AN EMPIRICAL STUDY ON SOME INFINITELY DIVISIBLE AND MIXTURES OF DISCRETE DISTRIBUTIONS

A Thesis submitted in partial fulfillment of the requirements
for the award of the Degree of Doctor of Philosophy

Rousan Ara Begum

Registration No. 146 of 2001

The School of Science and Technology

Department of Mathematical Sciences

Tezpur University

Napaam, Tezpur-784028, Assam, India

July 2006

ABSTRACT

Introduction

The theory of infinite divisible distributions, developed primarily during the period from 1920 to 1950, plays a fundamental role in several parts of theoretical probability. Loosely speaking, divisibility of a random variable X is the property that X can be divided into independent parts having the same distribution. A full account of this theory and its applications, as it had been developed by the late 40's, were presented in the monographs of Levy (1937), Gnedenko and Kolmogorov (1968), Loeve (1960) and Steutel and Harn (2004).

Infinite divisibility plays a significant role in the solution of limit problems for sums of independent random variables. It is well known that, stochastic processes, more specifically by the processes with stationary independent increments generate all infinitely divisible distributions. In fact, all Poisson stopped sum distributions (compound Poisson distributions) belong to the important class of infinitely divisible distributions (Johnson, et al, 1992). The converse of the result is also true. De Finetti (1931) has proved that all infinitely divisible distributions are limiting forms of generalized Poisson distributions (See also Lukacs, 1970).

Elementary infinitely divisible distributions, which are formulated on the basis of simple models, seem to be inadequate to describe the situations which may occur in a number of phenomena. In the last few years' research various infinite divisible distributions have been derived. There are some applied processes that give rise to classes of infinite divisible distributions. The best known are convolution and compounding. A general compounding theorem is due to Feller (1957). It has been observed that certain families of probability distributions maintain their infinite divisibility under repeated mixing and convolution.

Infinite divisibility

A random variable is said to be infinite divisible if and only if it has a characteristic function, $\phi(t)$ that can be represented for every positive integer 'n' as the n^{th} power of some characteristic function $\phi_n(t)$, ie.

$$\phi(t) = \{\phi_n(t)\}^n \ . \tag{1.1}$$

For example, in case of discrete distribution the Poisson, geometric, negative binomial, logarithmic-series, discrete Pareto and Borel distribution etc., are infinitely divisible. Conditions for a discrete distribution to be infinitely divisible are discussed in Katti (1967), Warde and Katti (1971), and Chang (1989).

In this study, we have selected the following two models (1.2) and (1.3), studied by Klebanov, Maniya and Melamed (1984) and Steutel (1990) under the name of geometrically infinitely divisible distributions.

$$G(t) = \frac{(1-\omega)g(t)}{1-\omega g(t)}, \qquad 0 < \omega < 1. \tag{1.2}$$

$$G(t) = \frac{\omega}{\omega + 1 - g(t)}, \quad \omega > 0 \tag{1.3}$$

where $g(t)$ is the pgf of the component distribution used.

The aim of this thesis is to derive some generalized distributions of Poisson-Lindley (Sankaran, 1970), three parameter Charlier (Jain and Gupta, 1975) and Gegenbauer (Plankett and Jain, 1975), by considering them as the component distribution in the above models (1.2) and (1.3). The generalized distributions have been studied in the presence of their recurrence relations for their probabilities, factorial moments and cumulants. The parameters are estimated and a few sets of reported data have been considered for the fitting of the distributions, and the fits are compared with that of other distributions.

The thesis consists of seven chapters. The **first chapter** is an introductory one which highlights the literatures on univariate discrete distributions and some Lagrangian distributions. The **second chapter** is a review on some probability distributions, such as Poisson Lindley distribution (Sankaran, 1970), three parameter Charlier distribution (Jain and Gupta, 1975) and Gegenbauer distribution (Plankett and Jain, 1975). In the succeeding chapters we have studied some of their generalized form of the distributions.

In **chapter 3**, we have studied certain properties of generalized Poisson-Lindley distributions of Type-I and Type-II, based on the models (1.2) and (1.3) studied by Klebanov, Maniya and Melamed (1984) and Steutel (1990) under the name geometrically infinitely divisible distributions. Certain properties of Infinite Divisible distributions of Type-I and Type-II, have been studied in **chapters 4** and **5** respectively, by considering three parameter Charlier and Gegenbauer as the generalizing distribution. A class of Charlier family of Lagrangian discrete probability distributions has been considered in **chapter 6**. In the last **Chapter 7**, we have made an attempt to test the fitting of Gegenbauer distribution to some well known published data on Ball Games.

Poisson-Lindley Distribution

Poisson Lindley distribution (Sankaran, 1970) is a one-parameter compound Poisson distribution, which has wide applications in the theory of accident proneness. The probability generating function of Poisson-Lindley distribution may be defined as

$$g(t) = \frac{\theta^2(\theta+2-z)}{(\theta+1)(\theta+1-z)^2}, \quad \theta > 0 \qquad \text{(See Sankaran, 1970)} \qquad (1.4)$$

Three parameter Charlier distribution

The probability generating function of three parameter Charlier distribution (Jain and Gupta, 1975) is given by

$$g(t) = e^{-\alpha}(1-\beta)^{\lambda} e^{\alpha t}(1-\beta t)^{-\lambda}, \quad \alpha, \beta, \lambda \geq 0 \qquad (1.5)$$

As particular limiting cases of three-parameter Charlier distribution, Poisson distribution and negative binomial distribution may be obtained by putting $\beta = 0$ and $\alpha = 0$ respectively.

Gegenbauer distribution

The Gegenbauer distribution Plunkett and Jain (1975), has the probability generating function of the form

$$g(t) = (1-\alpha-\beta)^{\lambda}(1-\alpha t-\beta t^2)^{-\lambda} \qquad (1.6)$$

The limiting distributions of Gegenbauer distribution are negative binomial ($\beta = 0$) and Hermite distribution ($\alpha \to 0$, $\beta \to 0$ and $\lambda \to \infty$, such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$) as particular limiting cases.

Lagrangian distribution

Consul and Shenton (1972) gave us a new generation of distributions having interesting properties associated with the queueing processes. A class of discrete probability distributions under the title 'Lagrangian Distributions' had been introduced into the literature by Consul and Shenton (1972, 1973, 1975). They used the particular title on account of the generation of these probability distributions by the well known Lagrange expansion of a function $g(s)$ as a power series in y when $y = s/g(s)$. Using Lagrange's expansion for the derivative of the probabilities of certain discrete distributions Consul and Shenton (1972, 1973, and 1975) and their co-workers derived the functional form of the distributions and studied different properties also. A detailed study on these mixtures of discrete probability distributions and their properties can be found in the works of Gurland (1957, 1958, 1965), Haight (1961), Janardan and Rao (1983), Consul (1989), Everitte and Hand (1981) and Johnson Kotz and Kemp (1992). The probability mass function (pmf) of Lagrange distributions (Consul and Shenton, 1972) of first kind (LD1) may be given as

$$P_r(X=x) = \frac{1}{x!}\left[\frac{\delta^{x-1}}{\delta s^{x-1}}\{g(s)\}^x f'(s)\right]_{S=0}, \qquad for\ x = 1,2,3,... \tag{1.7}$$

where $P_r(X=0) = f(0)$.

The probability mass function (pmf) of Lagrange distributions (Janardan and Rao (1983) of second kind (LD2) may be given as

and $$P_r(X=x) = \frac{1-g'(1)}{x!}\left[\frac{\delta^{x}}{\delta s^{x}}\{g(s)\}^x f(s)\right]_{S=0}, \quad for\ x = 0,1,2,3,... \tag{1.8}$$

$$= 0, \qquad otherwise$$

In chapter 6, a class of Charlier Family of Lagrangian distributions $type-I$ and $type-II$ have been derived like the other authors (Consul et al., 1973 and Janardan et al., 1983) by taking different choice of three parameter charlier, Poisson, negative binomial, Logarithmic series and delta distributions as $f(s)$ and $g(s)$. The main objective of this thesis is to investigate the probabilistic structures of Charlier Family of Lagrangian distributions of $type-I$ and $type-II$, and discuss some of their important properties and applications.

Estimation of parameters

The estimation of parameters plays a very important role in fitting of probability distributions. Of all the procedures of estimating the parameters, the method of moments is perhaps the oldest and the simplest. In many cases it may lead to tractable operations. Although, the method of maximum likelihood is considered to be more accurate for fitting a probability distribution than all other methods of estimation, in our cases it involves much more computational works than the method of moments. It is mainly from this reason, moment estimators are used.

In **chapter 3**, the parameters of generalized Poisson-Lindley distributions of type-I and type-II (GPL1 and GPL2), have been estimated by considering a method in which one parameter is estimated by using Newton-Raphson method and the other parameter is estimated by the method of moment.

In **chapter 4** and **chapter 5**, the parameters of generalized Charlier (GCD) and generalized Gegenbauer (GGD) distributions of type-I and type-II have been estimated by an adhoc method of using first two sample moments and ratios of first three frequencies for rapid prediction, because of complexity of maximum likelihood method of estimation.

It is to be very important to note that, when the frequency for the zero class in the sample is larger than most of the other class frequencies or when the graph of the sample distribution is approximately L-shaped, one would like to give more weight to this larger frequency value of the zero class than to the statistic of sample variance which is more affected by the frequencies of the higher classes (Anscombe, F. W., 1950). It is for this reason, the parameters of basic Lagrangian negative binomial distribution has been estimated by the method of using the ratio of first two frequencies and mean in **chapter 6**. The method of moment is also used and both methods give us satisfactory results. In case of basic Lagrangian Poisson distribution, the parameters are estimated by using maximum likelihood method and using the first sample frequency. Again for testing the goodness of fit of general Lagrangian distributions in **chapter 6**, one composite method of using the first two sample moments and ratio of first three frequencies is used for estimation of parameters.

Applications

The negative-binomial and the Poisson distributions are commonly used in ecological and biological problems. It is for that reason, the generalized distributions of Poisson-Lindley (GPL1 and GPL2), three parameter Charlier (GCD1 and GCD2) and Gegenbauer (GGD1 and GGD2) are fitted to some numerical data in different fields of **biology, ecology, social information** for which various modified forms of Poisson distribution were fitted by different authors. In all cases, observing the observed and expected frequencies it is clearly seen that our fitted distributions describe the data very well.

The basic Lagrangian negative binomial and Lagrangian Poisson distribution have been fitted to some data collected by Williams (1944) on the numbers of papers published by authors in a certain Journal for which Plunkett and Jain (1975) fitted generalized geometric distribution. It has been found that the basic LNB distribution gives better fit than the other distributions compared. The general Lagrangian negative binomial Poisson (LNBP), Lagrangian Poisson negative binomial (LPNB) and Lagrangian Poisson logarithmic (LPL) distributions have been also fitted to some well known data on **natural laws in social sciences, accidents, home injuries, biological and ecological** and the data **on the number of publication of research papers** etc. In all cases, the expected frequencies obtained are more satisfactory than the distributions compared.

In chapter 7, the Gegenbauer distribution has been fitted to the distribution of scores of teams and individuals in several sports involving ball games. For example, the distribution of runs scored in the completed innings in test matches by some famous batsmen at **cricket** (Reep and Bengamin, 1971), the distribution of scores of individual teams in U.S. Collegiate **football games** (Pollard, 1973 and Morney, 1956), and also the distribution of goals per match scored by individual teams in national Hockey League 1966-67 (Reep et al, 1971) have been used to fit the Gegenbauer distribution. It is observed that, in most of the above cases Gegenbauer distribution provides better fit than that of the negative binomial distribution.

Ph : +91-3712-267012(O) -236571(R), Fax : +91-3712-267006
E-mail : mborah@tezu.ernet.in

Tezpur University

Prof. Munindra Borah
Head, Department of Mathematical Sciences
Napaam, Tezpur-784028, Dist-Sonitpur, Assam, India

TO WHOM IT MAY CONCERN

It is to certified that the research work in the thesis entitled **"AN EMPIRICAL STUDY ON SOME INFINITELY DIVISIBLE AND MIXTURES OF DISCRETE DISTRIBUTIONS"** submitted by **Rousan Ara Begum** for the degree of Doctor of Philosophy was carried out under my supervision. The results presented here are of her own work and the thesis or part of it has not been submitted to any other University or Institute for any research degree or diploma.

She has fulfilled all the requirements under rules and regulation for the award of the degree of Doctor of Philosophy of Tezpur University, Napaam, Tezpur.

M Borah 26/7/2006

(Prof. M. Borah)
Professor & Head
Department of Mathematical Sciences
Tezpur University

Dept. of Mathematical Science, Tezpur University, TEZPUR ASSAM

Preface

The present study has been undertaken at the suggestion and supervision of Dr. M. Borah, Professor and Head of the department of Mathematical Sciences, Tezpur University, Tezpur, Assam. In the beginning of my research, he asked me to go through *certain technical papers on discrete probability distributions. Throughout my studies,* I have been highly influenced by the learned works of authors like F.W. Steutel, P.C. Consul, S.K. Katti, C.D. Kemp, L.R. Shenton, N.L. Johnson, S. Kotz, G.C. Jain, R.K. Gupta, I.G. Plunkett, M. Sankaran, M. Fisz, K.G. Janardan, J. Gurland, L.B. Klebanov, G.M. Maniya, I.A. Melamed, and A. W. Kemp. After going through the papers, my knowledge and interest towards Infinite divisibility of discrete distributions and Lagrangian type of distributions has increased and inspired me to continue this thesis. I am thankful to all of them.

In 1997, I had presented my first paper on the title "A class of discrete Probability Distributions applicable in Queueing theory", in the national seminar on Recent Trends and Advances of Mathematics and Statistics in Engineering and Technology, organized by the Department of Applied Mathematics, Indian School of Mines, Dhanbad, India. The publication of this paper in the book "*Mathematics and Statistics in Engineering and Technology*", under Narosa Publishing House, inspired me further. I have participated and presented another paper "Some Properties of Certain Infinitely Divisible Discrete Distribution." in an international conference organized by the Department of Statistics, University of Mysore, Manasgangothri, Mysore 570006, India, from December 28-30, 1998, on 'Combinatorics, Statistics, Pattern Recognition & Related Areas'.

In 1999, another paper had been published with the title "*Some properties Of Poisson Mixing Infinitely Divisible Distributions*" in the proceedings of the Annual Technical Session, Assam Science Society.

Again, in November 2001, I got a chance to add a chapter entitled "*Certain Infinitely Divisible Discrete Probability Distribution and Its Application*" based on my works in a book "Mathematics and Statistics in Engineering, Biotechnology and Science" edited by Mr. Dipak K. Sen, of R. S. More College, Govindpur, Dhanbad, Jharkhand. I am thankful to him, as he had given me a chance to establish few of my ideas. My forth paper entitled " *Some properties of Poisson-Lindley and its derived distributions*" is published in *Journal of Indian Statistical Association* of University of Pune, Vol-40, No. 1 in the year 2002.

I would also like to mention with thanks that U.G.C. has given me the Teachers Fellowship under 10th plan, which was a boon in carrying out my whole thesis.

I gratefully acknowledge my sincere thanks and gratitude to my supervisor and guide Prof. M. Borah, the Head of the department of Mathematical Sciences, Tezpur University, Tezpur, Assam for his encouragement and invaluable guidance which I received from him from the very beginning of the present investigation till to the submission of the thesis. I must admit that without his skillful guidance, constant support, inspiration and valuable suggestion, this research work would have never been completed. My study involves a lot of computer works on FORTRAN programming in which my supervisor Prof. M. Borah and my colleague Mr. Rajan Sarma have helped me to carry out the programming on computer. I am thankful to both of them.

I would like to offer my thanks to the faculty members of the department of Mathematical Sciences, Tezpur University, for their suggestion and encouragement in my research work. I am grateful to the authorities of Tezpur University for their co-operation and support during this period.

Finally, I am thankful to my colleagues in the Department of Statistice, Darrang College, Tezpur, Assam whose suggestion and encouragement which helped me in completing my whole research work.

Department of Mathematical Sciences
Tezpur University, Napaam
Tezpur-784028, Assam, India
Date: 26.7.2006

(Rousan Ara Begum)

Rousan Ara Begum

Contents

Chapter 4: Generalized Infinitely Divisible Distributions of Type-1

Chapter 5: Generalized Infinitely Divisible distributions of Type-2

A: List of Tables

B: List of figures

Chapter 1

Introduction

1.1 Introduction

The theory of discrete probability distributions is an important branch of Statistics with varieties of useful applications. The field of discrete probability distribution is originated from the works of James Bernoulli (1713) and Poisson (1837). In recent years, the mixtures of the basic discrete distributions such as Poisson, binomial, negative binomial and Logarithmic series have become an extremely important part of modern Statistics, with wide utilities in the fields of social sciences, physical sciences, biological and medical sciences, operation research, engineering and so on. A detailed study on these discrete mixtures of probability distributions and their properties can be found in the works of Gurland (1957, 1958, 1965), Khatri (1959, 1961), Haight (1961a, 1967), Patil (1961, 1962a, 1962b, 1964), Katti (1967, 1977), Katti and Gurland (1961, 1962a, 1962b), Johnson and Kotz (1969), Consul (1975, 1989), Consul and Shenton (1972, 1973, 1975), Steutel (1968, 1973, 1990), Everitte and Hand (1981), and Johnson et al (1992) and Steutel and Harn (2004) etc.

Before going to the main steps that we have studied, let us first discuss some important terms related to the theory of discrete probability distributions. It is very important to note that, the scope of this thesis is basically restricted to univariate discrete distributions only.

1.2 Mixture Distribution

A mixture of distributions is a mechanism which helps us to construct new distributions from the given ones for which empirical justification must be sought. According to Medgyessy (1977), a mixture of distributions is a superimposition of distributions with different functional forms or different parameters, in specified proportions.

A continuous mixture of discrete distribution arises when a parameter corresponding to some features of a model for a discrete distribution can be regarded as a random variable taking continuous values. Different mixtures of Poisson distributions, where the mixing distributions are countable or continuous are discussed in details in the book by Johnson et al (1992).

1.3 Generalized distributions

If $g_1(t)$ is the probability generating function (pgf) of a distribution function F_1, and if the argument 't' is replaced by the pgf $g_2(t)$ of another (or the same) distribution function F_2, then the resultant function $g_1\{g_2(t)\}$, is also a pgf, which is a polynomial function of 't' with non negative coefficients. The probability distribution corresponding to the new pgf $g_1\{g_2(t)\}$, is called a generalized F_1 distribution, and this can be written in the symbolic form $F_1 \vee F_2$.

In case of mixture distribution there are three important theorems due to Levy (1937b, 1954), Maceda (1948) and Gurland (1957) on the methods of generalization.

a) Levy's Theorem (see Feller, 1957)

A discrete probability distribution on non negative integers is infinitely divisible if and only if (iff) its pgf can be written in the form

$$G(s) = \exp[\lambda\{g(s) - 1\}],$$

where $\lambda > 0$ and $g(s)$ is another pgf.

The implication of Levy's Theorem is that an infinitely divisible distribution with non-negative support can be interpreted as

1. a stopped sum of Poisson distributions (i.e., as the sum of Y random variables with pgf g(z), where Y has a Poisson distribution).

2. a convolution (sum) of a Poisson singlet, Poisson doublet, triplet etc., where the successive parameters are proportional to the probabilities given by $g(s)$.

b) Maceda's theorem (see Maceda, 1948)

This theorem states that if we consider a mixture of Poisson distributions where the mixing distribution has non-negative support, then the resultant distribution is infinitely divisible iff the mixing distribution is infinitely divisible. As infinitely divisible discrete distributions can be interpreted as Poisson-stopped-sum distributions, the implication from Levy's and Maceda's theorems is that mixing Poisson distributions using an infinitely divisible distribution yields a Poisson-stopped-sum distribution. Furthermore, mixing a Poisson-stopped-sum distribution using an infinitely divisible distribution gives rise to another Poisson-stopped-sum distribution.

c) Gurland's theorem (see Gurland, 1957)

According to Gurland's theorem theorem, a distribution with pgf of the form $G_1\{G_2(z)\}$ will be called a generalized F_1 distribution, more precisely a F_1 distribution generalized by the generalizing F_2 distribution. This theorem can be stated symbolically as

$$F_1 \vee F_2 \sim F_2 \wedge F_1,$$

provided that

$$G_2(z/k\phi) = [G_2(z/\phi)]^k. \tag{1.3.1}$$

where k and ϕ denote the parameters of the distributions F_1 and F_2 respectively. The Poisson, binomial and negative binomial distributions all have pgf's of the form (1.3.1), therefore it follows that the discrete mixture distributions of Poisson, binomial and negative binomial distributions are also generalized distributions in the above sense.

1.4 Review on certain literatures of discrete mixture distributions

The origin of the theory of discrete probability distributions developed primarily during the works of James Bernoulli and Poisson. James Bernoulli, the Swiss mathematician derived the binomial distribution and published it in the year 1713. In 1837, Poisson distribution was derived by a French mathematician Simeon D. Poisson as a limiting form of the binomial distribution. Greenwood and Yule obtained negative binomial distribution in 1920 as a consequence of certain simple assumptions in accident proneness models. As time goes, it is seen that the simple basic distributions such as binomial, Poisson, negative binomial, logarithmic series etc. have been found to be describe some sets of data, which leads to construct new generalized or mixture models of the basic distributions. Univariate mixtures obtained by combining two or more of the elementary distributions with the help of the process of compounding or generalizing have become an important and extremely useful branch of Statistics. In this way a large number of discrete distributions were derived by different authors, which are classified as generalized, modified and contagious distributions. The generalized distributions can satisfactorily describe the data and slowly in the recent years the field of discrete distribution has been expanded. These distributions have important applications in biological sciences, medical sciences, social sciences, physical sciences, operation research, engineering and so on. From the books of Johnson and Kotz (1969), Everitt and Hand (1981), Consul(1989) and Johnson et al (1992) we may get a detailed information regarding the vast area of discrete mixture distributions and their properties.

Lindley (1958) derived a distribution known as Lindley distribution based on Baye's theorem. Sankaran (1970), generalized Lindley distribution by mixing with Poisson distribution, which is known as Poisson-Lindley distribution with application to mistakes in copying groups of random digits (data from Kemp and Kemp, 1965) and accidents to 647 women on high explosive shell in 5 weeks (data from Greenwood and Yule, 1920). Borah and Begum (2002) studied some properties of Poisson-Lindley and its derived distributions. They fitted the generalized distributions of Poisson-Lindley to some biological and ecological data for comparison. Some mixtures of Poisson Lindley distribution using Gurland's generalization (1957) were also studied by Borah and Deka Nath (2001). They also studied and fitted the inflated Poisson Lindley distribution to some well-known data, for empirical comparison.

The Hermite distribution which is a Poisson mixture of Bernoulli distribution was studied by Kemp and Kemp (1965). He applied Hermite distribution to the fields of biological sciences, physical science and operational research. The pgf of Hermite distribution can be conveniently expressed in terms of modified Hermite polynomials. By Gurland's theorem, the Hermite distribution can also be regarded as a Poisson mixture of binomial distributions which was used for biometrical data by Skellam (1952) and McGuire et al. (1957). The Hermite distribution which is a Poisson-stopped-sum (generalized Poisson) distribution can also be derived as the sum $X = Y_1 + Y_2$, where Y_1 is a Poisson random variable with parameter a_1 and Y_2 is a Poisson random variable with parameter a_2.

Plunkett and Jain (1975) derived the Gegenbauer distribution by mixing the Hermite distribution with the Gamma distribution. This distribution has a long history in the theory of stochastic processes. Further they have studied some of its limiting cases and its goodness of fit by estimating the parameters of the distribution by using the factorial moments. They fitted the distribution to a set of accident data using the method of moments. Borah (1984), studied the probability and moment properties of the three parameter Gegenbauer distribution, and also studied the estimation of the parameters by using the first two sample moments and the ratio of the first two sample frequencies. They fitted the distribution to a set of accident data using the method of moments. Medhi and Borah (1984) studied the four parameter generalized Gegenbauer distribution with different estimation techniques. Kemp (1979) studied the Gegenbauer distribution which can be obtained by convoluting binomial and pseudo binomial variables. Kemp (1992a, 1992b) also studied various forms of estimation, including ML estimation.

The Charlier polynomials were investigated by Doetsch (1934), Meixner (1934, 1938) and Berg (1985), which are associated with the Poisson distribution of rare events. McBride (1971) discussed two parameter Charlier polynomials and some of their properties. Jain and Gupta (1975) defined the generalized Charlier polynomial. Later on Medhi and Borah (1986) studied generalized four parameter Charlier distribution which includes Negative binomial (Bliss and Fisher, 1953), Hermite (Kemp and Kemp, 1965), Gegenbauer (Plankett and Jain, 1975) and three parameter Charlier (Jain and Gupta, 1975) distributions as particular limiting cases. They investigated different properties of this distribution and also discussed the methods for fitting of four parameter Charlier distribution.

The class of discrete probability distributions under the name of 'Lagrangian distributions' had been introduced into the literature by Consul and Shenton (1972, 1973, 1975) and Mohanty (1966), Consul and Jain (1973) and Janardan and Rao (1983) in the following years. The particular title was chosen by Consul and Shenton (1972) on account of the generation of these probability distributions by the well known Lagrange expansion of a function $g(s)$ as a power series in y when $y = s / g(s)$. The Lagrangian binomial distribution was obtained by Mohanty (1966), Jain and Consul (1971) derived Lagrangian negative binomial distribution, Jain and Gupta (1975) introduced into the literature the Lagrangian Logarithmic distribution. Consul and Jain (1973a) obtained Lagrangian Poisson distribution as a limiting form of the Lagrangian negative binomial distribution. Consul in his book (1989) studied the properties of Lagrangian Poisson distribution. Lagrangian Katz family of distribution was studied by Consul and Felix (1996) with estimation of parameters and applications. Borah and Begum (1997) have studied the probabilistic structures of Charlier Family of Lagrangian (CFL) distributions (Chapter 6) of $type-I$ and $type-II$, by using the Lagrange expansion to this Charlier distribution. A class of Hermite type Lagrangian distribution has been studied by Borah and Deka Nath (2000).

These CFL distributions give rise to a large number of discrete distributions. Most of the basic Lagrangian distributions viz., generalized negative binomial (Jain and Consul, 1971), generalized Poisson (Consul and Jain, 1973a), Borel-Tanner (Tanner, 1961), Haight (Haight, 1961), etc., may be obtained as its limiting cases. All these distributions are found to be of relevance in queueing theory and possess with some interesting properties [Consul and Shenton, 1973].

1.5 Infinite Divisibility of Probability Distributions

A random variable is said to be infinitely divisible iff it has a characteristic function (cf), $\phi(t)$ that can be represented for every positive integer 'n' as the n^{th} power of some cf $\phi_n(t)$, i.e.

$$\phi(t) = \{\phi_n(t)\}^n \ . \qquad (1.5.1)$$

In non-technical terms what it means is that there exist independently and identically distributed (iid) random variables X_{ni} $(i = 1,2,3,...,n)$ such that the distribution of $\sum X_{ni}$ is the same as the given distribution. For example, in case of discrete distribution, the Poisson, geometric, negative binomial, logarithmic-series, discrete Pareto and Borel distribution are infinitely divisible.

All Poisson stopped sum distributions (generalized Poisson distributions) having the pgf's of the form

$$G(t) = \exp[\lambda\{g(t)-1\}],$$

where $g(t)$ is the pgf of the generalizing distribution, belong to the important class of ID distributions (Johnson, et al, 1992). The converse of the result is also true (See Feller, 1957). De Finetti (1931) has proved that all ID distributions are limiting forms of generalized Poisson distributions (See also Lukacs, 1970). In has been observed that in case of continuous distribution the characteristic function of Normal distribution and in case of discrete distribution the characteristic function of Poisson distribution can be easily put in the form (1.5.1), but it may not always be possible to express the characteristic function of an infinite divisible distribution as in the form given in (1.5.1).Conditions for a discrete distribution to be infinitely divisible are discussed in Katti (1967), Warde and Katti (1971), and Chang (1989).

It is well known that all infinitely divisible distributions are generated by stochastic processes, more specifically by processes with stationary independent increments. There are number of methods to construct new infinitely divisible distributions from given ones. The best known are convolution and compounding. A general compounding theorem is due to Feller (1957). There are some applied processes, however, that give rise to classes of infinitely divisible distributions. It has been observed that certain families of probability distribution functions maintain their infinite divisibility (Goldie, 1967 and Steutel, 1968) under repeated mixing and convolution.

Condition for a Distribution to be Infinitely Divisible

Let us suppose that $p_0, p_1, p_2, \ldots,$ are the probabilities of 0,1,2,..., with $p_0 \neq 0, p_1 \neq 0$. Then according to Katti (1967) the necessary and sufficient condition for a distribution to be infinitely divisible is that for each value of i

$$\pi_i = \frac{iP_i}{P_0} - \sum_{j=1}^{i-1} \pi_{i-j} \frac{p_j}{p_0} \geq 0, \qquad \text{for } i = 1,2,..., \tag{1.5.2}$$

Note that for a given distribution function, one can numerically compute a number of $\frac{p_i}{p_0}$, to see if they are positive and if they are, then one can use this information along with his algebraic calculation to generate an inductive proof of infinitely divisibility.

Geometric Infinite Divisibility

The concept of geometric infinite divisibility (gid) was introduced by Klebanov, Maniya and Melamed (1984). According to Klebanov et al. (1984) a random variable

'X' is said to be geometrically infinitely divisible if there exists an independently identically distributed (iid) sequence of random variables $X_p^{(j)}, j = 1,2,3,..., N_p$ such that for every $p \in (0,1)$,

$$X \stackrel{d}{=} \sum_{j=1}^{N_p} X_p^{(j)} \tag{1.5.3}$$

where $P(N_p = k) = p(1-p)^{k-1}, k = 1,2,3,....$ and X, N_p and $X_p^{(j)}$ are independent.

Here the symbol $\stackrel{d}{=}$ stands for equality of distributions.

Pillai ans Sandhya (1990) have shown that the class of geometrically infinitely divisible distributions is a proper sub-class of infinitely divisible distributions. In terms of characteristic function (1.5.3) can be expressed as (1.6.1) where $G(t)$ and $g(t)$ are Laplace transform of X and $X_j^{(p)}$ respectively.

A more detailed description of geometrically infinitely divisible random variable is based on the fact that, a random variable Y with cf $f(t)$ is geometrically infinitely divisible iff

$$\phi(t) = \exp\{1 - \frac{1}{f(t)}\}, \tag{1.5.4}$$

represents an infinitely divisible cf (See Klebanov et al.,1984). Goldie (1967) proved that the product of two independent non-negative random variables is infinitely divisible, if one of the two is exponentially distributed or, equivalently, mixtures (with positive weights) of exponential random variables are infinitely divisible.

Certain Compound Distributions

The following two classes of infinitely divisible characteristic functions are of special interest in discrete probability theory. They are

a) Compound-Poisson Distribution

A distribution with cf of the form

$$\phi(t) = \exp[\lambda\{g(t) - 1\}], \quad \lambda > 0 \tag{1.5.5}$$

in which g(t) is also a cf, is always infinitely divisible. This is also known as Poisson-stopped-sum distribution. They arise as the distribution of the sum of a Poisson number of independently and identically distributed random variables with cf $g(t)$. Because of their infinite divisibility these distributions have very great importance in discrete distribution theory. They are known by different names. Feller (1943) used the term generalized Poisson, Galliher et al. (1959) and Kemp (1967) called them stuttering Poisson and Feller (1950, 1957, 1968) and Lloyd (1980), used the term compound Poisson.

b) Compound geometric Distribution

A probability distribution, with the cf of the form

$$\phi(t) = \frac{\lambda}{\lambda + 1 - g(t)}, \quad \lambda > 0 \tag{1.5.6}$$

in which g(t) is an arbitrary cf on non-negative integers, is always infinitely divisible. It is known as compound geometric distribution (Lukacs, E, 1960). In fact (1.5.6) is a special case of (1.5.5), i.e., the class of compound geometric distribution is a proper subclass of the class of compound Poisson distribution. However many infinitely divisible cf's appear in the special form (1.5.6) (Lukacs, E, 1960).

The negative binomial distribution with pgf

$$g(t) = (1-p)^k (1-pt)^{-k},$$

is compound geometric iff $k < 1$.

The compound exponential distribution having pgf of the form

$$\phi(t) = \frac{1}{1 - \log f(t)} \tag{1.5.7}$$

where $f(t)$ is also a infinitely divisible pgf, is infinitely divisible. The compound exponential distribution with pgf given by (1.5.7), coincide with the compound geometric form (1.5.6) and hence it is infinitely divisible (F.W. Steutel and K.V. Harn, 2004).

1.6 Review of the literatures on Infinitely Divisible Discrete Distributions

The theory of infinitely divisible distributions, developed primarily during the period from 1920 to 1950, has played a very important role in a variety of problems of probability theory and has been carried out along many lines. It plays a significant role in the solution of limit problems for sums of independent random variables. A full account of this theory and its applications, as it had been developed by the late 40's, were presented in the monographs of Levy (1937), Gnedenko and Kolmogorov (1968), and Loeve (1960).

Elementary infinitely divisible distributions which are formulated on the basis of simple models seem to be inadequate to describe under certain situations which may occur in a number of phenomena. In the last few years' research, various infinitely divisible distributions have been derived. Numerous new results have been obtained and entirely new applications have been found. In 1962, Mark Fisz gave a survey on recent developments in infinite divisibility. F. W. Steutel (1973) also surveyed on some recent results in infinitely divisibility and gave ideas on some basic theorems in infinite divisibility. Steutel (1968) also discussed methods of constructing infinitely divisible distributions mainly by mixing. Godambe and Patil (1969, 1975) consider a mixture of Poisson distributions where the mixing distributions have non-negative support. The importance of the property of infinite divisibility in modeling was stressed by Steutel (1983); see also the monograph by Steutel (1970) Pillai (1990), Pillai et al (1990, 1994). Some properties of infinitely divisible discrete distributions are, given by Johnson, et al (1992) and F.W. Steutel and K.V. Harn (2004). Begum and Borah (2003) studied certain Infinitely Divisible Discrete Probability Distributions and Applications.

The following two forms of pgf's of discrete distributions

$$G(t) = \frac{(1-\omega)g(t)}{1-\omega g(t)}, \qquad 0 < \omega < 1 \tag{1.6.1}$$

$$G(t) = \frac{\omega}{\omega + 1 - g(t)}, \quad \omega > 0 \tag{1.6.2}$$

were studied by Klebanov, Maniya and Melamed (1984), Steutel (1990) under the name of geometrically infinitely divisible distributions. The function (1.6.2) is an ID characteristic function (Steutel, 1968) which can be easily seen by writing

$$H(t) = \left[\frac{\omega}{\omega + 1 - g(t)}\right]^{1/n},$$

as a linear combination of cf's. In the above two forms (1.6.1) and (1.6.2), $g(t)$ is the pgf of the component distribution. Begum and Borah (2003) studied certain infinitely Divisible Discrete Probability Distributions and Applications.

The aim of this thesis is to derive, some generalized distributions of Poisson-Lindley (Sankaran, 1970), three parameter Charlier (Jain and Gupta, 1975) and Gegenbauer (Plankett and Jain, 1975) by considering them as the component distribution in the above models (1.6.1) and (1.6.2). The generalized distributions have been studied in the presence of their recurrence relations for their probabilities, factorial moments and cumulants with applications in different fields of social sciences, biological and ecological sciences, home injuries and accidents etc. The parameters are estimated and a few sets of reported data have been considered for the fitting of the distributions, and the fits are compared with that of other distributions.

Attempt has been also made to study the important properties and applications of some Lagarangian type of discrete probability distributions.

1.7 Synopsis of the Thesis

The thesis entitled by "An Empirical Study on Some Infinitely Divisible and Mixtures of Discrete Distributions" consists of seven chapters. The first chapter is an introductory one which gives an account of the relevant works done earlier by different authors in the theory of univariate discrete probability distributions. A brief description of the literatures on infinite divisible discrete probability distributions is also discussed.

The second chapter is a review on some well known discrete probability distributions, such as Poisson Lindley distribution (Sankaran, 1970), three parameter Charlier distribution (Jain and Gupta, 1975) and Gegenbauer distribution (Plankett and Jain, 1975). The above distributions are further investigated to study some of their important properties like probability recurrence relation, factorial moment recurrence relation and cumulant recurrence relation etc., with the estimation techniques of the parameters used by the respective authors. In the succeeding chapters we have made an attempt to study some of their generalized distributions by using the models (1.6.1) and (1.6.2) studied by Klebanov, Maniya and Melamed (1984) and Steutel (1990) under the name geometrically infinitely divisible distributions.

In chapter 3, two generalized distributions of Poisson-Lindley distribution viz., Generalized Poisson-Lindley Distribution of Type 1 (GPL1) and Generalized Poisson-Lindley Distribution of Type 2 (GPL2) have been obtained by using the models (1.6.1) and (1.6.2) respectively. Attempt has been made to derive the recurrence relations for probabilities, factorial moments and cumulants etc., for both of the generalized distributions. The problems of estimation of parameters and fitting of the distributions have been considered. For estimating the parameters of GPL1 and GPL2, we have used a composite method where one parameter is estimated by using Newton Raphson method and the other parameter is estimated by the method of moment. The derived distributions have been applied in different fields of biology, ecology and social information for empirical justifications where various modified forms of Poisson distribution were fitted by different authors. In all cases, our fitted distributions describe the data very well.

In chapter 4, using the model (1.6.1), two infinitely divisible distributions have been derived by considering three parameter Charlier (Jain and Gupta, 1975) and gegenbauer (Plankett and Jain, 1975) distributions respectively as the generalizing distribution. The derived distributions are denoted by Generalized Charlier Distribution of Type 1 (GCD1) and Generalized Gegenbauer Distribution of Type 1 (GGD1) respectively. Attempt has been made to study certain important properties of the derived distributions, such as the recurrence relations for probabilities, factorial moments and cumulants etc. The parameters of GCD1 and GGD1 have been estimated by an adhoc method of using first two sample moments and ratios of first three frequencies for rapid prediction, because of complexity of maximum likelihood method of estimation. A few sets of reported data on ecology and biology have been considered for empirical fitting of these distributions with satisfactory results. Some generalized distributions of Poisson, negative-binomial and Hermite distributions may be obtained as particular limiting cases of GCD1 and GGD1. These distributions have been also studied for some of their important properties with estimation of parameters. The fitting of the distributions have been also considered by using some published data in different fields of accidents, home injuries, biology and ecology.

In chapter 5, considering the second model (1.6.2), two infinitely divisible distributions viz., Generalized Charlier Distribution of Type 2 (GCD2) and Generalized Gegenbauer Distribution of Type 2 (GGD2) have been derived by using three parameter Charlier (Jain and Gupta, 1974) and Gegenbauer (Plankett and Jain, 1975) distributions respectively as the component distribution. Certain properties of the derived distributions, such as the recurrence relations for probabilities, moments, factorial moments and cumulants are studied. The parameters of GCD2 and GGD2 are estimated. Different applications of GCD2 and GGD2 are discussed. The fitted distributions are compared with that of the other distributions. The generalized distributions of Poisson, negative-binomial and Hermite may be obtained as particular limiting cases of GCD2 and GGD2. They have been also studied with estimation of their parameters and fitting of the distributions considering some published data on accidents, home injuries and also using some ecological and biological problems.

In chapter 6, we have made an attempt to derive a class of discrete probability distributions under the title, "Charlier family of Lagrangian (CFL) discrete probability distribution" by using the well known Lagrange's expansions, having wide flexibility and important implications in queueing theory. The probabilistic structures of general Lagrangian Charlier distributions of $type-I$ and $type-II$ have been derived by using Langrange expansion of first kind (LD1) and Lagrange expansion of second kind (LD2) respectively (according to Janardan and Rao's terminology). The basic Lagrangian Poisson (LP) and basic Lagrangian negative binomial (LNB) distributions have been investigated for some of their important properties. The parameters are estimated by using different methods. The fitting of the basic distributions have been considered for testing the validity of the estimates of the parameters. Further, the general Lagrangian Poisson negative binomial (LPNB), Lagrangian negative binomial Poisson (LNBP) and Lagrangian Poisson Logarithmic (LPL) distributions of $type-I$ and $type-II$ are also investigated. Some ad hoc methods are used to estimate the parameters of the distributions. It is also conceivable that discrete data occurring in ecology, epidemiology, and meteorology could be statistically modeled on the distributions considered in this investigation. It is found that Lagrangian probability distributions give better fit than their classical forms.

The Gegenbauer distribution (GD), which is a gamma-mixed Hermite distribution, is a very wider class of discrete probability distribution. The negative binomial and Hermite distributions may be obtained as its particular limiting cases. In the last chapter 7, attempt has been made to test the fitting of Gegenbauer distribution to the distribution of scores of teams and individuals in several sports involving ball games such as cricket, football and Hockey. It is found that in all cases, the Gegenbauer distribution provides better fit than the negative-binomial, Poisson, Hermite distribution.

Chapter 2

A Review on some discrete probability distributions

2.1 Introduction

The present chapter is a review on some well known discrete probability distributions, such as Poisson Lindley distribution by Sankaran (1970), three parameter Charlier distribution by Jain and Gupta (1975) and Gegenbauer distribution by Plankett and Jain (1975). These distributions have been discussed for some of their important properties like probability recurrence relation, moment recurrence relation, factorial moment recurrence relation and cumulant recurrence relation etc., with estimation techniques of the parameters used by the respective authors. In the succeeding chapters we have made an attempt to study some of their generalized geometrically infinitely divisible distributions based on the models (1.6.1) and (1.6.2) stated in chapter 1, studied by Klebanov, Maniya and Melamed (1984) and Steutel (1990).

2.2 Poisson-Lindley Distribution

Lindley (1958) derived a distribution known as Lindley distribution based on Baye's theorem with pmf

$$P_x(\theta) = \frac{\theta^2(\theta+2+x)}{(\theta+1)^{x+3}},\ x = 0,1,2,..., \tag{2.2.1}$$

Sankaran (1970) generalized Lindley distribution by mixing with Poisson distribution which is known as Poisson-Lindley distribution with application to mistakes in copying groups of random digits (data from Kemp and Kemp, 1965) and accidents to 647 women on high explosive shell in 5 weeks (data from Greenwood and Yule, 1920). In both cases, Poisson-Lindley distribution gives a better fit to the data than the Poisson distribution.

Borah and Begum (2002) studied certain properties of Poisson-Lindley and its derived distributions (Chapter 3). They fitted the generalized distributions of Poisson-Lindley to some biological and ecological data for empirical comparison. Poisson-Lindley distribution is a special case of Bhattacharya's (1966) more complicated mixed Poisson distribution. The Poisson-Lindley distribution (See Sankaran, 1970) has the pgf of the form

$$g(t)=\frac{\theta^2(\theta+2-t)}{(\theta+1)(\theta+1-t)^2},\quad \theta>0 \tag{2.2.2}$$

Its recurrence relation for probabilities may be written as

$$P_{r+1}=\frac{2(\theta+1)P_r-P_{r-1}}{(\theta+1)^2}, for\quad r\geq 1.\ \text{(Sankaran, 1970)} \tag{2.2.3}$$

where $P_0=\frac{\theta^2(\theta+2)}{(\theta+1)^3}$, $P_1=\frac{(\theta+3)\theta^2}{(\theta+1)^4}$.

The moment generating function (mgf) of Poisson Lindley distribution is

$$M(t)=\frac{\theta^2(\theta+2-e^t)}{(\theta+1)(\theta+1-e^t)^2} \tag{2.2.4}$$

Its moment recurrence relation may be written as

$$\mu'_{r+1}=\frac{2}{\theta^2}\left[\sum_{j=1}^{r}\binom{r+1}{j}\mu'_j\{(\theta+1)-2^{r-j}\}+\{(\theta+1)-2^r\}\right]-(\theta+1)^{-1}, r\geq 1 \tag{2.2.5}$$

where $\mu'_1=\frac{\theta+2}{\theta(\theta+1)}$.

The factorial moment generating function (fmgf) of Poisson Lindley distribution is

$$m(t)=\frac{\theta^2(\theta+1-t)}{(\theta+1)(\theta-t)^2}, \tag{2.2.6}$$

The recurrence relation for factorial moments may be written as

$$\mu_{(r+1)}=\frac{(r+1)\{2\theta\mu_{(r)}-r\mu_{(r-1)}\}}{\theta^2},\ r\geq 1 \tag{2.2.7}$$

where $\mu_{(1)}=\dfrac{\theta+2}{\theta(\theta+1)}$.

Again, the cumulant generating function (cgf) of the distribution is

$$K(t)=\log M(t)=\log\frac{\theta^2}{\theta+1}+\log(\theta+2-e^t)-2\log(\theta+1-e^t) \tag{2.2.8}$$

Its recurrence relation for cumulants may be written as

$$K_{r+1}=\frac{1}{\theta(\theta+1)}\left[\sum_{j=1}^{r}\binom{r}{j}K_{r-j+1}\{(2\theta+3)-2^j\}+(\theta+3)-2^r\right],r\geq 1 \tag{2.2.9}$$

where $K_1=\dfrac{\theta+2}{\theta(\theta+1)}$,

$$K_2=\frac{\theta^3+4\theta^2+6\theta+2}{(\theta+1)^2\theta^2}.$$

Estimation of parameter

The single parameter θ of Poisson-Lindley distribution may be estimated by using the method of moment (see Sankaran, 1970) remembering θ to be positive.

$$\theta=\frac{-(\mu-1)+\sqrt{(\mu-1)^2+8\mu}}{2\mu},\theta>0,\mu\neq 0 \tag{2.2.10}$$

2.3 Three Parameter Charlier distribution

Doetsch (1934), Meixner (1934, 1938) and Berg (1985) investigated the Charlier polynomials defined by the generating function

$$e^t(1-\beta t)^{-\lambda} \tag{2.3.1}$$

which are associated with the Poisson distribution of rare events. McBride (1971) discussed two parameter Charlier polynomials and some of their properties. Jain and Gupta (1975) defined the generalized Charlier polynomial by the generating function

$$e^{\alpha t}(1-\beta t^m)^{-\lambda}$$

Medhi and Borah (1986) studied generalized four parameter Charlier distribution with probability generating function

$$g(t)=e^{-\alpha}(1-\gamma-\beta)^{\lambda}e^{\alpha t}(1-\gamma t-\beta t^{m})^{-\lambda},\quad \alpha,\beta,\lambda\geq 0 \textit{ and } m=1,2,3,\ldots$$

This distribution includes negative binomial (Bliss and Fisher, 1953), Hermite (Kemp and Kemp, 1965), Gegenbauer (Plankett and Jain, 1975) and three parameter Charlier (Jain and Gupta, 1975) distributions as its particular limiting cases. Medhi and Borah (1986) investigated different properties of this distribution and also discussed the methods for fitting of four parameter Charlier distribution.

