\lambda_1^3\lambda_2^2+18\lambda_1^3\lambda_2^3-11\lambda_1^3\lambda_2^4+15\lambda_1^2\lambda_2^2 \\
\text{(2.41.d)} \quad & -38\lambda_1^2\lambda_2^3+66\lambda_1^2\lambda_2^4-43\lambda_1^2\lambda_2^5+10\lambda_1^2\lambda_2^6+18\lambda_1\lambda_2^3-42\lambda_1\lambda_2^4+72\lambda_1\lambda_2^5-48\lambda_1\lambda_2^6 \\
& +12\lambda_1\lambda_2^7] .
\end{aligned}$$

Thus, on solving (2.34.a) for $b_1(\hat{\lambda}_1)$ and $b_2(\hat{\lambda}_2)$, we have

$$b_1(\hat{\lambda}_1) = (L_2P_{12} - L_1P_{22})/2|I|^2 + 0(N^{-2})$$

(2.42)

$$b_2(\hat{\lambda}_2) = (L_1P_{12} - L_2P_{11})/2|I|^2 + 0(N^{-2})$$

where L_1 and L_2, which are defined in (2.34.b) can be explicitly obtained on using the calculated values of $(P_{\alpha,ij})$, $(\alpha,i,j = 1,2)$.

2.4 Joint Asymptotic Efficiency of Moment Estimators.

We shall determine the asymptotic relative efficiency of the moment estimators relative to the M.L. estimators. The variance covariance determinant of the moment estimators up to the second order accuracy is given by

$$\text{(2.43)} \qquad \begin{vmatrix} \mathrm{Var}(\lambda_1^*) & \mathrm{Cov}(\lambda_1^*,\lambda_2^*) \\ \mathrm{Cov}(\lambda_1^*,\lambda_2^*) & \mathrm{Var}(\lambda_2^*) \end{vmatrix} \approx \begin{vmatrix} \frac{v_1}{n} + n^{-2}\sum_{i=1}^{11}\alpha_i & \frac{c_1}{n} + n^{-2}\sum_{i=1}^{11}\omega_j \\ \frac{c_1}{n} + n^{-2}\sum_{i=1}^{11}\omega_j & \frac{v_1^*}{n} + n^{-2}\sum_{i=1}^{11}\alpha_i^* \end{vmatrix} .$$

where ν_1, ν_1^*, c_1, α_2, α_i^*, ω_j are explicitly stated in (2.16.b), (2.18), and (2.20), respectively.

To make a realistic comparison between the two types of estimators it is necessary that the values of the covariance matrices of both types of estimators, be computed up to the same degree of accuracy. Since the determination of the second order terms of the variances and covariance of the M.L. estimators is laborious and include huge numbers of terms, we shall confine ourselves to the first order accuracy, and obtain the first order asymptotic efficiency.

Fisher's measure of efficiency is given by

(2.44) $$e_{ff} = \frac{1}{n^2|I||D|}$$

where $|I|$ is given by (2.30), and $|D|$ is the generalized variance of the moment estimators computed up to the first order accuracy. Katti (37), has shown that

(2.45) $$|D| = \frac{|C(\mu_1',\mu_2)|}{|J|^2}$$

where

(2.46) $$|C(\mu_1',\mu_2)| = \begin{vmatrix} \mathrm{Var}(m_1') & \mathrm{Cov}(m_1',m_2) \\ \mathrm{Cov}(m_1',m_2) & \mathrm{Var}(m_2) \end{vmatrix}$$

$$= \frac{1}{n^2}(\mu_2\mu_4-\mu_3^2-\mu_2^3) = \frac{2\lambda_1^2}{n^2(1-\lambda_2)^{10}}(\lambda_1+2\lambda_2+\lambda_2^2-\lambda_1\lambda_2)$$

and

(2.47) $$|J| = \begin{vmatrix} \partial\mu_1'/\partial\lambda_1 & \partial\mu_1'/\partial\lambda_2 \\ \partial\mu_2/\partial\lambda_1 & \partial\mu_2/\partial\lambda_2 \end{vmatrix} = 2\lambda_1/(1-\lambda_2)^5 .$$

Thus

$$|D| = \frac{1}{2n^2}(\lambda_1+2\lambda_2+\lambda_2^2-\lambda_1\lambda_2) ,$$

and one can show that the asymptotic efficiency is given as

(2.48) $$e_{ff} = \frac{\lambda_1(1-\lambda_2)+2\lambda_2(1-\lambda_2)}{\lambda_1(1-\lambda_2)+2\lambda_2+\lambda_2^2} > 0$$

or

(2.49) $$e_{ff} = 1 - \frac{3\lambda_2^2}{\lambda_1(1-\lambda_2)+2\lambda_2+\lambda_2^2} < 1 \ .$$

Solving (2.49) for λ_1, one gets

(2.50) $$\lambda_1 = \frac{3\lambda_2^2}{(1-\lambda_2)(1-e_{ff})} - \frac{2\lambda_2(1+\lambda_2)}{(1-\lambda_2)} \ .$$

Equation (2.50) is very helpful in constructing the contours of efficiencies (see Fig. 1). From equation (2.49), it can be easily realized that e_{ff} is always less than one, and it is a monotonic increasing function of λ_1 if λ_2 is held constant. Moreover, for constant λ_1, e_{ff} is a monotonic decreasing function of λ_2.

Table (2.1) gives tabulations for the efficiencies for some selected values of the parameter space of (λ_1,λ_2).

Figure 1

Contours of the Asymptotic Efficiency of the Method of Moments

Relative to the M.L. Method.

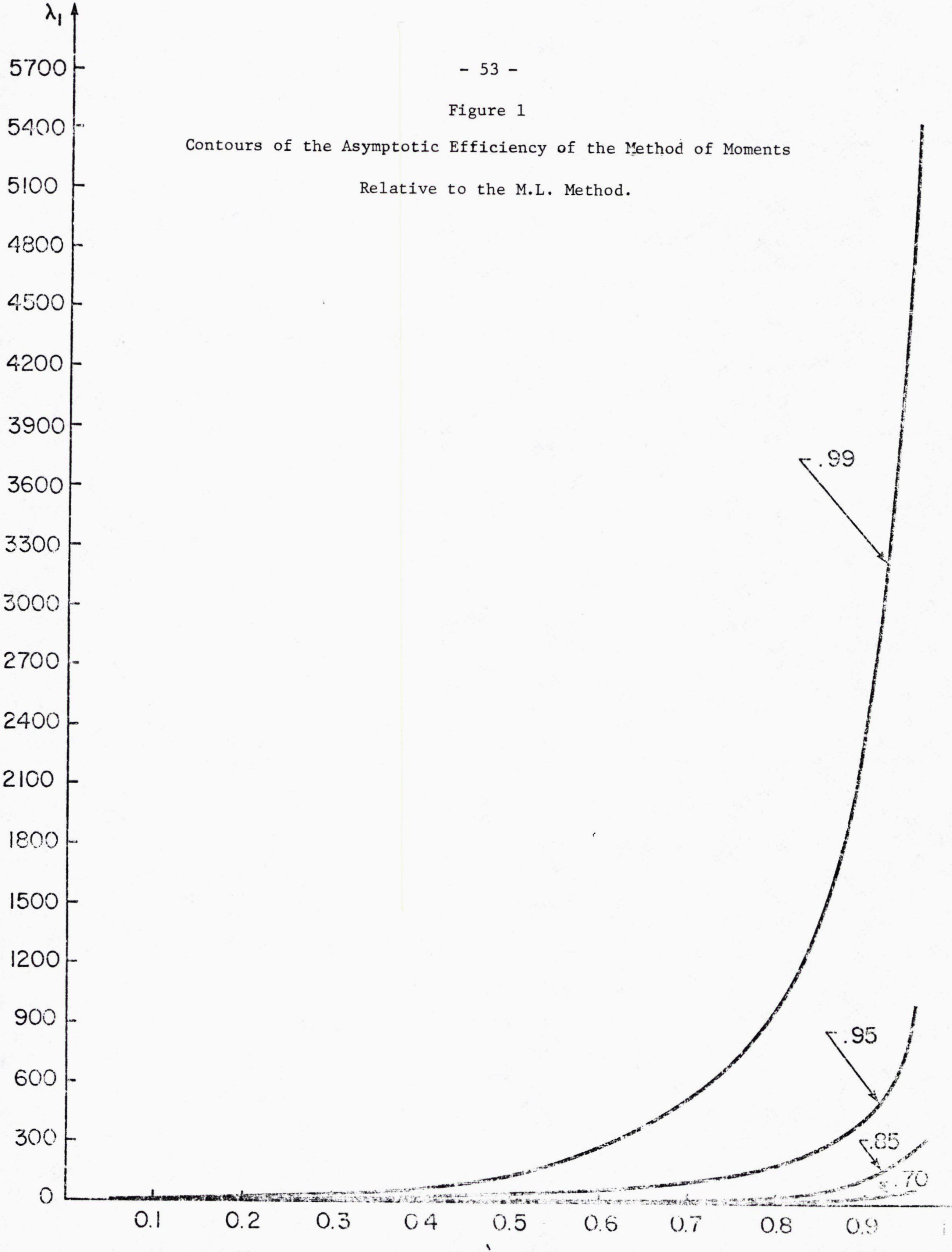

Table (2.1)
Asymptotic Efficiency of the Moment Estimators
Relative to M.L. Estimators

λ_1 \ λ_2	.1	.2	.3	.4	.5	.6	.7	.8
.5	.956	.857	.740	.619	.500	.386	.379	.179
1	.973	.903	.806	.692	.571	.449	.329	.213
2	.985	.941	.871	.778	.667	.542	.410	.273
4	.992	.967	.923	.857	.769	.658	.524	.368
4.5	.993	.970	.930	.869	.786	.679	.546	.388
6	.995	.977	.945	.895	.824	.727	.602	.442
8	.996	.983	.957	.917	.857	.773	.657	.500
10	.997	.986	.965	.931	.880	.806	.699	.547

Table (2.1) reveals the regions of poor and high efficiencies. For small values of λ_1 and large values of λ_2 (e.g. $\lambda_1 < 4$ and $\lambda_2 \geq .3$), the efficiency is poor and decreases when λ_2 gets larger. On the other hand, for large values of λ_1 and small values of λ_2 ($\lambda_2 < .3$), efficiency is appreciable.

2.5 Maximum Likelihood Estimation From Truncated Samples.

In section 2.3, we investigated the M.L. estimators of λ_1 and λ_2 when the sample observation is permitted over the full range of the complete distribution. When the sample observations are truncated, as for example, when the number of zero observations is unknown, or when the observations of higher counts are pooled, the estimation problem increases in complexity.

In what follows, we present an investigation into the precision of the samples in the estimation of the parameters in the case of decapitated and double truncated samples of size n relative to a complete sample of the same size. It should be realized that the precision of a truncated sample of size n relative to a complete sample of the same size depends on the form of the frequency function, the parameters to be estimated, and the points of truncation.

The probability function of a GPD truncated on the left at $x = c$ and on the right at $x = d$, (c and d are non-negative integers, and $c < d$), may be written as

$$(2.51) \quad P_x^* = P_x^*(\lambda_1,\lambda_2) = \begin{cases} 0 & x<d \\ (Q_c(\lambda_1,\lambda_2)-Q_{d+1}(\lambda_1,\lambda_2))^{-1}P_x(\lambda_1,\lambda_2) & c\leq x\leq d \\ 0 & x>d \end{cases}$$

where

$$(2.52) \qquad Q_a(\lambda_1,\lambda_2) = \sum_{x=a}^{\infty} P_x(\lambda_1,\lambda_2) \quad .$$

The truncated distribution (2.51) is thus normalized so that

$$(2.53) \qquad \sum_{x=0}^{\infty} P_x^* = \sum_{x=c}^{d} P_x^* = 1 \ .$$

A measure of the accuracy of M.L. estimators from a truncated sample of size n relative to the M.L. estimators obtained from a complete sample of the same size, was given by Swamy (77) and is defined by

(2.54) $$e^*_{ff} = |I^*|/|I|$$

where $|I|$ is defined in (2.26.c), and $|I^*|$ is the amount of informatin supplied by a single observation obtained from the truncated distribution (2.51), and is defined as

(2.55) $$|I^*| = P^*_{11}P^*_{22} - (P^*_{12})^2$$

where

$$P^*_{ij} = E\left(P^{*-2}\left(\frac{\partial P^*}{\partial \lambda_i}\,\frac{\partial P^*}{\partial \lambda_j}\right)\right), \qquad (i,j = 1,2)$$

and P^* are the probabilities given in (2.51). To calculate the elements P^*_{ij}, we differentiate (2.51) partially w.r.t. λ_i. Thus the partial derivative of P^* w.r.t. λ_i can be written as

(2.56) $$\frac{\partial P^*}{\partial \lambda_i} = \frac{\partial P/\partial \lambda_i}{Q_c - Q_{d+1}} - \frac{\partial Q_c/\partial \lambda_i - \partial Q_{d+1}/\partial \lambda_i}{(Q_c - Q_{d+1})^2}\,P$$

where $Q_a(\lambda_1,\lambda_2)$ is written as Q_a for abbreviation. Thus, using (2.56), along with the recurrence relations given in (2.21.a) and (2.21.b), we have

(2.57.a) $$P^*_{11} = P_{11} + \frac{(\partial Q_c/\partial \lambda_1 - \partial Q_{d+1}/\partial \lambda_1)^2}{(Q_c - Q_{d+1})^2}\;,$$

(2.57.b) $$P^*_{22} = P_{11} + \frac{(\partial Q_c/\partial \lambda_2 - \partial Q_{d+1}/\partial \lambda_2)^2}{(Q_c - Q_{d+1})^2}\;,$$

$$P^*_{12} = P_{12} + \frac{(\partial Q_c/\partial\lambda_1 - \partial Q_{d+1}/\partial\lambda_2)(\partial Q_c/\partial\lambda_2 - \partial Q_{d+1}/\partial\lambda_2)}{(Q_c - Q_{d+1})^2}\ , \tag{2.57.c}$$

and

$$\begin{aligned} |I^*| = |I| &+ P_{11}\left[\frac{\partial Q_c/\partial\lambda_2 - \partial Q_{d+1}/\partial\lambda_2}{Q_c - Q_{d+1}}\right]^2 \\ &- 2P_{12}\left[\frac{(\partial Q_c/\partial\lambda_1 - \partial Q_{d+1}/\partial\lambda_1)(\partial Q_c/\partial\lambda_2 - \partial Q_{d+1}/\partial\lambda_2)}{Q_c - Q_{d+1}}\right]^2 \\ &+ P_{22}\left[\frac{\partial Q_c/\partial\lambda_1 - \partial Q_{d+1}/\partial\lambda_1}{Q_c - Q_{d+1}}\right]^2 \ . \end{aligned} \tag{2.57.d}$$

By the definition of e^*_{ff} given in (2.54), we have

$$\begin{aligned} e^*_{ff} = 1 &+ \mathrm{Var}(\hat\lambda_2)\left(\frac{Q^*_2}{Q}\right)^2 + 2\mathrm{Cov}(\lambda^*_1,\hat\lambda_2)\left(\frac{Q^*_1}{Q}\right)\left(\frac{Q^*_2}{Q}\right) \\ &+ \mathrm{Var}(\hat\lambda_1)\left(\frac{Q^*_1}{Q}\right)^2 = 1 + \tau(\lambda_1,\lambda_2) \end{aligned} \tag{2.58}$$

where

$$\begin{aligned} Q = Q_c - Q_{d+1} &= \sum_{x=c}^{\infty} P_x(\lambda_1,\lambda_2) - \sum_{x=d+1}^{\infty} P_x(\lambda_1,\lambda_2) \\ &= \sum_{x=c}^{d} P_x(\lambda_1,\lambda_2) \ . \end{aligned} \tag{2.59.a}$$

$$Q_1^* = \frac{\partial}{\partial\lambda_1}\left[Q_c - Q_{d+1}\right] = \sum_{x=c}^{d} \frac{\partial P_x(\lambda_1,\lambda_2)}{\partial\lambda_1}$$

(2.59.b)
$$= \sum_{x=c}^{d} \left[P_{x-1}(\lambda_1+\lambda_2,\lambda_2) - P_x(\lambda_1,\lambda_2)\right]$$

$$= \sum_{x=c}^{d} P_{x-1}(\lambda_1+\lambda_2,\lambda_2) - Q \quad .$$

$$Q_2^* = \frac{\partial}{\partial\lambda_2}\left[Q_c - Q_{d+1}\right] = \sum_{x=c}^{d} \frac{\partial P_x(\lambda_1,\lambda_2)}{\partial\lambda_2} = \frac{(1-\lambda_2)}{\lambda_2}\sum_{x=c}^{d} xP_x(\lambda_1,\lambda_2)$$

(2.59.c)
$$- \frac{\lambda_1}{\lambda_2}\sum_{x=c}^{d} P_{x-1}(\lambda_1+\lambda_2,\lambda_2) \quad .$$

The investigation of the behaviour of the function $\tau(\lambda_1,\lambda_2)$ seems to be difficult, since $\mathrm{Cov}(\hat{\lambda}_1,\hat{\lambda}_2)$ is always negative as can be seen from (2.33). Thus, the gain or the loss in the efficiency will depend upon the sign of $\tau(\lambda_1,\lambda_2)$, whose value, in turn, will depend upon the values of the parameters λ_1, λ_2, and the points of truncation. However, we shall show that in the case of decapitated samples, there is a certain gain in the efficiency.

For $c = 1$ and $d = \infty$, i.e. in the case where the distribution is truncated at zero, we have

$$(2.60.a) \qquad Q = \sum_{x=1}^{\infty} P_x(\lambda_1,\lambda_2) = 1 - P_0(\lambda_1,\lambda_2) = 1 - e^{-\lambda_1}$$

$$(2.60.b) \qquad Q_1^* = \sum_{x=1}^{\infty} P_{x-1}(\lambda_1+\lambda_2,\lambda_2) - Q = e^{-\lambda_1}$$

and

$$Q_2^* = \frac{(1-\lambda_2)}{\lambda_2} \sum_{x=1}^{\infty} xP_x(\lambda_1,\lambda_2) - \frac{\lambda_1}{\lambda_2} \sum_{x=1}^{\infty} P_{x-1}(\lambda_1+\lambda_2,\lambda_2)$$

$$= \frac{1-\lambda_2}{\lambda_2}\left(\frac{\lambda_1}{1-\lambda_2}\right) - \frac{\lambda_1}{\lambda_2} = 0 \ .$$

Thus

$$e_{ff}^* = 1 + \frac{\mathrm{Var}(\lambda_1)}{(e^{\lambda_1}-1)^2} \quad ,$$

and from (2.23)

$$(2.61) \qquad e_{ff}^* = 1 + \frac{\lambda_1(2+\lambda_1)}{2(e^{\lambda_1}-1)^2} \quad ,$$

which approaches unity, for sufficiently large values of λ_1.

We conclude then, from (2.61), that there is a definite gain in the efficiency of M.L. estimators if the parent population is left truncated at the origin.

2.6. Sampling Properties of the M.L. and the Moment Estimators of the GPD.

In this section we compare the performance of the M.L. and the moment estimator, through numerical tabulations for first order terms of the biases only, because we have not derived higher order terms in the biases of the M.L. estimators. However, we shall show by some numerical examples that the second order terms of the moment estimators are of negligible importance for relatively large samples, and the appropriately chosen subregion of the parameter space.

Table 2.2.

In this table we give the biases of the M.L. (2.42) and the moment estimator (2.10) of the parameter λ_1. The first entry of each cell is the coefficient of n^{-1} in the bias of the M.L. estimator, and the second entry is the coefficient of n^{-1} in the bias of the moment estimator. The selected parameter values are

λ_1 = 0.1, 0.2, 0.4, 0.6, 0.8

λ_2 = 0.5, 1.5, 3.0, 6.0, 10.0.

(Table 2.2)

Biases of M.L. and Moment Estimators of λ_1

$\lambda_1 \backslash \lambda_2$	.1	.2	.4	.6	.8
0.5	16.41 .82	4.37 1.11	1.50 2.23	0.93 4.90	0.68 13.83
1.5	55.41 2.07	14.77 2.36	4.77 3.48	2.81 6.15	2.01 15.08
3.0	225.18 3.94	80.62 4.24	28.82 5.35	14.12 8.03	7.06 16.95
6.0	2777.31 7.69	1189.99 7.99	439.09 9.10	195.52 11.78	75.65 20.70
10.0	21691.21 12.69	9526.74 12.99	3530.45 14.10	1555.95 16.78	581.76 25.70

For all the selected values of λ_1 and λ_2, the biases of the moment estimator of λ_1 increase monotonically by increasing λ_1 and λ_2. But for the M.L. estimator, the bias increases by increasing λ_1 only, and decreases by increasing λ_2. Over the interval $\lambda_2 < .4$ and $.5 \le \lambda_1 \le 1.5$, the moment estimator is less biased than the M.L. estimator, but for $\lambda_2 \ge .4$ and the same interval of λ_1, the moment estimator is more biased than that of the M.L. For $\lambda_1 \ge 3$ and all the tabulated values of λ_2, the moment estimator is less biased than the M.L. estimator. In fact, the region $\lambda_2 < .4$ and $\lambda_1 \ge 3.0$, where the moment estimator is less biased than the M.L. estimator, is the region where the moment estimators are highly efficient as can be seen from Table (2.1).

Table 2.3.

The first order bias of λ_2 on a per observation basis, calculated from (2.42) are shown in Table (2.3). The selected values of the parameters are exactly the same as in Table (2.2).

(Table 2.3)

Biases of M.L. and Moment Estimators of λ_2

λ_1 \ λ_2	.1	.2	.4	.6	.8
0.5	-39.79 -1.67	-11.53 -2.18	-4.11 -3.47	-2.69 -5.12	-2.24 -7.13
1.5	-32.27 -1.31	-7.65 -1.39	-2.18 -1.66	-1.23 -2.04	-.84 -2.54
3.0	-43.85 -1.22	-12.03 -1.20	-3.25 -1.20	-1.30 -1.27	-.54 -1.40
6.0	-294.7 -1.17	-111.00 -1.10	-30.70 -.98	-9.20 -.89	-1.84 -.82
10	-1588.59 -1.15	-619.38 -1.06	-172.15 -.89	-50.59 -.73	-9.46 -.59

As can be seen, the biases of both kind of estimators are negative throughout the tabulated region. However, we shall compare the absolute values of the biases of these estimators. For $\lambda_2 < .4$ and all the tabulated values of λ_1, the bias of the moment estimator of λ_2 is less than that of the M.L. estimator. In fact, the M.L. estimator of λ_2 is less biased than that of the moment estimator in a relatively small subregion, which is described in the table by the rectangle $.5 \le \lambda_1 \le 1.5$ and $.6 \le \lambda_2 \le .8$.

We shall give now some of the values of the second order term in the biases of moment estimators.

(i) For the parameter λ_1, suppose that (2.10) is written as

$$\text{bias}(\lambda_1^*) \doteqdot K_1/n + \sum_{i=1}^{11} \pi_i/n^2 + \ldots$$

Thus, when $\lambda_1 = .5$, $\lambda_2 = .2$,

$$n^2 \sum_{i=1}^{11} \pi_i = -.01 \ .$$

When $\lambda_1 = 3$, $\lambda_2 = .4$,

$$n^2 \sum_{i=1}^{11} \pi_i = +.14 \ .$$

And when $\lambda_1 = 10$, $\lambda_2 = .8$,

$$n^2 \sum_{i=1}^{11} \pi_i = +.82.$$

(ii) Similarly, if we write (2.13) in the form

$$\text{bias}(\lambda_2^*) \doteqdot K_1^*/n + \sum_{i=1}^{11} \pi_i^*/n^2 + \ldots$$

Then, for $\lambda_1 = .5$, $\lambda_2 = .2$,

$$n^2 \sum_{i=1}^{11} \pi_i^* = +.14$$

When $\lambda_1 = 3$, $\lambda_2 = .4$,

$$n^2 \sum_{i=1}^{11} \pi_i^* = .01.$$

And when $\lambda_1 = 10$, $\lambda_2 = .8$,

$$n^2 \sum_{i=1}^{11} \pi_i^* = .02.$$

2.7 Concluding Remarks.

In this chapter we have studied two types of estimators for the parameters λ_1 and λ_2 of the GPD family. The M.L. equations do not give an explicit solution, and the calculation of the M.L. estimators will be obtained as approximate values, by using the Newton-Raphson iteration technique. The derivation of the first order biases and variances of the M.L. estimator was very tedious and we have presented the calculations for ready reference and to show the amount of work involved in the calculations. As we can see from the tables of biases of the M.L. estimators, there are some large values for the biases of $\hat{\lambda}_1$ and $\hat{\lambda}_2$. For example, from Tables (2.2) and (2.3)

$$n \cdot b_1(\hat{\lambda}_1) = 21691.21 \qquad \text{when } \lambda_1 = 10, \quad \lambda_2 = .1$$

$$n \cdot b_2(\hat{\lambda}_2) = -1588.59 \qquad \text{when } \lambda_1 = 10, \quad \lambda_2 = .1$$

which means a very large sample size is needed to reduce the effect of these inflated values. From truncated samples, the derivation of the M.L. estimators and their biases and variances is exactly similar to the derivation in the case of complete samples of the same size, so we have not derived such quantities. The accuracy of these estimators is measured in the sense of their relative efficiency with respect to the M.L. method of estimation from a complete sample of the same size, by using the definition of accuracy given by Swamy. As was realized by Cohen (9), the zero class in many cases, and specially for the Poisson distribution, contains large amounts of recording errors. We have confirmed that such a class should be removed when we start estimating by the method of M.L. depending upon the claim given by Swamy (77).

The problem of estimation using the sample moments is also a difficult one. We presented the general form of the Taylor expansion of a function of the two sample moments and we derived its asymptotic bias, variance and its covariance with another function of the same sample moments that has the same general form of the Taylor expansion. The results were given asymptotically and accurate up to order n^{-2}, and we specialized these results for the moment estimators λ_1^*, λ_2^* of the parameters λ_1 and λ_2. We have also given some comparisons between the asymptotic biases of the M.L. estimators and those of the moment estimators up to order n^{-1}, where we have given few numerical examples to show that higher order terms in the asymptotic biases of moment estimators are of negligible importance for moderately large samples.

The information obtained from the moment estimators, which is represented by its joint asymptotic efficiency, relative to the M.L. estimators, seems to be reliable for smaller values of λ_2 (i.e.

when the GPD is very close to the Poisson distribution) and for large values of λ_1 (i.e. when we are close to normality). It was also realized that regions of high biases of the M.L. estimators are those of high efficiency for the moment estimators.

CHAPTER III

ESTIMATION OF PARAMETERS FOR THE GENERALIZED NEGATIVE BINOMIAL DISTRIBUTION

3.1 Introduction.

The generalized negative binomial distribution (GNBD) is defined by the probability function

$$(3.1)\qquad P(X=x) = \begin{cases} \dfrac{n\Gamma(n+\beta x)}{x!\,\Gamma(n+\beta x-x+1)}\,\theta^{x}(1-\theta)^{n+\beta x-x}, & x = 0,1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$$

where $0 \le \theta \le 1$, $n > 0$ and $\beta = 0$ or $1 \le \beta < \theta^{-1}$.

The distribution was first introduced by Jain and Consul (34), with the conditions $|\theta\beta| < 1$ and $P(X=x) = 0$ for all $x \ge m$ if $n+m\beta < 0$. Consul and Shenton (13), gave the probability generating functions for the GNBD as a particular case of the Lagrangian class of discrete probability distributions, deleted the conditions given by Jain and Consul (34), and took $\beta \ge 0$. Subsequently, Nelson (52) pointed out a negative value of β does not give a true probability distribution as the sum of the probabilities over the domain of x is less than unity. Consul and Gupta (15), have recently shown that the value of β must either be zero or $1 \leq \beta \leq \theta^{-1}$ for the existence of a true probability distribution.

The distribution has its applications in the theory of queues. Mohanty ((50), (51)) has shown that, in a queue initiated with n customers, where the arrivals are assumed to be Poissonian, and each customer is served exponentially, according to first come, first served, λ is the traffic intensity and the customers arrive in batches of r, then the probability that exactly $rx+n$ customers will be served before the queue first vanishes is

$$P(X=x) = \frac{n}{n+(r+1)x} \binom{n+(r+1)x}{x} \zeta^{x}(1-\zeta)^{n+rx}$$

where

$$\zeta = \frac{\lambda}{1+\lambda}$$

which is a GNBD with $\beta = r+1$ and $\theta = \zeta$.

