THE UNIVERSITY OF CALGARY

FAMILIES OF DISCRETE PROBABILITY DISTRIBUTIONS ASSOCIATED WITH POWER SERIES EXPANSIONS

by

MOHAMED M. SHOUKRI

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

JUNE 1977

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

"Families of Discrete Distributions Associated
With Power Series Expansions"

submitted by Mohamed M. Shoukri in partial fulfillment of the requirements for the degree of Master of Science.

P. C. Consul

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

J R Collins

J.R. Collins
Department of Mathematics
& Statistics

E. S. Edgington

E.S. Edgington
Department of Psychology

July 20, 1977
Date

ABSTRACT

The study of systems of probability distributions started in 1895 with Karl Pearson's celebrated family of curves, which still plays a major role in modern statistics. Many systems of distributions, both continuous and discrete, have been developed in succeeding years, with a wide range of aims and usage. Applications occur in such diverse fields as economics, medicine and sport, to name but a few. Trying to keep track of such a wide-ranging and rapidly expanding literature is probably a hopeless task. Rather, I have tried to look at major systems specially those associated with series representation.

An introduction to the different types of families of univariate discrete distributions has been given in Chapter I.

A very wide family of discrete distributions with the title "Lagrangian Probability Distributions" has been introduced to statistical literature by Consul and Shenton (1972). The method to generate such families of distributions and their important properties are given in detail in Chapter II. Some important distributions of this class such as the Lagrangian Binomial Distribution, Lagrangian Poisson Distribution, the Borel-Tanner Distribution and Consul Distribution, which are generalizations of some classical distributions, are found to be the distribution of the number of customers served in the first busy period (FBP).

A modified form of the power series distributions by the name "Modified Power Series Distributions", (MPSD), has been exhibitted by Gupta (1974, (1)). It will be shown in Chapter III that the class of MPSD is a sub-class of Lagrangian Probability Distributions. Some

important properties of a MPSD will be studied in detail within that chapter. The distribution of the sum of independent decapitated MPSD will also be derived, because of its usefulness in estimation.

In Chapter IV the problem of estimation for a MPSD is discussed. Gupta (1975) introduced the method of maximum likelihood, considering θ as the only unknown parameter involved in the density function of MPSD. The writer gave a U.M.V.U.E. for a function $g(\theta)$ of the unknown parameter θ.

The majority of the Lagrangian and the MPSD are multiparameters distributions, but to avoid the complicated problems which may arise in the multiparametric situation, θ has been considered to be unknown and all the other parameters to be known constants. Studying the multiparametric situation is beyond this stage of research.

ACKNOWLEDGEMENT

I am truely grateful to Dr. P.C. Consul, who guided my choice to the subject of this thesis. I am deeply indebted to his generous help, continuous guidance, and for his perpetual encouragement.

Also I would like to thank Mrs. T. Hayashitani and Mrs. E. Muench for their patient and careful typing of this project.

TABLE OF CONTENTS

CHAPTER I

STATISTICAL DISTRIBUTIONS: trends and prospects

1.0 Introduction

The statistical distributions arose initially in connection with specific real situations. Once their relevance was established there was little further interest in theoretical analysis as they were used mainly 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) became important. Both of these topics are still attracting considerable interest, with more and more complicated distributions being derived.

In univariate theory, there are relatively few new distributions being studied. Interest is focused more on the classes of distributions, e.g.

(a) Construction of systems embracing many known distributions.

(b) Stable and infinitely divisible classes.

The discrete probability distributions can also be classified on the basis of their common properties regarding; moments, estimation, applications, etc.

The series representations of probability distributions are now becoming popular, but questions of speed convergence appear to be neglected.

Karl Pearson (1895) was the pioneer of the study of the systems of probability distributions, which plays a major role in modern statistics. He obtained that for the hypergeometric distribution, the ratio

$\frac{P(X=x+1) - P(X=x)}{P(X=x+1) + P(X=x)}$ is of the form (linear function of x)/(quadratic function of x). His observation led to the differential equation defining Pearson's system of curves.

Noack (1950) introduced the discrete analogue of the exponential family of curves into the statistical literature by the name of "power series distribution". The generalized power series distribution has been studied in detail by Khatri (1959), Patil (1962), (1963), and many others. Ord (1967) developed a system of discrete distributions, using a difference equation analogous to Pearson's differential equations.

My purpose in writing this chapter was not, therefore, to provide an encyclopaedic work, but to provide a brief discussion of the classical and recent discrete classes of probability distributions, which have grown considerably in importance in recent years.

1.1 Mixtures and early generalizations of discrete distributions

Mixtures of distributions have received an increasing amount of attention in recent statistical literature, due to interest in the mathematical aspects of mixtures as well as a considerable number of applied problems where mixtures are encountered. The term 'compound distribution' has also been used by some authors.

Gurland (1957) defines a 'compound distribution' as follows:

Let the random variable X_1 have the distribution function $F_1(x_1 \mid \theta)$ for a given value of the parameter θ. Suppose now that θ is regarded as a random variable X_2, say, with distribution function $F_2(x_2)$. Denote by $X_1 \wedge X_2$ the random variable with distribution function

$$\int_{-\infty}^{\infty} F_1(x_1 \mid cx_2)\, dF_2(x_2),$$

where c is a constant which is arbitrary or restricted in some prescribed sense. Then the random variable $X_1 \wedge X_2$ (uniquely defined here except for the constant c) is called a compound X_1 variable with respect to the 'Compounder' X_2.

Let X_1 be a random variable with probability generating function (p.g.f.) $g_1(z)$, representing the number of egg-masses laid in a sampling unit. Suppose that each egg-mass gives rise to X_2 larvae with p.g.f. $g_2(z)$. Then the distribution of X_2 averaged over X_1 is called 'X_1 generalized X_2 distribution' and the random variable is denoted by $X_1 \vee X_2$. It can be shown that the p.g.f. of $X_1 \vee X_2$ is given by $g_1\{g_2(z)\}$. Though this illustration is given in terms of egg masses and larvae, it is clear that with suitable terminology it applies as a definition to any situation in which there is an original population and each member of the population gives rise to a new population.

Katti(1970) gave some interrelations among such generalized distributions and their components.

(i) Let $g(z) = g_1(g_2(z))$, and let $\mu_{(i)}$, ${}_1\mu_{(i)}$ and ${}_2\mu_{(i)}$ denote the i-th factorial moments of $g(z)$, $g_1(z)$ and $g_2(z)$ respectively. Note that

$$\mu_{(i)} = \overset{(i)}{g}(z)\Big|_{z=1} \qquad i = 1,2,\dots , \qquad (1\text{-}1),$$

where $\overset{(i)}{g}(z)$ denotes the i-th derivative of $g(z)$ $w.r.$ to z. Let $t = g_2(z)$. On writing,

$$g(z) = g_1(t) \qquad (1\text{-}2)$$

and on differentiating (1-2) successively $w.r.$ to z, we have,

$$\overset{(1)}{g}(z) = \overset{(1)}{g_1}(t)\,\overset{(1)}{g_2}(z)$$

$$\overset{(2)}{g}(z) = \overset{(2)}{g_1}(t)\left\{\overset{(1)}{g_2}(z)\right\}^2 + \overset{(1)}{g_1}(t)\,\overset{(2)}{g_2}(z) \qquad (1\text{-}3)$$

$$\overset{(3)}{g}(z) = \overset{(3)}{g_1}(t)\left\{\overset{(1)}{g_2}(z)\right\}^3 + 3\,\overset{(2)}{g_1}(t)\,\overset{(1)}{g_2}(z)\,\overset{(2)}{g_2}(z) + \overset{(1)}{g_1}(t)\,\overset{(3)}{g_2}(z),$$

and so on. At $z = 1$, $t = g_2(z) = 1$, hence on substituting $z = 1$ in (1-3) we have

$$\mu_{(1)} = {}_1\mu_{(1)} \cdot {}_2\mu_{(1)}$$

$$\mu_{(2)} = {}_1\mu_{(2)}\ {}_2\mu_1^2 + {}_1\mu_{(1)}\ {}_2\mu_{(2)}$$

$$\mu_{(3)} = {}_1\mu_{(3)}\ {}_2\mu_1^3 + 3\ {}_1\mu_{(2)}\ {}_2\mu_{(2)}\ {}_2\mu_{(1)} + {}_1\mu_{(1)}\ {}_2\mu_{(3)},$$

and so on.

(ii) Also the factorial cumulants generating functions $K(u)$ is related to the p.g.f. $g(z)$ by the formula

$$K(u) = \log g(1 + u),$$

and the i-th factorial cumulants $K_{(i)}$ is given by the formula

$$K_{(i)} = \frac{\partial^i K(u)}{\partial u^i}\bigg|_{u=0}$$

and we can get the relations,

$$K_{(1)} = {}_1K_{(1)}\,{}_2\mu_{(1)}$$

$$K_{(2)} = {}_1K_{(2)}\,{}_2\mu_{(1)}^2 + {}_1K_{(1)}\,{}_2\mu_{(2)}$$

$$K_{(3)} = {}_1K_{(3)}\,{}_2\mu_{(1)}^3 + 3\ {}_1K_{(2)}\,{}_2\mu_{(2)} + {}_1K_{(1)}\,{}_2\mu_{(3)}$$

where ${}_1K_{(j)}$ denotes the factorial cumulants of the distribution given by $g_1(z)$.

Finally, if the random variable X corresponding to $g_1(g_2(z))$ takes only positive integer values, the relation between the probabilities and the factorial moments given by Katti (1970), is

$$P_i = \sum_{r=i}^{\infty} (-1)^{r-1} \binom{r}{i} \frac{\mu_{(r)}}{r!}$$

1.2 Contagious distributions:

The term "contagious distributions" was used by Neyman (1939) for describing a model of the distribution of larvae in a randomly chosen area in a field. The model was constructed by assuming that the variation (in number of groups of eggs per unit area) could be represented by a Poisson distribution, while the number of larvae developing from a group could be represented by independent random variables, each having a Common Poisson distribution.

Actually it is a model representing heterogeneity. However, heterogeneity should be distinguished from a 'true contagion', i.e. from situations where the events under observation depend on the pattern of previous occurances of the events.

As an illustrative example, we shall discuss the Neyman type A distribution and its generalization.

The Neyman type A distribution (N.T.A.D.) is obtained by compounding a Poisson distribution by a Poisson distribution. Suppose the field is divided into plots of equal areas. On the assumption that the insects providing larvae eggs are distributed over these plots, the probability that exactly ν insects are represented on a plot under observation is

Supposedly given by a Poisson with parameter m_1

$$e^{-m_1}\frac{(m_1)^{\nu}}{\nu!}$$

Now suppose the probability that exactly n eggs are laid by the group of insects on any plot is given by a Poisson distribution with parameter λ, say;

$$e^{-\lambda}\,\frac{\lambda^{n}}{n!}\ . \qquad (1\text{-}3)$$

Out of the n eggs laid by the group, some may be unfertilized, and some may get rotten so that the number of learvae produced will usually be less than n, when the counting is done. Let X, the number of larvae counted on the plot, be given by the binomial distribution with parameters n and P. i.e. the probability of observing x larvae is given by

$$\binom{n}{x}\,P^{x}(1-P)^{n-x} \qquad (1\text{-}4)$$

The p.g.f. of the distribution of larvae is that of a generalized Poisson given by

$$e^{m_1\{g(t)\,-\,1\}},$$

where $g(t)$ is the p.g.f. of the binomial distribution (1-4) compounded with the Poisson distribution (1-3) on the parameter n. In this case the parameter P in (1-4) is regarded as fixed for all egg-masses; hence $g(t)$ is given by

$$g(t) = \sum_{n=0}^{\infty}[1+P(t-1)]^{n}e^{-\lambda}\,\frac{\lambda^{n}}{n!} = e^{\lambda P(t-1)} \qquad (1\text{-}5)$$

If in (1-4) we regard P as a random variable having Beta distribution with parameters α and β, the result of compounding the distribution represented by (1-5) through values of P yields the following p.g.f.

$$g_1(t) = \frac{1}{B(\alpha,\beta)}\int_0^1 e^{\lambda P(t-1)}P^{\alpha-1}(1-P)^{\beta-1}\,dP$$

$$= {}_1F_1(\alpha,\alpha+\beta,\lambda(t-1))$$

where ${}_1F_1(\cdot)$ is the confluent hypergeometric function.

It follows that the overall distribution of the total number of larvae produced by the insects on any plot has probability generating function

$$H(t) = exp[-m_1]exp[m_1 \cdot {}_1F_1\{\alpha;\alpha+\beta;\lambda(t-1)\}] \qquad (1\text{-}6)$$

which for $\beta = 0,1$ and 2 reduces respectively to the p.g.f. of Neyman's types A, B and C distributions.

Let P_r denote the probability of an observation of r events, setting,

$${}_1F_1\{\alpha;\alpha+\beta;\lambda(t-1)\} = f(t), \text{ then}$$

$$H(t) = exp[-m_1]exp[m_1 f(t)]$$

$$= \sum_{r=o}^{\infty} P_r t^r \qquad \text{and,}$$

$$P_r = \frac{1}{r!} H^{(r)}(0) \; .$$

Gurland (1958) has proved that

$$P_{r+1} = \frac{m_1}{r+1} \sum_{k=o}^{r} F_k \, P_{r-k}$$

where

$$F_k = \frac{1}{k!} \, f^{(k+1)}(0)$$

$$= \frac{\lambda^{k+1}}{k!} \, \frac{\alpha(\alpha+1)\ldots(\alpha+k)}{(\alpha+\beta)(\alpha+\beta+1)\ldots(\alpha+\beta+k)} \cdot {}_1F_1(\alpha+k+1;\alpha+\beta+k+1;-\lambda).$$

Also, he considered the following six limiting forms according to the variations in α and β.

(a) $\beta \to \infty$, α fixed

$${}_1F_1\{\alpha,\alpha+\beta,\lambda(t-1)\} = \sum_{r=0}^{\infty} \frac{\Gamma(\alpha+r)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+r)} \frac{(\lambda(t-1))^r}{r!} = f(t)$$

as $\beta \to \infty$, $f(t) \to 1$, and $H(t) \to 0$,

which is the p.g.f. of a degenerate distribution corresponding to the constant zero.

(b) $\alpha \to \infty$, β fixed,

$${}_1F_1\{\alpha,\alpha+\beta,\lambda(t-1)\} \to e^{\lambda(t-1)}, \quad \text{and} \quad H(t)$$

approaches the p.g.f. of the N.T.A.D. given by $exp[m_1\{exp(\lambda(t-1) - 1\}]$

(c) $\alpha \to \infty$, $\beta \to \infty$, and $\frac{\beta}{\alpha} = \gamma$

$${}_1F_1\{\alpha,\alpha+\beta,\lambda(t-1)\} \to exp[\lambda(t-1)/1+\gamma] \quad \text{and,}$$

$$H(t) \to exp\left\{m_1\left[exp\left\{\frac{\lambda}{1+\gamma}(t-1)\right\} - 1\right]\right\} \quad \text{the}$$

p.g.f. of a N.T.A.D.

In the other 3 cases where; (i) $\alpha \to \infty$, $\beta \to \infty$ with μ_1' and μ_2 fixed; (ii) $\beta \to \infty$ with α, μ_1' and μ_2 fixed; (iii) $\alpha \to \infty$ with β, μ_1' and μ_2 fixed, we get the N.T.A.D.

Although a considerable amount of work has been done in the area of obtaining statistical distributions of the spread of plants and insects, the distribution theory is presently unable to cope up with all the problems due to the following causes:

(1) No single theoretical distribution has been found to describe any large scale data.

(2) For a given set of data, there could be two or more theoretical distributions that fit equally well and there is no way to distinguish between them based on the fits only.

(3) Two or more physical models could lead to the same final statistical distribution and hence the estimates of the parameters of the distributions may not be meaningful to those physical models.

1.3 Generalized Power Series Distributions (GPSD):

A random variable X is said to have GPSD with parameter θ and range T if its probability density function is given by

$$P(X=x) = \begin{cases} a(x)\,\dfrac{\theta^x}{f(\theta)} & , \quad x \in T \\ 0 & \quad \text{elsewhere} \end{cases} \tag{1-7}$$

where $a(x)$ is independent of θ and is positive for $x \in T$, a countable subset of the set of non-negative integers, and the series function $f(\theta) = \sum_{x \in T} a(x)\theta^x$ is positive, finite and differentiable over $\Omega = \{\theta : 0 \leq \theta < R\}$, where R is the radius of convergence of the power series $f(\theta)$. It may be noted that the GPSD given by Patil (1962) reduces to the PSD (Noack, 1950), when T is the entire set of non-negative integers. It can be easily seen that a proper choice of T and $f(\theta)$ reduces the GPSD to the Binomial, Negative Binomial, Poisson, logarithmic series distributions and their truncated forms, since any truncated GPSD is itself a GPSD in its own right and thus possesses all the properties of a GPSD.