The pgf of three parameter Charlier (TPC) distribution (Jain and Gupta, 1975) is given by

$$g(t)=e^{-\alpha}(1-\beta)^{\lambda}e^{\alpha t}(1-\beta t)^{-\lambda},\quad \alpha,\beta,\lambda\geq 0 \tag{2.3.2}$$

Its probability recurrence relation may be obtained as

$$P_{r+1}=\frac{(\alpha+r\beta+\lambda\beta)P_r-\alpha\beta P_{r-1}}{(r+1)},\ r\geq 1. \tag{2.3.3}$$

where $\qquad P_0=e^{-\alpha}(1-\beta)^{\lambda}$

and $\qquad P_1=(\alpha+\lambda\beta)P_0$.

The mgf of three parameter Charlier distribution is

$$M(t)=e^{-\alpha}(1-\beta)^{\lambda}(1-\beta e^{t})^{-\lambda}\ e^{\alpha e^{t}} \tag{2.3.4}$$

Its moment recurrence relation may be written as

$$\mu'_{r+1}=\frac{1}{1-\beta}\left[\beta\sum_{j=1}^{r}\binom{r}{j}\mu'_{r-j+1}+(\alpha+\lambda\beta-\alpha\beta 2^{r})+\sum_{j=1}^{r}\binom{r}{j}\{\alpha+\lambda\beta-\alpha\beta 2^{j}\}\mu'_{j}\right],$$

$$r\geq 1 \tag{2.3.5}$$

where $\qquad \mu'_1=\alpha+\dfrac{\lambda\beta}{1-\beta}$.

The fmgf of three parameter Charlier distribution is

$$m(t)=(1-at)^{-\lambda}e^{\alpha t},\quad a=\frac{\beta}{1-\beta}. \tag{2.3.6}$$

Its recurrence relation for factorial moments may be obtained as

$$\mu_{(r+1)} = \frac{\{\beta(r+\lambda)+\alpha(1-\beta)\}\mu'_{(r)} - r\alpha\beta\mu'_{(r-1)}}{(1-\beta)} \ , \ r \geq 1 \tag{2.3.7}$$

where $\mu_{(1)} = \alpha + \dfrac{\lambda\beta}{1-\beta}$.

Taking logarithm on both sides of (2.3.4), the cgf may be written as

$$K(t) = \log M(t) = -\alpha + \alpha e^{t} + \lambda \log(1-\beta) - \lambda \log(1-\beta e^{t}) \tag{2.3.8}$$

Its cumulants recurrence relation may be written as

$$K_{r+1} = \frac{1}{(1-\beta)}\left\{\beta \sum_{j=1}^{r} \binom{r}{j} k_{r-j+1} + (\alpha + \lambda\beta) - \alpha\beta 2^{r}\right\}, \ r \geq 1 \tag{2.3.9}$$

where $$K_1 = \alpha + \frac{\lambda\beta}{1-\beta},$$

$$K_2 = \frac{\alpha + \lambda\beta - 2\alpha\beta + \beta\mu}{(1-\beta)} .$$

2.4 Gegenbauer Distribution

Plunkett and Jain (1975) defined Gegenbauer polynomials corresponding to the generating function

$$g(t) = (1 - \alpha t - \beta t^{2})^{-\lambda} \ , \ \alpha + \beta < 1, \ \lambda > 0 .$$

Further, they studied some of its limiting forms of the corresponding Gegenbauer distribution and the goodness of fit by estimating the parameters of the distribution by using the factorial moments. They fitted the distribution to a set of accident data. This distribution has a long history in the theory of stochastic processes.

The Gegenbauer distribution (Plunkett and Jain, 1975) obtained by mixing the Hermite distribution with the gamma distribution has the pgf of the form

$$g(t) = (1 - \alpha - \beta)^{\lambda}(1 - \alpha t - \beta t^{2})^{-\lambda}, \alpha + \beta < 1, \lambda > 0 . \tag{2.4.1}$$

The probabilities of Gegenbauer distribution can be expressed in terms of Gegenbauer polynomials (see Rainville, 1960). As particular cases of Gegenbauer distribution with pgf (2.4.1), Hermite distribution may be obtained by taking limit as $\lambda \to \infty$, $\alpha \to 0$ and $\beta \to 0$ such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$, and putting $\beta = 0$, negative binomial distribution is obtained. Borah (1984), has also studied the probability and moment properties of the three parameter Gegenbauer distribution (2.4.1), and has used estimation via the first two sample moments and the ratio of the first two sample frequencies. Medhi and Borah's (1984) studied four parameter generalized Gegenbauer distribution with the pgf

$$g(t) = (1-\alpha-\beta)^{\lambda}(1-\alpha t-\beta t^{m})^{-\lambda}, \tag{2.4.2}$$

where 'm' denotes integer number. They estimated the parameters by the method of moments and some ad hoc methods via sample mean, sample variance, and ratio of first two frequencies, assuming a known small integer value for m.

Kemp (1979) studied the Gegenbauer distribution with pgf of the form

$$g(t) = \frac{(1-Q_1 t)^{U_1}(1-Q_2 t)^{U_2}}{(1-Q_1)^{U_1}(1-Q_2)^{U_2}}, \tag{2.4.3}$$

which can be obtained with various restrictions on Q_1, Q_2, U_1 and U_2 by convoluting binomial and pseudo binomial variables. Kemp (1992a, 1992b) studied various forms of estimation, including Maximum Likelihood (ML) estimation. Factorizing the quadratic expressions in equation (2.4.2), the pgf can be expressed as

$$g(t) = \left[\frac{(1-a)(1+b)}{(1-at)(1+bt)}\right]^{\lambda}, \tag{2.4.4}$$

where $a-b=\xi$, $ab=\eta$ and $0<b<a<1$. This factorized form makes the distribution easy to handle, especially its moment properties. Mckendrick (1926) had also obtained the distribution with pgf (2.4.4) as the outcome of a non-homogeneous birth-and-death process with λ initial individuals.

The Gegenbauer distribution with the pgf (2.4.1), has the probability recurrence relation of the form

$$P_{r+1} = \frac{\{\alpha(r+\lambda)P_r + \beta(2\lambda+r-1)P_{r-1}\}}{(r+1)}, \quad r \geq 1 \tag{2.4.5}$$

where $\quad P_0 = (1-\alpha-\beta)^{\lambda}$,

and $\quad P_1 = \alpha\lambda P_0$.

The mgf of Gegenbauer distribution is

$$M(t)=\left[\frac{(1-\alpha-\beta)}{(1-\alpha e^{t}-\beta e^{2t})}\right]^{\lambda}$$

Its moment recurrence relation is

$$\mu'_{r+1}=A\left[\lambda\sum_{j=1}^{r}\binom{r}{j}(\alpha+2^{r-j+1}\beta)\mu'_{j}++\lambda(\alpha+\beta 2^{r+1})+\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^{j}\}\mu'_{r-j+1}\right],$$

$$r\geq 1 \tag{2.4.6}$$

where $A=\dfrac{1}{1-\alpha-\beta}$,

$$\mu'_{1}=\frac{\lambda(\alpha+2\beta)}{(1-\alpha-\beta)}.$$

The factorial moment generating function of Gegenbauer distribution is

$$m(t)=(1-at-bt^{2})^{-\lambda}, \tag{2.4.7}$$

where $a=\dfrac{\alpha+2\beta}{1-\alpha-\beta}$

and $b=\dfrac{\beta}{1-\alpha-\beta}$.

Its factorial moment recurrence relation is

$$\mu_{(r+1)}=\frac{\{(\alpha+2\beta)(r+\lambda)\mu_{(r)}+r\beta(r+2\lambda-1)\mu_{(r-1)}\}}{(1-\alpha-\beta)},\ r\geq 1 \tag{2.4.8}$$

where $\mu_{(1)}=\dfrac{\lambda(\alpha+2\beta)}{(1-\alpha-\beta)}$.

The cgf of Gegenbauer distribution is

$$K(t)=\log M(t)=\lambda\log(1-\alpha-\beta)^{\lambda}-\lambda\log(1-\alpha e^{t}-\beta e^{2t})^{-\lambda}. \tag{2.4.9}$$

Its cumulant recurrence relation is

$$k_{r+1}=\frac{\sum_{j=1}^{r}{}^{r}C_{j}k_{r-j+1}(\alpha+2^{j}\beta)+\lambda(\alpha+2^{r+1}\beta)}{(1-\alpha-\beta)},r\geq 1 \tag{2.4.10}$$

where $k_{1}=\dfrac{\lambda(\alpha+2\beta)}{(1-\alpha-\beta)}$,

and $k_{2}=\dfrac{(\alpha+2\beta)k_{1}+\lambda(\alpha+4\beta)}{(1-\alpha-\beta)}$.

Estimation procedures

The estimation procedures of three parameter Charlier (Medhi and Borah, 1986) and Gegenbauer (Medhi and Borah, 1984) distributions are shown in Table 2.1.

Table 2.1

Estimation procedures of $G_1^{\alpha,\beta,\lambda}(t)$ and $G_2^{\alpha,\beta,\lambda}(t)$

Dist.	Method of moments	Mean, variance and ratio of first two frequencies
$G_1^{\alpha,\beta,\lambda}(t)$	$\beta^{\wedge}=\dfrac{\mu_3+2\mu_1'-3\mu_2}{\mu_3-\mu_2}$, $\alpha^{\wedge}=\dfrac{\mu_1'-(1-\beta^{\wedge})\mu_2}{\beta^{\wedge}}$, $\lambda^{\wedge}=\dfrac{(\mu_1'-\alpha^{\wedge})(1-\beta^{\wedge})}{\beta^{\wedge}}$.	$\beta^{\wedge}=\dfrac{\mu_2+\theta-2\mu_1'}{\mu_2-\mu_1'}$, $\alpha^{\wedge}=\dfrac{\mu_1'(\mu_1'-\theta)-\theta(\mu_2-\mu_1')}{2\mu_1'-\theta-\mu_2}$ $\lambda^{\wedge}=\dfrac{(\mu_1'-\alpha)(1-\beta^{\wedge})}{\beta^{\wedge}}$, $\theta=\dfrac{f_1}{f_0}$.
$G_2^{\alpha,\beta,\lambda}(t)$	$\lambda^{\wedge}=\dfrac{3\mu(\mu_2-\mu)\pm\mu\sqrt{\{9(\mu-\mu_2)^2-4A\mu\}}}{2A}$, $\beta^{\wedge}=\dfrac{\lambda^{\wedge}(\mu_2-\mu)+\mu^2}{D}$, $\alpha=\dfrac{2\lambda^{\wedge}(2\mu-\mu_2)+2\mu^2}{D}$, where $A=\mu_3+2\mu-3\mu_2$, $D=2\lambda^2+3\lambda\mu-\mu^2-\lambda\mu_2$.	$f(\lambda)=A\lambda^2+B\lambda+C$, $\Rightarrow\hat{\lambda}=\dfrac{-B\pm\sqrt{B^2-4AC}}{2A}$ $A=2\{\theta-(2\mu-\mu_2)\}$, $B=(3\mu-\mu_2)\theta-2\mu^2$, $C=\theta\mu^2$ and $\theta=\dfrac{f_1}{f_0}=\alpha\lambda$. $\beta^{\wedge}=\dfrac{\lambda^{\wedge}(\mu_2-\mu)+\mu^2}{D}$, $\alpha=\dfrac{2\lambda^{\wedge}(2\mu-\mu_2)+2\mu^2}{D}$,

Here, $G_1^{\alpha,\beta,\lambda}(t)$ and $G_2^{\alpha,\beta,\lambda}(t)$ denotes respectively three parameter Charlier and Gegenbauer distributions.

2.5 Hermite distribution

Kemp and Kemp (1965) studied Hermite distribution as a Poisson mixture of Bernoulli distribution. By using the Gurland's theorem, the Hermite distribution can also be regarded as a Poisson mixture of binomial distributions which was used for biometrical data by Skellam (1952) and McGuire et al. (1957).

The Hermite distribution which is a Poisson-stopped-sum (generalized Poisson) distribution can also be derived as the sum $X = Y_1 + Y_2$, where Y_1 is a Poisson random variable with parameter a_1 and Y_2 is a Poisson random variable with parameter a_2. Mckendrick (1914, 1926) derived the distribution as the sum of two correlated Poisson random variables. Mckendrick (1926) fitted the distribution using estimation by the method of moments to the distribution to count of bacteria in leucocytes and obtained a very much better fit than with a Poisson distribution.

The pgf of Hermite distribution is

$$G(t) = \exp\{\alpha_1(t-1) + \alpha_2(t^2-1)\},\ \alpha_1, \alpha_2 > 0. \tag{2.5.1}$$

Its probability recurrence relation is

$$P_{r+1} = \frac{\alpha_1 P_r + 2\alpha_2 P_{r-1}}{(r+1)},\ r \geq 1. \tag{2.5.3}$$

where $\qquad P_0 = \exp\{-(\alpha_1 + \alpha_2)\},\ P_1 = \alpha_1 P_0.$

The cgf of the distribution is

$$K(t) = \log G(e^t) = \alpha_1(e^t - 1) + \alpha_2(e^{2t} - 1) \tag{2.5.4}$$

Its cumulant recurrence relation is

$$K_{r+1} = \alpha_1 + \alpha_2 2^{r+1},\ r \geq 1. \tag{2.5.6}$$

The first two cumulants are respectively

$$K_1 = \alpha_1 + 2\alpha_2,\ K_2 = K_1 + 2\alpha_2.$$

The Poisson and negative binomial distributions may be obtained as particular limiting cases of three parameter charlier (Jain and Gupta, 1975) and Gegenbauer distribution (Plankett and Jain, 1975).

Poisson distribution

The Poisson distribution which is a power series distribution with infinite non-negative integer support belongs to the exponential family of distributions. It has the pgf

$$g(t) = e^{\alpha(t-1)},\ \alpha \geq 0 \tag{2.5.7}$$

Its probability recurrence relation may be written as

$$P_{r+1} = \frac{\alpha P_r}{(r+1)}, \; r \geq 1. \tag{2.5.8}$$

where $\qquad P_0 = e^{-\alpha}, \; P_1 = \alpha P_0$.

Similarly, the factorial moment recurrence relation may be written as

$$\mu_{(r+1)} = \alpha \mu_{(r)} \;, \; r \geq 1, \;\; \mu_{(1)} = \alpha \; . \tag{2.5.9}$$

Its recurrence relation for moments may be obtained as

$$\mu'_{r+1} = \alpha \left\{ 1 + \sum_{j=1}^{r} \binom{r}{j} \mu'_j \right\}, \; r \geq 1, \; \mu'_1 = \alpha \, . \tag{2.5.10}$$

All cumulants of Poisson distribution are equal, each being equal to α .

i.e., $\qquad K_r = \alpha, \; r \geq 1$. $\qquad$ (2.5.11)

Negative binomial distribution

The negative binomial distribution has the pgf

$$g(t) = (1-\beta)^{\lambda}(1-\beta t)^{-\lambda}, \quad \beta, \lambda \geq 0 \, . \tag{2.5.12}$$

Its probability recurrence relation may also be written as

$$P_{r+1} = \frac{\beta(r+\lambda)P_r}{(r+1)}, \; r \geq 1. \tag{2.5.13}$$

where $\qquad P_0 = (1-\beta)^{\lambda}, \; P_1 = \lambda \beta P_0$.

The factorial moment recurrence relation may be written as

$$\mu_{(r+1)} = \frac{\beta(r+\lambda)\mu_{(r)}}{(1-\beta)} \;, \;\; r \geq 1 \tag{2.5.14}$$

where $\mu_{(1)} = \dfrac{\lambda\beta}{1-\beta}$.

The moment recurrence relation of negative binomial distribution is

$$\mu'_{r+1} = A\left[\lambda + \sum_{j=1}^{r} \binom{r}{j} \lambda \mu'_j + \sum_{j=1}^{r} \binom{r}{j} \mu'_{r-j+1} \right], r \geq 1 \tag{2.5.15}$$

where $\qquad \mu'_1 = \dfrac{\lambda\beta}{1-\beta}$ and $A = \dfrac{\beta}{1-\beta}$.

The cumulant recurrence relation may be written as

$$K_{r+1} = \frac{\beta}{(1-\beta)}\left\{\sum_{j=1}^{r}\binom{r}{j}k_{r-j+1} + \lambda\right\},\ r \geq 1 \tag{2.5.16}$$

where the first two cumulants of the distribution may be obtained as

$$K_1 = \frac{\lambda\beta}{1-\beta},\ K_2 = \frac{\beta(\lambda + K_1)}{(1-\beta)}.$$

In succeeding chapters, i.e., in chapters 3, 4 and 5 it will be shown that all the distributions and their recurrence relations mentioned above may be obtained as a limiting case of our generalized geometrically infinitely divisible distributions.

Chapter 3

Poisson-Lindley and its derived distributions

3.1 Introduction

The Poisson-Lindley distribution (Sankaran, 1970) is a one-parameter compound Poisson distribution which has wide applications in the theory of accident proneness. Sankaran (1970) generalized Lindley distribution (Lindley, 1958) by mixing with Poisson distribution which is known as Poisson-Lindley distribution. It has pgf

$$g(t) = \frac{\theta^2(\theta+2-t)}{(\theta+1)(\theta+1-t)^2}, \quad \theta > 0 \tag{3.1.1}$$

In the preceding chapter, we have already discussed properties of Poisson-Lindley distribution (Sankaran, 1970). In this present chapter, Poisson-Lindley distribution has been further studied under two known forms of geometric infinite divisibility [(1.6.1) and (1.6.2)], discussed in chapter1. The pgf of these generalized distributions are denoted by the symbols GPL1 and GPL2 respectively. Some properties of GPL1 and GPL2, such as the recurrence relations for probabilities, factorial moments and cumulants are also investigated. The problems of estimation of parameters and fitting of these distributions have been also considered. For estimating the parameters of GPL1 and GPL2, we have considered a composite method where one parameter is estimated by using Newton Raphson method and the other parameter is estimated by the method of moment. The distributions GPL1 and GPL2 are fitted to some numerical data in different fields of biology, ecology and social information for which various modified forms of Poisson distribution were fitted by different authors. In all cases, observing the values of chi squares and also comparing the observed and expected frequencies it is clearly seen that our fitted distributions describe the data very well.

3.2 Generalized Poisson-Lindley Distribution 1 (GPL1)

a) Probability recurrence relation

The pgf of the generalized Poisson-Lindley (GPL1) distribution derived from the model (1.6.1) may be written as

$$G(t) = \frac{(1-\omega)(\theta+2-t)\theta^2}{(\theta+1)(\theta+1-t)^2 - \omega\theta^2(\theta+2-t)}, 0 < \omega < 1,\ \theta > 0. \qquad (3.2.1)$$

Differentiating (3.2.1) w.r.t. 't' and then equating the coefficients of t^r on both sides we may obtain the probability recurrence relation as

$$P_{r+1} = \frac{\{2(\theta+1)^2 - \omega\theta^2\}P_r - (\theta+1)P_{r-1}}{(\theta+1)^3 - \omega\theta^2(\theta+2)}, \ r \geq 1. \qquad (3.2.2)$$

where $P_0 = A_1(1-\omega)(\theta+2)\theta^2$,

$P_1 = A_1[\{2(\theta+1)^2 - \omega\theta^2\}P_0 - (1-\omega)\theta^2$,

and $A_1 = \dfrac{1}{[(\theta+1)^3 - \omega\theta^2(\theta+2)]}$.

Putting $r = 1,2,3,\ldots$, in equation (3.2.2), higher order probabilities may be obtained as

$$P_2 = \frac{\{2(\theta+1)^2 - \omega\theta^2\}P_1 - (\theta+1)P_0}{(\theta+1)^3 - \omega\theta^2(\theta+2)},$$

$$P_3 = \frac{\{2(\theta+1)^2 - \omega\theta^2\}P_2 - (\theta+1)P_1}{(\theta+1)^3 - \omega\theta^2(\theta+2)},$$

$$P_4 = \frac{\{2(\theta+1)^2 - \omega\theta^2\}P_3 - (\theta+1)P_2}{(\theta+1)^3 - \omega\theta^2(\theta+2)}, \text{ etc.}$$

b) Factorial Moment recurrence relation

The factorial moment recurrence relation may be derived from its fmgf given by

$$m(t) = \frac{(1-\omega)\theta^2(1+\theta+t)}{(1+\theta)(\theta-t)^2 - \omega\theta^2(1+\theta-t)} \qquad (3.2.3)$$

Differentiating (3.2.3) w.r.t. 't' and then equating the coefficients of $\frac{t^r}{r!}$ on both sides, we may obtain the factorial moment recurrence relation as

$$\mu_{(r+1)} = \frac{(r+1)\left[\{2\theta(\theta+1) - \omega\theta^2\}\mu_{(r)} - r(\theta+1)\mu_{(r-1)}\right]}{(\theta+1)\theta^2(1-\omega)} \quad ,r \geq 1 \tag{3.2.4}$$

The first factorial moment of the distribution is

$$\mu_{(1)} = \frac{(\theta+2)}{\theta(1+\theta)(1-\omega)}.$$

Putting $r = 1,2,3,...$, in equation (3.2.4) higher order probabilities may be derived as

$$\mu_{(2)} = \frac{2\{2\theta(\theta+1) - \omega\theta^2\}\mu_{(1)} - 2(\theta+1)}{(\theta+1)(1-\omega)\theta^2},$$

$$\mu_{(3)} = \frac{3\{2\theta(\theta+1) - \omega\theta^2\}\mu_{(2)} - 6(\theta+1)\mu_{(1)}}{(\theta+1)(1-\omega)\theta^2},$$

$$\mu_{(4)} = \frac{4\{2\theta(\theta+1) - \omega\theta^2\}\mu_{(2)} - 12(\theta+1)\mu_{(1)}}{(\theta+1)(1-\omega)\theta^2}.$$

where $\mu_{(r)}$ denotes the r^{th} ordered factorial moment of the distribution. Thus mean and variance of the distribution are respectively given by

$$\mu = \frac{(\theta+2)}{(\theta+1)\theta(1-\omega)}, \tag{3.2.5}$$

and $$\sigma^2 = \frac{2\theta\mu + 3\theta^2\mu - 2\omega\theta^2\mu + \theta^2 - 2}{(1+\theta)\theta^2(1-\omega)}. \tag{3.2.6}$$

It is noted that the mean is less than the variance $\sigma^2 = C\mu + D$, when $C, D \neq 0$.

where $$C = \frac{\{2\theta + 3\theta^2 - 2\omega\theta^3\}}{(1+\theta)(1-\omega)\theta^2}$$

and $$D = (\theta^2 - 2)(1+\theta)(1-\omega)\theta^2.$$

c) Estimation of Parameters

From the equation (3.2.5), the parameter ω can be expressed as

$$\omega = 1 - \frac{(\theta+2)}{(1+\theta)\theta\mu}, \tag{3.2.7}$$

Again, from equation (3.2.5) and (3.2.6), the parameter ω is expressed in terms of θ, as

$$\omega = \frac{A\theta^2 + B\theta + 2}{2\mu\theta^2}, \tag{3.2.8}$$

where $\quad A = 3\mu + 1 - \dfrac{\sigma^2}{\mu}$ and $B = 2\left(\mu - \dfrac{\sigma^2}{\mu}\right)$.

Eliminating ω from equations (3.2.7) and (3.2.8), a functional equation for θ in terms of μ and σ^2 may be obtained as

$$f(\theta) = \theta^3 + 3\theta^2 + 2\theta - \frac{2}{T}, \tag{3.2.9}$$

where $\quad T = \mu + 1 - \dfrac{\sigma^2}{\mu}$.

The parameter θ may be estimated by the Newton-Raphson method. When 'θ' is known, 'ω' may be estimated either by using (3.2.5) or (3.2.6) as

$$\omega = \frac{\theta(1+\theta)\mu - (\theta+2)}{(\theta+1)\theta\mu},$$

or $$\omega = \frac{\theta(2+3\theta)\mu + \theta^2 - 2 - \theta(1+\theta)\sigma^2}{2\mu\theta^2 - \theta(1+\theta)\sigma^2} \text{ respectively}.$$

where μ and σ^2 respectively denote the mean and variance of the distribution.

3.3 Generalized Poisson-Lindley Distribution 2 (GPL2)

a) Probability recurrence relation

The pgf of Poisson-Lindley distribution derived from the model (1.6.2) may be given as

$$G(t) = \frac{\omega(\theta+1)(\theta+1-t)^2}{(1+\omega)(\theta+1)(\theta+1-t)^2 - (\theta+2-t)\theta^2}, \omega > 0, \theta > 0 \tag{3.3.1}$$

Differentiating (3.3.1) w.r.t. 't' and then equating the coefficients of t^r on both sides, the recurrence relation for probabilities may be obtained as

$$P_{r+1} = \frac{\{2(\omega+1)(\theta+1)^2 - \theta^2\}P_r - (\omega+1)(\theta+1)P_{r-1}}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)}, \; r > 1 \tag{3.3.2}$$

where $\quad P_0 = \dfrac{\omega(\theta+1)^3}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)}$,

$$P_1 = \frac{\{2(\omega+1)(\theta+1)^2 - \theta^2\}P_0 - 2\omega(\theta+1)^2}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)},$$

and $$P_2 = \frac{\{2(\omega+1)(\theta+1)^2 - \theta^2\}P_1 - (\omega+1)(\theta+1)P_0 + \omega(\theta+1)}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)}.$$

Putting $r = 2,3,4,...,$ in the equation (3.3.2), we have

$$P_3 = \frac{\{2(\omega+1)(\theta+1)^2 - \theta^2\}P_2 - (\omega+1)(\theta+1)P_1}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)},$$

$$P_4 = \frac{\{2(\omega+1)(\theta+1)^2 - \theta^2\}P_3 - (\omega+1)(\theta+1)P_2}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)}, \text{ etc.}$$

b) Factorial Moment recurrence relation

Factorial moment recurrence relation may be derived from its fmgf given as

$$m(t) = \frac{\omega(\theta+1)(\theta-t)^2}{(\omega+1)(1+\theta)(\theta-t)^2 - \theta^2(1+\theta-t)} \tag{3.3.3}$$

Differentiating (3.3.3) w.r.t. 't' and equating the coefficients of $\frac{t^r}{r!}$ on both sides, we have obtained the factorial moment recurrence relation as

$$\mu_{(r+1)} = \frac{(r+1)\{2\theta(\omega+1)(\theta+1) - \theta^2\}\mu_{(r)} - r(\omega+1)(\theta+1)\mu_{(r-1)}}{(\theta+1)\omega\theta^2}, \; for \; r > 1, \tag{3.3.4}$$

where $$\mu_{(1)} = \frac{(\theta+2)}{\omega\theta(1+\theta)},$$

and $$\mu_{(2)} = \frac{2\{2\theta(\omega+1)(\theta+1) - \theta^2\}\mu_{(1)} - 2(\theta+1)}{(\theta+1)\omega\theta^2}.$$

Putting $r = 2,3,...,$ in equation (3.3.4) we have

$$\mu_{(3)} = \frac{3\{2\theta(\omega+1)(\theta+1) - \theta^2\}\mu_{(2)} - 2(\omega+1)(\theta+1)\mu_{(1)}}{(\theta+1)\omega\theta^2},$$

$$\mu_{(4)} = \frac{4\{2\theta(\omega+1)(\theta+1) - \theta^2\}\mu_{(3)} - 3(\omega+1)(\theta+1)\mu_{(2)}}{(\theta+1)\omega\theta^2}, \text{ etc.}$$

Hence mean μ and variance σ^2 of the distribution are may be expressed as

$$\mu = \frac{(\theta+2)}{(\theta+1)\theta\omega}, \tag{3.3.5}$$

and
$$\sigma^2 = \frac{4\omega\theta\mu(\theta+1)+\theta\mu(\theta+2)+\theta^2-2}{\omega(\theta+1)\theta^2}. \tag{3.3.6}$$

c) Estimation of Parameters

From equation (3.3.5), the parameter ω can be expressed as

$$\omega = \frac{\theta+2}{\theta(\theta+1)\mu}, \tag{3.3.7}$$

Again from equation (3.3.5) and (3.3.6) the parameter ω can be expressed as

$$\omega = \frac{\theta^2\mu+2\theta\mu+\theta^2-2}{\theta(\theta\sigma^2-4\mu)(\theta+1)}. \tag{3.3.8}$$

Eliminating ω from (3.3.7) and (3.3.8), θ may be obtained from the equation

$$f(\theta) = A\theta^2 + B\theta + C, \tag{3.3.9}$$

The parameter θ is estimated by noting that

$$\theta = \frac{-B \pm \sqrt{B^2-4AC}}{2A},$$

where $A = \sigma^2 - \mu^2 - \mu$, $B = 2(\sigma^2 - \mu^2 - 2\mu)$ and $C = -6\mu$.

It may be noted that θ is positive. Hence, ω may be estimated either from

$$\omega = \frac{(\theta+2)}{\mu\theta(\theta+1)},$$

or
$$\omega = \frac{\mu\theta(\theta+2)+\theta^2-2}{\theta(\theta+1)(\theta\sigma^2-4\mu)}.$$

It is to be very important to note that, the Poisson-Lindley distribution (Sankaran, 1970) may be obtained as particular limiting case of GPL1 distribution, when we put $\omega = 0$ in (3.2.1).

3.4 Goodness of fit

The negative Binomial, the Poisson and the Neyman's distributions are commonly used in ecological and biological problems, the Neyman's and negative-Binomial distributions represent model in which the non-randomness is attributed to contagion. In this investigation, the parameter θ of generalized Poisson-Lindley distribution is estimated by Newton-Raphson method, whereas the parameter ω is estimated by the method of moments. The distributions GPL1 and GPL2 are fitted to some numerical data for which various modified forms of Poisson distribution were fitted by different authors.

In Table 3.1, we consider the data on the *Pyrausta nubilalis*; to which generalized Poisson distribution was fitted by Jain (1975) and Neyman type-A distribution was fitted by Beall and Rescia (1953). Observing the values of chi squares and also comparing the observed and expected frequencies it is clearly seen that generalized Poisson-Lindley distributions give better fit.

The Distribution of corn borers is considered in Table 3.2, to which Negative-Binomial and the Neyman's type-A distributions were fitted by Bliss and Fisher (1953). It will be seen that the data agree excellently with generalized Poisson-Lindley fit.

The generalized Poisson-Lindley distributions, i.e. GPL1, and GPL2 are also fitted to some data on the number of European red mites on apple leaves collected by P. Garman (see Bliss et al, 1953) in Table 3.3 for which Jain and Consul fitted generalized negative binomial distribution and Medhi and Borah (1984), fitted generalized Gegenbauer distribution. In this case also our fitted distributions are found to be satisfactory.

In Table 3.4, we considered the problem of accidents to 647 women on high explosive shalls in 5 week period (Data by Greenwood and Yule, 1920) for which Poisson-Lindley was fitted by Sankaran (1970).

The problem of mistakes in copying groups of random digits (Data by Kemp and Kemp, 1965) is considered in the Table 3.5, for which Poisson-Lindley was fitted by Sankaran (1970). Comparing the observed and expected frequencies it is clearly seen that our fitted distributions describe the data very well.

Table 3.1: Comparison of observed and fitted Poisson-Lindley, GPL1 and GPL2 and generalized Poisson Distributions

(Pyrausta nubilslis in 1937, data by Beall and Rescia 1940)

No.of insects	Observed frequency $\bar{x} = 0.75$ $s^2 = 1.2946$	Fitted distributions			Generalized Poisson (Jain ,1975)
		GPL1 $\hat{\omega}$=0.5276 $\hat{\theta}$=3.4556	GPL2 $\hat{\omega}$=0.1123 $\hat{\theta}$=12.7400	Poisson-Lindley (Sankaran,1970) $\hat{\theta}$=1.8082	
0	33	31.86	33.10	31.49	32.46
1	12	13.84	12.49	14.16	13.45
2	6	5.92	5.50	6.09	5.60
3	3	2.52	2.49	2.54	2.42
4	1	1.06	1.13	1.04	1.08
5	1	0.80	1.29	0.42	0.97
Total	56	56.00	56.00	56.00	56.00
	χ^2	0.3743	0.0667	0.6532	0.2500
	df	1	1	2	
	$p-value$	> 0.54	> 0.79	> 0.72	

Table 3.2: Comparison of observed and fitted Poisson-Lindley, GLP1 and GLP2 distributions (Corn Borers data of Beall and Rescia, 1940)

No. of Insects	Observed Frequency $\bar{x} = 1.4833$ $s^2 = 3.1664$	Fitted distributions		
		GPL1 $\hat{\theta}$=3.4556 $\hat{\omega}$=0.7611	GPL2 $\hat{\theta}$=6.6400 $\hat{\omega}$=0.1148	Poisson-lindley (Sankaran,1970) $\hat{\theta}$=1.0096
0	43	48.05	49.11	45.36
1	35	29.04	24.83	30.07
2	17	17.36	15.30	18.70
3	11	10.35	9.84	11.16
4	5	6.16	6.38	6.48
5	4	3.67	4.14	3.68
6	1	2.18	2.69	2.06
7	2	1.30	1.75	1.14
8	2	1.89	1.14	0.62
Total	120	120.00	120.00	120.00
	χ^2	1.7289	1.7251	0.0667
	df	4	4	5
	$p-value$	> 0.78	> 0.78	> 0.99

GPL1: Generalized Poisson-Lindley Distribution of Type 1

GPL2: Generalized Poisson-Lindley Distribution of Type 2

Table 3.3: Comparison of observed and fitted Poisson-Lindley, GLP1, GLP2 and generalized negative binomial distributions to the Count of the number of European red mites on apple leaves

(Bliss et al 1953)

No.of mites per leaf	Leaves (observed) $\bar{x} = 1.1467$ $s^2 = 2.2585$	Fitted distributions Poisson lindley (Sankaran,1970) $\hat{\theta}=1.258$	GPL1 $\hat{\theta}=1.3921$ $\hat{\omega}=0.1117$	GPL2 $\hat{\theta}=8.5800$ $\hat{\omega}=0.1022$	Gen.Neg.Bin. (Jain and Consul, 1971)
0	70	67.19	67.62	70.89	71.48
1	38	38.89	38.68	33.35	33.98
2	17	21.26	21.04	18.70	19.80
3	10	11.21	11.09	10.84	11.59
4	9	5.76	5.70	6.73	6.57
5	3	2.90	2.92	3.69	3.55
6	2	1.44	1.47	2.15	1.80
7	1	0.71	0.74	1.26	0.84
8	0	0.34	0.71	0.73	0.39
Total	150	150.00	150.00	150.00	150.00
	χ^2	3.0136	2.8491	2.4433	2.0700
	df	4	3	3	3
	$p-value$	> 0.55	> 0.41	> 0.48	> 0.55

Table 3.4: Comparison of observed and fitted Poisson-Lindley, GLP1, GLP2 and negative binomial, distributions to the Accidents to 647 women on high explosive shells in 5 weeks

(Data by Greenwood and Yule, 1920)

No. of accidents	Observed frequency $\bar{x} = 0.4652$ $s^2 = 0.6903$	Fitted distributions GPL1 $\hat{\omega}=0.8145$ $\hat{\theta}=12.45$	GPL2 $\hat{\omega}=0.1189$ $\hat{\theta}=18.99$	Poisson-Lindley (Sankaran,1970) $\hat{\theta} = 2.728$	Neg-binomial (Plunkett and Jain, 1975)
0	447	441.52	449.03	441	445.89
1	132	140.27	130.21	143	134.90
2	42	44.52	44.57	45	44.00
3	21	14.13	15.26	14	14.69
4	3	4.48	5.22	4	4.94
5	2	2.08	2.71	1	2.56
Total	647	647.00	647.00	647	647.00
	χ^2	4.0948	3.4236	4.62	3.6315
	df	2	2	3	2
	$p-value$	> 0.12	> 0.18	> 0.20	> 0.15

GPL1: Generalized Poisson-Lindley Distribution of Type 1

GPL2: Generalized Poisson-Lindley Distribution of Type 2

Table 3.5: Distribution of mistakes in copying groups of random digits with expected frequencies obtained by fitting Poisson-Lindley, GLP1 and GLP2 distributions.

(Sankaran, 1970) (Data from Kemp and Kemp, 1965)

No.of insects	Observed frequency $\bar{x} = 0.7833$ $s^2 = 1.2364$	Fitted distributions		
		GPL1 $\hat{\omega}$=0.8329 $\hat{\theta}$=8.4500	GPL2 $\hat{\omega}$=0.1916 $\hat{\theta}$=7.4500	Poisson-Lindley (Sankaran,1970) $\hat{\theta} = 1.74$
0	35	33.63	35.67	33.0
1	11	14.79	12.59	15.3
2	8	6.50	6.08	6.8
3	4	2.85	2.93	2.9
4	2	2.23	2.73	1.2
Total	60	60.00	60.00	60.00
	χ^2	1.5398	0.8401	2.4219
	df	1	1	2
	$p-value$	> 0.21	> 0.35	> 0.29

GPL1: Generalized Poisson-Lindley Distribution of Type 1

GPL2: Generalized Poisson-Lindley Distribution of Type 2

From the above Tables it is clear that there is some improvement however small it may be in fitting these distributions over the other distributions considered earlier. The fitting of these distributions as indicated here may be used in other situations also.

Chapter 4

Generalized Infinitely Divisible Distributions of Type 1

4.1 Introduction

Elementary infinitely divisible distributions which are formulated on the basis of simple models seem to be inadequate to describe the situations which may occur in a number of phenomena. In the last few years' research a number of infinitely divisible distributions have been derived. Numerous new results have been obtained and entirely new applications have been found.

In this chapter 4, an attempt has been made to derive some generalized distributions of three parameter Charlier and Gegenbauer by considering them as the component distribution $g(t)$, in the model

$$G(t) = \frac{(1-\omega)g(t)}{1-\omega g(t)}, \qquad 0 < \omega < 1 \tag{4.1.1}$$

studied by Steutel (1968, 1990) and Klebanov, Maniya and Melamed (1984) under the name of geometrically infinitely divisible distribution, where $G(t)$ denotes the pgf of the generalized distribution. The generalized distributions of Charlier and Gegenbauer derived from the model (4.1.1) are denoted by the symbols GCD1 and GGD1 respectively. Further, certain important properties of these newly derived distributions such as the recurrence relations for probabilities, factorial moments and cumulants are investigated. The parameters of GCD1 and GGD1 are estimated by considering a composite method of using first two sample moments and ratios of first three frequencies.

To illustrate various applications of these distributions, they are fitted to some well known data e.g., the number of plants per quadrant of Lespedeza capitate (data of Beall and Rescia, 1953), the distributions of purchases of two different kinds of brands (brand K and D) of products (data of Chatfield, 1969) and data on chemically induced chromosome aberrations in cultures of human leukocytes in which Loeschcke and Kohler (1976) recommended the use of negative binomial distribution (NBD), while Janardan and Schaeffer (1977) have used a modified Poisson distribution and generalized Poisson distribution (GPD) is used by Consul (1989). In all the cases our distributions provide better fit to the observed data than the earlier ones.

It is seen that some generalized infinitely divisible distributions of Poisson, negative binomial and Hermite mixing may be obtained as particular limiting cases of GCD1 and GGD1. Certain properties of these limiting distributions are also investigated and the distributions are fitted to some published data in different fields of biology, ecology, home injuries, accidents and social information.

4.2 Generalized Charlier Distribution of Type 1(GCD1)

a) Probability recurrence relation

The probability generating function (pgf) of generalized three parameter Charlier distribution of type 1 (GCD1) derived from the model (4.1.1) may be written as

$$G(t)=\frac{(1-\omega)e^{-\alpha}(1-\beta)^{\lambda}e^{\alpha t}(1-\beta t)^{-\lambda}}{1-\omega e^{-\alpha}(1-\beta)^{\lambda}e^{\alpha t}(1-\beta t)^{-\lambda}}, \quad 0<\omega<1 \tag{4.2.1}$$

Differentiating both sides of (4.2.1) with respect to (w.r.t) 't', we get

$$(1-\beta t)G'(t)=\frac{\omega}{1-\omega}\left[(\alpha+\lambda\beta)G(t)G(t)-\alpha\beta tG(t)G(t)\right]+$$
$$\{(\alpha+\lambda\beta)G(t)-\alpha\beta tG(t)\} \tag{4.2.2}$$

Expanding (4.2.2) as a power series in 't' and equating the coefficients of t^r on both sides, the recurrence relation for probabilities may be obtained as

$$G_{r+1}=\frac{\omega}{(1-\omega)(r+1)}\left\{(\alpha+\lambda\beta)G_0G_r+\sum_{j=1}^{r}\{(\alpha+\lambda\beta)G_j-\alpha\beta G_{j-1}\}G_{r-j}\right\}+$$

$$\frac{(\alpha+\lambda\beta+r\beta)G_r-\alpha\beta G_{r-1}}{(r+1)}\ ,\ r\geq 1. \tag{4.2.3}$$

where $$G_0=\frac{(1-\omega)e^{-\alpha}(1-\beta)^{\lambda}}{1-\omega e^{-\alpha}(1-\beta)^{\lambda}}, \tag{4.2.4}$$

and $$G_1=(\alpha+\lambda\beta)\left\{1+\frac{\omega G_0}{1-\omega}\right\}G_0. \tag{4.2.5}$$

b) Factorial moment recurrence relation

The factorial moment generating function (fmgf) of GCD1 may be written as

$$m(t)=\frac{(1-\omega)e^{\alpha t}(1-ae^{t})^{-\lambda}}{1-\omega e^{\alpha t}(1-ae^{t})^{-\lambda}},\ 0<\omega<1 \tag{4.2.6}$$

where $a=\dfrac{\beta}{1-\beta}$.

Differentiating the equation (4.2.6) w.r.t.'t', we get

$$(1-at)m'(t)=\frac{\omega}{1-\omega}\{(\alpha+\lambda a)-a\alpha t\}m(t)m(t)+\{(\alpha+\lambda a)-a\alpha t\}m(t) \tag{4.2.7}$$

Expanding (4.2.7) as a power series in 't' and equating the coefficients of $\dfrac{t^r}{r!}$ on both sides the recurrence relation for factorial moments may be obtained as

$$\mu_{(r+1)}=\frac{\omega}{(1-\omega)}\left[(\alpha+\lambda a)\sum_{j=1}^{r}\binom{r}{j}\mu_{(j)}\mu_{(r-j)}-a\alpha\sum_{j=1}^{r}\binom{r}{j}\mu_{(j-1)}\mu_{(r-j)}-ar\mu_{(r)}\right]$$
$$+\frac{\{\alpha+a(\lambda+r)\}\mu_{(r)}-ar\alpha\mu_{(r-1)}}{(1-\omega)},\ r\geq 1 \tag{4.2.8}$$

where $\mu_{(1)}=\dfrac{\alpha(1-\beta)+\lambda\beta}{(1-\omega)(1-\beta)}$.

Hence mean and variance of GCD1 are respectively

$$\mu=\frac{\alpha(1-\beta)+\lambda\beta}{(1-\omega)(1-\beta)}, \tag{4.2.9}$$

$$\sigma^2=\mu+\omega\mu^2+\frac{\lambda\beta^2}{(1-\omega)(1-\beta)^2}, \tag{4.2.10}$$

c) Moment recurrence relation

The moment generating function (mgf) of GCD1 may be written as

$$M(t)=\frac{(1-\omega)e^{-\alpha}(1-\beta)^{\lambda}e^{\alpha e^{t}}(1-\beta e^{t})^{-\lambda}}{1-\omega e^{-\alpha}(1-\beta)^{\lambda}e^{\alpha e^{t}}(1-\beta e^{t})^{-\lambda}},\ 0<\omega<1 \tag{4.2.11}$$

Differentiating the equation (4.2.11) w.r.t.'t', we get

$$(1-\beta e^{t})M'(t)=\frac{\omega}{1-\omega}\{(\alpha+\lambda\beta)-\alpha\beta e^{t}\}e^{t}M(t)M(t)+\{(\alpha+\lambda\beta)-\alpha\beta e^{t}\}e^{t}M(t)$$

Expanding and equating the coefficients of $\frac{t^r}{r!}$ on both sides, the recurrence relation for moments may be obtained as

$$\mu'_{r+1}=A\left[(\alpha+\lambda\beta-\alpha\beta 2^{r})+\sum_{j=1}^{r}\binom{r}{j}(\alpha+\lambda\beta-\alpha\beta 2^{r-j})\mu'_j+\beta\sum_{j=1}^{r}\binom{r}{j}\mu'_{r-j+1}\right]$$

$$+B\left[\sum_{i=0}^{r}\binom{r}{i}(\alpha+\lambda\beta-\alpha\beta 2^{r-i})\sum_{j=0}^{i}\binom{i}{j}\mu'_j\mu'_{i-j}\right],\ r\geq 1 \tag{4.2.12}$$

where $A=\frac{1}{1-\beta},\ B=\frac{\omega}{(1-\beta)(1-\omega)}$

and $\mu'_1=\frac{\alpha(1-\beta)+\lambda\beta}{(1-\omega)(1-\beta)}.$

d) Cumulant recurrence relation

The cumulant generating function (cgf) of GCD1 may be obtained by taking logarithm on both sides of (4.2.11) as

$$K(t)=\log(1-\omega)+\log m(t)-\log\{1-\omega m(t)\} \tag{4.2.13}$$

where $m(t)=e^{-\alpha}(1-\beta)^{\lambda}e^{\alpha e^{t}}(1-\beta e^{t})^{-\lambda}$ is the mgf of three parameter Charlier distribution.