Rewriting the probability distribution (3.1) in the form

$$(3.2) \qquad P(X=x) = \frac{n\Gamma(n+\beta x)}{x!\Gamma(n+\beta x-x+1)} \frac{(\theta(1-\theta)^{\beta-1})^{x}}{(1-\theta)^{-n}} ,$$

it can be treated as a particular family of the MPSD class defined in (1.10). Gupta (21), has shown that the moments of an MPSD satisfy the recurrence relation

$$(3.3.a) \qquad \mu_{r+1} = \frac{g(\theta)}{g'(\theta)} \left[\frac{d\mu_{r}}{d\theta} + r\,\frac{d\mu_1'}{d\theta}\,\mu_{r-1}\right] , \qquad r = 1,2,\ldots$$

$$(3.3.b) \qquad \mu_1' = \frac{g(\theta)}{g'(\theta)} \frac{\partial \ln f}{\partial\theta} .$$

where $g(\theta) = \theta(1-\theta)^{\beta-1}$, and $f = (1-\theta)^{-n}$.

Thus, for the GNBD, the first six central moments are

(3.3.c) $$\text{Mean} = \mu_1' = n\theta/(1-\beta\theta)$$

(3.3.d) $$\text{Variance} = \mu_2 = n\theta(1-\theta)/(1-\beta\theta)^3$$

(3.3.e) $$\mu_3 = n\theta(1-\theta)(1-\beta\theta)^{-5}[1-2\theta+\beta\theta(2-\theta)]$$

(3.3.f) $$\begin{aligned}\mu_4 &= n\theta(1-\theta)(1-\beta\theta)^{-7}[3n\theta(1-\theta)(1-\beta\theta)+1-6\theta+6\theta^2\\ &\quad +2\beta\theta(4-9\theta+\theta^2)+\beta^2\theta^2(6-6\theta+\theta^2)]\end{aligned}$$

(3.3.g) $$\begin{aligned}\mu_5 &= n\theta(1-\theta)(1-\beta\theta)^{-9}[10n\theta(1-\theta)(1-\beta\theta)\{1-2\theta+\beta\theta(2-\theta)\}\\ &\quad + 1-14\theta+36\theta^2-24\theta^3+\beta\theta(22-93\theta+100\theta^2-19\theta^3-8\theta^4)\\ &\quad + \beta^2\theta^2(18-48\theta+35\theta^2-6\theta^3)+\beta^3\theta^3(24-36\theta+14\theta^2-\theta^3)]\end{aligned}$$

(3.3.h) $$\begin{aligned}\mu_6 &= n\theta(1-\theta)(1-\beta\theta)^{-11}[15n^2\theta^2(1-\theta)^2(1-\beta\theta)^2\\ &\quad + 5n\theta(1-\theta)(1-\beta\theta)\{5-26\theta+26\theta^2+2\beta\theta(16-37\theta+16\theta^2)\\ &\quad + \beta^2\theta^2(26-26\theta+5\theta^2)\}+1-30\theta(1-\theta)(1-2\theta)^2\\ &\quad + \beta\theta(52-510\theta+1360\theta^2-1350\theta^3+444\theta^4)+\beta^2\theta^2(328-1650\theta\\ &\quad + 2650\theta^2-1650\theta^3+328\theta^4)+\beta^3\theta^3(444-1350\theta+1260\theta^2\\ &\quad - 510\theta^3+52\theta^4)+\beta^4\theta^4(129-240\theta+150\theta^2-30\theta^3+\theta^4)] \ .\end{aligned}$$

Assuming the parameters n and β to be known constants, the problem of estimation of the parameter θ of the GNBD and the estimators of functions of θ, has been studied by many authors. Charalambides (8), considered the problem of finding the MVU estimator for the parametric function θ^m. He derived the distribution of the sufficient statistic for θ in the case of left truncated sample with known and unknown truncation points. Gupta (22) obtained the M.L. estimator of the parameter θ of the MPSD, and derived the first order approximation for its bias and variance. Kumar and Consul (44) derived a recurrence formula for the higher order negative moments of an MPSD, and used the negative moments of the GNBD to establish bounds for the bias of the M.L. estimator of the parameter θ.

Our main interest is the simultaneous estimation of the parameters n and θ, when β is known, of the probability distribution defined in (3.1), by using the methods of moments and maximum likelihood and to study some of the asymptotic properties of these estimators. We shall exclude the boundary points θ=0, θ=1, as the distribution becomes degenerate at those points.

3.2 Maximum Likelihood Estimation for the GNBD.

Let $X_1, X_2, \ldots, X_N$ be a random sample of size N, taken from the GNBD (3.1). The likelihood function is

$$(3.4) \qquad L = n^N \prod_{i=1}^{N} \left[\frac{(n+\beta x_i-1) \ \ldots \ (n+\beta x_i-x_i+1)}{x_i!}\right] \theta^{N\bar{x}}(1-\theta)^{nN+(\beta-1)N\bar{x}} .$$

Taking the logarithm of both sides of (3.4), we get

(3.5)
$$\log L = N \log n + N\bar{x} \log \theta + [nN+(\beta-1)N\bar{x}] \log (1-\theta)$$
$$+ \sum_{i=1}^{N} \sum_{j=1}^{x_i-1} \log (n+\beta x_i-j) - \sum_{i=1}^{N} \log x_i! - \sum_{i=1}^{N} \log (n+\beta x_i)$$

where $\sum_{j=1}^{x_i-1} \log (n+\beta x_i) = 0$ for $x_i = 0$, $i = 1,2,\ldots,N$. On differentiating $\ln L$ partially w.r.t. θ and n, and equating to zero, we get

(3.6.a)
$$\frac{\partial \log L}{\partial \theta} = \frac{N\bar{x}}{\theta} - \frac{Nn + (\beta-1)N\bar{x}}{1-\theta} = 0$$

(3.6.b)
$$\frac{\partial \log L}{\partial n} = \frac{N}{n} + N \log (1-\theta) + \sum_{i=1}^{N} \sum_{j=1}^{x_i-1} \frac{1}{n + \beta x_i - j} = 0 \quad .$$

The last two equations can be simplified to give

(3.7)
$$\hat{\theta} = \bar{X}(\hat{n}+\beta\bar{X})^{-1}$$

(3.8)
$$\hat{n} = -\left[\log (1-\hat{\theta}) + N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{x_i-1} \frac{1}{\hat{n} + \beta x_i - j}\right]^{-1} ,$$

where $\hat{\theta}$ and $\hat{n}$ are the M.L. estimators of θ and n respectively, provided that the matrix of the second order partial derivatives of the log likelihood function is negative definite. Since the above two equations are rather complicated in structure, it is difficult to give explicit algebraic expressions for $\hat{\theta}$ and $\hat{n}$.

Equation (3.8) can be written in the form

$$\hat{\theta} = 1 - \exp\left[-\left(\frac{1}{\hat{n}} + \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{x_i-1}\frac{1}{\hat{n}+\beta x_i-j}\right)\right]$$

and by using (3.7), it gives

$$(3.9) \qquad \overline{X}(\hat{n}+\beta\overline{X})^{-1} = 1 - \exp\left[-\left(\frac{1}{\hat{n}} + \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{x_i-1}\frac{1}{\hat{n}+\beta x_i-j}\right)\right] \quad .$$

The last equation can be rewritten as

$$\hat{n} = \overline{X}(1-\beta) + (\hat{n}+\beta\overline{X})\ \exp\left[-\left(\frac{1}{\hat{n}} + \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{x_i-1}\frac{1}{\hat{n}+\beta x_i-j}\right)\right] \quad ,$$

which is of the same form as $n = f(n)$, considered by Henrici (32), and so the Newton-Raphson iterative technique may be tried to solve equation (3.9).

3.2.1 Asymptotic Biases of the M.L. Estimators.

Since the regularity conditions given by Shenton and Bowman which are stated in section (1.4) are satisfied for the GNBD, the first order terms in the biases, variances and covariance of the M.L. estimators can be obtained by using the formulae given in (2.34.a) and (2.26.a) respectively.

We shall now give a computational procedure to obtain the first order terms in the biases of the M.L. estimators. The method of obtaining these for simultaneous estimation of several parameters was given by Shenton and Wallington (70). Using their notation, we denote the biases of $\hat{\theta}$ and $\hat{n}$ be $b_1(\hat{\theta})$ and $b_2(\hat{n})$ respectively. These biases can be obtained by solving the equations

$$(3.10)\qquad \begin{aligned} b_1(\hat{\theta})P_{11} + b_2(\hat{n})P_{12} &= -A_1/2\Delta N \\ b_1(\hat{\theta})P_{21} + b_2(\hat{n})P_{22} &= -A_2/2\Delta N \end{aligned}$$

where

$$\Delta = P_{11}P_{22} - (P_{12})^2 \quad ,$$

$$A_1 = P_{22}\, P_{1,11} - 2P_{12}\, P_{1,12} + P_{11}\, P_{1,22} \ ,$$

$$A_2 = P_{22}\, P_{2,11} - 2P_{12}\, P_{2,12} + P_{11}\, P_{2,22}$$

and

$$P_{1,11} = E\left(P^{-2}\frac{\partial P}{\partial\theta}\frac{\partial^2 P}{\partial\theta^2}\right), \qquad P_{1,12} = E\left(P^{-2}\frac{\partial P}{\partial\theta}\frac{\partial^2 P}{\partial\theta\,\partial n}\right),$$

$$(3.11)\qquad P_{1,22} = E\left(P^{-2}\frac{\partial P}{\partial\theta}\frac{\partial^2 P}{\partial n^2}\right), \qquad P_{2,11} = E\left(P^{-2}\frac{\partial P}{\partial n}\frac{\partial^2 P}{\partial\theta^2}\right),$$

$$P_{2,12} = E\left(P^{-2}\frac{\partial P}{\partial n}\frac{\partial^2 P}{\partial\theta\,\partial n}\right), \qquad P_{2,22} = E\left(P^{-2}\frac{\partial P}{\partial n}\frac{\partial^2 P}{\partial n^2}\right)$$

From Kendall and Stuart (39), we know that

(3.12)
$$P_{11} = E\left(P^{-2}\left(\frac{\partial P}{\partial \theta}\right)^2\right), \qquad P_{12} = P_{21} = E\left(P^{-2}\left(\frac{\partial P}{\partial \theta}\,\frac{\partial P}{\partial n}\right)\right),$$
$$P_{22} = E\left(P^{-2}\left(\frac{\partial P}{\partial n}\right)^2\right) \quad .$$

We shall now express the values given in (3.11) and (3.12) in terms of the parameters n, β, θ.

Since

$$NP_{11} = E\left[\frac{(\partial \log L)}{\partial \theta}\right]^2 = -E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right] \quad ,$$

by differentiating equation (3.6.a) w.r.t. θ, we get

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{N\overline{X}}{\theta^2} - \frac{Nn + (\beta-1)N\overline{X}}{(1-\theta)^2} \quad .$$

Thus

$$E\left(\frac{\partial^2 \log L}{\partial \theta^2}\right) = \frac{-nN}{\theta(1-\beta\theta)} - \frac{nN(1-\beta\theta) + nN\theta(\beta-1)}{(1-\theta)^2(1-\beta\theta)}$$

$$= \frac{-Nn}{\theta(1-\beta\theta)} - \frac{Nn}{(1-\theta)(1-\beta\theta)}$$

and

(3.13)
$$P_{11} = \frac{n}{\theta(1-\theta)(1-\beta\theta)} \quad .$$

Also, differentiating equation (3.6.a) w.r.t. n, we have

$$\frac{\partial^2 \log L}{\partial \theta \partial n} = -N(1-\theta)^{-1}$$

which gives the value

$$P_{12} = (1-\theta)^{-1} \quad . \tag{3.14}$$

Similarly, by differentiating equation (3.6.b) w.r.t. n, and taking expectations, we get

$$P_{22} = -\frac{1}{N} E\left(\frac{\partial^2 \log L}{\partial n^2}\right) = \frac{1-P_0}{n^2} + \sum_{x=2}^{\infty} \left[\left(\sum_{j=1}^{x-1} \frac{1}{(n+x-j)^2}\right) P\right] \quad . \tag{3.15}$$

Since the probabilities P on the tail (for large x) become very small, the function P_{22} can be robustly computed by ignoring the tail-end probabilities P less than a pre-assigned small value say 10^{-10}, for given values of n, β and θ. However, we shall use the orthogonal polynomial approximation to the function $\frac{\partial P}{\partial n}$, for computing P_{22}.

It can be easily seen that the probability distribution given in (3.1) satisfies the recurrence relation

$$\frac{\partial P}{\partial \theta} = (x-\mu_1') \frac{(1-\beta\theta)}{\theta(1-\theta)} P \quad . \tag{3.16}$$

By differentiating (3.16) w.r.t. θ,

$$\frac{\partial^2 P}{\partial \theta^2} = (x-\mu_1')^2 \frac{(1-\beta\theta)^2}{\theta^2(1-\theta)^2} P + \left[\frac{2\theta-\beta\theta^2-1}{\theta^2(1-\theta)^2} (x-\mu_1') - \frac{n}{\theta(1-\theta)(1-\beta\theta)}\right] P \tag{3.17}$$

and, the product

$$\frac{\partial P}{\partial \theta}\frac{\partial^2 P}{\partial \theta^2} = (x-\mu_1')^3 \frac{(1-\beta\theta)^3}{\theta^3(1-\theta)^3} P^2 + \left[\frac{2\theta-\beta\theta^2-1}{\theta^2(1-\theta)^2}\frac{(1-\beta\theta)}{\theta(1-\theta)} (x-\mu_1')^2 \right.$$

$$\left. - (x-\mu_1') \frac{(1-\beta\theta)}{\theta(1-\theta)} \frac{n}{\theta(1-\theta)(1-\beta\theta)}\right] P^2 \quad .$$

Now, by taking the expectation of $P^{-2} \frac{\partial P}{\partial \theta}\frac{\partial^2 P}{\partial \theta^2}$, (3.18.a) $$P_{1,11} = E\left(P^{-2}\,\frac{\partial P}{\partial\theta}\,\frac{\partial^2 P}{\partial\theta^2}\right) = \mu_3\,\frac{(1-\beta\theta)^3}{\theta^3(1-\theta)^3} + \mu_2\,\frac{(1-\beta\theta)(2\theta-\beta\theta^2-1)}{\theta^3(1-\theta)^3} .$$

Thus

(3.18.b) $$P_{1,11} = \frac{2n\beta}{\theta(1-\theta)(1-\beta\theta)^2} .$$

On differentiating (3.16) w.r.t. n, we have

$$\frac{\partial^2 P}{\partial n\partial\theta} = \frac{(1-\beta\theta)}{\theta(1-\theta)}\left[-\frac{\partial\mu_1'}{\partial n}\cdot P + (x-\mu_1')\,\frac{\partial P}{\partial n}\right]$$

$$= \frac{-P}{1-\theta} + (x-\mu_1')\,\frac{(1-\beta\theta)}{\theta(1-\theta)}\,\frac{\partial P}{\partial n} ,$$

and thus

$$\frac{\partial P}{\partial\theta}\,\frac{\partial^2 P}{\partial\theta\partial n} = (x-\mu_1')^2\,\frac{(1-\beta\theta)^2}{\theta^2(1-\theta)^2}\,P\,\frac{\partial P}{\partial n} - (x-\mu_1')\,\frac{(1-\beta\theta)}{\theta(1-\theta)^2}\,P^2 .$$

Therefore,

(3.19) $$E\left(P^{-2}\,\frac{\partial P}{\partial\theta}\,\frac{\partial^2 P}{\partial\theta\partial n}\right) = \frac{(1-\beta\theta)^2}{\theta^2(1-\theta)^2}\sum_{x=0}^{\infty}(x-\mu_1')^2\,\frac{\partial P}{\partial n} .$$

By using the relation

(3.20) $$\sum_{x=0}^{\infty}(x-\mu_1')^l\,\frac{\partial P}{\partial n} = \frac{\partial\mu_l}{\partial n} + l\mu_{l-1}\,\frac{\partial\mu_1'}{\partial n} ,$$

and for $l = 2$, (3.19) becomes

$$P_{1,12} = \frac{1}{\theta(1-\theta)(1-\beta\theta)}$$

Also, by equation (3.17)

$$E\left(P^{-2}\,\frac{\partial P}{\partial n}\,\frac{\partial^2 P}{\partial \theta^2}\right) = \frac{(1-\beta\theta)^2}{\theta^2(1-\theta)^2}\sum_{x=0}^{\infty}(x-\mu_1')^2\,\frac{\partial P}{\partial n} + \frac{(2\theta-\beta\theta^2-1)}{\theta^2(1-\theta)^2}\sum_{x=0}^{\infty}(x-\mu_1')\,\frac{\partial P}{\partial n}$$

(3.21) $$- \frac{n}{\theta(1-\theta)(1-\beta\theta)}\sum_{x=0}^{\infty}\frac{\partial P}{\partial n} \quad ,$$

and on using (3.20), we have

(3.22) $$P_{2,11} = \frac{1}{\theta(1-\theta)(1-\beta\theta)} + \frac{2\theta-\beta\theta^2-1}{\theta(1-\theta)^2(1-\beta\theta)} = (1-\theta)^{-2} \quad .$$

Since (3.16) gives

$$\frac{\partial P}{\partial \theta}\,\frac{\partial^2 P}{\partial n^2} = (x-\mu_1')\,\frac{(1-\beta\theta)}{\theta(1-\theta)}\,P\,\frac{\partial^2 P}{\partial n^2} \quad ,$$

(3.23.a) $$E\left(P^{-2}\,\frac{\partial P}{\partial \theta}\,\frac{\partial^2 P}{\partial n^2}\right) = \frac{(1-\beta\theta)}{\theta(1-\theta)}\sum_{x=0}^{\infty}(x-\mu_1')\,\frac{\partial^2 P}{\partial n^2} \quad .$$

However

$$\frac{\partial^2}{\partial n^2}\,[(x-\mu_1')P] = \frac{\partial}{\partial n}\left[-\frac{\partial \mu_1'}{\partial n}\cdot P + (x-\mu_1')\,\frac{\partial P}{\partial n}\right]$$

$$= \frac{\partial^2 \mu_1'}{\partial n^2}\,P - \frac{\partial \mu_1'}{\partial n}\,\frac{\partial P}{\partial n} - \frac{\partial \mu_1'}{\partial n}\,\frac{\partial P}{\partial n} + (x-\mu_1')\,\frac{\partial^2 P}{\partial n^2} \quad ,$$

which gives

(3.23.b) $$(x-\mu_1')\,\frac{\partial^2 P}{\partial n^2} = \frac{\partial^2}{\partial n^2}\,[(x-\mu_1')P] + \frac{\partial^2 \mu_1'}{\partial n^2}\cdot P + 2\,\frac{\partial \mu_1'}{\partial n}\,\frac{\partial P}{\partial n} \quad .$$

Since the function P given by (3.1) is such that the range of X is independent of the parameters, the operations of summation and differentiation can be interchanged and $\partial^2\mu_1'/\partial n^2 = 0$, accordingly, on taking the summation over x in (3.23.b), we have

(3.23.c) $$\sum_{x=0}^{\infty} (x-\mu_1') \frac{\partial^2 P}{\partial n^2} = \frac{\partial^2}{\partial n^2} \sum_{x=0}^{\infty} (x-\mu_1')P = 0 \quad ..$$

Thus

(3.23.d) $$P_{1,22} = 0 \quad .$$

The determination of the exact values of $P_{2,12}$ and $P_{2,22}$ seems to be intractable. So we shall have to use the asymptotic expansion for P^{-1} $\partial P/\partial n$ in the form of orthogonal polynomials. As stated in Cramér (16), if $P(x)$ is a distribution function with finite moments μ_γ of all orders, then the point x_0 is a point of increase for $P(x)$, if $P(x_0+h) > P(x_0-h)$ for every $h > 0$. If the function P has at least r points of increase, Cramér has proved that there exists a sequence of polynomials $G_0(x), G_1(x), \ldots$ uniquely determined by the following conditions

a) $G_n(x)$ is of degree n, and the coefficient of x^n in $G_n(x)$ is positive;

b) $G_n(x)$ satisfy the orthogonality conditions

$$\sum_{x=0}^{\infty} G_r(x)G_s(x)P(x) = E(G_r^2(x)) \qquad \text{if } r = s$$
$$= 0 \qquad r \neq s \; (r,s = 0,1,2,\ldots).$$

Szegö (78) stated the formal Fourier expansion of a continuous function $h(x)$ in terms of the set of orthogonal polynomials as

(3.24) $$h(x) = a_0G_0(x) + a_1G_1(x) + a_2G_2(x) + \ldots \quad .$$

In studying the efficiency of the moment estimators of the

Neyman-type A distribution, Shenton (69) has shown that, if we write

$$\frac{1}{P_x}\frac{\partial P_x}{\partial n} = a_0G_0(x) + a_1G_1(x) + a_2G_2(x) + \ldots$$

where $a_0, a_1, a_2, \ldots$ are chosen so that

$$(3.25) \qquad \sum_{x=0}^{\infty} [\frac{1}{P}\frac{\partial P}{\partial n} - (a_0G_0(x) + a_1G_1(x) + \ldots)]^2 P$$

is minimum, then

$$a_0 = 0, \qquad a_1 = \frac{\partial\mu_1'/\partial n}{E(G_1^2(x))}, \qquad a_2 = \frac{1}{E(G_2^2(x))}\left(\frac{\partial\mu_2}{\partial n} - \frac{\mu_3}{\mu_2}\frac{\partial\mu_1'}{\partial n}\right)$$

where

$$G_0(x) = 1, \qquad G_1(x) = x-\mu_1', \qquad G_2(x) = (x-\mu_1')^2 - \frac{\mu_3}{\mu_2}(x-\mu_1') - \mu_2 \quad ,$$

are the orthogonal polynomials associated with the probability distribution P_x.

Now, if we assume that, for the GNBD

$$(3.26) \qquad \frac{1}{P}\frac{\partial P}{\partial n} = a_0G_0(x) + a_1G_1(x) + a_2G_2(x) + \ldots$$

$$(3.27) \qquad \frac{1}{P}\left(\frac{\partial P}{\partial n}\right)^2 = \sum_{r=0}^{\infty} [a_r^2G_r^2(x) + \sum_{r\neq s} a_ra_sG_r(x)G_s(x)]P \quad .$$

Since

$$\frac{\partial^2 P}{\partial\theta\partial n} = (x-\mu_1')\frac{(1-\beta\theta)}{\theta(1-\theta)}\frac{\partial P}{\partial n} - \frac{1}{1-\theta}P \quad ,$$

then

$$\frac{\partial P}{\partial n}\frac{\partial^2 P}{\partial\theta\partial n} = (x-\mu_1')\frac{(1-\beta\theta)}{\theta(1-\theta)}\left(\frac{\partial P}{\partial n}\right)^2 - \frac{1}{1-\theta}P\frac{\partial P}{\partial n} \quad ,$$

and

$$(3.28.a) \qquad E\left(P^{-2}\,\frac{\partial P}{\partial n}\,\frac{\partial^2 P}{\partial\theta\partial n}\right) = \frac{(1-\beta\theta)}{\theta(1-\theta)}\sum_{x=0}^{\infty}\,(x-\mu_1')\,\frac{1}{P}\left(\frac{\partial P}{\partial n}\right)^2 \quad .$$

Substituting (3.27) in (3.28.a) we get, on using the first three terms in the expansion (3.26),

$$P_{2,12} = \frac{(1-\beta\theta)}{\theta(1-\theta)}\sum_{x=0}^{\infty}(x-\mu_1')(a_1^2G_1^2(x) + a_2^2G_2^2(x) + 2a_1a_2G_1(x)G_2(x))P$$

$$(3.28.b) \qquad = \frac{(1-\beta\theta)}{\theta(1-\theta)}\left[a_1^2\sum_{x=0}^{\infty}(x-\mu_1')G_1^2(x)P + a_2^2\sum_{x=0}^{\infty}(x-\mu_1')G_2^2(x)P\right.$$

$$\left. + 2a_1a_2\sum_{x=0}^{\infty}(x-\mu_1')G_1(x)G_2(x)P\right]$$

$$P_{2,12} = \frac{(1-\beta\theta)}{\theta(1-\theta)}\left[a_1^2\mu_3 + a_2^2\left(\mu_5 + \frac{\mu_3^3}{\mu_2^2} - \frac{2\mu_3\mu_4}{\mu_2}\right) + 2a_1a_2\left(\mu_4 - \frac{\mu_3^2}{\mu_2} - \mu_2^2\right)\right]$$

since

$$E(G_1^2(x)) = E((x-\mu_1')^2) = \mu_2 = \frac{n\theta(1-\theta)}{(1-\beta\theta)^3} \quad ,$$

and

$$E(G_2^2(x)) = \left(\mu_4 - \frac{\mu_3^2}{\mu_2} - \mu_2^2\right) \quad ,$$

then

$$a_1 = \frac{\partial\mu_1'/\partial n}{E(G_1^2(x))} = \frac{(1-\beta\theta)^2}{n(1-\theta)} \qquad a_2 = \frac{(\partial\mu_2/\partial n - \mu_3/\mu_2\ \partial\mu_1'/\partial n)}{(\mu_4 - \mu_3^2/\mu_2 - \mu_2^2)} \quad .$$

Thus, substituting the values of α_1 and α_2 in (3.28.b), we can write

$$P_{2,12} = D_1 + D_2 + D_3 \quad ,$$

where

$$D_1 = [1-2\theta+\beta\theta(2-\theta)]/n(1-\theta)^2$$

$$(3.29) \qquad D_2 = \left[\frac{\mu_2 \partial\mu_2/\partial n - \mu_3 \partial\mu_1'/\partial n}{\mu_2\mu_4 - \mu_3^2 - \mu_2^3}\right]^2 [\mu_5 + \mu_3^3/\mu_2^2 - 2\mu_3\mu_4/\mu_2]$$

$$D_3 = \left[\frac{2(1-\beta\theta)^5}{n^2\theta(1-\theta)^2}\right]\left[\mu_2 \frac{\partial\mu_2}{\partial n} - \mu_3 \frac{\partial\mu_1'}{\partial n}\right] \quad .$$

To find $P_{2,22}$, we assume that

$$(3.30) \qquad \frac{1}{P}\frac{\partial^2 P}{\partial n^2} = \{b_0 G_0(x) + b_1 G_1(x) + b_2 G_2(x) + \ldots\} \quad ,$$

where $b_0, b_1, \ldots$ are so determined that

$$\sum_{x=0}^{\infty} \left[\frac{1}{P}\frac{\partial^2 P}{\partial n^2} - (b_0 G_0(x) + b_1 G_1(x) + \ldots)\right]^2 P$$

is minimum. Multiplying both sides of (3.30) by $G_k(x)P$, and summing over all values of x, we get, on using the orthogonality condition $\sum_{x=0}^{\infty} \{G_r(x) G_s(x)\} P = 0$, $r \neq s$, that

$$b_k = \frac{1}{E(G_k^2(x))} \sum_{x=0}^{\infty} G_k(x) \frac{\partial^2 P}{\partial n^2} \quad .$$

For $k = 0,1,2$, we get

$$b_0 = 0,$$

$$b_1 = \frac{1}{E(G_1^2(x))} \sum_{x=0}^{\infty} (x-\mu_1') \frac{\partial^2 P}{\partial n^2} = 0$$

(from 3.23.c), and

$$b_2 = \frac{1}{E(G_2^2(x))} \left[\sum_{x=0}^{\infty} \{(x-\mu_1')^2 - \frac{\mu_3}{\mu_2} (x-\mu_1') - \mu_2\} \frac{\partial^2 P}{\partial n^2} \right]$$

(3.31)

$$= \frac{1}{E(G_2^2(x))} \sum_{x=0}^{\infty} (x-\mu_1')^2 \frac{\partial^2 P}{\partial n^2} \quad .$$

Since

(3.32.a)
$$\frac{\partial^2}{\partial n^2} [(x-\mu_1')^2 P] = 2\left(\frac{\partial \mu_1'}{\partial n}\right)^2 P + 2 \frac{\partial^2 \mu_1'}{\partial n^2} (x-\mu_1') P$$
$$+ 2 \frac{\partial \mu_1'}{\partial n} \frac{\partial P}{\partial n} - 2 \frac{\partial \mu_1'}{\partial n} (x-\mu_1') \frac{\partial P}{\partial n} + (x-\mu_1')^2 \frac{\partial^2 P}{\partial n^2} \quad .$$

Summing both sides of (3.32.a) over all values of $x = 0,1,2...$, we get

$$\frac{\partial^2}{\partial n^2} \sum_{x=0}^{\infty} (x-\mu_1')^2 P = \frac{2\theta^2}{(1-\beta\theta)^2} - \frac{2\theta^2}{(1-\beta\theta)^2} + \sum_{x=0}^{\infty} (x-\mu_1')^2 \frac{\partial^2 P}{\partial n^2}$$

or

(3.32.b)
$$\sum_{x=0}^{\infty} (x-\mu_1')^2 \frac{\partial^2 P}{\partial n^2} = 0 \quad .$$

On substituting (3.32.b) in (3.31), we get

$$b_2 = 0 \quad ,$$

and hence

(3.32.c)
$$P_{2,22} \approx 0 \quad .$$

Using (3.27) and the values of a_0, a_1, a_2, $E(G_1^2(x))$, and $E(G_2^2(x))$, it can be similarly shown

$$P_{22} \approx \frac{\theta(1-\beta\theta)}{n(1-\theta)} + \frac{(\partial\mu_2/\partial n - \mu_3/\mu_2 \;\partial\mu_1'/\partial n)^2}{(\mu_4 - \mu_3^2/\mu_2 - \mu_2^2)}$$

Substituting the values of $P_{i,j}$ and $P_{\alpha,ij}$ $(\alpha,i,j = 1,2)$ in equations (3.10), and solving for $b_1(\hat{\theta})$ and $b_2(\hat{n})$, we get

$$Nb_1(\hat{\theta}) \approx \frac{\theta^2(1-\theta)^3(1-\beta\theta)^2}{2[n(1-\theta)P_{22}-\theta(1-\beta\theta)]^2}\left[\frac{P_{22}}{(1-\theta)^2} - \frac{2P_{2,12}}{(1-\theta)}\right.$$

$$\left. + \frac{2P_{22}}{\theta(1-\theta)(1-\beta\theta)} - \frac{2n\beta P_{22}^2}{\theta(1-\beta\theta)^2}\right]$$

and

$$Nb_2(\hat{n}) \approx \frac{\theta^2(1-\theta)^2(1-\beta\theta)}{2[n(1-\theta)P_{22} - (1-\beta\theta)]^2}\left[\frac{2n\beta P_{22}}{\theta(1-\beta\theta)}\right.$$

$$\left. - \frac{2}{\theta(1-\theta)} + \frac{2nP_{22}}{\theta} - \frac{nP_{22}}{(1-\theta)}\right] .$$

Tabulation for the biases of the M.L. estimators will be given in the last section of this chapter, for some selected values of the parameters n, β, and θ, where we compare the performance of the biases of the M.L. estimator with the corresponding biases of the moment estimators.