The p.g.f. of a GPSD is given by

$$\phi(t) = \frac{1}{f(\theta)} \sum_{x \in T} a(x)(t\theta)^x = \frac{f(t\theta)}{f(\theta)} \quad ,$$

and the moment generating function is

$$M(t) = \frac{f(e^t\theta)}{f(\theta)}$$

Thus

$$\mu = E(X) = \theta \frac{d}{d\theta} \log f(\theta)$$

$$\mu_2 = Variance(X) = \mu + \theta^2 \frac{d^2}{d\theta^2} \log f(\theta).$$

The recurrence relation among the non-central moments is given by

$$\mu'_{r+1} = \theta \frac{d\mu'_r}{d\theta} + \mu'_1 \mu'_r, \quad \text{where}$$

$$\mu'_r = \sum_{x \in T} x^r a(x) \frac{\theta^x}{f(\theta)}$$

whence, the recurrence relation among central moments of a GPSD becomes

$$\mu_{r+1} = \theta \frac{d\mu r}{d\theta} + r\, \mu_2 \mu_{r-1}$$

1.3.1 Estimation For GPSD

Let $X_1, X_2, \ldots X_n$ be a random sample of size n taken from the GPSD defined by (1-7). The logarithm of the likelihood function L is

$$\log L = \sum_{i=1}^{n} \log a(x_i) + \sum_{i=1}^{n} x_i \log \theta - n \log f(\theta),$$

so that the efficient score for θ is

$$\psi(\theta) = \sum_{i=1}^{n} \frac{x_i}{\theta} - n \frac{f'(\theta)}{f(\theta)} = \frac{n}{\theta}(\bar{x} - \mu).$$

The likelihood equation $\psi(\hat{\theta}) = 0$ for estimating θ reduces to

$$\bar{X} = \mu(\hat{\theta})$$

The asymptotic variance of the estimate of θ is

$$V(\hat{\theta}) = \frac{\theta^2/n}{\mu_2} \quad ,$$

and the amount of bias becomes

$$b(\hat{\theta}) = -\frac{\theta}{2n}\frac{\mu_3 - \mu_2}{\mu_2^2}$$

The uniformly minimum variance unbiassed (U.M.V.U) estimator of θ for GPSD has been given by Roy and Mitra (1957) as

$$u_1(t) = \begin{cases} 0 & t < 1 \\ \dfrac{c(t-1,n)}{c(t,n)} & t \geq 1 \end{cases}$$

where n is the sample size, $t = \sum_{i=1}^{n} X_i$ and $c(i,n)$ is the coefficient of θ^i in the expansion of $(f(\theta))^n$.

However, Charalambos A. Charalambides (1974 a) gave the U.M.V.U estimator for certain functions of the parameters involved in the left-truncated Logarithmic series, Poisson, Binomial, and Negative Binomial distributions with known or unknown truncation points. The derivation of all M.V.U. estimators depends essentially on Rao-Black well and Lehmann-Scheffé theorems using sufficiency and completeness of suitable statistics.

The distribution of the sum $z = \sum_{i=1}^{n} X_i$ of the observations when the truncation point is known and the joint distribution of $y = \min(X_1, X_2, \ldots X_n)$ and $z = \sum_{i=1}^{n} X_i$, when the truncation point is unknown, are derived. These distributions, as well as the corresponding M.V.U. estimators, have been expressed in terms of certain generalized Stirling numbers of the first kind in the Logarithmic series case, and in terms of generalized Stirling numbers of the second kind in the Poisson case. The results for the Binomial and Negative Binomial distributions are given in terms of a generalized C-numbers (Charalambides, 1974 b).

Before we finish this section, it should be noted that Kemp (1968) has shown that a large number of discrete probability functions may be expressed in terms of the generalized hypergeometric series. Dacey (1972) extended Kemp's results by constructing a method which identifies all members of this hypergeometric families of discrete probability laws given by the p.g.f.'s of the form

$$G(t) = C\ {}_pF_q(a_1, a_2, \ldots a_p;\ b_1, b_2, \ldots b_q;\ \lambda t),$$

where the normalizing constant $C = 1/{}_pF_q\ (a_1, a_2, \ldots a_p;\ b\ , \ldots b_q;\ \lambda)$.

All the members of the GPSD family are implied in this wider class of discrete probability distributions; e.g.

(i) Binomial: $G(t) = C\ {}_1F_0(-n;\ -\lambda t),\quad P = \dfrac{\lambda}{1 + \lambda}$

(ii) Poisson $G(t) = C\ {}_0F_0(\lambda t)$

(iii) Hyper-Poisson: $G(t) = C\ {}_1F_1(1;\ \lambda;\ \theta t)$

(iv) Negative-Binomial: $G(t) = C\ {}_1F_0(K,\ qt)$

1.4 The General Dirichlet's Series Distributions.

Siromoney (1964) has introduced another general class of probability distributions called the general Dirichlet's series distribution (GDSD), defined by the random variable X whose distribution is given by the probability function

$$P(X{=}x) = \begin{cases} a(x) exp(-\lambda_x \theta)/f(\theta), & x = 1,2,\ldots \\ 0 & \text{elsewhere} \end{cases}$$

The GPSD is a particular family of this class. Also the transformation $\lambda x = log\ x$ gives the following family

$$P(X{=}x) = a(x)\ x^{-\theta}/f(\theta), \qquad x = 1,2,\ldots$$

and

$$E(X^r) = \sum_{x=1}^{\infty} x^r\ P(X{=}x) = \frac{f(\theta - r)}{f(\theta)}$$

1.5 Systems Defined By Difference Equations

Analogous to Pearson's differential equation, used to develop families of continuous distributions, Ord (1967)employed the difference equation

$$\Delta f_{j-1} = f_j - f_{j-1} = \frac{(a-j)f_{j-1}}{b_o + b_1 j + b_2 j(j-1)} \tag{1-8}$$

to define a class of discrete distributions, based on a lattice of unit width. It was noticed that the density function will depend on the roots of the denominator in equation (1-8). The criterion $K = (b_1-b_2-1)^2/\{4b_2(b_o+2)\}$ is used to distinguish the distribution, when

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

which is an alternative form of (1-8). This last form is used to avoid infinite values when $b_o = 0$.

Katz (1968) 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,

$$f(x+1)/f(x) = \frac{P(x)}{Q(x)}, \qquad \text{where } P(x)$$

and $Q(x)$ are polynomials which are suitably chosen to provide the various classes of discrete probability distributions defined on non-negative integers.

When $P(x) = \alpha + \beta x$ and $Q(x) = x + 1$, we have

$$f(x+1)/f(x) = \frac{\alpha + \beta x}{x + 1}, \qquad 0 < \beta < 1,$$

$$x = 0,1,2,\ldots$$

On multiplication by t^x and summation over x from 0 to ∞, we have the differential equation,

$$g'(t) = \alpha g(t) + \beta t g'(t)$$

$$\frac{d}{dt} \log g(t) = \frac{\alpha}{1 - \beta t}, \qquad \text{where } g(t) = \sum_{x=0}^{\infty} t^x f(x) \text{ is the}$$

p.g.f. of X. Solving we get

$$g(t) = A(1 - \beta t)^{-\alpha/\beta} .$$

Since $\quad g(1) = 1,$

then $$g(t) = \left(\frac{1 - \beta t}{1 - \beta}\right)^{-\alpha/\beta} \tag{1-9}$$

Thus,

$$Mean = E(X) = \frac{\alpha}{1 - \beta}$$

$$Variance = \mu_2 = \frac{\alpha}{(1-\beta)^2}$$

As special cases:

(i) If $\beta = 0$

$\therefore \lim\limits_{\beta \to 0} g(t) = e^{\alpha(t-1)}$, which is the p.g.f. of a Poisson distribution.

(ii) If $\beta < 0$, $\frac{-\alpha}{\beta} = n$, ($n$ is an integer) and $\frac{-\beta}{1 - \beta} = p$

$\therefore g(t) = (q+pt)^n$, which is the p.g.f. of a Binomial distribution.

(iii) If $0 < \beta < 1$, $\frac{\alpha}{\beta} = n$, and $\frac{\beta}{1 - \beta} = p$

$\therefore g(t) = (q-pt)^{-n}$, which is the p.g.f. of a Negative-Binomial distribution.

Define G_r to be the r-th cumulant of the density function given by the p.g.f. in (1-9). Now

$$G_r = \left\{\frac{d^r}{dt^r} \log \frac{(1-\beta e^t)^{-\alpha/\beta}}{(1-\beta)^{-\alpha/\beta}}\right\}_{t=0}$$

$$= \frac{-\alpha}{\beta}\left\{\frac{d^r}{dt^r} \log(1-\beta e^t)\right\}_{t=0}$$

$$\frac{dG_r}{d\beta} = \frac{\alpha}{\beta^2}\left\{\frac{d^r}{dt^r} \log(1-\beta e^t)\right\}_{t=0} - \frac{\alpha}{\beta}\left\{\frac{d^r}{dt^r} \frac{d}{d\beta} \log(1-\beta e^t)\right\}_{t=0}$$

$$= -\frac{1}{\beta} G_r + \frac{\alpha}{\beta}\left\{\frac{d^r}{dt^r} \frac{e^t}{1-\beta e^t}\right\}_{t=0}$$

$$= -\frac{1}{\beta} G_r - \frac{\alpha}{\beta^2}\left\{\frac{d^r}{dt^r} \cdot \frac{d}{dt} \log(1-\beta e^t)\right\}_{t=0}$$

$$= -\frac{1}{\beta} G_r + \frac{1}{\beta} G_{r+1}$$

Hence, the recurrence relation between the cumulants becomes

$$G_{r+1} = G_r + \beta \frac{dG_r}{d\beta}$$

CHAPTER II

LAGRANGIAN PROBABILITY DISTRIBUTION

A new class of discrete probability distributions, under the title "Lagrangian Probability Distributions" was recently introduced into the statistical literature by Consul and Shenton (1971, '74) the particular title was chosen by them on account of the generation of these probability distributions by the well known Lagrange expansion of a function $f(t)$ as a power series in u when $u = \frac{t}{g(t)}$. The general class consists of many families and the Double Binomial Family of these discrete Lagrangian distributions has, as particular cases, many interesting members such as the Borel-Tanner distribution (Tanner, 1961), Haight distribution (Haight, 1961), Generalized Negative Binomial distribution (Jain and Consul, 1971), Generalized Poisson distribution (Consul and Jain, 1972). All these distributions are found to be of relevance in queuing theory and possess a number of interesting properties. The present Chapter is mostly based upon the works of Consul and Shenton (1971, '74).

2.1 USE OF LAGRANGIAN EXPANSION FOR GENERATING GENERALIZED DISCRETE PROBABILITY DISTRIBUTIONS.

Let $g(t)$ be a p.g.f. defined on any subset of non-negative integers and be given by

$$g(t) = \sum_{s=0}^{\infty} P(s)\, t^{s}, \qquad P(0) \neq 0 \tag{2-1}$$

so that

$$g(1) = \sum_{s=0}^{\infty} P(s) = 1 \quad \text{and} \quad g(0) \neq 0 \tag{2-2}$$

Consider the transformation

$$t = ug(t) \tag{2-3}$$

which gives $u = 0$ for $t = 0$ and $u = 1$ for $t = 1$. Since u is a function of t, having a non-zero derivative at $t = 0$, it is reasonable to suppose that there exists a power-series expansion for t in terms of u, given by the Lagrange expansion.

If $f(t)$ is another p.g.f. defined on any subset of non-negative integers, it can be expanded in powers of u, by Lagrange's expansion as

$$f(t) = f(0) + \sum_{k=1}^{\infty} \frac{u^k}{k!} \frac{\partial^{k-1}}{\partial t^{k-1}} \left[(g(t))^k f'(t) \right]_{t=0} \tag{2-4}$$

The Lagrange expansion is usually defined for functions $f(z)$ and $g(z)$ which are analytic on and within a contour C surrounding a point $z = a$ such that

$$|u\, g(z)| < |z - a|.$$

In our case, this suggests the condition $|u \cdot g(t)| < |t|$ or $|u \cdot g(t)| < 1$, for the convergence of the p.g.f. $f(t)$. When $\phi_1(t)$ and $\phi_2(t)$ are two p.g.f.'s, it is well known that $\phi_1[\phi_2(t)]$ is also a p.g.f., so that the power series (2-4) must be another p.g.f. in u. Thus a new probability distribution, defined on some subset of non-negative integers, is given by

$$p(X=x) = \frac{1}{x!} D_t^{x-1} \left[(g(t))^x f'(t) \right]_{t=0} \tag{2-5}$$

where

$$p(X=0) = f(0) \, , \quad \text{and} \quad D_t^{x-1} = \frac{\partial^{x-1}}{\partial t^{x-1}}$$

Since (2-5) provides us with many families of discrete distributions as $g(t)$ and $f(t)$ are replaced by different sets of p.g.f.'s, it is called the class of Lagrangian distributions.

2.2 BASIC LAGRANGE PROBABILITY DISTRIBUTIONS

When $f(t) = t$, (2-4) gives

$$t = \tau(u) = \sum_{x=1}^{\infty} \frac{u^x}{x!} \left[D_t^{x-1} \left(g(t)\right)^x \right] \Big|_{t=0} = E[U^x] \qquad (2\text{-}6)$$

The above is referred as basic Lagrangian p.g.f. and the corresponding discrete probability distribution is given by,

$$P(X=x) = \frac{1}{x!} \left[D_t^{x-1} \left(g(t)\right)^x \right] \Big|_{t=0} \qquad x \in N \qquad (2\text{-}7)$$

where N is the set of positive integers. This probability distribution is called the family of basic Lagrangian distribution defined on N. This family consists of many well known probability distributions such as the Geometric distribution, the Borel-Tanner distribution, the Haight distribution; the Consul distribution which are respectively obtained when

(i) $g(t) = p + qt$, $q + p = 1$, $0 < p < 1$

(ii) $g(t) = e^{m(t-1)}$

(iii) $g(t) = p(1 - qt)^{-1}$ $q + p = 1$ $0 < p < 1$, and

(iv) $g(t) = (1 - \theta - \theta t)^m$ $0 < \theta < 1$, m is a positive integer.

Other choices for $g(t)$, such as

(v) $g(t) = (Q - Pt)^{-k}$ $Q - P = 1$

(vi) $g(t) = \dfrac{\log(1-\theta t)}{\log(1-\theta)}$ $0 < \theta < 1$

will provide new distributions.

THEOREM (i)

Let $X_1, X_2, \ldots X_n$ be n independently and identically distributed random variables whose probability function is given by the p.g.f. $g(t)$. If X is a basic Lagrangian random variable generated by $g(t)$, then

$$P(X=n) = n^{-1} \times P(X_1+X_2+\ldots X_n = n-1)$$

PROOF:

The p.g.f. of the n-fold convolution of the probability distribution (p.d.) given by the p.g.f. $g(t)$ is $(g(t))^n$. Let

$$g(t) = g_0 + tg_1 + t^2g_2 + \ldots. + t^r g_r + \ldots.$$

and

$$(g(t))^n = g_0^{\,n} + tg_1^{\,n} + t^2g_2^{\,n} + \ldots. \, t^r g_r^{\,n} + \ldots.$$

so that

$$P(X_1 + X_2 + \ldots X_n = r) = g_r^n$$

$$g_r^n = \frac{1}{r!} D_t^r \, (g(t))^n \Big|_{t=0} \quad , \text{ or } \quad D_t^r \, (g(t))^n \Big|_{t=0} = r! \, g_r^n$$

since,

$$P(X=n) = \frac{1}{n!} D_t^{n-1} \left[(g(t))^n\right] \Big|_{t=0}$$

$$= \frac{(n-1)!}{n!} \, g_{n-1}^n$$

$$= \frac{1}{n} \, g_{n-1}^n = \frac{1}{n} P(X_1 + X_2 + \ldots X_n = n-1) \qquad (2\text{-}8)$$

THEOREM (ii)

The value of $P(X=n)$ in the basic LPD represents the probability of a regular (uninterrupted) sequence consisting of n events only when the observer is on the other side of a resisting medium or a barrier to that the counting starts with the observation of the first unit and the individual events are processed one by one through the resisting medium. To prove this theorem we shall need the following lemma.

LEMMA

Let g^{n+1} and $\partial^n g$ denote the functions $[g(t)]^{n+1}$ and $\frac{d^n}{dt^n} g(t)$ respectively. Then

$$(n+1)^{-1}(\partial^n g^{n+1}) = (\partial g)(\partial^{n-1} g^n) + (\partial^n g)g^n + \sum_{r_1=1}^{n-2} \binom{n}{r_1} (\partial^{n-r_1} g) .$$

$$\Big[(\partial^{r_1} g^{n-r_1})g^{r_1} + \sum_{r_2=1}^{r_1-1} \binom{r_1}{r_2} (\partial^{r_1-r_2} g^{n-r_1})\{(\partial^{r_2} g^{r_1-r_2})g^{r_2}$$

$$+ \sum_{r_3-1}^{r_2-1} \binom{r_2}{r_3} (\partial^{r_2-r_3} g^{r_1-r_2})\{(\partial^{r_3} g^{r_2-r_3})g^{r_3} + \ldots\}\Big] ,$$

where the values of the summation is zero if the upper range in the summation is less than the lower range.