Differentiating (4.2.13) w.r.t. 't' and equating the coefficients of $\frac{t^r}{r!}$ on both sides, the cumulant recurrence relation of GCD1 may be obtained as

$$K_{r+1}=A\sum_{j=1}^{r}\binom{r}{j}K_{r-j+1}+B\left[\sum_{j=1}^{r}\binom{r}{j}\{(\alpha+\lambda\beta)-\alpha\beta 2^{r-j}\}\mu'_j+\{\alpha+\lambda\beta-\alpha\beta 2^{r}\}\frac{1}{\omega}\right],$$

$$r\geq 1 \tag{4.2.14}$$

where $A=\frac{\beta}{1-\beta}$ and $B=\frac{\omega}{(1-\beta)(1-\omega)}.$

Here $\mu_r^{'}$ denotes the r^{th} moment of the distribution.

The first and second cumulants of the distribution are respectively

$$K_1 = \frac{\alpha(1-\beta)+\lambda\beta}{(1-\omega)(1-\beta)},$$

and $$K_2 = K_1 + \omega K_1^2 + \frac{\lambda\beta^2}{(1-\omega)(1-\beta)^2}.$$

e) Particular cases of GCD1

The distributions obtained from GCD1 as its particular cases are given below in a tabular form.

Sl.No.	Parameter values	Name of Dist.	pgf $g(t)$
1.	$\omega = 0$	Three parameter Charlier	$e^{\alpha(t-1)}(1-\beta)^{\lambda}(1-\beta t)^{-\lambda}$, $\alpha,\beta,\lambda \geq 0$
2.	$\omega = 0, \alpha = 0$	Negative binomial	$(1-\beta)^{\lambda}(1-\beta t)^{-\lambda}, \lambda,\beta \geq 0,$
3.	$\omega = 0, \beta = 0$	Poisson	$e^{\alpha(t-1)}, \alpha > 0$

From the above Table it is clear that from (4.2.1), Poisson mixing infinitely divisible distribution (GPD1) may be obtained by putting $\beta = 0$ and negative binomial mixing infinitely divisible distribution (GNBD1) may be obtained putting $\alpha = 0$.

1) Three parameter Charlier distribution

The three parameter Charlier distribution (Jain and Gupta, 1975) has the pgf

$$g(t) = e^{-\alpha}(1-\beta)^{\lambda} e^{\alpha t}(1-\beta t)^{-\lambda}, \quad \alpha,\beta,\lambda \geq 0$$

This distribution may also be obtained as a limiting case of GCD1, when we put $\omega = 0$ in (4.2.1). The probability recurrence relation of three parameter Charlier distribution is

$$P_{r+1} = \frac{(\alpha + r\beta + \lambda\beta)P_r - \alpha\beta P_{r-1}}{(r+1)}, \; r \geq 1.$$

where $P_0 = e^{-\alpha}(1-\beta)^{\lambda}$, $P_1 = (\alpha + \lambda\beta)P_0$.

This is obtained by putting $\omega = 0$ in (4.2.3).

Similarly putting $\omega = 0$ in (4.2.8), the factorial moment recurrence relation of three parameter charlier distribution may be obtained as

$$\mu_{(r+1)} = \frac{\{\beta(r+\lambda)+\alpha(1-\beta)\}\mu_{(r)} - r\alpha\beta\mu_{(r-1)}}{(1-\beta)} \quad , \quad r \geq 1$$

where $\mu_{(1)} = \alpha + \frac{\lambda\beta}{1-\beta}$.

By putting $\omega = 0$, in (4.2.12) the moment recurrence relation of three parameter charlier distribution may be obtained as

$$\mu'_{r+1} = A\left[(\alpha+\lambda\beta-\alpha\beta 2^r) + \sum_{j=1}^{r}\binom{r}{j}(\alpha+\lambda\beta-\alpha\beta 2^{r-j})\mu'_j + \beta\sum_{j=1}^{r}\binom{r}{j}\mu'_{r-j+1}\right]$$

where $\qquad \mu'_1 = \alpha + \frac{\lambda\beta}{1-\beta}$, $A = \frac{1}{1-\beta}$.

Again, putting $\omega = 0$, in (4.2.14) the cumulant recurrence relation of three parameter charlier distribution may be obtained as

$$K_{r+1} = \frac{1}{(1-\beta)}\left\{\beta\sum_{j=1}^{r}\binom{r}{j}k_{r-j+1} + (\alpha+\lambda\beta) - \alpha\beta 2^r\right\}, \; r \geq 1$$

where $\qquad K_1 = \alpha + \frac{\lambda\beta}{1-\beta}$ and $K_2 = \frac{\alpha+\lambda\beta-2\alpha\beta+\beta\mu}{(1-\beta)}$.

2) Negative binomial distribution

The negative binomial distribution has the pgf

$$g(t) = (1-\beta)^{\lambda}(1-\beta t)^{-\lambda}, \quad \beta, \lambda \geq 0 .$$

This distribution may also be obtained as a limiting case of GCD1 when we put $\omega = 0$ and $\alpha = 0$ in (4.2.1). The probability recurrence relation of negative binomial distribution may also be obtained from (4.2.3), by putting $\omega = 0$, $\alpha = 0$ as

$$P_{r+1} = \frac{\beta(r+\lambda)P_r}{(r+1)}, \; r \geq 1.$$

where $\qquad P_0 = (1-\beta)^{\lambda}$ and $P_1 = \lambda\beta P_0$.

Similarly putting $\omega = 0$ and $\alpha = 0$, the factorial moment recurrence relation of negative binomial distribution may be obtained from (4.2.8) as

$$\mu_{(r+1)} = \frac{\beta(r+\lambda)\mu_{(r)}}{(1-\beta)} \quad , \quad r \geq 1$$

where $\mu_{(1)} = \frac{\lambda\beta}{1-\beta}$.

The moment recurrence relation of negative binomial distribution is

$$\mu'_{r+1} = A\left[\lambda + \sum_{j=1}^{r}\binom{r}{j}\lambda\mu'_j + \sum_{j=1}^{r}\binom{r}{j}\mu'_{r-j+1}\right], r \geq 1$$

where $\qquad \mu'_1 = \dfrac{\lambda\beta}{1-\beta}$

and $\qquad A = \dfrac{\beta}{1-\beta}$.

This may be obtained from (4.2.12) by putting $\omega = 0$ and $\alpha = 0$.

Putting $\omega = 0$ in (4.2.14), the cumulant recurrence relation of negative binomial distribution may be obtained as

$$K_{r+1} = \frac{\beta}{(1-\beta)}\left\{\sum_{j=1}^{r}\binom{r}{j}k_{r-j+1} + \lambda\right\},\ r \geq 1$$

where $\qquad K_1 = \dfrac{\lambda\beta}{1-\beta}$

and $\qquad K_2 = \dfrac{\beta(\lambda + K_1)}{(1-\beta)}$.

3) Poisson distribution

The Poisson distribution which is a power series distribution with infinite non-negative integer support has the pgf

$$g(t) = e^{\alpha(t-1)},\quad \alpha \geq 0$$

This distribution belongs to the exponential family of distributions, may be obtained from (4.2.1) as a limiting form of GCD1 by putting $\omega = 0$ and $\beta = 0$.

Again by putting $\omega = 0$ and $\beta = 0$, in (4.2.3) the probability recurrence relation of Poisson distribution may be obtained as

$$P_{r+1} = \frac{\alpha P_r}{(r+1)},\ r \geq 1.$$

where $\qquad P_0 = e^{-\alpha}$,

and $\qquad P_1 = \alpha P_0$.

Similarly putting $\omega = 0$ and $\beta = 0$, in (4.2.8) the factorial moment recurrence relation of Poisson distribution may be obtained as

$$\mu_{(r+1)} = \alpha\mu_{(r)} \ , \ r \geq 1,$$

where $$\mu_{(1)} = \alpha \ .$$

Again by putting $\omega = 0$ and $\beta = 0$, in (4.2.12) the recurrence relation for moments may be obtained as

$$\mu'_{r+1} = \alpha\left\{1 + \sum_{j=1}^{r}\binom{r}{j}\mu'_j\right\}, \ r \geq 1.$$

where $$\mu'_1 = \alpha \ .$$

Putting $\omega = 0$ and $\beta = 0$ in (4.2.14) we get all cumulants of Poisson distribution are equal to α.

i.e., $$K_r = \alpha \ , \ r \geq 1.$$

4.3 Generalized Gegenbauer Distribution of Type 1(GGD1)

a) Probability recurrence relation

The pgf of generalized Gegenbauer distribution of type 1 (GGD1), derived from the model (4.1.1) may be written as

$$G(t) = \frac{(1-\omega)(1-\alpha-\beta)^{\lambda}(1-\alpha t-\beta t^2)^{-\lambda}}{1-(1-\alpha-\beta)^{\lambda}(1-\alpha t-\beta t^2)^{-\lambda}} \ , \ 0 < \omega < 1 \tag{4.3.1}$$

Differentiating (4.3.1) w.r.t. 't' we get

$$(1-\alpha t-\beta t^2)G'(t) = \frac{\omega\lambda}{1-\omega}\left[(\alpha+2\beta t)G(t)G(t)\right] + \lambda(\alpha+2\beta t)G(t) \tag{4.3.2}$$

Expanding (4.3.2) and equating the coefficients of t^r both sides, the recurrence relation for probabilities may be obtained as

$$G_{r+1} = \frac{\omega\lambda}{(1-\omega)(r+1)}\left\{\alpha G_0 G_r + \sum_{j=1}^{r}\{\alpha G_j + 2\beta G_{j-1}\}G_{r-j}\right\} +$$

$$\frac{\alpha(r+\lambda)G_r + \beta(2\lambda+r-1)G_{r-1}}{(r+1)} \ , \ r \geq 1. \tag{4.3.3}$$

where $$G_0 = \frac{(1-\omega)(1-\alpha-\beta)^{\lambda}}{1-\omega(1-\alpha-\beta)^{\lambda}}, \tag{4.3.4}$$

and $$G_1 = \frac{\lambda\alpha G_0}{1-\omega(1-\alpha-\beta)^{\lambda}} = \alpha\lambda\left\{1 + \frac{\omega G_0}{1-\omega}\right\}G_0 . \tag{4.3.5}$$

b) Factorial moment recurrence relation

Corresponding to the pgf (4.3.1), the fmgf may be written in the following form

$$m(t) = \frac{(1-\omega)(1-at-bt^2)^{-\lambda}}{1-\omega(1-at-bt^2)^{-\lambda}}, \quad 0<\omega<1 \tag{4.3.6}$$

where $$a = \frac{\alpha+2\beta}{1-\alpha-\beta}$$

and $$b = \frac{\beta}{(1-\alpha-\beta)}.$$

Differentiating the equation (4.3.6) w.r.t 't', we get

$$(1-at-bt^2)m'(t) = \frac{\omega\lambda}{1-\omega}\{a+2bt\}m(t)m(t) + \lambda\{a+2bt\}m(t), \tag{4.3.7}$$

Now equating the coefficients of $\frac{t^r}{r!}$ on both sides of (4.3.7), the factorial moment recurrence relation may be obtained as

$$\mu_{(r+1)} = \frac{\omega\lambda}{1-\omega}\left[a\sum_{j=0}^{r}\binom{r}{j}\mu_{(j)}\mu_{(r-j)} + 2b\sum_{j=1}^{r}\binom{r}{j}j\mu_{(j-1)}\mu_{(r-j)}\right]$$

$$+a(\lambda+r)\mu_{(r)} + rb\{2\lambda+(r-1)\}\mu_{(r-1)}, \; r\geq 1 \tag{4.3.8}$$

The first factorial moment of the distribution is

$$\mu_{(1)} = \frac{\lambda(\alpha+2\beta)}{(1-\omega)(1-\alpha-\beta)}.$$

The mean and the variance of the distribution are respectively

$$\mu = \frac{\lambda(\alpha+2\beta)}{(1-\omega)(1-\alpha-\beta)}, \tag{4.3.9}$$

and $$\sigma^2 = \omega K_1^2 + \frac{(1+\beta)\mu}{(1-\alpha-\beta)} + \frac{2\lambda\beta}{(1-\omega)(1-\alpha-\beta)}. \tag{4.3.10}$$

c) Moment recurrence relation

The mgf of GGD1, corresponding to the pgf (4.3.1) may be written as

$$M(t) = \frac{(1-\omega)(1-\alpha-\beta)^{\lambda}(1-\alpha e^t-\beta e^{2t})^{-\lambda}}{1-\omega(1-\alpha-\beta)^{\lambda}(1-\alpha e^t-\beta e^{2t})^{-\lambda}}, \quad 0<\omega<1 \tag{4.3.11}$$

Differentiating both sides of (4.3.11) w.r.t 't', we get

$$(1-\alpha e^t-\beta e^{2t})M'(t)=\frac{\omega\lambda}{1-\omega}\{(\alpha+2\beta e^t\}e^t M(t)M(t)+\lambda\{(\alpha+2\beta e^t\}e^t M(t)$$

Equating the coefficients of $\frac{t^r}{r!}$ on both sides, the recurrence relation for moments may be obtained as

$$\mu'_{r+1}=A\left[\lambda(\alpha+\beta 2^{r+1})+\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^j)\mu'_{r-j+1}+\lambda\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^{r-j+1})\mu'_j\right]$$

$$+B\left[\lambda\sum_{i=0}^{r}\binom{r}{i}(\alpha+\beta 2^{r-i+1})\sum_{j=0}^{i}\binom{i}{j}\mu'_j\mu'_{i-j}\right],\ r\geq 1 \qquad (4.3.12)$$

where $\mu'_1=\frac{\lambda(\alpha+2\beta)}{(1-\omega)(1-\alpha-\beta)}$,

$$A=\frac{1}{1-\alpha-\beta},\quad B=\frac{\omega}{(1-\alpha-\beta)(1-\omega)}$$

d) Cumulant recurrence relation

Taking logarithm on both sides of (4.3.11) we get the cgf of GGD1 as

$$K(t)=\log(1-\omega)+\log m(t)-\log\{1-\omega m(t)\} \qquad (4.3.13)$$

where $m(t)=(1-\alpha-\beta)^{\lambda}(1-\alpha e^t-\beta e^{2t})^{-\lambda}$,

Differentiating (4.3.13) w.r.t 't' and then expanding and equating the coefficients of $\frac{t^r}{r!}$ on both sides, the cumulant recurrence relation may be obtained as

$$K_{r+1}=A\left[\lambda(\alpha+\beta 2^{r+1})+\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^j)K_{r-j+1}\right]+B\lambda\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^{j+1})\mu'_{r-j},$$

$$r\geq 1 \qquad (4.3.14)$$

where $A=\frac{1}{(1-\alpha-\beta)}$,

$$B=\frac{\omega}{(1-\alpha-\beta)(1-\omega)}$$

Here μ'_r denotes the r^{th} moment of the distribution.

The first and second cumulants of the distribution are respectively

$$K_1=\frac{\lambda(\alpha+2\beta)}{(1-\omega)(1-\alpha-\beta)},$$

and $$K_2 = \omega K_1^2 + \frac{(1+\beta)\mu}{(1-\alpha-\beta)} + \frac{2\lambda\beta}{(1-\omega)(1-\alpha-\beta)} .$$

e) Particular cases of GGD1

The distributions which may be derived from the GGD1, as its particular limiting cases are given below in a tabular form.

Sl.No.	Parameter values	Name of distribution	Pgf $g(t)$
1.	$\omega = 0$	Gegenbauer	$(1-\alpha-\beta)^{\lambda}(1-\alpha t-\beta t^2)^{\lambda}$ $\alpha,\beta,\lambda \ge 0$
2.	$\omega = 0, \beta = 0$	Negative binomial	$(1-\alpha)^{\lambda}(1-\alpha t)^{-\lambda}$, $\lambda > 0, 0 < \alpha < 1$
3	$\omega = 0, \lambda \to \infty$, $\alpha \to 0$ and $\beta \to 0$ such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$	Hermite	$\exp\{\alpha_1(t-1) + \alpha_2(t^2-1)\}$, $\alpha_1, \alpha_2 > 0$.

From the table B it is clear that, by putting $\beta = 0$, we may obtain negative binomial mixing ID distribution and for $\lambda \to \infty$, $\alpha \to 0$ and $\beta \to 0$ such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$, Hermite mixing ID distribution may be obtained.

1) Gegenbauer distribution

The Gegenbauer distribution (Plunkett and Jain, 1975) has the pgf

$$g(t) = (1-\alpha-\beta)^{\lambda}(1-\alpha t-\beta t^2)^{-\lambda}, \alpha+\beta < 1, \lambda > 0 .$$

This distribution may be obtained from (4.3.1) as a limiting distribution of GGD1 by putting $\omega = 0$. Putting $\omega = 0$ in (4.3.3) the probability recurrence relation of Gegenbauer distribution may be obtained as

$$P_{r+1} = (r+1)^{-1}\{\alpha(r+\lambda)P_r + \beta(2\lambda + r - 1)P_{r-1}\}, \quad r \ge 1$$

where $$P_0 = (1-\alpha-\beta)^{\lambda},$$

and $$P_1 = \alpha\lambda P_0 .$$

Again putting $\omega = 0$ in (4.3.8), the factorial moment recurrence relation of Gegenbauer distribution may be obtained as

$$\mu_{(r+1)} = \frac{\{(\alpha+2\beta)(r+\lambda)\mu_{(r)} + r\beta(r+2\lambda-1)\mu_{(r-1)}\}}{(1-\alpha-\beta)}, \ r \geq 1$$

where $\mu_{(1)} = \dfrac{\lambda(\alpha+2\beta)}{(1-\alpha-\beta)}$.

Similarly by putting $\omega = 0$ in (4.3.12), the moment recurrence relation of Gegenbauer distribution may be obtained as

$$\mu'_{r+1} = A\left[\lambda\sum_{j=1}^{r}\binom{r}{j}(\alpha+2^{r-j+1}\beta)\mu'_j + +\lambda(\alpha+\beta 2^{r+1}) + \sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^j\}\mu'_{r-j+1}\right], \ r \geq 1$$

where $\mu'_1 = \dfrac{\lambda(\alpha+2\beta)}{(1-\alpha-\beta)}$,

and $A = \dfrac{1}{1-\alpha-\beta}$.

Again putting $\omega = 0$ in (4.3.14), the cumulant recurrence relation of Gegenbauer distribution may be obtained as

$$k_{r+1} = (1-\alpha-\beta)^{-1}\left[\sum_{j=1}^{r}\binom{r}{j}k_{r-j+1}(\alpha+2^j\beta) + \lambda(\alpha+2^{r+1}\beta)\right], r \geq 1$$

where $k_1 = \dfrac{\lambda(\alpha+2\beta)}{(1-\alpha-\beta)}$

and $k_2 = \dfrac{(\alpha+2\beta)k_1 + \lambda(\alpha+4\beta)}{(1-\alpha-\beta)}$.

2) Hermite Distribution

Hermite distribution (Kemp and Kemp, 1965) which is a Poisson mixture of Bernoulli distribution has the pgf

$$G(t) = \exp\{\alpha_1(t-1) + \alpha_2(t^2-1)\}, \ \alpha_1, \alpha_2 > 0.$$

The Hermite distribution may be obtained from (4.3.1) as a limiting form of GGD1, taking limits as $\omega = 0$, $\lambda \to \infty$, $\alpha \to 0$ and $\beta \to 0$ such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$.

The probability recurrence relation of Hermite distribution is

$$P_{r+1} = \frac{\alpha_1 P_r + 2\alpha_2 P_{r-1}}{(r+1)}, \ r \geq 1.$$

where $P_0 = \exp\{-(\alpha_1 + \alpha_2)\}$

and $P_1 = \alpha_1 P_0$.

This is obtained from (4.3.3) taking limits as $\omega = 0$, $\lambda \to \infty$, $\alpha \to 0$ and $\beta \to 0$ such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$. Again the factorial moment recurrence relation of Hermite distribution obtained from (4.3.8) as its limiting case, is

$$\mu_{(r+1)} = \{(\alpha_1 + 2\alpha_2)\mu_{(r)} + 2r\alpha_2\mu_{(r-1)}\},\ r \geq 1$$

where $\mu_{(1)} = \alpha_1 + 2\alpha_2$.

Similarly taking limits as $\omega = 0$, $\lambda \to \infty$, $\alpha \to 0$ and $\beta \to 0$ such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$ in (4.3.12) the moment recurrence relation of Hermite distribution may be obtained as

$$\mu'_{r+1} = (\alpha_1 + 2\alpha_2) + \sum_{j=1}^{r}\binom{r}{j}(\alpha_1 + 2^{r-j+1}\alpha_2)\mu'_j,\ r \geq 1$$

where $\mu'_1 = \alpha_1 + 2\alpha_2$.

Again from (4.3.14) the cumulant recurrence relation of Hermite distribution may be obtained as

$$K_{r+1} = \alpha_1 + \alpha_2 2^{r+1},\ r \geq 1.$$

where $K_1 = \alpha_1 + 2\alpha_2$ and $K_2 = K_1 + 2\alpha_2$,

4.4 Estimation of parameters

To estimate the parameters of Generalized Charlier Distribution of Type 1(GCD1) and Generalized Gegenbauer Distribution of Type 1(GGD1), we shall use the simplest adhoc method of using sample mean, sample variance and the ratio of first three frequencies f_1/f_0 and f_2/f_0, as the other methods are found to be computationally complicated. In order to make the calculation easier, we shall first transform the probabilities (G_i's) of the Infinitely Divisible distributions in terms of the probabilities (P_i's) of the component distributions used. Then the probability recurrence relations of GCD1 and GGD1 may be written as

$$G_{r+1} = \frac{\omega}{1-\omega P_0}\left\{\sum_{j=1}^{r} G_{r-j+1}P_j + P_{r+1}G_0 + \frac{(1-\omega)P_{r+1}}{\omega}\right\},\ r \geq 1. \tag{4.4.1}$$

where $G_0 = \dfrac{(1-\omega)}{1-\omega P_0}$, (4.4.2)

$$G_1 = \frac{(1-\omega)P_1}{(1-\omega P_0)^2} \tag{4.4.3}$$

Putting $r = 1,2,3...$, in (4.4.1), we get

$$G_2 = \frac{(1-\omega)\{\omega P_1^2 + (1-\omega P_0)P_2\}}{(1-\omega P_0)^3}, \text{ so on.} \tag{4.4.4}$$

Therefore, the ratios of first three frequencies of the distribution corresponding to (4.1.1) are respectively

$$\theta_1 = \frac{f_1}{f_0} = \frac{P_1}{(1-\omega P_0)P_0}, \tag{4.4.5}$$

$$\theta_2 = \frac{f_2}{f_0} = \frac{\omega P_1^2 + (1-\omega P_0)P_2}{(1-\omega P_0)^2 P_0} = \omega\theta_1^2 P_0 + \frac{P_2}{P_1}\theta_1. \tag{4.4.6}$$

Adding equations (4.4.5) and (4.4.6) we get the a second degree equation in 'ω' as

$$A\omega^2 + B\omega + C = 0, \tag{4.4.7}$$

where $A = P_0^3\theta_1^2$,

$B = P_0 P_2 \theta_1 - P_0^2\{(\theta_1 + \theta_2) + \theta_1^2\}$

and $C = P_0(\theta_1 + \theta_2) - P_1 - \theta_1 P_2$.

Here $P_0 = \frac{f_0}{N}$, $P_1 = \frac{f_1}{N}$ and $P_2 = \frac{f_2}{N}$ are respectively first three probabilities of the component distributions used and $N = \sum f_i$ is the total frequency.

Solving equation (4.4.7), we get the value of ω as

$$\hat{\omega} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}, \quad 0<\omega<1 \tag{4.4.8}$$

For obtaining the remaining parameters of the distributions GCD1 and GGD1 we proceed as follows

a) GCD1 (α, β, λ)

The sample mean $\bar{x}$, sample variance s^2 and the ratio of first two frequencies f_0 and f_1 of GCD1 are respectively

$$\bar{x} = \frac{\alpha(1-\beta) + \lambda\beta}{(1-\omega)(1-\beta)}, \tag{4.4.9}$$

$$s^2 = \omega\bar{x}^2 + \bar{x} + \frac{\beta\bar{x}}{(1-\beta)} - \frac{\alpha\beta}{(1-\omega)(1-\beta)}, \tag{4.4.10}$$

$$\theta = \frac{f_1}{f_0} = (\alpha + \lambda\beta)\left\{1 + \frac{\omega G_0}{1-\omega}\right\}, \tag{4.4.11}$$

Eliminating α, λ from (4.4.9), (4.4.10) and (4.4.11) one by one and putting

$$T_1 = (s^2 - \bar{x} - \hat{\omega}\bar{x}^2)\ ,\ T_2 = \theta\left\{1 + \frac{\omega G_0}{1-\omega}\right\}^{-1} \text{ and } G_0 = \frac{f_0}{N}.$$

the estimated value of β may be obtained as

$$\hat{\beta} = 1 - \frac{(1-\hat{\omega})\bar{x} - T_2}{(1-\hat{\omega})T_1}, \tag{4.5.12}$$

Eliminating α from (4.4.9) and (4.4.10) and substituting the value of $\hat{\beta}$ we get the estimated value of λ as

$$\hat{\lambda} = \frac{1}{\hat{\beta}^2}(1-\hat{\omega})(1-\hat{\beta})^2 T_1, \tag{4.4.13}$$

Finally putting the values of $\hat{\beta}$ and $\hat{\lambda}$ in (4.4.9), we get

$$\hat{\alpha} = (1-\hat{\omega})\bar{x} - \frac{\hat{\lambda}\hat{\beta}}{(1-\hat{\beta})}. \tag{4.4.14}$$

b) GGD1 (α, β, λ)

The sample mean $\bar{x}$, the sample variance s^2 and the ratios of first two frequencies f_0 and f_1 of the distribution GCD1 are respectively

$$\bar{x} = \frac{(\alpha + 2\beta)\lambda}{(1-\omega)(1-\alpha-\beta)}, \tag{4.4.15}$$

$$s^2 = \omega K_1^2 + \frac{(1+\beta)\mu}{(1-\alpha-\beta)} + \frac{2\lambda\beta}{(1-\omega)(1-\alpha-\beta)}. \tag{4.4.16}$$

$$\theta = \frac{f_1}{f_2} = \alpha\lambda\left\{1 + \frac{\omega G_0}{1-\omega}\right\}, \tag{4.4.17}$$

Eliminating α, β one by one from (4.4.15), (4.4.16) and (4.4.17) we get a second degree equation in λ as

$$P\lambda^2 + Q\lambda + R = 0 \tag{4.4.18}$$

where $P = 4T_1 - 2(T_2 + T_3)$,

$Q = 2T_1^2 + (T_2 - 3T_1)T_3$

and $R = -T_1^2 T_3$.

writing $T_1 = \bar{x}(1-\hat{\omega})$, $T_2 = (1-\hat{\omega})(s^2 - \omega\bar{x}^2)$, $T_3 = \theta\left\{1 + \frac{\omega G_0}{1-\omega}\right\}^{-1}$

and $G_0 = \frac{f_0}{N}$.

The quadratic equation (4.4.18) may be solved either by using Newton Raphson method or the estimated value of λ may be obtained as

$$\hat{\lambda} = \frac{-Q \pm \sqrt{Q^2 - 4PR}}{2P}, \tag{4.4.19}$$

Putting the estimated value of λ in (4.4.15) and (4.4.17), after eliminating α we get the estimated value of β as

$$\hat{\beta} = \frac{\hat{\lambda}(T_1 - T_3) - T_1 T_3}{2\hat{\lambda}^2 + \hat{\lambda}T_1}, \tag{4.4.20}$$

Finally, putting the estimated values of β and λ in (4.4.15), we get

$$\hat{\alpha} = \frac{T_1 - \hat{\beta}(T_1 + 2\hat{\lambda})}{\hat{\lambda} + T_1}. \tag{4.4.21}$$

4.5 Applications of GCD1 and GGD1

It is believed that generalized Charlier distribution (GCD1) and generalized Gegenbauer distribution (GGD1) should give a reasonably good fit to some numerical data in different fields of biology, ecology and social information for which various modified forms of Poisson and negative binomial distributions were suggested by different statisticians. Therefore we have tried to fit GCD1 and GGD1 to some published data and compared them with other distribution on the basis of χ^2 criterion. The derived distributions GCD1 and GGD1 have been fitted by an adhoc method for rapid prediction, because of complexity of maximum likelihood method of estimation.

For the application of GCD1 and GGD1, we consider the example of the number of plants per quadrant of Lespedeza capitate in Table 4.1, well known data of Beall and

Rescia, 1953 to which Medhi and Borah (1986) fitted generalized Charlier distribution. For GGD1, using (4.4.8) we get $\hat{\omega} = 0.2475$, using (4.4.12), (4.4.13) and (4.4.14) we get $\hat{\beta} = 0.6408$, $\hat{\lambda} = 0.0437$ and $\hat{\alpha} = 0.0024$. For GGD1, using (4.4.19), (4.4.20) and (4.4.21), we get $\hat{\beta} = -0.0390$, $\hat{\lambda} = 0.0433$ and $\hat{\alpha} = 0.7023$ for the same value of $\hat{\omega} = 0.2475$. Observing the expected values obtained by GCD1 and GGD1, it is found that our fitted distributions agree with the observed data. The Poisson-Pascal distribution by Katti and Gurland (1961) is also compared with our fitted distribution and our distributions are found to be satisfactory.

In Tables 4.2 and 4.3, we have considered the data on the observed frequencies of distributions of purchases of two different kinds of brands of products (brand K and D) (Chatfield, 1969) where GPD model is fitted to these data by Consul (1989). For GCD1 in Table 4.2, we get $\hat{\omega}$ =0.3112, $\hat{\beta}$ = 0.6707, $\hat{\lambda}$ =0.0315 and $\hat{\alpha}$ =-0.0031 and in Table 4.3, we get $\hat{\omega}$ =0.2345, $\hat{\beta}$ = 0.7229, $\hat{\lambda}$ =0.0562 and $\hat{\alpha}$ =0.0159. The GGD1 is also fitted to the observed data of brand K of products in Table 4.2. Here for the same value of $\hat{\omega}$ =0.3112 we get $\hat{\beta}$ = 0.0628, $\hat{\lambda}$ =0.0311 and $\hat{\alpha}$ =0.5786. Considering the χ^2 values and a comparison of observed frequencies with the expected frequencies it is seen that the GCD1 and GGD1 both are as good as the GPD model for the data fitted by Consul (1989).

In the analysis of data observed on chemically induced chromosome aberrations in cultures of human leukocytes, Loeschcke and Kohler (1976) recommended the NBD, while Janardan and Schaeffer (1977) have used a modified Poisson distribution and also GPD by Consul (1989). In Table 4.4, the GCD1 is fitted to the distribution of number of Chromatid Aberrations, and it is also seen that our model fits the data best. Using (4.4.8) we get $\hat{\omega}$ =0.2032, using (4.4.12), (4.4.13) and (4.4.14) we get $\hat{\beta}$ = 0.6120, $\hat{\lambda}$ =0.1647 and $\hat{\alpha}$ =0.1765. From all these discussion it is clear that in all cases the data sets are very well described by the GCD1 model.

Table 4.1

Observed and fitted distributions of GCD1 and GGD1
(Observed frequencies of Lespedeza Capitate, data from Beall and Rescia, 1953)
(Medhi and Borah, 1986)

No of plants	Observed frequency $\bar{x}=0.1068$ $s^2=0.2944$	Fitted distributions			Poisson-Pascal distribution Katti and Gurland (1961)
		GCD1 $\omega\hat{}=0.2475$ $\beta\hat{}=0.6408$ $\lambda\hat{}=0.0437$ $\alpha\hat{}=0.0024$	GGD1 $\omega\hat{}=0.2475$ $\beta\hat{}=-0.0390$ $\lambda\hat{}=0.0433$ $\alpha\hat{}=0.7023$	TPCD (Medhi and Borah, 1986)	
0	7178	7179.18	7179.08	7179.13	7185.0
1	286	286.06	286.06	286.05	276.0
2	93	91.73	94.73	91.52	94.5
3	40	40.26	41.23	40.28	41.5
4	24	19.71	20.21	20.02	20.2
5	7	10.25	10.52	10.33	10.4
6	5	5.53	5.69	5.61	5.6
7	1	3.07	3.16	3.12	3.1
8	2	1.73	1.78	1.77	1.7
9	1	0.99	1.03	1.02	1.0
10	2	0.58	0.59	0.45	0.6
11	1	0.91	0.32	0.45	0.3
Total	7640	7640.00	7640.00	7640.00	7640.0
	χ^2	2.05	2.04	1.59	9.58
	$d.f.$	7	7	8	8
	$p-value$	> 0.95	> 0.95	> 0.99	> 0.29

GCD1: Generalized Charlier Distribution of Type 1
GGD1: Generalized Gegenbauer Distribution of Type 1
TPCD: Three parameter Charlier Distribution

Table 4.2

Observed and fitted distributions of purchases of a product of brand K, (Chatfield, 1969) by the Number of Consumers over a number of weeks (Consul, 1989)

No of consumers	Observed frequency Chatfield (brand K) $\bar{x} = 0.0886$ $s^2 = 0.2807$	Fitted distributions		Generalized Poisson Distribution (Consul, 1989) $\hat{\theta} = 0.0463$ $\hat{\lambda} = 0.4770$
		GCD1 $\hat{\omega} = 0.3112$ $\hat{\beta} = 0.6707$ $\hat{\lambda} = 0.0315$ $\hat{\alpha} = -0.0031$	GGD1 $\hat{\omega} = 0.3112$ $\hat{\beta} = 0.0628$ $\hat{\lambda} = 0.0311$ $\hat{\alpha} = 0.5786$	
0	1671	1671.49	1671.36	1670.78
1	43	43.02	43.05	48.04
2	19	16.74	16.25	14.91
3	9	7.66	7.37	6.73
4	2	3.91	3.77	3.57
5	3	2.12	2.05	2.08
6	1	1.18	1.16	1.28
7	0	0.69	0.67	0.82
8	0	0.41	0.40	0.54
9	2	2.78	3.87	1.25
Total	1750	1750.00	1750.00	1750.00
	χ^2	1.4006	2.1151	2.66
	$d.f.$	1	1	2
	$p-value$	> 0.23	> 0.14	> 0.26

GCD1: Generalized Charlier Distribution of Type 1

GGD1: Generalized Gegenbauer Distribution of Type 1

Table 4.3

Observed and fitted distributions of purchases of a product of brand D, (Chatfield, 1969) by the Number of Consumers over a number of weeks (Consul, 1989)

No of consumers	Observed frequency Chatfield (brand D) $\bar{x} = 0.21225$ $s^2 = 0.7223$	Fitted dist. GCD1 $\hat{\omega} = 0.2345$ $\hat{\beta} = 0.7229$ $\hat{\lambda} = 0.0562$ $\hat{\alpha} = 0.0159$	Generalized Poisson Distribution (Consul, 1989) $\hat{\theta} = 0.1133$ $\hat{\lambda} = 0.4663$
0	875	874.84	875.05
1	63	62.98	62.18
2	19	19.11	20.40
3	10	9.42	9.33
4	4	5.23	4.95
5	4	3.09	2.87
6	1	1.89	1.75
7	2	1.19	1.12
8	0	0.76	0.73
9	1	0.49	0.49
10	0	0.33	0.33
11	0	0.22	0.23
12	1	0.45	0.57
Total	980	980.00	980.00
	χ^2	0.36	0.44
	$d.f.$	2	3
	$p-value$	> 0.83	> 0.93

GCD1: Generalized Charlier Distribution of Type 1

Table 4.4

Observed and fitted distribution of GCD1
(Number of Chromatid Aberrations (0.02 g Chinon I, 24 Hours)
data by Loeschcke and Kohler, 1976 (Consul, 1989)

No of aberrations	Observed Frequency $\bar{x} = 0.5475$ $s^2 = 1.1227$	Fitted dist. GCD1 $\hat{\omega} = 0.2032$ $\hat{\beta} = 0.6120$ $\hat{\lambda} = 0.1647$ $\hat{\alpha} = 0.1765$	Generalized Poisson Distribution (Consul, 1989) $\hat{\theta} = 0.3928$ $\hat{\lambda} = 0.2826$
0	268	267.59	270.1
1	87	86.85	80.0
2	26	26.22	28.9
3	9	10.41	11.6
4	4	4.91	5.0
5	2	2.50	2.3
6	1	1.32	1.0
7	3	0.20	0.1
Total	400	400.00	400.00
	χ^2	1.34	2.37
	$d.f.$	1	3
	$p-value$	> 0.24	> 0.49

GCD1: Generalized Charlier Distribution of Type 1

4.6 Properties of particular cases of GCD1 and GGD1

It is seen that some generalized infinitely divisible distributions of Poisson, negative binomial (Begum and Borah, 1999, 2003) and Hermite distributions may be obtained as particular limiting cases of GCD1 and GGD1. Attempt has been also made to study the properties of these distributions together with estimation of their parameters. Some well known data sets in different fields of biology, ecology, social information, home injuries, accident data etc., are considered for fitting of the derived distributions.

a) Poisson mixing infinitely divisible distribution of Type 1 (GPD1)

The pgf of Poisson mixing infinitely divisible distribution (GPD1) derived from (4.2.1) by putting $\beta = 0$ may be written as

$$G(t) = \frac{(1-\omega)e^{\alpha(t-1)}}{1-\omega e^{\alpha(t-1)}}, \qquad 0 < \omega < 1 \tag{4.6.1}$$

Poisson mixing infinitely divisible distribution has the probability recurrence relation

$$G_{r+1} = \frac{1}{(r+1)}\left[\alpha G_r + \frac{\alpha\omega}{1-\omega}\left\{G_0 G_r + \sum_{j=1}^{r} G_j G_{r-j}\right\}\right] \quad , \; for\ r \geq 1 \tag{4.6.2}$$

where $G_0 = \dfrac{(1-\omega)e^{-\alpha}}{1-\omega e^{-\alpha}}$,

$$G_1 = \alpha\left\{1 + \frac{\omega G_0}{1-\omega}\right\}G_0 .$$

Its factorial moment recurrence relation is

$$\mu_{(r+1)} = \frac{\alpha^{r+1}}{(1-\omega)} + \frac{\omega(r+1)}{(1-\omega)}\sum_{j=1}^{r}\binom{r}{r-j+1}\frac{\alpha^{r-j+1}}{j}\mu_{(j)}, \; for\ r \geq 1 \tag{4.6.3}$$

where $\mu_{(1)} = \dfrac{\alpha}{1-\omega}$, is the first factorial moment of the distribution.

Therefore mean and variance of the distribution are respectively

$$\mu = \frac{\alpha}{1-\omega}$$

and $$\sigma^2 = \omega\mu^2 + \mu .$$

Estimation of parameters

For estimating the parameters of Poisson mixing infinitely divisible distribution, as method of maximum likelihood will be very cumbersome in this case the method of moment has been used. The first two sample moments of the distribution are respectively

$$\bar{x} = \frac{\alpha}{1-\omega} \tag{4.6.4}$$

$$s^2 = \omega\bar{x}^2 + \bar{x} . \tag{4.6.5}$$

Solving the equations (4.6.4) and (4.6.5) we may have the estimated values of ω and α respectively as

$$\hat{\omega} = \frac{s^2 - \bar{x}}{\bar{x}^2} \text{ and } \hat{\alpha} = \bar{x} + 1 - \frac{s^2}{\bar{x}}.$$

Graphical Representation of GPD1

To study the behaviour of the GPD1 model for varying values of ω and λ, the probabilities for possible values of x are computed. It is clear from the graphs in figures 4.1, 4.2 and 4.3 that when ω remains fixed at $\omega = 0.25, \omega = 0.5$ and $\omega = 0.8$ respectively, and λ varies, the probability curves become unimodel and positively skewed, but its flatness increases with increasing value of λ.

In figures 4.1, λ takes values 0.5, 1.5, 3.5, 5.5, 7 and 9.5. For figures 4.2, and 4.3, λ takes values 0.5, 1.5, 2.5, 3.5, 5.5, 7 and 9.5. Again, in figure 4.4 for fixed $\lambda = 5.5$, when ω takes values 0.25, 0.5 and 0.8 then for smaller value of ω, the curve is more bell-shaped and it becomes more flattened when ω increases.

Figure 4.1

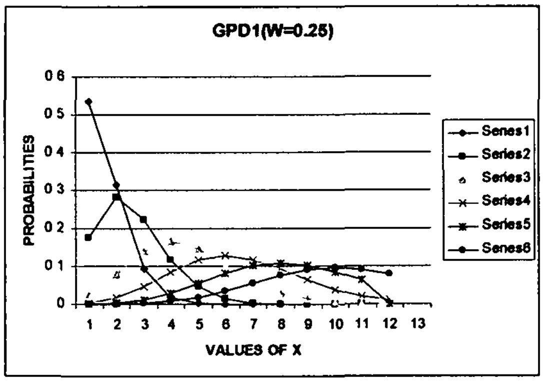


Figure 4.2

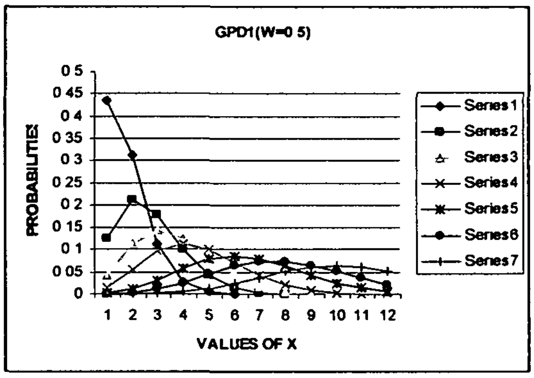


Figure 4.3

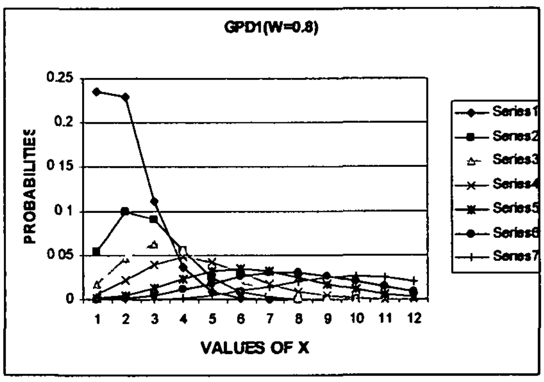
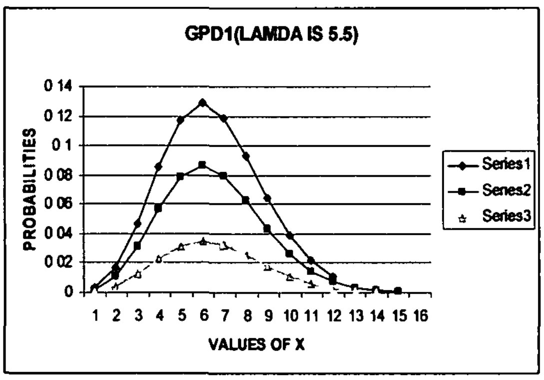


Figure 4.4



b) Negative binomial mixing infinitely divisible distribution of Type 1 (GNBD1)

The pgf of negative binomial mixing infinitely divisible distribution (GNBD1) derived from (4.2.1) by putting $\alpha = 0$ may be written as

$$G(t) = \frac{(1-\omega)(1-\beta)^{\lambda}(1-\beta t)^{-\lambda}}{1-\omega(1-\beta)^{\lambda}(1-\beta t)^{-\lambda}}, \quad 0<\omega<1 \tag{4.6.6}$$

It has the probability recurrence relation is

$$G_{r+1} = \frac{1}{r+1}\left[\beta(r+\lambda)G_r + \frac{\lambda\beta\omega}{1-\omega}\left\{G_0G_r + \sum_{j=1}^{r} G_j G_{r-j}\right\}\right], \ for \ r \geq 1 \tag{4.6.7}$$

where $\quad G_0 = \dfrac{(1-\omega)(1-\beta)^{\lambda}}{1-\omega(1-\beta)^{\lambda}}.$

and $\quad G_1 = \dfrac{\beta\lambda G_0}{1-\omega(1-\beta)^{\lambda}} = \lambda\beta\left\{1+\dfrac{\omega G_0}{1-\omega}\right\}G_0.$

Its factorial moment recurrence relation is

$$\mu_{(r+1)} = A\left\{b^{r+1}\binom{\lambda+r}{r+1} + \omega\sum_{j=1}^{r}\binom{\lambda+r-j}{r-j+1}b^{r-j+1}\frac{\mu_j}{j!}\right\}, r \geq 1 \tag{4.6.8}$$

where $A = \dfrac{(r+1)!}{(1-\omega)}, \ b = \dfrac{\beta}{1-\beta}.$

The first factorial moment of the distribution is $\mu_{(1)} = \dfrac{nb}{1-\omega}$.