However, it should be pointed out that the computed values for the biases may not be very reliable on account of the approximations used for the different equations in the above work. The results will be more reliable if the approximations are taken up to the third and fourth degree of approximation, but that task is so gigantic that it is outside the scope of the present work. Shenton and Bowman (75) have studied such terms for the particular case of $\beta = 1$, and have shown that this task is a very difficult one.

3.2.2. Asymptotic Covariances of the M.L. Estimators.

In this section we shall derive the first order terms in the covariances of the M.L. estimators for the parameters θ and n. The derivation of second order terms is rather laborious because their number is very large, and the expressions for the covariances are expected to be very complicated. Bowman and Shenton (6) have given the asymptotic expansions of the covariances of the M.L. estimators as follows:

$$\mathrm{Var}(\hat{\theta}) = t_1/N + T_1/N^2 + \ldots$$

(3.35) $$\mathrm{Cov}(\hat{\theta},\hat{n}) = t_{12}/N + T_{12}/N^2 + \ldots$$

$$\mathrm{Var}(\hat{n}) = t_2/N + T_2/N^2 + \ldots$$

where

$$t_2 = P_{11}/\Delta, \qquad t_{12} = -P_{12}/\Delta, \qquad t_1 = P_{22}/\Delta \quad ,$$

T_1, T_{12}, T_2 are the coefficients of N^{-2} in the expansion of Var $(\hat{\theta})$, Cov $(\hat{\theta},\hat{n})$, and Var $(\hat{n})$ respectively, and

$$\Delta = P_{11}P_{22} - P_{12}^2$$

$$\approx \frac{n}{\theta(1-\theta)(1-\beta\theta)} \frac{(\partial\mu_2/\partial n - \mu_3/\mu_2 \;\partial\mu_1'/\partial n)^2}{(\mu_4 - \mu_3^2/\mu_2 - \mu_2^2)}$$

which can be shown to be strictly positive for all n, θ, and β in the admissible parameter space. Thus, the asymptotic variances and covariance to order N^{-1}, become

$$N\cdot\mathrm{Var}(\hat{\theta}) \approx \frac{\theta^2(1-\beta\theta)^2}{n^2} \frac{(\mu_4 - \mu_3^2/\mu_2 - \mu_2^2)}{(\partial\mu_2/\partial n - \mu_3/\mu_2 \;\partial\mu_1'/\partial n)} + \frac{\theta(1-\theta)(1-\beta\theta)}{n}$$

(3.36) $$N\cdot\mathrm{Var}(n) \approx \frac{(\mu_4 - \mu_3^2/\mu_2 - \mu_2^2)}{(\partial\mu_2/\partial n - \mu_3/\mu_2 \;\partial\mu_1'/\partial n)^2}$$

$$N\cdot\mathrm{Cov}(\hat{n},\hat{\theta}) \approx -\frac{\theta(1-\beta\theta)(\partial\mu_2/\partial n - \mu_3/\mu_2 \;\partial\mu_1'/\partial n)^2}{n(\mu_4 - \mu_2^2/\mu_2 - \mu_2^2)} \quad .$$

3.3 The Moment Estimation Problem.

The moment estimators for the parameters θ, n and β of the GNBD, were initially given by Jain and Consul (34), by equating the mean, the variance and the third moment of the sample, about the mean, with the corresponding population values, but no study of the biases, variances and covariances of those estimators was made by them. The problem of evaluating the variance-covariance matrix is highly involved, specially when the moment estimators are complex in structure, and the derivation of exact expressions appears to be an impossible task.

Since the domain of θ is restricted by the values of β, we assume β to be known. We shall use the bivariate Taylor expansion of a function $f(m_1', m_2)$ of the sample moments as in Chapter II. Expressions for the biases and variances will be given in terms of the notations for the values of the partial derivatives of the moment estimator function.

Based upon a random sample of size N, we define the moment estimator of θ and n, assuming β to be known, as the estimators which satisfy the equations obtained by equating the sample mean and variance with the corresponding population mean and variance. The two equations are therefore

$$n\theta/(1-\beta\theta) = m_1'$$

and

$$n\theta(1-\theta)/(1-\beta\theta)^3 = m_2 \quad ,$$

where m_1' is the sample mean, and m_2 is the sample variance.

Since Binet (4) has studied the moment estimation problem for the binomial distribution ($\beta = 0$), and Katti and Gurland (37), and Anscombe (1) have studied the moment estimation problem for the negative binomial distribution ($\beta = 1$), we shall consider the domain $1 < \beta < \theta^{-1}$ only for the GNBD.

By eliminating n between the last two equations

$$\text{(3.37.a)} \qquad \frac{m_2}{m_1'} = \frac{1-\theta}{1-\beta\theta} \cdot \frac{1}{1-\beta\theta} \quad .$$

Since $\theta < \beta\theta < 1$, the expression on the right of the above equation is always greater than unity. Thus, if $m_1' \geq m_2$, the above equation will give an absurd estimator for θ. Accordingly, the moment estimators for n and θ will be inadmissible when $m_1' \geq m_2$.

Solving (3.37.a) for θ, we have

$$\text{(3.37.b)} \quad \theta = \frac{(2\beta - m_1'/m_2) \pm \{(2\beta - m_1'/m_2)^2 - 4\beta^2(1 - m_1'/m_2)\}^{\frac{1}{2}}}{2\beta^2} \quad .$$

When $m_1' < m_2$, it can be easily shown that the value θ, obtained by taking the positive sign in (3.37.b), does not satisfy the equation

$$n = m_1'\left(\frac{1}{\theta} - \beta\right)$$

for n, as it gives a negative value for n. Thus the positive sign in (3.37.b) becomes inadmissible. Hence the moment estimators for θ and n become

$$\text{(3.38)} \qquad \theta^* = \frac{(2\beta - m_1'/m_2) - \{(m_1'/m_2)^2 + 4\beta(\beta-1)m_1'/m_2\}^{\frac{1}{2}}}{2\beta^2}$$

(3.39) $$n^* = m_1'\left(\frac{1}{\theta^*} - \beta\right)$$

respectively.

However, one should remember that there is a positive probability for m_1' to be greater than m_2 even though $1 < \beta < \theta^{-1}$ and in such cases the moment estimators will be bad.

3.3.1 Asymptotic Biases of Moment Estimators.

In order to expand θ^* and n^* in the form (2.5), we need the partial derivatives $\partial^{r+s}/\partial m_1^r \partial m_2^s$, $(r+s \leq 4,\ r,s = 0,1,2,3,4)$ of θ^*, and n^*, and to evaluate these derivatives at $m_1' = \mu_1$, $m_2 = \mu_2$. Thus, if we write $\theta^* = f(m_1',m_2) = f$, the partial derivatives of f w.r.t. m_1' and m_2 calculated at $m_1 = \mu_1$ and $m_2 = \mu_2$ $(m_1' = m_1,\ \mu_1' = \mu_1)$, with $a = 4\beta(\beta-1)$ and $Y = \mu_1^2 + a\mu_1\mu_2$ will be given as

(3.40.a) $$\frac{\overline{\partial} f}{\partial m_1} = A = (2\beta^2\mu_2)^{-1} - (2\mu_1+a\mu_2)/4\beta^2\mu_2 Y^{1/2}$$

(3.40.b) $$\frac{\overline{\partial} f}{\partial m_2} = B = -\mu_1/\mu_2 \cdot A$$

(3.40.c) $$\frac{\overline{\partial}^2 f}{\partial m_1^2} = C = a^2\mu_2/8\beta^2 Y^{3/2}$$

(3.40.d) $$\frac{\overline{\partial}^2 f}{\partial m_1' \partial m_2} = D = (2\beta^2\mu_2^2)^{-1} + (4\mu_1^3+6a\mu_1^2\mu_2+a^2\mu_1\mu_2^2)/8\beta^2\mu_2^2 Y^{3/2}$$

(3.40.e) $$\frac{\overline{\partial}^2 f}{\partial m_2^2} = E = -\mu_1(\beta^2\mu_2^3)^{-1} - (8\mu_1^4+12a\mu_1^3\mu_2+3a^2\mu_1^2\mu_2^2)/8\beta^2\mu_2^3 Y^{3/2}$$

(3.40.f) $$\frac{\overline{\partial}^2 f}{\partial m_1^3} = F = -3a^2\mu_2(2\mu_1+a\mu_2)/16\beta^2 Y^{5/2}$$

(3.40.g) $$\frac{\overline{\partial}^3 f}{\partial m_1^2 \partial m_2} = G = \alpha^2(2\mu_1^2-\alpha\mu_1\mu_2)/16\beta^2 Y^{5/2}$$

(3.40.h) $$\frac{\overline{\partial}^3 f}{\partial m_1 \partial m_2^2} = H = -(\beta^2\mu_2^3)^{-1}$$
$$- (16\mu_1^5+40\alpha\mu_1^4\mu_2+30\alpha^2\mu_1^3\mu_2^2+3\alpha^3\mu_1^2\mu_2^3)/16\beta^2\mu_2^3 Y^{5/2}$$

(3.40.i) $$\frac{\overline{\partial}^3 f}{\partial m_2^3} = I = 3\mu_1(\beta^2\mu_2^4)^{-1}$$
$$+ (48\mu_1^6+120\alpha\mu_1^5\mu_2+90\alpha^2\mu_1^4\mu_2^2+15\alpha^3\mu_1^3\mu_2^3)/16\beta^2\mu_2^4 Y^{5/2}$$

(3.40.j) $$\frac{\overline{\partial}^4 f}{\partial m_1^4} = J = 3\alpha^2\mu_2(16\mu_1^2+16\alpha\mu_1\mu_2+5\alpha^2\mu_2^2)/32\beta^2 Y^{7/2}$$

(3.40.k) $$\frac{\overline{\partial}^4 f}{\partial m_1^3 \partial m_2} = K = -3\alpha^2(4\mu_1^3-2\alpha\mu_1^2\mu_2-\alpha^2\mu_1\mu_2^2)/32\beta^2 Y^{7/2}$$

(3.40.1) $$\frac{\overline{\partial}^4 f}{\partial m_1^2 \partial m_2^2} = L = -(12\alpha^3\mu_1^3-3\alpha^4\mu_1^2\mu_2)/32\beta^2 Y^{7/2}$$

(3.40.m) $$\frac{\overline{\partial}^4 f}{\partial m_1 \partial m_2^3} = M = 3(\beta^2\mu_2^4)^{-1}$$
$$+ (96\mu_1^7+336\alpha\mu_1^6\mu_2+420\alpha^2\mu_1^5\mu_2^2+210\alpha^3\mu_1^4\mu_2^3+15\alpha^4\mu_1^3\mu_2^4)/32\beta^2\mu_2^4 Y^{7/2}$$

(3.40.z) $$\frac{\overline{\partial}^4 f}{\partial m_2^4} = Z = -12\mu_1(\beta^2\mu_2^5)^{-1}-(384\mu_1^3+1344\alpha\mu_1^7\mu_2+276\alpha^2\mu_1^6\mu_2^2$$
$$+840\alpha^3\mu_1^5\mu_2^3+105\alpha^4\mu_1^4\mu_2^4)/32\beta^2\mu_2^5 Y^{7/2} \quad .$$

On the other hand, if we write $n^* = f^*(m_1', m_2) = f^*$, the partial derivatives of f^* w.r.t. m_1, m_2 calculated at $m_1 = \mu_1$ and $m_2 = \mu_2$, with $l=(2\beta-1)/2$ and $W = \mu_1^4 + \alpha\mu_1^3\mu_2$, will be given as

$$\frac{\overline{\partial} f^*}{\partial m_1} = A^* = l(\mu_2-\mu_1)^{-2}(2\mu_1\mu_2-\mu_1^2) \tag{3.41.a}$$
$$+ (4\mu_1^3\mu_2-2\mu_1^4-a\mu_1^3\mu_2+3a\mu_1^2\mu_2^2)/4(\mu_2-\mu_1)^2W^{1/2}$$

$$\frac{\overline{\partial} f^*}{\partial m_2} = B^* = -l\mu_1^2(\mu_2-\mu_1)^{-2} - (a\mu_1^4+2\mu_1^4+a\mu_1^3\mu_2)/4(\mu_2-\mu_1)^2W^{1/2} \tag{3.41.b}$$

$$\frac{\overline{\partial}^2 f^*}{\partial m_1^2} = C^* = 2l\mu_2^2(\mu_2-\mu_1)^{-3} + [\mu_1^6(8\mu_2^2+4a\mu_2^2-a^2\mu_2^2) \tag{3.41.c}$$
$$+ \mu_1^5(12a\mu_2^3+6a^2\mu_2^3) + 3a^2\mu_1^4\mu_2^4]/8(\mu_2-\mu_1)^3W^{3/2}$$

$$\frac{\overline{\partial}^2 f^*}{\partial m_1 \partial m_2} = D^* = -2l\mu_1\mu_2(\mu_2-\mu_1)^{-3} \tag{3.41.d}$$
$$+ (a^2\mu_1^7\mu_2-3a^2\mu_1^5\mu_2^3-6a^2\mu_1^6\mu_2^2-4a\mu_1^7\mu_2-12a\mu_1^6\mu_2^2-8\mu_1^7\mu_2)/8(\mu_2-\mu_1)^3W^{3/2}$$

$$\frac{\overline{\partial}^2 f^*}{\partial m_2^2} = E^* = 2l\mu_1^2(\mu_2-\mu_1)^{-3} \tag{3.41.e}$$
$$+ [\mu_1^8(8+4a-a^2)+\mu_1^7(12a\mu_2+6a^2\mu_2)+3a^2\mu_1^6\mu_2^2]/8(\mu_2-\mu_1)^3W^{3/2}$$

$$\frac{\overline{\partial}^3 f^*}{\partial m_1^3} = F^* = 6l\mu_2^2(\mu_2-\mu_1)^{-4} + [\mu_1^{10}(48\mu_2^2+24a\mu_2^2-6a^2\mu_2^2) \tag{3.41.f}$$
$$+ \mu_1^9(120a\mu_2^3+60a^2\mu_2^3-3a^3\mu_2^3)+\mu_1^8(90a^2\mu_2^4+27a^3\mu_2^4)$$
$$+ \mu_1^7(27a^3\mu_2^5)-3\mu_1^6a^3\mu_2^6]/16(\mu_2-\mu_1)^4W^{5/2}$$

$$\frac{\overline{\partial}^3 f^*}{\partial m_1^2 \partial m_2} = G^* = -2l(\mu_2^2+2\mu_1\mu_2)(\mu_2-\mu_1)^{-4} + [\mu_1^{11}(4a^2\mu_2-16a\mu_2-32\mu_2) \tag{3.41.g}$$
$$+ \mu_1^{10}(a^3\mu_2^2-38a^2\mu_2^2-88a\mu_2^2-16\mu_2^2)-\mu_1^9(13a^3\mu_2^3+80a^2\mu_2^3+40a\mu_2^3)$$
$$- \mu_1^8(33a^3\mu_2^4+30a^2\mu_2^4)-\mu_1^7(3a^3\mu_2^5)]/16(\mu_2-\mu_1)^4W^{5/2}$$

$$\frac{\partial^3 f^*}{\partial m_1 \partial m_2^2} = H^* = 2l\mu_1(2\mu_2+\mu_1)(\mu_2-\mu_1)^{-4}+[\mu_1^{12}(16+8a-2a^2)$$

(3.41.h)
$$+ \mu_1^{11}(32\mu_2+56a\mu_2+16a^2\mu_2+a^3\mu_2)+\mu_1^{10}(8a\mu_2^2+70a^2\mu_2^2-a^3\mu_2^2)$$

$$+ \mu_1^9(60a^2\mu_2^3+39a^3\mu_2^3)+\mu_1^8(9a^3\mu_2^4)]/16(\mu_2-\mu_1)^4W^{5/2}$$

$$\frac{\partial^3 f^*}{\partial m_2^3} = I^* = -6l\mu_1^2(\mu_2-\mu_1)^{-4}+[\mu_1^{12}(6a^2-3a^3-24a-48)$$

(3.41.i)
$$+ \mu_1^{11}(15a^3\mu_2-60a^2\mu_2-120a\mu_2)-45a^2\mu_1^{10}\mu_2^2(2+a)$$

$$- 15\mu_1^9a^3\mu_2^3]/16(\mu_2-\mu_1)^4W^{5/2}$$

$$\frac{\partial^4 f^*}{\partial m_1^4} = J^* = 24l\mu_2^2(\mu_2-\mu_1)^{-5}+[\mu_1^9(48\mu_2^2+24a\mu_2^2-6a^2\mu_2^2)(3a\mu_1^4\mu_2$$

$$+ 8\mu_1^5+5a\mu_1^3\mu_2^2)+\mu_1^8(120a\mu_2^3+60a^2\mu_2^3-3a^3\mu_2^3)(3a\mu_1^3\mu_2^2+10\mu_1^5$$

$$+ 5a\mu_1^4\mu_2-2\mu_1^4\mu_2)+\mu_1^7(90a^2\mu_2^4+27a^3\mu_2^4)(a\mu_1^3\mu_2^2+12\mu_1^5$$

(3.41.j)
$$+ 7a\mu_1^4\mu_2-4\mu_1^4\mu_2)+\mu_1^6(243a^4\mu_1^4\mu_2^6-162a^3\mu_1^4\mu_2^6+378a^3\mu_1^5\mu_2^5$$

$$- 27a^4\mu_1^3\mu_2^7)+\mu_1^5(24a^3\mu_1^4\mu_2^7-33a^4\mu_1^4\mu_2^7-48a^3\mu_1^5\mu_2^6$$

$$+ 9a^4\mu_1^3\mu_2^8)]/32(\mu_2-\mu_1)^5W^{7/2}$$

$$\frac{\partial^4 f^*}{\partial m_1^3 \partial m_2} = K^* = -12l(\mu_2^2+\mu_1\mu_2)(\mu_2-\mu_1)^{-5}+[\mu_1^{10}(4\alpha^2\mu_2-16\alpha\mu_2$$

$$- 32\mu_2)(2\mu_2\mu_1^4+6\mu_1^5+7\alpha\mu_1^3\mu_2^2+\alpha\mu_1^4\mu_2)+\mu_1^9(\alpha^3\mu_2^2-38\alpha^2\mu_2^2-88\alpha\mu_2^2$$

$$- 16\mu_2^2)(5\alpha\mu_1^3\mu_2^2+8\mu_1^5+3\alpha\mu_1^4\mu_2)+\mu_1^8(13\alpha^3\mu_2^3+80\alpha^2\mu_2^3$$

(3.41.k)

$$+ 40\alpha\mu_2^3)(2\mu_1^4\mu_2-10\mu_1^5-5\alpha\mu_1^4\mu_2-3\alpha\mu_1^3\mu_2^2)+\mu_1^7(30\alpha^2\mu_2^4$$

$$+ 33\alpha^3\mu_2^4)(4\mu_2\mu_1^4-12\mu_1^5-\alpha\mu_1^3\mu_2^2-7\alpha\mu_1^4\mu_2)+\mu_1^6(18\alpha^3\mu_1^4\mu_2^6$$

$$- 42\alpha^3\mu_1^5\mu_2^5+3\alpha^4\mu_1^3\mu_2^7-27\alpha^4\mu_1^4\mu_2^6)]/32(\mu_2-\mu_1)^5W^{7/2}$$

$$\frac{\partial^4 f^*}{\partial m_1^2 \partial m_2^2} = L^* = 4l(\mu_2^2+4\mu_1\mu_2+\mu_1^2)(\mu_2-\mu_1)^{-5}+[\mu_1^{16}(64+32\alpha-8\alpha^2)$$

$$+ \mu_1^{15}(256\mu_2+352\alpha\mu_2+80\alpha^2\mu_2+8\alpha^3\mu_2)+\mu_1^{14}(64\mu_2^2+928\alpha\mu_2^2$$

(3.41.l)

$$- 8\alpha^3\mu_2^2+720\alpha^2\mu_2^2)+\mu_1^{13}(224\alpha\mu_2^3+1232\alpha^2\mu_2^3+712\alpha^3\mu_2^3+4\alpha^4\mu_2^3)$$

$$+ \mu_1^{12}(280\alpha^2\mu_2^4+680\alpha^3\mu_2^4+190\alpha^4\mu_2^4)+\mu_1^{11}(144\alpha^3\mu_2^5+180\alpha^4\mu_2^5)$$

$$+ 9\mu_1^{10}\alpha^4\mu_2^6)]/32(\mu_2-\mu_1)^5W^{7/2}$$

$$\frac{\partial^4 f^*}{\partial m_1 \partial m_2^3} = M^* = -12l\mu_1(\mu_1+\mu_2)(\mu_2-\mu_1)^{-5}+[\mu_1^{16}(-192-96\alpha+24\alpha^2-12\alpha^3)$$

$$+ \mu_1^{15}(-192\mu_2-768\alpha\mu_2-312\alpha^2\mu_2+72\alpha^3\mu_2+3\alpha^4\mu_2)$$

(3.41.m)

$$+ \mu_1^{14}(-672\alpha\mu_2^2-1176\alpha^2\mu_2^2-336\alpha^3\mu_2^2-12\alpha^4\mu_2^2)+\mu_1^{13}(-840\alpha^2\mu_2^3$$

$$- 840\alpha^3\mu_2^3-30\alpha^4\mu_2^3)+\mu_1^{12}(-420\alpha^3\mu_2^4-300\alpha^4\mu_2^4)$$

$$+ \mu_1^{11}(-45\alpha^4\mu_2^5)]/32(\mu_2-\mu_1)^5W^{7/2}$$

$$\frac{\partial^4 f^*}{\partial m_2^4} = Z^* = 24l\mu_1^2(\mu_2-\mu_1)^{-5}+[\mu_1^{12}(6a^2-3a^3-24a-48)(46\mu_1^4$$

$$- 13a\mu_1^3\mu_2)+\mu_1^{11}(15a^3-60a^2-120a)(3a\mu_1^4\mu_2-6\mu_1^4\mu_2-2\mu_1^5$$

(3.41.z) $$- 11a\mu_1^3\mu_2^2)+\mu_1^{10}(45a^3+90a^2)(4\mu_1^4\mu_2^2-a\mu_1^4\mu_2^2+4\mu_1^5\mu_2$$

$$+ 9a\mu_1^3\mu_2^3)+\mu_1^9(30a^3\mu_1^4\mu_2^3+15a^4\mu_1^4\mu_2^3+90a^3\mu_1^5\mu_2^2$$

$$+ 105a^4\mu_1^3\mu_2^4)]/32(\mu_2-\mu_1)^5W^{7/2} \quad .$$

Thus, by using the general form of the bias given in (2.8.b), the bias $B_1(\theta^*) = E[\theta^*-\theta]$ of the moment estimator θ^* of the parameter θ in the GNBD can be written as

(3.42.a) $$B_1(\theta^*) = k_1/N + N^{-2}\sum_{i=1}^{11}\pi_i + O(N^{-3})$$

where k_1 and π_i ($i = 1,2,\ldots,11$) are given in (2.8.c) and (2.8.d) respectively with the difference, that the values of the partial derivatives $(B,C,D,\ldots,Z)$ are those given in (3.40.b - 3.40.z). Thus

$$B_1(\theta^*) = N^{-1}\left\{\frac{\mu_1}{2\beta^2\mu_2} - \frac{\mu_1(2\mu_1+a\mu_2)}{4\beta^2\mu_2(\mu_1^2+a\mu_1\mu_2)^{1/2}} + \frac{a^2\mu_2^2}{16\beta^2(\mu_1^2+a\mu_1\mu_2)^{3/2}}\right.$$

(3.42.b) $$+\frac{\mu_3}{2\beta^2\mu_2^2} + \frac{\mu_3(4\mu_1^3+6a\mu_1^2\mu_2+a^2\mu_1\mu_2^2)}{8\beta^2\mu_2^2(\mu_1^2+a\mu_1\mu_2)^{3/2}} - \frac{(\mu_4-\mu_2^2)}{2}\left[\frac{\mu_1}{\beta^2\mu_2^3}\right.$$

$$\left.\left.+\frac{(8\mu_1^4+12a\mu_1^3\mu_2+3a^2\mu_1^2\mu_2^2)}{8\beta^2\mu_2^3(\mu_1^2+a\mu_1\mu_2)^{3/2}}\right]\right\} + N^{-2}\sum_{i=1}^{11}\pi_i + O(N^{-3})$$

On the other hand, if we replace $A, B, C, D, \ldots, Z$ by $A^*, B^*, \ldots, Z^*$, given in (3.41.a - 3.41.z) in formulae (2.8.c) and (2.8.d), the bias $B_2(n^*)$ of the moment estimator n^* can be written as

$$\begin{aligned}
B_2(n^*) \doteqdot N^{-1}\Bigg\{ & \frac{l\mu_2\mu_1^2}{(\mu_2-\mu_1)^2} + \mu_2(a\mu_1^4+2\mu_1^4+a\mu_1^3\mu_2)/4(\mu_2-\mu_1)^2(\mu_1^4 \\
& + a\mu_1^3\mu_2)^{1/2} + \frac{l\mu_2^3}{(\mu_2-\mu_1)^3} + [\mu_1^6(8\mu_2^3+4a\mu_2^3-a^2\mu_2^3) \\
& + \mu_1^5(12a\mu_2^4+6a^2\mu_2^4)+3a^2\mu_1^4\mu_2^5)]/16(\mu_2-\mu_1)^3(\mu_1^4+a\mu_1^3\mu_2)^{3/2} \\
& - \frac{2l\mu_1\mu_2\mu_3}{(\mu_2-\mu_1)^3} + \mu_3(a^2\mu_1^7\mu_2-3a^2\mu_1^5\mu_2^3-6a^2\mu_1^6\mu_2^2-4a\mu_1^7\mu_2 \qquad (3.43) \\
& - 12a\mu_1^6\mu_2^2-8\mu_1^7\mu_2)/8(\mu_2-\mu_1)^3(\mu_1^4+a\mu_1^3\mu_2)^{3/2} + \frac{l\mu_1^2(\mu_4-\mu_2^2)}{(\mu_2-\mu_1)^3} \\
& + (\mu_4-\mu_2^2)[\mu_1^8(8+4a-a^2)+\mu_1^7(12a\mu_2+6a^2\mu_2) \\
& + 3a^2\mu_1^6\mu_2^2]/16(\mu_2-\mu_1)^3(\mu_1^4+a\mu_1^3\mu_2)^{3/2}\Bigg\} + N^{-2}\sum_{j=1}^{11}\pi_j^* \ ,
\end{aligned}$$

where π_j^* $(j = 1, \ldots, 11)$ have the same form given in (2.8.d) with the difference that the values of the partial derivatives $(B, C, D, \ldots)$ are replaced by $B^*, C^*, D^*, \ldots$, given in (3.41.b - 3.41.z).