PROOF:

According to the LIEBNITZ theorem,

$$\frac{\partial^n}{\partial x^n} (u(x) \cdot v(x)) = \sum_{r=0}^{n} \binom{n}{r} \left(\frac{\partial^r}{\partial x^r} v(x)\right)\left(\frac{\partial^{n-r}}{\partial x^{n-r}} u(x)\right) .$$

Now, $(n+1)^{-1}(\partial^n g^{n+1}) = (n+1)^{-1} \partial^{n-1}(\partial g^{n+1}) = (n+1)^{-1} \partial^{n-1} (n+1) g^n \partial g$

$$= \partial^{n-1}(g^n \cdot \partial g)$$

$$= \sum_{r_1=0}^{n-1} \binom{n-1}{r_1} \partial^{r_1}(g^n)\, \partial^{n-1-r_1} (\partial g)$$

Putting $r_1 = 0$ and $r = n-1$, i.e. writing down the first and the last terms. and on using the relations

$$\binom{n-1}{r}\partial g^{n} = \binom{n}{r} g^{r}\, \partial g^{n-r}$$

and $\binom{n-1}{r}(\partial^{n-r}g)(\partial^{r_1}g^{n}) = \binom{n}{r}(\partial^{n-r_1}g)\, \partial^{r_1-1}(g^{r_1}\partial g^{n-r_1})$

we get,

$$(n+1)^{-1}(\partial^{n}g^{n+1}) = (\partial g)(\partial^{n-1}g^{n}) + (\partial^{n}g)\, g^{n}$$

$$+ \sum_{r_1=1}^{n-2}\binom{n}{r_1}(\partial^{n-r_1}g)\, \partial^{r_1-1}(g^{r_1}\partial g^{n-r_1})$$

$$= (\partial g)(\partial^{n-1}g^{n}) + (\partial^{n}g)g^{n} + \sum_{r_1=1}^{n-2}\binom{n}{r_1}(\partial^{n-r_1}g) \left[\sum_{r_2=0}^{r_1-1}\binom{r_1-1}{r_2}\left(\partial^{r_1-r_2} g^{n-r_1}\right)\left(\partial^{r_2} g^{r_1}\right)\right]$$

Now, we separate the first term in the second summation and modify the form of the derivative slightly to obtain .

$$(n+1)^{-1}(\partial^{n}g^{n+1}) = (\partial g)(\partial^{n-1}g^{n}) + (\partial^{n}g)\, g^{n}$$

$$+ \sum_{r_1=1}^{n-2}\binom{n}{r_1}(\partial^{n-r_1}g)\left[(\partial^{r_1}g^{n-r_1})g^{r_1} + \sum_{r_2=1}^{r_1-1}\binom{r_1}{r_2}(\partial^{r_1-r_2}g^{n-r_1})\, \partial^{r_2-1}(g^{r_2}\partial g^{r_1-r_2})\right] .$$

The last derivative under the second summation can be again written in the form of a summation and the process of separating the first term and modifying the derivative in the others can be continued further. Repeating these operations till such conversions become impossible, we can obtain the required result. Since

$$g_r = \frac{1}{r!}\left[\partial^r g(t)\right]_{t=0}, \quad g_r^{k} = \frac{1}{r!}\left[\partial^r (g(t))^k\right]_{t=0}$$

The result of the lemma at the point $t=0$ can be simplified to the form

$$\frac{1}{(n+1)!}\left(\partial^n g^{n+1}\right)\Big|_{t=0} = g_1 \frac{1}{n!}\left(\partial^{n-1} g^n\right)_{t=0} + g_n g_0^{n} +$$

$$\sum_{r_1=1}^{n-2} g_{n-r_1}\left[g_{r_1}^{n-r_1} g_0^{r_1} + \sum_{r_2=1}^{r_1-1} g_{r_1-r_2}^{n-r_1}\{g_{r_2}^{r_1-r_2} g_0^{r_2} + \sum_{r_3=1}^{r_2-1} g_{r_2-r_3}^{r_1-r_2}\{g_{r_3}^{r_2-r_3} g_0^{r_3} + \ldots\}\}\right] \tag{2-9}$$

which is true for all integral values of n.

<u>PROOF of the theorem:</u>

When the resisting medium is uniform, each unit will take the same period of time to get through it and the observer will find the units appearing sequentially at regular intervals without interruption unless there was a break in the flow. The length of these uninterrupted regular sequences of events will be 1,2,3,... n,... units with gaps of longer intervals between the different sequences.

The present problem is very similar to that of recurrent events or a simple queuing process where an external observer, who is unable to see the length of the queue on account of some barrier, is trying to determine the probability that an observed regular sequence (busy period) will consist of n events (customers) when (i) the passage through the barrier is like an unknown contagion having the same effect on every event and takes the same amount of time and (ii) the process does not stop at any time so that any new regular sequence of events begins without a waiting period, i.e. each sequence is always initiated by one unit.

The problem has been studied by several mathematicians [Feller, 1968, Vol. 1]. Therefore, what is required to be proved is that the value of the probability (2-7) for $X=n$ is precisely equal to the probability, obtained by other methods, that the observed regular sequence will consist of n events when (i) and (ii) are true.

Let $g(t)$ be the p.g.f. of the p.d. $\{g_r\}$, $r = 0,1,\ldots$ of r units materializing during the period of time taken by one unit for passage through the barrier. Evidently, $\left(g(t)\right)^{k}$ is the k-fold convolution and its distribution is $\overset{k}{g}_r$, $r = 0,1,2,\ldots$ where $\overset{k}{g}_r$ is the probability of arrival of r units during the passage of k units through the barrier.

Let the system be in state k when $(k-1)$ units are waiting for passage and the k-th unit is getting through the passage or is being processed. Let $P_k(X=x)$ denote the probability of the regular sequence having x more events when the system is in state k. Also, let a regular sequence consist of $(n+1)$ events. Now

$$P(X=n+1) = \sum_{r_1=0}^{n-1} \{\text{Prob. of } (n-r_1) \text{ arrivals during the passage of the first}\} \cdot P_{n-r_1}(X=n)$$

$$\therefore P_1(X=n+1) = \sum_{r_1=0}^{n-1} g_{n-r_1} P_{n-r_1}(X=n)$$

Separating the first and the last term, i.e. putting $r_1=0$ and $r_1=n-1$,

$$P_1(X=n+1) = g_1 P_1(X=n) + g_n P_n(X=n) + \sum_{r_1=1}^{n-2} g_{n-r_1} P_{n-r_1}(X=n)$$

$$= g_1 P_1(X=n) + g_n \overset{n}{g}_0 + \sum_{r_1=1}^{n-2} g_{n-r_1} \left[\sum_{r_2=0}^{r_1-1} \{\text{Prob. of } (r_1-r_2) \text{ arrivals during the passage of } (n-r_1)\} \times P_{r_1-r_2}(X=r_1) \right]$$

$$= g_1 P_1(X=n) + g_n g_0^n + \sum_{r_1=1}^{n-2} g_{n-r_1} \left[\sum_{r_2=0}^{r_1-1} g_{r_1-r_2}^{n-r_1} P_{r_1-r_2}(X=r_1) \right]$$

$$= g_1 P_1(X=n) + g_n g_0^n + \sum_{r_1=1}^{n-2} g_{n-r_1} \left[g_{r_1}^{n-r_1} g_0^{r_1} + \sum_{r_2=1}^{r_1-1} g_{r_1-r_2}^{n-r_1} P_{r_1-r_2}(X=r_1) \right]$$

This process can be continued onwards successively till the summation reduces to a single value. Thus

$$P_1(X=n+1) = g_1 P_1(X=n) + g_n g_0^n + \sum_{r_1=1}^{n-2} g_{n-r_1} \left[g_{r_1}^{n-r_1} g_0^{r_1} + \sum_{r_2=1}^{r_1-1} g_{r_1-r_2}^{n-r_1} \left\{ g_{r_2}^{r_1-r_2} g_0^{r_2} + \sum_{r_3=1}^{r_2-1} g_{r_2-r_3}^{r_1-r_2} \left\{ g_{r_3}^{r_2-r_3} g_0^{r_3} + \sum_{r_4=1}^{r_3-1} \cdots \right\} \right\} \right] \tag{2-10}$$

The above expression is true for all integral values of n. By subtracting (2-10) from (2-9) in our lemma, one finds that

$$P_1(X=n+1) - \frac{1}{(n+1)!} (\partial^n g^{n+1})_{t=0} = g_1 \left[P_1(X=n) - \frac{1}{n!} (\partial^{n-1} g^n)_{t=0} \right] \tag{2-11}$$

for all integral values of n. Recursively, the above formula gives

$$P(X=n+1) - \frac{1}{(n+1)!} (\partial^n g^{n+1})_{t=0} = (g_1)^n [P_1(X=1) - g_0]$$

$$= (g_1)^n [g_0 - g_0] = 0$$

Thus,

$$P(X=n+1) = \frac{1}{(n+1)!} (\partial^n g^{n+1})_{t=0}$$

Hence the probability $P_1(X=n+1)$, given by (2-10) of a regular sequence consisting of $n+1$ events only is equal to the value $P(X=n+1)$ in the basic

LPD and the theorem is true for all values of n. Some particular cases of (2-10) are

$$P_1(X=1) = g_0 \quad , \quad P_1(X=2) = g_1 g_0 = (2!)^{-1}(\partial g^2)_{t=0}$$

$$P_1(X=3) = g_2 g_0^{\,2} + g_1 g_1 g_0 = (3!)^{-1}(\partial^2\, g^3)_{t=0}$$

$$P_1(X=4) = g_3 g_0^{\,3} + 3g_2 g_1 g_0^{\,2} + (g_1)^3 g_0 = (4!)^{-1}(\partial^3\, g^4)_{t=0}$$

2.3 LAGRANGIAN DISTRIBUTIONS IN GENERAL

If $g(t)$ is the p.g.f. of the transformation $t = ug(t)$ and $f(t)$ is the transformed p.g.f., the general Lagrangian probability distributions were defined by

$$P(X=x) = \frac{1}{x!}\, D_t^{x-1} \left\{ \left(g(t)\right)^x \frac{\partial f(t)}{\partial t} \right\}_{t=0} \quad , x = 0,1,2,\ldots \tag{2-12}$$

The following theorems clarify the probabilistic structure of LPD's.

Theorem (iii)

The Lagrangian distributions defined by (2-12) for $f(t) = t^n$ are the n-th fold convolutions of the basic LPD given by (2-7).

The proof follows trivially by considering the p.g.f.'s of the two families of distributions.

Theorem (iv)

The general Lagrangian probability distributions given by (2-12), are obtained by randomizing the index parameter n in theorem (iii) according to the p.d. generated by the p.g.f. $f(t)$.

PROOF: If $\{f_r\}$, $r = 0,1,2,\ldots$ represent the successive probabilities in the p.d. of the p.g.f. $f(t)$, then

$$f(t) = \sum_{n=0}^{\infty} f_n t^n = \sum_{n=0}^{\infty} f_n \left(\psi(u)\right)^n$$

$$= f(0) + \sum_{n=1}^{\infty} f_n \left[\sum_{x=n}^{\infty} \frac{u^x}{x!} \left(\frac{\partial}{\partial t}\right)^{x-1} \left\{\left(g(t)\right)^x n\, t^{n-1}\right\}\Big|_{t=0}\right] \tag{2-13}$$

Also,

$$f(t) = f(0) + \sum_{n=1}^{\infty} f_n \left[\sum_{x=n}^{\infty} \frac{u^x}{x!} \binom{x-1}{n-1} n! \left(\frac{\partial}{\partial t}\right)^{x-n} \left(g(t)\right)^x\Big|_{t=0}\right]$$

$$= f(0) + \sum_{n=1}^{\infty} f_n \sum_{x=n}^{\infty} \frac{n u^x}{x(x-n)!} \left(\frac{\partial}{\partial t}\right)^{x-n} \left(g(t)\right)^x\Big|_{t=0} \tag{2-14}$$

The result (2-13) can also be expressed as

$$f(t) = f(0) + \sum_{x=0}^{\infty} \frac{u^x}{x!} \left(\frac{\partial}{\partial t}\right)^{x-1} \left(g(t)\right)^x \left\{\sum_{n=1}^{x} n f_n\, t^{n-1}\right\}\Big|_{t=0} \tag{2-15}$$

The theorem follows easily from (2-14) and (2-15).

Theorem (v)

The negative binomial distribution is a Lagrangian p.d. of the simple binomial element given by $g(t) = q + Pt$.

PROOF: Let $f(t) = (q + Pt)^n$ where

$$t = ug(t) = u(q + Pt)$$

$$\therefore t = uq(1 - uP)^{-1}$$

and the p.g.f. of the corresponding Lagrangian distributions becomes,

$$\begin{aligned} \psi(u) &= f(t) = [q + Puq(1 - uP)^{-1}]^n \\ &= q^n(1 - uP)^{-n} \end{aligned}$$

which is the p.g.f. of the negative binomial distribution.

Theorem (vi)

All Lagrangian probability distributions are closed under convolution if the transformation $t = ug(t)$ remains the same.

PROOF: Let X_1 and X_2 be two independent Lagrangian random variables, but not identically distributed, whose distributions are given by $f_1(t)$ and $f_2(t)$ under the transformation $t = ug(t)$ so that their p.g.f.'s are

$$\psi_1(u) = f_1(t) = \sum_{h=0}^{\infty} \frac{u^h}{h!} D_t^{h-1} \left\{ (g(t))^h \frac{\partial}{\partial t} f_1(t) \right\} \Big|_{t=0}$$

and

$$\psi_2(u) = f_2(t) = \sum_{k=0}^{\infty} \frac{u^k}{k!} D_t^{k-1} \left\{ (g(t))^k \frac{\partial}{\partial t} f_2(t) \right\} \Big|_{t=0}$$

The sum $X_1 + X_2$ has the generating function $\psi_1(u)\,\psi_2(u)$. On account of the transformation $t = ug(t)$ the product $\psi_1(u)\,\psi_2(u) = f_1(t)\,f_2(t)$ of the above two expansions can be easily written in the form

$$\sum_{r=0}^{\infty} \frac{u^r}{r!} D_t^{r-1} \left\{ (g(t))^r \frac{\partial}{\partial t} \{f_1(t)\cdot f_2(t)\} \right\} \Big|_{t=0}$$

which is a Lagrangian p.g.f. The result can be easily extended to the sum of any number of independent Lagrangian variates.

2.4 CUMULANTS AND MOMENTS OF LAGRANGE DISTRIBUTIONS

The following notations will be used in general:

Distribution	Lagrange	Transformer	Transformed
p.g.f.	$\gamma(u)$	$g(t)$	$f(t)$
cumulants	L_r	G_r	F_r
non-central moments	μ_r	g_r	f_r
$P(X=x)$	$P_l(x)$	$P_g(x)$	$P_f(x)$

Cumulants of the Lagrange Distributions

The relations for the cumulants of the L.D. are derived by the mapping $t = exp(T)$, $u = exp(\beta)$ so that (2-3) becomes

$$T = \beta + \log g(e^T) \qquad \text{(after taking logarithm).} \tag{2-16}$$

Writing

$$f(t) = \sum_{x=0}^{\alpha} u^x P_l(x) = \gamma(u) \ , \tag{2-17}$$

and putting $t = e^T$ and $u = e^\beta$ in (2-17),

the last relation becomes

$$\log f(e^T) = \sum_{r=1}^{\infty} L_r \beta^r/r! \, . \tag{2-18}$$

Relations (2-16) and (2-18) can be written as

$$T = \beta + \sum_{r=1}^{\infty} G_r T^r/r! \tag{2-19}$$

and

$$\sum_{r=1}^{\infty} F_r T^r/r! = \sum_{r=1}^{\infty} L_r \frac{\beta^r}{r!} \, . \tag{2-20}$$

The process of solving these two equations for the cumulants L_1, L_2, $L_3, \ldots$ is rather complicated, but the task becomes somewhat simpler by introducing the cumulants $\{D_r\}$, $r=1,2,\ldots$ of the Lagrange-Dirac distribution defined by (2-19) and

$$T = \sum_{r=1}^{\infty} D_r \beta^r / r! \tag{2-21}$$

Thus, the relation (2-20) can be written as

$$\sum_{r=1}^{\infty} L_r \beta^r / r! = \sum_{m=1}^{\infty} F_m \left[\sum_{s=1}^{\infty} D_s \beta^s / s! \right]^m \bigg/ m! \tag{2-22}$$

so that

$$L_r = \frac{\partial^r}{\partial \beta^r} \left[\sum_{m=1}^{\infty} F_m \left(\sum_{s=1}^{\infty} D_s \beta^s / s! \right)^m \bigg/ m! \right] \Bigg|_{\beta=0} \tag{2-23}$$

$$= \sum_{m=1}^{r} \left[F_m \frac{\partial^r}{\partial \beta^r} \left(\sum_{s=1}^{r} D_s \beta^s / s! \right)^m \bigg/ m! \right] \Bigg|_{\beta=0}$$

The coefficient of F_m, $m=1,2,3,\ldots$ can be obtained by expanding the multinomial $\left(\sum_{s=1}^{r} D_s \beta^s / s! \right)^m$ and choosing the coefficient of β^r. This can be achieved in a systematic manner by considering all possible partitions of m. Thus, the cumulants L_r, $r=1,2,3,\ldots$ of the Lagrange distribution become

$$L_r = \sum_{m=1}^{r} F_m \left[\frac{r!}{\pi_1 ! \, \pi_2 ! \, \ldots \, \pi_r !} \prod_{i=1}^{r} (\phi_i)^{\pi_i} \right] \tag{2-24}$$

where the second summation is taken over all partitions $\pi_1, \pi_2, \ldots \pi_r$ of m such that

$$\pi_1 + 2\pi_2 + \ldots r\pi_r = r \tag{2-25}$$

and

$$\phi_i = \frac{D_i}{i!} \tag{2-26}$$

As particular cases the first five cumulants of L.D. can now be written in the form

$$L_1 = F_1D_1 \tag{2-27}$$
$$L_2 = F_1D_2 + F_2D_1^2$$
$$L_3 = F_1D_3 + 3F_2D_1D_2 + F_3D_1^3$$
$$L_4 = F_1D_4 + 3F_2D_2^2 + 4F_2D_1D_3 + 6F_3D_1^2D_2 + F_4D_1^4$$
$$L_5 = F_1D_5 + 5F_2D_1D_4 + 10F_2D_2D_3 + 10F_3D_1^2D_3 + 15F_3D_1D_2^2 + 10F_4D_1^3D_2 + F_5D_1^5$$

Though the cumulants $\{D_r\}$ are not yet determined and so are $\{L_r\}$, one nice thing about the relations (2-24) and (2-27) is that they are similar to the expressions for non-central moments in terms of cumulants given on pages 68-69 by Kendall and Stuart (1963). The expression (2-24) resembles Faa de Bruno's expression (1855) for the n-th derivative of a composite function and can be easily programmed for a digital computer, assuming that an expression for the D's is available, and all higher cumulants can be determined.