Hence mean and variance of the distribution are respectively

$$\mu = \frac{\lambda\beta}{(1-\omega)(1-\beta)}, \tag{4.6.9}$$

and $$\sigma^2 = \omega\mu^2 + \frac{\mu}{(1-\beta)}. \tag{4.6.10}$$

Estimation of parameters

The parameters λ, α and ω of the negative binomial mixing infinitely divisible distribution may be estimated in the following three methods

i) Method of using first three factorial moments

The first three sample factorial moments of the distribution are respectively

$$m_{(1)} = \frac{\lambda b}{1-\omega},$$

$$m_{(2)} = m_{(1)}\{b(\lambda+1) + 2\omega m_{(1)}\},$$

and $$m_{(3)} = (\lambda+2)bm_{(2)} + 4\omega m_{(1)}m_{(2)} - 2\omega m_{(1)}\{\omega m_{(1)} + b\}m_{(1)}$$

writing $R = m_{(2)} - 2m^2_{(1)}$, $Q = m_{(3)}m_{(1)} - 6m_{(1)}^2(R + m_{(1)}^2)$ and $\beta = \dfrac{b}{1+b}$

we have $$\hat{\lambda} = \frac{2R^2 - Q}{R^2 - Q},$$

$$\hat{\beta} = \frac{Q - R^2}{Q + \{\mu_{(1)} - R\}R}$$

and $$\hat{\omega} = 1 - \frac{Q - 2R^2}{R\mu^2_{(1)}}.$$

ii) Method of using first two sample moments and ratio of first two frequencies

First two sample moments and the ratio of first two frequencies are respectively

$$\bar{x} = \frac{\lambda\beta}{(1-\omega)(1-\beta)},$$

$$s^2 = \omega\mu^2 + \frac{\mu}{(1-\beta)},$$

and $$\theta = \frac{f_1}{f_0} = \lambda\beta\left\{1 + \frac{\omega G_0}{1-\omega}\right\}G_0.$$

Solving the above three equations of sample mean $\bar{x}$, sample variance s^2 and the ratio of first two frequencies $\theta = \dfrac{f_1}{f_0}$, we obtained the estimated values of β, ω and λ respectively as

$$\hat{\beta} = \frac{\bar{x}^2 + \bar{x} - s^2}{(\lambda-1)\bar{x}},$$

$$\hat{\omega} = \frac{\lambda(s^2 - \bar{x}) - \bar{x}^2}{(\lambda-1)\bar{x}^2},$$

$$\hat{\lambda} = \frac{\bar{x}^2 P_0 - \theta(s^2 - \bar{x}^2)}{(\bar{x}^2 + \bar{x} - s^2) + (s^2 - \bar{x})P_0 - \bar{x}\theta},$$

where $$P_0 = \frac{f_0}{N}$$

iii) Method of using first sample moment and ratio of first three frequencies

The first sample moment $\bar{x}$ and the ratio of first three frequencies $\theta_1 = \frac{f_1}{f_0}$ and $\theta_2 = \frac{f_2}{f_0}$ of the distributions are respectively

$$\bar{x} = \frac{\lambda\beta}{(1-\omega)(1-\beta)},$$

$$\theta_1 = \frac{f_1}{f_0} = \lambda\beta\left\{1 + \frac{\omega G_0}{1-\omega}\right\}G_0,$$

and $\quad \theta_2 = \frac{f_2}{f_0} = \frac{1}{2}\beta(\lambda+1)\theta_1 + \beta\lambda\theta_1\frac{\omega G_0}{1-\omega}.$

writing $\quad R = 2(\theta_2 - \theta_1^2)$ and $Q = (\bar{x}P_0 - \theta_1)\theta_1$.

we have $\quad \hat{\lambda} = \frac{Q - \bar{x}RP_0}{(1-P_0)R + Q},$

$$\hat{\beta} = \frac{\bar{x}P_0 - \theta_1}{\lambda(1-P_0) + \bar{x}P_0},$$

$$\hat{\omega} = \frac{\bar{x} - \beta(\bar{x}+\lambda)}{\bar{x}(1-\beta)}.$$

c) Hermite mixing infinitely divisible distribution of Type 1 (GHMD1)

The pgf of Hermite mixing infinitely divisible distribution derived from the model (4.3.1) by considering the limits as $\lambda \to \infty$, $\alpha \to 0$ and $\beta \to 0$ such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$ $\alpha = 0$ may be written as

$$G(t) = \frac{(1-\omega)\exp\{\alpha_1(t-1) + \alpha_2(t^2-1)\}}{1-\omega\exp\{\alpha_1(t-1) + \alpha_2(t^2-1)\}}, \beta, \lambda \geq 0. \quad (4.6.11)$$

The distribution has the probability recurrence relation

$$G_{r+1} = \frac{\alpha_1 G_r + 2\alpha_2 G_{r-1}}{(r+1)} + \frac{\omega}{(r+1)(1-\omega)}\left[\alpha_1 G_0 G_r + \sum_{j=1}^{r}\{\alpha_1 G_j + 2\alpha_2 G_{j-1}\}G_{r-j}\right],$$

$$\text{for } r \geq 1 \quad (4.6.12)$$

where $\quad G_0 = \frac{(1-\omega)\exp\{-(\alpha_1+\alpha_2)\}}{1-\omega\exp\{-(\alpha_1+\alpha_2)\}}$ and $G_1 = \alpha_1\left\{1 + \frac{\omega G_0}{1-\omega}\right\}G_0$.

Its cumulant recurrence relation is

$$K_{r+1} = \{\alpha_1 + \alpha_2 2^{r+1}\} + \frac{\omega}{(1-\omega)}\left[\{\alpha_1 + \alpha_2 2^{r+1}\} + \sum_{j=1}^{r}\binom{r}{j}\{\alpha_1 + \alpha_2 2^{r-j+1}\}\mu_j'\right],$$

$$for\ r \geq 1 \qquad (4.6.13)$$

where first two cumulants of the distribution are respectively

$$K_1 = \frac{\alpha_1 + 2\alpha_2}{1-\omega}$$

and $\quad K_2 = \omega K_1^2 + K_1 + \dfrac{2\alpha_2}{1-\omega}$.

Its factorial moment recurrence relation is

$$\mu_{(r+1)} = \{(\alpha_1 + 2\alpha_2)\mu_{(r)} + 2r\alpha_2\mu_{(r-1)}\} + \frac{\omega}{(1-\omega)}\left[(\alpha_1 + 2\alpha_2)\sum_{j=0}^{r}\binom{r}{j}\mu_{(j)}\mu_{)r-j)} + \right.$$

$$\left. 2\alpha_2\sum_{j=1}^{r} j\binom{r}{j}\mu_{(j-1)}\mu_{)r-j)}\right], \quad for\ r \geq 1 \qquad (4.6.14)$$

where $\mu_{(1)} = \dfrac{\alpha_1 + 2\alpha_2}{1-\omega}$, is the first factorial moment of the distribution.

Therefore mean and variance of the distribution are respectively

$$\mu = \frac{\alpha_1 + 2\alpha_2}{1-\omega} \qquad (4.6.15)$$

and $$\sigma^2 = \omega\mu^2 + \mu + \frac{2\alpha_2}{1-\omega}. \qquad (4.6.16)$$

Estimation of parameters

For estimating the parameters of Hermite mixing infinitely divisible distribution the following methods have been used.

i) first two sample moments and ratio of first two frequencies

The sample mean, sample variance and ratio of first two frequencies are respectively

$$\bar{x} = \frac{\alpha_1 + 2\alpha_2}{1-\omega}, \qquad (4.6.17)$$

$$s^2 = \omega\bar{x}^2 + \bar{x} + \frac{2\alpha_2}{1-\omega}, \qquad (4.6.18)$$

$$\theta = \frac{f_1}{f_0} = \alpha_1\left\{\frac{1 + \omega(G_0 - 1)}{1-\omega}\right\}. \qquad (4.6.19)$$

Eliminating, α_2 from (4.6.17) and (4.6.18) we have

$$\alpha_1 = \frac{\theta(1-\omega)}{1+\omega(G_0-1)} \tag{4.6.20}$$

Again eliminating, α_1 from (4.6.19) and (4.6.20) we may have a second degree equation in ω as

$$A\omega^2 + B\omega + C = 0 \tag{4.6.21}$$

where $\quad A = \bar{x}^2(G_0-1)$, $B = (G_0-1)(2\bar{x}-s^2)+\bar{x}^2$

and $\quad C = (2\bar{x}-s^2)-\theta$.

Solving equation (4.6.21), we get

$$\hat{\omega} = \frac{-B \pm \sqrt{B^2-4AC}}{2A}, \tag{4.6.22}$$

Putting the value of $\hat{\omega}$ in (4.6.20), we get

$$\hat{\alpha}_1 = \frac{\theta(1-\hat{\omega})}{1+\hat{\omega}(G_0-1)} \tag{4.6.23}$$

Substituting the values of $\hat{\omega}$ and $\hat{\alpha}_1$ in (4.6.18), we get

$$\hat{\alpha}_2 = \frac{1}{2}(s^2-\bar{x}-\hat{\omega}\bar{x}^2)(1-\hat{\omega}). \tag{4.6.24}$$

ii) first sample moment and ratio of first three frequencies

The sample mean and the ratios of first three frequencies are respectively

$$\bar{x} = \frac{\alpha_1+2\alpha_2}{1-\omega}, \tag{4.6.25}$$

$$\theta_1 = \frac{f_1}{f_0} = \alpha_1\left\{\frac{1+\omega(G_0-1)}{1-\omega}\right\}, \tag{4.6.26}$$

$$\theta_2 = \theta_1^2 + \frac{\alpha_2}{\alpha_1}\theta_1 - \frac{1}{2}\theta_1\alpha_1. \tag{4.6.27}$$

Eliminating, α_2 from (4.6.25) and (4.6.27) we have

$$\omega = \frac{\bar{x}-(2T+1)-\alpha_1^2}{\bar{x}}. \tag{4.6.28}$$

where $\quad T = \dfrac{\theta_2-\theta_1^2}{\theta_1}$,

From (4.6.26) and (4.6.28), eliminating ω we may have a second degree equation in α_1 as

$$A\alpha_1^2 + B\alpha_1 + C = 0 \tag{4.6.29}$$

where $\qquad A = (G_0 - 1), G_0 = \frac{f_0}{N}, \; B = (G_0 - 1)(2T + 1) + \theta_1$

and $\qquad C = \theta_1(2T + 1) - \bar{x}G_0$.

Solving (4.6.29) as an ordinary second degree equation we get

$$\hat{\alpha}_1 = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}, \tag{4.6.30}$$

Putting the value of $\hat{\alpha}_1$ in (4.6.26), we get

$$\hat{\omega} = \frac{\theta - \hat{\alpha}_1}{\theta_1 + (G_0 - 1)\hat{\alpha}_1} \tag{4.6.31}$$

Again putting the value of $\hat{\alpha}_1$ and $\hat{\omega}$ in (4.6.25), we get

$$\hat{\alpha}_2 = \frac{1}{2}\{\bar{x}(1 - \hat{\omega}) - \hat{\alpha}_1\}. \tag{4.6.32}$$

4.7 Applications

The estimation of parameters in fitting probability distributions plays a very important role. Of all the procedures of estimating the parameters, the method of moments is perhaps the oldest and the simplest. In many cases it leads to tractable operations. The method of maximum likelihood is considered to be more accurate for fitting a probability distribution on given data, but it involves much more computational work than is required by the method of moments. It is mainly from this reason moment estimators are used for estimating the parameters of generalized Poisson (GPD1) and generalized negative binomial (GNBD1) distributions. In case of generalized Hermite distribution (GHMD1), the first two sample moments and the ratio of first two frequencies are used to estimate the parameters.

The negative-binomial and the Poisson distributions are commonly used in different ecological and biological problems. Therefore, we have considered some reported data sets in different fields of biology, ecology, social information and accidents for fitting generalized Poisson, generalized negative binomial and generalized Hermite distributions.

In Table 4.5, we have considered the Student's historic data on Haemocytometer counts of yeast cells for which Hermite distribution was fitted by Kemp and Kemp (1966) and Gegenbauer distribution was fitted by Borah (1984). Here it is seen that the expected frequencies computed by our derived distributions match satisfactorily than the Hermite distribution (Kemp and Kemp, 1966). In Table 4.6, we have considered the data of Beall and Rescia (1940), for which generalized Poisson distribution (GPD) was fitted by Jain (1975) and Neyman type A and Neyman Type-B were fitted by McGuire et al (1957). In this case also our distributions give good fit in comparison.

In Table 4.7 and 4.8, we consider two sets of data of Adelstein (1952) on number of accidents (home injuries) of 122 experienced men in six years (1937-1942) and eleven years (1937-1947) periods respectively. Again in Table 4.9, we consider the data of first year shunting accidents and a five year record of experienced 170 men for the age group 21-25 years. It is clear that in all the above three cases there is some improvement, however small it may be, in fitting these distributions over the other distributions considered earlier.

TABLE 4.5

Observed and fitted Generalized Poisson Distribution (GPD1), Generalized negative binomial Distribution (GNBD1) and Generalized Hermite Distribution (GHMD1) (Haemocytometer Counts of Yeast Cells) (Borah, 1984)

No. of Yeast cells Per square	Observed Frequency $\bar{x}=0.6825$ $s^2=0.8117$	Fitted GPD1 λ^=0.4932 ω^=0.2774 (MM)	Fitted GNBD1 $\hat{\lambda}$=2.7463, $\hat{\beta}$=0.2202 ω^=0.8799 (MFM)	Fitted GHMD1 $\hat{\omega}=0.3882$ $\hat{\alpha}_1=0.4491$ $\hat{\alpha}_2=-0.0158$ (MVR)	Gegenbauer Distribution (Borah,1984)	Hermite Distribution (Kemp and Kemp, 1966)
0	213	212.51	214.77	212.04	214.15	213.12
1	128	126.18	121.52	127.26	123.00	122.91
2	37	43.81	45.39	43.32	44.88	46.71
3	18	12.63	13.74	12.45	13.36	13.31
4	3	3.50	3.58	3.64	3.55	3.16
5	1	0.97	0.81	1.02	0.86	0.64
6	0	0.27	0.05	0.27	0.20	0.15
Total	400	400.00	400.00	400.00	400.00	400.00
	χ^2	2.3201	2.7136	3.5802	2.8342	3.8825
	$d.f.$	2	1	1	1	2
	$p-value$	>0.31	>0.09	>0.05	>0.09	>0.14

MM= Method of Moment

MFR= Method of factorial moment

MVR= Method of mean, variance and ratio of first two frequencies.

TABLE 4.6

Observed and fitted Generalized Poisson Distribution (GPD1), Generalized negative binomial Distribution (GNBD1) and Generalized Hermite Distribution (GHMD1) (Data from the paper by Beall and Rescia, 1940)

No. of Insects	Observed Frequency $\bar{x}$=0.75 s^2=1.2946	Fitted GPD1 λ^=0.1403 ω^=0.8089 (MM)	Fitted GNBD1 n^=1.00068 a^=0.9724, ω^=0.0213 (MFM)	Fitted GHMD1 $\hat{\omega}$ = 0.4137 $\hat{\alpha}_1$ = 0.2568 $\hat{\alpha}_2$ = 0.0914 (MVR)	Generalized Poisson Distribution (Jain, 1975)
0	33	31.00	31.98	32.74	32.46
1	12	14.86	13.73	11.88	13.47
2	6	6.05	5.89	7.01	5.60
3	3	2.69	2.52	2.57	2.42
4	1	0.99	1.08	0.97	1.08
5	1	0.41	0.46	0.83	0.97
Total	56	56.00	56.00	56.00	56.00
	χ^2	0.8824	0.6389	0.24	0.25
	df	1			
	$p-value$	> 0.34			

TABLE 4.7

Comparison of observed frequencies for Home injuries of 122 experienced men during (1937-1942) with the fitted Generalized Poisson Distribution (GPD1), Generalized negative binomial Distribution (GNBD1) and Generalized Hermite Distribution (GHMD1) (Consul, 1989)

No.of injuries	Observed $\bar{x}$=0.5409 s^2=0.60897	Fitted GPD1 λ^=0.4153 ω^=0.2323 (MM)	Fitted GNBD1 $\hat{\lambda}$=5.9314, $\hat{\beta}$=0.0777 ω^=0.0767 (MFM)	Fitted GHMD1 $\hat{\omega}$ = 0.5747 $\hat{\alpha}_1$ = 0.2727 $\hat{\alpha}_2$ = −0.0213 (MVR)	GPD Consul,1989
0	73	73.03	73.21	72.96	72.23
1	36	35.82	35.41	35.98	35.32
2	10	10.13	10.34	10.02	10.41
3	2	2.35	2.41	2.35	3.04
4	1	0.67	0.63	0.69	0.00
Total	122	122.00	122.00	122.00	122.00

MM= Method of Moment,

MFR= Method of factorial moment,

MVR= Method of mean, variance and ratio of first two frequencies.

TABLE 4.8

Comparison of observed frequencies for Home injuries of 122 experienced men during 11 years (1937-1947) with the fitted Generalized Poisson Distribution (GPD1) and Generalized Hermite DistributionGHMD1 (Consul, 1989)

No.of injuries	Observed Frequency $\bar{x}=0.9836$ $s^2=1.5571$	Fitted Distributions GPD1 $\hat{\lambda}=0.4006$ $\hat{\omega}=0.5928$ (MM)	GHMD1 $\hat{\omega}=0.2913$ $\hat{\alpha}_1=0.4904$ $\hat{\alpha}_2=0.1034$ (MVR)	GPD Consul,1989
0	58	55.21	56.90	57.22
1	34	36.68	33.25	34.41
2	14	17.03	18.29	16.64
3	8	7.41	7.92	7.59
4	6	3.21	3.17	6.14
5	2	1.39	2.47	----
Total	122	122.00	122.00	122.00
	χ^2	3.4359	2.0327	1.09
	$d.f.$	2	1	2
	$p-value$	>0.17	>0.15	>0.57

TABLE 4.9

Comparison of Observed Frequencies for First-Year Shunting Accidents and for a Five Year Record of Experienced men for the age group 21-25 years with fitted Generalized Hermite Distribution (GHMD1) (Consul,1989)

No.of Accidents	Observed Frequency $\bar{x}=0.7529$ $s^2=0.6801$	Fitted GHMD1 $\hat{\omega}=0.8986$ $\hat{\alpha}_1=0.1354$ $\hat{\alpha}_2=-0.0295$ (MVR)	GPD Consul,1989
0	80	80.89	76.40
1	56	57.13	65.03
2	30	24.04	23.37
3	4	7.32	5.20
≥ 4	0	0.62	----
Total	170	170.00	170.00

MVR= Method based on mean, variance and ratio of first two frequencies.

In the above Tables 4.6, 4.7 and 4.9, in fitting GNBD1 and GHMD1 where the number of observations is 6 or less and we have three parameters to be estimated, $\chi 2$ and corresponding $p-value$ are not provided as degrees of freedom (df) is very negligible.

In Table 4.10, we have considered the data collected by P. Garman (see Bliss et al 1953) on the Count of the number of European Red Mites on Apple Leaves. Comparing the χ^2 values obtained from GPD1 with that of the other distributions compared, it has been observed that our distribution is found to be satisfactory.

TABLE 4.10: Observed and fitted Generalized Poisson Distribution (GPD1) (Count of the number of European Red Mites on Apple Leaves) (Jain and Consul, 1971)

No. of mites per leaf	Leaves (observed) $\bar{x}=1.1467$ $s^2=2.2585$	Fitted Distributions		Gen.Neg.Bin. (Jain and Consul, 1971)
		GPD1 $\hat{\lambda}=0.1638$ $\hat{\omega}=0.8571$(MM)	Poisson lindley (Sankaran,1970) $\hat{\theta}=1.258$	
0	70	66.79	67.19	71.48
1	38	40.17	38.89	33.98
2	17	20.87	21.26	19.80
3	10	10.75	11.21	11.59
4	9	5.76	5.76	6.57
5	3	2.95	2.90	3.55
6	2	1.57	1.44	1.80
7	1	0.75	0.71	0.84
8	0	0.39	0.34	0.39
Total		150	150.00	150.00
	χ^2	2.8640	3.0136	2.0700
	$d.f.$	3	4	3
	$p-value$	>0.41	>0.55	>0.55

MM= Method of Moment,

In Table 4.11, we considered the observed data of Kendall (1961), on the number of strikes in 4-week periods in two leading industries in U.K. during 1948-1959 and concluded that the aggregate data for the two industries of Vehicle manufacturing and Ship building agree with Poisson law. The distributions corresponding to GPD1 and GHMD1 have been fitted to the observed data for the two industries. The results are given in Table 4.11, along with the expected frequencies of GPD (Consul, 1989). Based on observed and expected frequencies it is clear that the pattern of strikes in vehicle manufacturing and ship building data describe the models very closely.

TABLE 4.11

Comparison of Observed Frequencies of the Number of Outbreaks of Strike in U.K. during 1948-1959 with the Expected frequencies of GPD1 and GHMD1 (Consul, 1989)

Vehicle manufacturing Industries					Ship-building Industries			
		Fitted Distributions				Fitted Distributions		
No. of outbreaks	Observed Frequency $\bar{x}=0.4103$ $s^2=0.5496$	GPD1 $\hat{\lambda}=0.0705$ $\hat{\omega}=0.8281$ (MM)	GHMD1 $\hat{\omega}=0.4361$ $\hat{\alpha}_1=0.1941$ $\hat{\alpha}_2=0.0186$ (MVR)	GPD Consul,1989 $\hat{\theta}=0.351$ $\hat{\lambda}=-0.144$	Observed Frequency $\bar{x}=0.3269$ $s^2=0.4124$	GPD1 $\hat{\lambda}=0.0656$ $\hat{\omega}=0.7993$ (MM)	GHMD1 $\hat{\omega}=0.4361$ $\hat{\alpha}_1=0.1941$ $\hat{\alpha}_2=0.0186$ (MVR)	GPD Consul,1989 $\hat{\theta}=0.29$ $\hat{\lambda}=-0.113$
0	110	109.47	109.84	109.82	117	116.60	117.14	116.74
1	33	33.82	32.93	33.36	29	30.43	29.05	30.22
2	9	9.25	9.83	9.24	9	6.94	7.93	6.97
3	3	2.52	2.53	3.58	0	1.57	1.52	0.88
4	1	0.69	0.87	0.00	1	0.36	0.36	0.00
Total	156	156.00	156.00	156.00	156	156.00	156.00	156.00
	χ^2	0.2236		0.06	χ^2	1.1281		1.19
	$d.f.$	1		1	$d.f.$	1		1
	$p-value$	>0.63		>0.8	$p-value$	>0.28		>0.27

GPD1: Generalized Poisson Distribution of Type 1
GHMD1: Generalized Hermite Distribution of Type 1
GPD: Generalized Poisson Distribution

Chapter 5

Generalized Infinitely Divisible Distributions of Type 2

5.1 Introduction

In this chapter 5, we have made an attempt to derive generalized distributions of three parameter Charlier and Gegenbauer by considering them as the component distribution $g(t)$, in the geometrically infinitely divisible model

$$G(t) = \frac{\omega}{(\omega+1) - g(t)} \quad , \quad \omega > 0 \tag{5.1.1}$$

studied by Klebanov, Maniya and Melamed (1984). The generalized distributions of Charlier and Gegenbauer are denoted by the symbols GCD2 and GGD2 respectively. Further, certain important properties of these newly derived distributions are investigated. The parameters are estimated and distributions are fitted to some published data in biology, ecology and social information etc.

It has been observed that some generalized infinitely divisible distributions of Poisson and negative binomial may be obtained as particular limiting cases of GCD2 and GGD2. The important properties of these limiting distributions have been also investigated and the distributions are fitted to some well known published data in biology and ecology.

5.2 Generalized Charlier Distribution of Type 2 (GCD2)

a) Probability recurrence relation

The pgf of generalized Charlier distribution of type 2 (GCD2), derived from the model (5.1.1) may be written as

$$G(t) = \frac{\omega}{(\omega+1) - e^{-\alpha}(1-\beta)^{\lambda} e^{\alpha t}(1-\beta t)^{-\lambda}}, \quad \omega > 0 \tag{5.2.1}$$

Differentiating both sides of (5.2.1) with respect to (w.r.t.) 't' we get

$$(1-\beta t)G'(t) = \frac{\omega+1}{\omega}\left[(\alpha+\lambda\beta)G(t)G(t) - \alpha\beta t G(t)G(t)\right] - \left\{(\alpha+\lambda\beta)G(t) - \alpha\beta t G(t)\right\}$$

Now equating the coefficients of t^r on both sides, the recurrence relation for probabilities may be obtained as

$$G_{r+1} = \frac{(\omega+1)}{\omega(r+1)}\left\{(\alpha+\lambda\beta)G_0 G_r + \sum_{j=1}^{r}\{(\alpha+\lambda\beta)G_j - \alpha\beta G_{j-1}\}G_{r-j}\right\}$$
$$-\left[\frac{(\alpha+\lambda\beta-r\beta)G_r - \alpha\beta G_{r-1}}{(r+1)}\right] \quad r \geq 1. \tag{5.2.2}$$

where $$G_0 = \frac{\omega}{(\omega+1) - e^{-\alpha}(1-\beta)^{\lambda}}, \tag{5.2.3}$$

and $$G_1 = (\alpha+\lambda\beta)\left\{\frac{(\omega+1)}{\omega}G_0 - 1\right\}G_0. \tag{5.2.4}$$

b) Factorial moment recurrence relation

Corresponding to the pgf (5.2.1), the fmgf of GCD2 may be written as

$$m(t) = \frac{\omega}{\omega+1-e^{\alpha t}(1-ae^{t})^{-\lambda}}, \quad \omega > 0 \tag{5.2.5}$$

where $$a = \frac{\beta}{1-\beta}.$$

Differentiating the equations (5.2.5) w.r.t.'t', we get

$$(1-at)m'(t) = \frac{\omega+1}{\omega}\{(\alpha+\lambda a) - a\alpha t\}m(t)m(t) - \{(\alpha+\lambda a) - a\alpha t\}m(t)$$

Equating the coefficients of $\frac{t^r}{r!}$ on both sides the recurrence relation for factorial moments may be obtained as

$$\mu_{(r+1)} = \frac{\omega+1}{\omega}\left[(\alpha+\lambda a)\sum_{j=0}^{r}\binom{r}{j}\mu_{(j)}\mu_{(r-j)} - a\alpha\sum_{j=1}^{r}\binom{r}{j}(j)\mu_{(j-1)}\mu_{(r-j)}\right]$$
$$-\left[\{\alpha + a\lambda - ar)\}\mu_{(r)} - ar\alpha\mu_{(r-1)}\right],\ r \geq 1 \qquad (5.2.6)$$

The first factorial moment of the distribution is

$$\mu_{(1)} = \frac{\alpha(1-\beta)+\lambda\beta}{\omega(1-\beta)},$$

Hence mean and variance of GCD2 are respectively

$$\mu = \frac{\alpha(1-\beta)+\lambda\beta}{\omega(1-\beta)},$$

$$\sigma^2 = \frac{\omega\mu - \alpha\beta}{\omega(1-\beta)} + (\omega+1)\mu^2.$$

c) Moment recurrence relation

The mgf of GCD2 corresponding to the pgf (5.2.1), may be written as

$$M(t) = \frac{\omega}{(\omega+1) - e^{-\alpha}(1-\beta)^{\lambda} e^{\alpha e^t}(1-\beta e^t)^{-\lambda}}, \qquad \omega > 0 \qquad (5.2.7)$$

Differentiating (5.2.7) w.r.t. 't' we get

$$(1-\beta e^t)M'(t) = \frac{\omega}{\omega+1}\{(\alpha+\lambda a) - \alpha\beta e^t\}e^t M(t)M(t) - \{(\alpha+\lambda a) - \alpha\beta e^t\}e^t M(t)$$

Expanding and equating the coefficients of $\frac{t^r}{r!}$ on both sides the moment recurrence relation may be obtained as

$$\mu'_{r+1} = B\left[\sum_{i=0}^{r}\binom{r}{i}\{(\alpha+\lambda\beta) - \alpha\beta 2^{r-i}\}\sum_{j=0}^{i}\binom{i}{j}\mu'_j\mu'_{i-j}\right] - A\left[\sum_{j=1}^{r}\binom{r}{j}\{\alpha+\lambda\beta-\alpha\beta 2^{r-j}\}\mu'_j + \right.$$
$$\left.\{\alpha+\lambda\beta) - \alpha\beta 2^r\} - \beta\sum_{j=1}^{r}\binom{r}{j}\mu'_{r-j+1}\right], r \geq 1 \qquad (5.2.8)$$

where $A = \frac{1}{1-\beta}$, $B = \frac{\omega+1}{(1-\beta)\omega}$ and $\mu'_1 = \frac{\alpha(1-\beta)+\lambda\beta}{\omega(1-\beta)}$.

d) Cumulant recurrence relation

Taking logarithm on both sides of (5.2.7), the cgf of GCD2 may be written as

$$K(t) = \log\omega - \log\{(\omega+1) - m(t)\} \qquad (5.2.9)$$

where $m(t) = e^{-\alpha}(1-\beta)^{\lambda} e^{\alpha e^t}(1-\beta e^t)^{-\lambda}$, is the mgf of the three parameter Charlier Distribution.

Differentiating (5.2.9) w.r.t. 't' and equating the coefficients of $\frac{t^r}{r!}$ on both sides, the cumulant recurrence relation of GCD2 may be obtained as

$$K_{r+1} = A\sum_{j=1}^{r}\binom{r}{j}K_{r-j+1} + B\sum_{j=1}^{r}\binom{r}{j}\{(\alpha+\lambda\beta)-\alpha\beta 2^{r-j}\}\mu_j' + \frac{(\alpha+\lambda\beta)-\alpha\beta 2^r}{\omega(1-\beta)},$$

$$r \geq 1. \qquad (5.2.10)$$

where $A = \frac{\beta}{1-\beta}$ and $B = \frac{(\omega+1)}{(1-\beta)\omega}$.

Here μ_r' denotes the r^{th} order 'raw' moment of the distribution.

The first and second cumulants of the distribution are respectively

$$K_1 = \frac{\alpha(1-\beta)+\lambda\beta}{\omega(1-\beta)},$$

and $$K_2 = (1+\omega)K_1^2 + \frac{K_1}{(1-\beta)} - \frac{\alpha\beta}{\omega(1-\beta)}.$$

It has been observed that from (5.2.1), Poisson mixing infinitely divisible distribution (GPD2) may be obtained by putting $\beta = 0$ and negative binomial mixing infinitely divisible distribution (GNBD2) may be obtained by putting $\alpha = 0$.

5.3 Generalized Gegenbauer Distribution of type 2 (GGD2)

a) Probability recurrence relation

The pgf of generalized Gegenbauer distribution of type 2 (GGD2) derived from the model (5.1.1) may be written as

$$G(t) = \frac{\omega}{(\omega+1)-(1-\alpha-\beta)^{\lambda}(1-\alpha t-\beta t^2)^{-\lambda}}, \qquad \omega > 0 \qquad (5.3.1)$$

Differentiating both sides of (5.3.1) w.r.t. 't' we get

$$(1-\alpha t-\beta t^2)G'(t) = \frac{(\omega+1)\lambda}{\omega}[(\alpha+2\beta t)G(t)G(t)] - \lambda\{(\alpha+2\beta t)G(t)\}$$

Equating the coefficients of t^r on both sides the recurrence relation for probabilities may be obtained as

$$G_{r+1}=\frac{(\omega+1)\lambda}{\omega(r+1)}\left\{\alpha G_0G_r+\sum_{j=1}^{r}\{\alpha G_j+2\beta G_{j-1}\}G_{r-j}\right\}+$$

$$\frac{\alpha(r-\lambda)G_r+\beta(r-1-2\lambda)G_{r-1}}{(r+1)},\quad r\geq 1. \tag{5.3.2}$$

where
$$G_0=\frac{\omega}{(\omega+1)-(1-\alpha-\beta)^{\lambda}}, \tag{5.3.3}$$

and
$$G_1=\lambda\alpha\left\{\frac{(\omega+1)G_0}{\omega}-1\right\}G_0. \tag{5.3.4}$$

b) Factorial moment recurrence relation

The fmgf of GGD2, corresponding to the pgf (5.3.1) may be written as

$$m(t)=\frac{\omega}{\omega+1-(1-at-bt^2)^{-\lambda}},\ \omega>0 \tag{5.3.5}$$

where $a=\dfrac{\alpha+2\beta}{1-\alpha-\beta}$ and $b=\dfrac{\beta}{(1-\alpha-\beta)}$.

Differentiating both sides of (5.3.5) w.r.t 't', we get

$$(1-at-bt^2)m'(t)=\frac{(\omega+1)\lambda}{\omega}\{a+2bt\}m(t)m(t)-\lambda\{a+2bt\}m(t),$$

Now equating the coefficients of $\dfrac{t^r}{r!}$ on both sides the factorial moment recurrence relation may be obtained as

$$\mu_{(r+1)}=\frac{(\omega+1)\lambda}{\omega}\left[a\sum_{j=0}^{r}\binom{r}{j}\mu_{(j)}\mu_{(r-j)}+2b\sum_{j=1}^{r}\binom{r}{j}j\mu_{(j-1)}\mu_{(r-j)}\right]$$

$$+\left[a(\lambda-r)\mu_{(r)}-rb\{2\lambda-(r-1)\}\mu_{(r-1)}\right],\ r\geq 1 \tag{5.3.6}$$

where
$$\mu_{(1)}=\frac{\lambda(\alpha+2\beta)}{\omega(1-\alpha-\beta)}.$$

Hence mean and variance of the derived distribution are respectively

$$\mu=\frac{\lambda(\alpha+2\beta)}{\omega(1-\alpha-\beta)},$$

and
$$\sigma^2=(\omega+1)\mu^2+\mu+\frac{\omega}{\lambda}\mu^2+\frac{2\lambda\beta}{\omega(1-\alpha-\beta)}.$$

c) Moment recurrence relation

Corresponding to the pgf (5.3.1), the mgf may be written as

$$M(t)=\frac{\omega}{\omega+1-(1-\alpha-\beta)^{\lambda}(1-\alpha e^{t}-\beta e^{2t})^{-\lambda}},\ \omega>0 \tag{5.3.7}$$

Differentiating both sides of (5.3.7) w.r.t 't', we get

$$(1-\alpha e^{t}-\beta e^{2t})M'(t)=\frac{\omega+1}{\omega}\lambda\{(\alpha+2\beta e^{t}\}e^{t}M(t)M(t)-\lambda\{(\alpha+2\beta e^{t}\}e^{t}M(t)$$

Equating the coefficients of $\frac{t^{r}}{r!}$ on both sides the recurrence relation for moments may be obtained as

$$\mu'_{r+1}=A\left[\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^{j})\mu'_{r-j+1}-\lambda\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^{r-j+1})\mu'_{j}-\lambda(\alpha+\beta 2^{r+1})\right]$$

$$+B\left[\sum_{i=0}^{r}\binom{r}{i}(\alpha+\beta 2^{r-i+1})\sum_{j=0}^{i}\binom{i}{j}\mu'_{j}\mu'_{i-j}\right],\ r\geq 1 \tag{5.3.8}$$

where $A=\frac{1}{1-\alpha-\beta}$, $B=\frac{(\omega+1)\lambda}{(1-\alpha-\beta)\omega}$

and $\mu'_{1}=\frac{\lambda(\alpha+2\beta)}{\omega(1-\alpha-\beta)}$.

d) Cumulant recurrence relation

Taking logarithm on both sides of (5.3.7) we get the cgf as

$$K(t)=\log\omega-\log\{\omega+1-m(t)\} \tag{5.3.9}$$

where $m(t)=(1-\alpha-\beta)^{\lambda}(1-\alpha e^{t}-\beta e^{2t})^{-\lambda}$ is the mgf of Gegenbauer distribution.

Differentiating (5.3.9) w.r.t 't' and equating the coefficients of $\frac{t^{r}}{r!}$ on both sides the cumulant recurrence relation may be obtained as

$$K_{r+1}=A\left[\frac{\lambda}{\omega}(\alpha+\beta 2^{r+1})+\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^{j})K_{r-j+1}\right]+B\sum_{j=1}^{r}\binom{r}{j}(\alpha+\beta 2^{j})\mu'_{j},$$

$$r\geq 1 \tag{4.3.10}$$

where $A=\frac{1}{1-\alpha-\beta}$, $B=\frac{(\omega+1)\lambda}{(1-\alpha-\beta)\omega}$.

Here μ_r' denotes the r^{th} order 'raw' moment of the distribution.

The first two cumulants of the distribution are respectively

$$K_1 = \frac{\lambda(\alpha + 2\beta)}{\omega(1-\alpha-\beta)},$$

and
$$K_2 = (\omega+1)K_1^2 + K_1 + \frac{\omega}{\lambda}K_1^2 + \frac{2\lambda\beta}{\omega(1-\alpha-\beta)}.$$

It has been observed that from (5.3.1) negative binomial mixing infinitely divisible distribution (GNBD2) may be obtained by putting $\beta = 0$.

5.4 Estimation of parameters

To estimate the parameters of Generalized Charlier Distribution of Type 2 (GCD2) and Generalized Gegenbauer Distribution of Type 2 (GGD2) we shall use the simple adhoc method of using sample mean $\bar{x}$, sample variance s^2 and the ratio of first three frequencies $\theta_1 = f_1/f_0$ and $\theta_2 = f_2/f_0$, as the other methods are found to be computationally complicated. In order to make the calculation easier, we shall first transform the probabilities G_i's of GCD2 and GGD2 in terms of the probabilities P_i's of the component distributions $g(t)$ used. Then probability recurrence relation of GCD2 and GGD2 corresponding to the pgf (5.1.1) may be written as

$$G_{r+1} = \frac{1}{\omega+1-P_0}\left[\sum_{j=1}^{r} G_j P_{r-j+1} + G_0 P_{r+1}\right], \ r \geq 1 \tag{5.4.1}$$

where $$G_0 = \frac{\omega}{\omega+1-P_0} \tag{5.4.2}$$

$$G_1 = \frac{\omega P_1}{(\omega+1-P_0)^2}. \tag{5.4.3}$$

Then, the ratios of first three frequencies of the distribution corresponding to (5.1.1) are respectively

$$\theta_1 = \frac{f_1}{f_0} = \frac{P_1}{\omega+1-P_0}, \tag{5.4.4}$$

$$\theta_2 = \frac{f_2}{f_0} = \frac{P_1^2 + (\omega+1-P_0)P_2}{(\omega+1-P_0)^2} = \theta_1^2 + \frac{P_2}{(\omega+1-P_0)}. \tag{5.4.5}$$

Adding equations (5.4.4) and (5.4.5) we get the estimated value of ω as

$$\hat{\omega} = \frac{P_1 + P_2}{\theta_1 + \theta_2 - \theta_1^2} - (1 - P_0) \ , \ \omega > 0 \tag{5.4.6}$$

For obtaining the remaining parameters of the distributions GCD2 and GGD2 we may proceed as follows

a) GCD2

The sample mean $\bar{x}$, sample variance s^2 and the ratios of first two frequencies f_0 and f_1 of the distribution GCD2 are respectively

$$\bar{x} = \frac{\alpha(1-\beta) + \lambda\beta}{\omega(1-\beta)}, \tag{5.4.7}$$

$$s^2 = (\omega+1)\bar{x}^2 - \frac{\alpha\beta}{\omega(1-\beta)} + \frac{\bar{x}}{(1-\beta)}, \tag{5.4.8}$$

$$\theta_1 = (\alpha + \lambda\beta)\left\{\frac{(\omega+1)G_0}{\omega} - 1\right\}, \theta_1 = \frac{f_1}{f_0}. \tag{5.4.9}$$

Eliminating α and then λ, one by one from the above three equations, we get the estimated value of β as

$$\hat{\beta} = 1 + \frac{\theta_1 - \bar{x}T_2}{T_1 T_2} \tag{5.4.10}$$

Again eliminating α from (5.4.7) and (5.4.8), and putting the value of $\hat{\beta}$ we get

$$\hat{\lambda} = \frac{1}{\hat{\beta}^2}\hat{\omega}(1-\hat{\beta})^2 T_1 \tag{5.4.11}$$

Substituting the estimated values of λ and β in (5.4.7) we get

and $$\hat{\alpha} = \hat{\omega}\bar{x} - \frac{\hat{\lambda}\hat{\beta}}{(1-\hat{\beta})}. \tag{5.4.12}$$

where $T_1 = \{s^2 - (\hat{\omega}+1)\bar{x}^2 - \bar{x}\}$,

$T_2 = (\omega+1)G_0 - \omega$

b) GGD2

The sample mean $\bar{x}$, sample variance s^2 and the ratio of first two frequencies f_0 and f_1 of the distribution corresponding to (5.3.1) are respectively

$$\bar{x} = \frac{(\alpha + 2\beta)\lambda}{\omega(1-\alpha-\beta)}, \tag{5.4.13}$$

$$s^2 = (\omega+1)\bar{x}^2 + \frac{(1+\beta)\bar{x}}{(1-\alpha-\beta)} + \frac{2\lambda\beta}{\omega(1-\alpha-\beta)}, \tag{5.4.14}$$

$$\theta_1 = \alpha\lambda\left\{\frac{(\omega+1)G_0}{\omega} - 1\right\},\ \theta_1 = \frac{f_1}{f_0}. \tag{5.4.15}$$

Eliminating λ and then β one by one from (5.4.13), (5.4.14) and (5.4.15) we get a second degree equation in α as

$$P\alpha^2 + Q\alpha + R = 0 \tag{5.4.16}$$

Solving this quadratic equation we get the estimated value of α as

$$\hat{\alpha} = \frac{-Q \pm \sqrt{Q^2 - 4PR}}{2P}, \tag{5.4.17}$$

where $P = -A\bar{x}^2$,

$Q = 2A^2\bar{x}^2 - 2A\bar{x}\theta_1 + A\theta_1\{s^2 - \bar{x} - (\omega+1)\bar{x}^2\}$,

$R = 2\theta_1[A\bar{x} - \theta_1 - A\{s^2 - \bar{x} - (\omega+1)\bar{x}^2\}]$,

and $A = \{(\omega+1)G_0 - \omega\}$.

Eliminating λ from (5.4.13) and (5.4.15) and substituting the value of $\hat{\alpha}$, we get

$$\hat{\beta} = \frac{\hat{\alpha}(1-\hat{\alpha})A\bar{x} - \hat{\alpha}\theta_1}{2\theta_1 + \hat{\alpha}A\bar{x}} \tag{5.4.18}$$

Finally putting the values of $\hat{\alpha}$ and $\hat{\beta}$ in equation (5.4.13), we get

$$\hat{\lambda} = \frac{\hat{\omega}(1-\hat{\alpha}-\hat{\beta})\bar{x}}{\hat{\alpha}+2\hat{\beta}}. \tag{5.4.19}$$

5.5 Applications of GCD2 and GGD2

To illustrate the applications of GCD2 and GGD2, in Table 5.1, we consider the well known data of Beall and Rescia (1953), on the number of plants per quadrant of Lespedeza capitate for which generalized Charlier distribution was fitted by Medhi and Borah (1986).

In Table 5.1, it has been observed that when GCD2 and GGD2 are applied to the frequency distribution of Lespedeza Capitate (data of Beall and Rescia, 1953), GCD2 provides a good fit with a χ^2 value of 2.09 than GGD2 which provides a χ^2 value of 15.32. Here, using the equation (5.4.6) we get $\hat{\omega} = 0.9065$ and for GCD2, using (5.4.10), (5.4.11) and (5.4.12) we get $\hat{\beta} = 0.6291$, $\hat{\lambda} = 0.0522$ and $\hat{\alpha} = 0.0082$. For the same value of $\omega\hat{} = 0.9065$ using the equations (5.4.17), (5.4.18) and (5.4.19), we get $\hat{\alpha} = 0.3237$, $\hat{\beta} = 0.0704$ and $\hat{\lambda} = 0.1263$ for GGD2. The Poisson-Pascal distribution by Katti and Gurland (1961) is also compared with our fitted distribution and it is found to be satisfactory.

In Tables 5.2 and 5.3, we consider the data of Chatfield (1969) on the observed frequencies of distributions of purchases of two different kinds of brands of products (brand K and D) where generalized Poisson distribution (GPD) was fitted by Consul (1989). From Table 5.2 we get $\hat{\omega} = 0.9272$ and for GCD2 we get $\hat{\beta} = 0.6595$, $\hat{\lambda} = 0.0438$ and $\hat{\alpha} = -0.0026$. In case of GGD2, for the same value of $\hat{\omega} = 0.9272$, we get $\hat{\beta} = 0.1224$, $\hat{\lambda} = 0.1022$ and $\hat{\alpha} = 0.2553$. Here also GCD2 provides a good fit with a χ^2 value of 0.561 than GGD2 which provides a χ^2 value of 2.8748.

Again Table 5.3 gives $\hat{\omega} = 0.8381$ and for GCD2 we get $\hat{\beta} = 0.7161$, $\hat{\lambda} = 0.0563$ and $\hat{\alpha} = 0.0359$. For GGD2, we get $\hat{\beta} = 0.0614$, $\hat{\lambda} = 0.1977$ and $\hat{\alpha} = 0.3800$ for the same value of $\hat{\omega} = 0.8381$. Considering the χ^2 values obtained, it is observed that the GCD2 is as good as the GPD model (Consul, 1989) for both brands but GGD2 is good for the brand K only.