3.3.2 Asymptotic Variances and Covariance of Moment Estimators.

If we write $\theta^* = f(m_1', m_2)$, the general formula for the asymptotic variance of the moment estimator, given in (2.16.a), will give, on using (2.16.b)

$$\mathrm{Var}(\theta^*) \doteqdot N^{-1}\left[\frac{1}{2\beta^2\mu_2} - \frac{(2\mu_1+a\mu_2)}{4\beta^2\mu_2(\mu_1^2+a\mu_2)^{1/2}}\right]^2\left[\mu_2 - \frac{2\mu_1\mu_3}{\mu_2} - \frac{\mu_1^2\mu_4}{\mu_2^2} - \mu_1^2\right] \tag{3.44}$$
$$+ N^{-2}\sum_{i=1}^{12}\alpha_i$$

where α_i ($i = 1,2,\ldots,12$) are as defined in (2.16.c), and $A,B,C,\ldots,Z$ are the values of the partial derivatives given in (3.40.a - 3.40.z).

By the same argument, the asymptotic variance of $n^* = f^*(m_1', m_2)$ is given by

$$\mathrm{Var}(n^*) \doteqdot N^{-1}\left\{\mu_2\left[\frac{l(2\mu_1\mu_2-\mu_1^2)}{(\mu_2-\mu_1)^2} + \frac{(4\mu_1^3\mu_2-2\mu_1^4-a\mu_1^3\mu_2+3a\mu_1^2\mu_2^2)}{4(\mu_2-\mu_1)^2(\mu_1^4+a\mu_1^3\mu_2)^{1/2}}\right]^2\right.$$
$$- 2\mu_3\left[\frac{l(2\mu_1\mu_2-\mu_1^2)}{(\mu_2-\mu_1)^2} + \frac{(4\mu_1^3\mu_2-2\mu_1^4-a\mu_1^3\mu_2+3a\mu_1^2\mu_2^2)}{4(\mu_2-\mu_1)^2(\mu_1^4+a\mu_1^3\mu_2)^{1/2}}\right]\left[\frac{l\mu_1^2}{(\mu_2-\mu_1)^2}\right. \tag{3.45}$$
$$\left.+ \frac{(a\mu_1^4+2\mu_1^4+a\mu_1^3\mu_2)}{4(\mu_2-\mu_1)^2(\mu_1^4+a\mu_1^3\mu_2)^{1/2}}\right] + (\mu_4-\mu_2^2)\left[\frac{l\mu_1^2}{(\mu_2-\mu_1)^2}\right.$$
$$\left.\left.+ \frac{a\mu_1^4+2\mu_1^4+a\mu_1^3\mu_2}{4(\mu_2-\mu_1)^2(\mu_1^4+a\mu_1^3\mu_2)^{1/2}}\right]\right\} + N^{-2}\sum_{j=1}^{12}\alpha_j^*$$

and α_j^* $(j = 1,2,\ldots,12)$ are obtained on replacing $A,B,C,$ by $A^*,B^*,C^*,\ldots,Z^*$ in (2.16.c).

Substituting the values of the partial derivatives given in (3.40.a - 3.40.n) and (3.41.a - 3.41.z) in (2.19.c), the covariance between the moment estimators $\theta^* = f(m_1',m_2)$ and $n^* = f^*(m_1',m_2)$ may be written as

$$
\begin{aligned}
\mathrm{Cov}(\theta^*,n^*) = N^{-1}\Bigg\{ & \left[\frac{1}{2\beta^2\mu_2} - \frac{(2\mu_1+a\mu_2)}{4\beta^2\mu_2(\mu_1^2+a\mu_1\mu_2)^{1/2}}\right]\Bigg[\left(\mu_2 - \frac{\mu_1\mu_3}{\mu_2}\right) \\
& \times \left(\frac{l(2\mu_1\mu_2-\mu_1^2)}{(\mu_2-\mu_1)^2} + \frac{(4\mu_1^3\mu_2-2\mu_1^4-a\mu_1^3\mu_2+3a\mu_1^2\mu_2^2)}{4(\mu_2-\mu_1)^2(\mu_1^4+a\mu_1^3\mu_2)^{1/2}}\right) \\
& - \left(\mu_3 - \frac{\mu_1(\mu_4-\mu_2^2)}{\mu_2}\right)\left(\frac{l\mu_1^2}{(\mu_2-\mu_1)^2} + \frac{(a\mu_1^4+2\mu_1^4+a\mu_1^3\mu_2)}{4(\mu_2-\mu_1)^2(\mu_1^4+a\mu_1^3\mu_2)^{1/2}}\right)\Bigg]\Bigg\} \\
& + N^{-2}\sum_{j=1}^{11}\omega_j + O(N^{-3})
\end{aligned}
\tag{3.46}
$$

and ω_j $(j = 1,2,\ldots,11)$ are given in (2.19.d).

3.4. Sampling Properties of M.L. and Moment Estimators.

We shall now provide some tables to compare, at least numerically, between the performances of the two types of estimators. We shall confine our comparisons between first order terms of the biases for each of the estimated parameters θ and n. Although we have provided the expressions of the first and the second order terms in the biases,

Table (3.1)

Biases of $\hat{\theta}$ and θ^*, ($\beta = 2$).

θ	.01	.06	.11	.16	.21	.26	.31	.36	.41	.46
n \ $\beta\theta$	.02	.12	.22	.32	.42	.52	.62	.72	.82	.92
1	- .0103	- .0696	- .1429	- .2285	- .3212	- .4120	- .4893	- .5410	- .5581	- .5383
	-3.6886	-3.6066	-3.5131	-3.4079	-3.2910	-3.1623	-3.0216	-2.8687	-2.7035	-2.526
3	.0031	.0087	- .0155	- .0826	- .1870	- .2945	- .3585	- .3559	- .3001	- .2238
	-1.8318	-1.7305	-1.6274	-1.5227	-1.4166	-1.3095	-1.2016	-1.0933	- .9849	- .87714
5	.006	.0264	- .0043	- .1362	- .3504	- .5218	- .5500	- .4508	- .2990	- .1697
	-1.4604	-1.3552	-1.2502	-1.1456	-1.0417	- .9389	- .8376	- .7382	- .6412	- .5474
15	.0096	.0272	- .7398	-3.4282	-17.9398	-12.2805	- .8607	- .2417	- .0835	-1.3943
	-1.0891	- .9800	- .8731	- .7686	- .6669	- .5684	- .4736	- .3831	- .2975	- .2176
21	.0106	- .0413	-2.6763	-36.8865	-3.0176	- .3157	- .1016	- .0427	.0075	.3087
	-1.0360	- .9264	- .8192	- .7147	- .6133	- .5154	- .4216	- .3324	- .2484	- .1705
23	.0109	- .086	-3.8398	-2341.08	-1.1547	- .1863	- .0718	- .032	.0106	.2386
	-1.0245	- .9148	- .8075	- .7030	- .6017	- .5039	- .4103	- .3213	- .2377	- .1602
25	.0112	- .1462	-5.4939	-46.633	- .5737	- .124	- .005	- .0254	.0123	.1984
	-1.0148	- .9050	- .7976	- .6932	- .5919	- .4944	- .4008	- .3121	- .2288	- .1516
31	.0120	- .454	-23.0089	-1.4989	- .1483	- .0562	- .0327	- .0155	.0140	.1384
	- .9932	- .8832	- .7758	- .6713	- .5701	- .4727	- .3796	- .2915	- .2088	- .1325
35	.0125	- .8052	-520.15	- .5094	- .0845	- .0407	- .0262	- .0122	.0140	.1174
	- .9829	- .8728	- .7653	- .6608	- .5597	- .4625	- .3696	- .2816	- .1993	- .1234
39	.0131	-1.3104	-26.5886	- .2343	- .0564	- .0323	- .0221	- .0099	.0137	.127
	- .9748	- .8646	- .7571	- .6526	- .5515	- .4544	- .3616	- .2738	- .1917	- .1161

variances and the covariance of the moment estimators, the investigation of the behaviour of these functions can be achieved by using the computer facilities.

(i) <u>Tables of biases of $\hat{\theta}$ and θ^*</u>.

Table 3.1: Table giving the asymptotic biases of $\hat{\theta}$ and θ^* to order N^{-1}. The first entry in each cell is for $\hat{\theta}$ and the second for θ^*, $(\beta = 2)$.

For most of the selected values of the parameters, the bias of the M.L. estimator is negative. At some values of θ and n the bias changes sign which means that if the first order term is a sufficient approximation to the value of the bias, then, that bias reaches zero within some interval for the changing parameter. As an example, for $3 \leq n \leq 15$ and $.06 \leq \theta \leq .11$; for $21 \leq n \leq 39$ and $.36 \leq \theta \leq .41$, the bias of the M.L. estimator attains the value zero.

For the moment estimator θ^*, the bias is negative for all the selected values of the parameters. The bias of the M.L. estimator and that of the moment estimator for the parameter θ increase as we increase the values of n and θ. In terms of absolute values, the bias of the moment estimator is higher than that of the M.L. estimator. Since the general term in the expansion of the function of bias of the M.L. estimator is difficult to determine analytically, the pattern of convergence of that series is not known to us. Although

Table (3.2)

Biases of $\hat{\theta}$ and θ^*, (β = 15).

θ	.01	.015	.02	.025	.030	.035	.04	.045	.05	.055
n \ $\beta\theta$	.150	.225	.300	.375	.450	.525	.600	.675	.750	.825
1	- .1425	- .2152	- .2884	- .3618	- .4351	- .5078	- .5796	- .6502	- .7191	- .7862
	-2.9082	-3.0497	-3.1899	-3.3287	-3.4663	-3.6024	-3.7373	-3.8707	-4.0028	-4.1356
3	- .0435	- .0675	- .0926	- .1185	- .1448	- .1711	- .1968	- .2216	- .2451	- .2671
	-1.0186	-1.0613	-1.1035	-1.1454	-1.1868	-1.2278	-1.2684	-1.3086	-1.3483	-1.3876
5	- .0236	- .0380	- .0537	- .0704	- .0875	- .1047	- .1213	- .1369	- .1512	- .1638
	- .6407	- .6636	- .6863	- .7087	- .7309	- .7529	- .7746	- .7962	- .8174	- .8385
15	- .0037	- .0089	- .0165	- .0259	- .036	- .0457	- .0537	- .0594	- .0626	- .0637
	- .2627	- .2659	- .26897	- .27202	- .2750	- .2779	- .2809	- .2837	- .2865	- .2893
21	- .0008	- .0053	- .0130	- .0232	- .0341	- .0437	- .0505	- .0536	- .0535	- .0512
	- .2087	- .2091	- .2094	-2.096	- .2099	- .2101	- .2103	- .2105	- .2107	- .2108
23	- .0002	- .0047	- .0127	- .0234	- .0349	- .0447	- .0510	- .0533	- .0521	- .0468
	- .19701	- .1967	- .1964	- .1961	- .1957	- .1954	- .1950	- .1946	- .1942	- .1938
25	.0003	- .0042	- .0126	- .0241	- .0362	- .0462	- .0521	- .0535	- .0512	- .0468
	- .18715	- .1863	- .1855	- .1847	- .1838	- .1830	- .1821	- .1813	- .1804	- .1795
31	.0014	- .0036	- .014	- .0284	- .0429	- .0537	- .0581	- .0565	- .0507	- .0431
	- .1652	- .1632	- .1613	- .1593	- .1574	- .1554	- .1534	- .1515	- .1495	- .1476
35	.0019	- .0038	- .0160	- .0330	- .0496	- .0607	- .0639	- .0600	- .0516	- .0419
	- .1548	- .1523	- .1497	- .1473	- .1448	- .1423	- .1398	- .1373	- .1349	- .1324
39	.0022	- .0044	- .0189	- .0389	- .0578	- .0692	- .0709	- .0644	- .0533	- .0413
	- .1465	- .1435	- .1406	- .1377	- .1347	- .1318	- .1289	- .1261	- .1232	- .1203

the region of poor performance for the bias of the moment estimators seems to be wider than that of the M.L. estimator, there are some explosive values for the bias of the M.L. estimator, which can be explained as being due to some singularities. As an example, $n = 23$, $\theta = .16$,

$$b_1(\hat{\theta}) \approx -2341.08/N \quad .$$

We may also note, if we rely upon the first order approximation, we will get misleading asymptotics and this point will be discussed in the following section.

Table 3.2: Table giving the asymptotic biases of M.L. and moment estimators for θ, ($\beta = 15$).

It should be noticed that the chosen value for β is relatively large. Most of the values of the bias of $\hat{\theta}$ are negative, and at some values the sign changes which also means if we accept the first order term of the bias as a sufficient approximation, the function of the bias attains zero within some interval for the changing parameter. As an example, for $25 \le n \le 39$ and $.01 \le \theta \le .05$, the bias reaches zero.

For the moment estimator θ^*, the bias is negative for all the selected values of n and θ. In terms of absolute values, and for most of the selected values of θ and n, the bias of M.L. estimators is much smaller than that of the moment estimator.

For all values of θ and $1 \le n \le 21$, both biases decrease as we increase the values of θ and n. When $n \ge 23$, and $.01 \le \theta \le .045$,

Table (3.3)

Biases of $\hat{n}$ and n^*, (β = 2).

θ	.06	.11	.16	.21	.26	.31	.36	.41	.46
n \ $\beta\theta$	.12	.22	.32	.42	.52	.62	.72	.82	.92
1	.1139	.2456	.4083	.5945	.7902	.9750	1.1235	1.2096	1.2130
	-882.85	-107.35	-17.66	3.34	10.51	14.24	17.87	24.25	45.57
3	.5147	1.4385	2.8931	4.7074	6.3858	7.3341	7.2666	6.3451	4.0036
	-748.67	-64.96	3.44	16.14	19.17	20.53	22.39	27.97	48.54
5	1.2675	4.4842	10.5100	17.9276	23.0030	23.5187	20.2548	15.0186	10.0051
	-614.5	-22.58	24.54	28.94	27.84	26.81	27.38	31.69	51.50
15	21.741	175.6763	525.1646	2289.4868	1422.0229	94.8669	25.8556	6.0046	513.2433
	-56.39	189.33	130.05	92.97	71.18	58.21	51.18	50.28	66.34
21	75.5515	755.5711	7326.9374	477.85	40.4924	9.9817	1.3769	-11.1395	-193.6934
	468.92	316.49	193.35	131.38	97.19	77.05	65.46	61.44	75.23
23	109.7145	1154.5883	496632.9372	189.862	23.6719	6.4757	.0224	-12.1444	-165.904
	593.100	358.87	214.45	144.19	105.85	83.33	70.22	65.16	78.20
25	156.4794	1756.2975	10458.9239	96.5363	15.4105	4.5232	-.8025	-12.8732	-151.1613
	727.27	401.25	235.55	156.99	114.52	89.61	74.98	68.87	81.17
31	411.546	8663.8329	375.4026	25.0314	6.3540	2.0485	-2.0699	-14.3689	-132.2802
	1129.81	528.41	298.85	195.41	140.53	108.45	89.26	80.03	90.07
35	727.8247	213941.7223	131.7262	13.8438	4.3730	1.4010	-2.5291	-15.0608	-127.1374
	1398.16	613.17	341.05	221.02	157.87	121.01	98.77	87.47	96.00
39	1220.7434	16142.7974	60.9256	8.8942	3.3666	1.0248	-2.862	-15.6028	-124.228
	1666.52	697.94	383.26	246.63	175.20	133.57	108.29	94.91	101.93

the bias of the M.L. estimator decreases and starts to increase at $\theta \geq .05$. Also, for the moment estimator θ^*, and for the values of $n \geq 23$, and $.01 \leq \theta \leq .055$, the bias increases monotonically.

This table is strikingly different from table 3.1 in the sense that the explosive values that appeared in table 3.1 in the values of the bias of $\hat{\theta}$ disappeared when we chose a large value for β. A more detailed study to higher order terms is required so that one can establish an idea about how reliable the first order term is, and we shall discuss this point in the last section.

(ii) Tables of Biases of n and n^*.

Table 3.3: This table gives the asymptotic bias of M.L. and moment estimators. The two entries in a cell relate to order N^{-1} in the biases of n and n^* respectively, ($\beta = 2$).

For all values of θ, and for values of $1 \leq n \leq 15$, the bias of $\hat{n}$ is positive, and for those values of $n \geq 21$, that bias starts to change sign at different places in the table. As an example, for $21 \leq n \leq 23$, $36 \leq \theta \leq 41$, and for $25 \leq n \leq 39$ and $.31 \leq \theta \leq .36$, the bias of n changes its sign from negative to positive, which means it reaches zero within some of the previously described intervals. On the other hand, and unlike the bias of $\hat{n}$, the bias of n^* changes sign for all values of θ, and for $n < 15$. For those values of $n \geq 15$ it is always positive.

For $n = 1$ and all values of θ, the biases of $\hat{n}$ and n^* increase monotonically, and for all $n \geq 3$, both biases reach a maximum and start to decline again.

We notice in this table and as in table (3.1), the bias of $\hat{n}$ has an explosive value for $\theta = .16$ and $n = 23$, where

$$b_2(\hat{n}) \approx 496632.9372/N \quad .$$

It is in fact, as in the other tables, difficult to determine the subregion of the parameter space, where one of the estimators is superior to the other estimator, because we need, in fact, some more information about the behaviour of second order terms. But we can see, in general, that the biases of n^* are much higher than those of $\hat{n}$, and we shall give further investigation to the moments of n^* at the end of this chapter.

Table 3.4: Table giving the asymptotic biases of M.L. and moment estimators for the parameter n, ($\beta = 15$).

We realize from this table that the bias of $\hat{n}$ is positive throughout the entire table, where the bias of n^* changes sign at some points of the selected values of the parameters. As an example; for $15 \leq n \leq 25$, and $.02 \leq \theta \leq .025$, the bias of n^* changes sign from negative to positive, which means that it reaches zero at some values within the prescribed subregion, which is the narrow interval $(.02, .025)$ of the parameter θ. The minimum value of the bias of n^* is attained when $n = 1$, and $\theta = .01$, where its maximum value is reached at $n = 39$, and $\theta = .055$. On the other hand, the minimum value of the bias of $\hat{n}$ is also attained when $n = 1$, and $\theta = .01$, where it reaches its maximum for $n = 39$, and $\theta = .03$.

Table (3.4)

Biases of $\hat{n}$ and n^*, (β = 15).

θ	.01	.015	.02	.025	.030	.035	.04	.045	.05	.055
n \ βθ	.150	.225	.300	.375	.400	.525	.600	.675	.750	.825
1	.1508	.2296	.3099	.3910	.4722	.5528	.6321	.7094	.7838	.8548
	-2824.22	-727.01	-222.23	-42.92	39.30	86.97	122.92	158.90	205.80	284.41
3	.4840	.7557	1.0398	1.3295	1.6178	1.8971	2.1601	2.3996	2.6089	2.7825
	-2766.82	-699.33	-205.45	-35.41	47.83	93.63	128.31	163.39	209.62	287.73
5	.8628	1.3805	1.9352	2.5065	3.0717	3.6070	4.0888	4.4958	4.8102	5.0188
	-2709.43	-671.66	-188.67	-19.90	56.36	100.28	133.70	167.88	213.45	291.05
15	3.5932	6.4602	9.8785	13.5513	17.0762	20.0252	22.0402	22.908	22.5891	21.2001
	-2422.46	-533.28	-104.77	37.66	98.99	133.55	160.66	190.35	232.59	307.63
21	6.0886	11.6896	18.7364	26.4310	33.6035	39.0553	41.9418	39.4369	41.9833	34.9192
	-2250.28	-450.25	-54.43	72.19	124.58	153.52	176.83	203.83	244.09	317.58
23	7.1001	13.9238	22.6495	32.2104	41.0130	47.4591	50.5041	46.1605	49.9133	40.14
	-2192.89	-422.57	-37.64	83.71	133.11	160.17	182.22	208.32	247.89	320.90
25	8.2134	16.4466	27.1386	38.8826	49.5467	57.0483	58.6853	60.1376	53.4571	45.6840
	-2135.49	-394.89	-20.86	95.22	141.63	166.83	187.61	212.82	251.72	324.22
31	12.2452	26.0498	44.7368	65.2825	83.0247	93.8743	96.1018	78.9722	90.4183	64.3496
	-1963.31	-311.87	29.48	129.75	167.22	186.79	203.79	226.29	263.20	334.17
35	15.5996	34.4932	60.6810	89.338	113.1000	126.1063	126.6442	99.3125	116.5525	78.6669
	-1848.53	-256.52	63.04	152.77	184.27	200.10	214.57	235.28	270.86	340.80
39	19.5720	44.9421	80.8198	119.7300	150.5526	163.1959	165.436	147.2921	122.7895	94.7747
	-1733.74	-201.16	96.60	175.80	201.33	213.41	225.36	244.27	278.51	347.44

It should also be realized that, for all values of $n \geq 1$, and $.01 \leq \theta \leq .055$, the bias of n^* increases by increasing the values of both θ and n, where, only for those values of $1 \leq n \leq 15$, the bias of $\hat{n}$ increases monotonically. For $n \geq 15$, the bias of $\hat{n}$ is no longer monotonic, where we realize for such an interval, that bias increases by increasing θ, until it reaches a maximum and starts to decline again.

We would like to point out the similarity of the behaviour of the function of the bias of $\hat{n}$ given in this table and the behaviour of the function of the bias of $\hat{\theta}$ in table (3.2), in the sense that explosive values do not occur unlike tables (3.1) and (3.2), when we take $\beta = 2$. One may conjecture that singularities in the biases of M.L. estimators for n and θ are severe when β is small.

3.5 Concluding Remarks.

In addition to the previous comments on the given tables, we would like to call attention to some characteristics of the sample estimators of the parameters θ and n of the GNBD.

(i) We have shown that joint sufficient statistics for the parameters n and θ do not exist, which made the M.L. estimators, not only difficult to evaluate, but also that they may not exist. Consequently, we should not rely upon the first order terms for the biases, variances and the covariance of the M.L. estimators, and more effort should be made to derive the second order terms, although the task seems to be extremely difficult. The unreliability of the first order terms of the moments of the M.L. estimators can be exhibited by the following values:

(1) For $\beta = 2$, $\theta = .36$, $n = 11$

$$\mathrm{Var}(\hat{\theta}) \approx .0254/N \ ,$$

(2) For $\beta = 2$, $\theta = .36$, $n = 15$

$$\mathrm{Var}(\hat{\theta}) \approx .0181/ N,$$

(3) For $\beta = 2$, $\theta = .46$, $n = 11$

$$\mathrm{Var}(\hat{\theta}) \approx .0033/N \ ,$$

(4) For $\beta = 2$, $\theta = .46$, $n = 15$

$$\mathrm{Var}(\hat{\theta}) \approx .0027/N \ .$$

The above values show that the variance of $\hat{\theta}$ decreases by increasing θ and n for constant β.

On the other hand, the variance of $\hat{n}$ has very large values at some points of the parameter space. For example, for $\beta = 2$, $\theta = .11$, $n = 33$

$$\mathrm{Var}(\hat{n}) \sim 20540.531/N,$$

and for $\beta = 2$, $\theta = .11$, $n = 35$

$$\mathrm{Var}(\hat{n}) \approx 22971.195/N,$$

and the question about the largeness of the sample size which is required to reduce the effect of these inflated values can only be answered if at least the second order term in the variance of $\hat{n}$ is known, as will be seen in the next point.

(ii) From tables (3.1) and (3.2) in the previous section, we have seen that the values of the biases of the moment estimator θ^* are reliable for certain regions in the parameter space, and on using the second order term of the bias of that estimator, one can determine the sample size required to reduce the effect of that term. The idea was given by Shenton and Bowman (73) as follows.

Let us write

$$B_1(\theta^*) = \frac{k_1}{N} + \frac{1}{N^2}\sum_{i=1}^{11} \pi_i + O\left(\frac{1}{n^3}\right) ,$$

then for $0 < a < 1$,

$$N = \sum_{i=1}^{11} \pi_i / a k_1$$

is the required sample size to make the second order term in the bias of θ^* a certain proportion of the first order term, which can be easily obtained on using the terms given in (3.42.b). On the other hand, one should expect a larger sample size for this purpose in the case of the moment estimator n^*

Shenton and Bowman have indicated that (for $\beta = 1$), the potential instability in the moments of n^* is due to the presence of singularities, which in turn are related to the probability of occurence of inadmissible parameter estimates.

It should be realized that, for the GNBD, the moment estimator n^* suffers from this unfortunate situation, where the stochastic difference $(m_2 - m_1)$, which was called "over dispersion", by Bliss (5), appears in the denominator of n^* as given in (3.39).

The biases of the moment estimator of the parameter n, as shown in tables (3.3) and (3.4) are very large, and the reason is due to the instability of that estimator. Shenton and Bowman (73), postulated that the quantity $(m_2 - m_1')$, being in the denominator, will create singularity in n^*. Their suggested method for finding a safe sample size will not depend upon the value of the second order term in the bias, but it will depend upon the size of the coefficient of variation of the stochastic difference (m_2-m_1'). Thus, if we want the coefficient of variation v, where

$$v = \sqrt{\mathrm{Var}(m_2-m_1')}/E(m_2-m_1') \quad ,$$

to be some value between zero and one, which means $E(m_2-m_1')$ should be large in comparison to $[\mathrm{Var}(m_2-m_1')]^{\frac{1}{2}}$, the required sample size will be given as

$$N = (-b + \sqrt{b^2-4ac})/2a \quad , \tag{3.47.a}$$

where

$$a = v^2(\mu_2-\mu_1')^2 \tag{3.47.b}$$

$$b = -[2v^2\mu_2(\mu_2-\mu_1')+\mu_4-2\mu_3+\mu_2-\mu_2^2] \tag{3.47.c}$$

$$c = v^2\mu_2^2 - (4\mu_2^2+2\mu_3-2\mu_4) \quad . \tag{3.47.d}$$

This sample size should be large enough to obtain reliable asymptotes for θ^* also. Some values of the sample sizes are given in tables (3.5) and (3.6).

Table (3.5)

$\beta = 2,\ v = .2$

Estimated Sample Size N

n \ θ	.01	.16	.26	.31	.41	.46
1	237384	2221	1525	1522	2343	4833
3	116459	907	579	561	818	1645
7	81909	531	309	286	383	734
15	68089	382	201	177	209	370

Table (3.6)

$\beta = 15,\ v = .2$

Estimated Sample Size N

n \ θ	.01	.015	.03	.04	.05	.055
1	38205	23403	14457	15480	21698	30000
3	13194	8021	4890	5208	7270	10035
7	6048	3627	2156	2273	3148	4331
15	3190	1869	1063	1099	1500	2050

One should note the impractical sample size which is needed at some values of the parameter space, to make the coefficient of variation of the overdispersion (0.2).

(iii) Finally, we shall provide some tables of the asymptotic relative efficiency. Using Fisher's definition of efficiency,

$$E_f = \frac{1}{N^2\Delta|D|}$$

where

(2.48.a)
$$\Delta = P_{11}P_{22} - P_{12}^2$$
$$= [n(1-\theta)P_{22}-\theta(1-\beta\theta)]/\theta(1-\theta)^2(1-\beta\theta)$$

and $|D|$ is the generalized variance of the moment estimators defined in (2.45), where

(2.48.b) $$|D| = \frac{2(1-\theta)^3}{N^2\theta[1-\beta(2-\theta)]^2}[n(1-\beta\theta)+2\beta(1-\theta)+\beta^2\theta-1] \quad .$$

Hence

(3.48.c) $$E_f = \frac{\theta^2(1-\beta\theta)[1-\beta(2-\theta)]^2}{2(1-\theta)[n(1-\theta)P_{22}-\theta(1-\beta\theta)][n(1-\beta\theta)+2\beta(1-\theta)+\beta^2\theta-1]} \quad .$$

Tables (3.7) will provide some values for the asymptotic efficiency of the moment estimators relative to the method of M.L. As can be seen for small values of θ and constant β, E_f approaches zero as n tends to infinity, and a good approximation to the efficiency is reached when β is much larger than n, and when $\beta\theta$ is very close to .5.