Since the cumulants, given by (2-24) and (2-27), of the general L.D. depend upon the cumulants D_s, $s=1,2,3,\ldots$ of the basic L.D., we shall now derive simple expressions for D_s. For BLD equations,

$$T = \beta + \log g(e^T) \tag{2-28}$$
$$= \beta + \sum_{r=1}^{\infty} G_r \frac{T^r}{r!}$$

while relation (2-18) changes to

$$T = \sum_{r=1}^{\infty} L_r \frac{\beta^r}{r!} \tag{2-29}$$

On eliminating T from (2-28) by using the value of T from (2-29), we get

$$\sum_{s=1}^{\infty} D_s \ \beta^s/s! = \beta + \sum_{r=1}^{\infty} G_r \left(\sum_{s=1}^{\infty} D_s \ \beta^s/s!\right)^r \Big/ r! \tag{2-30}$$

If we define

$$D^*_s = D_s \ , \qquad s = 2,3,4,\ldots$$

$$D^*_1 = D_1 - 1$$

Then (2-30) can be expressed as

$$\sum_{s=1}^{\infty} D_s^* \ \beta^s/s\,! = \sum_{r=1}^{\infty} G_r\left(\sum_{s=1}^{\infty} D_s \ \beta^s/s\,!\right)^r \Big/ r! \ ,$$

which is precisely the relation between $\{L_r\}$ and $\{F_r\}$ given in (2-22).

As an illustration we can write some particular cases as

$$D^*_1 = G_1D_1$$

$$D_2 = G_1D_2 + G_2D_1{}^2$$

$$D_3 = G_1D_3 + 3G_2D_1D_2 + G_3D_1{}^3$$

$$D_4 = G_1D_4 + 3G_2D_2{}^2 + 4G_2D_1D_3 + 6G_3D_1{}^2D_2 + G_4D_1{}^4$$

$$D_5 = G_1D_5 + 10G_2D_2D_3 + 5G_2D_1D_4 + G_3(10D_1{}^2D_3 + 5D_1D_2{}^2 + 5D_1D_2{}^3\ )$$
$$+ 10G_4D_1{}^3D_2 + G_5D_1{}^5$$

which can be solved successively to yield

$$D_1 = 1/(1 - G_1)$$

$$D_2 = G_2/(1 - G_1)^3$$

$$D_3 = G_3/(1 - G_1)^4 + 3G_2{}^2/(1 - G_1)^5$$

$$D_4 = G_4/(1 - G_1)^5 + 10G_3G_2/(1 - G_1)^6 + 15G_2{}^3/(1 - G_1)^7$$

$$D_5 = G_5/(1 - G_1)^6 + 15G_4G_2/(1 - G_1)^7 + 10G^2{}_3/(1 - G_1)^7$$
$$+ 105G_3G_2{}^2/(1 - G_1)^8 + 105G_2{}^4\ /(1 - G_1)^9$$

From (2-27) and (2-31) one may write easily the first four cumulants of L.P.D. in terms of G_r and F_r, writing $\lambda = (1 - G_1)^{-1}$

$$L_1 = F_1\lambda$$

$$L_2 = F_1G_2\lambda^3 + F_2\lambda^2 \qquad (2\text{-}32)$$

$$\begin{aligned} L_3 &= F_1(G_3\lambda^4 + 3G_2{}^2\lambda^5) + 3F_2G_2\lambda^4 + F_3\lambda^3 \\ L_4 &= F_1(G_4\lambda^5 + 10F_1G_2G_3\lambda^6 + 15G_2{}^3\lambda^7) + 3F_2G_2{}^2\lambda^6 \\ &\quad + 4F_2(G_3\lambda^5 + 3G_2{}^2\lambda^6) + 6F_3G_2\lambda^5 + F_4\lambda^4 \end{aligned} \qquad (2\text{-}32)$$

2.5 SOME IMPORTANT LAGRANGIAN DISTRIBUTIONS

(i) The Double Binomial Family

When $g(t) = (q + Pt)^m$ for $P + q = 1$ such that $mP < 1$, and $f(t) = (q' + P't)^n$ for $P' + q' = 1$, the Lagrangian probability distribution takes the form

$$P(X=0) = (q')^n$$

$$P(X=x) = \frac{n}{x}(q')^n (Pq^{m-1})^x \sum_{r=0}^{k} \binom{n-1}{r}\binom{mx}{x-r-1}\left(\frac{qP'}{q'P}\right)^{r+1} \qquad (2\text{-}33)$$

$x = 1,2,3,\ldots$, where $k = \min(x-1, n-1)$.

The Lagrangian probability distribution obtained in (2-33) can be described as a true double binomial when $0 < P, P' < 1$ and m, n are positive integers.

The first four central moments of the Lagrange distribution (2-33) can be easily written down by (2-32) and are given by

Mean $= \mu'_1 = nP'(1 - mP)^{-1}$

Variance $= \mu_2 = nP'\{q' + mP(q - q')\}(1 - mP)^{-3}$

$$\mu_3 = \mu_2\left\{\frac{3mPq}{(1-mP)^2} + \frac{2q + 2q' - 1}{1-mP}\right\} - \frac{2nqP'q'}{(1-mP)^4} ,$$

$$\begin{aligned} \mu_4 &= 3\mu^2{}_2 + \frac{\mu^2}{(1-mP)^2}\left[1 - \frac{mPq(4-6q)}{1-mP} + \frac{mPq^2(8+7mP)}{(1-mP)^2}\right] \\ &\quad - 6mnPP'q'(q-q')(P'+q)(1-mP)^{-5} - 6nP'[P'q'^2 + \\ &\quad mP(Pq^2-P'q'^2)](1-mP)^{-6} . \end{aligned}$$

The following are some particular cases of the double binomial family of LPD's.

(ii) <u>The Lagrangian Poisson Distribution (LP)</u>

The Lagrangian Poisson distribution is defined by

$$P_x(\lambda_1,\lambda_2) = \lambda_1(\lambda_1 + \lambda_2 x)^{x-1}\,\frac{e^{-(\lambda_1+\lambda_2 x)}}{x!}\,,\qquad \begin{array}{l}\lambda_1 > 0\\ |\lambda_2| < 1\,,\quad x = 0,1,2,\ldots\,,\end{array} \tag{2-34}$$

such that,

$$P_x(\lambda_1,\lambda_2) = 0 \qquad \text{for } x \geq m \text{ if } \lambda_1 + m\lambda_2 \leq 0.$$

It is given by (2-12) when $g(t) = e^{\lambda_2(t-1)}$ and $f(t) = e^{\lambda_1(t-1)}$.

If we allow λ_2 to be negative, then we do not get a "distribution" since $P_x(\lambda_1,\lambda_2)$ will be negative for sufficiently large X. We can get "probabilities" that come close to unity when $\lambda_1 > 4|\lambda_2|$, and that these "approximate distributions" are sometimes useful.

The first four cumulants of the LP are given by (2-32) in the form

$$L_1 = \lambda_1(1-\lambda_2)^{-1}$$
$$L_2 = \lambda_1(1-\lambda_2)^{-3}$$
$$L_3 = \lambda_1(1+2\lambda_2)(1-\lambda_2)^{-5}$$
$$L_4 = 3\lambda_1^{\,2}(1-\lambda_2)^{-6} + \lambda_1(1+8\lambda_2+6\lambda_2^{\,2})(1-\lambda_1)^{-7}.$$

Consul and Shenton (1974) have shown that the recurrence relation between the consecutive cumulants is given by:

$$(1-\lambda_2)L_{k+1} = \lambda_2\frac{\partial L_k}{\partial\lambda_2} + \lambda_1\frac{\partial L_k}{\partial\lambda_1}\,,\quad k = 1,2,3,\ldots\,,$$

with $L_1 = \lambda_1(1-\lambda_2)^{-1}$.

(iii) The Lagrangian Binomial Distribution (LBD)

The LBD is a special case of the double binomial family given by

$$P' = P = \theta \ , \ g(t) = (1-\theta+\theta t)^{\beta} \text{ and } f(t) = (1-\theta+\theta t)^{n} \ .$$

Thus, for $n > 0$, $0 < \theta < 1$ and $|\theta\beta| < 1$, the p.d.f. of the LBD as given by Jain and Consul (1971) is,

$$P(X=x) = \frac{n}{n+\beta x}\binom{n+\beta x}{x}\theta^{x}(1-\theta)^{n+\beta x-x} \qquad (2\text{-}35)$$

$$x = 0,1,2,\ldots$$

$$\beta \geq 0 \ .$$

The classical negative binomial distribution with parameters θ and n is a special case of the LBD with parameters θ, n and β, and is obtained when $\beta = 1$. The binomial distribution is also a particular case of the LBD and is obtained when $\beta = 0$.

The moments of the distribution are given by

$$\text{Mean} = n\theta/(1-\beta\theta)$$

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

The recursive relation between the cumulants (L_s, $s=1,2,\ldots$) is given by:

$$(1-\beta\theta)L_{s+1} = \theta(1-\theta)\left(\frac{\partial L_s}{\partial\theta}\right) \qquad \text{(Consul and Shenton, 1974).}$$

Jain and Consul (1971) have studied some of the properties of the distribution.

If θ is very small and β, n are very large such that $n\theta = \lambda_1$ and $\beta\theta = \lambda_2$, where λ_1 is finite, positive and $|\lambda_2| < 1$, the LBD can be approximated by Lagrangian Poisson distribution.

As $n \to 0$, the zero-truncated LBD with parameters n, θ and β tends to the Lagrangian Logarithmic distribution with parameters θ and β. Its probability function is given by:

$$P(X=x) = [-\log(1-\theta)]^{-1} \frac{\theta^{x}}{x} \frac{\Gamma(x\beta)}{\Gamma(x)\Gamma(\beta x-x+1)} (1-\theta)^{\beta x-x} \tag{2-36}$$

which gets reduced to the logarithmic series distribution when $\beta = 1$. It should be noted that taking,

$$f(t) = -\log(1-\theta t)/-\log(1-\theta),\ g(t) = \left(\frac{1-\theta t}{1-\theta}\right)^{-\beta+1},\ u = \frac{t}{g(t)}$$

$$f(t) = f(0) + \sum_{x=1}^{\infty} \frac{1}{x!} \left(\frac{t}{g(t)}\right)^{x} \left[\left(\frac{d}{dt}\right)^{x-1} \left(g(t)\right)^{x} f'(t)\right] \Bigg|_{t=0}$$

$$\frac{-\log(1-\theta)}{-\log(1-\theta)} = \sum_{x=1}^{\infty} \frac{1}{x!} \frac{\partial^{x-1}}{\partial t^{x-1}} \left[\frac{(1-\theta)^{\beta x-x}\theta}{-\log(1-\theta)}(1-\theta t)^{-\beta x+x-1}\right] \Bigg|_{t=0}$$

or,

$$-\log(1-\theta) = \sum_{x=1}^{\infty} \frac{\theta^{x}}{x!} \frac{\Gamma(\beta x)}{\Gamma(\beta x-x+1)} (1-\theta)^{\beta x-x}$$

$$1 = \sum_{x=1}^{\infty} \frac{\theta^{x}}{x!} \frac{\Gamma(\beta x)}{\Gamma(\beta x-x+1)} \frac{(1-\theta)^{\beta x-x}}{[-\log(1-\theta)]}$$

(iv) <u>Lagrangian Gamma Distribution (LG)</u>

Consider a Lagrangian Poisson process with density function given by (2-34). Define a and θ so that

$$\lambda_1 = at \ , \quad \lambda_2 = a\theta t$$

where $t \geq 0$ is the time between occurrences of the Lagrangian Poisson process, and $a > 0$. With this notation, the cumulative distribution of the waiting time T, the time until the r-th occurrence of the Lagrangian Poisson process is given by,

$$F(t;a,\theta,r) = P(T \leq t)$$

$$= 1 - \sum_{k=0}^{r-1} at[(1+k\theta)at]^{k-1} \frac{e^{-(1+k\theta)at}}{k!} .$$

Differentiation yields the density function of T.

$$f(t;a,\theta,r) = \sum_{k=0}^{r-1} \frac{a[(1+k\theta)at]^{k-1}}{k!} e^{-(1+k\theta)at}[(1+k\theta)at-k] \qquad t \geq 0$$

where $a > 0$; $r = 1,2,\ldots$; $\theta > \frac{-1}{r+1}$

which is known as the Lagrangian Gamma density function [Nelson and Consul, 1974].

Irrespective of the value of θ, the special case $r = 1$ yields the classical exponential distribution

$$f(t;a,\theta,1) = a\, e^{-at} .$$

The special case $\theta = 0$ yields

$$f(t;a,0,r) = \frac{a(at)^{r-1}}{(r-1)!} e^{-at} ,$$ which is the classical Gamma distribution for integer r.

The n-th non-central moment is given by,

$$E(T^n) = \int_0^\infty t^n f(t;a,\theta,r)dt$$

$$= \sum_{k=0}^{r-1} \frac{a^k}{k!} (1+k\theta)^{k-1} \int_0^\infty t^{n+k-1} e^{-(1+k\theta)at}[(1+k\theta)at-k]dt$$

$$E(T^n) = \frac{1}{a^n} \sum_{k=0}^{r-1} \frac{n(n+k-1)!}{k!(1+k\theta)^{n+1}} .$$

For $n = 1$

$$E(T) = \frac{1}{\alpha} \sum_{k=0}^{r-1} \frac{1}{(1+k\theta)^2}$$

For $n = 2$

$$E(T^2) = \frac{2}{\alpha} E(T) + \frac{2(1-\theta)}{\alpha^2} \sum_{k=0}^{r-1} \frac{k}{(1+k\theta)^3}$$

The LG distribution should find wide application in waiting-line models and queuing theory, where the arrival process can be described by a Generalized Poisson (including binomial and negative binomial) Process.

Considering the Lagrangian Binomial process given by

$$P(x;n,\beta,\alpha) = \frac{n\Gamma(n+\beta x)}{x!\ \Gamma(n+\beta x-x+1)} \alpha^x (1-\alpha)^{n+\beta x-x} \qquad x = 0,1,\ldots$$

$$n > 0,\ 0 < \alpha < 1 \text{ and}$$

$$|\beta\alpha| < 1$$

changing α to αt, $t \geq 0$ is the time between occurrences of the Lagrangian Binomial process. With this notation, the cumulative distribution of the random variable T, which is the time until the r-th occurrence of the Lagrangian Binomial process, is given by

$$F(t;n,\alpha,\beta,r) = P(T \leq t)$$

$$= 1 - \sum_{x=0}^{r-1} \frac{n}{n+\beta x} \binom{n+\beta x}{x} (\alpha t)^x (1-\alpha t)^{n+\beta x-x} .$$

Differentiating w.r. to t, we get the density function of T

$$f(t;n,\alpha,\beta,r) = \sum_{x=0}^{r-1} \frac{n}{n+\beta x} \binom{n+\beta x}{x} \alpha(\alpha t)^{x-1} (1-\alpha t)^{n+\beta x-x-1} [\alpha t(n+\beta x)-x]$$

$$0 < t < \frac{1}{\alpha}$$

known as the Lagrangian Beta density function.

The following properties can be easily proved.

If β and α are small, letting $n \to \infty$, such that $n\alpha = c$ (constant), the special case $r = 1$ yields the classical exponential density function:

$$f(t;n,\alpha,\beta,1) = c\,e^{-ct}$$

If $\beta = 1$

$$f(t,n,\alpha,1,r) = \frac{\alpha\Gamma(n+r)}{\Gamma n\,\Gamma r}\,(\alpha t)^{r-1}(1-\alpha t)^{n-1}$$

which is the p.d.f. of Beta distribution. The Lagrangian Beta should then give a closer approximation to the true distribution of many phenomena that can be described by the usual Beta distribution.

Letting $n \to \infty$, with $\alpha = \frac{a}{n}$ and $\beta = n\theta$, the LB can be approximated to

$$F(t;a,\theta,r) = 1 - \sum_{x=0}^{r-1} \frac{1}{x!}\,e^{-(1+x\theta)at}\,at[(1+x\theta)at]^{x-1},$$

which is the cumulative distribution of a Lagrangian gamma random variable.

2.6 APPLICATIONS OF LAGRANGIAN PROBABILITY DISTRIBUTIONS

Occurrence in queuing theory:

The basic approach in the mathematical study of queuing theory is to assume that the arrivals follow a Poisson distribution while servicing follows an exponential or Erlang distribution. If this assumption does not hold, the problem becomes complicated.

Consul and Shenton (1974) have shown that the Lagrangian Distribution provides us with the distribution of the number of customers served in a busy period when the queues are generated by different arrival patterns.

If a single queue is initiated by n customers which have constant service time t and the arrival of more customers, joining the queue on first come first served basis is Bernoullian from a source when units are in use in each interval t, and P is the probability of needing service by a unit during the service interval t of each unit, the probability distribution of the customers served in the First Busy Period (FBP) is Lagrangian. The authors have considered the following situation.