Table 5.1

Observed and fitted distributions of GCD2 and GGD2 to the frequency distribution of Lespedeza Capitate (Beall and Rescia, 1953) (Medhi and Borah, 1986)

No. of Plants per quadrant	Observed Frequency $\bar{x} = 0.1068$ $s^2 = 0.2944$	Fitted distributions		TPCD (Medhi and Borah, 1986)	Poisson-Pascal Katti and Gurland (1961)
		GCD2 $\omega\hat{} = 0.9065$ $\beta\hat{} = 0.6291$ $\lambda\hat{} = 0.0522$ $\alpha\hat{} = 0.0082$	GGD2 $\omega\hat{} = 0.9065$ $\beta\hat{} = 0.0704$ $\lambda\hat{} = 0.1263$ $\alpha\hat{} = 0.3237$		
0	7178	7179.00	7155.95	7179.13	7185.0
1	286	287.62	283.76	286.05	276.0
2	93	89.82	124.67	91.52	94.5
3	40	40.15	43.80	40.28	41.5
4	24	19.93	18.26	20.02	20.2
5	7	10.45	6.95	10.33	10.4
6	5	5.67	2.85	5.61	5.6
7	1	3.14	1.19	3.12	3.1
8	2	1.78	0.51	1.77	1.7
9	1	1.02	0.22	1.02	1.0
10	2	0.57	0.10	0.45	0.6
11	1	0.85	1.74	0.45	0.3
Total	7640	7640.00	7640.00	7640.00	7640.0
	χ^2	2.09	15.32	1.59	9.58
	$d.f.$	7	7	8	8
	$p-value$	>0.95	>0.03	>0.99	>0.29

GCD2: Generalized Charlier distribution of type 2

GGD2: Generalized Gegenbauer distribution of type 2

TPCD: Three parameter Charlier distribution

Table 5.2

Observed and fitted distributions of GCD2 and GGD2 to the Number of Units r of Different Brands Bought by Number of Consumers Over a number of weeks (Consul, 1989)

No of Units	Observed frequency Chatfield (brand K) $\bar{x} = 0.0886$ $s^2 = 0.2807$	Fitted distributions		Generalized Poisson Distribution Consul(1989) $\hat{\theta} = 0.0463$ $\hat{\lambda} = 0.4770$
		GCD2 $\hat{\omega} = 0.9272$ $\hat{\beta} = 0.6595$ $\hat{\lambda} = 0.0438$ $\hat{\alpha} = -0.0026$	GGD2 $\hat{\omega} = 0.9272$ $\hat{\beta} = 0.1224$ $\hat{\lambda} = 0.1022$ $\hat{\alpha} = 0.2553$	
0	1671	1671.39	1665.00	1670.78
1	43	43.28	42.49	48.04
2	19	17.37	27.43	14.91
3	9	8.18	8.17	6.73
4	2	4.24	3.94	3.57
5	3	2.32	1.43	2.08
6	1	1.31	0.63	1.28
7	0	0.76	0.27	0.82
8	0	0.45	0.12	0.54
9	2	0.70	0.52	1.25
Total	1750	1750.00	1750.00	1750.00
	χ^2	0.5610	2.8748	2.66
	$d.f.$	1	1	2
	$p - value$	> 0.45	> 0.08	> 0.26

GCD2: Generalized Charlier distribution of type 2

GGD2: Generalized Gegenbauer distribution of type 2

Table 5.3

Observed and fitted distributions of GCD2 and GGD2 to the Number of Units r of Different Brands Bought by Number of Consumers Over a number of weeks (Consul, 1989)

No of Units	Observed frequency Chatfield (brand D) $\bar{x}=0.21225$ $s^2=0.7223$	Fitted distributions		Generalized Poisson Distribution $\hat{\theta}=0.1133$ $\hat{\lambda}=0.4663$
		GCD2 $\hat{\omega}=0.8381$ $\hat{\beta}=0.7161$ $\hat{\lambda}=0.0563$ $\hat{\alpha}=0.0359$	GGD2 $\hat{\omega}=0.8381$ $\hat{\beta}=0.0614$ $\hat{\lambda}=0.1977$ $\hat{\alpha}=0.3800$	
0	875	874.34	867.45	875.05
1	63	63.75	61.35	62.18
2	19	19.15	28.21	20.40
3	10	9.20	12.08	9.33
4	4	5.14	5.65	4.95
5	4	3.06	2.43	2.87
6	1	1.89	1.09	1.75
7	2	1.19	0.51	1.12
8	0	0.77	0.24	0.73
9	1	0.50	0.11	0.49
10	0	0.33	0.05	0.33
11	0	0.22	0.03	0.23
12	1	0.46	0.80	0.57
total	980	980.00	980.00	980.00
	χ^2	0.3729	6.6162	0.44
	$d.f.$	2	2	3
	$p-value$	>0.82	>0.03	>0.93

GCD2: Generalized Charlier distribution of type 2

GGD2: Generalized Gegenbauer distribution of type 2

5.6 Properties of particular cases

The limiting distributions of GCD2 and GGD2 are some generalized distributions of Poisson and negative binomial distribution. Attempt has been made to study certain important properties of these distributions along with their estimation of parameters and fitting of distributions.

a) Poisson mixing infinitely divisible distribution of Type 2 (GPD2)

The pgf of Poisson mixing infinitely divisible distribution (GPD2) obtained from (5.2.1) by putting $\beta = 0$ may be written as

$$G(t) = \frac{\omega}{(\omega+1) - e^{\alpha(t-1)}}\ , \quad \omega > 0 \tag{5.6.1}$$

It has the probability recurrence relation

$$G_{r+1} = \frac{1}{r+1}\left[\frac{\alpha(\omega+1)}{\omega}\left\{G_0 G_r + \sum_{j=1}^{r} G_j G_{r-j}\right\} - \alpha G_r\right],\ for\ r \geq 1 \tag{5.6.2}$$

$$= \frac{e^{-\alpha}}{\omega + 1 - e^{-\alpha}}\left[\frac{\alpha^{r+1} G_0}{(r+1)!} + \sum_{j=1}^{r}\frac{\alpha^j}{j!} G_{r-j+1}\right],\ for\ r \geq 1.$$

where $G_0 = \dfrac{\omega}{(\omega+1) - e^{-\alpha}}$ and $G_1 = \alpha G_0\left\{\dfrac{\omega+1}{\omega} G_0 - 1\right\}$.

Its factorial moment recurrence relation is

$$\mu_{(r+1)} = \frac{\alpha^{r+1}}{\omega} + \frac{r+1}{\omega}\sum_{j=1}^{r}\binom{r}{j-1}\frac{\alpha^j}{j}\mu_{(r-j+1)}\ ,\ for\ r \geq 1 \tag{5.6.3}$$

where $\mu_{(1)} = \dfrac{\alpha}{\omega}$ is the first factorial moment of the distribution.

Therefore mean and variance of the distribution are respectively

$$\mu = \frac{\alpha}{\omega} \text{ and } \sigma^2 = \{(\omega+1)\mu + 1\}\mu.$$

Estimation of parameters

For estimating the parameters of Poisson mixing infinitely divisible distribution the method of moment may be used, which gives

$$\hat{\alpha} = \frac{s^2}{\bar{x}} - \bar{x} - 1\ ,\quad \hat{\omega} = \frac{s^2 - \bar{x} - \bar{x}^2}{\bar{x}^2}.$$

where $\bar{x}$ and s^2 denote respectively the sample mean and sample variance of the distribution. Again using the first sample moment $\bar{x}$ and ratio of first two frequencies $\theta = \frac{f_1}{f_0}$, we may have

$$\hat{\omega} = \frac{\theta - \bar{x}G_0}{\bar{x}(G_0 - 1)} \text{ and } \hat{\alpha} = \bar{x}\hat{\omega}$$

b) Negative binomial mixing infinitely divisible distribution of Type 2 (GNBD2)

The pgf of negative binomial mixing infinitely divisible distribution (GNBD2) derived from the model (5.2.1) by putting $\alpha = 0$ may be written as

$$G(t) = \frac{\omega}{\omega + 1 - (1-\beta)^{\lambda}(1-\beta t)^{-\lambda}}, \quad \omega > 0 \qquad (5.6.4)$$

It has the probability recurrence relation

$$G_{r+1} = \frac{1}{r+1}\left[\beta(r-\lambda)G_r + \frac{(\omega+1)\lambda\beta}{\omega}\sum_{j=0}^{r} G_j G_{r-j}\right], \quad for\ r \geq 1 \qquad (5.6.5)$$

$$= \frac{(1-\beta)^n}{\omega + 1 - (1-\beta)^n}\left\{\binom{\lambda + r}{r+1} b^{r+1} G_0 + \sum_{j=0}^{r}\binom{\lambda + r - 1}{r - j + 1} b^{r-j+1} G_j\right\},$$

where $\quad P_0 = \dfrac{\omega}{\omega + 1 - (1-\beta)^n}, \quad P_1 = n\beta(1-\beta)^n \dfrac{P_0^2}{\omega}$

and $b = \dfrac{\beta}{1-\beta}$

Its factorial moment recurrence relation is

$$\mu_{(r+1)} = \frac{(r+1)!}{\omega}\left\{\binom{\lambda + r}{r+1} b^{r+1} + \sum_{j=0}^{r}\binom{\lambda + r - j}{r - j + 1} b^{r-j+1}\frac{\mu_{(j)}}{j!}\right\}, \quad r \geq 1 \qquad (5.6.6)$$

where $b = \dfrac{\beta}{1-\beta}$, $\mu_{(1)} = \dfrac{\lambda\beta}{\omega(1-\beta)}$,

Hence mean and variance are respectively

$$\mu = \frac{\lambda\beta}{\omega(1-\beta)}$$

and $\qquad \sigma^2 = \mu + (\omega+1)^2\mu + \dfrac{\beta\mu}{(1-\beta)}.$

Estimation of parameters

The parameters (λ, α and ω) of the negative binomial mixing infinitely divisible distribution may be estimated by using the following three methods

i) Method of using first three factorial moments

The first three sample factorial moments of the distribution are respectively

$$m_{(1)} = \frac{\lambda b}{\omega},$$

$$m_{(2)} = m_{(1)}\{b(\lambda+1) + 2m_{(1)}\},$$

and $$m_{(3)} = (\lambda+2)bm_{(2)} + 4m_{(1)}m_{(2)} - 2\{m_{(1)} + b\}m_{(1)}^2,$$

where $b = \dfrac{\beta}{(1-\beta)}$.

writing $$R = \mu_{(2)} - 2\mu^2_{(1)},$$

$$Q = \mu_{(3)}\mu_{(1)} - 4\mu_{(1)}^2\mu_{(2)} - R\mu_{(2)} + 2\mu_{(1)}^4.$$

we have $$\hat{\lambda} = \frac{R^2 - Q}{Q},\ \hat{\beta} = \frac{Q}{Q + R\mu_{(1)}} \text{ and } \hat{\omega} = \frac{R^2 - Q}{R\mu^2_{(1)}}.$$

ii) Method of first two sample moments and ratio of first two frequencies

First two sample moments and the ratio of first two frequencies are respectively

$$\bar{x} = \frac{\lambda\beta}{\omega(1-\beta)},$$

$$s^2 = (\omega+1)\bar{x}^2 + \frac{\bar{x}}{(1-\beta)},$$

and $$\theta = \frac{f_1}{f_0} = \lambda\beta\left\{\frac{(\omega+1)G_0}{\omega} - 1\right\}.$$

Solving the above three equations we obtained

$$\hat{\omega} = \frac{(s^2 - \bar{x})\theta - \bar{x}^2 P_0}{\{(P_0 - 1) + \theta\}\bar{x}^2},$$

$$\hat{\beta} = 1 - \frac{\bar{x}}{(s^2 - \bar{x}^2) - \hat{\omega}\bar{x}^2},$$

and $$\hat{\lambda} = \frac{\bar{x}(1-\hat{\beta})\hat{\omega}}{\hat{\beta}}.$$

where $P_0 = \frac{f_0}{N}$.

iii) Method of first sample moment and ratio of first three frequencies

First sample moment $\bar{x}$ and the ratio of first three frequencies $\theta_1 = \frac{f_1}{f_0}$ and $\theta_2 = \frac{f_2}{f_0}$ are respectively

$$\bar{x} = \frac{\lambda\beta}{\omega(1-\beta)},$$

$$\theta_1 = \frac{f_1}{f_0} = \lambda\beta\left\{\frac{(\omega+1)G_0}{\omega} - 1\right\},$$

and $$\theta_2 = \frac{f_2}{f_0} = \frac{\beta(1-\lambda)\theta_1}{2} + \frac{\lambda\beta(\omega+1)\theta_1 G_0}{\omega}.$$

Solving the above three equations we have

$$\hat{\lambda} = \frac{B \pm \sqrt{B^2 - 4AC}}{2A},$$

where $A = \{\theta_1(G_0 - 1)\}^2$,

$$B = -2(G_0 - 1)\{\bar{x}G_0 + (G_0 - 1)\}\theta_1^2,$$

and $C = \{(G_0 - 1)(\theta_1 - 2\theta_2) + G_0\theta_1^2\}\{(2\bar{x} - \theta_1)G_0\theta_1 + (G_0 - 1)(\theta_1 + 2\theta_2)\}$.

and $$\hat{\beta} = \frac{(\bar{x} - \theta_1)G_0 + T(G_0 - 1)}{(G_0 - 1) - \lambda(G_0 - 1) + \bar{x}P_0},$$

$$\hat{\omega} = \frac{(\theta_1 - \bar{x}G_0) + \beta\bar{x}G_0}{\bar{x}(G_0 - 1)(1-\beta)}.$$

5.7 Application

For the applications of generalized negative binomial Distribution of Type 2 i.e., for GNBD2 we have considered the Student's historic data on Haemocytometer counts of yeast cells in Table 5.4 where Hermite distribution was fitted by Kemp and Kemp (1966) and Gegenbauer distribution was fitted by Borah (1984). In Table 5.4, on the basis of $\chi 2$ criterion it can be said that the expected frequencies computed by our derived distribution match satisfactorily as the other distributions compared.

TABLE 5.4

Observed and fitted Generalized negative binomial distribution of Type 2 (GNBD2)
(Haemocytometer Counts of Yeast Cells)

No. of Insects	Observed frequency $\bar{x}=0.6825$ $s^2=0.81169$	Fitted dist. GNBD2 $\hat{\lambda}=2.7463$, $\hat{\beta}=0.2202$ $\hat{\omega}=0.8799$ (MFM)	Gegenbauer Distribution (Borah,1984)	Hermite Distribution (Kemp and Kemp, 1965)
0	213	214.77	214.15	213.12
1	128	121.52	123.00	122.91
2	37	45.39	44.88	46.71
3	18	13.74	13.36	13.31
4	3	3.58	3.55	3.16
5	1	0.81	0.86	0.64
6	0	0.19	0.20	0.15
Total	400	400.00	400.00	400.00
	$\chi 2$	3.27	2.8342	3.8825
	df	1	1	2
	$p-value$	>0.07	>0.09	>0.14

MFM: method of factorial moments

Chapter 6

A Class of Lagrangian Discrete Probability Distributions

6.1 Introduction

Lagrangian expansions for the derivation of expressions for the probabilities of certain discrete distributions have been used for many years. The potential of this technique for deriving distributions and their properties has been systematically exploited by Consul and Shenton (1972, 1973 and 1975), Consul and Jain (1973), Janardan and Rao (1983) and their co-workers. Consul and Felix (1996) studied the Lagrangian Katz family of distributions with estimation of parameters and applications. Consul in his book (1989) on generalized Poisson distributions offers a systematic study of Lagrangian Poisson distribution.

The objective of this chapter is to investigate the probabilistic structures of a class of Charlier Family of Lagrangian (CFL) discrete probability distributions having wide flexibility. Using Lagrange's expansions of $type-I$ and $type-II$, the CFL distributions has been derived like the authors Consul and Shenton (1972, 1973) and Janardan and Rao (1983) by taking different choice of the pgf's of three parameter charlier, Poisson, negative binomial, Logarithmic series and delta distributions as $f(s)$ and $g(s)$.

The pmf of CFL distributions of the first kind and second kind are presented in Tables **6.1** and **6.2** respectively (see also Appendix A).

The basic Lagrangian Poisson (LP) and Lagrangian negative binomial (LNB) distributions together with estimation of their parameters and fitting of the distributions have been also studied. These basic distributions have been fitted to some published data collected by Williams (1944) for which generalized geometric distribution had been fitted by Plunkett and Jain (1975). It has been found that the basic LNB distribution gives better fit than the other distributions compared.

The general Lagrangian distributions such as Lagrangian Poisson negative binomial, Lagrangian Poisson Logarithmic and Lagrangian negative binomial Poisson distributions of first kind are fitted to some secondary data sets, where generalized Poisson distribution (GPD) was fitted by Consul (1989).

6.2 Lagrangian Distributions

A class of discrete probability distributions under the title 'Lagrangian Distributions' had been introduced into the literature by Consul and Shenton (1972, 1973, 1975). 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 $g(s)$ as a power series in u when $u = s/g(s)$. Since Lagrange expansion seems to be associated with queueing processes, they used Lagrange expansion for generating new families of discrete generalized distributions which satisfy the convolution property.

If $g(s)$ and $f(s)$ are two probability generating functions (pgf) defined on some or all non-negative integers, such that $g(0) \neq 0$, Consul and Shenton (1972) used Lagrange's expansion to define families of discrete generalized probability distributions which is called Lagrange distributions of first kind (LD1) according to the terminology of Janardan and Rao (1983). Lagrange distribution of this kind has the probability mass function (pmf)

$$P_r(X = x) = \frac{1}{x!}\left[\frac{\delta^{x-1}}{\delta s^{x-1}}\{g(s)\}^x f'(s)\right]_{S=0}, \qquad for\ x = 1,2,3,... \qquad (6.2.1)$$

where $P_r(X = 0) = f(0)$.

Again using Lagrange's expansion of second kind, Janardan and Rao (1983) investigate another class of discrete distributions called Lagrange distributions of second kind (LD2) with pmf

$$P_r(X=x)=\frac{1-g'(1)}{x!}\left[\frac{\delta^x}{\delta s^x}\{g(s)\}^x f(s)\right]_{S=0} \quad , \ for\ x=0,1,2,3,... \tag{6.2.2}$$

$$=0, \qquad otherwise$$

Consul and Shenton (1972), and Janardan and Rao (1983) generate different families of Lagrange distributions by various choice of the functions $g(s)$ and $f(s)$.

All Lagrangian distributions (LD) corresponding to the given transformation $s=ug(s)$, are closed under convolution. Consul and Shenton (1973) also proved that under one set of limiting conditions all discrete Lagrangian distributions tend to normality, and that under another set of limiting conditions they tend to inverse Gaussian distribution.

a) Properties of Moments and Cumulants of general Lagrangian distributions

The general Lagrangian Probability Distribution (LPD) possesses, some important properties (see Consul and Shenton, 1973). The r^{th} cumulants K_r, for $r=1,2,3,...$, of the general LPD may be expressed as

$$K_r=\sum_{m=1}^{r}F_m\left[\sum\frac{r!}{\pi_1!\pi_2!...\pi_r!}\sum_{i=1}^{r}\left(\frac{D_i}{i!}\right)^{\pi_i}\right] \tag{6.2.3}$$

where the second summation is taken over all partitions $\pi_1,\pi_2,...,\pi_r$. of m such that $\pi_1+2\pi_2+3\pi_3+...+r\pi_r=r$. Similar expression is also available for moments (Consul and Shenton, 1972). Hence, these formulae can be used to write down for higher moments and cumulants of any generalized Lagrangian probability distributions (LPD).

For simplicity, let F_r be the r^{th} cumulant for the pgf $f(s)$, and let D_r be the r^{th} cumulant for the basic Lagrangian distribution obtained from $g(s)$. Then the first few cumulants of LPD can be written down as particular cases of (6.2.3) in the form

$$K_1=F_1D_1,$$

$$K_2=F_1D_2+F_2D_1^2,$$

$$K_3=F_1D_3+3F_2D_1D_2+F_3D_1^3,$$

$$K_4=F_1D_4+3F_2D_2^2+4F_2D_1D_3+6F_3D_1^2D_2+F_4D_1^4, \tag{6.2.4}$$

[see Consul and Shenton (1975)].

Hence the parameters of LPD can be estimated in terms of its cumulants. As the method of maximum likelihood will be very cumbersome in this case due to the complexity of the pmf, the moments may be used to estimate the parameters. Let us consider the most basic Lagrangian distribution.

b) The Basic Lagrangian Distribution

The discrete distribution represented by

$$P_r(X=x)=\frac{1}{x!}\left[\frac{\delta^{x-1}}{\delta s^{x-1}}\{g(s)\}^x\right]_{S=0}, \quad for\ x=1,2,3,... \tag{6.2.5}$$

$$=0, \qquad otherwise$$

is called basic Lagrangian probability, where $g(s)$ is a pgf defined on non-negative integers such that $g(0)\neq 0$. Examples of such basic Lagrangian type distributions are Boral-Tanner distributions (Tanner, 1961), Haight distributions (Haight, 1961), consul distributions (Consul and Shenton, 1975 and Consul, 1983) and geometric distributions. Consul and Shenton (1975) showed that all Lagrangian distributions of this basic type are closed under convolution.

c) Cumulants of Basic Lagrangian Distribution

We have seen that the cumulants given by (6.2.3) of the general LPD depend upon the cumulants $D_i, i=1,2,3,...,$ of the basic LPD, we may derive simple expressions for D_i. If G_r be the r^{th} cumulant ($r=1,2,3...$) of the distribution taken as $g(s)$, the first four cumulants of basic Lagrangian Distribution can be written as

$$D_1=\frac{1}{1-G_1},$$

$$D_2=\frac{G_2}{(1-G_1)^3},$$

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

$$D_4=\frac{G_4}{(1-G_1)^5}+10\frac{G_3G_2}{(1-G_1)^6}+15\frac{G_2{}^3}{(1-G_1)^7}. \tag{6.2.6}$$

(see Consul and Shenton, 1975)

d) Derivation of Charlier Family of Lagrangian Distributions

The pgf of three parameter Charlier distribution is given by

$$g(s) = G_1^{\alpha,\beta,\lambda}(s) = e^{-\alpha}(1-\beta)^{\lambda} e^{\alpha s}(1-\beta s)^{-\lambda}, \quad \alpha,\beta,\lambda \geq 0 \tag{6.2.7}$$

Its probability mass function (pmf) is

$$P_r^{\alpha,\beta,\lambda} = \frac{1}{r!}\frac{\delta^r}{\delta s^r}\{G_1^{\alpha,\beta,\lambda}(s)\} \tag{6.2.8}$$

$$= \frac{e^{-\alpha}(1-\beta)^{\lambda}}{r!}\sum_{j=0}^{r} {}^rC_j \frac{\beta^j \alpha^{r-j}\Gamma(\lambda+j)}{\Gamma(\lambda)}, \quad for\ r = 1,2,3,...$$

where $\quad P_0^{\alpha,\beta,\lambda} = e^{-\alpha}(1-\beta)^{\lambda}$.

As the ranges for the skewness and kurtosis of the three parameter Charlier distribution covers the ranges of the other basic distributions, viz., Poisson, negative binomial etc., as demonstrated in the literature, this family of Lagrangian distribution may be considered to be more flexible.

Similarly, the pgf of Poisson, negative binomial, Logarithmic series and delta distributions may be given by

$$G_2^{\gamma}(s) = \exp\{\gamma(s-1)\},$$

$$G_3^{\beta,N}(s) = (1-\beta)^N(1-\beta s)^{-N},$$

$$G_4^{\gamma}(s) = \log(1-\beta s)/\log(1-\beta)$$

and $\quad G_5^{n}(s) = s^n \qquad (6.2.9)$

respectively, where prefixes (i.e. $\alpha,\beta,\lambda,\gamma,n,N$) denote the parameter/s of the corresponding distribution.

A new class (Charlier Type) of Lagrangian distributions has been derived by selecting various pgf given in equations (6.2.7) and (6.2.9) at random as $g(s)$ and $f(s)$ respectively, and putting them in the expression (6.2.1) and (6.2.2) just like the other authors [Consul et al., (1973), and Janardan et al., (1983)]. For example, let us take $g(s)$ as pgf of three parameter Charlier distribution, i.e. $G_1^{\alpha,\beta,\lambda}(s)$ and $f(s)$ as pgf of Poisson distribution, i.e. $G_2^{\gamma}(s)$. Then the pmf of Lagrangian Charlier Poisson (LCP1) distribution of first kind will be given by

$$P(X=x)=\frac{1}{x!}\frac{\delta^{x-1}}{\delta s^{x-1}}\left[\{G_1^{\alpha,\beta,\lambda}(s)\}^x\frac{\delta}{\delta s}\{G_2^{\gamma}(s)\}\right]_{s=0}, for\ r=1,2,3,... \tag{6.2.10}$$

where $P(X=0)=f(0)$.

Or we may write equation (6.2.10) as

$$P_r^{\alpha,\beta,\lambda}=\frac{\gamma(1-\beta)^{\lambda x}e^{-(\alpha x+\gamma)}(\alpha x+\gamma)^{x-1}}{x!}{}_2F_0(1-x,\lambda x,-\beta/(\alpha x+\gamma)),\ for\ x=1,2,3,....$$

where ${}_2F_0(a,b;;x)$ denotes Hypergeometric function.

This equation may be put in a simpler form as

$$P(X=x)=\frac{\gamma(1-\beta)^{\lambda x}}{x!}e^{-(\alpha x+\gamma)}\sum_{j=0}^{r}\binom{r}{j}(\lambda x)_j(\alpha x+\gamma)^{r-j}\beta^j,\ \alpha,\beta,\lambda>0. \tag{6.2.11}$$

$$for\ r=1,2,3,...,\ r=x-1$$

where $P(X=0)=Ae^{-\gamma}$.

and $(\lambda x)_j=\lambda x(\lambda x+1)(\lambda x+2)...(\lambda x+j-1)$.

Similarly, Lagrangian Charlier Poisson distribution of second kind may be obtained as

$$P_r^{\alpha,\beta,\lambda}=\frac{(1-\beta)^{\lambda x}e^{-(\alpha x+\gamma)}(\alpha x+\gamma)^{x-1}}{x!}\left\{1-\left(\alpha+\frac{\lambda\beta}{1-\beta}\right)\right\}{}_2F_0(-x,\lambda x,\beta/(\alpha x+\gamma)),$$

$$x\geq 1 \tag{6.2.12}$$

This may also be written in a simpler form as

$$P(X=x)=\frac{A(1-\beta)^{\lambda x}e^{-(\alpha x+\gamma)}}{x!}\sum_{j=0}^{x}\binom{x}{j}(\lambda x)_j(\alpha x+\gamma)^{x-j}\beta^j, x\geq 0 \tag{6.2.13}$$

where $A=\left\{1-\left(\alpha+\frac{\lambda\beta}{1-\beta}\right)\right\}$ and $(\lambda x)_j=\lambda x(\lambda x+1)(\lambda x+2)...(\lambda x+j-1)$.

The basic Lagrangian Charlier distributions of first kind may be obtained by putting $n=1$ in serial number 7, in Table **6.1** as

$$P_r^{\alpha,\beta,\lambda}=\frac{(1-\beta)^{\lambda x}e^{-\alpha x}}{x!}(\alpha x)^{x-1}{}_2F_0(1-x,\lambda x,-\beta/\alpha x),\ x\geq 1 \tag{6.2.14}$$

This may be written in a simpler form as

$$P(X=x)=\frac{(1-\beta)^{\lambda x}e^{-\alpha x}}{x!}\sum_{j=0}^{r}\binom{r}{j}(\lambda x)_j(\alpha x)^{r-j}\beta^j, \alpha,\beta,\lambda>0,\ \beta<1 \tag{6.2.15}$$

$$for\ x=1,2,3...\ and\ r=x-1$$

where $\qquad (\lambda x)_j = \lambda x(\lambda x+1)(\lambda x+2)...(\lambda x+j-1)$

Similarly, the basic Lagrangian Charlier distributions of second kind may be obtained as

$$P_r^{\alpha,\beta,\lambda} = \frac{(1-\beta)^{\lambda x} e^{-\alpha x}}{(x-1)!}\left\{1-\left(\alpha+\frac{\lambda\beta}{1-\beta}\right)\right\}(\alpha x)^{x-1}\,{}_2F_0(1-x,\lambda x,\beta/\alpha x)\,,x\geq 1 \qquad (6.2.16)$$

by putting $n=1$ in serial number 14, in Table **6.2**

This may also be written as

$$P(X=x) = \frac{A(1-\beta)^{\lambda x} e^{-\alpha x}}{(x-1)!}\sum_{j=0}^{r}\binom{r}{j}(\lambda x)_j(\alpha x)^{r-j}\beta^j\,, \qquad (6.2.17)$$

for $\quad x=1,2,3...,$ *and* $r=x-1$.

where $(\lambda x)_j = \lambda x(\lambda x+1)(\lambda x+2)...(\lambda x+j-1)$

and $\qquad A=\left\{1-\left(\alpha+\frac{\lambda\beta}{1-\beta}\right)\right\},$

It is very interesting to note that the generalized Poisson distribution (see Consul, 1973) is a particular case of Charlier family of Lagrange distributions. All these distributions shown in Table 6.1 and 6.2 will be relevance in queueing theory and process. The interesting properties were discussed by Consul and Shenton (1973) in the case of generalized Poisson distribution. In general, it is also conceivable that discrete data occurring in ecology, epidemiology, and meteorology could be statistically modeled on one of the distributions considered in this investigation, see for example Jain (1975).

Table 6.1: Charlier Family of Lagrange Distributions of first kind

Sl. No.	$g(s)$	$f(s)$	$LD1(g;f;x)$
1	$G_1^{\alpha,\beta,\lambda}(s)$	$G_2^{\gamma}(s)$	$\gamma(1-\beta)^{\lambda x} e^{-(\alpha x+\gamma)}(\alpha x+\gamma)^{x-1} / x!\, {}_2F_0\{1-x, \lambda x, -\beta/(\alpha x+\gamma)\}, x \geq 1.$
2	$G_1^{\alpha,\beta,\lambda}(s)$	$G_3^{\beta,N}(s)$	$N\beta(1-\beta)^{\lambda x+N} e^{-\alpha x}(\alpha x)^{x-1} / x!\, {}_2F_0\{1-x, \lambda x+N+1, -\beta/\alpha x\}, x \geq 1.$
3	$G_2^{\gamma}(s)$	$G_3^{\beta,N}(s)$	$N\beta(1-\beta)^{N} e^{-\gamma x}(\gamma x)^{x-1} / x!\, {}_2F_0\{1-x, N+1, -\beta/\gamma x\}, x \geq 1.$
4	$G_1^{\alpha,\beta,\lambda}(s)$	$G_4^{\beta}(s)$	$a\beta(1-\beta)^{\lambda x} e^{-\alpha x}(\alpha x)^{x-1} / x!\, {}_2F_0\{1-x, \lambda x+1, -\beta/\alpha x\}, x \geq 1,\ a=\{-\log(1-\beta)\}^{-1}.$
5	$G_2^{\gamma}(s)$	$G_4^{\beta}(s)$	$a\beta e^{-\gamma x} / x!\, {}_2F_0\{1-x, 1, -\beta/\gamma x\}, x \geq 1,\ a=\{-\log(1-\beta)\}^{-1}.$
6	$G_3^{\beta,N}(s)$	$G_2^{\gamma}(s)$	$(1-\beta)^{Nx} e^{-\gamma}(\gamma)^{x} / x!\, {}_2F_0\{1-x, Nx, -\beta/\gamma\}, x \geq 1.$
7	$G_1^{\alpha,\beta,\lambda}(s)$	$G_5^{n}(s)$	$(n/x)(1-\beta)^{\lambda x} e^{-\alpha x}(\alpha x)^{x-n} / (x-n)!\, {}_2F_0\{n-x, \lambda x, -\beta/\alpha x\}, x \geq n.$

$G_1^{\alpha,\beta,\lambda}(s)$, $G_2^{\gamma}(s)$, $G_3^{\beta,N}(s)$, $G_4^{\beta}(s)$ and $G_5^{n}(s)$ denote the probability generating function of three parameter Charlier, Poisson, negative binomial, Logarithmic series and delta distribution respectively. ${}_2F_0(a,b;x)$ denotes the Hypergeometric function.

Table 6.2: Charlier Family of Lagrange Distributions of second kind

Sl. No.	$g(s)$	$f(s)$	$LD2(g;f;x)$
8	$G_2^{\gamma}(s)$	$G_3^{\beta,N}(s)$	$(1-\gamma)(1-\beta)^N e^{-\gamma x}(\gamma x)^{x-1}/x!_2F_0\{-x,N,-\beta/\gamma x\}, x\geq 1$.
9	$G_3^{\beta,N}(s)$	$G_2^{\gamma}(s)$	$\{1-N\beta/(1-\beta)\}(1-\beta)^{Nx}e^{-\gamma}\gamma^x/x!_2F_0\{-x,Nx,-\beta/\gamma\}, x\geq 1$.
10	$G_2^{\gamma}(s)$	$G_1^{\alpha,\beta,\lambda}(s)$	$(1-\gamma)(1-\beta)^{\lambda}e^{-\gamma x}(\gamma x+\alpha)^x/x!_2F_0\{-x,\lambda,-\beta/(\gamma x+\alpha)\}, x\geq 1$.
11	$G_3^{\beta,N}(s)$	$G_1^{\alpha,\beta,\lambda}(s)$	$\{1-N\beta/(1-\beta)\}(1-\beta)^{Nx+\lambda}e^{-\alpha}\alpha^x/x!_2F_0\{-x,Nx+\lambda,-\beta/\alpha\}, x\geq 1$.
12	$G_1^{\alpha,\beta,\lambda}(s)$	$G_3^{\beta,N}(s)$	$[1-\{\alpha+\lambda\beta/(1-\beta)\}](1-\beta)^{\lambda x+N}e^{-\alpha x}(\alpha x)^x/x!_2F_0\{-x,\lambda x+N,-\beta/\alpha x\}, x\geq 1$.
13	$G_1^{\alpha,\beta,\lambda}(s)$	$G_2^{\gamma}(s)$	$[1-\{\alpha+\lambda\beta/(1-\beta)\}](1-\beta)^{\lambda x}e^{-(\alpha x+\gamma)}(\alpha x+\gamma)^x/x!_2F_0\{-x,\lambda,-\beta/(\alpha x+\gamma)\}, x\geq 1$.
14	$G_1^{\alpha,\beta,\lambda}(s)$	$G_5^{n}(s)$	$[1-\{\alpha+\lambda\beta/(1-\beta)\}](1-\beta)^{\lambda x}e^{-\alpha x}(\alpha x)^x/(x-n)!_2F_0\{n-x,\lambda x,-\beta/\alpha x\}, x\geq n$.
15	$G_1^{\alpha_1,\beta,\lambda_1}(s)$	$G_1^{\alpha_2,\beta,\lambda_2}(s)$	$[1-\{\alpha_1+\lambda_1\beta/(1-\beta)\}](1-\beta)^{\lambda_1 x+\lambda_2}e^{-(\alpha_1 x+\alpha_2)}(\alpha_1 x+\alpha_2)^x/x!$ ${}_2F_0(-x,\lambda_1 x+\lambda_2,-\beta/(\alpha_1 x+\alpha_2))$, $x\geq 1$

$G_1^{\alpha,\beta,\lambda}(s)$, $G_2^{\gamma}(s)$, $G_3^{\beta,N}(s)$, $G_4^{\beta}(s)$ and $G_5^{n}(s)$ denote the probability generating function of three parameter Charlier, Poisson, negative binomial, logarithmic series and delta distribution respectively. ${}_2F_0(a,b;x)$ denotes the Hypergeometric function.

6.3 Basic Lagrangian Negative binomial (LNB) Distribution

Taking $g(s) = (1-\beta)^N (1-\beta s)^{-N}$, the pgf of negative binomial distribution in (6.2.5), the pmf of basic Lagrangian negative binomial (LNB) distribution may be written as

$$P(X = x) = \frac{(Nx)_r}{x!}\beta^r (1-\beta)^{Nx}, \ for \quad x = 1,2,3... \quad and \quad r = x-1 \tag{6.3.1}$$

where $(Nx)_r = Nx(Nx+1)(Nx+2)...(Nx+r-1)$, and $(Nx)_0 = 1$.

a) Cumulants of basic LNB Distribution

The first four cumulants of negative binomial distribution are respectively

$$G_1 = \frac{N\beta}{(1-\beta)}$$

$$G_2 = \frac{N\beta}{(1-\beta)^2}$$

$$G_3 = (1+\beta)\frac{N\beta}{(1-\beta)^3}$$

$$G_4 = \{(1+\beta)^2 + 2\beta\}\frac{N\beta}{(1-\beta)^4}. \tag{6.3.2}$$

Then, using Consul and Shenton's (1975) general formula for cumulants given in (6.2.6) the first four cumulants of basic LNB distribution are respectively

$$D_1 = \frac{(1-\beta)}{(1-\beta)-N\beta}, \tag{6.3.3}$$

$$D_2 = \frac{N\beta(1-\beta)}{(1-\beta-N\beta)^3}, \tag{6.3.4}$$

$$D_3 = (1-\beta)(1+\beta)\frac{N\beta}{(1-\beta-N\beta)^4} + \frac{3(N\beta)^2(1-\beta)}{(1-\beta-N\beta)^5}, \tag{6.3.5}$$

$$D_4 = \{(1+\beta)^2 + 2\beta\}\frac{N\beta(1-\beta)}{(1-\beta-N\beta)^5} + \frac{10(N\beta)^2(1-\beta^2)}{(1-\beta-N\beta)^6} + \frac{(N\beta)^2(1-\beta)(1+\beta)^2}{(1-\beta-N\beta)^7}. \tag{6.3.6}$$

From the cumulants the moments of the distribution may be easily obtained.

Thus the indices of skewness and kurtosis are respectively

$$\beta_1 = \frac{A + 2N\beta B + N^2\beta^2 C}{N\beta K}, \tag{6.3.7}$$

and $$\beta_2 = 3 + \frac{A_1 - NB_1 - N^2C_1}{N\beta K}. \tag{6.3.8}$$

where $K = (1-\beta)(1-\beta-N\beta)$,

$A = 1 - 2\beta^2 + \beta^4$,

$B = (2 - \beta - 2\beta^2 + \beta^3)$,

$C = (4 - 4\beta + \beta^2)$,

$A_1 = 1 + 2\beta - 6\beta^2 + 2\beta^3 + \beta^4$,

$B_1 = 9\beta - 4\beta^2 - 3\beta^3 + 2\beta^4$,

and $C_1 = 9\beta^2 + 6\beta^3 - \beta^4$.

b) Estimation of parameters

In order to estimate the parameters of basic LNB distribution the following methods may be used.

i) Method of moment

The sample mean ($\bar{x}$) and sample variance (m_2) of the distribution are respectively

$$\bar{x} = \frac{(1-\beta)}{(1-\beta) - N\beta}, \tag{6.3.9}$$

and $$m_2 = \frac{N\beta(1-\beta)}{(1-\beta-N\beta)^3}, \tag{6.3.10}$$

By eliminating N from (6.3.9) and (6.3.10) we get the estimated value of β as

$$\beta^{\wedge} = \frac{m_2 - (\bar{x}-1)\bar{x}^2}{m_2}, \tag{6.3.11}$$

By substituting this value of β in equation (6.3.9) the estimated value of N may be obtained as

$$N^{\wedge} = \frac{(\bar{x}-1)(1-\beta^{\wedge})}{\bar{x}\beta^{\wedge}}. \tag{6.3.12}$$

ii) Ratio of first two frequencies and mean

If f_1 and f_2 denote respectively the first two sample frequencies of basic LNB distribution then their ratio may be written as

$$\theta = \frac{f_2}{f_1} = N\beta P_1, \tag{6.3.13}$$

where $P_1 = \frac{f_1}{f}$ and f is the total frequency.

Eliminating N from (6.3.9) and (6.3.13) we get

$$\beta^{\wedge} = \frac{(\bar{x}-1)P_1 - \bar{x}\theta}{(\bar{x}-1)P_1}, \tag{6.3.14}$$

Substituting this value of β in equation (6.3.9) the estimated value of N may be obtained as

$$N^{\wedge} = \frac{(\bar{x}-1)(1-\beta^{\wedge})}{\bar{x}\beta^{\wedge}}. \tag{6.3.15}$$

6.4 Basic Lagrangian Poisson (LP) Distribution

Taking $g(s) = e^{\gamma(t-1)}$, the pgf of Poisson distribution in (6.2.2) the pmf of basic Lagrangian Poisson (LP) distribution may be derived as

$$P(X=x) = \frac{e^{-\gamma x}}{x!}(\gamma x)^{x-1}, \; for \quad x = 1,2,3... \tag{6.4.1}$$

a) Cumulants of basic LP Distribution

The first four cumulants of Poisson distribution are respectively

$$G_1 = G_2 = G_3 = G_4 = \lambda \tag{6.4.2}$$

Then, using Consul and Shenton's (1975) general formula given in (6.2.6) the first four cumulants of basic LPD distribution are respectively

$$D_1 = \frac{1}{(1-\gamma)}, \tag{6.4.3}$$

$$D_2 = \frac{\gamma}{(1-\gamma)^3}, \tag{6.4.4}$$

$$D_3 = \frac{\gamma}{(1-\gamma)^4} + \frac{3\gamma^2}{(1-\gamma)^5}, \tag{6.4.5}$$

$$D_3 = \frac{\gamma}{(1-\gamma)^5} + \frac{10\gamma^2}{(1-\gamma)^6} + \frac{15\gamma^2}{(1-\gamma)7}. \tag{6.4.6}$$

From these cumulants we can easily obtain the moments of the distribution.

The indices of skewness and kurtosis are respectively

$$\beta_1 = \frac{(2\gamma+1)^2}{\gamma(1-\gamma)}, \tag{6.4.7}$$

$$\beta_2 = 3 + \frac{(1-23\gamma-9\gamma^2)}{\gamma(1-\gamma)}. \tag{6.4.8}$$

b) Estimation of parameters

In order to estimate the single parameter λ of basic LPD the following methods may be used.

i) Method of maximum likelihood

The likelihood function of LPD is

$$L = \prod_{i=0}^{k} \frac{1}{x_i!} e^{-\lambda x_i} (\lambda x_i)^{x_i - 1}, \tag{6.4.9}$$

Differentiating partially the log likelihood function of (6.4.9) and then solving for λ we get the estimated value of λ as

$$\hat{\lambda} = \frac{\bar{x}-1}{\bar{x}} \tag{6.4.10}$$

ii) Method of using first sample frequency

Equating the first probability of basic LPD to $\frac{f_1}{f}$, where f_1 and f represent respectively the first and the total frequencies we may have

$$\hat{\lambda} = -\log\frac{f_1}{f}. \tag{6.4.11}$$

6.5 Application

For the application of both basic LNB and LP distributions, we have considered the well known data collected by Williams (1944) about the numbers of papers published by authors in a certain Journal in the year 1935, for which geometric distribution was fitted by Williams (1944) and generalized geometric distribution was fitted by Plunkett and Jain (1975). The expected frequencies are given in Table **6.3** with that of geometric (Williams, 1944) and generalized geometric (Plunkett and Jain, 1975) distributions and it is seen that basic LNB distribution provides a good fit.

From Table **6.3**, we have the sample mean $\bar{x} = 1.4610$ and sample variance $s^2 = 0.812987$. In case of basic LNB distribution using the method of moment we have $\beta^{\wedge} = -0.2105$ and $N^{\wedge} = -1.8145$. Again by using the ratio of first two frequencies and mean we have $\beta^{\wedge} = -0.1715$ and $N^{\wedge} = -2.1553$. Comparing the χ^2 values obtained from LNB distribution (using both methods of estimation) with that of other distributions compared, it has been observed that LNB distribution gives better fit than the earlier ones.

It will be seen from the observed data that first three frequencies are large in comparison to the remaining ones so the method of using the first sample moment and ratio of first two frequencies gives better fit than method of moment.

In order to estimate the parameters of basic LP distribution, both maximum likelihood (ML) method and the method of using the first sample frequency (FF) both have been used. In this case, we get $\lambda^{\wedge} = 0.3677$ by using the first frequency and $\lambda^{\wedge} = 0.3552$, by the method of maximum likelihood. It is seen that LP distribution is better fitted than geometric distribution by Williams (1944). For the application of basic LP distribution, we have considered another set of data in Table **6.4**, collected by Williams (1944) about the numbers of papers published by authors in a certain Journal in the year 1936. The basic LP distribution is compared with geometric distribution (Williams, 1944) and generalized Logarithmic series distribution (Jain, 1975). Expected frequencies for basic LP distribution are given in Table **6.4** with that of generalized logarithmic series distribution by Jain (1975).

It is apparent from all these discussions that the fitting of basic LP distribution is not satisfactory compared to generalized geometric distribution by Plunkett and Jain (1975) in Table **6.3** and generalized Logarithmic series distribution by Jain (1975) in Table **6.4**.

Table 6.3

Comparison of observed frequencies of Publications of research papers in review of applied Mycology 1935 (Plunkett and Jain, 1975) with the Expected Lagrangian Poisson distribution (LPD) and Lagrangian negative binomial distribution (LNBD)

No. of papers published	No. of Authors $\bar{x} = 1.4610$	Fitted Distributions				Generalized Geometric dist. Plunkett and Jain (1975)
		LNBD $\hat{\beta} = -0.2105$ $\hat{N} = -1.8145$ (MM)	LNBD $\hat{\beta} = -0.1715$ $\hat{N} = -2.1553$ (FM)	LPD $\hat{\lambda} = 0.3417$ (FF)	LPD $\hat{\lambda} = 0.3156$ (ML)	
1	1085	1079.66	1085.58	1085.03	1113.72	1079.78
2	285	291.60	285.31	263.44	256.36	291.47
3	96	96.43	95.08	95.95	88.52	96.40
4	31	35.43	35.68	41.41	36.22	35.43
5	21	13.89	14.38	19.64	16.28	13.90
6	5	5.69	6.08	9.89	7.77	5.70
7	3 }	2.41 }	2.66 }	5.19	3.87	2.41
8	1 }	1.89 }	2.23 }	6.45	4.26	1.91
Total	1527	1527.00	1527.00	1527.00	1527.00	1527.00
	χ^2	4.4757	4.0248	12.4	10.3708	4.4549
	$d.f.$	4	4	6	6	4
	$p-value$	> 0.34	> 0.40	> 0.05	> 0.10	> 0.34

ML: method of maximum likelihood
FF: method of first frequency
MM: method of moment
FM: method based on first frequency and mean

Table 6.4

Comparison of observed frequencies of Publications of research papers in the review of applied Mycology 1936 (Jain, 1975) with the Expected Lagrangian Poisson distribution (LPD)

No. of papers per author	Observed frequency $\bar{x} = 1.5509$	Fitted Distribution LPD $\hat{\lambda} = 0.3677$ (FF)	Fitted Distribution LPD $\hat{\lambda} = 0.3552$ (ML)	Generalized logarithmic Series Jain (1975)
1	1062	1062.03	1075.39	1052.72
2	263	270.36	267.78	287.52
3	120	103.24	100.02	107.10
4	50	46.72	44.28	45.10
5	22	23.23	21.53	20.83
6	7	12.26	11.12	10.00
7	6	6.75	5.98	4.97
8	2	3.83	3.32	2.53
9	0	2.22	1.89	1.31
10	1	1.31	1.09	0.70
11	1	2.05	1.60	1.81
Total	1534	1534.00	1534.00	1534.00
	χ^2	8.67	8.44	5.14
	$d.f.$	6	6	4
	$p-value$	> 0.19	> 0.20	> 0.27

ML: maximum likelihood
FF: first frequency

6.6 General Lagrangian Poisson negative binomial (LPNB) Distribution

Taking $g(s)=\exp\{\gamma(s-1)\}$ and $f(s)=(1-\beta)^N(1-\beta s)^{-N}$ in equation (6.2.1), the pmf of Lagrangian Poisson negative binomial (LPNB) distribution of $type-I$ may be written as

$$P(X=x)=\frac{N\beta e^{-\gamma x}(1-\beta)^N}{x!}\sum_{j=0}^{r}\binom{r}{j}\beta^j(\gamma x)^{r-j}(N+1)_j\,,\quad for\ \ x\geq 1 \tag{6.6.1}$$

where $P_0=(1-\beta)^N$, $r=x-1\ and\ (N+1)_j=(N+1)(N+2)...(N+j)$.