Tables (3.7)

Asymptotic Relative Efficiency of Moment Estimators

Relative to M.L. Estimators

Case I: $\beta = 2.$

n \ θ	.01	.02	.05	.07	.09	.10
1	.7165	.8133	.9316	.9617	.9782	.9837
4	.0354	.0584	.1268	.1693	.2080	.2258
7	.0095	.0168	.0400	.0555	.0703	.077
10	.0042	.0078	.0196	. 276	.0354	.0392

Case II: $\beta = 9.$

n \ θ	.01	.02	.05	.07	.09	.10
1	.9862	.9957	.9999	.9981	.9891	.9362
4	.4664	.6141	.7551	.7563	.6345	.2691
7	.1968	.3052	.4379	.4219	.2769	.0737
10	.1019	.1694	.2601	.2428	.1411	.0328

Case III: $\beta = 15.$

n \ θ	.001	.003	.005	.008	.009	.01
5	.2159	.34	.4372	.5364	.5581	.5796
8	.0865	.1557	.2147	.2866	.3068	.3254
11	.0445	.0847	.1213	.1690	.1831	.1963
14	.0265	.0523	.0766	.1094	.1192	.1286

Case IV: $\beta = 25$.

n \ θ	.001	.003	.005	.008	.009	.01
5	.3839	.5705	.6678	.7466	.7636	.7777
8	.1794	.3187	.4132	.5049	.5266	.5454
11	.0986	.1911	.2624	.3392	.3587	.3756
14	.0685	.1244	.1765	.2361	.2518	.2656

In concluding, since the M.L. estimators may not exist, and further investigations to the second order terms of the moments of the M.L. estimators seem to be a very difficult task, we would recommend using the moment estimators given in (3.38) and (3.39) for fitting the GNBD to numerical data, as they can be calculated easily. Moreover, an investigation to higher order terms for the moments of the moment estimators has become very handy using the corresponding expressions provided in this chapter, and hence one can determine the safe sample size and compromise between the required efficiency and the cost of sampling, if we are determined to use moment estimators.

CHAPTER IV

A BIVARIATE GENERALIZATION OF A CLASS OF POWER SERIES DISTRIBUTIONS "THE BIVARIATE MODIFIED POWER SERIES DISTRIBUTION"

4.1 Introduction.

The bivariate forms of many important discrete probability distributions have been studied by many statisticians. The trinomial, the double Poisson, the bivariate negative binomial, and the bivariate logarithmic series distributions are in fact the bivariate generalizations of the well-known univariate distributions. A systematic account of various families of distributions of bivariate discrete random variables have been given by Patil and Joshi (61), Johnson and Kotz (35), and Mardia (49) in their books. The class of bivariate Lagrange distribution (BLD), and the bivariate Borel-Tanner distribution (BBTD) were introduced by Shenton and Consul (74).

In this chapter we shall define a class of bivariate discrete distributions, under the title "Bivariate Modified Power Series Distribution", (BMPSD); and study some of its properties.

Definition: A BMPSD is defined by a bivariate discrete random variable (X,Y) having the probability distribution function

(4.1)

$$P(X=x,Y=y) = \begin{cases} a(x,y)(g(\theta_1,\theta_2))^x(h(\theta_1,\theta_2))^y/f(\theta_1,\theta_2) & (x,y) \in S \\ 0 & \text{otherwise} \end{cases}$$

where, S is a subset of the cartesian product of the set of non-negative integers with itself, $a(x,y) > 0$, $\theta_1,\theta_2 \geq 0$, $g(\theta_1,\theta_2)$, $h(\theta_1,\theta_2)$ and $f(\theta_1,\theta_2)$ are finite, positive, and differential functions of θ_1 and θ_2, $(\theta_1,\theta_2) \in \Omega$, (where $\Omega = \{(\theta_1,\theta_2):t_1<\theta_1<t_2,\ t_3<\theta_2<t_4,\ t_i$ are real numbers $i = 1,\ldots,4\}$). It is clear that Ω is a subset of the domain of convergence of the power series

(4.2)
$$f(\theta_1,\theta_2) = \sum_S a(x,y)g^x(\theta_1,\theta_2)h^y(\theta_1,\theta_2) \quad .$$

It can be easily seen that the BMPSD class includes, among others, the trinomial distribution (Mardia; (49)), the bivariate negative binomial (Lundberg; (48)), the double Poisson (Patil and Joshi; (61)), the bivariate logarithmic series (Mardia; (49)), the BBTD, and the generalized negative binomial distribution GBNBD (Mohanty; (51)). Many important families of BMPSD can be generated by using Poincaré generalization of the Lagrange expansion, as given in Goursat (17), either by the expansion of $f(\theta_1,\theta_2)$ in powers of $g(\theta_1,\theta_2)$ and $h(\theta_1,\theta_2)$ under the transformations $\theta_1 = g(\theta_1,\theta_2)\chi_1$, $\theta_2 = h(\theta_1,\theta_2)\chi_2$, where χ_1 and χ_2 are functions of θ_1 and θ_2 with the condition $\chi_i(0,0) \neq 0$ $(i = 1,2)$, or by the Lagrange expansion of the probability generating function (pgf) $\phi(t_1,t_2)$ in powers of u and v, under the transformations

$$t_1 = u \cdot l(t_1,t_2), \qquad t_2 = v \cdot k(t_1,t_2) \quad ,$$

where $l(t_1,t_2)$ and $k(t_1,t_2)$ are pgf's as suggested by Shenton and Consul (74).

Section 2 of this chapter contains the notations that will be used generally in this and the following chapter. Moments, cumulants and related properties will be studied in section 3. In section 4 we give the convolution property and we characterize the double Poisson and the bivariate negative binomial distributions in the class of BMPSD. In the last two sections we give some particular families of BMPSD and discuss some of their properties.

4.2 Notations.

For the sake of brevity and convenience we shall use the following symbols.

(i) The functions $g(\theta_1,\theta_2)$, $h(\theta_1,\theta_2)$ and $f(\theta_1,\theta_2)$ will be denoted by the letters g, h and f respectively.

(ii) The differential operator $\frac{\partial}{\partial\theta_i}$ will be abbreviated by ∂_i, and the functions $\partial_i \log g$, $\partial_i \log h$, $\partial_i \log f$ by g_i, h_i and f_i respectively. Similarly, $g_{ij} = \partial^2_{ij} \log g$ and so on, where i, $j = 1,2$.

(iii) Unless otherwise stated, $\sum$ will stand for the two-fold summation over all points $(x,y) \in S$, where S is as previously defined.

(iv) $\Delta = g_1h_2 - g_2h_1$, and for all non-negative integers r and s,

$$(4.3) \qquad \mu'_{r,s} = E[X^rY^s], \qquad \mu_{r,s} = E[(X-\mu'_{10})^r(Y-\mu'_{01})^s]$$

4.3 Moments and Cumulants of BMPSD.

Since $\sum P(X=x,Y=y)$ is always equal to unity, we have

$$f = \sum a(x,y) g^x h^y .$$

On differentiating the above equation partially w.r.t. θ_1 and θ_2 respectively, dividing by f and on summation, we get the equations

$$f_1 = g_1 \mu'_{10} + h_1 \mu'_{01}$$

$$f_2 = g_2 \mu'_{10} + h_2 \mu'_{01} \quad .$$

Solving for μ'_{10} and μ'_{01}, we have

(4.4.a) $$\mu'_{10} = (f_1 h_2 - f_2 h_1)/\Delta$$

(4.4.b) $$\mu'_{01} = (f_2 g_1 - f_1 g_2)/\Delta$$

To obtain a recurrence relation among the higher non-central product moments, we write,

$$\mu'_{r,s} = \sum x^r y^s \; a(x,y) \; g^x h^y / f \quad .$$

On differentiating partially w.r.t. θ_1 and θ_2 respectively, and on simplification, we obtain

$$\partial_1 \mu'_{r,s} = g_1 \mu'_{r+1,s} + h_1 \mu'_{r,s+1} - f_1 \mu'_{r,s}$$

$$\partial_2 \mu'_{r,s} = g_2 \mu'_{r+1,s} + h_2 \mu'_{r,s+1} - f_2 \mu'_{r,s} \quad .$$

By solving the above two equations simultaneously for $\mu'_{r+1,s}$ and $\mu'_{r,s+1}$ and on using the relations (4.4.a) and (4.4.b), we get the recurrence relations

$$\mu'_{r+1,s} = (h_2\cdot\partial_1\mu'_{r,s} - h_1\cdot\partial_2\mu'_{r,s})/\Delta + \mu'_{r,s}\mu'_{10} \tag{4.5.a}$$

$$\mu'_{r,s+1} = (g_1\cdot\partial_2\mu'_{r,s} - g_2\cdot\partial_1\mu'_{r,s})/\Delta + \mu'_{r,s}\mu'_{01}. \tag{4.5.b}$$

Similarly, by differentiating $\mu_{r,s}$ partially w.r.t. θ_1 and θ_2 respectively and on simplification one can get the following two recurrence relations between the central product moments,

$$\begin{aligned}\mu_{r+1,s} = \Delta^{-1}[(h_2\partial_1-h_1\partial_2)\mu_{r,s} &+ r\mu_{r-1,s}(h_2\partial_1-h_1\partial_2)\mu'_{10}\\ &+ s\mu_{r,s-1}(h_2\partial_1-h_1\partial_2)\mu'_{01}]\end{aligned} \tag{4.6.a}$$

and

$$\begin{aligned}\mu_{r,s+1} = \Delta^{-1}[(g_1\partial_2-g_2\partial_1)\mu_{r,s} &+ r\mu_{r-1,s}(g_1\partial_2-g_2\partial_1)\mu'_{10}\\ &+ s\mu_{r,s-1}(g_1\partial_2-g_2\partial_1)\mu'_{01}] \quad .\end{aligned} \tag{4.6.b}$$

Thus, by a proper choice of the integers r and s, the above formulae can be easily manipulated to obtain the marginal moments and the coefficient of correlation between X and Y. The recurrence relations among the factorial moments $\mu_x^{[r]} = E[X^{[r]}]$ and $\mu_y^{[s]} = E[Y^{[s]}]$ of the BMPSD can also be similarly obtained and are given by

$$(4.7.a) \qquad \mu_x^{[r+1]} = \Delta^{-1}(h_2\partial_1 - h_1\partial_2)\mu_x^{[r]} + (\mu'_{10} - r)\mu_x^{[r]}$$

$$(4.7.b) \qquad \mu_y^{[s+1]} = \Delta^{-1}(g_1\partial_2 - g_2\partial_1)\mu_y^{[s]} + (\mu'_{01} - s)\mu_y^{[s]}$$

where $\mu_x^{[1]} = \mu'_{10}$ and $\mu_y^{[1]} = \mu'_{01}$.

The joint cumulants can be obtained from the joint moments by employing the symbolic operator as suggested on page 83 by Kendall and Stuart (39). The simplification given by Harvey (31) to that technique is rather elegant and more straightforward as will be shown in the sequel. Consider an operator D with the following rules.

$D(c) = 0$, $\quad$ c is constant

$D(ck_{ij}) = cD(k_{ij})$, $\quad$ where k_{ij} is the ijth cumulant

$D(k_{ij}^r) = rk_{ij}^{r-1}D(k_{ij})$

$D(U+V) = D(U) + D(V)$ $\quad$ (U and V are polynomials in the joint

$D(UV) = UD(V) + VD(U)$ $\quad$ cumulants)

$D_1(k_{ij}) = k_{i+1,j}$

$D_2(k_{ij}) = k_{i,j+1}$.

With D so defined, one operates on μ'_{ij}, the ijth moment about an arbitrary origin, to obtain $\mu'_{i+1,j}$ and $\mu'_{i,j+1}$ as

$$\mu'_{i+1,j} = D_1(\mu'_{ij}) + k_{10}\mu'_{ij}$$

$$\mu'_{i,j+1} = D_2(\mu'_{ij}) + k_{01}\mu'_{ij}$$

(with the initial condition $\mu'_{00} = 1$).

As an example:

$$\mu'_{10} = D_1(1) + k_{10} = k_{10}$$

$$\mu'_{20} = D_1(\mu'_{10}) + k_{10}\mu'_{10}$$

$$= D_1(k_{10}) + k_{10}^2$$

$$= k_{20} + k_{10}^2 \quad ,$$

$$\mu'_{30} = D_1(\mu'_{20}) + k_{10}\mu'_{20}$$

$$= D_1(k_{20}+k_{10}^2) + k_{10}(k_{20}+k_{10}^2)$$

$$= k_{30} + 3k_{10}k_{20} + k_{10}^3 \quad ,$$

$$\mu'_{11} = D_2(\mu'_{10}) + k_{01}\mu'_{10}$$

$$= D_2(k_{10}) + k_{01}\mu'_{10}$$

$$= k_{11} + k_{01}k_{10} \quad ,$$

$$\mu'_{21} = k_{21} + k_{20}k_{01} + 2k_{11}k_{10} + k_{10}^2k_{01} \quad ,$$

and so on.

4.4 Some Properties of BMPSD.

4.4.1 Convolution Property of BMPSD.

Let (X_i,Y_i), $i = 1,2,\ldots,N$ be a random sample of size N taken from the BMPSD given by (4.1) and let $Z_1 = X_1+X_2+\ldots+X_N$ and $Z_2 = Y_1+Y_2+\ldots+Y_N$. When the functions h and g are zeros at $\theta_1 = 0$ and $\theta_2 = 0$, due to the properties of the power series functions, the joint probability function of (Z_1,Z_2) can be easily written in the form

$$(4.8) \qquad P(Z_1=z_1,Z_2=z_2) = b(Z_1,Z_2,N)g^{Z_1}h^{Z_2}/f^N$$

where $b(Z_1,Z_2,N) = \sum \prod_{i=1}^{N} a(x_i,y_i)$, and the summation extends over all the ordered N-tuples $\{(x_1,y_1),\ldots,(x_N,y_N)\}$ of non-negative integers of the set S under the conditions $x_1+x_2+\ldots+x_N = Z_1$ and $y_1+y_2+\ldots+y_N = Z_2$. Though it seems to be a very difficult summation to find, in actual practice, the function $b(Z_1,Z_2,N)$ can be easily obtained as a coefficient of $g^{Z_1}h^{Z_2}$ by expanding f^N in powers of g and h with the help of the bivariate Lagrange expansion (64), and equals

$$(Z_1!Z_2!)^{-1}\partial^{Z_1-1}\partial^{Z_2-1}[\chi_1^{Z_1}\chi_2^{Z_2}\partial_{12}(f^N)$$

$$(4.9) \qquad + \chi_1^{Z_1}\partial_1(\chi_2^{Z_2})(\partial_2 f^N) + \chi_2^{Z_2}(\partial_2\chi_1^{Z_1})(\partial_1 f^N)]\Big|_{\theta_1=\theta_2=0}$$

where $\chi_1 = \theta_1 g^{-1}$ and $\chi_2 = \theta_2 h^{-1}$.

4.4.2 Two Characterization Theorems.

Theorem 4.1. *The means* μ'_{10} *and* μ'_{01} *of a BMPSD with* $f(0,0) = 1$ *are proportional to the parametric functions* g *and* h *respectively if and only if it is a double Poisson probability distribution.*

Proof: Let the means μ'_{10} and μ'_{01} be proportional to the functions

g, h, and let c_1, c_2 be the constants of proportionality.

By equations (4.4.a) and (4.4.b) we obtain

$$f_1h_2 - f_2h_1 = c_1 g\Delta$$

$$f_1g_2 - f_2g_1 = -c_2 h\Delta \quad .$$

Solving for f_1 and f_2 one gets

(4.10.a) $f_1 = c_1gg_1 + c_2hh_1 = c_1\partial_1 g + c_2\partial_1 h$

(4.10.b) $f_2 = c_1gg_2 + c_2hh_2 = c_1\partial_2 g + c_2\partial_2 h$

the solution of equation (4.10.a) is $\ln f = c_1g + c_2h + A(\theta_2)$ and of (4.10.b) is $\ln f = c_1g + c_2h + B(\theta_1)$, i.e. $f = A_1(\theta_2)e^{c_1g+c_2h}$ and $f = B_1(\theta_1)e^{c_1g+c_2h}$ respectively. If the two relations are to hold in any domain of θ_1 and θ_2, then $A_1(\theta_2)$ and $B_1(\theta_1)$ must be, not only independent of θ_1 and θ_2 but also, must be equal. Since $f(0,0) = 1$, the function f must be of the form $f = e^{\phi_1+\phi_2}$ where

$$\phi_1 = c_1\{g-g(0,0)\} \qquad \text{and} \qquad \phi_2 = c_2\{h-h(0,0)\} \quad .$$

By the uniqueness of the series expansion, $a(x,y) = (x!y!)^{-1}$ and the BMPSD becomes a double Poisson.

Conversely; let (X,Y) be a double poisson random vector, with probability function

$$p(X=x,Y=y) = \frac{1}{x!y!}\,\theta_1^x\theta_2^y/e^{\theta_1+\theta_2} \quad .$$

Thus, on using (4.4.a) and (4.4.b) one can show that

$$\mu'_{10} = \theta_1 \qquad \text{and} \qquad \mu'_{01} = \theta_2 \quad ,$$

which establishes the required result.

Theorem 4.2. *The means* μ'_{10} *and* μ'_{01} *of a BMPSD with* $f(0,0) = 1$ *are equal to* $cg(1-g-h)^{-1}$ *and* $ch(1-g-h)^{-1}$ *respectively, where* c *is any real number, and* g, h *are the parametric functions of the BMPSD, if and only if it is a bivariate negative binomial distribution.*

Proof: Let

$$\mu'_{10} = \frac{f_1h_2-f_2h_1}{g_1h_2-g_2h_1} = cg(1-g-h)^{-1}$$

and

$$\mu'_{01} = \frac{f_2g_1-f_1g_2}{g_1h_2-g_2h_1} = ch(1-g-h)^{-1} \quad .$$

Solving the last two equations for f_1 and f_2 we get

$$f_1 = c(1-g-h)^{-1}[\partial_1 g+\partial_1 h]$$

$$f_2 = c(1-g-h)^{-1}[\partial_2 g+\partial_2 h]$$

i.e.

$$f = k_1(\theta_2)(1-g-h)^{-c} \qquad \text{and} \qquad f = k_2(\theta_1)(1-g-h)^{-c}$$

respectively. By the same argument given in Theorem (1), $k_1(\theta_2)$ and $k_2(\theta_1)$ are equal and independent of θ_1 and θ_2. Since $f(0,0) = 1$, the function f must be of the form $f = (1-\chi_1-\chi_2)^{-c}$, where

$$\chi_1 = \frac{g-g(0,0)}{1-g(0,0)-h(0,0)} \quad \text{and} \quad \chi_2 = \frac{h-h(0,0)}{1-g(0,0)-h(0,0)}$$

the converse of this theorem can be easily proved and one can show that for the bivariate negative binomial distribution, whose probability function is given as $(x+y+c-1)!(x!y!c!)^{-1}\theta_1^x\theta_2^y(1-\theta_1-\theta_2)^c$, $\mu'_{10} = c\theta_1(1-\theta_1-\theta_2)^{-1}$ and $\mu'_{01} = c\theta_2(1-\theta_1-\theta_2)^{-1}$.

Corollary: *By taking $h(\theta_1,\theta_2) = 0$ in the above two theorems, one gets the corresponding characterizations for the Poisson and the negative binomial distributions, respectively,* [*see* (18)].

4.5 Some Particular Families of BMPSD.

4.5.1 The Generalized Double Poisson Distribution (GDPD).

The bivariate random vector (X,Y) has a GDPD with parameters $(m_1,m_2,\theta_1,\theta_2)$ which we shall write as $(X,Y) \sim$ GDPD, if its probability distribution function is given by

$$P(X=x,Y=y) = \frac{(1+m_1x+m_2y)^{x+y-1}}{x!y!}\,\theta_1^x\theta_2^y\,\exp\{-(\theta_1+\theta_2)(1+m_1x+m_2y)\}$$

(4.11.a)

$$x,y \geq 0(1)$$
$$\theta_i,m_i > 0 \ (i = 1,2)$$
$$0 < \theta_1m_1 + \theta_2m_2 < 1 \quad .$$

The distribution can be obtained by expanding the bivariate probability generating function $\phi(t_1,t_2) = \exp[\theta_1(t_1-1)+\theta_2(t_2-1)]$ using the bivariate Lagrange expansion formula (64), under the transformations

$$t_1 = u\cdot l(t_1,t_2), \qquad t_2 = v\cdot k(t_1,t_2) \quad ,$$

in powers u and v, where

$$l(t_1,t_2) = \exp\{m_1\theta_1(t_1-1) + m_1\theta_2(t_2-1)\}$$

$$k(t_1,t_2) = \exp\{m_2\theta_1(t_1-1) + m_2\theta_2(t_2-1) \quad ,$$

are two probability generating functions.

The GDPD is a BMPSD, where

$$g = \theta_1 \exp[-m_1\theta_1-m_1\theta_2], \qquad h = \theta_2 \exp[-m_2\theta_1-m_2\theta_2],$$

(4.11.b) $$\text{and } f = \exp[\theta_1+\theta_2] \quad .$$

4.5.2 <u>The Generalized Bivariate Negative Binomial Distribution (GBNBD).</u>

The bivariate random vector (X,Y) has a GBNBD with parameters $(n,\beta_1,\beta_2,\theta_1,\theta_2)$, which we shall write $(X,Y) \sim$ GBNBD, if its probability distribution function is given by

$$P(X=x,Y=y) = \frac{n\Gamma(n+\beta_1x+\beta_2y)}{x!y!\Gamma(n+\beta_1x+\beta_2y-x-y+1)}\,\theta_1^x\theta_2^y(1-\theta_1-\theta_2)^{n+\beta_1x+\beta_2y-x-y}$$

$$x,y \geq 0(1)$$

$$0 < \theta_1,\theta_2 < 1$$

(4.12.a) $$0 < \theta_1\beta_1+\theta_2\beta_2 < 1 \quad .$$

The distribution can be obtained by expanding the bivariate probability generating function $\phi(t_1,t_2) = (1-\theta_1-\theta_2+\theta_1t_1+\theta_2t_2)^n$, using the same bivariate Lagrange expansion formula, under the transformations

$$t_1 = u \cdot l(t_1, t_2), \qquad t_2 = v \cdot k(t_1, t_2) \quad ,$$

in powers of u and v, where

$$l(t_1, t_2) = (1-\theta_1-\theta_2+\theta_1 t_1+\theta_2 t_2)^{\beta_1}$$

$$k(t_1, t_2) = (1-\theta_1-\theta_2+\theta_1 t_1+\theta_2 t_2)^{\beta_2}$$

are two probability generating functions.

The GBNBD is a BMPSD, with parametric functions,

$$g = \theta_1(1-\theta_1-\theta_2)^{\beta_1-1}, \qquad h = \theta_2(1-\theta_1-\theta_2)^{\beta_2-1},$$

(4.12.b) and $f = (1-\theta_1-\theta_2)^{-n}$

4.5.3 <u>The Generalized Bivariate Logarithmic Series Distribution (GBLSD).</u>

The bivariate random vector (X,Y) has a GBLSD, with parameters $(\beta_1, \beta_2, \theta_1, \theta_2)$, which we shall write $(X,Y) \sim$ GBLSD, if its probability distribution function is given by

$$P(X=x, Y=y) = \frac{\Gamma(\beta_1 x+\beta_2 y)(-\ln(1-\theta_1-\theta_2))^{-1}}{x!y!\Gamma(\beta_1 x+\beta_2 y-x-y+1)} \theta_1^x \theta_2^y (1-\theta_1-\theta_2)^{\beta_1 x+\beta_2 y-x-y}$$

$$x, y \geq 1(1)$$

$$0 < \theta_1, \theta_2 < 1$$

(4.13.a) $$0 < \theta_1\beta_1+\theta_2\beta_2 < 1$$

Also, the GBLSD can be obtained by expanding the bivariate probability generating function $\phi(t_1,t_2) = \ln(1-\theta_1 t_1-\theta_2 t_2)/\ln(1-\theta_1-\theta_2)$, by using the bivariate Lagrange expansion formula, under the transformations

$$t_1 = u\cdot l(t_1,t_2), \qquad t_2 = v\cdot k(t_1,t_2) \quad ,$$

in powers of u and v, where

$$l(t_1,t_2) = \left(\frac{1-\theta_1 t_1-\theta_2 t_2}{1-\theta_1-\theta_2}\right)^{-\beta_1+1}$$

$$k(t_1,t_2) = \left(\frac{1-\theta_1 t_1-\theta_2 t_2}{1-\theta_1-\theta_2}\right)^{-\beta_2+1}$$

are two probability generating functions.

It is clear that the GBLSD is a BMPSD, where

$$g = \theta_1(1-\theta_1-\theta_2)^{\beta_1-1}, \qquad h = \theta_2(1-\theta_1-\theta_2)^{\beta_2-1},$$

(4.13.b) $\qquad$ and $\quad f = -\ln(1-\theta_1-\theta_2)$

4.5.4 The Bivariate Borel-Tanner Distribution (BBTD).

The bivariate random vector (X,Y) has a BBTD (see Shenton and Consul (74)) with parameters (n,m,θ_1,θ_2), which we shall write $(X,Y) \sim$ BBTD, if its probability distribution function is given by

$$P(X=x,Y=y) = \frac{(m+n)(x+y)^{x+y-m-n-1}}{(x-m)!\ (y-n)!}\ \theta_1^{x-m\ y-n}\ \exp\{-(\theta_1+\theta_2)(x+y)\}$$

$x \geq m(1)$

$y \geq n(1)$

$\theta_1,\theta_2 > 0,$

(4.14) $\qquad$ m, n are positive integers.

The BBTD can be obtained by expanding the bivariate probability generating function $\phi(t_1,t_2) = t_1^m t_2^n$, on using the bivariate Lagrange expansion formula (64), under the transformations

$$t_1 = u\cdot l(t_1,t_2), \qquad t_2 = v\cdot k(t_1,t_2),$$

in powers of u and v, where

$$k(t_1,t_2) = l(t_1,t_2) = \exp\{\theta_1(t_1-1)+\theta_2(t_2-1)\} \quad .$$

The BBTD is also a BMPSD with,

$$g = \theta_1 e^{-\theta_1}, \qquad h = \theta_2 e^{-\theta_2}, \qquad \text{and} \qquad f = \theta_1^m \theta_2^n \ .$$

4.5.5 The Bivariate Binomial Delta Distribution (BBDD).

The bivariate random vector (X,Y) is said to have a BBDD, with parameters (n,m,θ_1,θ_2), which we shall write $(X,Y) \sim$ BBDD, if its probability distribution function is given by

$$P(X=x,Y=y) = \frac{(m\beta_1+n\beta_2)\Gamma(\beta_1 x+\beta_2 y)\theta_1^{x-m}\theta_2^{y-n}}{(x-m)!(y-n)!\Gamma(\beta_1 x+\beta_2 y-x-y+m+n+1)} (1-\theta_1-\theta_2)^{\beta_1 x+\beta_2 y+m+n-x-y}$$

$$x \geq m(1)$$

$$y \geq n(1)$$

$$0 < \theta_1,\theta_2 < 1$$

(4.15.a) $$0 < \beta_1\theta_1+\beta_2\theta_2 < 1 \quad .$$

This distribution can be generated by expanding the bivariate probability generating function $\phi(t_1,t_2) = t_1^m t_2^n$, by using the same bivariate Lagrange expansion formula, under the transformations

$$t_1 = u \cdot l(t_1,t_2), \qquad t_2 = v \cdot k(t_1,t_2) \quad ,$$

in powers of u and v, where

$$l(t_1,t_2) = (1-\theta_1-\theta_2+\theta_1 t_1+\theta_2 t_2)^{\beta_1}$$

$$k(t_1,t_2) = (1-\theta_1-\theta_2+\theta_1 t_1+\theta_2 t_2)^{\beta_2} \ .$$

Clearly, the BBDD is a BMPSD with,

$$g = \theta_1(1-\theta_1-\theta_2)^{\beta_1-1}, \qquad h = \theta_2(1-\theta_1-\theta_2)^{\beta_2-1},$$

$$\text{and} \quad f = \theta_1^m \theta_2^n (1-\theta_1-\theta_2)^{-(m+n)} \tag{4.15.b}$$

4.6 Properties and Applications of Some BMPSD Families.

In this section we shall discuss some of the properties, and applications of the GBNBD family and the GDPD family, as they possess many interesting properties, and they have a wide variety of applications.