Let us suppose that a number of costly electronic machines, which frequently need exactly the same type of service (constant service time t) are sold by a company in two adjacent cities A and B, which are serviced by one electronic engineer who lives in city A. All the machines (say, m) needing service in city B are brought every day in the early morning to city A and wait for service, which begins when the engineer starts his day's work. All the machines needing service in city A are to be brought to the engineer after each service interval t and are to be serviced in the order of their arrival but after the n machines of city B have been serviced. The FBP of this engineer would fit exactly with this problem. Let P be the probability of any machine in city A needing service during the service interval t of one machine and let a fixed number (m), of machines be in use in the factory. If any machine or machines break down in any service interval t, they are replaced immediately before the next service interval starts so that the work will not suffer. As the service times are constant and the machines breaking down in each interval are independent of the other intervals, the probability distribution of different number of breakdowns during the service period of n machines of city B will be given by the successive terms of the expansion $(q + P)^{mn}$ where $q = 1 - P$.

Though this method of generating a Bernoullian input seems to be somewhat artificial but it is more logical and leads to a more natural generation of the Poisson input (when P is small and m is large such that mP is constant). The distribution of the number of customers (machines) (say, Nn) serviced in the FBP when n is fixed is given by

$$P(Nn=x) = \frac{n}{x}\binom{mx}{x-n} P^{x-n} q^{(m-1)x+n}$$

$$x = n, n+1, \ldots .$$

Thus, the Lagrangian Binomial-Delta distribution is a queuing model.

However, the number of customers, n, which initiates a queue when service begins may be considered binomial, Poisson or negative binomial variate (Consul and Shenton, ibid). Let the probability of the FBP being initiated by n customers be given by $\binom{j}{n} P^n Q^{j-n}$, where P is the probability of a customer needing service in city B which has a total of (j) potential customers and $Q = 1 - P$. Evidently, the probability that the FBP consists of x customers in this case is given by

$$P(X=x) = \frac{j}{mx+1}\binom{mx+1}{x}\left(\frac{Pq}{pQ}\right)\left(Pq^{m-1}\right) Q^{j}\,{}_2F_1\left(1-x, 1-j; mx-x+2; \frac{Pq}{pQ}\right)$$

$$x = 1, 2, \ldots$$

which provides us with the double binomial family of Lagrangian distributions.

The Ballot Problem:

Suppose that, in a ballot, candidate A scores m votes and candidate B scores n votes, where $m > n$. The probability that throughout the counting there are always more votes for A than for B equals

$$\frac{m-n}{m+n} \qquad \text{(Feller, 1968).}$$

The generalized form of the ballot problem can be stated as follows.

If candidate A scores m votes and candidate B scores n votes, where $m > rn+a$, $r \geq 0$ and $a \geq 0$ are integers what is the probability that at each instant A's votes are always greater than $rn+a$.

The probability generating function of this distribution is given by

$$G(t;r,a) = \sum_{n=0}^{\infty} \frac{a}{(r+1)n+a} \binom{(r+1)n+a}{n} p^{rn+a} (1-p)^{n} t^{(r+1)n+a}$$

which is the LBD in a slightly modified form (Mohanty, 1966).

The above presentation has its application in the queuing theory (Takacs, 1962).

Let us suppose that customers arrive at a counter in batches of size m, in accordance with a Poisson process of density λ. The customers are served individually by a single server. The server is idle if and only if there is no customer in the system. The service times are identically distributed, mutually independent random variables with distribution function

$$H(x) = \begin{cases} 1 - e^{-\mu x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

and independent of arrival times. The busy period is defined as the time interval during which the counter is continuously busy. Let f_n be the probability that a busy period consists of nm services. We have

$$f_n = \frac{m}{n(m+1)-1}\binom{n(m+1)-1}{n-1}\left(\frac{\lambda}{\lambda+\mu}\right)^{n-1}\left(\frac{\mu}{\lambda+\mu}\right)^{nm},$$

which gives a LBD.

Biological researches:

In many enumerative studies in Biology referred to by Haight (1967) it has been noticed that the distribution of the random counts are not exactly Poisson law and have been explained by various modifications and generalizations of the Poisson distribution. The Lagrangian Poisson, which has only two parameters and the capability of greater changes in variance than the mean is found to be more suitable in such biometric studies (Consul and Jain, 1973).

CHAPTER III

MODIFIED POWER SERIES DISTRIBUTIONS (MPSD)

3.1 Introduction

Let X be a discrete random variable with probability function

$$P(X=x) = \begin{cases} a(x)\,\dfrac{(g(\theta))^x}{f(\theta)} & x \in T \\ 0 & \text{elsewhere} \end{cases} \tag{3-1}$$

where T is a subset of the set of non-negative integers, $a(x)$; $g(\theta)$ and $f(\theta)$ are positive, finite, and differentiable.

Gupta (1974),(1)) has called (3-1) a MPSD. In case $g(\theta)$ is invertable, it reduces to Patil's (1962) generalized power series distribution and if T is the entire set of non-negative integers it reduces to power series distribution first given by Noack (1950). Thus the GPSD is a sub-class of the MPSD.

3.2 Moments and Cumulants of MPSD

We have,

$$f(\theta) = \sum_{x \in T} a(x)\ (g(\theta))^x .$$

Differentiating w.r. to θ we get;

$$f'(\theta) = \sum_{x \in T} x\ a(x)\ (g(\theta))^{x-1}\ g'(\theta)$$

$$f'(\theta) = \frac{f(\theta)\ g'(\theta)}{g(\theta)}\ E(X)$$

or

$$E(X) = \frac{g(\theta)}{g'(\theta)}\ \frac{d}{d\theta}\ log\ f(\theta) \tag{3-2}$$

Since,

$$\mu_r = \frac{1}{f(\theta)} \sum_{x \in T} (x-\mu_1')^r \, a(x)(g(\theta))^x .$$

Differentiating w.r. to θ we get;

$$g(\theta)\,\frac{d\mu_r}{d\theta} = \frac{g'(\theta)}{g(\theta)} \sum_{x \in T}\left[(x-\mu_1')^r \, a(x)(g(\theta))^x\left\{x - \frac{g(\theta)f'(\theta)}{f(\theta)g'(\theta)}\right\}\right] - rg(\theta)\,\frac{d\mu_1'}{d\theta}\,\mu_{r-1}$$

or,

$$g(\theta)\,\frac{d\mu_r}{d\theta} = g'(\theta)\mu_{r+1} - rg(\theta)\mu_{r-1}\,\frac{d\mu_1'}{d\theta}$$

Thus,

$$\mu_{r+1} = \frac{g(\theta)}{g'(\theta)}\left[\frac{d\mu_r}{d\theta} + r\,\frac{d\mu_1'}{d\theta}\,\mu_{r-1}\right]$$

or

$$\mu_{r+1} = \frac{g(\theta)}{{}'g(\theta)} + r\,\mu_2\mu_{r-1} . \tag{3-3}$$

By differentiating both sides of the identity,

$$\sum_{r=1}^{\infty} \frac{k_r}{r!}\,t^r = log \sum_{u=0}^{\infty} \frac{\mu'_u}{u!}\,t^u$$

with respect to t and identifying the coefficients in t^{r-1}, we get

$$\mu_r' = \sum_{j=1}^{r} \binom{r-1}{j-1} \mu'_{r-j}\,K_j . \tag{3-4}$$

Differentiation of (3-4) w.r. to θ yields

$$\frac{d\mu_r'}{d\theta} = \sum_{j=1}^{r} \binom{r-1}{j-1}\left[\frac{d\mu'_{r-j}}{d\theta}\,K_j + \mu'_{r-j}\,\frac{dK_j}{d\theta}\right] . \tag{3-5}$$

Combining (3-4) and (3-5) and noting that

$$\mu'_{r+1} = \frac{g(\theta)}{g'(\theta)}\,\frac{d\mu_r'}{d\theta} + \mu_r'\,\mu_1' \tag{3-6}$$

we get;

$$\sum_{j=1}^{r+1}\begin{bmatrix} r \\ j-1\end{bmatrix}\mu'_{r+1-j}\;K_j = \frac{g(\theta)}{g'(\theta)}\left[\sum_{j=1}^{r}\binom{r-1}{j-1}\left\{\frac{d\mu'_{r-j}}{d\theta}\,K_j+\mu'_{r-j}\,\frac{dK_j}{d\theta}\right\} + \sum_{j=1}^{r}\binom{r-1}{j-1}\mu'_{r-j}\;K_j\right]\mu'_1$$

$$= \sum_{j=1}^{r}\binom{r-1}{j-1}\left[\left\{\frac{g(\theta)}{g'(\theta)}\;\frac{d\mu'_{r-j}}{d\theta} + \mu'_1\mu'_{r-j}\right\}K_j + \frac{g(\theta)}{g'(\theta)}\mu'_{r-j}\;\frac{dK_j}{d\theta}\right] .$$

Making use of equation (3-6) again, and on simplifying, the $(r+1)$th cumulant will be given by

$$K_{r+1} = \frac{g(\theta)}{g'(\theta)}\;\sum_{j=1}^{r}\binom{r-1}{j-1}\mu'_{r-j}\;\frac{dK_j}{d\theta} - \sum_{j=2}^{r}\binom{r-1}{j-2}\mu'_{r+1-j}\;K_j \qquad (3\text{-}7)$$

3.3 Properties of the MPSD

"The variance of a random variable with the probability function (3-1) is equal to, greater or less than its mean, if and only if $K =, >$ or < 0 respectively and

$$f(\theta) = exp\left\{d + R\int g'(\theta)e^{k\psi(\theta)}\;d\theta\right\},$$

where $\psi'(\theta) = P(\theta)\left[\frac{g'(\theta)}{g\;(\theta)}\right]^2 \Big/ \frac{d}{d\theta}$ log $f(\theta)$, while d and R are some arbitrary constants".

PROOF:

Let the variance of X be equal to {mean + $KP(\theta)$}, where $P(\theta)$ is positive monotonic increasing function of θ, so that the variance is greater than, equal to or less than the mean as $K >, =$ or < 0.

Since,

$$E(X) = \mu = \frac{g(\theta)}{g'(\theta)}\,\frac{d}{d\theta}\,log\;f(\theta) \qquad (3\text{-}8)$$

$$E\{(X-\mu)^2\} = \mu_2 = \frac{g(\theta)}{g'(\theta)}\,\frac{d\mu}{d\theta} \qquad (3\text{-}9)$$

Differentiating (3-8) w.r. to θ, we get,

$$\frac{d\mu}{d\theta} = \frac{g(\theta)}{g'(\theta)} \frac{d^2}{d\theta^2} \log f(\theta) + \left(\frac{d}{d\theta} \log f(\theta)\right)\left(\frac{d}{d\theta} \frac{g(\theta)}{g'(\theta)}\right) .$$

Substituting in (3-9),

$$\mu_2 = \left(\frac{g(\theta)}{g'(\theta)}\right)^2 \left(\frac{d^2}{d\theta^2} \log f(\theta)\right) + \frac{g(\theta)}{g'(\theta)} \frac{d}{d\theta}\left(\frac{g(\theta)}{g'(\theta)}\right) \left(\frac{d}{d\theta} \log f(\theta)\right)$$

$$= \left(\frac{g(\theta)}{g'(\theta)}\right)^2 \left(\frac{d^2}{d\theta^2} \log f(\theta)\right) + \mu \frac{d}{d\theta}\left(\frac{g(\theta)}{g'(\theta)}\right) .$$

By assumption

$$\mu_2 = \mu + KP(\theta) \text{, therefore,}$$

$$\left(\frac{g(\theta)}{g'(\theta)}\right)^2 \frac{d^2}{d\theta^2} \log f(\theta) + \mu \frac{d}{d\theta}\left(\frac{g(\theta)}{g'(\theta)}\right) = \frac{g(\theta)}{g'(\theta)} \frac{d}{d\theta} \log f(\theta) + KP(\theta)$$

or,

$$\frac{\frac{d^2}{d\theta^2} \log f(\theta)}{\frac{d}{d\theta} \log f(\theta)} = \frac{g'(\theta)}{g(\theta)} + \frac{KP(\theta)}{\frac{d}{d\theta} \log f(\theta)} \left(\frac{g'(\theta)}{g(\theta)}\right)^2 - \frac{\mu \frac{d}{d\theta}\left(\frac{g(\theta)}{g'(\theta)}\right)}{\left(\frac{d}{d\theta} \log f(\theta)\right)\left(\frac{g(\theta)}{g'(\theta)}\right)^2}$$

$$= \frac{g'(\theta)}{g(\theta)}\left[1 - \frac{\mu}{\frac{d}{d\theta} \log f(\theta)} \frac{g'(\theta)}{g(\theta)} \cdot \frac{d}{d\theta}\left(\frac{g(\theta)}{g'(\theta)}\right)\right] + K\psi'(\theta)$$

$$= \frac{g'(\theta)}{g(\theta)}\left[1 - \frac{d}{d\theta}\left[\frac{g(\theta)}{g'(\theta)}\right]\right] + K\psi'(\theta)$$

where,

$$\psi'(\theta) = P(\theta)\left(\frac{g'(\theta)}{g(\theta)}\right)^2 \Big/ \frac{d}{d\theta} \log f(\theta) .$$

On integrating the above differential equation we get,

$$\mathrm{Log}\left(\frac{d}{d\theta} \log f(\theta)\right) = \log g(\theta) - \log\left(\frac{g(\theta)}{g'(\theta)}\right) + K\psi(\theta) + c$$

$$= \log g'(\theta) + K\psi(\theta) + c.$$

Thus,

$$\frac{d}{d\theta}(\log f(\theta)) = g'(\theta)exp[c+K\psi(\theta)],$$

which gives on integration and simplification

$$f(\theta) = exp\{d + R \int g'(\theta)e^{K\psi(\theta)}d\theta\} \tag{3-10}$$

The converse of the theorem can be proved very easily.

Corollary:

A MPSD random variable with non-zero probabilities at all non-negative integers must be a Poisson random variable if, and only if the mean equals to variance.

PROOF:

The mean equals the variance when $k = 0$. Thus the above result becomes

$$f(\theta) = exp\{d+Rg(\theta)\} = Ae^{Rg(\theta)} = A \sum_{i=o}^{\alpha} \frac{(Rg(\theta))^i}{i!}, \text{ if } A = e^d.$$

For the modified power series

$$\sum_x P(X=x) = \sum_x a(x) \frac{(g(\theta))^x}{f(\theta)} = 1 \qquad x \in T,$$

and so, by the uniqueness of the series expansion

$$P(X=x) = \frac{AR^x}{x!} \frac{(g(\theta))^x}{Ae^{Rg(\theta)}} = e^{-Rg(\theta)} \frac{(Rg(\theta))^x}{x!}, \qquad x = 0,1,\ldots$$

having the mean $Rg(\theta)$.

However, if we write (3-10) in the following form

$$f(\theta) = exp\{\log f(\theta)\} = \sum_{x=o}^{\infty} \frac{1}{x!}(\log f(\theta))^x$$

the MPSD can always be expressed in terms of Poisson probabilities with mean $\log f(\theta)$.

THEOREM: The MPSD is uniquely determined by its first two moments when they are given as a function of some variable P.

PROOF:

Let the MPSD which determines a distribution be given by $f(\theta)$ and the first two moments be given in the parameter P, where $\theta = h(P)$ is unknown.

Since,

$$\mu_1 = \frac{g(\theta)}{g'(\theta)} \frac{d}{d\theta} \log f(\theta)$$

But
$$\frac{g(\theta)}{g'(\theta)} \frac{d}{d\theta} \log f(\theta) = \frac{g(\theta)}{g'(\theta)} \frac{d}{dP} \log f(\theta) \cdot \frac{dP}{d\theta}$$

$$= \frac{d}{dP} \log f(\theta) \Big/ \frac{g'(\theta)/g(\theta)}{dP/d\theta}$$

$$\mu_1 = \frac{d}{dP} \log y(P) \Big/ \frac{d \log g(\theta)}{dP} \tag{3-11}$$

$$y(P) = f(h(P))$$

i.e. $$\frac{d \log g(\theta)}{dP} = \frac{\frac{d}{dP} \log y(P)}{\mu_1} \tag{3-12}$$

Similarly,

$$\mu_2 = \frac{d\mu_1}{dP} \Big/ \frac{d \log g(\theta)}{dP} \tag{3-13}$$

which can be written as

$$\frac{d \log g(\theta)}{dP} = \frac{d\mu_1}{dP} \Big/ \mu_2 \tag{3-14}$$

On integrating we get

$$\log g(\theta) = \int \left[\left(\frac{d\mu_1}{dP}\right) \Big/ \mu_2\right] dP + c_3$$

$$= \log g_1(P) + \log c_1 \qquad \text{(Say)}$$

i.e. $g(\theta) = c_1 g_1(P)$

Combining (3-14) and (3-12) we get

$$\frac{\frac{d}{dp}\, log\; y(P)}{\mu_1} = \frac{d\mu_1}{dP}\Big/\mu_2 \;.$$

Thus,

$$log\; y(P) = \int\left[\left(\frac{d\mu_1}{dp}\Big/\mu_2\right)\mu_1\right]dP + c_4$$

$$= log\; y_1(P) + log\; c_2 \qquad \text{(Say)}$$

$$\therefore y(P) = c_2 y_1(P) \tag{3-15}$$

where c_1, c_2, c_3 and c_4 are constants.