Similarly, considering the equation (6.2.2), the pmf of Lagrangian Poisson negative binomial distributions of $type-II$ may be written as

$$P(X=x)=\frac{(1-\gamma)(1-\beta)^N}{x!}e^{-\gamma x}\sum_{j=0}^{r}\binom{r}{j}(N)_j(\gamma x)^{r-j}\beta^j\,,for\ \ x\geq 0 \tag{6.6.2}$$

where $r=x\ and\ (N)_j=N(N+1)(N+2)...(N+j-1)$

a) Cumulants of LPNB Distribution of Type-I

The first four cumulants of negative binomial distribution are respectively

$$F_1=\frac{N\beta}{(1-\beta)},$$

$$F_2=\frac{N\beta}{(1-\beta)^2},$$

$$F_3=(1+\beta)\frac{N\beta}{(1-\beta)^3},$$

and

$$F_4=\{(1+\beta)^2+2\beta\}\frac{N\beta}{(1-\beta)^4}. \tag{6.6.3}$$

Then using Consul and Shenton (1975) formula (6.2.4), the first three cumulants of general LPNB distribution may be written as

$$K_1=\frac{N\beta}{(1-\beta)(1-\gamma)}, \tag{6.6.4}$$

$$K_2=\frac{N\beta\gamma}{(1-\beta)(1-\gamma)^3}+\frac{N\beta}{(1-\beta)^2(1-\gamma)^2}, \tag{6.6.5}$$

$$K_3 = \frac{N\beta\gamma}{(1-\beta)(1-\gamma)^4}\left\{1+\frac{3\gamma}{(1-\gamma)}+\frac{3}{(1-\beta)}\right\}+(1+\beta)\frac{N\beta}{(1-\beta)^3(1-\gamma)^3}. \tag{6.6.6}$$

Hence mean and variance of LPNB distribution of type-I are respectively

$$\mu = \frac{\beta N}{(1-\beta)(1-\gamma)}, \tag{6.6.7}$$

and
$$\sigma^2 = \frac{\mu^2(1-\gamma\beta)}{(1-\gamma)N\beta}. \tag{6.6.8}$$

b) Estimation of parameters

The parameters of LPNB distribution of $type-I$ may be estimated by using the following methods.

i) First two sample moments and the ratio of first two frequencies

The sample mean $\bar{x}$ and sample variance s^2 of the distribution are respectively

$$\bar{x} = \frac{\beta N}{(1-\beta)(1-\gamma)}, \tag{6.6.11}$$

$$s^2 = \frac{\bar{x}\gamma}{(1-\gamma)^2}+\frac{\bar{x}^2}{N\beta}. \tag{6.6.12}$$

The ratio of first two frequencies may be written as

$$\theta = \frac{f_1}{f_0} = \frac{P_1}{P_0} = N\beta e^{-\gamma}, \tag{6.6.13}$$

Eliminating N and β one by one from (6.6.11), (6.6.12) and (6.6.13) we get a transcendental equation in γ as

$$f(\gamma) = e^{-\gamma}(1-\gamma)^2+\frac{\theta}{\bar{x}}\gamma-\frac{\theta s^2}{\bar{x}^2}(1-\gamma)^2, \tag{6.6.14}$$

This equation can be easily solved by using the Newton Raphson method. Putting the estimated value of γ, from equations (6.6.11) and (6.6.13) we get

$$\beta^{\wedge} = 1-\frac{e^{\gamma}\theta}{(1-\gamma)\bar{x}}, \tag{6.6.15}$$

By substituting the estimated values of β and γ in equation (6.6.11) we get

$$N^{\wedge} = \frac{\bar{x}(1-\hat{\gamma})(1-\beta^{\wedge})}{\beta^{\wedge}}. \tag{6.6.16}$$

ii) First three sample moments

The first three sample moments of the distribution are respectively

$$\bar{x} = \frac{\beta N}{(1-\beta)(1-\gamma)}, \tag{6.6.17}$$

$$m_2 = \frac{\bar{x}\gamma}{(1-\gamma)^2} + \frac{\bar{x}^2}{N\beta}. \tag{6.6.18}$$

$$m_3 = \frac{\bar{x}\gamma}{(1-\gamma)^3} - \frac{3\gamma m_2}{(1-\gamma)^2} + (1+\beta)\frac{\bar{x}m_2}{N\beta} - (1+\beta)\frac{\bar{x}^2\gamma}{N\beta 1-\gamma)^2}. \tag{6.6.19}$$

Eliminating N from the first pair we get

$$\frac{(1-\beta)}{\bar{x}(1-\gamma)} = \frac{1}{m_2(1-\gamma)^2 - \bar{x}\gamma} \tag{6.6.20}$$

Again eliminating N from the second pair we get

$$\frac{(1-\beta)}{\bar{x}(1-\gamma)} = \frac{m_3(1-\gamma)^3 - \gamma\bar{x} - 3\gamma m_2(1-\gamma)}{\{m_2(1-\gamma)^2 - \bar{x}\gamma\}^2} \tag{6.6.21}$$

Adding (6.6.20) and (6.6.21) we get a transcendental equation in γ as

$$f(\gamma) = A(1-\gamma)^4 + B(1-\gamma)^2 + C\gamma = 0 \tag{6.6.22}$$

where $A = \frac{2m_2^2}{\bar{x}} - m_3$, $B = -m_2$, $C = 2\bar{x}$.

The equation (6.6.22) can be solved by using the Newton Raphson method. Substituting the estimated value of γ in (6.6.20), we may estimate β as

$$\hat{\beta} = 1 - \left\{\frac{\bar{x}(1-\gamma)}{m_2(1-\gamma)^2 - \bar{x}\gamma}\right\}, \tag{6.6.23}$$

Putting the estimated values of γ and β in (6.6.17), the value of N may be estimated as

$$\hat{N} = \frac{\bar{x}(1-\beta)(1-\gamma)}{\beta}, \tag{6.6.24}$$

6.7 General Lagrangian Poisson Logarithmic (LPL) Distributions

Considering $g(s) = \exp\{\gamma(s-1)\}$ and $f(s) = \frac{\log(1-\beta t)}{\log(1-\beta)}$, the pgfs of Poisson and Logarithmic distributions respectively in equation (6.2.1), the pmf of Lagrangian Poisson Logarithmic (LPL) distributions of $type-I$ may be written as

$$P(X=x)=\frac{a\beta e^{-\gamma x}}{x!}\sum_{j=0}^{r}\frac{r!}{j!}\beta^{r-j}(\gamma x)^{j},\ \ for\ \ x\geq 1\ \ and\ r=x-1. \tag{6.7.1}$$

where $P_1 = ae^{-\gamma}\beta$ and $a=-\dfrac{1}{\log(1-\beta)}$.

Similarly, taking $g(s)=\exp\{\gamma(s-1)\}$ and $f(s)=\dfrac{\log(1-\beta t)}{\log(1-\beta)}$, in equations (6.2.2) the pmf of Lagrangian Poisson Logarithmic (LPL) distributions of $type-II$ may be written as

$$P(X=x)=a(1-\gamma)e^{-\gamma x}\sum_{j=0}^{r}\frac{(\gamma x)^{j}\beta^{x-j}}{(x-j)},for\ \ x\geq 1\ and\ \ r=x-1 \tag{6.7.2}$$

a) Cumulants of LPL Distribution of Type-I

The first four cumulants of Logarithmic distribution are respectively

$$F_1=\frac{a\beta}{(1-\beta)},$$

$$F_2=\frac{a\beta(1-a\beta)}{(1-\beta)^2},$$

$$F_3=a\beta\frac{(1+\beta-3a\beta+2a^2\beta^2)}{(1-\beta)^3},$$

$$F_4=\{(1+4\beta+\beta^2-4a\beta(1+\beta)+6a^2\beta^2-3a^3\beta^3\}\frac{a\beta}{(1-\beta)^4}. \tag{6.7.3}$$

According to Consul and Shenton (1975) formula, the first three cumulants of general Lagrangian Poisson Logarithmic distribution may be written as

$$K_1=\frac{a\beta}{(1-\beta)(1-\gamma)}, \tag{6.7.4}$$

$$K_2=\frac{a\beta\gamma}{(1-\beta)(1-\gamma)^3}+\frac{a\beta(1-a\beta)}{(1-\beta)^2(1-\gamma)^2}, \tag{6.7.5}$$

$$K_3=\frac{a\beta\gamma}{(1-\beta)(1-\gamma)^4}\left\{1+\frac{3\gamma}{(1-\gamma)}\right\}+\frac{3a\beta\gamma(1-a\beta)}{(1-\beta)^2(1-\gamma)^4}+$$
$$\frac{a\beta(1+\beta-3a\beta+2a^2\beta^2)}{(1-\beta)^3(1-\gamma)^3} \tag{6.7.6}$$

Hence mean and variance of the derived distribution are respectively

$$\mu = \frac{a\beta}{(1-\beta)(1-\gamma)}, \qquad (6.7.7)$$

$$\sigma^2 = \frac{\mu\gamma}{(1-\gamma)^2} + \frac{\mu}{(1-\beta)(1-\gamma)} - \mu^2. \qquad (6.7.8)$$

b) Estimation of parameters

The parameters of LPL distribution of $type-I$ may be estimated by using the following two methods.

i) First two sample moments

The first two sample moments of the distribution are respectively

$$\bar{x} = \frac{a\beta}{(1-\beta)(1-\gamma)}, \qquad (6.7.9)$$

$$m_2 = \frac{\bar{x}\gamma}{(1-\gamma)^2} + \frac{\bar{x}}{(1-\gamma)(1-\beta)} - \bar{x}^2. \qquad (6.7.10)$$

where $$a = -\frac{1}{\log(1-\beta)},$$

Eliminating γ from (6.7.9) and (6.7.10), and writing $k = \frac{m_2 + \bar{x}^2}{\bar{x}^2}$, we have an equation in β of the form

$$f(\beta) = \bar{x}(1-\beta)^2 + a\beta^2(1-ak), \qquad (6.7.11)$$

This equation can be solved by using the Newton-Raphson method. After getting β, the estimated value of γ may be obtained by using (6.7.9) as

$$\gamma = 1 - \frac{a\hat{\beta}}{\bar{x}(1-\hat{\beta})} \qquad (6.7.12)$$

ii) First sample moment and first frequency

Equating the first sample frequency of LPLD with f_1, we have

$$\frac{f_1}{f} = e^{-\gamma}a\beta, \qquad (6.7.13)$$

where f denotes the total frequency.

Taking logarithm on both sides of (6.7.13), we get

$$\log\frac{f_1}{f} = -\gamma + \log a\beta \qquad (6.7.14)$$

Eliminating γ from (6.7.9) and (6.7.14), we get a transcendental equation in β as

$$f(\beta) = \log(a\beta) + \frac{a\beta}{(1-\beta)\bar{x}} - \left\{1 + \log\frac{f_1}{f}\right\} \qquad (6.7.15)$$

Solving this equation by using the Newton-Raphson method, the estimated value of β may be obtained. By substituting this value of β in (6.7.9) we get the estimated value of γ as

$$\gamma = 1 - \frac{a\hat{\beta}}{\bar{x}(1-\hat{\beta})} \qquad (6.7.16)$$

6.8 General Lagrangian negative binomial Poisson (LNBP) Distributions

Taking $g(s) = (1-\beta)^N(1-\beta s)^{-N}$ and $f(s) = \exp\{\gamma(s-1)\}$, in (6.2.1), the pmf of general Lagrangian negative binomial Poisson (LNBP) distributions of *type – I* may be obtained as

$$P(X = x) = \frac{e^{-\gamma}(1-\beta)^{Nx}}{x!}\sum_{j=0}^{r}\binom{r}{j}\beta^j\gamma^{r-j+1}(Nx)_j\,, \qquad (6.8.1)$$

$$for \quad x = 1,2,3..., \; r = x-1 \; and \; (Nx)_j = (Nx)(Nx+1)...(Nx+j-1)$$

where $\qquad P_0 = e^{-\gamma} \qquad (6.8.2)$

Again, considering $g(s) = (1-\beta)^N(1-\beta s)^{-N}$ and $f(s) = \exp\{\gamma(s-1)\}$ in (6.2.2) the pmfs of general Lagrangian negative binomial Poisson distributions of *type – II* may be written as

$$P(X = x) = \frac{(1-\beta-N\beta)}{x!}e^{-\gamma}(1-\beta)^{Nx-1}\sum_{j=0}^{r}\binom{r}{j}(Nx)_j\gamma^{r-j}\beta^j\,, \qquad (6.8.3)$$

$$for \; x \ge 0 \; and \; r = x$$

a) Cumulants of LNBP Distribution of Type-I

The first four cumulants of Poisson distribution are $F_1 = F_2 = F_3 = F_4 = \gamma$. Therefore according to Consul and Shenton (1975) formula (6.2.4), the first three cumulants of general Lagrangian negative binomial Poisson distribution of $type-I$ may be written as

$$K_1 = \frac{(1-\beta)\gamma}{1-\beta-N\beta}, \tag{6.8.4}$$

$$K_2 = \frac{N\beta(1-\beta)\gamma}{(1-\beta-N\beta)} + \frac{(1-\beta)^2\gamma}{(1-\beta-N\beta)^2}, \tag{6.8.5}$$

$$K_3 = \frac{N\beta(1-\beta^2)\gamma}{(1-\beta-N\beta)^4} + \frac{3(N\beta)^2(1-\beta)\gamma}{(1-\beta-N\beta)^5} + \frac{3(N\beta)(1-\beta)^2\gamma}{(1-\beta-N\beta)^4} + \frac{(1-\beta)^3\gamma}{(1-\beta-N\beta)^3}. \tag{6.8.6}$$

Hence mean and variance of the LNBPD are respectively

$$\mu = \frac{(1-\beta)\gamma}{1-\beta-N\beta}, \tag{6.8.7}$$

$$\sigma^2 = \frac{N\beta(1-\beta)\gamma}{(1-\beta-N\beta)} + \frac{(1-\beta)^2\gamma}{(1-\beta-N\beta)^2}. \tag{6.8.8}$$

b) Estimation of parameters

The parameters of LNBP distribution of $type-I$ may be estimated by using the following two methods.

i) First two sample moments and first frequency

The sample mean and sample variance of the distribution are respectively

$$\bar{x} = \frac{(1-\beta)\gamma}{1-\beta-N\beta}, \tag{6.8.9}$$

$$s^2 = N\bar{x}\beta + \frac{(1-\beta)\bar{x}}{(1-\beta-N\beta)}. \tag{6.8.10}$$

By equating the first probability with $\frac{f_0}{f}$, using equation (6.8.2), we obtain the estimated value of γ as

$$\hat{\gamma} = -\log\frac{f_0}{f}, \tag{6.8.11}$$

where f_0 and f represent respectively the first and the total frequencies of the distribution. Eliminating N from the equations (6.8.9) and (6.8.10) we obtain the estimate of β as

$$\beta^{\wedge} = 1 - \frac{\hat{\gamma}s^2 - \bar{x}^2}{(\bar{x} - \hat{\gamma})\hat{\gamma}}, \tag{6.8.12}$$

Substituting this value of $\hat{\beta}$ in equation (6.8.9) we may estimate

$$N^{\wedge} = \frac{(\bar{x} - \gamma)(1 - \beta^{\wedge})}{\bar{x}\beta^{\wedge}}. \tag{6.8.13}$$

ii) First two frequencies and mean

By equating the first two probabilities of LNBP distribution with $\frac{f_0}{f}$ and $\frac{f_1}{f}$ respectively we may write

$$\frac{f_0}{f} = e^{-\gamma}, \tag{6.8.14}$$

$$\frac{f_1}{f} = \gamma(1 - \beta)^N e^{-\gamma}, \tag{6.8.15}$$

so that
$$\theta = \frac{f_1}{f_0}\gamma(1 - \beta)^N \tag{6.8.16}$$

where f_0, f_1 and f, represent respectively the first two sample frequencies and the total frequency of LNBP distribution. Taking log on both sides of (6.8.14), we may obtained

$$\hat{\gamma} = -\log\frac{f_0}{f}, \tag{6.8.17}$$

Eliminating N from the sample mean (6.8.9), and from the equation obtained by taking log on both sides of (6.8.16), we may have a transcendental equation on β as

$$f(\beta) = (1 - \beta)\log(1 - \beta) - A\beta, \tag{6.8.18}$$

where
$$A = (\bar{x} - \gamma^{\wedge})^{-1}\{\log\theta - \log\hat{\gamma}\}\bar{x}.$$

Using the Newton-Raphson Method the equation (6.8.18) can be solved for β. Putting the estimated values of γ and β, the third parameter N may be obtained by either by using (6.8.16) or (6.8.9) as

$$\hat{N} = \frac{\log\theta - \log\gamma}{\log(1 - \beta)}. \tag{6.8.19}$$

or
$$\hat{N} = \frac{(1 - \beta)(\bar{x} - \gamma)}{\bar{x}\beta} \tag{6.8.20}$$

6.9 Lagrangian negative binomial Logarithmic (LNBL) Distribution

Taking $g(s) = (1-\beta)^N (1-\beta s)^{-N}$ and $f(s) = \dfrac{\log(1-\beta t)}{\log(1-\beta)}$ in equation (6.2.1) the pmf of Lagrangian negative binomial logarithmic (LNBL) distributions of *type – I* may be written as

$$P(X = x) = \frac{a(1-\beta)^{Nx}}{x!}\beta^{r+1}(Nx+1)_r, \, for \;\; x = 1,2,3\ldots, \;\; r = x-1 \tag{6.9.1}$$

where $a = -\dfrac{1}{\log(1-\beta)}$ and $(Nx+1)_j = (Nx+1)(Nx+2)\ldots(Nx+j)$

Similarly, using equation (6.2.2), the pmf of Lagrangian negative binomial logarithmic distributions of *type – II* may be written as

$$P(X = x) = \left\{1 - \frac{N\beta}{1-\beta}\right\} a(1-\beta)^{Nx}\beta^x \sum_{j=0}^{x-1}\frac{(Nx)_j}{j!(x-j)!}, \tag{6.9.2}$$

$$for \;\; x = 1,2,3\ldots, \;\; (Nx)_j = (Nx)(Nx+1)\ldots(Nx+j-1)$$

a) Cumulants of LNBL distribution of Type I

The first four cumulants of Logarithmic distribution are respectively

$$F_1 = \frac{a\beta}{(1-\beta)},$$

$$F_2 = \frac{a\beta(1-a\beta)}{(1-\beta)^2},$$

$$F_3 = a\beta\frac{(1+\beta-3a\beta+2a^2\beta^2)}{(1-\beta)^3},$$

$$F_4 = \{(1+4\beta+\beta^2-4a\beta(1+\beta)+6a^2\beta^2-3a^3\beta^3\}\frac{a\beta}{(1-\beta)^4}. \tag{6.9.3}$$

According to Consul and Shenton (1975) general formula, using equation (6.2.4) the first three cumulants of general LNBL distribution of Type-I may be written as

$$K_1 = \frac{a\beta}{(1-\beta-N\beta)}, \tag{6.9.4}$$

$$K_2 = \frac{a\beta^2 N}{(1-\beta-N\beta)^3} + \frac{a\beta(1-a\beta)}{(1-\beta-N\beta)^2}, \tag{6.9.5}$$

$$K_3 = \frac{(1+\beta)a\beta^2 N}{(1-\beta-N\beta)^4} + \frac{3a\beta(N\beta)^2}{(1-\beta-N\beta)^5} + \frac{3a\beta(N\beta)(1-a\beta)}{(1-\beta-N\beta)^4} +$$

$$\frac{a\beta(1+\beta-3a\beta+2a^2\beta^2)}{(1-\beta-N\beta)^3}, \tag{6.9.6}$$

Hence mean and variance of the LNBL distribution are respectively

$$\mu = \frac{a\beta}{(1-\beta-N\beta)}, \tag{6.9.7}$$

$$\sigma^2 = \frac{\mu\beta N}{(1-\beta-N\beta)^2} + \frac{(1-a\beta)\mu}{(1-\beta-N\beta)}. \tag{6.9.8}$$

b) Estimation of parameters

The parameters of LNBL distribution of $type-I$ may be estimated by using the methods discussed below.

a) First two sample moments

The first two sample moments of the distribution are respectively

$$\bar{x} = \frac{a\beta}{(1-\beta-N\beta)}, \tag{6.9.9}$$

$$m_2 = \frac{\bar{x}\beta N}{(1-\beta-N\beta)^2} + \frac{(1-a\beta)\bar{x}}{(1-\beta-N\beta)}. \tag{6.9.10}$$

where $$a = -\frac{1}{\log(1-\beta)},$$

Eliminating N from (6.9.9) and (6.9.10), and writing $k = \frac{m_2 + \bar{x}^2}{\bar{x}^3}$, we have an equation in β of the form

$$f(\beta) = (1-\beta) - a^2\beta^2 k, \tag{6.9.11}$$

This equation can be solved by using the Newton-Raphson method. After getting β, the value of N may be estimated by using (6.9.9) as

$$\hat{N} = \frac{\bar{x}(1-\hat{\beta}) - a\hat{\beta}}{\bar{x}\hat{\beta}} \tag{6.9.12}$$

b) First sample moment and ratio of first two frequencies

Equating the first two sample frequencies of LPL distribution of Type-I with f_1

and f_2 respectively we may write

$$\frac{f_1}{f} = a\beta(1-\beta)^N, \tag{6.9.13}$$

and
$$\frac{f_2}{f} = (1-\beta)^{2N}(2N+1)\frac{a\beta^2}{2}. \tag{6.9.14}$$

where f is the total frequency. From (6.9.13) and (6.9.14) we get

$$\frac{f_2 f}{f_1^2} = K = \frac{2N+1}{2a} \tag{6.9.15}$$

Eliminating N from sample mean (6.9.9) and (6.9.15), we get a transcendental equation in β as

$$f(\beta) = (2-\beta)\bar{x} - 2a\beta(1+\bar{x}K) \tag{6.9.16}$$

Solving this equation by Newton-Raphson method we may obtain the estimated value of β. Substituting the value of $\hat{\beta}$ in (6.9.9) we may get the estimated value of γ as

$$N = \frac{\bar{x}(1-\hat{\beta}) - a\hat{\beta}}{\bar{x}\hat{\beta}} \tag{6.9.17}$$

6.10 Applications

It is to be very important to note that, when the frequency for the zero class in the sample is larger than most of the other class frequencies or when the graph of the sample distribution is approximately L-shaped, one would like to give more weight to this larger frequency value of the zero class than to the statistic of sample variance which is more affected by the frequencies of the higher classes (Anscombe, F. W., 1950). It is for this reason the method of first two sample frequencies and sample mean is generally used. For conducting the empirical comparison of general Lagrangian distributions studied, we have considered some reported observed data for fitting of the distributions.

In the first Table **6.5**, we have considered Biological data used by Janardan et al. (1979) for Generalized Poisson Model (see Consul, 1989), who have considered Cole's (1946) classic sets of data on spiders and sow bugs. Here it can be shown that our models LNBP1 and LPNB1 fit the data very well.

In Table **6.6**, we have considered the observed frequencies for Home injuries (Consul, 1989) of 122 Experienced Men during 5 years (1937-1942) with the expected frequencies of LPNB1 distribution which is found to be satisfactory.

Table 6.5

Comparison of observed frequencies of distribution of 102 spiders under 240 Boards with the expected frequencies of LNBP and LPNB Distributions (Consul,1989)

No. of Spiders/Boards	Observed frequency	Fitted LNBP1	Fitted LPNB1	(Consul,1989) GPD
0	159	159.00	158.81	157.2
1	64	63.47	63.90	66.5
2	13	14.63	14.30	14.2
3	4	2.51	2.97	2.0
Total	240	240.00	240.00	240.0
Estimates		$\lambda^{\wedge} = 0.4117$ $\beta^{\wedge} = -0.0566$ $N^{\wedge} = -0.5831$	$\lambda^{\wedge} = 0.0692$ $\beta^{\wedge} = -0.0905$ $N^{\wedge} = -4.7660$	$\hat{\theta} = 0.4114$ $\hat{\lambda} = 0.0320$

Table 6.6

Comparison of observed frequencies for Home injuries of 122 Experienced Men during 5 years (1937-1942) with the Expected frequencies of general LPNB Distribution (Consul, 1989)

No. of injuries	Observed frequency	Fitted LPNB1	GPD (Consul, 1989)
0	73	72.98	73.23
1	36	35.99	35.32
2	10	9.95	10.41
3	2	1.39	3.04
4	1	1.69	
Total	122	122.00	122.00
Estimates		$\lambda^{\wedge} = -0.0924$ $\beta^{\wedge} = 0.2391$ $N^{\wedge} = 1.8807$	$\hat{\theta} = 0.51$ $\hat{\lambda} = 0.06$

LNBP: Lagrangian negative binomial Poisson
LPNB: Lagrangian Poisson negative binomial
GPD: Generalized Poisson distribution

The LPNB1 distribution is also fitted in Table **6.7** for comparing the observed frequencies for first year Shunting Accidents for different age groups (Consul, 1989). Observing the expected frequencies obtained from LPNB1 distribution it can be conclude that the distribution is highly satisfactory in this case.

Table 6.7

Comparison of observed frequencies for First year Shunting Accidents with the Expected general LPNB Distribution for different age groups (Consul, 1989)

No. of Accidents	Age 26-30			Age 31-35		
	Observed frequency	Fitted LPNB1	GPD (Consul, 1989)	Observed frequency	Fitted LPNB1	GPD (Consul, 1989)
0	121	121.18	126.42	80	79.99	80.23
1	85	85.12	74.49	61	60.99	60.41
2	19	17.50	21.45	13	13.06	13.48
3	1	0.89	4.02	1	0.96	0.88
≥ 4	1	2.31	0.62	0	0.00	
Total	227	227.00	227.00	155	155.00	155.00
Estimates		$\lambda^{\wedge} = -0.2022$ $\beta^{\wedge} = 0.1665$ $N^{\wedge} = 3.4474$	$\hat{\theta} = 0.585$ $\hat{\lambda} = -0.007$		$\lambda^{\wedge} = -0.1604$ $\beta^{\wedge} = 0.361$ $N^{\wedge} = 18.0234$	$\hat{\theta} = 0.658$ $\hat{\lambda} = -0.134$

LPNB: Lagrangian Poisson negative binomial
GPD: Generalized Poisson distribution

In the forth Table **6.8**, we have considered the well-known data on natural laws in social sciences by Kendall (1961) who considered the observed data on the number of outbreaks of strike in three leading industries in United Kingdom during 1948-1959. Generalized LNBP1 model has been fitted to these data using the sample mean, sample variance and first frequency and it is clear from the observed and the expected frequencies that the data fit our model well.

In Table **6.9**, we have considered the Genetics data (Consul, 1989) on the variations in the exposure, D and R cells, it is seen that as the exposure to radiation increased, the value of the parameter $\lambda^\wedge$ increases in our model. When D and R cells are decreased from 20/400 to 36/200, the value of $\beta^\wedge$ increases and the value of $N^\wedge$ increased or decreased with the quantity of D cells is increased or decreased. Comparing the observed and expected values this Table shows that our model LNBP1 is representing the effect of processes very well.

In this Chapter 6, in fitting Lagrangian negative binomial Poisson (LNBP1) and Lagrangian Poisson negative binomial (LPNB1) distributions for the data where the number of observations is 5 or less (e.g., Tables 6.5, 6.6, 6.7, 6.8 and 6.9) and we have more than two parameters to be estimated, the $\chi 2$ and corresponding $p-value$ are not provided as degrees of freedom (df) is very negligible.

For the application of general LPL1 distribution, we consider the well known data collected by Williams (1944) about the numbers of papers published by authors in a certain Journal for which Plunkett and Jain (1975) fitted generalized geometric distribution. Expected frequencies for general LPL1 distribution are given in Table **6.10** with that of geometric (Williams, 1944) and generalized geometric distribution (Plunkett and Jain, 1975) and it is seen that our distribution provides a good fit. By using the first sample moment and first sample frequency of the distribution we have $\beta^\wedge = 0.4873$ and $\gamma^\wedge = 0.0263$. Comparing the χ^2 values obtained from general LPL1 distribution with that of other distributions compared, it has been observed that our distribution gives better fit than the earlier ones.

In Table **6.11**, we consider the data on the number of papers published per author for whom geometric distribution and logarithmic series distribution were fitted by Williams (1944) and generalized logarithmic series distribution by Jain (1975). Here also the method of using the first sample mean and first frequency is used to estimate the parameters. In this case we have $\beta^\wedge = 0.6002$ and $\gamma^\wedge = -0.0559$. From the expected data obtained from LPL1 distribution and also observing the χ^2 values, it may be concluded that our distributions is found to be highly satisfactory.

Table 6.8

Comparison of observed frequencies of the number of outbreaks of strike in four leading industries in the U.K. during 1948-1959 with the expected frequencies of general LNBP1 distribution (Kendall, 1961)

No. of outbreaks	Vehicle manufacture			ship building			Transport		
	Observed frequency	Fitted LNBP1	GPD Consul (1989)	Observed frequency	Fitted LNBP1	GPD Consul, (1989)	Observed frequency	Fitted LNBP1	GPD Consul, (1989)
0	110	109.99	109.82	117	117.00	116.74	114	113.99	114.84
1	33	32.86	33.36	29	29.78	30.22	35	31.91	33.88
2	9	9.56	9.24	9	7.08	6.97	4	9.60	7.27
3	3	2.64	3.58	0	1.65	0.88	2	0.00	2.01
≥ 4	1	0.70		1	0.38		1	0.71	9.69
Total	156	156.00	156.00	156	156.00	156.00	156	156.00	156.00
Estimates		$\lambda^{\wedge} = 0.3494$ $\beta^{\wedge} = -0.1152$ $N^{\wedge} = -1.437$	$\hat{\theta} = 0.351$ $\hat{\lambda} = -0.144$		$\lambda^{\wedge} = 0.2877$ $\beta^{\wedge} = -0.0406$ $N^{\wedge} = -3.0742$	$\hat{\theta} = 0.29$ $\hat{\lambda} = -0.113$		$\lambda^{\wedge} = 0.3137$ $\beta^{\wedge} = -1.371$ $N^{\wedge} = -0.132$	$\hat{\theta} = 0.310$ $\hat{\lambda} = 0.098$

LNBP: Lagrangian negative binomial Poisson

GPD: Generalized Poisson distribution

Table 6.9

Comparison of observed frequencies of the number of human Cytogenic Dosimetry: Radiation from 241 Am. with the expected frequencies of LNBP distribution (Consul, 1989)

Exposure(rad)	D and R Cells	frequencies of cell with D and R 0	1	2	3	Total
0.85 (Observed Frequency)	20/400	385	11	3	1	400
Fitted LNBP1 $\lambda^{\wedge} = 0.03822$ $\beta^{\wedge} = -0.5849$ $N^{\wedge} = -0.7370$		385.00	10.48	3.36	0.90	400
GPD (Consul, 1989) $\hat{\theta} = 0.0383$ $\hat{\lambda} = 0.2334$		384.96	11.68	2.34	1.02	400
1.71 (Observed Frequency)	17/200	187	10	2	1	200
Fitted LNBP1 $\lambda^{\wedge} = 0.0672$ $\beta^{\wedge} = -0.1396$ $N^{\wedge} = -1.7092$		187.00	10.05	2.19	0.55	200
GPD (Consul, 1989) $\hat{\theta} = 0.0674$ $\hat{\lambda} = 0.2074$		186.97	10.24	2.00	0.79	200
3.42 (Observed Frequency)	36/200	176	15	6	3	200
Fitted LNBP1 $\lambda^{\wedge} = 0.1278$ $\beta^{\wedge} = 0.1538$ $N^{\wedge} = 1.5949$		176.00	17.24	4.08	1.45	200
GPD (Consul, 1989) $\hat{\theta} = 0.1297$ $\hat{\lambda} = 0.2797$		175.68	17.22	4.49	2.61	200

LNBP: Lagrangian negative binomial Poisson
GPD: Generalized Poisson Distribution

Table 6.10: Comparison of observed frequencies of the Publications of research papers in review of applied Mycology 1935 with the expected frequencies of LPL Distribution (Plunkett and Jain, 1975)

No. of papers published	No. of Authors $\bar{x} = 1.4610$	Fitted LPL1 $\beta^{\wedge} = 0.4873$ $\gamma^{\wedge} = 0.0263$.	Generalized Geometric (Plunkett and Jain, 1975)
1	1085	1085.00	1079.78
2	285	285.24	291.47
3	96	95.71	96.40
4	31	35.98	35.43
5	21	14.42	13.90
6	5	6.02	5.70
7	3	2.58	2.41
8	1	2.05	1.91
Total	1527	1527.00	1527.00
	χ^2	3.9514	4.4549
	$d.f.$	4	4
	$p-value$	> 0.41	> 0.34

Table 6.11: Comparison of observed frequencies of the Publications of research papers in review of applied Mycology 1936 with the expected frequencies of LPL Distribution (Jain, 1975)

No. of papers per author	Observed frequency $\bar{x} = 1.5509$	Fitted LPL1 $\beta^{\wedge} = 0.6002$ $\gamma^{\wedge} = -0.0559$	Generalized Logarithmic Series dist. Jain (1975)
1	1062	1062.01	1052.72
2	263	274.25	287.52
3	120	108.33	107.10
4	50	46.72	45.10
5	22	21.64	20.83
6	7	10.43	10.00
7	6	5.17	4.97
8	2	2.61	2.53
9	0	1.34	1.31
10	1	0.70	0.70
11	1	0.80	1.81
Total	1534	1534.00	1534.00
	χ^2	3.6019	5.14
	$d.f.$	5	4
	$p-value$	> 0.60	> 0.27

LPL: Lagrangian Poisson logarithmic

Chapter 7

Applications of Gegenbauer Distribution to Ball Games

7.1 Introduction

The Gegenbauer distribution was derived by Plunkett and Jain (1975), by mixing Hermite distribution with gamma distribution. The probabilities of Gegenbauer distribution can be expressed in terms of Gegenbauer polynomials (see Rainville, 1960). This distribution has a long history in the theory of stochastic processes. It has been observed in the past that the negative binomial distribution (NBD) produce close fit to the distributions of scores of teams and individuals in several sports involving ball games. In this present chapter, it is presupposed that Gegenbauer distribution would provide better fit than the negative binomial, Poisson or Hermite distribution, as the later distributions may be obtained as particular cases of the former.

Wood (1945) attempted to fit a probability distribution to individual batsmen's scores. Morney (1956) showed that the distribution of the goals of football matches gave good fit to the NBD. While analyzing the passing move distributions in association football matches, Reep and Bengamin (1968), have shown that the distribution give good fit to NBD. They have described the association football matches in details and have discussed the chance mechanisms in the game which lead to the NBD. Following the works done earlier, Reep et al. (1971) have shown that the NBD is also applicable to certain movements and performances in other ball games viz. crickets, ice-hockey, base ball and lawn tennis. They have also shown that the distributions of numbers of goals scored by individual football teams in individual matches provide good fit to the NBD. Polland (1973) has fitted the NBD successfully to the frequency distribution of group scores of individual teams in U.S. Collegiate football games.

In this chapter we have attempted to fit the Gegenbauer distribution to the runs scored in the completed innings in test matches by some famous batsmen at cricket, to the distribution of scores of individual teams in U.S. Collegiate football games, and also to the distributions of goals per match scored by individual teams in national Hockey League 1966-67. In most of the cases it is seen that Gegenbauer distribution gives the better fit.

The Gegenbauer distribution derived by Plunkett and Jain (1975) has the pgf of the form

$$g(t) = (1-\alpha-\beta)^{\lambda}(1-\alpha t-\beta t^2)^{-\lambda} \tag{7.1.1}$$

They fitted Gegenbauer distribution to a set of accident data using the method of moments. The limiting distributions of Gegenbauer distribution are negative binomial (when $\beta = 0$) and Hermite distribution (as $\lambda \to \infty$, $\alpha \to 0$ and $\beta \to 0$ such that $\lambda\alpha = \alpha_1$ and $\lambda\beta = \alpha_2$). Factorizing the quadratic expressions in equation (7.1.1), the pgf can be expressed as

$$G(t) = \left[\frac{(1-a)(1+b)}{(1-at)(1+bt)}\right]^{\lambda}, \quad 0 < b < a < 1 \tag{7.1.2}$$

where $a - b = \xi$, $ab = \eta$ and $0 < b < a < 1$.

This factorized form makes the distribution easy to handle, especially its moment properties.

7.2 Properties of Gegenbauer Distribution

a) Recurrence relation for probabilities

The Gegenbauer distribution of the form (7.1.2) has the recurrence relations for probabilities as

$$P_{r+1} = \frac{(a-b)(r+\lambda)P_r + ab(2\lambda+r-1)P_{r-1}}{(r+1)}, \quad r \geq 1 \tag{7.2.1}$$

This is obtained by differentiating the equation (7.1.2), with respect to't' and then equating the coefficient of t^r on both sides. The first two probabilities of the distribution are respectively

$$P_0 = (1-a)^{\lambda}(1+b)^{\lambda} \text{ and } P_1 = \lambda(a-b)P_0,$$

b) Recurrence relation for cumulants

The mgf of Gegenbauer distribution corresponding to the pgf (7.1.2) is of the following form

$$M(t) = \left[\frac{(1-a)(1+b)}{(1-ae^t)(1+be^t)}\right]^{\lambda}, \tag{7.2.2}$$

The cumulant generating function is obtained by taking log on both sides of (7.2.2), as

$$K(t) = \log M(t) = \lambda \log\{(1-a)(1+b)\} - \lambda \log(1-ae^t) - \lambda \log(1+be^t). \tag{7.2.3}$$

Its recurrence relation for cumulants may be written as

$$K_{r+1} = \frac{\lambda\{(a-b)+ab2^{r+1}\}}{(1-a)(1+b)} + \sum_{j=1}^{r}\{(a-b)+ab2^j\}^r C_j K_{r-j+1}, r \geq 1 \tag{7.2.4}$$

This is obtained by differentiating (7.2.3), w.r.t.'t' and then equating the coefficients of $t^r / r!$, on both sides. First two cumulants of the distribution are respectively

$$K_1 = \frac{\lambda a}{(1-a)} - \frac{\lambda b}{(1+b)}$$

and $$K_2 = \frac{\lambda a}{(1-a)^2} - \frac{\lambda b}{(1+b)^2},$$

From the cumulants the moments can be easily obtained. Hence the mean and variance of the distribution are respectively

$$\mu = \frac{\lambda(a-b)+2ab}{1-(a-b)-ab} \tag{7.2.5}$$

and $$\sigma^2 = \mu + \frac{2ab\lambda + (a-b)\mu + 2ab\mu}{1-(a-b)-ab}. \tag{7.2.6}$$

7.3 Estimation of parameters

It is well known that of all of the procedures of estimation, the method of moments is perhaps the oldest and the simplest and in many cases it leads to tractable operations whereas the method of maximum Likelihood is very cumbersome in comparison. Therefore, in order to estimate the parameters of the Gegenbauer distribution the following techniques may be used. They are by using

- The first three sample moments
- The first two sample moments and the ratio of the first two frequencies

a) The first three sample moments

The first three sample moments of the distribution may be written respectively as

$$\bar{x} = \frac{\lambda(a-b)+2ab}{1-(a-b)-ab}, \tag{7.3.1}$$

$$s^2 = \bar{x} + \frac{2ab\lambda + (a-b)\bar{x} + 2ab\bar{x}}{1-(a-b)-ab}, \tag{7.3.2}$$

$$m_3 = s^2 + \frac{2ab(\bar{x}+2\lambda)+(a-b)s^2+2abs^2}{1-(a-b)-ab}, \tag{7.3.3}$$

Eliminating $(a-b)$ and ab respectively from (7.3.1), (7.3.2) and (7.3.3) we may have a second degree equation in 'λ' as

$$A\lambda^2 + B\lambda + C = 0 \tag{7.3.4}$$

where $A = 3s^2 - m_3 - 2\bar{x}$,

$B = \bar{x}(2s^2 - 3\bar{x})$

and $C = -\bar{x}^3$.

The equation (7.3.4) may be solved for λ, either by using Newton Raphson method or solving as

$$\hat{\lambda} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}, \lambda > 0 \tag{7.3.5}$$

Putting the estimated value of λ in (7.3.1) and (7.3.2) and eliminating a, we may obtain another second degree equation in b as

$$Pb^2 + Qb + R = 0 \tag{7.3.6}$$

where $P = (\bar{x}^2 - \hat{\lambda}s^2) + \hat{\lambda}(3\bar{x} + 2\hat{\lambda})$,

$Q = 2(\bar{x}^2 - \hat{\lambda}s^2) + 4\hat{\lambda}\bar{x}$

and $R = \bar{x}^2 - (s^2 - \bar{x})\hat{\lambda}$.

Solving (7.3.6) for b, we get

$$\hat{b} = \frac{-Q \pm \sqrt{Q^2 - 4PR}}{2P}. \tag{7.3.7}$$

Putting the value of λ and b, the estimated value of a may be obtained either from (7.3.1) or (7.3.2), as

$$\hat{a} = \frac{(1+\hat{b})\bar{x} + \hat{\lambda}\hat{b}}{(1+2\hat{b})\hat{\lambda} + (1+\hat{b})\bar{x}}, \tag{7.3.8}$$

or
$$\hat{a} = \frac{(1+\hat{b})s^2 - \bar{x}}{\hat{b}(2\hat{\lambda} + \bar{x} + s^2) + s^2}. \tag{7.3.9}$$

b) The first two sample moments and the ratio of the first two frequencies

Taking the ratio of first two sample frequencies of the distribution as

$$\theta = \frac{f_1}{f_0} = \frac{P_1}{P_0} = (a-b)\lambda \tag{7.3.10}$$

Using θ, the equation (7.3.1) can be written as

$$\bar{x} = \frac{\theta + 2ab\lambda}{1-(a-b)-ab} \tag{7.3.11}$$

Eliminating $(a-b)$ and ab respectively from (7.3.10), (7.3.11) and (7.3.2) we may obtain a second degree equation in λ as

$$A\lambda^2 + B\lambda + C = 0 \tag{7.3.12}$$

where $A = 2(s^2 + \theta - 2\bar{x})$,

$B = -(2\bar{x}^2 - 3\bar{x}\theta + s^2\theta)$

and $C = \bar{x}^2\theta$.

The equation (7.3.12) can be solved either by using Newton Raphson method or the value of λ may be obtained as

$$\hat{\lambda} = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}, \tag{7.3.13}$$

Putting the estimated value of λ and eliminating a from (7.3.10) and (7.3.11), we get another second degree equation in b as

$$Pb^2 + Qb + R = 0 \tag{7.3.14}$$

where $P = \hat{\lambda}(2\hat{\lambda} + \bar{x})$,

$$Q = (2\hat{\lambda} + \bar{x})\theta$$

and $$R = \bar{x}\theta + (\theta - \bar{x})\hat{\lambda}.$$

Equation (7.3.14) may be solved for b either by using Newton-Raphson method or we get

$$\hat{b} = \frac{-Q \pm \sqrt{Q^2 - 4PR}}{2P}, \tag{7.3.15}$$

Putting the value of $\hat{\lambda}$ and $\hat{b}$ in (7.3.10), we get

$$\hat{a} = \frac{\theta}{\hat{\lambda}} + \hat{b}\ . \tag{7.3.16}$$

7.4 Applications

For the application of Gegenbauer distribution, we present some observed frequency distributions and the corresponding expected frequency distributions obtained from the fitted Gegenbauer distribution (GD) and negative binomial distribution (NBD). We give the values of sample mean ($\bar{x}$), sample variance (s^2) and also the calculated values of chi-square (χ^2) with degrees of freedom (d.f) for each distribution. For estimating the parameters (a, b, λ), the first two sample moments and the ratio of first two sample frequencies have been used.

In Table 7.1, we have tried to examine whether the distribution of catches taken in the completed innings at cricket tests provide good fit to the Gegenbauer distribution. For this purpose, the distribution of catches taken by Sobers (one of the best all rounders in the cricket world) is used. The expected as well as the observed frequencies of Gegenbauer distribution together with NBD are given in Table 7.1. Comparing the χ^2 values obtained from Gegenbauer distribution with that of the negative binomial distribution, it has been observed that our distribution gives better fit than the NBD.

Table 7.1

Applications of Gegenbauer distribution (GD) and negative binomial distribution (NBD) to the distribution of catches taken by Sobers in the completed innings at cricket tests (Sinha et al., 1987)

Catches	Observed Frequency $\bar{x} = 0.675$ $s^2 = 1.4444$	Expected frequency	
		NBD (Sinha et al., 1987)	GD $\hat{\lambda} = 1.0765$ $b = 0.2429$ $a = 0.4513$
0	107	102.1	105.97
1	24	32.0	23.77
2	14	13.6	17.67
3	9	6.3	6.51
4	2	3.0	3.39
5	3	1.5	1.45
6	1	1.5	1.26
Total	160	160.00	160.00
	χ^2	3.40	1.72
	$d.f.$	2	3
	$p-value$	> 0.10	> 0.63

GD: Gegenbauer distribution
NBD: negative binomial distribution

Reep *et al.* (1971) have fitted the NBD to the distribution of runs scored by Cowdrey and Barrington in the completed innings at cricket tests. Here we have attempt to fit the Geganbauer distribution to the data. The observed as well as the expected frequencies on the basis of the two distributions under consideration are given in Table 7.2. It is observed from Table 7.2 that the Gegenbauer distribution provides better fit than NBD for the distribution of the runs scored by Barrington than the runs scored by Cowdrey.