4.6.1 Some Properties and Applications of the GBNBD.

The GBNBD was first introduced in 1972 by Mohanty (51) as a queueing model. He considered a queueing process initiated by n_1 customers of type I and n_2 customers of type II. Let the customers of type i ($i = 1,2$) arrive in batches of size r_i with Poisson mean rate α_i. Assume that each customer is served exponentially with mean μ. Then the probability that exactly n_1+r_1x customers of type I and n_2+r_2y customers of type II will be served before the queue first vanishes is given by (4.12.a) with

$$n = n_1+n_2, \qquad \beta_i = 1+r_i, \qquad \text{and} \qquad \theta_i = \frac{r_i}{\mu+\alpha_1+\alpha_2} \qquad (i = 1,2).$$

Since $f_1 = f_2 = n(1-\theta_1-\theta_2)^{-1}$, $g_2 = (1-\beta_1)(1-\theta_1-\theta_2)^{-1}$, $h_1 = (1-\beta_2)(1-\theta_1-\theta_2)^{-1}$ and

$$g_1 = (1-\theta_2-\beta_1\theta_1)/\theta_1(1-\theta_1-\theta_2), \qquad h_2 = (1-\theta_1-\beta_2\theta_2)/\theta_2(1-\theta_1-\theta_2),$$

and

$$\Delta = (1-\beta_1\theta_1-\beta_2\theta_2)/\theta_1\theta_2(1-\theta_1-\theta_2) \quad ,$$

the two means and the variances are given by the formulae (4.4.a), (4.4.b), (4.6.a) and (4.6.b) in the form

$$(4.16) \qquad \mu'_{10} = n\theta_1(1-\beta_1\theta_1-\beta_2\theta_2)^{-1}, \qquad \mu'_{01} = n\theta_2(1-\beta_1\theta_1-\beta_2\theta_2)^{-1},$$

$$(4.17) \qquad \mu_{20} = n\theta_1(1-\theta_2-2\theta_2\beta_2+\beta_2\theta_2^2+\beta_2^2\theta_1\theta_2)/(1-\beta_1\theta_1-\beta_2\theta_2)^3$$

and

$$\mu_{02} = n\theta_2(1-\theta_2-2\beta_1\theta_1+\beta_1^2\theta_1^2+\beta_1^2\theta_1\theta_2)/(1-\beta_1\theta_1-\beta_2\theta_2)^3$$

(where $0 < \beta_1\theta_1+\beta_2\theta_2 < 1$).

Property 4.1. If $\beta_1 = \beta_2 = \beta$, the probability distribution of the random variable $Z = X+Y$ is a GNBD and is given by

$$P(Z=z) = \frac{n\Gamma(n+\beta Z)}{Z!\,\Gamma(n+\beta Z-Z-1)}\,\theta^Z(1-\theta)^{n+\beta Z-Z}, \qquad \text{where } \theta = \theta_1+\theta_2.$$

Property 4.2. As $n \to 0$, the origin truncated GBNBD tends to the GBLSD given as (4.13.a)

Property 4.3. If (X_i,Y_i), $i = 1,2,\ldots,N$ is a random sample of size N taken from the GBNBD, the probability distribution of the sums $Z_1 = X_1+X_2+\ldots+X_N$ and $Z_2 = Y_1+Y_2+\ldots+Y_N$ is given as

$$P(Z_1=z_1,Z_2=z_2) = \frac{nN\Gamma(nN+\beta_1 Z_1+\beta_2 Z_2)}{Z_1!Z_2!\Gamma(nN+\beta_1 Z_1+\beta_2 Z_2-Z_1-Z_2+1)}\,\theta_1^{Z_1}\theta_2^{Z_2}(1-\theta_1-\theta_2)^{nN+\beta_1 Z_1+\beta_2 Z_2-Z_1-Z_2} \tag{4.18}$$

which can be obtained directly by utilizing (4.8) and (4.9).

Property 4.4. The joint probability distribution of the random variables X^*, Y^*, where

$$X^* = m + \sum_{i=1}^{N} X_i \qquad \text{and} \qquad Y^* = n + \sum_{i=1}^{N} Y_i \quad ,$$

so that $N = m+n$, is the BBDD given by (4.15.a) (the proof is straightforward).

Property 4.5. The marginal probability distribution of X is a GNBD with parameters (n,β_1,θ_1) and so is Y with parameters (n,β_2,θ_2), and the conditional distribution of Y for given value of X is

$$P(Y/x) = \frac{\Gamma(n+\beta_1 x+\beta_2 y)\Gamma(n+\beta_1 x-x+1)}{y!\Gamma(n+\beta_1 x)\Gamma(n+\beta_1 x+\beta_2 y-x-y+1)}\;\frac{(\theta_2(1-\theta_1-\theta_2)^{\beta_2-1})^y}{(1-\theta_2)^{-n-\beta_2 x}} \tag{4.19}$$

which is in the form of an MPSD.

Property 4.6. The regression equation of Y for a given x is

$$(4.20)\qquad E(Y/x)\begin{cases} \dfrac{n\theta_2(1-\theta_1-\theta_2)}{(1-\theta_2)(1-\theta_1-\beta_2\theta_2)} & \text{for } x = 0 \;, \\[2ex] \dfrac{n\theta_2(1-\theta_1-\theta_2)}{(1-\theta_2)(1-\theta_1-\beta_2\theta_2)} + \dfrac{x\,(1-\theta_1-\theta_2)}{(1-\theta_2)(1-\theta_1-\beta_2\theta_2)} & \text{for } x > 0 \;. \end{cases}$$

The proof of property (4.5) is similar to the case of the GDPD and will be given in the next section, and the proof of property (4.6) is obtained as the mean of an MPSD (3.3.b).

4.6.2. Some Properties and Applications of the GDPD.

Following Shenton and Consul (74), the GDPD represents the probability distribution of the number of customers of type I and type II served by a single server in busy periods and of no customers being in a queue, when the input is Poisson and the rate of arrival of customers of the ith type (i = 1,2) is $m_1\theta_i$ from channel I and is $m_2\theta_i$ from channel II, where θ_1 and θ_2 are the constant service rates of type I and type II respectively. Moreover, if $(X,Y) \sim$ GBNBD, where θ_i is very small and n, β_i are very large (i = 1,2), so that

$$(4.21)\qquad \begin{aligned} n\theta_i &\to \alpha_i \\ \beta_1\theta_i &\to m_1\alpha_i \\ \beta_2\theta_i &\to m_2\alpha_i \end{aligned}$$

then, the GBNBD with parameters $(n,\beta_1,\beta_2,\theta_1,\theta_2)$ can be approximated by the GDPD with parameters $(m_1,m_2,\alpha_1,\alpha_2)$. For $m_1 = m_2 = 0$ the GDPD is reduced to the double Poisson distribution.

For the GDPD given by (4.11.a), one can show that

$$
(4.22)\qquad \left.\begin{aligned}
\Delta &= (1-\theta_1 m_1-\theta_2 m_2)/\theta_1\theta_2 \\
\mu'_{10} &= \theta_1/(1-\theta_1 m_1-\theta_2 m_2) \\
\mu_{20} &= \frac{\theta_1[(1-\theta_2 m_2)^2+\theta_1\theta_2 m_2^2]}{(1-\theta_1 m_1-\theta_2 m_2)^3}
\end{aligned}\right\} \qquad 0 < \theta_1 m_1+\theta_2 m_2 < 1,
$$

and the values of μ'_{01} and μ_{02} can be written down by symmetry. The coefficient of correlation between X and Y is given by

$$
(4.23)\qquad \rho = \frac{\theta_1\theta_2[m_1(1-\theta_2 m_2)+m_2(1-\theta_1 m_1)]}{[\theta_1\theta_2(1-2\theta_1 m_1+\theta_1^2 m_1^2+\theta_1\theta_2 m_1^2)(1-2\theta_2 m_2+\theta_2^2 m_2^2+\theta_1\theta_2 m_2^2)]^{\frac{1}{2}}}
$$

Property 4.7. If $m_1 = m_2 = m$, the probability distribution of the random variable $Z = X+Y$ is GPD, and is given by

$$
(4.24)\qquad P(Z=z) = \frac{(1+m_1 Z)^{Z-1}}{Z!}\ \frac{(\theta e^{-\theta m_1})^Z}{e^{\theta}}, \qquad \text{where } \theta = \theta_1+\theta_2 .
$$

Property 4.8. If $(X,Y) \sim$ GDPD, then X and Y are stochastically independent if and only if $m_1 = m_2 = 0$.

Proof: If we put $m_1 = m_2 = 0$ in (4.11), each of X and Y will be distributed independently as Poisson. On the other hand, the equality of (4.23) to zero is satisfied only when $m_1 = m_2 = 0$.

Property 4.9. If $(X,Y) \sim$ GDPD, then X and Y cannot be perfectly correlated.

Proof: The equality of ρ^2 to unity will imply that $(1-\theta_1 m_1-\theta_2 m_2)^2 = 0$, which is not true unless $m_1\theta_1+m_2\theta_2=1$, and this is a contradiction to the condition enforced by the strict inequality $0<m_1\theta_1+m_2\theta_2<1$.

Property 4.10. If (X_i,Y_i), $i=1,2,\ldots,N$ is a random sample taken from GDPD family, then the joint probability distribution of the sums $Z_1 = X_1+X_2+\ldots+X_N$ and $Z_2 = Y_1+Y_2+\ldots+Y_N$ is also a GDPD given by

$$P(Z_1=z_1,Z_2=z_2)$$

$$= \frac{N(N+m_1Z_1+m_2Z_2)^{Z_1+Z_2-1}}{Z_1!Z_2!}\theta_1^{Z_1}\theta_2^{Z_2}\exp[-(\theta_1+\theta_2)(N+m_1Z_1+m_2Z_2)] \quad . \tag{4.25}$$

Property 4.11. The joint probability distribution of the random variables V_1^*, V_2^*, where

$$V_1^* = m^* + \sum_{i=1}^{N} X_i, \qquad V_2^* = n^* + \sum_{i=1}^{N} Y_i,$$

so that $n^*+m^*=N$ is the BBTD defined by (4.14).

Property 4.12. The marginal distribution of X is a GPD with parameters (m_1,θ_1) and so is Y with parameters (m_2,θ_2), and the regression equation of Y on X is given by

$$E(Y/x) = \begin{cases} \theta_2/(1-\theta_2 m_2) & \text{for } x=0 \\[2ex] \dfrac{\theta_2}{(1-\theta_2m_2)^2} + \dfrac{\theta_2 m_1}{(1-\theta_2m_2)^2}\,x & \text{for } x>0\,. \end{cases} \tag{4.26}$$

Proof: Since $P(X=x) = \sum_y a(x,y) g^x h^y / f$, then

(4.27) $$f = \sum_x g^x \sum_y a(x,y) h^y \quad .$$

Using the univariate Lagrange expansion, to expand the function $f = e^{\theta_1+\theta_2}$, under the transformation $l(\theta_1,\theta_2) = \theta_1/u$, in powers of u, and $l(\theta_1,\theta_2) = e^{m_1(\theta_1+\theta_2)}$, we get

$$f = f(0,\theta_2) + \sum_{x=1}^{\infty} (\theta_1 e^{-m_1(\theta_1+\theta_2)})^x (x!)^{-1} D_{\theta_1}^{x-1} [l^x(\theta_1,\theta_2) e^{\theta_1+\theta_2}] \Big|_{\theta_1=0} \quad .$$

Thus,

$$e^{\theta_1+\theta_2} = e^{\theta_2} + \sum_{x=1}^{\infty} \frac{1}{x!} (1+m_1 x)^{x-1} e^{\theta_2(1+m_1 x)} \, g^x$$

(4.28) $$1 = \sum_{x=0}^{\infty} \frac{(1+m_1 x)^{x-1}}{x!} e^{\theta_2(1+m_1 x)} \frac{g^x}{f} \quad .$$

Comparing (4.27) with (4.28), and by the uniqueness of the power series exapnsion, we get

$$\sum_{y=0}^{\infty} a(x,y) \frac{h^y}{f} = \frac{(1+m_1 x)^{x-1}}{x!} \frac{e^{\theta_2(1+m\ x)}}{e^{\theta_1+\theta_2}}$$

Therefore

$$P(X=x) = (\theta_1 e^{-m_1(\theta_1+\theta_2)})^x \frac{(1+m_1 x)^{x-1}}{x!} \frac{e^{\theta_2(1+m_1 x)}}{e^{\theta_1+\theta_2}}$$

(4.29) $$= \frac{1}{x!} (1+m_1 x)^{x-1} \frac{(\theta_1 e^{-\theta_1 m_1})^x}{e^{\theta_1}}$$

Therefore

$$(4.30) \qquad P(Y=y/X=x) = \frac{1}{y!}\,\frac{(1+m_1x+m_2y)^{x+y-1}}{(1+m_1x)^{x-1}}\,\frac{(\theta_2 e^{-m_2(\theta_1+\theta_2)})^y}{e^{\theta_2(1+m_1x)}}$$

which is in the form of an MPSD. On using the moment properties of an MPSD, one gets

$$E(Y/x) = \frac{\theta_2(1+m_1x)}{1-\theta_2m_2}, \qquad x = 0,1,2,\ldots$$

which is the required result.

4.6.3. Goodness of Fit of the GDPD.

It seems logical that the GDPD should give a reasonably good fit to some numerical data for which statisticians have suggested various forms. Accordingly, we shall consider the data regarding the number of accidents among 122 experienced shunters during the eleven years period (1937-1947) which is given in table (4.1). The data has been arranged according to the periods 0-6 years and 6-11 years. Arbous and Kerrick (2), fitted the bivariate negative binomial distribution to the data based upon the compound Poisson and the contagious hypothesis. In order to test the goodness of fit of the GDPD we shall use the chi-square test. It is not known that the chi-square test is valid for this type of data or not, and although the validity or the accuracy of that test is an open problem and for lack of anything superior, we shall use the classical chi-square test. Based upon the heuristic method of moments, and assuming that $m_1 = m_2 = m$, to avoid the complexity involved in the calculations, the values of the sample estimators in terms of sample moments are given as

$$(4.31)\qquad \theta_1^* = \frac{\mu_{10}^*}{1 + m^*(\mu_{10}^*+\mu_{01}^*)}$$

$$(4.32)\qquad \theta_2^* = \frac{\mu_{01}^*}{1 + m^*(\mu_{10}^*+\mu_{01}^*)}$$

$$(4.33)\qquad m^{*2}(\mu_{10}^*+\mu_{01}^*)\mu_{10}^*\mu_{01}^* + 2m^*\mu_{10}^*\mu_{01}^* - \mu_{11}^* = 0 \ .$$

From equation (4.33), the admissible moment estimator for m is

$$(4.34)\qquad m^* = \frac{[(\mu_{10}^*\mu_{01}^*)^2 + (\mu_{10}^*+\mu_{01}^*)\mu_{11}^*\mu_{10}^*\mu_{01}^*]^{\frac{1}{2}} - \mu_{10}^*\mu_{01}^*}{(\mu_{10}^*+\mu_{01}^*)\mu_{10}^*\mu_{01}^*} \ .$$

Hence

$$(4.35)\qquad \theta_1^* = \frac{\mu_{10}^{*2}\mu_{01}^*}{[(\mu_{10}^*\mu_{01}^*)^2 + (\mu_{10}^*+\mu_{01}^*)\mu_{11}^*\mu_{10}^*\mu_{01}^*]^{\frac{1}{2}}} \ ,$$

$$(4.36)\qquad \theta_2^* = \frac{\mu_{01}^{*2}\mu_{10}^*}{[(\mu_{10}^*\mu_{01}^*)^2 + (\mu_{10}^*+\mu_{01}^*)\mu_{11}^*\mu_{10}^*\mu_{01}^*]^{\frac{1}{2}}} \ ,$$

where μ_{11} is the covariance between X and Y.

From table (4.1) one can show that $\mu_{01}^* = 1.2705$, $\mu_{10}^* = .9754$, $\mu_{11}^* = .3755$, $\rho^* = .2575$, $m^* = .1320$, $\theta_1^* = .7524$ and $\theta_2^* = .9800$. Table (4.2) provides the expected frequencies, where the adjoining cells are pooled, if necessary, so that the expected frequency for the group of pooled cells is at least five, and at the end of the table we provide a summary of the numerical quantities. By a comparison of the set of expected frequencies with the observed frequencies it is clear that the GDPD gives almost the same frequencies for this particular set of data.

In concluding, it may be worth while to investigate in more details the GDPD in the light of any biological data that may be available and throw some light on the closeness of fit using other methods of estimations, some of which will be discussed in the next chapter.

Table (4.1).

1937-42 (6 years)

1943-47 (5 years)

x \ y	0	1	2	3	4	5	6	TOTAL
0	21	18	8	2	1	0	0	50
1	13	14	10	1	4	1	0	43
2	4	5	4	2	1	0	1	17
3	2	1	3	2	0	1	0	9
4	0	0	1	1	0	0	0	2
5	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0
7	0	1	0	0	0	0	0	1
TOTAL	40	39	26	8	6	2	1	122

Table (4.2).

Observed Frequencies	Expected Frequencies
21.0	21.59
18.0	16.84
8.0	8.28
3.0	4.98
13.0	12.93
14.0	14.95
10.0	8.64
5.0	5.72
5.0	5.61
5.0	6.64
6.0	7.04
5.0	5.71
6.0	5.30
3.0	2.99
Total 122	Total 127.22

χ^2_{10} (α=.05) = 18.3   Calculated χ^2 = 2.06*.

*The value of χ^2, as calculated by Arbous and Kerrich for the distribution of the sum $X+Y$, assuming negative binomial is 2.82.

CHAPTER V

ESTIMATION OF THE PARAMETERS OF THE BIVARIATE MODIFIED POWER SERIES DISTRIBUTION

5.1 Introduction.

In Chapter IV we have defined, in form, the class of BMPSD as a parametric family. The probability distribution given by (4.1) can be written as

$$(5.1) \qquad P(X=x,Y=y) = \exp\left[\sum_{j=1}^{2} \tau_j(\theta_1,\theta_2)K_j(x,y) + C(x,y) + q(\theta_1,\theta_2)\right]$$

where $K_1(X,Y) = X$, $K_2(X,Y) = Y$, $\tau_1(\theta_1,\theta_2) = \log g(\theta_1,\theta_2)$, $\tau_2(\theta_1,\theta_2) = \log h(\theta_1,\theta_2)$, $C(X,Y) = \log a(X,Y)$, and $q(\theta_1,\theta_2) = \log f(\theta_1,\theta_2)$, and thus, the BMPSD belongs to the exponential class. Let us assume the following regularity conditions:

(i) the set ${}_0I_0 = \{(x,y):x \geq 0,\ y \geq 0\}$ does not depend upon the parameters θ_1, θ_2.

(ii) $\tau_1(\theta_1,\theta_2)$, and $\tau_2(\theta_1,\theta_2)$ are nontrivial, continuous functions of θ_1, θ_2 and are independent of each other.

(iii) $K_1(X,Y)$ and $K_2(X,Y)$ are continuous for all $(X,Y) \in {}_0I_0$, and are not linear homogeneous functions of each other.

(iv) $C(X,Y)$ is a continuous function of X and Y.

Koopman (41) has proved that, any probability distribution belonging to the exponential class like (5.1), shall have sufficient statistics for the parameters θ_1, θ_2, and based upon a random sample of size N, the joint sufficient statistics are $\sum_{i=1}^{N} K_j(X_i,Y_i)$, $j = 1,2$. Thus, for the BMPSD, $Z_1 = \sum_{i=1}^{N} X_i$, $Z_2 = \sum_{i=1}^{N} Y_i$ are jointly sufficient statistics for θ_1 and θ_2. The completeness of the sufficient statistics will follow from the following theorem which was proved by Lehmann (47).

Theorem 5.1: *Let X be a random vector with probability distribution*

$$dP_\theta(x) = C(\theta)\ \exp\left[\sum_{j=1}^{k} \theta_j T_j(x)\right] d\mu(x)$$

and let P^T be the family of distributions of $T = (T_1(x), T_2(x), \ldots, T_k(x))$ as θ ranges over the set Ω. Then P^T is complete provided Ω contains a k-dimensional rectangle.

Now, we have assumed that the parameter space of the BMPSD contains a 2-dimensional rectangle, and since $\tau_1(\theta_1,\theta_2)$ and $\tau_2(\theta_1,\theta_2)$ are continuous by hypothesis, so the new parameter space does contain 2-dimensional rectangles (see; Silvey (76)). Thus, it follows from the above theorem that (Z_1,Z_2) is complete.

In this chapter, we shall derive the M.L. estimators for the parameters θ_1 and θ_2, and the biases, variances and covariances of these estimators. We shall also discuss the necessary and sufficient conditions for the existence of a minimum variance unbiased (MVU) estimator for any real valued parametric function $k(\theta_1,\theta_2)$, of the parameters θ_1 and θ_2. Applications of these results will be given to the GDPD and GBNBD families.

5.2 Notations.

For the sake of brevity and convenience, we shall use the following symbols.

(i) The set of positive integers $\{(x,y):x \geq r,\ y \geq s\}$ of a two-dimensional space will be denoted by ${}_rI_s$, where r and s are non-negative integers.

(ii) A subset U_N of the set ${}_0I_0$ is said to be the index set of the function f^N if

$$f^N = \sum b(Z_1,Z_2,N)g^{Z_1}h^{Z_2}, \qquad b(Z_1,Z_2,N) > 0 \qquad \text{for } (Z_1,Z_2) \in U_N \subset {}_0I_0,$$

and $b(Z_1,Z_2,N)$ is defined in (4.8).

Clearly, the range of the BMPSD will be the index set U_1.

(iii) A real valued parametric function $k(\theta_1,\theta_2)$ is said to be MVU estimable if it possesses a MVU estimator based on a random sample of size N. Moreover, if for the sample size N we have

$$k(\theta_1,\theta_2)f^N = \sum C(Z_1,Z_2,N)g^{Z_1}h^{Z_2}, \tag{5.2}$$

where $C(Z_1,Z_2,N) \neq 0$ for $(Z_1,Z_2) \in U_N^* \subset {}_0I_0$, then U_N^* is the index set of the function $k(\theta_1,\theta_2)f^N$.

(iv) $\overline{X} = Z_1/N$ and $\overline{Y} = Z_2/N$ are the sample means, calculated from a random sample of fixed size N.

5.3 Maximum Likelihood Estimation for BMPSD.

The logarithm of the likelihood function L is given by

$$\ln L = \text{constant} + \sum_{i=1}^{N} x_i \ln g + \sum_{i=1}^{N} y_i \ln h - N \ln f \quad .$$

On differentiating partially w.r.t. θ_1 and θ_2 and equating to zero, the M.L. equations become

$$\frac{Z_1}{N} g_1 + \frac{Z_2}{N} h_1 - f_1 = 0 \tag{5.3}$$

$$\frac{Z_1}{N} g_2 + \frac{Z_2}{N} h_2 - f_2 = 0 \quad . \tag{5.4}$$

The solution of these equations for θ_1 and θ_2 is not straightforward because they are involved in the functions g_1, g_2, h_1, h_2, f_1 and f_2. However, one can easily solve them for $\overline{X}$ and $\overline{Y}$ to get

(5.5) $$\overline{X} = (f_1h_2-f_2h_1)/(g_1h_2-g_2h_1)$$

(5.6) $$\overline{Y} = (f_2g_1-f_1g_2)/(g_1h_2-g_2h_1)$$

which are precisely the same as the values of μ'_{10} and μ'_{01} in (4.4.a) and (4.4.b) respectively.

It has been shown by Huzurbazar (33), that under the regularity conditions (*i*-*iv*) given in section (5.1), the likelihood equations

$$\frac{\partial \log L}{\partial \theta_i} \qquad i = 1,2$$

have a unique solution for every sample of any size, and that the solution does make the likelihood function a maximum. Moreover, under the same regularity conditions, and for a sufficiently large sample, the variance-covariance matrix of the M.L. estimators can be approximated by Fisher's information matrix. That is

$$\text{Cov}(\hat{\theta}_1,\hat{\theta}_2) = \left[-E\left[\frac{\partial^2 \log L}{\partial \theta_i \partial \theta_j}\right]\right]^{-1} \qquad i,j = 1,2.$$

It is quite likely that the equations (5.5) and (5.6) may not give an explicit solution for θ_1 and θ_2. In such situations one has to use the iterative method given by Rao (67), to obtain a convergent solution starting with some initial set of values $(\theta_{10},\theta_{20})$.

An improved approximation is obtained by the relation

$$\theta_{1,n+1} = \theta_{1,n} + \delta(\theta_{1,n},\theta_{2,n})$$

(5.7)

$$\theta_{2,n+1} = \theta_{2,n} + \epsilon(\theta_{1,n},\theta_{2,n})$$

where

$$\delta(\theta_{1,n},\theta_{2,n}) = (G\partial_2 F - F\partial_2 G)/N^2 I$$

(5.8)

$$\epsilon(\theta_{1,n},\theta_{2,n}) = (F\partial_1 G - G\partial_1 F)/N^2 I$$

and

$$F = \partial_1(\ln L), \qquad G = \partial_2(\ln L) ,$$

$$|I| = E[(\partial_1 P)^2 P^{-2}]\cdot E[(\partial_2 P)^2 P^{-2}] - \{E(P^{-2}(\partial_1 P)(\partial_2 P))\}^2 ,$$

P being the probability function given in (4.1).

5.3.1. Approximation to the Biases and Covariances of the Maximum Likelihood Estimators.

The variances and the covariances of the M.L. estimators for θ_1 and θ_2 are the elements in the inverse of the matrix whose ijth element is given by

$$P_{ij} = E[P^{-2}(\partial_i P)(\partial_j P)], \qquad i,j = 1,2 .$$

On using the recurrence relation

(5.9) $$\frac{\partial P}{\partial\theta_i} = (Xg_i - Yh_i - f_i)P \qquad i = 1,2$$

one can easily show that

$$\mathrm{Var}(\hat{\theta}_1) = (NI)^{-1}(\mu_{20}g_2^2+\mu_{02}h_2^2+2\rho\sqrt{\mu_{20}\mu_{02}}\,g_2h_2+\xi_2^2) + O(N^{-2})$$

$$\mathrm{Var}(\hat{\theta}_2) = (NI)^{-1}(\mu_{20}g_1^2+\mu_{02}h_1^2+2\rho\sqrt{\mu_{20}\mu_{02}}\,g_1h_1+\xi_1^2) + O(N^{-2})$$

$$\mathrm{Cov}(\hat{\theta}_1,\hat{\theta}_2) = -(NI)^{-1}(\mu_{20}g_1g_2+\mu_{02}h_1h_2+2\rho\sqrt{\mu_{20}\mu_{02}}(g_2h_1+g_1h_2)+\xi_1\xi_2)+O(N^{-1}),$$

where $\xi_i = \mu'_{10}g_i + \mu'_{01}h_i - f_i$ and ρ is the coefficient of correlation between X and Y.

To find the amount of bias $b_1(\hat{\theta}_1)$ and $b_2(\hat{\theta}_2)$ in the M.L. estimators, on following Haldane (29), Shenton and Wallington (70), we have in the simultaneous estimation, the equations

$$\begin{bmatrix} P_{11} & P_{12} \\ & \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} b_1(\hat{\theta}_1) \\ \\ b_2(\hat{\theta}_2) \end{bmatrix} = \begin{bmatrix} L_1 \\ \\ L_2 \end{bmatrix} \quad ,$$

where

$$L_\alpha = -\frac{1}{2}\sum_{i,j} P_{\alpha,ij}P^{i,j}, \qquad \alpha,i,j = 1,2,$$

and $P^{i,j}$ is the ijth element in the inverse of the matrix P_{ij} and

$$P_{\alpha,ij} = E[P^{-2}(\partial_\alpha P)(\partial^2_{ij}P)] \quad .$$

Utilizing the recurrence relation (5.9), it can be shown that

$$P_{\alpha,ij} = \mu_{30}g_\alpha g_i g_j + \mu_{21}(h_\alpha g_i g_j + g_\alpha(g_i h_j + h_i g_j))$$

$$+ \mu_{12}(g_\alpha h_i h_j + h_\alpha(g_i h_j + h_i g_j)) + \mu_{03}h_\alpha h_i h_j$$

$$+ \mu_{20}(\xi_\alpha g_i g_j + g_\alpha(g_i \xi_j + g_j \xi_i) + g_\alpha g_{ij})$$

$$+ \rho\sqrt{\mu_{20}\mu_{02}}(\xi_\alpha(g_i h_j + h_i g_j) + h_\alpha(g_i \xi_j + g_j \xi_i)$$

$$+ g_\alpha(h_i \xi_j + h_j \xi_i) + g_\alpha h_{ij} + h_\alpha g_{ij})$$

$$+ \mu_{02}(\xi_\alpha h_i h_j + h_\alpha(h_i \xi_j + h_j \xi_i) + h_\alpha h_{ij}) + \xi_\alpha(\xi_i \xi_j + \xi_{ij}) \quad ,$$

where

$$\xi_{ij} = \mu'_{10}g_{ij} + \mu'_{01}h_{ij} - f_{ij} \quad .$$

Though it is very difficult to get the exact moments of the M.L. estimators in the general case of the BMPSD, and the work of getting the first order approximations is also quite awkward, one can obtain the exact biases, variances and covariances for some particular families.

We shall now derive these values for the GDPD family.