Let $a(x)$ be the coefficient of $(g_1(P))^x$ in the expansion of $y_1(P)$, then $c_2 c_1^{-x} a(x)$ will be the coefficient of $(g(\theta))^x$ in the expansion of $y(P)$. Hence the probability that the random variable X takes the value x is

$$c_2 c_1^{-x} a(x)\, \frac{(g(\theta))^x}{y(P)} = a(x)\, \frac{(g_1(P))^x}{y_1(P)} \;.$$

We may assume without loss of generality $c_1 = c_2 = 1$

$$\text{i.e.}\quad g(\theta) = g(P) = exp\left\{\left[\int \left(\frac{d\mu_1}{dP}\right)\Big/\mu_2\right]dP\right\}$$

$$f(\theta) = y_1(P) = exp\left\{\left[\int \left(\frac{d\mu_1}{dP}\Big/\mu_2\right)\mu_1\right]dP\right\} \tag{3-16}$$

3.4 Some important MPSD

This class of distributions includes among others, the binomial, the negative binomial, the Poisson, the logarithmic series, the LBD (Chapter II) and their truncated forms. It may be noted that a truncated

MPSD is also a MPSD in its own right.

(i) The LBD and its decapitated form

Let $f(\theta) = (1-\theta)^{-n}$, $0 < \theta < 1$, $g(\theta) = \theta(1-\theta)^{\beta-1}$; $|\theta\beta| < 1$.

The Lagrangian expansion for $f(\theta)$, under the transformation

$$u = \theta/(1-\theta)^{-(\beta-1)}$$

gives

$$P(X=x) = \frac{n\Gamma(n+\beta x)}{x!\,\Gamma(n+\beta x-x+1)} \frac{\left[\theta(1-\theta)^{\beta-1}\right]^x}{(1-\theta)^{-n}}, \quad x = 0,1,2,\ldots \quad (3\text{-}17)$$

and

$$\mu_1 = \frac{n\theta}{1-\beta\theta}$$

$$\mu_2 = \frac{n\theta(1-\theta)}{(1-\beta\theta)^3} .$$

Conversely, if μ_1 and μ_2 are given we have from (3-16)

$$\frac{d\mu_1}{d\theta} = \frac{n(1-\beta\theta) + n\theta\beta}{(1-\beta\theta)^2} = \frac{n}{(1-\beta\theta)^2}$$

$$g(\theta) = exp\left\{\int \frac{n(1-\beta\theta)^3}{(1-\beta\theta)^2 n\theta(1-\theta)}\, d\theta\right\}$$

$$= exp\left\{\int \frac{1-\beta\theta}{\theta(1-\theta)}\, d\theta\right\} = e^I$$

where,

$$I = \int \frac{1-\beta\theta}{\theta(1-\theta)}\, d\theta = \int\left\{\frac{1}{\theta} + \frac{1}{1-\theta} - \beta\left(\frac{1}{1-\theta}\right)\right\} d\theta$$

$$= log\ \theta - log(1-\theta) + \beta\ log(1-\theta)$$

$$= log\ \theta(1-\theta)^{\beta-1}$$

$$\therefore\ g(\theta) = \theta(1-\theta)^{\beta-1}$$

And;

$$f(\theta) = exp\left\{\int \frac{n}{1-\theta}\, d\theta\right\} = (1-\theta)^{-n}$$

which is the series function of the LBD.

Now let $f(\theta) = (1-\theta)^{-n} - 1$ and $g(\theta) = \theta(1-\theta)^{\beta-1}$ under the same transformation given in (3-17), the Lagrangian expansion of the function $f(\theta)$, gives

$$p_0(X=x) = \frac{n\Gamma(n+\beta x)}{x!\Gamma(n+\beta x-x+1)} \frac{[\theta(1-\theta)^{\beta-1}]^x}{(1-\theta)^{-n} - 1} \quad , \quad x = 1,2,3,\ldots$$

which is the p.d.f. of decapitated LBD.

The mean of the the decapitated LBD.

$$mean = \mu_1 = \frac{n\theta}{(1-\beta\theta)[1-(1-\theta)^n]}$$

and the variance,

$$\mu_2 = \frac{n\theta(1-\theta)}{(1-\beta\theta)^3[1-(1-\theta)^n]} - \frac{n^2\theta^2(1-\theta)^n}{(1-\beta\theta)^2[1-(1-\theta)^n]^2}$$

This reduces for $\beta = 0$, to

$$\frac{n\theta(1-\theta)}{1-(1-\theta)^n} - \frac{n^2\theta^2(1-\theta)^n}{[1-(1-\theta)^n]^2}$$

which is the variance of a decapitated binomial distribution as obtained by Stephen (1945).

Again the particular case $\beta = 1$, gives the variance of a decapitated negative binomial distribution

$$\frac{n\theta}{(1-\theta)^2[1-(1-\theta)^n]} - \frac{n^2\theta^2(1-\theta)^{n-2}}{[1-(1-\theta)^n]^2}$$

Stephen (1945) has explained why it may be useful to know the mean and variance of the reciprocal of a variance. Mendenhall and Lehman (1960) have shown their importance in estimation theory. In the following, we shall be interested in the inverse moments of a decapitated LBD. Taking $r = -1$ in equation (3-6), we obtain

$$\mu'_o = \frac{g(\theta)}{g'(\theta)} \frac{d\mu'_{-1}}{d\theta} + \mu'_{-1}\mu'_1$$

or

$$1 = \frac{\theta(1-\theta)}{1 - \beta\theta} \frac{d\mu'_{-1}}{d\theta} + \frac{n\theta}{(1-\beta\theta)[1-(1-\theta)^n]} \mu'_{-1} .$$

Letting $\mu'_{-1} = y$, we get the linear differential equation

$$\frac{dy}{d\theta} + \frac{n}{(1-\theta)[1-(1-\theta)^n]} y = \frac{1 - \beta\theta}{\theta(1-\theta)} . \qquad (3\text{-}18)$$

The integrating factor is,

$$exp\left[\int \frac{n}{(1-\theta)[1-(1-\theta)^n]} d\theta\right] = \frac{1 - (1-\theta)^n}{(1-\theta)^n}$$

Hence the solution of (3-18) is given by

$$y \cdot \frac{1 - (1-\theta)^n}{(1-\theta)^n} = \int_0^\theta \frac{1 - (1-\theta)^n}{(1-\theta)^n} \frac{1 - \beta\theta}{\theta(1-\theta)} d\theta \qquad (3\text{-}19)$$

The integrand can be written as

$$\frac{(1-\beta) + \beta(1-\theta)}{\theta(1-\theta)^{n+1}} [1 - (1-\theta)^n]$$

$$= \frac{1 - \beta}{\theta(1-\theta)^{n+1}} + \frac{\beta}{\theta(1-\theta)^n} - \frac{1 - \beta}{\theta(1-\theta)} - \frac{\beta}{\theta}$$

$$= (1-\beta)\left[\frac{1}{\theta} + \frac{1}{1-\theta} + \cdots \frac{1}{(1-\theta)^{n+1}}\right] + \beta\left[\frac{1}{\theta} + \frac{1}{1-\theta} + \cdots \frac{1}{(1-\theta)^{n}}\right]$$

$$- (1-\beta)\left[\frac{1}{\theta} + \frac{1}{1-\theta}\right] - \frac{\beta}{\theta}$$

$$= \frac{1}{(1-\theta)^2} + \cdots + \frac{1}{(1-\theta)^{n-1}} + \beta\left(\frac{1}{1-\theta} - \frac{1}{(1-\theta)^{n+1}}\right)$$

Therefore equation (3-19) can be written as

$$y \cdot \frac{1-(1-\theta)^n}{(1-\theta)^n} = \left\{\frac{1}{1-\theta} + \frac{1}{2(1-\theta)^2} + \cdots + \frac{1}{n(1-\theta)^n} - \beta\left[\log(1-\theta) + \frac{1}{n(1-\theta)^n}\right]\right\}\Bigg|_0^{\theta}$$

$$= \sum_{\rho=1}^{n} \frac{1}{\rho(1-\theta)^{\rho}} - \beta\left[\frac{1}{n(1-\theta)^n} + \log(1-\theta)\right] - \sum_{\rho=1}^{n} \frac{1}{\rho} + \frac{\beta}{n}$$

This gives

$$y = E\left(\frac{1}{X}\right) = \frac{(1-\theta)^n}{1-(1-\theta)^n} \sum_{\rho=1}^{n} \frac{1}{\rho}\left[\frac{1}{(1-\theta)^{\rho}} - 1\right] - \frac{\beta}{n} - \frac{(1-\theta)^n}{1-(1-\theta)^n}\,\beta \log(1-\theta) \qquad (3\text{-}20)$$

Again using equation (3-6) with $r = -2$, we obtain

$$\mu'_{-1} = \frac{g(\theta)}{g'(\theta)}\,\frac{d\mu'_{-2}}{d\theta} + \mu'_{-2}\mu'_1 \; .$$

Letting $\mu'_{-2} = z$ and using the value of μ'_{-1} as obtained in equation (3-20) we get

$$\frac{dz}{d\theta} + \frac{n}{(1-\theta)[1-(1-\theta)^n]}\, z = \frac{(1-\beta\theta)(1-\theta)^{n-1}}{\theta[1-(1-\theta)^n]}\left[\sum_{\rho=1}^{n} \frac{1}{\rho}\left(\frac{1-(1-\theta)^{\rho}}{(1-\theta)^{\rho}}\right)\right.$$

$$\left. - \frac{\beta}{n}\left(\frac{1-(1-\theta)^n}{(1-\theta)^n}\right) - \beta \log(1-\theta)\right] ,$$

and the solution is given by:

$$z\,\frac{1-(1-\theta)^n}{(1-\theta)^n} \doteq \int_0^\theta \frac{1-\beta\theta}{\theta(1-\theta)}\left[\sum_{\rho=1}^{n}\frac{1}{\rho}\left(\frac{1-(1-\theta)^\rho}{(1-\theta)\rho}\right)-\frac{\beta}{n}\left(\frac{1-(1-\theta)^n}{(1-\theta)^n}\right)-\beta\,log(1-\theta)\right]d\theta$$

We have already shown that the decapitated LBD reduces to decapitated binomial distribution for $\beta = 0$ and to decapitated negative binomial distribution for $\beta = 1$. We shall now obtain the distribution of $z = X_1 + X_2 + \cdots X_k$. And we shall start by modifying Patil's (1963) result as follows: Let X be a discrete random variable having the probability function

$$P(X{=}x) = a(x)\,\frac{(g(\theta))^x}{f(\theta)}\,,$$

and let $X_1, X_2, \ldots X_K$ be a random sample of size K from this distribution. Then the distribution of $z = \sum_{i=1}^{K} X_i$ is given by,

$$P(Z{=}z) = b(z,K)\,\frac{(g(\theta))^z}{(f(\theta))^K}$$

where

$$(f(\theta))^K = \sum_z b(z,K)(g(\theta))^z$$

For the decapitated LBD we have

$$g(\theta) = \theta(1-\theta)^{\beta-1}$$

$(f(\theta))^K = [(1-\theta)^{-n} - 1]^K$, which can be expanded using Lagrange's formula under the transformation

$$u = \frac{\theta}{h(\theta)}\,,\ \text{where } h(\theta) = (1-\theta)^{-\beta+1}$$

$$(f(\theta))^K = \sum_{z=K}^{\infty} \frac{1}{z!} \frac{d^{z-1}}{d\theta^{z-1}} (1-\theta)^{-\beta z+z} Kn(1-\theta)^{-n-1}[(1-\theta)^{-n}-1]^{K-1} \Big|_{\theta=0} \theta^z(1-\theta)^{(\beta-1)z}$$

$$= \sum_{z=K}^{\infty} \frac{Kn}{z!} \frac{d^{z-1}}{d\theta^{z-1}} (1-\theta)^{-\beta z+z-nK-1}[1-(1-\theta)^n]^{K-1} \Big|_{\theta=0} (\theta(1-\theta)^{\beta-1})^z$$

$$= \sum_{z=K}^{\infty} \frac{Kn}{z!} \frac{d^{z-1}}{d\theta^{z-1}}\left[\sum_{r=0}^{K-1} \binom{K-1}{r} (-1)^r (1-\theta)^{-\beta z+z-nK+nr-1}\right] \Big|_{\theta=0} (\theta(1-\theta)^{\beta-1})^z$$

$$= \sum_{z=K}^{\infty} nK\left[\sum_{r=0}^{K-1} \binom{K-1}{r} (-1)^r \frac{(\beta z+nK-nr-1)!}{z!(\beta z-z+nK-nr)!}\right] (\theta(1-\theta)^{\beta-1})^z$$

$$= \sum_{z=K}^{\infty} Kn\left[\sum_{r=0}^{K-1} \binom{K-1}{r} (-1)^r \frac{\binom{\beta z+nK-nr}{z}}{\beta z+nK-nr}\right] (\theta(1-\theta)^{\beta-1})^z$$

$$= \sum_{z=K}^{\infty}\left[\sum_{r=0}^{K} \frac{(-1)^r\binom{K}{r}n(K-r)}{(\beta z+nK-nr)} \binom{\beta z+nK-nr}{z}\right] (\theta(1-\theta)^{\beta-1})^z$$

Replacing $K-r$ by S in the last equation we get;

$$(f(\theta))^K = \sum_{z=K}^{\infty}\left[\sum_{S=0}^{K} \frac{(-1)^{K-S}\binom{K}{K-S}nS}{\beta z+nS} \binom{\beta z+nS}{z}\right] (\theta(1-\theta)^{\beta-1})^z$$

$$= \sum_{z=K}^{\infty}\left[\sum_{S=1}^{K} (-1)^{K-S}\frac{\binom{K}{S}nS}{\beta z+nS-z} \binom{\beta z+nS-1}{z}\right] (\theta(1-\theta)^{\beta-1})^z$$

Thus,

$$P(Z=z) = \sum_{S=1}^{K} \frac{nS(-1)^{K-S}\binom{K}{S}}{(\beta z+nS-z)} \binom{\beta z+ns-1}{z} \frac{(\theta(1-\theta)^{\beta-1})^z}{[(1-\theta)^{-n}-1]^K} \qquad (3\text{-}21)$$

For $\beta = 1$ equation (3-21) gives the same result obtained by Ahuja (1971), and for $\beta = 0$ we get the same result obtained by Ahuja (1969).

(ii) The Lagrangian Logarithmic Distribution (LLD)

In the probability function of the decapitated LBD, if n is increased without limit, then the limiting form of the distribution is LLD, whose probability function is given by;

$$P(X=x) = \frac{\Gamma(\beta x)}{x\Gamma x\Gamma(\beta x-x+1)} \frac{[\theta(1-\theta)^{\beta-1}]^{x}}{[-log(1-\theta)]}, \text{ where}$$

$$g(\theta) = \theta(1-\theta)^{\beta-1} \quad \text{and} \quad f(\theta) = -log(1-\theta), \quad x = 1,2,\ldots$$

This distribution gets reduced to the logarithmic series distribution when $\beta = 1$.

The first four crude moments of the distribution are given by:

$$\mu_1 = \frac{a\theta}{1-\beta\theta}$$

$$\mu_2' = \frac{a\theta(1-\theta)}{(1-\beta\theta)^3}$$

$$\mu_3' = \frac{a\theta(1-\theta)[1-2\theta+\beta\theta(2-\theta)]}{(1-\beta\theta)^5}$$

$$\mu_4' = \frac{a\theta[(1-4\theta)(1-3\theta) + 2\beta\theta(1-\theta)(4-9\theta+4\theta^2)\beta^2\theta^2(1-\theta)(6-6\theta+\theta^2)]}{(1-\beta\theta)^7}$$

where $a^{-1} = [-log(1-\theta)]$

(iii) The Lagrangian Poisson Distribution (LPD)

The probability density function of the LPD in a modified form is given by

$$P(X=x) = \frac{(1+\lambda x)^{x-1}}{x!} \frac{(\theta e^{-\lambda\theta})^{x}}{e^{\theta}}, \quad x = 0,1,2,\ldots \tag{3-22}$$

$\theta > 0 \quad 0 < \lambda\theta < 1$ and λ is a known constant.

In this case, $g(\theta) = \theta e^{-\lambda\theta}$ and $f(\theta) = e^{\theta}$

$$E(X) = \frac{\theta}{(1 - \theta\lambda)}$$

$$E(X^2) = \frac{\theta}{(1 - \theta\lambda)^3} + \frac{\theta^2}{(1 - \theta\lambda)^2}$$

$$\mathit{Variance} = \frac{\theta}{(1 - \theta\lambda)^3}$$

A random variable X is said to have a decapitated LPD, which is also MPSD if its probability density function is given by:

$$P(X=x) = \frac{(1 + \lambda x)^{x-1}}{x!} \frac{(\theta e^{-\lambda\theta})^{x}}{(e^{\theta} - 1)} \qquad x = 1,2,\ldots \qquad (3\text{-}23)$$

where $f(\theta) = (e^{\theta} - 1)$ and $g(\theta) = \theta e^{-\lambda\theta}$.

The moments of the distribution can be calculated using the recurrence relation given by (3-6).