Table 7.2

Applications of negative binomial distribution (NBD) and Gegenbauer distribution (GD) to the runs scored in the completed innings at cricket test Reep *et al.* (1971)

Cowdrey				Barrington		
Runs (units of 20)	Observed frequency $\bar{x} = 1.6923$ $s^2 = 4.3156$	Expected frequency NBD (Sinha et al., 1987)	Expected frequency GD $\hat{\lambda} = 1.5049$ $b = 0.2238$ $a = 0.5664$	Observed frequency $\bar{x} = 2.0948$ $s^2 = 4.8962$	Expected frequency NBD (Sinha et al., 1987)	Expected frequency GD $\hat{\lambda} = 3.1555$ $b = 0.3258$ $a = 0.4763$
0	64	56.1	60.10	40	31.0	36.69
1	33	37.3	30.99	19	27.3	17.42
2	16	23.7	24.77	13	20.0	23.41
3	12	14.8	15.17	15	13.6	12.64
4	11	9.2	9.78	14	8.9	10.48
5	12	5.7	6.00	3	5.7	6.90
6	0	3.5	3.68	6	3.6	4.00
7	6	2.2	2.22	3	2.3	2.27
8	1	1.3	1.33	2	1.4	1.39
9	1	2.1	1.96	1	2.2	0.80
Total	156	155.9	156.00	116	116.0	116.00
	χ^2	5.12	5.98	χ^2	10.02	8.80
	$d.f.$	4	3	$d.f.$	4	3
	$p-value$	> 0.25	> 0.11	$p-value$	> 0.03	> 0.03

GD: Gegenbauer distribution
NBD: negative binomial distribution

In Table 7.3, we have attempted to fit the Gegenbauer distribution to the distribution of runs scored by S.M. Gavaskar and G.R. Vishwanath, the two famous batsmen of cricket in the completed innings in the test matches and compared with NBD. The observed as well as the expected frequencies calculated on the basis of the two distributions under consideration are given in Table 7.3 and it is observed that Gegenbauer distribution provides better fit for the distribution of the runs scored by Gavaskar and the runs scored by Vishwanath.

Table 7.3

Applications of Gegenbauer distribution (GD) and negative binomial distribution (NBD) to the runs scored in the completed innings at cricket test (Sinha et al., 1987)

	Gavaskar			Vishwanath		
Runs (Units of 20)	Observed frequency $\bar{x} = 2.1177$ $s^2 = 6.0156$	Expected frequency		Observed frequency $\bar{x} = 1.5775$ $s^2 = 4.0327$	Expected frequency	
		NBD (Sinha et al., 1987)	GD $\hat{\lambda} = 1.6984$ $b = 0.3293$ $a = 0.5991$		NBD (Sinha et al., 1987)	GD $\hat{\lambda} = 1.2989$ $b = 0.1793$ $a = 0.5774$
0	48	42.1	46.70	58	55.1	57.47
1	22	30.4	21.40	30	33.7	29.72
2	21	21.2	23.44	21	20.7	21.33
3	15	14.4	13.98	12	12.6	13.03
4	6	9.2	10.67	7	7.7	8.11
5	11	6.5	6.81	6	4.7	4.93
6	6	7.3	4.65	4	2.9	2.98
7	1	2.9	2.99	1	1.8	1.79
8	2	1.9	1.95	2	1.1	1.07
9	1	1.3	1.25	0	0.7	0.64
10	1	0.8	0.80	0	0.4	0.38
11	2	1.5	1.36	1	0.6	0.55
Total	136	136.0	136.00	142	142.0	142.00
	χ^2	8.91	5.61	χ^2	1.05	0.53
	$d.f.$	5	4	$d.f.$	4	3
	$p-value$	> 0.10	> 0.23	$p-value$	> 0.90	> 0.91

GD: Gegenbauer distribution
NBD: negative binomial distribution

Table 7.4 gives the observed and expected frequencies obtained on the basis of Gegenbauer distribution and NBD (Reep et al, 1971), fitted to the distribution of goals per match scored by individual teams in national Hockey League 1966-67.

Table 7.4

The observed and the expected frequencies of the distribution of goals per match scored by individual teams in National Hockey League 1966-67 (Reep et al, 1971)

No of goals	Observed frequency $\bar{x} = 2.9786$ $s^2 = 3.5304$	Expected frequency NBD (Sinha et al., 1987)	GD $\hat{\lambda} = 99.098$ $b = 0.0369$ $a = 0.0616$
0	29	27.6	27.96
1	71	68.8	68.45
2	82	91.4	90.90
3	89	85.9	85.98
4	65	64.1	64.54
5	45	40.4	40.72
6	24	22.3	22.39
7	7	11.1	10.99
8	4	5.0	4.90
9	1	2.1	2.01
10+	3	1.3	1.16
Total	420	420.0	420.00
	χ^2 $d.f.$ $p-value$	3.43 6 > 0.70	3.13 5 > 0.67

GD: Gegenbauer distribution

NBD: negative binomial distribution

Finally, in Table 7.5, we have given the expected frequencies obtained from Gegenbauer distribution to the number of points per game scored by individual teams in U.S. Collegiate football games where NBD was fitted by Pollard (1973). From Table 7.4

and Table 7.5, it is clear that in both cases Gegenbauer distribution provides better fit than the NBD. The evidences reveal that Gegenbauer distribution may be used to describe random counts arisen from various situations in ball games.

Table 7.5

The observed and expected frequencies of the distribution of the number of points per game scored by individual teams in U.S. Collegiate football games (Pollard, 1973)

Number of points	Observed frequency $\bar{x} = 2.58$ $s^2 = 3.76$	Expected frequency: NBD (Sinha et al., 1987)	Expected frequency: GD $\hat{\lambda} = 4.919$ $b = -0.0339$ $a = 0.3286$
0	272	278.7	275.65
1	485	490.1	491.52
2	537	590.1	512.21
3	407	406.6	408.45
4	258	275.9	276.24
5	157	167.3	166.94
6	101	93.5	92.95
7	57	49.0	48.62
8	23	24.4	24.21
9	8	11.7	11.59
10	5	5.4	5.36
11	6	4.3	4.26
Total	2316	2316.0	2316.00
	χ^2	7.42	7.18
	$d.f.$	9	8
	$p-value$	> 0.50	> 0.51

GD: Gegenbauer distribution

NBD: negative binomial distribution

RFERENCES

1. Adelstein, A.M. (1952): Accident proneness: a criticism of the concept based upon an analysis of shunter' accidents, *Journal of the Royal Statistical Society, Series A,* 115, 334-410.
2. Anscombe, F.J. (1950): Sampling theory of the negative binomial and logarithmic series distributions, *Biometrika,* 37, 358-382.
3. Beall, G. and Rescia, R. (1953): A generalization of Neyman's contagious distribution. *Biometrics*, 9, 354-386.
4. Begum, R.A. and Borah, M. (2003): Certain Infinitely Divisible Discrete Probability Distribution and its Application, *Mathematics and Statistics in Engineering, Biotechnology and Science,* D. K. Sen and P.K. Mishra (editors), 88-111.
5. Berg, S. (1985): Generating discrete distributions from modified Charlier type B exapansions, *Contribution to probability and Statistics in Honour of Gunnar Blom,* edited by J. Lanke and G. Lindgren, 39-48, Lund: University of Lund.
6. Bhattacharya, S.K. (1966): Confluent Hypergeometric distributions of discrete and continuous type with applications to accident proneness, *Bulletin of the Calcutta Statistical Association,* 15, 20-31.
7. Bliss, C. I., and Fisher, R.A. (1953): Fitting the negative binomial distribution to biological data and note on the efficient fitting of the negative binomial, *Biometrics,* 9, 176-200.
8. Borah, M. (1984): The Gegenbauer distribution revised: Some recurrence relations for moments, cumulants. etc. estimation of parameters and its goodness of fit. *Journal of Indian Society of Agricultural Statistics*, 36, 72-78.
9. Borah, M. and Begum, R.A. (1997): A class of Discrete Probability distributions applicable in queueing Theory, *Mathematics and Statistics in Engineering and Technology,* edited by A. Chattopadhyay.

10. Borah, M. and Begum, R.A. (1999): Some properties of Poisson Mixing Infinitely Divisible distributions, *Proceedings of the Annual Technical Session, Assam Science Society*, Vol 1, 40-47.
11. Borah, M. and Begum, R.A (2002): Some properties of Poisson-Lindley and its derived distributions, *Journal of Indian Statistical Association*, Vol. 40, 13-25.
12. Borah, M. and Deka Nath, A. (2000): A class of Hermite type Lagrangian Distributions, *Recent Trends in Mathematical Sciences,* J.C. Misra and S.B. Sinha (Editors), Narosa Publishing house, 319-326.
13. Borah, M. and Deka Nath, A. (2001): Poisson-Lindley and some of its Mixture distributions, *Journal of Pure and Applied Mathematika Sciences,* Vol. 53, No. 1-2, 1-8.
14. Chang, D.K. (1989): On infinitely divisible discrete distributions, *Utilitas Mathematica,* 36, 215- 217.
15. Chatfield, C. (1969): On estimating the parameters of the logarithmic series and negative binomial distributions, *Biometrika,* 56, 411-414.
16. Consul, P.C. (1975): Some new characterizations of discrete Lagrangian distributions, *Statistical Distributions in Scientific Work,* 3: characterizations and Applications, edited by G.P. Patil, S. Kotz and J.K. Ord , 279-290.
17. Consul, P.C. (1989): *Generalized Poisson Distributions properties and Applications,* New York, Dekker.
18. Consul, P.C. and Felix, F. (1996): Lagrangian Katz family of distributions, *Communication in Statistics, Theory and Methods,* 25(2), 415-434.
19. Consul, P.C. and Jain, G.C. (1973a): A generalization of the Poisson distribution, *Technometrics,* 15,791-799.
20. Consul, P.C. and Jain, G.C. (1973b): On some interesting properties of the generalized Poisson distribution, *Biometrische Zeitschrift,* 15, 495-500.
21. Consul, P.C. and Shenton, L.R. (1972): Use of Lagrange expansion for generating discrete generalized probability distributions, *SIAM Journal of Applied Mathematics,* 23, 239-248.
22. Consul, P.C. and Shenton, L.R. (1973): Some interesting properties of Lagrangian distributions, *Communication in Statistics, Theory and Methods* 2, 263- 272.
23. Consul, P.C. and Shenton, L.R. (1975): On the probabilistic structure and properties of discrete Lagrangian distributions, *Statistical Distributions in Scientific Works,* 1: Models and Structures, edited by G.P. Patil, S. Kotz, and J.K. Ord , 41-48, Dordrecht: Reidel.

24. Everitt, B.S. and Hand, D.J. (1981): *Finite mixture distributions*, London: Chapman and Hall.

25. Feller, W. (1950): *An Introduction to Probability Theory and Its Applications*, Vol. I, New York, Wiley.

26. Feller, W. (1957): *An Introduction to Probability Theory and Its Applications*, Vol. I, (second Edition), New York, Wiley.

27. Feller, W. (1968): *An Introduction to Probability Theory and Its Applications*, Vol. I, (third Edition), New York, Wiley.

28. Feller, W. (1943): On a general class of 'contagious' distributions, *Annals of Mathematical Statistics*, 14, 389-400.

29. Fisz, M. (1962): Infinitely divisible distributions: recent results and applications. *Ann. Math. Statist.*, 33, 68- 84.

30. Galliher, H.P., Morse, P.M., and Simond, M. (1959): Dynamics of two classes of continuous rewiew inventory systems, *Operations Research*, 7, 362-384.

31. Garman, P. (1953): Original data on European red mites on apple leaves, *Connecticut*.

32. Gnedenko, B.V. and Kolmogorov, A.N. (1968): Limit distributions for sums of independent random variables, revised edition, *Addison-Wesley Publishing Co.*, Reading Massachusetts.

33. Godambe, A.V. and Patil, G.P. (1969): Infinite Divisibility and additivity of certain probability distributions with an application to mixtures and randomly stopped sums, Department of Statistics, *The Pennsylvania State University Technical Reports and preprints*, No. 20.

34. Godambe, A.V. and Patil, G.P. (1975): Some characterizations involving additivity and infinite divisibility and their applications to Poisson mixtures and Poisson sums, *Statistical Distributions in Scientific Work, 3, Characterizations and Applications,* G.P. Patil, S. Kotz, and J.K. Ord (editors), 339- 351. Dordrecht: Reidel.

35. Goldie Charles (1967): A class of infinitely divisible random variables, *Proceedings of Cambridge. Philosophical. Society.*

36. Gurland, J. (1957): Some interrelations among compound and generalized distributions *Biometrika,* 44, 265-268.

37. Gurland, J. (1958): A generalized class of contagious distributions, *Biometrics*, 14, 229-249.

38. Gurland, J. (1965): A method of estimation for some generalized Poisson distributions, Classical and Contagious Distributions, G.P. Patil (editor), 141-158, *Calcutta: Statistical Publishing Society*; Oxford: Pergamon Press.

39. Haight, F.A. (1961a): A distribution analogous to the Borel-Tanner, *Biometrika*, 48, 167-173.

40. Haight, F.A. (1967): *Handbook of the Poisson Distribution*, New York: Wiley.

41. Jain, G.C. (1975): A power series distribution association with Lagrange expansion, *Biometrische Zeitshrift*, 17, 85-97.

42. Jain, G.C. and Consul, P.C. (1971): A generalized negative binomial distribution, *SIAM Journal of Applied Mathematics*, 21, 501-513.

43. Jain, G.C. and Gupta, R.P. (1975): On a class of polynomials and associated probabilities, *Utilitas as Math*, 7, 363-381.

44. Janardan, K.G. and Rao, B.R. (1983): Lagrange distributions of the second kind and weighted distributions, *SIAM J. Applied Mathematics*, 43, 302- 313.

45. Janardan, K.G. and Schaeffer, D.J. (1977): A generalization of Markov-Polya distribution its extensions and applications, *Biometrical Journal*, 19, 87-106.

46. Janardan, K.G. and Schaeffer, D.J. (1977): Models for the analysis of chromosomal aberrations in human leukocytes, *Biometrical Journal*, 19(8), 599-612.

47. Janardan, K.G., Kerster, H.W., and Schaeffer, D.J. (1979): Biological applications of the Lagrangian Poisson distribution, *Bioscience*, 29, 599-602.

48. Johnson, N.L., and Kotz, S. (1969): *Discrete Distributions*, (first edition), Boston: Houghton Mifflin.

49. Johnson, N.L., Kotz, S. and Kemp, A.W. (1992): *Univariate Discrete Distributions*, (Second Edition), John Wiley & Sons, New York.

50. Katti, S.K. (1967): Infinite divisibility of integer valued random variables, *Annals of Mathematical Statistics*, 38, 1306- 1308.

51. Katti, S.K. (1977): Infinite divisibility of Discrete Distributions III, *Colloquia Mathematica societatis Janos Bolyai, 21. Analytic Function Methods in probability theory, Debrecen (Hungary)* 165-171

52. Katti, S.K. and Gurland, J. (1961): The Poisson Pascal distribution, *Biometrics*, 17, 527-538.

53. Katti, S.K. and Gurland, J. (1962a): Efficiency of certain methods of estimation for the negative binomial and the Neyman type A distributions, *Biometrika*, 49, 215-226.

54. Katti, S.K. and Gurland, J. (1962b): Some method of estimation for the Poisson binomial distribution, *Biometrics*, 18, 42-54.

55. Keilson, J.and Steutel, F.W. (1972): Families of infinitely divisible distributions closed under mixing and convolution. *The Annals of Mathematical Statistics,* 43(1), 242-250.

56. Klebanov, L.B., Maniya, G.M. and Melamed, I.A. (1984): A problem of zolotarev and analogs of infinitely divisible and stable distributions in a scheme of summing a random number of random variables. *Theor. prob. Applicn.* 4, 791- 794.

57. Kemp, A.W. (1979): Convolutions involving binomial pseudo-variables, *Sankhya*, Series A, 232-243.

58. Kemp, A. W. (1992a): Heine-Euler extensions of the Poisson distribution, *Communication in Statistics*, Theory and Methods, 21.

59. Kemp, A. W. (1992b): Steady state Markov chain models for the Heine and Euler distributions, *Journal of Applied Probability*, 29.

60. Kemp, A. W. (1992c): On counts of individuals able to signal the presence of an observer, *Biometrical Journal*, 34.

61. Kemp, A.W. and Kemp, C.D. (1966): An alternative deviation of the Hermite distribution, *Biometrika,* 53, 627-628.

62. Kemp, C. D. (1967): On contagious distribution suggested for accident data, *Biometrics,* 23, 241-255.

63. Kemp, C.D. and Kemp, A.W. (1965): Some properties of the Hermite distribution, *Biometrika,* 52, 381-394.

64. Kendall, M.G. (1961): Natural law in the social sciences, *Journal of the Royal Statistical Society,* Series A, 124, 1-18.

65. Khatri, C.G. (1959): On certain properties of power series distributions, *Biometrika,* 46, 486-490.

66. Khatri, C.G. (1961): On the distributions obtained by varying the number of trials in a binomial distribution, *Annals of the Institute of Statistical Mathematics*, Tokyo, 13, 47-51.

67. Levy, P. (1954): *Theorie de l' Addition des variables Alcatoires,* (second edition), Gauthier- Villars, Paris.

68. Lindley, D.V. (1958): Fiducial distribution and Bayes' Theorem, *Journal of the Royal Statistical Society,* Series B, 20, 102-107.

69. Loeschcke, V., and Kohler,W. (1976): Deterministic and stochastic models of the negative binomial distribution and the analysis of chromosomal aberrations in human leukocytes, *Biometrische Zeitschrift*, 18,427-451.

70. Loeve, M. (1960): *Probability theory*, Second Ed., Van Nostrand, Princeton, N.J.

71. Lloyd, E. L. (1980): *Hand book of Applicable Mathematics*, VIA: Statistics. Chichester: Wiley.

72. Lukacs,E. (1960) *Characteristics Functions* , London: Griffin.

73. Lukacs,E. (1970) *Characteristics Functions* (Second edition), London: Griffin.

74. Maceda, E.C. (1948) On the compound and generalized Poisson distributions, *Annals of Mathematical Statistics*, 19, 414- 416.

75. McBride, E.B. (1971): *Obtaining Generating Functions*, Springer, Berlin, New York.

76. Medgyessy, P. (1977): *Decomposition of Superpositions of Density Functions and Discrete Distributions*, Budapest: Akademia Kiado; Bristol: Adam Hilger.

77. Medhi, J. and Borah, M. (1986). On generalized four- parameter Charlier distribution, *Journal of Statistical Planning and Inference*, 14, 69- 77.

78. Medhi, J. and Borah, M. (1984): On generalized Gegenbauer polynomials and associated probabilities. *The Indian Journal of Statistics*, Series-B, Vol 46, Part 2, 157-165.

79. Mc Guire, J.U. Brindley, T.A. and Bancroft, T.A. (1957): The distribution of European corn borer Pyrausta Nubilalis (HBN) in field corn, *Biometrics,* 13, 65-78 (extra and extensions 14, (1958), 432-434.).

80. Mohanty, S.G. (1966): On a generalized two coin tossing problem, *Biometrische Zeitschrift,* 8, 266-272.

81. Morney, M.J. *(1956): Facts from figures* (3rd edition) 102, Penguin, London.

82. Patil, G.P. (1961): *Contributions to estimation in a class of Discrete Distributions, Ph.D. thesis,* Ann Arbor, MI: University of Michigan.

83. Patil, G.P. (1962a): Maximum-likelihood estimation for generalized power series distributions and its application to a truncated binomial distribution, *Biometrika*, 49, 227-237.

84. Patil, G.P. (1962b): Some methods of estimation for the logarithmic series distribution, *Biometrics,* 18, 68-75.

85. Patil, G.P. (1964): On certain compound Poisson and compound binomial distributions, *Sankhya,* series A, 26, 293-294.

86. Pillai, R.N. (1990): Harmonic mixtures and geometric infinite divisibility. *Journal of Indian Statistical Association.* 28, 87- 98.

87. Pillai, R.N. and Sandhya, E. (1990): Distributions with complete monotone derivative and geometric infinite divisibility. *Advanced Applied Probability*, 22, 751- 754.

88. Pillai, R.N. and Jose, K.K. (1994): Geometric infinite divisibility and autoregressive time series modelling, *Proceeding of the third Ramanujan Symposium on Stochastic Processes and their Applications,* University of Madras, Madras 17-19 January, 1994, 81- 87.

89. Plunkett, I. G., and Jain, G.C. (1975): Three generalized negative binomial distributions. *Biometrische Zeitschrift.* 17, 286-302.

90. Pollard, R. (1973): Collegiate football scores and the negative binomial distribution, *Journal of American Statistical Association*, 68, 315-352.

91. Rainville, E.R. (1960): Special Functions, New York: Macmillan.

92. Reep, C. and Benjamin, B. (1968): Skill and chance in association football, *Journal of Royal Statistical Society, A,* 131, 581-585.

93. Reep, C., Polllard, R. and Benjamin, B.(1971): Skill and chance in ball games, *Journal of Royal Statistical Society, A,* 134, 625-629.

94. Sandhya, E. (1991): Geometric Infinite Divisibility and Applications, *Ph. D. thesis* submitted to the University of Kerala, January, 1991.

95. Sankaran, M.(1970): The discrete Poisson Lindley distribution, *Biometrics*, 26, 145-149.

96. Sinha, A.K., Kumar, A. and Kumar, R. (1987): Some applications of the negative binomial and Neyman Type A distributions to Ball Games, *ASR,* Vol. 1, No. 1, 22-31.

97. Skellam, S.N. (1952): Studies in statistical ecology 1: Spatial pattern, *Biometrika*, 39, 346-362.

98. Steutel, F.W. (1983): Infinite divisibility, *Encyclopedia of Statistical Sciences*, 4, S. Kotz, N.L. Johnson and C.B. Read (editors), 114- 116, New York, Wiley.

99. Steutel, F.W. (1970): Preservation of infinite divisibility under mixing, *Mathematical Centre Tract,* 33, Amsterdam.

100. Steutel, F.W. (1990): The set of geometrically infinitely divisible distributions, *Technische Universiteit Eindhoven Memorandum* COSOR.

101. Steutel, F.W. (1968): A class of Infinitely Divisible Mixtures, *Annals of Mathematical Statistics*, 39, 1153-1157.

102. Steutel, F.W. (1973): Some recent results in infinite divisibility, *Stochastic Pocesses Appl.*, Vol. 1, 125- 143.

103. Steutel, F.W. and Harn, K.V. (2004): Infinite divisibility of probability distributions on the real line, Pure and Applied Mathematics, A Dekker series of Monographs and Textbooks.

104. Tanner, J. C. (1961): *Biometrika*, 48, 222-224.

105. Warde, W.D. and Katti, S.K. (1971) Infinite divisibility of discrete distributions II, *Annals of Mathematical Statistics*, 42, 1088- 1090.

106. Wood, G. H. (1945): "Cricket scores and geometric progression." *Journal of Royal Statistical Society,* A, Vol. 108, 12-22.

Appendix A

List of Publications

1. Borah M. & R.A. Begum (1997): A class of discrete probability distributions applicable in queuing theory in Mathematics and Statistics in Engineering and Technology (Edited by A. Chattopadhya), Narosa Pub. House, New Delhi, 180 - 198.

2. Borah, M. and R. Begum (1999): Some Properties of Poisson Mixing Infinitely Divisible Distribution, *Proceedings of the Annual Technical Session of Assam Science Society*, vol-1, 40- 47.

3. Borah, M. and R.A. Begum (2002): Some Properties of Poisson- Lindley and its derived distributions, *Journal of the Indian Statistical Association, University of Pune 411007,* Vol. 40, 13- 25.

4. Begum, R.A. and M. Borah (2003): Certain Infinitely divisible discrete probability distributions and its applications, *Mathematics and Statistics in Engineering Bio- Technology and Science*, edited by D.K.Sen and P.K.Mishra, Centre for Bio Mathematical Studies, Jharkhand, India, 88-111.

5. Begum, R.A. and M. Borah (2006): Application of Gegenbauer Distribution to Ball Games, *Assam Statistical Review, A Research Journal of Statistics,* Published By the Department of Statistics, Dibrugarh University, Vol. 20, No.1-2, 60-72.

Appendix B

Charlier family of Lagrangian distribution of first kind

$g(t)$	$f(t)$	Probability mass function. $P(X=x)$
$G_1^{\alpha,\beta,\lambda}(t)$	$G_2^{\gamma}(t)$	$\frac{(1-\beta)^{\lambda x}\gamma e^{-(\alpha x+\gamma)}}{x!}\sum_{j=0}^{r}\binom{r}{j}(\alpha x+\gamma)^{r-j}\beta^{j}(\lambda x)_j$, $for\ x\geq 1$, $r=x-1$, $(\lambda x)_j=\lambda x(\lambda x+1)(\lambda x+2)...(\lambda x+j-1)$, $P_0=e^{-\gamma}$.
$G_1^{\alpha,\beta,\lambda}(t)$	$G_3^{\beta,N}(t)$	$\frac{N\beta(1-\beta)^{N+\lambda x}e^{-\alpha x}}{x!}\sum_{j=0}^{r}\binom{r}{j}(\lambda x+N+1)_j(\alpha x)^{r-j}\beta^{j}$, $for\quad x=1,2,3...,r=x-1,P_0=(1-\beta)^{N}$, $and\quad (\lambda x+N+1)_j=(\lambda x+N+1)(\lambda x+N+2)...(\lambda x+N+j)$.
$G_2^{\gamma}(t)$	$G_3^{\beta,N}(t)$	$\frac{e^{-\gamma x}N\beta(1-\beta)^{N}}{x!}\sum_{j=0}^{r}\binom{r}{j}(\gamma x)^{r-j}(N+1)_j\beta^{j}$, $for\ x=1,2,3...$, $r=x-1,(N+1)_j=(N+1)(N+2)(N+3)...(N+j)$, $and\ P_0=(1-\beta)^{N}$.
$G_1^{\alpha,\beta,\lambda}(t)$	$G_4^{\beta}(t)$	$\frac{\beta(1-\beta)^{\lambda x}e^{-\alpha x}a}{x!}\sum_{j=0}^{r}\binom{r}{j}(\lambda x+1)_j(\alpha x)^{r-j}\beta^{j}$, $for\ x=1,2,3...,r=x-1,a=\{-\log(1-\beta)\}^{-1}\ and$ $(\lambda x+1)_j=(\lambda x+1)(\lambda x+2)...(\lambda x+j)$.
$G_2^{\gamma}(t)$	$G_4^{\beta}(t)$	$\frac{e^{-\gamma x}}{x!}a\beta\sum_{j=0}^{r}\frac{r!}{j!}(\gamma x)^{j}\beta^{r-j}$, $for\ x=1,2,3...,and\quad r=x-1$ $a=\{-\log(1-\beta)\}^{-1}$, $and\quad P_0=a\beta e^{-\gamma}$.
$G_3^{\beta,N}(t)$	$G_4^{\beta}(t)$	$\frac{(Nx+1)_r}{x!}a\beta^{r+1}(1-\beta)^{Nx}$, $for\ x=1,2,3...,\ r=x-1\ and$ $(Nx+1)_r=(Nx+1)(Nx+2)...(Nx+r)$, $a=\{-\log(1-\beta)\}^{-1}$.
$G_2^{\gamma}(t)$	$G_1^{\alpha,\beta,\lambda}(t)$	$\frac{(1-\beta)^{\gamma}e^{-(\alpha+\gamma x)}}{x!}\sum_{j=0}^{r}\binom{r}{j}(\alpha+\gamma x)^{r-j}\beta^{j}\{\alpha(\gamma)_j+\lambda\beta(\gamma+1)_j\}$ $for\quad x=1,2,3...,\ and\quad r=x-1$, $(Nx)_r=Nx(Nx+1)(Nx+2)...(Nx+r-1)$, $P_0=e^{-\alpha}(1-\beta)^{\lambda}$.
$G_3^{\beta,N}(t)$	$G_2^{\gamma}(t)$	$\frac{e^{-\gamma}\gamma(1-\beta)^{Nx}}{x!}\sum_{j=0}^{r}\binom{r}{j}\gamma^{r-j}(Nx)_j\beta^{j}$, $for\ x=1,2,3...$, $r=x-1$, $(Nx)_j=Nx(Nx+1)(Nx+2)...(Nx+j-1)$, $P_0=e^{-\gamma}$.
$G_3^{\beta,N}(t)$	$G_1^{\alpha,\beta,\lambda}(t)$	$\frac{(1-\beta)^{\lambda+Nx}e^{-\alpha}}{x!}\sum_{j=0}^{r}\binom{r}{j}\alpha^{r-j}\beta^{j}\{\alpha(\lambda+Nx)_j+\lambda\beta(\lambda+Nx+1)_j\}$, $for\ x=1,2,3...,\ r=x-1,P_0=e^{-\alpha}(1-\beta)^{\lambda}$. $(\lambda+Nx)_r=(\lambda+Nx)(\lambda+Nx+1)(\lambda+Nx+2)...(\lambda+Nx+r-1)$,

Charlier family of Lagrangian distribution of second kind

$g(t)$	$f(t)$	Probability mass function. $P(X=x)$
$G_2^{\gamma}(t)$	$G_3^{\beta,N}(t)$	$\frac{e^{-\gamma x}(1-\gamma)(1-\beta)^N}{x!}\sum_{j=0}^{x}\binom{x}{j}(\gamma x)^{x-j}(N)_j\beta^j$ *for* $x=0,1,2,3...,$ *and* $(N+1)_j=(N+1)(N+2)(N+3)...(N+j)$
$G_3^{\beta,N}(t)$	$G_2^{\gamma}(t)$	$\frac{e^{-\gamma}(1-\beta)^{Nx}}{x!}\left\{1-\left(\frac{N\beta}{1-\beta}\right)\right\}\sum_{j=0}^{x}\binom{x}{j}\gamma^{x-j}(Nx)_j\beta^j$ *for* $x=0,1,2,3...,$ *and* $(Nx)_j=Nx(Nx+1)(Nx+2)...(Nx+j-1),$
$G_2^{\gamma}(t)$	$G_1^{\alpha,\beta,\lambda}(t)$	$\frac{(1-\beta)^{\lambda}(1-\gamma)e^{-(\gamma x+\alpha)}}{x!}\sum_{j=0}^{x}\binom{x}{j}(\alpha+\gamma x)^{x-j}\beta^j(\lambda)_j$ *for* $x=0,1,2,3...,$ *and* $(Nx)_r=Nx(Nx+1)(Nx+2)...(Nx+r-1),$
$G_3^{\beta,N}(t)$	$G_1^{\alpha,\beta,\lambda}(t)$	$\frac{(1-\beta)^{\lambda+Nx}e^{-\alpha}}{x!}\left\{1-\left(\frac{N\beta}{1-\beta}\right)\right\}\sum_{j=0}^{x}\binom{x}{j}\alpha^{x-j}\beta^j(\lambda+Nx)_j$ *for* $x=0,1,2,3...,$ $(Nx+\lambda)_r=(Nx+\lambda)(Nx+\lambda+1)(Nx+\lambda+2)...(Nx+\lambda+r-1),$
$G_1^{\alpha,\beta,\lambda}(t)$	$G_3^{\beta,N}(t)$	$\frac{(1-\beta)^{N+\lambda x}e^{-\alpha x}}{x!}\left\{1-\left(\alpha+\frac{\lambda\beta}{1-\beta}\right)\right\}\sum_{j=0}^{x}\binom{x}{j}(\lambda x+N)_j(\alpha x)^{x-j}\beta^j$ *for* $x=0,1,2,3...,$ $(\lambda x+N)_j=(\lambda x+N)(\lambda x+N+1)(\lambda x+N+2)...(\lambda x+N+j-1),$
$G_1^{\alpha,\beta,\lambda}(t)$	$G_2^{\gamma}(t)$	$\left\{1-\left(\alpha+\frac{\lambda\beta}{\beta}\right)\right\}\frac{(1-\beta)^{\lambda x}\gamma e^{-(\alpha x+\gamma)}}{x!}\sum_{j=0}^{x}\binom{x}{j}(\alpha x+\gamma)^{x-j}\beta^j(\lambda x)_j$ *for* $x=0,1,2,3...,$ $(\lambda x+\gamma)_j=(\lambda x+\gamma)(\lambda x+\gamma+1)(\lambda x+\gamma+2)...(\lambda x+\gamma+j-1),$
$G_2^{\gamma}(t)$	$G_4^{\beta}(t)$	$a(1-\gamma)e^{-\gamma x}\sum_{j=0}^{x-1}\frac{1}{j!(x-j)}(\gamma x)^j\beta^{x-j}$ *for* $x=1,2,3...,\ a=\{-\log(1-\beta)\}^{-1},$ *and* $(\gamma x)_r=\gamma x(\gamma x+1)(\gamma x+2)...(\gamma x+r-1),$
$G_3^{\beta,N}(t)$	$G_4^{\beta}(t)$	$\left(1-\frac{N\beta}{1-\beta}\right)a\beta^x(1-\beta)^{Nx}\sum_{j=0}^{x-1}\frac{(Nx)_j}{(x-j)j!}$ *for* $x=1,2,3...,\ a=\{-\log(1-\beta)\}^{-1},$ *and* $(Nx)_r=Nx(Nx+1)(Nx+2)...(Nx+r-1).$

Appendix C

A Brief note on the methods of estimation

The method of Maximum Likelihood is the best and the most important method of estimating parameters of a distribution. It was first used by C. F. Guass in developing the theory of least squares and subsequently reintroduced by Prof. R. A. Fisher in 1912. Besides this there are other methods also. They are method of moments, method of minimum chi square, method of minimum variance, method of inverse Probability and least square method of estimation. Of all these methods, method of moments is perhaps the oldest and the simplest and it can also be used with a desired number of accuracy. Estimation by using the method of moments was first introduced by Karl Pearson (1900). This method sometimes yield estimates most easily in comparison with Maximum Likelihood estimates.

In certain situations where the method of Maximum Likelihood leads to intractable equations, then method of moments is the best method to estimate the parameters of a distribution. It is also true that method of moments do not lead us to the same estimate as the method of Maximum Likelihood. It has been observed that under certain general conditions the estimates obtained by the method of moments are asymptotically normal but not in general efficient. Generally method of moments yields less efficient estimates than those obtained from the method of Maximum Likelihood.

It has been observed that, in case of certain mixture distributions specially when there are three or more than three parameters, the solution of Maximum Likelihood equations for estimating the parameters is very cumbersome. For example, in our case, the generalized distributions of Poisson Lindley, three parameter Charlier and Gegenbauer distributions in Chapters 3, 4 and 5, we have faced such type of problems. Some authors like Fisher, 1941; Anscombe, 1950, Evans, 1953; Katti and Gurland, 1961, 1962 have suggested alternative methods which can lead to relatively simple equations to solve. These methods have been found to be very useful in the problem of fitting of the distributions. Katti and Gurland (1961, 1962) have also calculated the efficiencies of these methods. These methods are method of moments, method of minimum chi square and method of using first two moments and the proportion of zeros in estimating the parameters of Poisson Pascal and Poisson Binomial distributions. But it may be well to say that under certain situations these methods give maximum results.

Katti and Gurland (1961, 1962) used the method of moments when the mean and the variances are in large magnitude. When mean and the variances are moderate and the proportion of the zero class frequency is large they used the method of the first two moments and the proportion of the zero class frequency in estimating the parameters of Poisson Pascal distribution. Again when the first three frequencies are large in comparison with the remaining, the method of first two moments and the ratio of the first two frequencies are used in estimation.

In view of the above results of Katti and Gurland (1961, 1962), we have taken some ad hoc methods to estimate the parameters in different Chapters. The p values are also calculated for each Table and have shown along with the chi square values and their respective degrees of freedoms.

But in certain tables in Chapter 4, in fitting Generalized negative binomial and Generalized Hermite distributions of type 1 and in Chapter 6, in fitting Lagrangian negative binomial Poisson and Lagrangian Poisson negative binomial distributions as the number of observations in the data sets is less than equal to 5, the fitting of the distributions can not be compared on the basis of chi square criterion as the degrees of freedom is very negligible. That is why in those tables p-values are not provided.

31

A Class of Discrete Probability Distributions Applicable in Queueing Theory

Munindra Borah and Rousan Ara Begam*
Department of Mathematical Sciences, Tezpur University,
Sonitpur, Assam, India

*Department of Statistics, Darrang College, Tezpur, India

1. Introduction

In queueing theory, to obtain the expected length and variance of the random variable X, i.e. the number of customers served in the First Busy Period (FBP), the probability distribution of X is needed; but this is dependent upon the systems of arrivals and the service interval of each customer. Similar type of problem had been considered for Poisson arrivals at a constant service time by Borel and Tanner (Tanner 1961 and Prabhu, 1965). It has been well illustrated by the authors [Consul and Shen (1971)] that Lagrangian probability distribution (LPD) consist of many families and that has many interesting members such as Borel-Tanner distribution (Tanner, 1951). Haight distribution (Haight, 1961), generalized negative binomical distribution (Jain and Consul, 1971), and generalized Poisson distribution (Consul and Jain, 1972). Mohanty (1966) obtained the Lagrangian binomial distribution as the distribution of he number of failures x encountered in getting $\beta x + n$ successes given a sequence of n independent bernoulli trials. A queueing process interpretation has been given by Takacs (1962) and Mohanty (1966). The application of Lagrangian type distributions in the theory of random mappings has been studied by Berg and Mutafchiev (1990) and Berg and Nowicki (1991). Devroye (1992) has studied the computer generation of Lagrangian type variables. Consul's (1989) book on the Lagrangian Poisson distribution highlights how intensively it has been studied for its many properties, and all for its various mode of genesis. Here we have considered a more general problem where the input of the customers is a three parameter Charlier (where Poisson, or negative binomial may be obtained as a particular case), and the initial number of customers waiting for service in a queue is also another discrete random variable. A very simple and well known method to determine the probability distribution of the number of customers in the FBP has been considered

2. Lagrangian Distributions

A class of discrete probability distributions under the title "Lagrangian Distributions" had been introduced into the literature by Consul and Shenton (1971, 1973) The particular title was chosen by them on account of the generation of these probability distributions by using the well known Lagrange expansion of a function $f(x)$ as a power series in y when $y = x/g(x)$ Considering $g(s)$ and $f(s)$ as the two probability generating functions (p g f) defined on non-negative integers such that $g(0) \neq 0$, Consul and Shenton (1972) used Lagrange's expansion to define families of discrete generalized probability distributions which is called Langrange Distribution of first kind (LD1 according to the terminology of Janardan and Rao (1983) Lagrange disribution of this kind has the probability mass function (p m f)

$$Pr(X = 0) = f(0),$$

$$Pr(X = x) = \frac{1}{x!}\left[\frac{d^{x-1}}{ds^{x-1}}\{g(s)\}^{x} f'(s)\right]_{s=0}, \text{ for } x = 1, 2, 3, \qquad (1\ 1)$$

Using the Lagrange's expansion of second kind, Janardan and Rao (1983) investigate a new class of discrete distributions called lagrange distributions of the second kind (LD2) with the p m f

$$Pr(X = x) = \frac{1 - g'(1)}{x!}\frac{d^{x}}{ds^{x}}[\{g(s)\}^{x} f(s)]_{s=0}, \text{ for } x = 0, 1, 2, 3, \qquad (1\ 2)$$

Consul and Shenton (1972), and Janardan and Rao (1983) generate different families of Lagrange distribution by various choice of the functions $f(s)$ and $g(s)$

3. Charlier Type Distributions

Doetsch (1933), Meixner (1934, 1938) and Berg (1985) investigated the Charlier polynomials defined by the generating funct on

$$e^{s}(1 - \beta s)^{-N} \qquad (2\ 1)$$

Jain and Gupta (1975) defined the generalized Charlier polynomial $C^{\lambda}_{n\,m}(\alpha, \beta)$ of degree n by the generating function

$$e^{\alpha s}(1 - \beta s^{m})^{-\lambda} = \sum_{n=0}^{[n/m]} C^{\lambda}_{n\,m}(\alpha, \beta)\, s^{n}/n! \qquad (2\ 2)$$

Hence

$$C^{\lambda}_{n\,m}(\alpha, \beta) = \sum_{j=0}^{[n/m]} n!\,\frac{\beta^{j}}{j!}\,\frac{\alpha^{n-mj}}{(n - mj)!}\,\frac{\Gamma(\lambda + j)}{\Gamma(\lambda)} \qquad (2\ 3)$$

where $[K]$ denotes the integer part of K Medhi and Bora (1984) have also studied the probability, moment and cumulant properties of four parameter charlier distribution, and have used estimation via the first three cumulants and the ratio of the first two sample frequencies (f_1/f_0). Jain and Gupta (1975) have obtained some operational formulae and recurrence relations for probabilities and moments. They have extensively studied the various properties of he polynomials.

The p.g.f. of three parameter Charlier distribution (for $m = 1$) is given by

$$G_1^{\alpha,\beta,\lambda}(s) = e^{-\alpha}(1-\beta)^{\lambda} e^{\alpha s}(1-\beta s)^{-\lambda} \tag{2.4}$$

with the probability mass function (p.m.f.)

$$P_r^{\alpha,\beta,\lambda} = \frac{1}{r!}\frac{d^r}{ds^r}\{G_1^{\alpha,\beta,\lambda}(s)\}$$

$$= \frac{1}{r!} P_0^{\alpha,\beta,\lambda} \sum_{j=0}^{r} {}^rC_j \frac{\beta^j \alpha^{r-j}\Gamma(\lambda j)}{\Gamma(\lambda)}, \text{ for } r = 1, 2, 3, \ldots \tag{2.5}$$

where $P_0^{\alpha,\beta,\lambda} = e^{-\alpha}(1-\beta)^{\lambda}$, and ${}^rC_j = r!/\{j!\,(r-j)!\}$. As the ranges for the skewness and kurtosis of the three parameter Charlier distribution covers the ranges of the other basic distributions, viz. 'Poisson, Negative binomial etc. as demonstrated in the literature, this family of Lagrangian Distributions may be considered to be more flexible.

Similarly, the p.g.f. of Poisson, Negative binomial, Logarithmic series and delta distributions may be given by

$$G_2^{\gamma}(s) = \exp[\gamma(s-1)]$$

$$G_3^{\beta,N}(s) = (1-\beta)^N(1-\beta s)^{-N}$$

$$G_4^{\beta}(s) = \log(1-\beta s)/\log(1-\beta)$$

$$G_5^{n}(s) = s^n \tag{2.6}$$

respectively, where prefixes (i.e. α, β, λ, γ, n, N) denote the parameter/s of the corresponding distribution.

A new class (Charlier Type) of Lagrangian distributions has been derived by selecting various p.g.f. given in equations (2.4) and (2.6) at random as $g(s)$ and $f(s)$ respectively, and putting them in the expressions (1.1) and (1.2) just like the other authors [Consul et al. (1973), and Janardan et al. (1983)]. The Charlier family of Lagrange distributions of the first and second kinds derived from Lagrange's expansion of first kind (LD1) and second kind (LD2) is presented in Tables 1 and 2, respectively. The basic Lagrangian Charlier distributions of first kind may be obtained putting $n = 1$ in serial number 7 (in Table 1) as

$$P_r^{\alpha,\beta,\lambda} = (1/x!)(1-\beta)^{\lambda x} e^{-\alpha x}(\alpha x)_2^{x-1}\, F_0(1-x,\, \lambda x,\, -\beta/\alpha x),\ x \geq 1 \tag{2.7}$$

Similarly, Lagrangian Charlier distributions of second kind may be obtained as

Table 1. Charlier Family of Lagrange Distributions of first kind

No.	$g(s)$	$f(s)$	LD1$(g; f; x)$	
1	$G_1^{\alpha,\beta,\lambda}(s)$	$G_2^{\gamma}(s)$	$\gamma(1-\beta)^{\lambda x}e^{-(\alpha x+\gamma)}(\alpha x+\gamma)^{x-1}/x!$ ${}_2F_0(1-x, \lambda x, -\beta/(\alpha x+\gamma))$,	$x \geq 1$
2	$G_1^{\alpha,\beta,\lambda}(s)$	$G_3^{\beta,N}(s)$	$N\beta(1-\beta)^{\lambda x+N}e^{-\alpha x}(\alpha x)^{x-1}/x!$ ${}_2F_0(1-x, \lambda x+N+1, -\beta/\alpha x)$,	$x \geq 1$
3	$G_1^{\gamma}(s)$	$G_3^{\beta,N}(s)$	$N\beta(1-\beta)^{N}e^{-\gamma x}(\gamma x)^{x-1}/x!$ ${}_2F_0(1-x, N+1, -\beta/\gamma x)$,	$x \geq 1$
4	$G_1^{\alpha,\beta,\lambda}(s)$	$G_4^{\beta}(s)$	$A\beta(1-\beta)^{\lambda x}e^{-\alpha x}(\alpha x)^{x-1}/x!$ ${}_2F_0(1-x, \lambda x+1, -\beta/\alpha x)$, where $A = -1/\log(1-\beta)$	$x \geq 1$
5	$G_2^{\gamma}(s)$	$G_4^{\beta}(s)$	$A\beta\, e^{-\gamma x}/x!$ ${}_2F_0(1-x, 1, -\beta/\gamma x)$, where $A = -1/\log(1-\beta)$	$x \geq 1$
6	$G_3^{\beta,N}(s)$	$G_2^{\gamma}(s)$	$(1-\beta)^{Nx}e^{-\gamma}\gamma^{x}/x!$ ${}_2F_0(1-x, Nx, -\beta/\gamma)$,	$x \geq 1$
7	$G_1^{\alpha,\beta,\lambda}(s)$	$G_5^{n}(s)$	$(n/x)\,(1-\beta)^{\gamma x}e^{-\alpha x}(\alpha x)^{x-n}/(x-n)!$ ${}_2F_0(n-x, \lambda x, \beta/\alpha x)$,	$x \geq n$

$G_1^{\alpha,\beta,\lambda}(s)$, $G_2^{\gamma}(s)$, $G_3^{\beta,N}(s)$, $G_4^{\beta}(s)$ and $G_5^{n}(s)$ denote the probability generating function of three parameter Charlier, Poisson, Negative binomial, logarithmic series and delta distribution respectively. ${}_2F_0(a, b; x)$ denotes hypergeometric function.

$$P_r^{\alpha,\beta,\lambda} = \{1/(x-1)!\}\, A(1-\beta)^{\lambda x}e^{-\alpha x}(\alpha x)^{x-1}_{2}\, F_0(1-x, \lambda x, -\beta/\alpha x),\ x \geq 1 \tag{2.8}$$

where $A = [1 - \{\alpha + \lambda\beta/(1-\beta)\}]$, putting $n = 1$ in number 14 in Table 2. It is very interesting to note that the generalized Poisson distribution (see Consul, 1973) is a particular case of Charlier family of Lagrange distributions. All these distributions shown in Tables 1 and 2, will be of relevance in queueing theory and possess all the interesting properties discussed by Consul and Shenton (1973) in the case of generalized Poisson disribution. In general, it is also conceivable that discrete data occurring in the ecology, epidemiology, and meteorology could be statistically modelled on one of the distributions considered in this investigation, see for example Jain (1975).