5.3.2 M.L. Estimation of the GDPD.

If (X_i, Y_i), $i = 1,2,\ldots,N$ is a random sample taken from the GDPD family defined by (4.11.a), then by (5.5) and (5.6), the M.L. estimators for θ_1 and θ_2 will be given by the relations

$$Z_1/N = \theta_1(1-\theta_1 m_1-\theta_2 m_2)^{-1} \quad ,$$

$$Z_2/N = \theta_2(1-\theta_1 m_1-\theta_2 m_2)^{-1} \quad .$$

Solving for θ_1 and θ_2, the M.L. estimators are given explicitly as

$$\hat{\theta}_1 = Z_1(N+m_1Z_1+m_2Z_2)^{-1}$$

(5.10)

$$\hat{\theta}_2 = Z_2(N+m_1Z_1+m_2Z_2)^{-1} \quad .$$

By property (4.10), the joint probability distribution of Z_1 and Z_2 is given by (4.25). Thus

$$E(\hat{\theta}_1) = \sum \frac{Z_1}{N+m_1Z_1+m_2Z_2} \cdot \frac{N(N+m_1Z_1+m_2Z_2)^{Z_1+Z_2-1}}{Z!Z_2!} \frac{g^{Z_1}h^{Z_2}}{e^{N(\theta_1+\theta_2)}}$$

$$= \frac{Ng}{N+m_1} \sum (N+m_1) \frac{(N+m_1+m_1U+m_2Z_2)^{U+Z_2-1}}{U!Z_2!} \frac{g^U h^{Z_2}}{e^{N(\theta_1+\theta_2)}}$$

$$= \frac{N\theta_1 e^{-m(\theta_1+\theta_2)}}{(N+m_1)e^{N(\theta_1+\theta_2)}} \cdot e^{(N+m_1)(\theta_1+\theta_2)}$$

(5.11.a) $$= \frac{N\theta_1}{N+m_1}$$

and by symmetry one can show that

(5.11.b) $$E(\hat{\theta}_2) = \frac{N\theta_2}{N+m_2} \quad .$$

Thus one can easily derive unbiased estimators for θ_1 and θ_2, which are given respectively as

(5.11.c) $$\hat{\hat{\theta}}_1 = \frac{(N+m_1)Z_1/N}{N+m_1Z_1+m_2Z_2}$$

and

(5.11.d) $$\hat{\hat{\theta}}_2 = \frac{(N+m_2)Z_2/N}{N+m_1Z_1+m_2Z_2} \quad .$$

Consequently, the exact amount of the biases of $\hat{\theta}_1$ and $\hat{\theta}_2$ are $b_1(\hat{\theta}_1) = -m_1\theta_1(N+m_1)^{-1}$, $b_2(\hat{\theta}_2) = -m_2\theta_2(N+m_2)^{-1}$, respectively.

The exact variances and covariances of the M.L. estimators can be obtained very easily, once the following expectations are calculated.

$$E(\hat{\theta}_1\ \hat{\theta}_2) = \sum \frac{N(N+m_1Z_1+m_2Z_2)^{Z_1+Z_2-3}}{(Z_1-1)!(Z_2-1)!} \frac{g^{Z_1}h^{Z_2}}{e^{N(\theta_1+\theta_2)}}$$

$$= \frac{Nghe^{(N+m_1+m_2)(\theta_1+\theta_2)}}{(N+m_1+m_2)e^{N(\theta_1+\theta_2)}} \times$$

$$\times \sum \frac{(N+m_1+m_2)}{U_1!U_2!} (N+m_1+m_2+m_1U_1+m_2U_2)^{U_1+U_2-1} \frac{g^{U_1}h^{U_2}}{e^{(N+m_1+m_2)(\theta_1+\theta_2)}}$$

(5.13) $$= \frac{N\theta_1\theta_2}{N+m_1+m_2} \quad .$$

Since

$$\mathrm{Cov}(\hat{\theta}_1,\hat{\theta}_2) = E(\hat{\theta}_1\hat{\theta}_2) - E(\hat{\theta}_1)E(\hat{\theta}_2)$$

then from (5.11·a), (5.11.b) and (5.13)

(5.14) $$\mathrm{Cov}(\hat{\theta}_1,\hat{\theta}_2) = \frac{N\theta_1\theta_2}{N+m_1+m_2} - \frac{N^2\theta_1\theta_2}{(N+m_1)(N+m_2)} \quad .$$

$$E(\hat{\theta}_1^2) = E\left[\frac{Z_1(Z_1-1)+Z_1}{(N+m_1Z_1+m_2Z_2)^2}\right]$$

$$= \sum \frac{N(N+m_1Z_1+m_2Z_2)^{Z_1+Z_2-2}}{(Z_1-2)!Z_2!(N+m_1Z_1+m_2Z_2)} \frac{g^{Z_1}h^{Z_2}}{e^{N(\theta_1+\theta_2)}}$$

(5.15) $$+ \sum \frac{N(N+m_1Z_1+m_2Z_2)^{Z_1+Z_2-2}}{(Z_1-1)!Z_2!(N+m_1Z_1+m_2Z_2)} \frac{g^{Z_1}h^{Z_2}}{e^{N(\theta_1+\theta_2)}} \quad .$$

The first summation on the right hand side of (5.15) can be easily shown to be equal to $N\theta_1^2(N+2m_1)^{-1}$. Moreover, since

$$\frac{N+m_1}{N+m_1+m_1Z_1+m_2Z_2} = \left[1 - \frac{m_1Z_1}{N+m_1+m_1Z_1+m_2Z_2} - \frac{m_2Z_2}{N+m_1+m_1Z_1+m_2Z_2}\right] \quad ,$$

then

(5.16) $$E\left[\frac{1}{N+m_1+m_1Z_1+m_2Z_2}\right] = \frac{1}{N+m_1} - \frac{m_1\theta_1}{N+2m_1} - \frac{m_2\theta_2}{N+m_1+m_2} \quad .$$

Now, using the identity (5.16), one can show that the second summation on the right hand side of (5.15) is equal to

$$\frac{N\theta_1}{N+m_1}\left\{\frac{1}{N+m_1} - \frac{m_1\theta_1}{N+2m_1} - \frac{m_2\theta_2}{N+m_1+m_2}\right\} \quad .$$

Thus

$$(5.17)\qquad E(\hat{\theta}_1^2) = \frac{N\theta_1^2}{N+2m_1} + \frac{N\theta_1}{N+m_1}\left\{\frac{1}{N+m_1} - \frac{m_1\theta_1}{N+2m_1} - \frac{\theta_2 m_2}{N+m_1+m_2}\right\}$$

and the exact value of the variance of $\hat{\theta}_1$ will be given as

$$(5.18)\qquad \mathrm{Var}(\hat{\theta}_1) = \frac{N\theta_1}{(N+m_1)^2} - \frac{N^2 m_1\theta_1^2}{(N+m_1)^2(N+2m_1)} - \frac{Nm_2\theta_1\theta_2}{(N+m_1)(N+m_1+m_2)}\ .$$

Also, by symmetry, one can show that the exact variance of $\hat{\theta}_2$ is

$$(5.19)\qquad \mathrm{Var}(\hat{\theta}_2) = \frac{N\theta_2}{(N+m_2)^2} - \frac{N^2 m_2\theta_2^2}{(N+m_2)^2(N+2m_2)} - \frac{Nm_1\theta_1\theta_2}{(N+m_2)(N+m_1+m_2)}\ .$$

5.4 Minimum Variance Unbiased Estimation For a BMPSD.

Theorem 5.2: *The parametric functions g and h given in* (4.1) *are not MVU estimable if U_1 is bounded on the right.*

Proof: Let U_1 be bounded on the right, and let us assume, if possible, that an MVU estimator for g exists. This means there exists a function $M(Z_1,Z_2)$, of the complete sufficient statistics Z_1 and Z_2, such that

$$(5.20)\qquad E[M(Z_1,Z_2)] = g$$

or

$$(5.21)\qquad \sum M(Z_1,Z_2)b(Z_1,Z_2,N)g^{Z_1}h^{Z_2}/f^N = g$$

from which

$$\sum M(Z_1,Z_2)b(Z_1,Z_2,N)g^{Z_1}h^{Z_2} = gf^N \quad . \tag{5.22}$$

But from (5.2), relation (5.22) can be written as

$$\sum M(Z_1,Z_2)b(Z_1,Z_2,N)g^{Z_1}h^{Z_2} = \sum b(Z_1,Z_2,N)g^{Z_1+1}h^{Z_2} \tag{5.23}$$

where the summations on both sides of (5.23) are taken over all $(Z_1,Z_2) \in U_N$.

Since $M(Z_1,Z_2)$ and $b(Z_1,Z_2,N)$ are independent of θ_1 and θ_2, the polynomials on the two sides of (5.23) cannot be equal for any function g. Accordingly, the function g cannot have an MVU estimator, and the assumption is not valid.

Similarly, the function h cannot have an MVU estimator if U_1 is bounded.

The following theorem provides the necessary and sufficient conditions for the existence of an MVU estimator of any real valued parametric function $k(\theta_1,\theta_2)$.

Theorem 5.3: *The necessary and sufficient conditions for $k(\theta_1,\theta_2)$ to be MVU estimable on the basis of a random sample of size N taken from the BMPSD are that $k(\theta_1,\theta_2)\cdot f^N$ is analytic at the origin, and that $U_N^* \subseteq U_N$, where U_N^* and U_N are the index sets of the functions $k(\theta_1,\theta_2)\cdot f^N$ and f^N respectively. Also, when $k(\theta_1,\theta_2)$ is MVU estimable, its MVU estimator $l(Z_1,Z_2,N)$ is given by*

$$l(Z_1,Z_2,N) = \begin{cases} \dfrac{C(Z_1,Z_2,N)}{b(Z_1,Z_2,N)} & (Z_1,Z_2) \in U_N^* \\ \\ 0 & \text{otherwise} . \end{cases}$$

Proof: *Condition is necessary.*

Let $k(\theta_1,\theta_2)$ be MVU estimable for some N, i.e. there exists a function $l(Z_1,Z_2,N)$ of the complete and sufficient statistics (Z_1,Z_2) such that

$$E[l(Z_1,Z_2,N)] = k(\theta_1,\theta_2) \quad . \tag{5.24}$$

Thus

$$\sum_{U_N} l(Z_1,Z_2,N)b(Z_1,Z_2,N)g^{Z_1}h^{Z_2} = k(\theta_1,\theta_2)f^N \tag{5.25}$$

and $k(\theta_1,\theta_2)\cdot f^N$ must possess an expansion in powers of g and h, i.e. it must be analytic at the origin. Let

$$k(\theta_1,\theta_2)\cdot f^N = \sum_{U_N^*} C(Z_1,Z_2,N)g^{Z_1}h^{Z_2} \quad . \tag{5.26}$$

Equating (5.26) with the relation (5.25), we have

$$\sum_{U_N} l(Z_1,Z_2,N)b(Z_1,Z_2,N)g^{Z_1}h^{Z_2} = \sum_{U_N^*} C(Z_1,Z_2,N)g^{Z_1}h^{Z_2} \tag{5.27}$$

where $C(Z_1,Z_2,N) \neq 0$ for $(Z_1,Z_2) \in U_N^*$.

Now, for every $(Z_1,Z_2) \in U_N^*$, $b(Z_1,Z_2,N)$ must be > 0, i.e. $(Z_1,Z_2) \in U_N$, which implies that $U_N^* \subseteq U_N$.

To get the expression of the MVU estimator $l(Z_1,Z_2,N)$ of $k(\theta_1,\theta_2)$

we equate the coefficients of $g^{Z_1}h^{Z_2}$ in both sides of (5.27) for all $(Z_1,Z_2) \in U_N$. Thus, the MVU estimator for $k(\theta_1,\theta_2)$ becomes

$$\text{(5.28)} \qquad l(Z_1,Z_2,N) = \begin{cases} \dfrac{C(Z_1,Z_2,N)}{b(Z_1,Z_2,N)} & (Z_1,Z_2) \in U_N^* \\ \\ 0 & \text{otherwise .} \end{cases}$$

Condition is sufficient.

Let $U_N^* \subseteq U_N$ and let $k(\theta_1,\theta_2)\cdot f^N$ be analytic at the origin. Expanding $k(\theta_1,\theta_2)\cdot f^N$ in powers of g and h, we get

$$k(\theta_1,\theta_2)\cdot f^N = \sum_{U_N^*} C(Z_1,Z_2,N)g^{Z_1}h^{Z_2}$$

or

$$k(\theta_1,\theta_2) = \sum_{U_N^*} \frac{C(Z_1,Z_2,N)}{b(Z_1,Z_2,N)}\, b(Z_1,Z_2,N)\, \frac{g^{Z_1}h^{Z_2}}{f^N}$$

$$= \sum_{U_N} l(Z_1,Z_2,N)\cdot P(Z_1=z_1,Z_2=z_2) \quad ,$$

which implies that $l(Z_1,Z_2,N)$ is an unbiased estimator for $k(\theta_1,\theta_2)$. Since $l(Z_1,Z_2,N)$ is a function of the joint complete sufficient statistics (Z_1,Z_2), it must be the MVU estimator for $k(\theta_1,\theta_2)$.

<u>Corollary 5.1</u>. *The parametric function $g^{a_1}h^{a_2}/f^{a_3}$, where a_1, a_2 are any non-negative integers and a_3 is a positive integer, is MVU estimable for all sample sizes $N \geq a_3$ if and only if $U_{N-a_3} \subseteq U_N$, and in that case, the MVU estimator for $g^{a_1}h^{a_2}/f^{a_3}$ is*

$$b(Z_1-a_1,Z_2-a_2,N-a_3)/b(Z_1,Z_2,N) \qquad \text{for all } (Z_1,Z_2) \in U^*_{N-a_3} \quad .$$

Corollary 5.2. *If r_1 and r_2 are non-negative integers, then from Corollary* (5.1), *the MVU estimator for the probability $p(x=r_1,y=r_2)$ is* $a(r_1,r_2)\,\dfrac{b(Z_1-r_1,Z_2-r_2,N-1)}{b(Z_1,Z_2,N)}$.

5.5 Minimum Variance Unbiased Estimation for Some BMPSD.

5.5.1 Minimum Variance Unbiased Estimation for the GDPD.

Let (X_i,Y_i), $i = 1,2,\ldots,N$ be a random sample taken from the GDPD defined by (4.11.a). We have shown that the unbiased estimators for θ_1 and θ_2 derived by the method of M.L. are given by (5.11.c) and (5.11.d) as

$$\hat{\hat{\theta}}_1 = \frac{(N+m_1)Z_1/N}{N+m_1Z_1+m_2Z_2} \tag{5.29}$$

$$\hat{\hat{\theta}}_2 = \frac{(N+m_2)Z_2/N}{N+m_1Z_1+m_2Z_2} \quad . \tag{5.30}$$

In fact, these unbiased estimators are functions of the complete and sufficient statistics (Z_1,Z_2), and, by the uniqueness theorem, they must be the MVU estimators for θ_1 and θ_2 respectively. Moreover, the relations (5.14), (5.18) and (5.19) can be used trivially to derive the variances and the covariances of the MVU estimators $\hat{\hat{\theta}}_1$ and $\hat{\hat{\theta}}_2$.

We shall now use theorem (5.2) to find MVU estimator for the parametric function

$$k(\theta_1,\theta_2) = \theta_1^a\theta_2^b e^{c\theta_1+d\theta_2} \quad ,$$

where a, b are non-negative integers, and c, d are any real numbers. From (5.28), the MVU estimator is given by $l = C(Z_1,Z_2,N)/b(Z_1,Z_2,N)$, where $b(Z_1,Z_2,N)$ is defined by (4.9), and $C(Z_1,Z_2,N)$ is the coefficient of $g^{Z_1}h^{Z_2}$ in the bivariate Lagrange expansion of the function

(5.31) $$\phi = k(\theta_1,\theta_2)\cdot f^N = \theta_1^a\theta_2^b e^{\theta_1(N+c)+\theta_2(N+d)}$$

which is given as

(5.32)
$$\phi = \sum \frac{g^{Z_1}h^{Z_2}}{Z_1!Z_2!}\,\partial_{\theta_1}^{Z_1-1}\partial_{\theta_2}^{Z_2-1}\left[\chi_1^{Z_1}\chi_2^{Z_2}\frac{\partial^2\phi}{\partial\theta_1\partial\theta_2}\right.$$
$$\left.+\ \chi_1^{Z_1}\partial_{\theta_1}(\chi_2^{Z_2})\,\partial_{\theta_2}(\phi) + \chi_2^{Z_2}\partial_{\theta_2}(\chi_1^{Z_1})\,\partial_{\theta_1}(\phi)\right]\Bigg|_{\theta_1=\theta_2=0}$$

where $\chi_1 = \theta_1 g^{-1}$ and $\chi_2 = \theta_2 h^{-1}$

and the functions g, h are the parametric functions of the GDPD and are given in (4.11.b). Thus,

(5.33)
$$l = \frac{(Z_1-1)!(Z_2-1)!(N+c+m_1Z_1+m_2Z_2)^{Z_1-a-1}(N+d+m_1Z_1+m_2Z_2)^{Z_2-b-1}}{N(N+m_1Z_1+m_2Z_2)^{Z_1+Z_2-1}(Z_1-a-1)!(Z_2-b-1)!}\Bigg[(N+c)(N+d)$$
$$+\ m_1Z_1(N+c) + m_2Z_2(N+d) + \frac{a(N+d+m_1Z_1)(N+c+m_1Z_1+m_2Z_2)}{(Z_1-a)}$$
$$+\ \frac{b(N+c+m_2Z_2)(N+d+m_1Z_1+m_2Z_2)}{(Z_2-b)} + \frac{ab(N+c+m_1Z_1+m_2Z_2)(N+d+m_1Z_1+m_2Z_2)}{(Z_1-a)(Z_2-b)}\Bigg] .$$

5.5.2 Minimum Variance Unbiased Estimation for the GBNBD.

Based upon a random sample of size N taken from the GBNBD, whose probability distribution function is given by (4.11.b), and on using equations (5.5) and (5.6), the M.L. estimators for θ_1 and θ_2, are given respectively as

(5.34)
$$\hat{\theta}_1 = Z_1(nN+\beta_1 Z_1+\beta_2 Z_2)^{-1}$$
$$\hat{\theta}_2 = Z_2(nN+\beta_1 Z_1+\beta_2 Z_2)^{-1} \quad .$$

As can be seen, the M.L. estimators are functions of the complete and sufficient statistics (Z_1,Z_2), but they are not unbiased. The amount of biases are unobtainable by direct calculation, and instead we shall derive the MVU estimator for a real valued parametric function $k(\theta_1,\theta_2)$ of the parameters θ_1 and θ_2 of the GBNBD, using theorem (5.2).

Let $k(\theta_1,\theta_2) = \theta_1^{\gamma_1}\theta_2^{\gamma_2}$, where γ_1 and γ_2 are non-negative integers. Again we shall use the bivariate Lagrange expansion formula (5.32) to expand the function $\phi = k(\theta_1,\theta_2)\cdot f^N = \theta_1^{\gamma_1}\theta_2^{\gamma_1}(1-\theta_1-\theta_2)^{-nN}$, under the transformations $\chi_1 = \theta_1 g^{-1}$ and $\chi_2 = \theta_2 h^{-1}$, where the functions g and h are the parametric functions of the GBNBD given in (4.12.b). Utilizing the formula

$$\partial_{\theta_1}^{Z_1-1}\partial_{\theta_2}^{Z_2-1}[\theta_1^a\theta_2^b(1-\theta_1-\theta_2)^{-c}]\Big|_{\theta_1=\theta_2=0} = \frac{(Z_1-1)!(Z_2-1)!}{(Z_1-a-1)!(Z_2-b-1)!}\,\frac{(c+Z_1+Z_2-a-b-3)!}{(c-1)!}$$

one can show that the MVU estimator $l(Z_1,Z_2,N)$ for $k(\theta_1,\theta_2) = \theta_1^{\gamma_1}\theta_2^{\gamma_2}$

is given as

$$(5.35)\qquad l(Z_1,Z_2,N) = \begin{cases} \dfrac{Z_1!Z_2!\Gamma(nN+\beta_1 Z_1+\beta_2 Z_2-\gamma_1-\gamma_2)}{nN(Z_1-\gamma_1)!(Z_2-\gamma_2)!\Gamma(nN+\beta_1 Z_1+\beta_2 Z_2)}[nN+\gamma_1(\beta_1-1)+\gamma_2(\beta_2-1)] & Z_1 \geq \gamma_1,\ Z_2 \geq \gamma_2 \\ 0 & \text{otherwise} . \end{cases}$$

As may be realized from the last section of this chapter, our search for an MVU estimator of a parametric function was limited to functions which are polynomials of θ_1 and θ_2, and these are in fact the simplest form of parametric functions. It is worth trying to characterize the class of MVU estimable parametric functions, and different forms of $k(\theta_1,\theta_2)$, rather than the polynomial case, should be considered, although we expect the expressions of the MVU estimators to be very complicated.

APPENDIX I

PROVING THE NEGATIVE DEFINITESS OF THE MATRICES OF SECOND ORDER PARTIAL DERIVATIVES OF THE LOG LIKELIHOOD FUNCTION FOR THE GPD AND THE GNBD

(I-A). We shall prove that for the GPD the matrix

$$\Lambda = \begin{bmatrix} \dfrac{\partial^2 \log L}{\partial \lambda_1^2} & \dfrac{\partial^2 \log L}{\partial \lambda_1 \partial \lambda_2} \\[2ex] \dfrac{\partial^2 \log L}{\partial \lambda_1 \partial \lambda_2} & \dfrac{\partial^2 \log L}{\partial \lambda_2^2} \end{bmatrix}_{\substack{\lambda_1=\hat{\lambda}_1 \\ \lambda_2=\hat{\lambda}_2}}$$

is negative definite.

Since

$$\frac{\partial \log L}{\partial \lambda_1} = -n + \frac{n}{\lambda_1} + \sum_{x=0}^{k} n_x \frac{(x-1)}{\lambda_1+\lambda_2 x} \quad ,$$

$$\frac{\partial^2 \log L}{\partial \lambda_1^2} = -\left[\frac{n}{\lambda_1^2} + \sum_{x=0}^{k} n_x \frac{(x-1)}{(\lambda_1+\lambda_2 x)^2}\right] < 0 \quad ,$$

$$\frac{\partial \log}{\partial \lambda_2} = \sum_{x=0}^{k} n_x \left[-x + \frac{x(x-1)}{\lambda_1+x\lambda_2}\right]$$

$$\frac{\partial^2 \log L}{\partial \lambda_1 \partial \lambda_2} = -\sum_{x=0}^{k} n_x \frac{x(x-1)}{(\lambda_1+\lambda_2 x)^2} < 0 \quad ,$$

and

$$\frac{\partial^2 \log L}{\partial \lambda_2^2} = -\sum_{x=0}^{k} \left[n_x \frac{x^2(x-1)}{(\lambda_1+x\lambda_2)^2}\right] < 0 \quad .$$

Thus

$$|\Lambda| = \frac{n}{\hat{\lambda}_1^2} \sum_{x=0}^{k} n_x \frac{x^2(x-1)}{(\hat{\lambda}_1+\hat{\lambda}_2 x)} + \left[\left(\sum_{x=0}^{k} n_x \frac{(x-1)}{(\hat{\lambda}_1+\hat{\lambda}_2 x)^2}\right)\left(\sum_{y=0}^{k} n_y \frac{y^2(y-1)}{(\hat{\lambda}_1+\hat{\lambda}_2 y)^2}\right)\right]$$

$$- \left(\sum_{x=0}^{k} n_x \frac{x(x-1)}{(\hat{\lambda}_1+\hat{\lambda}_2 x)^2}\right)^2$$

$$|\Lambda| = \frac{n}{\hat{\lambda}_1^2} \sum_{x=0}^{k} n_x \frac{x^2(x-1)}{(\hat{\lambda}_1+\hat{\lambda}_2 x)^2}$$

$$+ \sum_{x \neq y} \left[\frac{n_x n_y (x-1) y^2 (y-1)}{(\hat{\lambda}_1+\hat{\lambda}_2 x)^2 (\hat{\lambda}_1+\hat{\lambda}_2 y)^2} - n_x n_y \frac{x(x-1)y(y-1)}{(\hat{\lambda}_1+\hat{\lambda}_2 x)^2 (\hat{\lambda}_1+\hat{\lambda}_2 y)^2}\right] .$$

By the Cauchy Schwartz inequality, the double fold summation is non-negative. Thus Λ is positive, and hence the matrix of the second order partial derivatives of the log likelihood function is negative definite.

(I.B). In the case of the GNBD, we find

$$\left(\frac{\partial^2 \log L}{\partial \theta^2}\right)^* = -\frac{N\bar{x}}{\hat{\theta}^2} - \frac{\hat{n}N+(\beta-1)N\bar{x}}{(1-\hat{\theta})^2} = -\frac{N\bar{x}(n+\beta\bar{x})}{\hat{\theta}^2(\hat{n}+(\beta-1)\bar{x})} < 0$$

$$\left(\frac{\partial^2 \log L}{\partial n^2}\right)^* = \frac{N}{\hat{n}^2} - \sum_{i=1}^{N} \sum_{j=1}^{x_i-1} \frac{1}{(\hat{n}+\beta x_i - j)^2} < 0$$

$$\left(\frac{\partial^2 \log L}{\partial\theta\partial n}\right)_* = -\frac{N}{1-\hat{\theta}} < 0,$$

where $(\quad)_*$ means $(\quad)_{\substack{\theta=\hat{\theta}\\ n=\hat{n}}}$.

Now

$$|V| = \left[\frac{\partial^2 \log L}{\partial\theta^2} \cdot \frac{\partial^2 \log L}{\partial n^2} - \left(\frac{\partial^2 \log L}{\partial\theta\partial n}\right)^2\right]\Bigg|_{\substack{\theta=\hat{\theta}\\ n=\hat{n}}}$$

$$= \frac{N^2\bar{x}(n+\beta\bar{x})^2}{\hat{n}^2\hat{\theta}^2(\hat{n}+(\beta-1)\bar{x})}\left[\frac{1}{\hat{n}^2} - \frac{\bar{x}}{(\hat{n}+(\beta-1)\bar{x})(\hat{n}+\beta\bar{x})} + \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{x_i-1}\frac{1}{(\hat{n}+\beta x_i-j)^2}\right].$$

One can easily see that $|V| > 0$ if $\bar{x} \leq 1$. We feel that $|V| > 0$ for all values of $\bar{x}$, but are unable to prove it. We hope that someone will be able to establish this result for the uniqueness of the M.L. estimators.

APPENDIX II

LEMMA ON THE ASYMPTOTIC BIASES AND COVARIANCES FOR THE M.L. ESTIMATORS

We shall derive the first order terms in the biases and covariances of the M.L. estimators, when the parent population depends on two unknown parameters, to be estimated. The derivation of these asymptotic moments was given by Haldane (29), and was generalized to the multi-parameter case by Shenton and Bowman (75). The expressions for the biases of the M.L. estimators derived by Haldane are similar to those obtained by Shenton and Wallington (70).

Consider a sample of N observations falling in $m > 3$ classes and drawn from a population in which the probability of the rth class is P_r, where $P_r = P_r(\lambda_1, \lambda_2)$ is a known function depending on two parameters λ_1 and λ_2, to be estimated.

Symbols:

(1) Let

$$h_1 = \hat{\lambda}_1 - \lambda_1 \qquad \text{and} \qquad h_2 = \hat{\lambda}_2 - \lambda_2,$$

where $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are the M.L. estimators for λ_1 and λ_2 respectively. We shall denote the biases of $\hat{\lambda}_1$ and $\hat{\lambda}_2$ by $b_1(\hat{\lambda}_1)$ and $b_2(\hat{\lambda}_2)$ respectively.

Thus

$$b_1(\hat{\lambda}_1) = E(\hat{\lambda}_1 - \lambda_1)$$
$$b_2(\hat{\lambda}_2) = E(\hat{\lambda}_2 - \lambda_2).$$

(2) $$C_{lk} = \frac{\partial^{l+k} P_r}{\partial\lambda_1^l \partial\lambda_2^k} \qquad l,k = 0,1,2,\ldots \quad l+k \neq 0 .$$

We shall denote P_r by a_r.