Now let $X_1, X_2, \ldots X_K$ be a random sample taken from the distribution (3-23). We shall find the distribution of the sum, $Z = \sum_{i=1}^{K} X_i$. Using the same technique as in the case of the decapitated LBD.

$$(f(\theta))^K = (e^{\theta} - 1)^K \quad \text{and} \quad h(\theta) = e^{\lambda\theta}$$

with the transformation $u = \frac{\theta}{h(\theta)}$, the Lagrangian expansion of $(f(\theta))^K$ is given by:

$$(f(\theta))^K = \sum_{z=K}^{\infty} \frac{1}{Z!} D_{\theta}^{z-1} \left\{ e^{\lambda\theta z} D_{\theta} (e^{\theta} - 1)^K \right\} \Big|_{\theta=o} (\theta e^{-\lambda\theta})^z$$

$$= \sum_{z=k}^{\infty} \frac{1}{z!} D_{\theta}^{z-1} \left\{ e^{\lambda\theta z} D_{\theta} \left(\sum_{n=k}^{\infty} \frac{K!}{n!} S_n^K \, \theta^n \right) \right\} \Big|_{\theta=o} (\theta e^{-\lambda\theta})^z$$

where S_n^K is the Stirling number of the second kind given by Jordan (1960)

$$(f(\theta))^K = \sum_{z=K}^{\infty} \frac{1}{z} D_\theta^{z-1}\left\{e^{\lambda\theta z} \sum_{n=k}^{\infty} \frac{K!}{n!} S_n^K\, n\, \theta^{n-1}\right\}\Bigg|_{\theta=o} (\theta e^{-\lambda\theta})^z$$

Put $n - 1 = m$

$$(f(\theta))^K = \sum_{z=K}^{\infty} \frac{1}{z!}\left\{\sum_{m=k-1}^{\infty} \frac{k!}{m!} S_{m+1}^K D_\theta^{z-1} (\theta^m e^{\lambda\theta z})\right\}\Bigg|_{\theta=o} (\theta e^{-\lambda\theta})^z \tag{3-24}$$

Leibnitz's applying rule for the higher derivatives of the product of two functions we get;

$$(f(\theta))^K = \sum_{z=k}^{\infty} \frac{1}{z!}\left\{\sum_{m=k-1}^{\infty} \frac{K!}{m!} S_{m+1}^K \left[\sum_{t=o}^{z-1} \binom{z-1}{t} D_\theta^t \theta^m D_\theta^{z-1-t} e^{\lambda\theta z}\right]\right\}\Bigg|_{\theta=o} \times (\theta e^{-\lambda\theta})^z$$

$$= \sum_{z=k}^{\infty} \frac{K!}{z!}\left\{\sum_{t=o}^{z-1} \binom{z-1}{t} (\lambda z)^{z-1-t} S_{t+1}^K\right\}(\theta e^{-\lambda\theta})^z$$

Thus we obtain the required distribution as

$$P(Z=z) = \frac{K!}{z!} \frac{\left\{\sum_{t=o}^{z-1} \binom{z-1}{t} (\lambda z)^{z-1-t}\ S_{t+1}^K\right\}}{(e^\theta - 1)^K} (\theta e^{-\lambda\theta})^z$$

If $\lambda = 0$, equation (3-24) can be written to the form,

$$(f(\theta))^K = \sum_{z=K}^{\infty} \frac{K!}{z!} S_z^K \theta^z,$$

which gives the sum of independent decapitated Poisson variable (Tate and Goen, 1958).

3.5 The relation between MPSD and LD.

The following theorem will establish the realtion between the class of MPSD and the class of LD.

THEOREM: An MPSD is a subclass of Lagrangian probability distributions but the converse is not true.

PROOF: Let $f(t,\theta)$ be a p.g.f., such that its series expansion in powers of t takes the form

$$f(t,\theta) = \sum_{i=k}^{e} c(\theta)t^{i}c_{i}(\theta) \tag{3-26}$$

where the integers K and e are in the domain $0 \leq K \leq e \leq \infty$.

Let $g(t,\theta)$ be another p.g.f. such that

$$g(t,\theta) = \sum_{s=0}^{m} tb(\theta)b_{s}(\theta), \tag{3-27}$$

where m is a positive integer. Now

$$(g(t,\theta))^{x} = (b(\theta))^{x}\sum_{s=0}^{mx} t^{s}d_{s}(\theta), \tag{3-28}$$

where $d_s(\theta)$ is non-negative function of θ and s.

The Lagrangian expansion of $f(t,\theta)$ under the transformation $t = ug(t)$, is given by

$$f(t,\theta) = \sum_{j=k}^{\infty} \frac{U^{j}}{j!}\frac{\partial^{j-1}}{\partial t^{j-1}}\left[(g(t,\theta))^{x}f'(t,\theta)\right]\Big|_{t=0}$$

so that the Lagrangian p.d.f. becomes

$$P(X=x) = \frac{1}{x!}\frac{\partial^{x-1}}{\partial t^{x-1}}(b(\theta))^{x}\left\{\sum_{s=0}^{mx} t^{s}d_{s}(\theta)\right\}c(\theta)\left[\sum_{i=k}^{e} it^{i-1}c_{i}(\theta)\right]\Bigg|_{t=0}$$

Putting $s + i = n$

$$P(X=x) = c(\theta)\,\frac{(b(\theta))^x}{x!}\,\frac{\partial^{x-1}}{\partial t^{x-1}}\left[\sum_{n=k}^{e+mx} t^{n-1}\left\{\sum_{i=k}^{z} ic_i(\theta)d_{n-i}(\theta)\right\}\right]\Bigg|_{t=o},$$

where $z = \min(n,e)$

$$P(X=x) = c(\theta)\frac{(b(\theta))^x}{x!}\,(x-1)!\left[\sum_{i=k}^{z} ic_i(\theta)d_{x-i}(\theta)\right] \qquad (3\text{-}29)$$

$x = K, K+1, \ldots$

$z = \min(e,x)$.

The above is the general form of a Lagrangian p.d.f. where $\left\{\sum_{i=k}^{z} ic_i(\theta)d_{x-i}(\theta)\right\}$ is generally not of the form $\alpha(x)(h(\theta))^x$. Thus, a Lagrangian p.d.f. will not always be an MPSD. However, when

$$c_i(\theta) = \phi_1(i)\phi_2(\theta)(h(\theta))^i$$

$$d_{x-i}(\theta) = \phi_3(x-i)\phi_4(\theta)(h(\theta))^{x-i}$$

the Lagrangian p.d.f. (3-29) takes the form

$$P(X=x) = \phi_2(\theta)\phi_4(\theta)c(\theta)(b(\theta)h(\theta))^x\left[\frac{1}{x}\sum_{i=k}^{z} i\phi_1(i)\phi_3(x-i)\right],$$

$x = K, K+1, \ldots$

$z = \min(e,x)$.

Setting

$$b(\theta)h(\theta) = g(\theta)$$

$$\phi_2(\theta)\phi_4(\theta)c(\theta) = \frac{1}{f(\theta)}, \text{ and}$$

$$\frac{1}{x}\sum_{i=k}^{z} i\phi_1(i)\phi_3(x-i) = a(x)$$

$$\therefore P(X=x) = a(x)\,\frac{(g(\theta))^x}{f(\theta)},$$

Thus an MPSD is a subclass of Lagrangian distributions. The following are two examples of Lagrangian p.d. which do not belong to the class of MPSD.

Ex. 1. $g(t) = e^{\theta(t-1)}$, $f(t) = (q+Pt)^n$ $\theta < 1$

and $P(X=x) = \frac{(\theta x)^{x-1}}{x!}\, e^{-\theta x}\, nPq^{n-1}\ {}_2F_0(1-x,\ 1-n;\ \frac{P}{\theta qx})$

(Consul & Shenton, 1972)

Ex. 2. $g(t) = (Q-Pt)^{-k}$, $f(t) = e^{M(t-1)}$ $KP < 1$

and $P(X=x) = \frac{e^{-M}M^x}{x!}\, Q^{-Kx}\ {}_2F_0(1-x,\ Kx;\ -\frac{P}{MQ})$.

CHAPTER IV

ESTIMATION FOR AN MPSD

In this chapter we introduce two methods for estimating the parameter θ involved in the p.d.f. given by (3-1). The first ~~method~~ is the maximum likelihood method (Gupta, 1975) and the other one is the minimum variance unbiassed estimation.

4.1 MAXIMUM LIKELIHOOD ESTIMATION (MLE) OF A MPSD

Let $X_1, X_2, \ldots X_n$ be a random sample of size n from the MPSD given by (3-1), and let L be the likelihood function. The efficient score for θ is,

$$\gamma(\theta) = \frac{1}{L}\frac{\partial L}{\partial \theta} = n\frac{g'(\theta)}{g(\theta)}\,[\bar{x} - \mu(\theta)] \tag{4-1}$$

where, $\mu(\theta) = E(X) = \dfrac{g(\theta)}{g'(\theta)}\dfrac{d}{d\theta}\log f(\theta)$, and $\bar{x}$ is the sample mean.

The solution of the likelihood equation $\gamma(\theta) = 0$ is therefore, given by

$$\bar{x} = \mu(\hat{\theta}) \tag{4-2}$$

The method of maximum likelihood and the method of moments lead to the same estimate in the case of a MPSD.

The amount of information, $I(\theta)$ contained in the sample is given by

$$I(\theta) = -E\left\{\frac{\partial^2}{\partial\theta^2}\log L\right\} = n\left(\frac{g'(\theta)}{g(\theta)}\right)^2 \mu_2(\theta) \tag{4-3}$$

where $\mu_2(\theta) = \dfrac{g(\theta)}{g'(\theta)}\dfrac{d\mu(\theta)}{d\theta}$, is the variance of a MPSD.

$$\mathrm{var}\,(\hat{\theta}) = \{I(\theta)\}^{-1} = \frac{1}{n}\,\frac{g(\theta)}{g'(\theta)} \Big/ \frac{d\mu(\theta)}{d\theta} \tag{4-4}$$

The solution of equation (4-2) for $\hat{\theta}$ is given by,

$$\hat{\theta} = \mu^{-1}(\bar{x}) \tag{4-5}$$

provided that $\mu(\theta)$ is invertable. The inversion of $\mu(\theta)$ can be justified at least when $g(\theta)$ is increasing and hence $\mu(\theta)$ is an increasing function of θ. If, however, (4-2) does not readily give an algebraic solution, one may use an iterative process of solution by starting with an approximation θ_0. An improved approximation θ_1 is then obtained from

$$\theta_1 = \theta_0 + \frac{\gamma(\theta_0)}{I(\theta_0)} = \theta_0 + \frac{g(\theta_0)}{g'(\theta_0)}\,\frac{\bar{x}-\mu(\theta_0)}{\mu_2(\theta_0)} \tag{4-6}$$

and this process is repeated until one obtains a sufficiently accurate solution (Rao, 1952).

The amount of bias $b(\hat{\theta})$ of the MLE (Haldane and Smith, 1956) is given by,

$$b(\hat{\theta}) = -\frac{1}{2n}\left(\frac{B_1}{A_1^{\,2}}\right) \tag{4-7}$$

where,

$$A_1 = \sum_{x \in T} \frac{1}{P_x}\left(\frac{dP_x}{d\theta}\right)^2 \tag{4-8}$$

and,

$$B_1 = \sum_{x \in T} \frac{1}{P_x}\left(\frac{dP_x}{d\theta}\right)\left(\frac{d^2P_x}{d\theta^2}\right) \tag{4-9}$$

For P_x given by (3-1), it may be verified that

$$A_1 = \left(\frac{g'(\theta)}{g(\theta)}\right)^2 \mu_2 \tag{4-10}$$

$$B_1 = \left(\frac{g'(\theta)}{g(\theta)}\right)^3 \left[\mu_3 + \frac{g(\theta)g''(\theta) - \{g'(\theta)\}^2}{\{g'(\theta)\}^2}\,\mu_2\right] \tag{4-11}$$

Thus,

$$b(\hat{\theta}) = -\frac{1}{2n\mu_2^2}\,\frac{g(\theta)}{g'(\theta)}\left[\mu_3 + \frac{g(\theta)g''(\theta) - \{g'(\theta)\}^2}{\{g'(\theta)\}^2}\,\mu_2\right] \qquad (4\text{-}12)$$

where $\mu_3 = \frac{g(\theta)}{g'(\theta)}\,\frac{d\mu_2}{d\theta}$, is the third central moment of the MPSD.

Gupta (1975) applied the method of maximum likelihood in the following cases:

(i) The LBD and its decapitated form

The p.d.f. of a LBD is given by

$$P(X=x) = \frac{k\Gamma(k+\beta x)}{x!\,\Gamma(k+\beta x-x+1)}\;\frac{\{\theta(1-\theta)^{\beta-1}\}^x}{(1-\theta)^{-k}}\;, \qquad x = 0,1,2,\ldots$$

$$0 < \theta < 1$$

$$|\theta\beta| < 1$$

Here $g(\theta) = \theta(1-\theta)^{\beta-1}$, $f(\theta) = (1-\theta)^{-k}$ and $\mu(\theta) = \frac{k\theta}{1-\beta\theta}$.

Equation (4-2), therefore, gives

$$\bar{x} = k\hat{\theta}/1-\beta\hat{\theta} \qquad \text{or} \qquad \hat{\theta} = \frac{\bar{x}}{k+\beta\bar{x}}$$

For binomial distribution $\beta = 0$ and hence $\hat{\theta} = \frac{\bar{x}}{k}$. For negative binomial distribution $\beta = 1$ and hence $\hat{\theta} = \frac{\bar{x}}{k+\bar{x}}$. From equation (4-4), var($\hat{\theta}$) becomes

$$\mathrm{var}(\hat{\theta}) = \frac{\theta(1-\theta)(1-\beta\theta)}{n} .$$

In particular, for the binomial, var($\hat{\theta}$) $= \frac{\theta(1-\theta)}{n}$ and for the negative binomial distribution, var($\hat{\theta}$) $= \frac{\theta(1-\theta)^2}{n}$. The bias of $\hat{\theta}$ is obtained from equation (4-12) as

$$b(\hat{\theta}) = -\beta\theta(1-\theta)/nk .$$

For the decapitated LBD given by

$$P_0(X=x) = \frac{k\Gamma(k+\beta x)}{x!\Gamma(k+\beta x-x+1)} \cdot \frac{\left(\theta(1-\theta)^{\beta-1}\right)^x}{\left((1-\theta)^{-k}-1\right)} , \qquad x = 1,2,3,\ldots$$

$$\mu(\theta) = \frac{k\theta}{(1-\beta\theta)[1-(1-\theta)^k]} .$$

Therefore, $\hat{\theta}$, the MLE of θ is given by

$$\bar{x} = \frac{k\hat{\theta}}{(1-\beta\hat{\theta})[1-(1-\hat{\theta})^k]} .$$

This reduces, for the zero truncated binomial distribution, to

$$\bar{x} = \frac{k\hat{\theta}}{1-(1-\hat{\theta})^k} ,$$

which agrees with Patil (1962), and for the zero truncated negative binomial distribution to

$$\bar{x} = \frac{k\hat{\theta}}{(1-\hat{\theta})[1-(1-\hat{\theta})^k]} .$$

The solution can be found by the iterative method.

(ii) The LP and its decapitated form

The p.d.f. of the LPD is given by

$$P(X=x) = \frac{(1+\lambda x)^{x-1}}{x!} \quad \frac{(\theta e^{-\lambda\theta})^x}{e^\theta} , \qquad x = 0,1,2,\ldots$$

$$\theta > 0$$

$$0 < \theta\lambda < 1$$

Here $g(\theta) = \theta e^{-\lambda\theta}$, $f(\theta) = e^\theta$ and $\mu(\theta) = \frac{\theta}{1-\lambda\theta}$.

Equation (4-2), therefore gives

$$\hat{\theta} = \frac{\bar{x}}{1+\lambda\bar{x}}$$

For the Poisson distribution $\lambda = 0$ and hence $\hat{\theta} = \bar{x}$.

Equation (4-4) gives, var $(\hat{\theta})$ as

$$\text{var}\ (\hat{\theta}) = \theta(1-\theta\lambda)/n \quad \text{and from (4-12)}$$

$$b(\hat{\theta}) = -\theta\lambda/n\ .$$

The p.d.f. of the decapitated LPD is given by

$$P_0(X=x) = \frac{(1+\lambda x)^{x-1}}{x!} \quad \frac{(\theta e^{-\lambda\theta})^x}{e^{\theta}(1-e^{-\theta})} \qquad x = 1,2,\ldots$$

Here $g(\theta) = \theta e^{-\lambda\theta}$, $f(\theta) = e^{\theta}-1$, and $\mu(\theta) = \dfrac{\theta}{(1-\theta\lambda)(1-e^{-\theta})}$.

Equation (4-2), therefore, yields

$$\bar{x} = \frac{\hat{\theta}}{(1-\hat{\theta}\lambda)(1-e^{-\hat{\theta}})}$$

For the Poisson distribution, the last equation reduces to

$$\bar{x} = \frac{\hat{\theta}}{1-e^{-\hat{\theta}}}\ .$$

This equation has been solved by Barton, David and Marrington (1960).

(iii) The MLE for the Lagrangian Logarithmic Distribution (LLD)

As has been mentioned in Chapter (III) the LLD can be obtained from the decapitated LBD, when k is infinitely large. The p.d.f. is given by

$$P(X=x) = \frac{\Gamma(x\beta)}{x\Gamma(x)\Gamma(x\beta-x+1)} \quad \frac{\{\theta(1-\theta)^{\beta-1}\}^x}{[-\log(1-\theta)]} \qquad , \ x = 1,2,\ldots$$

This conforms to the logarithmic series distribution for $\beta = 1$. Here $g(\theta) = \theta(1-\theta)^{\beta-1}$, $f(\theta) = -\log(1-\theta)$ and $\mu(\theta) = \frac{a\theta}{1-\theta\beta}$ where $a = [-\log(1-\theta)]^{-1}$.

Equation (4-2) thus gives

$$\bar{x} = \frac{\hat{\theta}}{[-\log(1-\hat{\theta})][1-\beta\hat{\theta}]} .$$

Equation (4-4) gives $\text{var}(\hat{\theta})$ as

$$\text{var}(\hat{\theta}) = \frac{\theta(1-\theta)^2(1-\theta\beta)}{na(1-\theta-a\theta+a\beta\theta^2)} .$$

4.2 MINIMUM VARIANCE UNBIASSED ESTIMATION FOR MPSD

We shall now derive a uniformly minimum variance unbiassed estimator for the function $g(\theta)$, assuming that θ is the only unknown parameter involved in the given probability distribution..