(3) Let $Z_r = \frac{n_r}{N} - a_r$, where n_r is the frequency in the rth class. It is known that

$$E(Z_r) = 0, \quad E(Z_r^2) = N^{-1}a_r(1-a_r), \quad \text{and} \quad E(Z_rZ_s) = -N^{-1}a_ra_s.$$

Expanding $\hat{P}_r = P_r(\hat{\lambda}_1,\hat{\lambda}_2)$, by using the bivariate Taylor's expansion given in Chapter I, we have

$$\hat{P}_r = a_r + c_{10}h_1 + c_{01}h_2 + \frac{1}{2}c_{20}h_1^2 + c_{11}h_1h_2 + \frac{1}{2}c_{02}h_2^2 + \frac{1}{3!}c_{30}h_1^3 + \cdots .$$

Differentiating P_r partially w.r.t. $\hat{\lambda}_1$ and $\hat{\lambda}_2$ respectively, we have

$$\text{(II-1)} \qquad \frac{\partial\hat{P}_r}{\partial\hat{\theta}_1} = c_{10} + c_{20}h_1 + c_{11}h_2 + \frac{1}{2}c_{30}h_1^2 + \cdots ,$$

$$\text{(II-2)} \qquad \frac{\partial\hat{P}_r}{\partial\hat{\theta}_2} = c_{01} + c_{02}h_2 + c_{11}h_1 + \frac{1}{2}c_{03}h^2 + \cdots .$$

(4) Let

$$P_{ij} = \sum_r \frac{1}{P_r}\left(\frac{\partial P_r}{\partial\lambda_i}\right)\left(\frac{\partial P_r}{\partial\lambda_j}\right)$$

$$(*) \qquad P_{\alpha,ij} = \sum_r \frac{1}{P_r}\left(\frac{\partial P_r}{\partial\lambda_\alpha}\right)\left(\frac{\partial^2 P_r}{\partial\lambda_i\partial\lambda_j}\right)$$

$$P_{ij}^{u,s} = \sum_r \frac{1}{P_r^{u+s-1}}\left(\frac{\partial P_r}{\partial\lambda_i}\right)^u\left(\frac{\partial P_r}{\partial\lambda_j}\right)^s$$

$$\alpha,i,j = 1,2$$

$$u,s = 0,1,2,\ldots$$

so that $P_{ij}^{0,0} = 1$, and $P_{ij}^{1,1} = P_{ij}$. Also, let

$$\tau_{k,ij}^{u,s} = \sum \frac{Z_r}{P_r^{u+s+k}} \left(\frac{\partial^i P_r}{\partial\lambda_1^i}\right)^u \left(\frac{\partial^j P_r}{\partial\lambda_2^j}\right)^s \left(\frac{\partial^{i+j} P_r}{\partial\lambda_1^i \partial\lambda_2^j}\right)^k .$$

$$k = 0,1,2,\ldots .$$

The necessary conditions for maximizing $n_r \log P_r$ are

$$\text{(II-3)} \qquad \sum_r n_r \frac{1}{\hat{P}_r} \frac{\partial \hat{P}_r}{\partial \hat{\lambda}_1} = 0$$

$$\sum_r n_r \frac{1}{\hat{P}_r} \frac{\partial \hat{P}_r}{\partial \hat{\lambda}_2} = 0 .$$

The first of equations (II-3) may be written as

$$\text{(II-4)} \qquad \sum (1+a_r^{-1}Z_r)(c_{10}+c_{20}h_1+c_{11}h_2+\tfrac{1}{2}c_{30}h_1^2+\ldots)(1+c_{10}a_r^{-1}h_1$$

$$+c_{01}a_r^{-1}h_2+\tfrac{1}{2}c_{20}h_1^2a_r^{-1}+c_{11}a_r^{-1}h_1h_2+\tfrac{1}{2}c_{02}a_r^{-1}h_2^2+\ldots)^{-1} = 0 .$$

Since all the terms after the first term in the expression of $\hat{P}_r$ are numerically less than a_r, the denominator in (II-4) can be expanded as follows:

$$(1+c_{10}a_r^{-1}h_1+c_{01}a_r^{-1}h_2+\ldots)^{-1}$$

$$\approx 1-(c_{10}a_r^{-1}h_1+c_{01}a_r^{-1}h_2+\tfrac{1}{2}c_{20}a_r^{-1}h_1^2+c_{11}a_r^{-1}h_1h_2+\tfrac{1}{2}c_{02}a_r^{-1}h_2^2+\ldots)$$

$$+ (c_{10}a_r^{-1}h_1+c_{01}a_r^{-1}h_2+\ldots)^2 \;\ldots .$$

Thus substituting the last expression in equation (II-4) and collecting the coefficients of h_1, h_2, h_1^2, h_1h_2 and h_2^2, we get

$$\begin{aligned}&\sum(1+a_r^{-1}Z_r)[c_{10}-(c_{10}^2a_r^{-1}-c_{20})h_1-(c_{10}c_{01}a_r^{-1}-c_{11})h_2\\ &+(c_{10}^3a_r^{-2}-\tfrac{3}{2}c_{10}c_{20}a_r^{-1}+\tfrac{1}{2}c_{30})h_1^2+(2c_{10}^2c_{01}a_r^{-2}-3c_{10}c_{11}a_r^{-1}-c_{01}c_{20}a_r^{-1}+c_{21})h_1h_2\\ &+(c_{10}c_{01}^2a_r^{-2}-\tfrac{1}{2}c_{10}c_{20}a_r^{-1}-c_{01}c_{11}a_r^{-1}+\tfrac{1}{2}c_{12})h_2^2+\dots]=0\ .\end{aligned}\tag{II.5.a}$$

By symmetry, one can write the second of equations (II-3) as

$$\begin{aligned}&\sum(1+a_r^{-1}Z_r)[c_{01}-(c_{01}^2a_r^{-1}-c_{02})h_2-(c_{01}c_{10}a_r^{-1}-c_{11})h_1\\ &+(c_{01}^3a_r^{-2}-\tfrac{3}{2}c_{01}c_{02}a_r^{-1}+\tfrac{1}{2}c_{03})h_2^2+(2c_{01}^2c_{10}a_r^{-2}-2c_{01}c_{11}a_r^{-1}-c_{10}c_{02}a_r^{-1}+c_{12})h_1h_2\\ &+(c_{01}c_{10}^2a_r^{-2}-\tfrac{1}{2}c_{01}c_{02}a_r^{-1}-c_{10}c_{11}a_r^{-1}+\tfrac{1}{2}c_{21})h_1^2+\dots]=0\ .\end{aligned}\tag{II.5.b}$$

Taking the summation over all classes in equations (II.5.a) and (II.5.b), and neglecting terms of order N^{-2} and of higher orders, we get

$$\begin{aligned}&-\tau^{1,0}_{0,11}+\frac{1}{\sqrt{N}}[P_{11}h_1+P_{12}h_2]+\frac{1}{N}[(\tau^{2,0}_{0,11}-\tau^{1,0}_{0,22})h_1-(2P^{2,1}_{11}-2P_{1,12}-P_{2,11})h_1h_2\\ &+(\tau^{1,1}_{0,11}-\tau^{0,0}_{1,11})h_2+(P^{3,0}_{11}-\frac{3}{2}P_{1,11})h_1^2-(P^{1,2}_{11}-\frac{1}{2}P_{2,22}-P_{2,12})h_2^2]\\ &+O(N^{-3/2})=0\end{aligned}\tag{II.6.a}$$

and

$$\begin{aligned}&-\tau^{0,1}_{0,11}+\frac{1}{\sqrt{N}}(P_{21}h_1+P_{22}h_2)+\frac{1}{N}[(\tau^{1,1}_{0,11}-\tau^{0,0}_{1,11})h_1+(\tau^{0,2}_{0,11}-\tau^{0,1}_{0,22})h_2\\ &-(P^{0,3}_{11}-\frac{3}{2}P_{2,22})h_2^2-(2P^{1,2}_{11}-2P_{2,12}-P_{1,22})h_1h_2\\ &-(P^{2,1}_{11}-\frac{1}{2}P_{1,11}-P_{1,12})h_1^2]+O(N^{-3/2})=0\quad ,\end{aligned}$$

respectively.

Thus as a first approximation we have

$$P_{11}h_1 + P_{12}h_2 - \tau^{1,0}_{0,11} + O(N^{-1}) = 0$$

$$P_{21}h_1 + P_{22}h_2 - \tau^{0,1}_{0,11} + O(N^{-1}) = 0$$

or

$$h_1 = \frac{\tau^{1,0}_{0,11}P_{22} - \tau^{0,1}_{0,11}P_{12}}{|I|} + O(N^{-1})$$

$$h_2 = \frac{\tau^{0,1}_{0,11}P_{11} - \tau^{1,0}_{0,11}P_{12}}{|I|} + O(N^{-1})$$

where $|I| = P_{11}P_{22} - P_{12}^2$.

Since

$$\tau^{1,0}_{0,11} = \sum \frac{\partial P_r}{\partial \lambda_1} a_r^{-1} Z_r \ ,$$

$$E(\tau^{1,0}_{0,11}) = N^{-1} \sum \frac{\partial P_r}{\partial \lambda_1} a_r^{-1} E(Z_r) = 0 \quad ,$$

and similarly

$$E(\tau^{0,1}_{0,11}) = 0 \quad .$$

Thus

$$E(h_1) = O + O(N^{-1})$$

and

$$E(h_2) = O + O(N^{-1}) \quad .$$

Moreover

$$E(h_1^2) = |I|^{-1}[P_{22}^2 E(\tau^{1,0}_{0,11})^2 - 2P_{12}P_{22}E(\tau^{1,0}_{0,11}\ \tau^{0,1}_{0,11})$$

$$+ P_{12}^2 E(\tau^{0,1}_{0,11})^2] + O(N^{-2}) \quad .$$

But

$$(\tau^{1,0}_{0,11})^2 = \left(\sum \frac{\partial P_r}{\partial \lambda_1} a_r^{-1} Z_r\right)^2$$

$$= \sum_r \left(\frac{\partial P_r}{\partial \lambda_1}\right)^2 a_r^{-2} Z_r^2 + 2 \sum_{r \neq s} \frac{\partial P_r}{\partial \lambda_1} \frac{\partial P_s}{\partial \lambda_2} a_r^{-1} a_s^{-1} Z_r Z_s$$

$$E(\tau^{1,0}_{0,11})^2 = N^{-1}\left[\sum \left(\frac{\partial P_r}{\partial \lambda_1}\right)^2 a_r^{-2} a_r(1-a_r) - 2 \sum \frac{\partial P_r}{\partial \lambda_1} \frac{\partial P_s}{\partial \lambda_2}\right]$$

$$= N^{-1}\left[\sum \left(\frac{\partial P_r}{\partial \lambda_1}\right)^2 a_r^{-1} - \left(\sum \frac{\partial P_r}{\partial \lambda_1}\right)^2\right]$$

$$= N^{-1} P_{11}.$$

And similarly, one can show that

$$E(\tau^{0,1}_{0,11})^2 = N^{-1} P_{22} \quad ,$$

$$E(\tau^{0,1}_{0,11} \tau^{1,0}_{0,11}) = N^{-1} P_{12} \quad .$$

Finally

$$E(h_1^2) = \frac{P_{22}}{N|I|} + O(N^{-2}) \quad ,$$

and one can easily show that

$$E(h_2^2) = \frac{P_{11}}{N|I|} + O(N^{-2}) \tag{II-7}$$

$$E(h_1 h_2) = \frac{-P_{12}}{N|I|} + O(N^{-2}) \quad .$$

The expressions in (II-7) are well known and usually obtained as the elements in the inverse matrix of Fisher's determinant of information.

To find $E(h_1)$ and $E(h_2)$ to order N^{-1} we shall use the approximations given in (II-7). Now, (II-6-a) may be written as

$$h_1 = \frac{\tau^{1,0}_{0,11} - P_{12}h_2 + (P^{1,2}_{11} - \tfrac{1}{2}P_{2,22} - P_{2,12})h_2^2 - (\tau^{1,1}_{0,11} - \tau^{0,0}_{1,11})h_2 + O(N^{-3/2})}{P_{11} - (P^{3,0}_{11} - \tfrac{3}{2}P_{1,11})h_1 - (2P^{2,1}_{11} - P_{2,11} - 2P_{1,12})h_2 + \tau^{2,0}_{0,11} - \tau^{1,0}_{0,22} + O(N^{-1/2})} \tag{II-8}$$

$$= \frac{k_1 - k_2}{I_0 - I_1}$$

where I_0 represents a term of order 1, k_1 and I_1 terms of order $N^{-1/2}$, and k_2 a term of order N^{-1}. Hence

$$\begin{aligned} E(h_1) = P^{-2}_{11}[&(P^{3,0}_{11} - \tfrac{3}{2}P_{1,11})\{E(h_1\tau^{1,0}_{0,11}) - P_{12}E(h_1h_2)\} \\ &+ (2P^{2,1}_{11} - P_{2,11} - 2P_{1,12})\{E(h_2\tau^{1,0}_{0,11}) - P_{12}E(h_2^2)\} - E(\tau^{2,0}_{0,11}\tau^{1,0}_{0,11}) \\ &+ E(\tau^{1,0}_{0,11}\tau^{1,0}_{0,22}) - P_{12}P_{11}E(h_2) + P_{12}\{E(h_2\tau^{2,0}_{0,11}) - E(h_2\tau^{1,0}_{0,22})\} \\ &+ P_{11}(P^{1,2}_{11} - \tfrac{1}{2}P_{2,22} - P_{2,12})E(h_2^2) - P_{11}E(h_2\tau^{1,1}_{0,11}) + P_{11}E(h_2\tau^{0,0}_{1,11})] + O(N^{-2}) \end{aligned} \tag{II-9}$$

Now the following expectations can be obtained after some manipulation for the definitions given in (4).

$$E(h_1\tau^{1,0}_{0,11}) = 1$$

$$E(h_2\tau^{1,0}_{0,11}) = 0$$

$$E(\tau^{2,0}_{0,11}\tau^{1,0}_{0,11}) = N^{-1}P^{3,0}_{11}$$

$$E(\tau^{1,0}_{0,11}\tau^{1,0}_{0,22}) = N^{-1}P_{1,11}$$

$$E(h_2\tau^{2,0}_{0,11}) = N^{-1}|I|^{-1}(P_{11}P^{2,1}_{11} - P_{12}P^{3,0}_{11})$$

$$E(h_2\tau^{1,0}_{0,22}) = N^{-1}|I|^{-1}(P_{11}P_{2,11}-P_{12}P_{1,11})$$

$$E(h_2\tau^{1,1}_{0,11}) = N^{-1}|I|^{-1}(P_{11}P^{1,2}_{11}-P_{12}P^{2,1}_{11})$$

$$E(h_2\tau^{0,0}_{1,11}) = N^{-1}|I|^{-1}(P_{11}P_{2,12}-P_{12}P_{1,12}) \quad .$$

Substituting the above expressions in (II-9) we get

$$\begin{aligned}P_{11}E(h_1)+P_{12}E(h_2) &= \frac{P^{-1}_{11}}{N|I|}[(P^{3,0}_{11}-\tfrac{3}{2}P_{1,11})(P_{11}P_{22}-P^2_{12}+P^2_{12}) \\ &+ (2P^{2,1}_{11}-P_{2,11}-2P_{1,12})(P_{11}P_{12}-P_{11}P_{12}-P_{11}P_{12}) \\ &- P^{3,0}_{11}(P_{11}P_{22}-P^2_{12})+P_{1,11}(P_{11}P_{22}-P^2_{12}) \\ &+ P_{12}(P_{11}P^{2,1}_{11}-P_{12}P^{3,0}_{11}-P_{11}P_{2,11}+P_{12}P_{1,11}) \\ &+ P^2_{11}(P^{1,2}_{11}-\tfrac{1}{2}P_{2,22}-P_{2,12})-P_{11}(P_{11}P^{1,2}_{11}-P_{12}P^{2,1}_{11}) \\ &+ P_{11}(P_{11}P_{2,12}-P_{12}P_{1,12})] + O(N^{-2})\end{aligned}$$

or $\quad P_{11}E(h_1) + P_{12}E(h_2)$

(II-10)
$$= \frac{-1}{2N|I|}[P_{11}P_{1,22}-2P_{12}P_{1,12}+P_{22}P_{1,11}] \quad .$$

Similarly, on using equation (II-6-b), we get the following expression for h_2.

$$h_2 = \frac{\tau^{0,1}_{0,11}-P_{12}h_1+(P^{2,1}_{11}-\tfrac{1}{2}P_{1,11}-P_{1,12})h^2_1-(\tau^{1,1}_{0,11}-\tau^{0,0}_{1,11})+O(N^{-3/2})}{P_{22}-(P^{0,3}_{11}-\tfrac{3}{2}P_{2,22})h_2-(2P^{1,2}_{11}-P_{1,22}-2P_{2,12})+\tau^{0,2}_{0,11}-\tau^{0,1}_{0,22}+O(N^{-1})} \quad .$$

From the last equation one can show that

(II-11) $$P_{12}E(h_1)+P_{22}E(h_2) = \frac{-1}{2N|I|}[P_{11}P_{2,22}-2P_{12}P_{2,12}+P_{22}P_{2,11}] \quad .$$

Solving equations (II-10) and (II-11) simultaneously for $E(h_1)$ and $E(h_2)$ we have

$$E(h_1) = b_1(\hat{\lambda}_1) = \frac{1}{2N|I|^2}\,[-P_{22}(P_{22}P_{1,11}-2P_{12}P_{1,12}+P_{11}P_{1,22})$$

$$+ P_{12}(P_{22}P_{2,11}-2P_{12}P_{2,12}+P_{11}P_{2,22})] + O(N^{-2})$$

$$E(h_2) = b_2(\hat{\lambda}_2) = \frac{1}{2N|I|^2}\,[P_{12}(P_{22}P_{1,11}-2P_{12}P_{1,12}+P_{11}P_{1,22})$$

$$- P_{11}(P_{22}P_{2,11}-2P_{12}P_{2,12}+P_{11}P_{2,22})] + O(N^{-2})$$

which are the biases of the M.L. estimators obtained by Shenton and Wallington (70).

REFERENCES

1. Anscombe, F.J. (1950). "Sampling theory of the negative binomial and logarithmic series distributions", Biometrika, 37, 358.

2. Arbous, A.G. and Kerrich, J.E. (1951). "Accident statistics and the concept of accident proness", Biometrics, 7, 340-432.

3. Bhat, B.R. and Kulkarni, N.V. (1966). "On efficient multinomial estimation", Jour. Roy. Stat. Soc., Sr. B. 28, 45.

4. Binet, F. (1953). "The fitting of the positive binomial distribution when both parameters are estimated from the sample", Annals of Eugenics, London, 18, 117-119.

5. Bliss, C.I. (1953). "Fitting the negative binomial distribution to biological data", Biometrics, 9, 176.

6. Bowman, K.O. and Shenton, L.R. (1965). "Asymptotic Covariances for the maximum likelihood estimators of the parameters of a negative binomial distribution", Report K-1643, Union Carbide Nuclear Division, Oak Ridge, Tenn.

7. Charalambides, Ch. (1974). "Minimum variance unbiassed estimation for a class of left-truncated discrete distributions", Sankhya, Vol. 36, Ser. A. Pt. 4, 397-418.

8. Charalambides, Ch. (1974). "The distribution of sufficient statistics of truncated generalized logarithmic series, Poisson and negative binomial distributions", The Canadian J. of Statistics, Vol. 2, No. 2, 261-275.

9. Cohen, A.C. (1960). "Estimation in truncated Poisson distribution when zeros and some ones are missing", Journal of the American Statistical Association, 55, 342-348.

10. Consul, P.C. and Jain, G.C. (1973). "A generalization of Poisson distribution", Technometrics, 15, (4), 791-799,

11. Consul, P.C. and Jain, G.C. (1973). "On some interesting properties of the generalized Poisson distribution", Biom. z, Bd. 15, H.Y.S., 495-500.

12. Consul, P.C. and Shenton, L.R. (1972). "Use of Lagrangian expansion for generating discrete generalized probability distributions", SIAM J. Appl. Math., Vol. 23, No. 2, 239-248.

13. Consul, P.C. and Shenton, L.R. (1973). "Some interesting properties of Lagrangian distributions", Communications in Statistics 2 (3) 263-272.

14. Consul, P.C. and Shenton, L.R. (1973). "On the multivariate generalization of the family of discrete Lagrange distributions", Multivariate Statistical Inference. Proceedings of the research seminar at Dalhousie University, Halifax, March 1972.

15. Consul, P.C. and Gupta, H.C. "The generalized negative binomial distribution and its characterization by zero-regression", SIAM, J. Appl. Math. (To appear).

16. Cramér, H. (1946). "Mathematical Methods of Statistics", Princeton University Press.

17. Goursat, E. (1916). "A course of mathematical analysis", II. Pt. I, (Functions of a complex variable), Ginn Company.

18. Govindarajulu, A. (1970). "Elementary characterization of discrete distributions", Random Counts in Models and Structures, Penn. State Univ. Press, 77-96.

19. Guldberg, A. (1931). "On discontinuous frequency functions and statistical series", Skandinavisk Aktuarietidskrift, 14, 167-187.

20. Guldberg, A. (1934). "On discontinuous frequency functions of two variables", Skandinavisk Aktuariedskrift, 17, No. 1, 89-117.

21. Gupta, R.C., "Modified power series distributions and some of its applications", Sankhya, Ser. B, Vol. 36, Pt. 3, 288-298.

22. Gupta, R.C., "Maximaum likelihood estimation of a modified power series distribution and some of its applications", Communications in Statistics, 4 (7), 689-697 (1975).

23. Gurland, J. (1958). "A generalized class of contagious distributions", Biometrics, 14, 229-249.

24. Guttman, I. (1958). "A note on a series solution of a problem in estimatinn", Biometrika 45, 565-567.

25. Haight, F.A. (1961). "Index to the distribution of mathematical statistics", Journal of Research of the National Bureau of Standards, 65, Sec. B, 23-60.

26. Haight, F.A. and Breuer, M.A. (1960). "The Borel-Tanner distribution", Biometrika 47, 143-150.

27. Haight, F.A. (1961). "A distribution analogous to the Borel-Tanner", Biometrika, 48, 167-173.

28. Haight, A. (1967). "Hand book of Poisson distribution", John Wiley & Sons, New York.

29. Haldane, J.B.S. (1953). "The estimation of two parameters from a sample", Sankhya, 12, 313-20.

30. Haldane, J.B.S. and Smith, S.M. (1956). "The sampline distribution of a maximum likelihood estimate", Biometrika, 43, 96-103.

31. Harvey, J.R. (1972). "On expressing moments in terms of cumulants and vice versa". The American Statistician, Vol. 26, 38-39.

32. Henrici, P. (1965). "Elements of numerical analysis", John Wiley.

33. Huzurbazar, V.S. (1949). "On property of distributions admitting sufficient statistics", Biometrika, 36, 71.

34. Jain, G.C. and Consul, P.C., "A generalized negative binomial distribution", SIAM J. Appl. Math., Vol. 21, No. 4, (1971), 501-513.

35. Johnson, N.L. and Kotz, S. (1969). "Discrete distributions", New York, John Wiley and Sons, Inc.

36. Joshi, S.W. and Park, C.J. (1974). "Minimum variance unbiased estimation for truncated power series distributions", Sankhya, A, 36, 305-314.

37. Katti, S.K. and Gorland, J. (1962). "Efficiency of certain methods of estimation for the negative binomial and the Neyman Type A distributions", Biometrika, 49, 215-226.

38. Katz, L. (1963). "Unified treatment of a broad class of discrete probability distributions", Proceedings of the International Symposium on Discrete Distributions, Montreal, 1975, 182.

39. Kendall, M.G. and Stuart, A. (1963). "The advanced theory of statistics", Vol. I, Charles Griffin & Co., London.

40. Khatri, C.G. (1959). "On certain properties of power-series distributions", Biometrika, 46, 486-90.

41. Koopman, B.O. (1936). "On distributions admitting a sufficient statistic", Trans. Amer. Math. Soc. 39, 399.

42. Kosambi, D.D. (1949). "Characteristic properties of series distributions", Proc. National Inst. Sci. India, 15, 109-113.

43. Krishnamoorthy, A.S. (1951). "Multivariate binomial and Poisson distributions", Sankhya, 11, 117.

44. Kumar, A. and Consul, P.C. (1979). "Negative moments of a modified power series distribution and bias of the maximum likelihood estimator", Communications in Statistics, A8 (2), 151-166.

45. Kumar, A. (1979). "Some estimation problems in power series distributions", Ph.D. Thesis, University of Calgary, Calgary, Alberta, Canada.

46. Lehmann, E.L. and Scheffé, H. (1950). "Completeness, similar regions and unbiased estimation", Pt. 1, Sankhya, 10, 305-340.

47. Lehmann, E.L. (1959). "Testing statistical Hypothesis", Wiley, New York.

48. Lundberg, O. (1940). "On random processes and their application to sickness and accident statistics", Uppsala: Alqvist and Wiksell (10.4).

49. Mardia, K.V. (1970). "Families of bivariate distributions", London Charles Griffin and Co.

50. Mohanty, S.G. (1966). "On a generalized two-coin tossing problem", Biometrische Zeitschrift, 8 (4), pp. 266-272.

51. Mohanty, S.G. (1972). "On queues involving patches", J. Appl. Prob. 9, 430-435.

52. Nelson, D.L. (1975). "Some remarks on generalizations of the negative binomial and Poisson distributions", Technometrics, Vol. 17, No. 1, 135-136.

53. Neyman, J. and Bates, G.C. (1952). "contributions to the theory of accident proness", University of California, Publications in Statistics, 1, 215.

54. Noack, A. (1950). "A class of random variables with discrete distributions", Ann. Amth. Stat., 21, 127-132.

55. Olkin, I. and Sobel, M. (1965). "Integral expressions for tail probabilities of the multinomial and negative multinomial distributions, Biometrika, 52, 167.

56. Ord, J.K. (1967). "Graphical method for a class of discrete distributions", J. Roy. Stat. Soc., Ser. A., 130, 232-238.

57. Ord, J.K. (1972). "Families of frequency distributions", Charles Griffin & Co., Ltd.

58. Park, C.J. (1973). "The power series distributions with unknown truncation parameter", Ann. Math. Stat., Vol. 1, No. 2, 395-399.

59. Patil, G.P. (1962). "Certain properties of the generalized power series distribution", Ann. Inst. Stat. Math., Tokyo, 14, 1979-182.

60. Patil, G.P. (1962). "Maximum likelihood estimation for GPSD and its application to a truncated binomial distribution", Biometrika, 49, 1 and 2, p. 227.

61. Patil, G.P. and Joshi, S.W. (1968). "A dictionary and bibliography of discrete distributions", Oliver and Boyd, Edinburgh.

62. Patil, G.P. (1968). "On sampling with replacement from populations with multiple characters", Sankhya Ser. B, 355-66.

63. Pearson, K. (1895). "Contributions to the mathematical theory of evolution II", Skew variation in homogeneous material. Philosophical transactions of the Royal Society of London, Ser. A, 186, 343-414.

64. Poincaré, H. (1886). "Sur les résidues des integrales doubles", Acta Math., 9, 321-80.

65. Raja Rao B. (1958). "On the relative efficiencies of BAN estimates based on doubly truncated and censored samples", Proceedings of the National Institute of Sciences of India 24, 366-76.

66. Rao, C.R. (1948). "Sufficient statistics and minimum variance estimates", Proc. Camb. Phil. Soc. 45, 215-218.

67. Rao, C.R. (1952). "Advanced statistical methods in biometric research", John Wiley & Sons, Inc., New York.

68. Roy, J. and Mitra, S. (1957). "Unbiassed minimum variance estimation in a class of discrete distributions", Sankhya, 18, 371-378.

69. Shenton, L.R. (1949). "On the efficiency of the method of moments and Neyman's type A distribution", Biometrika, 36, 450-454.

70. Shenton, L.R. and Wallington, P.A. (1962). "The bias of moment estimators with an application to the negative binomial distribution", Biometrika, 49, 193-204.

71. Shenton, L.R. and Myers, R. (1963). "Orthogonal statistics", Proceedings of the International Symposium on Discrete Distributions, Montreal, 445-458.

72. Shenton, L.R. and Bowman, K.O. (1963). "Higher moments of a maximum likelihood estimator", J. Roy. Statist. Soc. (B), 25, 305-17.

73. Shenton, L.R. and Bowman, K.O. (1967). "Remarks on large sample estimators for some discrete distributions", Technometrics, 9, 4, 587-598.

74. Shenton, L.R. and Consul, P.C., (1973). "On bivariate Lagrange and Borel-Tanner distributions and their use in queueing theory", Sankhya, Ser. A, Vol. 35, Pt. 2, 229-236.

75. Shenton, I.R. and Bowman, K.O. (1977). "Maximum likelihood estimation in small samples", Griffin's Statistical Monographs & Courses, No. 38.

76. Silvey, S.D. (1970). "Statistical inference", Library of University Mathematics. Penguin Books Ltd, Harmondsworth.

77. Swamy, P.S. (1962). "On the amount of information supplied by censored samples of grouped observations in the estimation of statistical parameters", Biometrika 49, 245-249.

78. Szegö, G. (1939). "Orthogonal polynomials", Amer. Math. Soc. Colloqium Publ. New York, Vol. 23.

79. Tanner, J.C. (1953). "A problem of interference between two queues", Biometrika, 40, 50-69.

80. Tanner, J.C. (1961). "A derivation of Borel distribution", Biometrika, 40, 222-224.

81. Tweedie, M.C.K. (1947). "Functions of a statistical variate with given means, with special reference to Laplacian distributions", Proc. Comb. Phil. Soc. 43, 41-49.

82. Widder, D.V. (1947). "Advanced Calculus", Prentice-Hall, Inc. New York.