Let $X_1, X_2, \ldots X_n$ be a random sample taken from a distribution that has a p.d.f. given by:

$$P(X=x) = a(x)\,\frac{(g(\theta))^x}{f(\theta)} \qquad x \in S \tag{4-13}$$

$\theta > 0$, $a(x) > 0$ does not involve θ, $f(\theta) = \sum_{x \in S} a(x)\,(g(\theta))^x$ and S is a countable subset of the set of non-negative integers. Without loss of generality, we shall assume $a(0) = 1$.

<u>Lemma</u>: $T = \sum_{i=1}^{n} x_i$ is a complete sufficient statistic for $g(\theta)$ in the sense of Lehmann-Scheffé and Rao Blackwell theorem.

As has been shown in Chapter (III), the statistic T has a p.d.f. given by:

$$P(T=t) = b(t,n)\,\frac{(g(\theta))^t}{(f(\theta))^n}$$

where $b(t,n)$ is the coefficient of $(g(\theta))^t$ in the expansion of $(f(\theta))^n$.

Define for any positive integer r

$$\begin{aligned} U_r(t) &= 0 && t < r \\ &= \frac{b(t-r,n)}{b(t,n)} && t \geq r \end{aligned}$$

$$\begin{aligned} E[U_r(T)] &= \sum_{t=r}^{\infty} \frac{b(t-r,n)}{b(t,n)}\; b(t,n)\; \frac{(g(\theta))^t}{(f(\theta))^n} \\ &= (g(\theta))^r \sum_{t=r}^{\infty} b(t-r,n)\; \frac{(g(\theta))^{t-r}}{(f(\theta))^n} \\ &= (g(\theta))^r . \end{aligned}$$

Thus, $U_1(T)$ is a uniformly minimum variance unbiassed estimator (U.M.V.U.E.) for $g(\theta)$. Thus,

$$\hat{g}(\theta) = \frac{b(t-1,n)}{b(t,n)} .$$

We shall apply the obtained results to the cases of LBD and LPD. Some of the known results will follow trivially as special cases.

(i) The LBD and its decapitated form

Let $X_1, X_2, \ldots X_n$ be a random sample taken from a distribution that has p.d.f.

$$P(X=x) = \frac{k}{k+\beta x}\binom{k+\beta x}{x}\frac{[\theta(1-\theta)^{\beta-1}]^{x}}{(1-\theta)^{-k}}$$

where k and β are known constants. The probability distribution of the complete sufficient statistics $T = \sum_{i=1}^{n} x_i$ is given by,

$$P(T=t) = b(t,n)\,\frac{(\theta(1-\theta)^{\beta-1})^{t}}{(1-\theta)^{-kn}} \quad ,$$

where $b(t,n)$ is the coefficient of $\theta(1-\theta)^{\beta-1}$ in the expansion of $(1-\theta)^{-kn}$. Using Lagrange's formula for the expansion of the function $(f(\theta))^{n} = (1-\theta)^{-kn}$, under the transformation $u = \theta/(1-\theta)^{-(\beta-1)}$ we get:

$$P(T=t) = \frac{nk}{nk+\beta t}\binom{nk+\beta t}{t}\frac{[\theta(1-\theta)^{\beta-1}]^{t}}{(1-\theta)^{-nk}}$$

Thus, $b(t,n) = \frac{nk}{nk+\beta t}\binom{nk+\beta t}{t}$. The U.M.V.U.E. for the function $g(\theta) = \theta(1-\theta)^{\beta-1}$ is then given by

$$\frac{b(t-1,n)}{b(t,n)} = \frac{nk+\beta t}{nk+\beta(t-1)}\,\frac{\binom{nk+\beta(t-1)}{t-1}}{\binom{nk+\beta t}{t}}$$

For $\beta = 1$, the LBD becomes a negative binomial distribution and, $g(\theta) = \theta$

$$\therefore \quad \hat{\theta} = \frac{nk+t}{nk+t-1}\,\frac{\binom{nk+t-1}{t-1}}{\binom{nk+t}{t}} = \frac{t}{nk+t-1}$$

which is the same result as given by Roy and Mitra (1957).

For $\beta = 0$, we get the binomial distribution and, $g(\theta) = \frac{\theta}{1-\theta}$.

The U.M.V.U.E. for the function $g(\theta)$ is given by

$$\frac{\hat{\theta}}{1-\hat{\theta}} = \frac{\binom{nk}{t-1}}{\binom{nk}{t}} = \frac{t}{nk-t+1} \; .$$

or

$$\hat{\theta} = \frac{t}{nk+1} \quad , \qquad (4\text{-}14)$$

The estimator given by (4-14) is biassed but has minimum variance. An unbiassed estimator for θ can be constructed by multiplying the right hand side of (4-14) by $\frac{nk+1}{nk}$. This result agrees with Corollary 1, p. 1052 (Patil 1963). Thus, there is no U.M.V.U.E. for θ, since the range of values of the binomial variable is finite.

For $\beta = 2$ we get, $g(\theta) = \theta(1-\theta)$ and the U.M.V.U.E. is given by

$$\hat{\theta}(1-\hat{\theta}) = \frac{nk+2t}{nk+2(t-1)} \; \frac{\binom{nk+2(t-1)}{t}}{\binom{nk+2t}{t}} = \phi(t) \; ,$$

and the existence of an estimator which attains minimum variance requires $\phi(t) \leq \frac{1}{4}$. For larger values of β an iterative technique can be used to find $\hat{\theta}$.

However, the distribution of the sum of n independent decapitated Lagrangian Binomial variates is given by

$$P(T=t) = \sum_{r=1}^{n} \frac{kr(-1)^{n-r}\binom{n}{r}}{(\beta t+kr-t)} \binom{\beta t+kr-1}{t} \frac{\left(\theta(1-\theta)^{\beta-1}\right)^{t}}{\left((1-\theta)^{-k}-1\right)^{n}} \; . \qquad (4\text{-}15)$$

Thus, $$b(t,n) = \sum_{r=1}^{n} \frac{kr(-1)^{n-r}\binom{n}{r}}{(\beta t+kr-t)} \binom{\beta t+kr-1}{t} \; .$$

For $\beta = 1$

$$b(t,n) = \sum_{r=1}^{n} (-1)^{n-r} \binom{n}{r} \binom{t+kr-1}{t} = n!\, k^n\, S^t_{k,n} \qquad \text{(say)}$$

which is the same result as obtained by Cacullous and Charalambides (1972). The UMVU estimator for θ' in this case, is given by

$$\hat{\theta} = \frac{b(t-1,n)}{b(t,n)} = \frac{t}{nk+t-1} \left(1 - \frac{S^{t-1}_{k,n-1}}{S^{t}_{k,n}}\right) \tag{4-16}$$

If it is required to find a U.M.V.U.E. for the non-truncated negative binomial distribution, we multiply the right hand side of (4-16) by $(1 - S^{t-1}_{k,n-1}/S^{t}_{k,n})^{-1}$.

(ii) <u>The LP Distribution</u>

Let $X_1, X_2, \ldots X_n$ be a random sample taken from a distribution that has p.d.f.

$$P(X=x) = \frac{(1+\lambda x)^{x-1}}{x!} (\theta e^{-\theta\lambda})^x / e^{\theta} \qquad x = 0,1,2,\ldots$$

where λ is known and $\theta > 0$ is unknown parameter. The probability distribution of the complete sufficient statistic $T = \sum_{i=1}^{n} x_i$ is given by

$$P(T=t) = \frac{n(n+\lambda t)^{t-1}}{t!} \; \frac{(\theta e^{-\lambda\theta})^t}{e^{n\theta}}$$

Here $g(\theta) = \theta e^{-\lambda\theta}$ and $b(t,n) = \frac{n(n+\lambda t)^{t-1}}{t!}$

Thus the U.M.V.U.E. for $g(\theta)$ is given by,

$$\frac{b(t-1,n)}{b(t,n)} = \frac{t(n+\lambda(t-1))^{t-2}}{(n+\lambda t)^{t-1}}$$

For $\lambda = 0$, we get the Poisson distribution, and the U.M.V.U.E. for $g(\theta) = \theta$ will be given by:

$$\hat{\theta} = \frac{t}{n} ,$$

which is the same result as obtained by Roy and Mitra (1957).

BIBLIOGRAPHY

1. AHUJA, J.C. (1969), "On the distribution of the sum of independent positive binomial variables", Canad. Math. Bull., 13, 151-152.

2. __________ (1970), "On a certain recurrence relation", Ann. Math. Stat., Vol. 41, No. 2, 665-666.

3. __________ (1971), "On the distribution of the sum of independent decapitated negative binomial variables", Ann. Math. Stat., 42, 383-384.

4. BARTON, D.E., DAVID, F.N. & MERRINGTON, M. (1960), "Tables for the solution of the exponential equation, exp(-a) + Ka = 1", Biometrika, 47, 439-445.

5. BLISCHKE, W.R. (1968), "Mixture of discrete distributions", Proc. Int. Symp. on classical and contagious discrete distributions (ed. G.P. Patil), Statistical Publishing Society, Calcutta and Pergman Press, London and New York.

6. CACOULLOS, T. & CHARALAMBIDES, CH. (1972), "On minimum variance unbiassed estimation for truncated binomial and negative binomial distributions", Ann. Inst. Stat. Math. Tokyo, 27, No. 2, 235-244.

7. CHARALAMBIDES, CH. (1974a), "Minimum variance unbiassed estimation for a class of left-truncated discrete distributions", Sankhya, Vol. 36, Ser. A. Pt. 4, 397-418.

8. ______________ (1974b), "The generalized Stirling and C numbers", Sankhya, Vol. 36, Ser. A, Pt. 4, 419-436.

9. CONSUL, P.C. & JAIN, G.C. (1973), "On some interesting properties of the generalized Poisson distribution", Biom. z, Bd., 15, H.7.S., 495-500.

10. ________________________ (1973), "Generalization of the Poisson distribution", Technometrics, Vol. 15, No. 4, 791-799.

11. CONSUL, P.C. & SHENTON, L.R. (1972), "Use of Lagrangian expansion for generating discrete generalized probability distributions", SIAM J. APPL. Math., Vol. 23, No. 2, 239-248.

12. __________________________ (1973), "Some interesting properties of Lagrangian distributions", Communications in statistics, 2(3) 263-272.

13. CONSUL, P.C. & SHENTON, L.R. (1974), "On the Probabilistic structure and properties of discrete Lagrangian distributions", Statistical distributions in scientific work, Vol. 1, D. Reidel Publishing Company.

14. DACEY, M. (1972), "A family of discrete probability distributions defined by the generalized hyper-geometric series", Sankhya, Ser. B 34, 243-250.

15. FAA DE BRUNO (1855), "Note Sur Une Nouvelle Formula Du Calcul Differential", Quart. Jour. Math., 1, 359-360.

16. FELLER, W. (1943), "On a general class of contagious distributions", Ann. Math. Stat. 4, 389-409.

17. FELLER, W. (1968), "An introduction to Probability theory and its applications", Vol. 1, John Wiley, New York.

18. GUPTA, R.C. (1974,(1)), "Modified power series distribution and some of its applications", Sankhya, 36, Ser. B, Pt. 3, 288-298.

19. ___________ (1974,(2)), "Distribution of the sum of independent decapitated generalized negative binomial variables", Sankhya, 36, Ser. B, Pt. 1, 67-69.

20. ___________ (1974,(3)), "Distribution of the sum of independent decapitated generalized Poisson variables", Sankhya, Vol. 36, Ser. B, Pt. 2, 212-214.

21. ___________ (1975), "Maximum likelihood estimation of a modified power series distribution and some of its applications", Communications in statistics, 4(7), 689-697.

22. GURLAND, J. (1957), "Some interrelations among compound and generalized distributions", Biometrics, 44, 265-268.

23. ___________ (1958), "A generalized class of contagious distributions", Biometrics, 14, 229-250.

24. GURLAND, J. & TRIPATHI, R. (1974), "Estimation on some extensions of the Katz family of discrete distributions", Statistical distributions in scientific work, Vol. 1, D. Reidel Publishing Company.

25. HAIGHT, F.A. & BREUER, M.A. (1960), "The Borel-Tanner distribution", Biometrika 47, 143-150.

26. HAIGHT, F.A. (1961), "A distribution analogous to the Borel-Tanner", Biometrika, 48, 167-173.

27. HAIGHT, A. (1967), "Hand book of Poisson distribution", John Wiley & Sons, New York.

28. HALDANE, J.B.S. & SMITH, S.M. (1956), "The sampling distribution of a maximum likelihood estimate", Biometrika, 43, 96-103.

29. JAIN, G.C. & CONSUL, P.C. (1971), "A generalized negative binomial distribution", SIAM J. Appl. Math., 21, 501-513.

30. JAIN, G.C. & GUPTA, R.P. (1973), "A logarithmic series type distribution", Trabajos De Estadistica, 99-105.

31. JAIN, G.C. (1974), "A Borel-Tanner distribution and its approximation to the negative binomial distribution", Math. operations forsch. U. Statist. 5, Heft 3, S., 207-212.

32. JORDAN, C.H. (1965), "Calculus of finite differences", Chelsea Publishing Company, New York.

33. KATTI, S.K. (1970), "Interrelations among generalized distributions and their components", Biometrics, 22, 44-52.

34. KATZ, L. (1968), "Unified treatment of a broad class of discrete probability distributions", Classical and Contagious discrete distributions, G.P. Patil (ed.), Pergman Press, New York.

35. KEMP, A. (1968), "A wide class of discrete distributions and the associated differential equations", Sankhya, Ser. A 30, 401-410.

36. KENDALL, M.G. & STUART, A. (1963), "The advanced theory of statistics", Vol 1, Charles Griffin and Con., London.

37. KHATRI, C.G. (1959), "On certain properties of power series distributions", Biometrika, 46, 486-490.

38. MENDENHALL, W. & LEHMANN, E.H. (1960), "An approximation to the negative moments of the positive binomial useful in life testing", Technometrics, 2, 227-242.

39. MOHANTY, S.G. (1966), "On a generalized two-coin tossing problem", Biometrische Zeitschrift, 8(4), pp. 266-272.

40. NELSON, D.L. & CONSUL, P.C. (1974), "A Lagrangian gamma distribution", Statistical distributions in scientific work, Vol 1, D Reidel Publishing Company.

41. NOACK, A. (1950), "A class of random variables with discrete distributions", Ann. Math. Stat., 21, 127-132.

42. ORD, J.K. (1967), "Graphical method for a class of discrete distributions", J. Roy. Stat. Soc., Ser. A, 130, 232-238.

43. _________ (1972), "Families of frequency distributions", Charles Griffin & Co., Ltd.

44. PARK, C.J. (1973), "The power series distributions with unknown truncation parameter", Ann. Math. Stat., Vol. 1, No. 2, 395-399.

45. PATEL, J., KAPADIA, C.H. & OWEN, D.B. (1976), "Hand book of statistical distributions", Marcel Dekker, New York and Basel.

46. PATIL, G.P. (1962), "Certain properties of the generalized power series distribution", Ann. Inst. Stat. Math., Tokyo, 14, 179-182.

47. PATIL, G.P. (1962), "Maximum likelihood estimation for GPSD and its application to a truncated binomial distribution", Biometrika, 49, 1 and 2, P. 227.

48. ____________ (1963), "Minimum variance unbiassed estimation and certain problems of additive number theory", Ann. Math. Stat., 34, 1050-1058.

49. PATIL, G.P. & WANI, J.K. (1965a), "Maximum likelihood estimation for the logarithmic series distribution with one parameter", Proc. Int. Symp. on classical and contagious discrete distributions (Ed., G.P. Patil), Statistical Publishing Society, Calcutta and Pergman Press, London and New York.

50. ____________ (1965b), "On certain structural properties of the logarithmic series distribution and first type Stirling distribution", Sankhya, A, 27, 271-280.

51. PATIL, G.P. & BILDIKAR, S. (1966), "On minimum variance unbiassed estimation for the logarithmic series distribution", Sankhya, A, 28, 239-250.

52. PRABHU, N.V. (1965), "Queues and inventories", John Wiley & Sons, New York.

53. RAO, C.R. (1952), "Advanced statistical methods in biometric research", John Wiley & Sons, New York.

54. ROY, J. & MITRA, S. (1957), "Unbiassed minimum variance estimation in a class of discrete distributions", Sankhya, 18, 371-378.

55. SINGH, J. (1975), "A note on the Stirling distribution of the second kind", Communications in statistics, 4 (8), 753-759.

56. SIROMONEY, G. (1964), "The general Dirichlet's series distribution", Journal of the Indian statistical association, 2, 69-74.

57. STEPHEN, F.F. (1945), "The expected value and variance of the reciprocal and other negative powers of a positive Bernoullian variate", Ann. Math. Stat., 16, 50-61.

58. TAKACS, L. (1962), "A generalization of the ballot problem and its application in the theory of queues", J.A.S.A., 57, 327-337.

59. TANNER, J.C. (1953), "A problem of interference between two queues", Biometrika, 40, 50-69.

60. ____________ (1961), "A derivation of Borel distribution", Biometrika, 40, 222-224.

61. TATE, R. & GOEN, R. (1959), "Minimum variance unbiassed estimation for a truncated Poisson distribution", Ann. Math. Stat., 29, 755-765.