4. Properties of Moments and Cumulants

The general Lagrangian Probability Distribution (LPD) possesses, some important properties (see Consul and Shenton, 1973). The cumulants K_r, $r = 1, 2, 3, \ldots$, of the general LPD become

$$K_r = \sum_{m=1}^{r} F_m\left[\sum \frac{r!}{\pi_1!\,\pi_2!\,\pi_3!\ldots\pi_r!}\sum_{i=1}^{r}(D_i/i!\,\pi_i)\right] \tag{3.1}$$

Table 2. Charlier Family of Lagrange Distributions of Second Kind

No.	$g(s)$	$f(s)$	LD2$(g; f; x)$	
8	$G_2^{\gamma}(s)$	$G_3^{\beta,N}(s)$	$(1-\gamma)(1-\beta)^N e^{-\gamma x}(\gamma x)^{x-1}/x!$ ${}_2F_0(-x, N, -\beta/\gamma x)$,	$x \geq 1$
9	$G_3^{\beta,N}(s)$	$G_2^{\gamma}(s)$	$\{1 - N\beta/(1-\beta)\}(1-\beta)^{Nx} e^{-\gamma}\gamma^x/x!$ ${}_2F_0(-x, Nx, -\beta/\gamma)$,	$x \geq 1$
10	$G_2^{\gamma}(s)$	$G_1^{\alpha,\beta,\lambda}(s)$	$(1-\gamma)(1-\beta)^{\lambda} e^{-\gamma x}(\gamma x + \alpha)^x/x!$ ${}_2F_0(-x, \gamma, -\beta/(\gamma x + \alpha))$,	$x \geq 1$
11	$G_3^{\beta,N}(s)$	$G_1^{\alpha,\beta,\lambda}(s)$	$\{1 - N\beta/(1-\beta)\}(1-\beta)^{Nx+\lambda} e^{-\alpha}\alpha^x/x!$ ${}_2F_0(-x, Nx + \lambda, -\beta/\alpha)$,	$x \geq 1$
12	$G_1^{\alpha,\beta,\lambda}(s)$	$G_3^{\beta,N}(s)$	$[1 - \{\alpha + \lambda\beta/(1-\beta)\}](1-\beta)^{\lambda x+N} e^{-\alpha x}(\alpha x)^x$ $/x!\ {}_2F_0(-x, \lambda x + N, -\beta/\alpha x)$,	$x \geq 1$
13	$G_1^{\alpha,\beta,\lambda}(s)$	$G_2^{\gamma}(s)$	$[1 - \{\alpha + \lambda\beta/(1-\beta)\}](1-\beta)^{\lambda x} e^{-(\alpha x+\gamma)}$ $(\alpha x + \gamma)^x/x!\ {}_2F_0(-x, \lambda,$ $-\beta/(\alpha x + \gamma))$,	$x \geq 1$
14	$G_1^{\alpha,\beta,\lambda}(s)$	$G_5^{n}(s)$	$[1 - \{\alpha + \lambda\beta/(1-\beta)\}]9(1-\beta)^{\lambda x} e^{-\alpha x}(\alpha x)^{x-n}$ $/(x-n)!\ {}_2F_0(n-x, \lambda x, -\beta/\alpha x)$,	$x \geq n$
15	$G_1^{\alpha_1,\beta,\gamma_1}(s)$	$G_1^{\alpha_2,\beta,\lambda_2}(s)$	$[1 - \{\alpha_1 + \lambda_1\beta/(1-\beta)\}](1-\beta)^{\lambda_1 x+\lambda_2} e^{-(\alpha_1 x+\alpha_2)}$ $(\alpha_1 x + \alpha_2)^x/x!\ {}_2F_0(-x, \lambda_1 x + \lambda_2,$ $-\beta/(\alpha_1 x + \alpha_2))$,	$x \geq 1$

$G_1^{\alpha,\beta,\lambda}(s)$, $G_2^{\gamma}(s)$, $G_3^{\beta,N}(s)$ and $G_5^{n}(s)$ denote the probability generating function of three parameter Charlier, Poisson, Negative binomial and delta distribution respectively. ${}_2F_0(a, b; x)$ denotes hypergeometric function.

where the second summation is taken ove all partitions $\pi_1, \pi_2, \pi_3, \ldots, \pi_r$ of m such that $\pi_1 + 2\pi_2 + 3\pi_3 + \ldots + r\,\pi_r = r$. Similar expression is also available for moments (Consul and Shenton, 1972). Hence, these formulae can be used to write down higher moments and cumulants of any generalized Lagrangian probability distributions (LPD).

For simpicity, let F_r be the rth cumulant for the pgf $f(s)$, and let D_r be the rth cumulant for the basic Lagrangian distribution obtained from $g(s)$. Then the first few cumulants of LPD can be written down as particular cases of (3.1) in the form:

$$K_1 = F_1 D_1,$$

$$K_2 = F_1 D_2 + F_2 D_1^2,$$

$$K_3 = F_1 D_3 + 3F_2 D_1 D_2 + F_3 D_1^3,$$

$$K_4 = F_1 D_4 + 3D_2 D_2^2 + 4F_2 D_1 D_3 + 6F_3 D_1^2 D_2 + F_4 D_1^4 \tag{3.2}$$

Hence the parameters of LPD can be estimated in terms its cumulants. Minimum variance unbiased estimation for Lagrangian distributions has been examined by Consul and Famoye (1989).

5. Conclusion

This paper defines a class of Charlier type of Lagrangian probability Distributions by using well known Lagrange's expansions, which are aplicable in queueing theory. It is believed that LPD should give better fit than their classical forms. It may be of interest to investigate in further.

6. Acknowledgement

First author expresses his thanks for the partial support provided by the Tezpur University under Minor Research Project grant (U.G.C.).

References

1. Berg, S. (1985). Generating Discrete Distributions from Modified Charlier Type B Expansions. Contributions to Probability and Statistics in Honour of Gunnar Blom, J. Lanke and G. Lindgren (editors), 39–48, Lund, Unviersity of Lund.
2. Consul, P.C. and Jain, G.C. (1973). A Generalization of Poisson Distribution, Technometrics, 15, 791–799.
3. Consul, P.C. and Shenton, L.R. (1972). Use of Lagrange Expansion for Generating Discrete Generalized Probability Distributions, SIAM J. Appl. Math., 23, 239–248.
4. Consul, P.C. and Shenton, L.R. (1973). Communication in Statistics, 2, 263–272
5. Doetsch, G. (1934) Die in der Statistic Seltener Ereignisse Auftretenden Charlierschen Polynome und eine Damit Zusammenhangende Differentialdiffer Enzengleichung. Mathematische Annalen, 109, 257–266.
6. Haight, F.A. (1961). A Distribution Analogous to the Borel-Tanner, Biometrika, 48, 167–173.
7. Jain, G.C. (1975). On Power Series Distribution Associated with Lagrange Expansion, Biomerishe Zeitschrif., 17, 85–97.
8. Jain, G.C. and Consul, P.C. (1971) A Generalized Negative Binomial distribution, SIAM J. Appl. Math., 21, 501–513.
9. Jain, G.C. and Gupta, R.P. (1975) On a Class of Polynomials and Associated Probabilities, Utilitas Mathematica, 7, 363–381.
10. Janardan, K.G. (1982). Weighted Lagrange Distributions and their Characterizations, SIAM J. Appl. Math. 47, 411–415.
11. Janardan, K.G. and Rao, B.R. (1983). Lagrange Distributions of the Second Kind and Weighted Distributions, SIAM J. Appl. Math., 43, 302–313.
12. Meixner, J. (1934). Orthogonals Polynomsysteme Mit Einer Besonderen Gestalt Der Erzeugenden, Funktion, J. London Math. Soc., 9, 6–13.
13. Meixner, J. 91938). Erzeugende Funktionen der Charlierschen Polynome, *Mathematische Zeitschrift*, 44, 531–535.
14. Medhi, J. and Borah, M. (1986). On Generalized Four-parameter Charlier Distribution, *J. Statistical Planning* and Inference, 14, 69–77.
15. Prabhu, N.U. (1965). Queue and Inventories, John Wiley, New York.
16. Tanner, J.C. (1961). Biometrika, 48, 222–224.

Journal of the Indian Statistical Association
Vol.40, 2002, 13- 25

SOME PROPERTIES OF POISSON - LINDLEY AND ITS DERIVED DISTRIBUTIONS

M. Borah,
Tezpur University, Assam

R. A. Begum
Darrang College, Tezpur, Assam

Abstract

Poisson-Lindley is a well known one-parameter compound Poisson distribution, which has wide applications in the theory of accident proneness In this paper, Poisson - Lindley distribution has been studied and under two known forms of geometric infinite divisibility, the three parameters of Poisson - Lindley infinitely divisible distributions have been obtained. Further, some properties such as the recurrence relation for probabilities, factorial moments and cumulants of these distributions are also investigated. The problem of estimation of parameters, their sensitivity and the fitting of the distribution have been studied.

Key Words : *Lindley and Poisson-Lindley Distribution, Infinite divisibility, Recurrence Relation, Estimation and Goodness of fit.*

1 Introduction

A random variable is said to be infinitely divisible if and only if for every integer n, its characteristics function $\phi(t)$ can be expressed as

*Received (Revised version) : June 2002.

the n-th power of some other characteristics function $\phi_n(t)$, i.e.,

$$\phi(t) = \{\phi_n(t)\}^n. \tag{1}$$

The importance of the property of infinitely divisibility in modeling was stressed by Steutel (1973, 1983). see also the monograph by Steutel (1970), Pillai and Sandhya (1990) and Pillai and Jose (1994) who have studied some properties of geometrically infinitely divisible discrete distributions. Some properties of infinitely divisible discrete distributions have been studied by Johnson et. al. (1992). The infinitely divisible discrete distributions with probability generating functions (pgf's) of the form

$$G(t) = \frac{(1-\omega)g(t)}{1-\omega g(t)}, \qquad 0 < \omega < 1 \tag{2}$$

$$G(t) = \frac{\omega}{\omega + 1 - g(t)}, \qquad \omega > 0 \tag{3}$$

were studied by Klebanov et. al. (1984) under the name of geometrically infinitely divisible distributions. Here, $g(t)$ corresponds to the pgf of the component distribution used. The resulting distribution is always infinitely divisible, no matter whether the component distribution is infinitely divisible (ID) or not. Hence, the above two forms (2) and (3) are ID irrespective of whether $g(t)$ is ID or not. The basic geometrically infinitely divisible distributions given by Klebanov et al (1984) and Keilson and Steutel (1972), are closed under mixing and convolution. In case of continuous distribution, the characteristic function (cf) of the Normal distribution and in case of discrete distribution, the cf of Poisson can be easily put in the form (1), but it may not always be possible to express the cf of a distribution as in the form (1). Conditions for a discrete distribution to be infinitely divisible are discussed in Katti (1967), Warde and Katti (1971), Chang (1989) and Godambe and Patil (1975). In this paper, two new forms of infinitely divisible distributions have been derived by using Poisson-Lindley distribution as a component distribution with the help of (2) and (3). The pgf of these distributions based on the model (2) and (3) are denoted by the symbols GPL1, GPL2 respectively.

Poisson-Lindley Distribution :

Lindley (1958) derived a distribution known as Lindley distribution based on Bayes Theorem. Sankaran (1970) generalized Lindley distribution by mixing with Poisson distribution which is known as Poisson-Lindley distribution. The pgf of Poisson-Lindley distribution obtained by compounding the Poisson distribution with one due to Lindley may be defined as

$$g(t) = \frac{(\theta + 2 - z)\theta^2}{(\theta + 1)(\theta + 1 - z)^2}\ , \qquad \theta > 0$$

The probability recurrence relation may be expressed as

$$P_{r+1} = \frac{P_r(\theta + 3 + r)}{(\theta + 1)(\theta + 2 + r)}, \quad \text{for} \quad r \geq 0,$$

where $P_0 = \frac{\theta^2(\theta+2)}{(\theta+1)^3}$. Here $\mu = \frac{\theta+2}{\theta(\theta+1)}$, and

$$\sigma^2 = \frac{\theta^3 + 4\theta^2 + 6\theta + 2}{\theta^2(\theta + 1)^2},$$

be respectively the mean and variance of the distribution. The parameter θ of the distribution is related to μ by $\theta = \frac{-(\mu-1)+\sqrt{(\mu-1)^2+8\mu}}{2\mu}$, remembering θ to be positive, see Sankaran (1970).

This distribution has been studied by Sankaran (1970), with applications to data on errors and accidents. It is a special case of Bhattacharya's (1966) more complicated mixed Poisson distribution.

2 Generalized Poisson-Lindley (GPL1) Distribution

The probability generating function (pgf) of the generalized Poisson-Lindley distribution derived from the model (GPL1) can be given as

$$G(t) = \frac{(1-\omega)(\theta+2-t)\theta^2}{(\theta+1)(\theta+1-t)^2-\omega\theta^2(\theta+2-t)}, \quad 0<\omega<1, \quad \theta>0. \quad (1)$$

From (1), we get

$$P_{r+1} = A_1[\{2(\theta+1)^2-\omega\theta^2\}\ P_r-(\theta+1)\ P_{r-1}], \quad (2)$$

where

$$A_1 = \frac{1}{[(\theta+1)^3-\omega\theta^2(\theta+2)]}$$

$$P_0 = A_1(1-\omega)(\theta+2)\theta^2$$

and

$$P_1 = A_1[\{2(\theta+1)^2-\omega\theta^2\}\ P_0-(1-\omega)\theta^2$$

Factorial moment recurrence relation may be derived from its factorial moment generationg function (fmgf) given by

$$m(t) = \frac{(1-\omega)\theta^2(1+\theta+t)}{(1+\theta)(\theta-t)^2-\omega\theta^2(1+\theta-t)} \quad . \quad (3)$$

From (3), we get the factorial moment recurrence relation to be

$$\mu_{r+1} = \frac{(r+1)[\{2\theta(1+\theta)-\omega\theta^2\}\mu_{(r)}-r(1+\theta)\mu_{(r-1)}]}{(1+\theta)\theta^2(1-\omega)} \quad (4)$$

where

$$\mu_{(1)} = \frac{(\theta+2)}{\theta(1+\theta)(1-\omega)} \quad (5)$$

Mean and variance of the distribution are respectively given by

$$\mu = \frac{(\theta + 2)}{(\theta + 1)\theta(1 - \omega)} , \tag{6}$$

and

$$\sigma^2 = \frac{2\theta\mu + 3\theta^2\mu - 2\omega\theta^2\mu + \theta^2 - 2}{(1 + \theta)\theta^2(1 - \omega)} . \tag{7}$$

It is noted that the mean is less than the variance $\sigma^2 = C\mu + D$, when $C, D \neq 0$ where $C = \{2\theta + 3\theta^2 - 2\omega\theta^3\}/(1 + \theta)(1 - \omega)\theta^2$ and $D = (\theta^2 - 2)(1 + \theta)(1 - \omega)\theta^2$.

Estimation of Parameters

From equation (6), the parameter ω can be expressed as

$$\omega = 1 - \frac{(\theta + 2)}{(1 + \theta)\theta\mu}, \tag{8}$$

Again, from equation (6) and (7), the parameter ω is expressed in terms of θ, as

$$\omega = \frac{A\theta^2 + B\theta + 2}{2\mu\theta^2}, \tag{9}$$

where

$$A = 3\mu + 1 - \frac{\sigma^2}{\mu}, \qquad B = 2(\mu - \frac{\sigma^2}{\mu}).$$

Eliminating ω from equations (8) and (9), a functional equation for θ in terms of μ may be obtained as

$$f(\theta) = \theta^3 + 3\theta^2 + 2\theta - \frac{2}{T}, \tag{10}$$

where

$$T = \mu + 1 - \frac{\sigma^2}{\mu}.$$

The parameter θ may be estimated by the Newton-Raphson method.

Hence, ω may be estimated either by

$$\omega = \frac{\{\theta(1+\theta)\mu - (\theta+2)\}}{(\theta+1)\theta\mu} \quad , \qquad (11)$$

or

$$\omega = \frac{\{\theta(2+3\mu+\theta^2 - 2 - \theta(1+\theta)\sigma^2\}}{2\mu\theta^2 - \theta(1+\theta)\sigma^2}. \qquad (12)$$

where μ and σ^2 respectively denote the mean and variance of the distribution.

3 Generalized Poisson-Lindley (GPL2) Distribution

The pgf of Poisson-Lindley distribution derived from the model (GPL2) can be given as

$$G(t) = \frac{\omega(\theta+1)(\theta+1-t)^2}{(1+\omega)(\theta+1)(\theta+1-t)^2 - (\theta+2-t)\theta^2}, \quad \omega > 0, \theta > 0 \quad (1)$$

Differentiating (1) w.r.t. t and equating the coefficients of t^2 on both sides, the recurrence relation for probabilities may be obtained as

$$P_{r+1} = \frac{\{2(\omega+1)(\theta+1)^2 - \theta^2\}P_r - (\omega+1)(\theta+1)P_{r-1}}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)} \quad , \qquad (2)$$

for $r > 1$ where

$$P_0 = \frac{\omega(\theta+1)^3}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)} \quad ,$$

$$P_1 = \frac{\{2(\omega+1)(\theta+1)^2 - \theta^2\}P_0 - 2\omega(\theta+1)^2}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)} \quad ,$$

and

$$P_2 = \frac{\{2(\omega+1)(\theta+1)^2 - \theta^2\}P_1 - (\omega+1)(\theta+1)P_0 + \omega(\theta+1)}{(\omega+1)(\theta+1)^3 - \theta^2(\theta+2)} \quad .$$

Factorial moment recurrence relation may be derived from its fmgf as

$$m(t) = \frac{\omega(\theta+1)(\theta-t)^2}{(\omega+1)(1+\theta)(\theta-t)^2 - \theta^2(1+\theta-t)} \tag{3}$$

which gives

$$\mu_{(r+1)} = \frac{(r+1)\{2\theta(\omega+1)(\theta+1)-\theta^2\}\mu_{(r)} - r(\omega+1)(\theta+1)\mu_{(r-1)}}{(\theta+1)\omega\theta^2}$$

for $r > 1$ where

$$\mu_{(1)} = \frac{(\theta+2)}{\omega\theta(1+\theta)} \quad ,$$

and

$$\mu_{(2)} = \frac{2\{2\theta(\omega+1)(\theta+1)-\theta^2\}\mu_{(1)} - 2(\theta+1)}{(\theta+1)\omega\theta^2}$$

Hence the mean and variance of the distribution may be given as

$$\mu = \frac{(\theta+2)}{(\theta+1)\theta\omega}, \tag{4}$$

and

$$\sigma^2 = \frac{4\omega\theta\mu(\theta+1) + \theta\mu(\theta+2) + \theta^2 - 2}{\omega(\theta+1)\theta^2}. \tag{5}$$

Estimation of Parameters

From equation (5) the parameter ω can be expressed as

$$\omega = \frac{\theta+2}{\theta(\theta+1))\mu}, \tag{6}$$

Again from equations (5) and (6), the parameter ω can be expressed as

$$\omega = \frac{\theta^2\mu + 2\theta\mu + \theta^2 - 2}{\theta(\theta\sigma^2 - 4\mu)(\theta+1)}. \tag{7}$$

Eliminating ω from (7) and (8), θ may be obtained from

$$f(\theta) = A\theta^2 + B\theta + C \quad . \tag{8}$$

The parameter θ is estimated by noting that

$$\theta = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} \tag{9}$$

where $A = \sigma^2 - \mu^2 - \mu, B = 2(\sigma^2 - \mu^2 - 2\mu), C = -6\mu$. It may be noted that θ is positive.

Hence, ω may be estimated either from

$$\omega = \frac{(\theta + 2)}{\mu\theta(\theta + 1)}, \quad \text{or} \quad \omega = \frac{\mu\theta(\theta + 2) + \theta^2 - 2}{\theta(\theta + 1)(\theta\sigma^2 - 4\mu)} \tag{10}$$

4 Goodness of Fit

The negative Binomial, the Poisson and the Neyman's distributions are commonly used in ecological and biological problems, the Neyman's and Negative-Binomial distribution represent model in which the non-randomness is attributed to contagion. In this investigation, the parameter θ of generalized Poisson-Lindley distribution is estimated by Newton-Raphson method, whereas the parameter ω is estimated by the method of moments. In Table 1, we consider the data on the *Pyrausta nubilalis*, to which generalized Poisson distribution was fitted by Jain (1975). It is observed that generalized Poisson-Lindley distribution give better fit. The Distribution of corn borers is considered in Table II, to which the Negative-Binomial and the Neyman's type A distributions were fitted by Bliss and Fisher (1953). It will be seen that the data agree excellently with generalized Poisson-Lindley fit. The generalized Poisson-Lindley distributions, i.e. GPL1 and GPL2 are also fitted in Table III to some data collected by P. Garman (see Bliss et al, 1953) and the fit is found to be satisfactory.

Table 1: ***Pyrausta nubilslis in 1937*** **(data by Beall and Rescia 1940)**

No. of insects	observed	Poisson-Lindley (by Sankaran, 1970) $\hat{\theta} = 1.8082$	GPL1 $\hat{\theta} = 3.4556$ $\hat{\omega} = 0.5276$	GPL2 $\hat{\theta} = 12.74$ $\hat{\omega} = 0.1123$	Generalized-Poisson (by Jain, 1975)
0	33	31 49	31.86	33.10	32.46
1	12	14.16	13.84	12.49	13.45
2	6	6.09	5.92	5 50	5.60
3	3	2.54	2.52	2.49	2.42
4	1	1.04	1.06	1.13	1 08
5	1	0.42	0.80	1.29	0 97
Total	56	56 00	56.00	56.00	56.00
χ^2		0.6532	0.3743	0.0667	0.25

Table 2 : Corn Borers data of Beall (1940)

No. of insects	observed	Poisson-Lindley (by Sankaran, 1970) $\hat{\theta} = 1.0096$	GPL1 $\hat{\theta} = 3.4556$ $\hat{\omega} = 0.7611$	GPL2 $\hat{\theta} = 6.64$ $\hat{\omega} = 0.1148$
0	43	45.36	48.05	49.11
1	35	30.07	29.04	24.83
2	17	18.70	17.36	15.30
3	11	11.16	10.35	9.84
4	5	6.48	6.16	6.38
5	4	3.68	3.67	4.14
6	1	2.06	2.18	2.69
7	2	1.14	1.30	1.75
8	2	0.62	1.89	1.14
Total	120	120.00	120.00	120.00
χ^2		0.0667	1.7289	1.7251

Table 3 : Count of the number of European red mites on apple leaves (data of Bliss et al 1953)

No of mites per leaf	Leaves (observed)	Poisson-Lindley (by Sankaran, 1970) $\hat{\theta}$=1.258	GPL1 $\hat{\theta}=1\ 3921$ $\hat{\omega}=0.1117$	GPL2 $\hat{\theta}=8\ 58$ $\hat{\omega}=0.1022$	Jain and Consul (1971)
0	70	67.19	67.62	70 89	71 48
1	38	38.89	38.68	33 35	33 98
2	17	21.26	21.04	18.70	19.80
3	10	11.21	11 09	10.84	11 59
4	9	5.76	5 7	6.73	6 57
5	3	2.90	2.92	3 69	3 55
6	2	1.44	1.47	2.15	1.80
7	1	0 71	0.74	1.26	0.84
8	0	0.34	0.71	0.73	0 39
Total	150	150.00	150.00	150 00	150 00
χ^2		3 0136	2.8491	2.4433	2.07

References

Beall, G. and Rescia (1940) The fit and significance of contagious distributions when applied to observations on Larval insects, *Ecology*, **21**, 460-474.

Bhattacharya, S. K. (1966) Confluent Hypergeometric distributions of discrete and continuous type with applications to accident proneness, *Bulletin of the Calcutta Statistical Association*, **15**, 20-31.

Bliss, C. I. and Fisher R. A. (1953) Fitting the negative binomial distribution to biological data and note on the efficient fitting of the binomial, *Biometrics*, **9**, 176-200.

Chang, D. K. (1989) On infinitely divisible discrete distribution, *Utilitas Mathematica*, **36**, 215-217.

Consul, P. C. and Shenton, L. R. (1972). Use of Lagrange expansion for generating discrete generalized probability distribution, *SIAM J. Appl. Math*, **23**, 239-248.

Feller, W. (1957) *An Introduction to Probability Theory and Its Applications,* Vol I, (Third Edition), New York, Wiley.

Fisz, M. (1962): Infinitely divisible distributions: recent results and applications, *Ann. Math. Statist.*, **33**, 68-84.

Godambe, A. V. and Patil, G. P. (1975) Some characterizations involving additivity and infinite divisibility and their applications to Poisson mixtures and Poisson sums, *Statistical Distributions in Scientific Work*, **3**, Characterizations and Applications, (G. P. Patil, S. Kotz, and J. K. Ord (editors),) 339-351. Dordrecht: Reidel.

Goldie Charles (1967): A class of infinitely divisible random variables, *Proc. Camb. Phil. Soc.*

Jain, G. C. (1975) On power series distributions associated with Lagrange expansion, *Bio. Z.* **17**, 85-97.

Johnson, N. L. Kotz, S. and Kemp, A. W. (1992): *Univariate Discrete Distributions,* Second Ed. John Wiley & Sons, Inc.

Katti, S. K. (1967) Infinite divisibility of integer valued random variables, *Ann. Math. Statist.*, **38**, 1306-1308.

Keilson, J. and Steutel, F. W. (1972): Families of infinitely divisible distributions closed under mixing and convolution, *Ann. Math. Statist.*, **43 (1)**, 242-250.

Klebanov, L. B. Maniya, G. M. and Melamed, I. A. (1984) A problem of Zolotarev and analogs of infinitely divisible and stable distributions in a scheme of summing a random number of random variables. *Theor. Prob. Applicn.* **4**, 791-794.

Lindley, D. V. (1958) Fiducial distribution and Bayes theorem, *Journal of The Royal Statistical Society, Series* B, **20**, 102-107.

Pillai, R. N. (1985) Semi α-Laplace distribution *Commun. Statist., Theory and Methods*, **14**, 991-1000.

Pillai, R. N. (1990): Harmonic mixtures and geometric infinite divisibility, *J. Indian Statist. Assoc.* **28**, 87-98.

Pillai, R. N. and Sandhya, E. (1990) Distributions with complete monotone derivative and geometric infinite divisibility, *Adv. Appl. Probab.*, **22**, 251-754.

Pillai, R. N. and Jose, K. K. (1994) Geometric infinite divisibility and autoregressive time series Modeling, *Proceeding of the Third Ramanujan Symposium on Stochastic Processes and their Applications,* University of Madras, Madras 17-19 January, 1994, 81-87.

Sandhya, E. (1991): *Geometric Infinite Divisibility and Applications*, Ph.D. thesis (unpublished), submitted to the University of Kerala, January, 1991.

Sankaran, M. (1970): The Discrete Poisson-Lindley distribution, *Biometrics*, **26**, 145-149.

Steutel, F. W. (1968): A class of Infinitely Divisible Mixtures, *Ann. Math. Statist.*, **39**, 1153-1157.

Steutel, F. W. (1970): *Preservation of infinite divisibility under mixing*, Mathematical Centre Tract, **33**, Amsterdam.

Steutel, F. W. (1973): Some recent results in infinite divisibility, *Stochastic Processes Appl.*, **1**, 125-143.

Steutel, F. W. (1983) Infinite divisibility, *Encyclopedia of Statistical Sciences,* 4, S. Kotz, N. L. Johnson and C. B. Read (editors), 114-116, New York, Wiley.

Steutel F. W. (1990) The set of geometrically infinitely divisible distributions, *Technische Universiteit Eindhoven Memorandum* COSOR.

Warde, W. D. and Katti, S. K. (1971) Infinite divisibility of discrete distribution II, *Ann. Math. Statist.*, **42**, 1088-1090.

M. Borah,
Department of Mathematical Sciences,
Tezpur University,
Napaam, Tezpur - 784028
Assam, India

R. A. Begum,
Department of Statistics,
Darrang College,
Tezpur - 784001,
Assam, India

CERTAIN INFINITELY DIVISIBLE DISCRETE PROBABILITY DISTRIBUTIONS AND ITS APPLICATIONS.

R. A. Begum., M. Borah.

@Department of Statistics.Darrang College, Tezpur, Assam.
*Department Mathematical Sciences. Tezpur University, Napaam, Tezpur, Assam.

Abstract

Elementary infinitely divisible distributions which are formulated on the basis of simple models seen to be inadequate to describe the situations which may occur in a number of phenomena. In the last few years a number of various infinitely divisible distributions have been derived. In this paper two forms of infinitely divisible discrete distributions have been studied. The recurrence relations for their probabilities and factorial moments are also investigated. As the method of maximum likelihood will be very cumbersome, some other adhoc methods have been also used to estimate the parameters. A few sets of reported data have been considered for the fitting of the distributions, and the fits are compared with that obtained with other distributions..

Key words : Infinitely divisible distributions, Recurrence Relations, Factorial moments.

Introduction

The theory of infinitely divisible distributions, developed primarily during the period from (1920), to (1950), has played a very important role in the solution of limit problems for sums of independent random variables. A full account of this theory and its applications as it had been developed by the late 40's, were presented in the monographs of Levy (1937), Gnedenko and Kolmogorov (1968), and Loeve (1960). In the last few years research in this field have shown that infinite divisible laws plays a significant role in a variety of problems of probability theory and has been carried out along many lines. Numerous new results have been obtained and entirely new applications have been found. In 1961, Mark Fisz gave a survey on recent developments in infinite divisibility only for the distributions on the real line. F. W. Steutel (1972), also surveyed on some recent results in infinitely divisibility.

1 Meaning Of Infinite Divisibility

A random variable is said to be infinitely divisible if and only if (iff) it has a characteristic function (cf), $\phi(t)$ that can be represented for every positive integer 'n' as the nth power of some cf $\phi_n(t)$, ie

$$\phi(t) = \{\phi_n(t)\}^n . \tag{1.1}$$

In case of continuous distribution the cf of Normal distribution and in case of discrete distribution the cf of Poisson distribution can be easily put in the form (1.1), but it may not be always possible to express the cf of an infinite divisible distribution as in the form given in (1.1). Conditions for a discrete distribution to be infinitely divisible are discussed in Katti (1967), Warde and Katti (1971), and Chang (1989).

In non-technical terms what it means is that, if there exist independently and identically distributed random variables $x_{n1}, x_{n2}, \ldots, x_{nn}$, such that the distribution of $x = x_{n1} + x_{n2} + \ldots + x_{nn}$ is the same as the given distribution.

In agriculture, we usually start with a large plot of land and we subdivide it into n-parts and assume that the yields in the n-parts are independent and since they are usually of equal size, they have identical distribution .Thus, if the distribution of yield over the entire plot is such as to permit this subdivision, then it needs to be divisible. The question of infiniteness in divisibility is really a theoretical idealization of the fact that the plot is subdivided quite extensively. Same is true in biology where a blood sample taken from a patient is subdivided in to a number of parts to test for different diseases. In testing for reliability of equipment over a period of time, we take an interval and subdivide the interval into many parts and act as though the distribution of failure in each part is independently and identically distributed. This again leads to infinite divisibility. Hence testing of infinite divisibility is an interesting part of statistical inference.

2 Condition For a Distribution To Be Infinitely Divisible

Suppose that $p_0, p_1, p_2, \ldots$ are probabilities of $0,1,2,3,\ldots,$ with $p_0 \neq 0, p_1 \neq 0$. Then according to Katti (1967) the necessary and sufficient condition for a distribution to be infinitely divisible is that for each value of i

$$\pi_i = \frac{iP_i}{P_0} - \sum_{j=1}^{i-1} \pi_{i-j} \frac{p_j}{p_0} \geq 0 , \qquad \text{for } i=1,2,.. \tag{2.1}$$

Note that for a given distribution function, one can numerically compute a number of p_i / p_0, to see if they are positive and if they are, then one can use this information along with his algebraic calculation to generate an inductive proof of infinitely divisibility.

The following two classes of infinitely divisible characteristic functions are of special interest.

i) Compound-Poisson Distribution: A distribution with characteristics function(cf) of the form

$$\phi(t) = e^{\lambda\{g(t)-1\}} , \quad \lambda \geq 0 \tag{2.2}$$

in which g(t) is also a cf, is always infinitely divisible. This is also known as Poisson-stopped-sum distribution. They arise as the distribution of the sum of a Poisson number of independently and identically distributed random variables with cf Because of their infinite divisibility these distributions have very great importance in discrete distribution theory. They are known by different names. Feller (1943) used the term generalized Poisson, Galliher et al. (1959) and Kemp (1967) called them stuttering Poisson. The term compound Poisson was used by Feller (1950,1957,1968) and Lloyd (1980).

ii) Compound-geometric Distribution: A distribution with cf of the form

$$\phi(t) = \frac{\lambda}{\lambda + 1 - g(t)} , \quad \lambda > 0 \tag{2.3}$$

in which g(t) is also a cf is always infinitely divisible.

It is well known that all infinitely divisible (inf. div.) distributions are generated by a stochastic processes, more specifically by processes with stationary independent increments There is some applied processes, however, that give rise to classes of infinitely divisible distributions Certain families of probability distribution functions maintain their infinitely divisibility under repeated mixing and convolution Godambe and Patil (1975) consider a mixture of Poisson distributions where the mixing distributions has non-negative support The importance of the property of infinite divisibility in modeling was stressed by Steutel (1973,1983); see also the monograph by Steutel (1970) Pillai (1990), Pillai et al (1990,1994). Some properties of infinitely divisible discrete distributions are given by Johnson, et al (1992).

It will turn out that all infinitely divisible distributions are limits of compound Poisson distributions. There is a number of a number of methods to construct new infinite divisible distributions from given one, the best known are convolution and compounding. A general compounding theorem is due to Feller (1957)

3 Geometric Infinite Divisibility

According to Klebanov (1984) a random variable X is said to be of geometrically infinitely divisible (gid) if there exists an independently identically distributed sequence of random variables $X_j^{(p)}$ j=1,2,----N_p such that for any $p \in (0,1)$

$$X \overset{d}{=} \sum_{j=1}^{N_p} X_j^{(p)}$$

where $P(N_p=k)=p(1-p)^{k-1}$, k= 1,2,..... and X, N_p and $X_j^{(p)}$ are independent $\overset{d}{=}$ stands for equality of distributions. The geometrically infinitely divisible distributions form a sub-class of infinitely divisible distributions. A more detailed description of geometrically infinitely divisible random variable is based on the fact that, a random variable Y with cf f(t) is geometrically infinitely divisible if and only if

$\phi(t) = \exp\{1 - \frac{1}{f(t)}\}$, is an infinitely divisible characteristic function.

Infinitely divisible discrete distributions with pgf's of the form

$$G(t) = \frac{(1-\omega)g(t)}{1-\omega g(t)}, \qquad 0 < \omega < 1$$

(3.1)

$$G(t) = \frac{\omega}{\omega + 1 - g(t)}, \qquad \omega > 0$$

(3.2)

respectively, were studied by Klebanov, Maniya and Melamed (1984), Steutel (1990) under the name geometrically infinitely divisible distributions. In the above two forms (3.1) AND (3.2), g(t) is the pgf of the component distribution. It is very important to note that the resultant distribution is always infinitely divisible no ma ter whether the component distribution is infinitely divisible or not.

Here, in this paper, taking pgf of Poisson, Poisson-Lindley, and negative binomial distributions, as the component distribution the above two forms of geometric infinitely divisibility have been studied. Further, the recurrence relations for probabilities and factorial moments are also investigated. The problems of estimation and the fitting of the distributions have also been considered.

Poisson Distribution

A random variable x is said to follow Poisson distribution, if it assumes only non negative values of x having the following probability function

$$P(x=k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k=0,1,2,\ldots, \quad \lambda > 0. \qquad (3.3)$$

with pgf g(t)= $e^{\lambda(t-1)}$, where λ is the only parameter of the distribution. The distribution is a power series distribution with infinite non-negative integer support. It belongs to the exponential family of distributions. Poisson (1837) published the derivation of the distribution by considering the limiting forms of the binomial distribution, which bears his name.

Negative-Binomial Distribution

A random variable x is said to follow Negative-binomial distribution with parameters k and p if its probability mass function is given by

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

(3.4)

where $0<p<1$, $k=1,2,3,\ldots$ It has the pgf $g(t) = (1-q)^k (1-qt)^{-k}$ $0<t<1$ and $p+q=1$.

Poisson-Lindley Distribution

Poisson-Lindley is a one parameter well known compound Poisson distribution, which has wide applications to the theory of accident proneness.

Lindley (1958) derived a distribution known as Lindley distribution based on Baye's theorem. Sankaran (1970) generalized Lindley distribution by mixing with Poisson distribution which is known as Poisson-Lindley distribution. The pgf of Poisson-Lindley distribution obtained by compounding the Poisson distribution with one due to Lindley. may be defined as

$$g(t)=\frac{(\theta+2-z)\theta^2}{(\theta+1)(\theta+1-z)^2}, \qquad \theta>0 \tag{3.5}$$

The Probability recurrence relation may be expressed as

$$P_{r+1}=\frac{P_r(\theta+3+r)}{(\theta+1)(\theta+2+r)}, for \quad r\geq 0 \tag{3.6}$$

where $$P_0=\frac{\theta^2(\theta+2)}{(\theta+1)^3},$$

$$\mu=\frac{\theta+2}{\theta(\theta+1)}, \text{ and } \sigma^2=\frac{\theta^3+4\theta^2+6\theta+2}{\theta^2(\theta+1)^2}$$

denote respectively the mean and the variance of the distribution. The parameter of the distribution is estimated as below

$$\theta=\frac{-(\mu-1)+\sqrt{(\mu-1)^2+8\mu}}{2\mu},$$

remembering θ to be positive. See Sankaran (1970). This distribution was studied by Sankaran (1970), with applications to errors and accidents. It is a special case of Bhattacharya,s (1966) more complicated mixed Poisson distribution.

4 Properties Of Poisson Mixing Infinitely Divisible Distribution

a) The pgf of Poisson mixing infinite divisible distribution derived from model (3.1) will be obtained as

$$G(t) = \frac{(1-\omega)e^{\lambda(t-1)}}{1-\omega e^{\lambda(t-1)}}, \quad \text{for } \omega > 0,\ \lambda>0, \qquad (4.1)$$

Its probability recurrence relation

$$P_{r+1} = \frac{e^{-\lambda}}{1-\omega e^{-\lambda}} \left\{ \frac{\lambda^{r+1}(1-\omega+P_0\omega)}{(r+1)!} + \omega \sum_{j=1}^{r} \frac{\lambda^j P_{r-j+1}}{j!} \right\}, \ r \geq 0 \qquad (4.2)$$

may be derived by expanding the pgf (4.1) and equating the coefficients of t^r on both sides, where

$$P_0 = \frac{(1-\omega)e^{-\lambda}}{1-\omega e^{-\lambda}}.$$

Now, putting r = 0, 1, 2, ... respectively in equation (4.2), we have

$$P_1 = \lambda P_0 \left\{ 1 + \frac{\omega P_0}{1-\omega} \right\},$$

$$P_2 = \lambda P_1 \left\{ \frac{1}{2} + \frac{\omega P_0}{1-\omega} \right\},$$

$$P_3 = P_2 \left\{ \frac{P_2}{P_1} - \frac{\lambda^2 P_1}{12 P_2} \right\}, \qquad \text{etc.}$$

Factorial moments recurrence relation

$$\mu_{(r+1)} = \frac{(r+1)\omega}{1-\omega} \sum_{j=1}^{r} \binom{r}{r-j+1} \frac{\lambda^{r-j+1}}{j} \mu_{(j)} + \frac{\lambda^{r+1}}{1-\omega}, \qquad (4.3)$$

may be derived from the factorial moment generating function (fmgf) corresponding to the pgf (4.1), where $\mu_{(1)} = \frac{\lambda}{1-\omega}$. The higher other factorial moments may be obtained by putting r = 1,2, 3 etc. in (4.3) respectively as

$$\mu_{(2)} = \mu_{(1)}\{\lambda + 2\,\omega\,\mu_{(1)}\}$$
$$\mu_{(3)} = 3\omega\mu_{(1)}\{\lambda\mu(1) + \mu_{(2)}\} + \lambda^2\mu_{(1)}$$
$$\mu_{(4)} = 4\omega\mu_{(1)}\{\mu_{(3)} + \frac{3}{2}\lambda\,\mu_{(2)} + \lambda^2\mu_{(1)}\} + \lambda^3\mu_{(1)},$$

where mean $\mu = \dfrac{\lambda}{1-\omega}$, variance $= \mu\,\{(\lambda+1) + \mu(2\,\omega - 1)\}$.

Estimation Of Parameters

For the Poisson mixing inf. dist. (4.1), using the first two central moments the parameters are estimated as

$$\omega = \frac{\sigma^2 - \mu}{\mu^2}$$

$$\lambda = \mu + 1 - \frac{\sigma^2}{\mu}$$

where μ and σ^2 are the mean and variance of the distribution (4.1).

b) The pgf of Poisson mixing infinite divisible distribution derived from model (3.2), will be given by

$$G(t) = \frac{\omega}{\omega + 1 - e^{\lambda(t-1)}}, \quad \text{for } 0<\omega<1,\ \lambda > 0. \tag{4.4}$$

Its Probability recurrence relation

$$P_{r+1} = F_3\left[\frac{\lambda^{r+1}P_0}{(r+1)!} + \sum_{j=1}^{r}\frac{\lambda^j}{j!}P_{r-j+1}\right] \tag{4.5}$$

may be obtained by expanding (4 4) with respect to 't' and equating the coefficients of 'tr' on both sides, where

$$P_0 = \frac{\omega}{\omega + 1 - e^{-\lambda}},$$

and
$$F_3 = \frac{e^{-\lambda}}{\omega + 1 - e^{-\lambda}}.$$

Putting r = 1,2, ... respectively, in the equation (4.5) the higher order probabilities may be obtained as

$$P_1 = \lambda \xi_3 P_0,$$

$$P_2 = \lambda P_1 \left\{ \frac{1}{2} + e^{-\lambda} \frac{P_0}{\omega} \right\},$$

$$P_3 = P_2 \left\{ \frac{P_2}{P_1} - \frac{\lambda^2 P_1}{12 P_2} \right\}$$

Similarly, the factorial moments recurrence relation

$$\mu_{(r+1)} = \frac{r+1}{\omega} \sum_{j=1}^{r} \binom{r}{j-1} \frac{\lambda^j}{j} \mu_{(r-j+1)} + \frac{\lambda^{r+1}}{\omega} \qquad (4.6)$$

may be obtained from the fmgf of the derived distribution (4.4), where

$$\mu_{(1)} = \frac{\lambda}{\omega}.$$

Hence, the higher order factorial moments may be obtained putting r = 1, 2, 3 respectively in the equation (4.6) as

$$\mu_{(2)} = \mu_{(1)} \{\lambda + 2\mu_{(1)}\}$$

$$\mu_{(3)} = \mu_{(1)} \{\lambda^2 + 3(\mu_{(2)} + \lambda\mu_{(1)})\}$$

$$\mu_{(4)} = 4\mu_{(1)} \left\{ \mu_{(3)} + \frac{3}{2}\lambda\mu_{(2)} + \lambda^2 \mu_{91)} \right\} + \lambda^3 \mu_{(1)}$$

mean $\mu = \dfrac{\lambda}{\omega}$, and

variance $\sigma^2 = \mu(1 + \mu + \lambda)$

Estimation of Parameters

Using the first two moments, the parameters ω and λ of the distribution (4.4) may be estimated as

$$\omega = \frac{\sigma^2 - \mu - \mu^2}{\mu^2}$$

$$\lambda = \frac{\sigma^2}{\mu} - \mu - 1.$$

$$m(t) = \frac{(1-\omega)\theta^2(1+\theta+t)}{(1+\theta)(\theta-t)^2 - \omega\theta^2(1+\theta-t)},$$

(5.3)

Differentiating (5.3) w.r.t. 't' and equating the coefficients of t'/r' on both sides, we get the factorial moment recurrence relation as

$$\mu_{(r+1)} = \frac{(r+1)[\{2\theta(1+\theta)-\omega\theta^2\}\mu_{(r)} - r(1+\theta)\mu_{(r-1)}]}{(1+\theta)\theta^2(1-\omega)}, \tag{5.4}$$

where $\mu_{(1)} = \dfrac{(\theta+2)}{\theta(1+\theta)(1-\omega)}$.

Putting r=1,2,3,---, respectively in equation (5.4) we have

$$\mu_{(2)} = \frac{2\{2\theta(\theta+1)-\omega\theta^2\}\mu_{(1)} - 2(\theta+1)}{(\theta+1)(1-\omega)\theta^2},$$

$$\mu_{(3)} = \frac{3\{2\theta(\theta+1)-\omega\theta^2\}\mu_{(2)} - 6(\theta+1)\mu_{(1)}}{(\theta+1)(1-\omega)\theta^2}$$

$$\mu_{(4)} = \frac{4\{2\theta(\theta+1)-\omega\theta^2\}\mu_{(2)} - 12(\theta+1)\mu_{(1)}}{(\theta+1)(1-\omega)\theta^2}$$

where $\mu_{(r)}$ denotes the r[th] order factorial moment of the distribution. Mean and variance of the distribution may be given as

$$\mu = \frac{(\theta+2)}{(\theta+1)\theta(1-\omega)} \tag{5.5}$$

and $\sigma^2 = \dfrac{2\theta\mu + 3\theta^2\mu - 2\omega\theta^2\mu + \theta^2 - 2}{(1+\theta)\theta^2(1-\omega)}$ respectively. (5.6)

It is noted that the mean is less than the variance $\sigma^2 = C\mu + D$, when C D $\succ$ 0,
where

$$C = \frac{\{2\theta + 3\theta^2 - 2\omega\theta^3\}}{(1+\theta)(1-\omega)\theta^2} \quad \text{and} \quad D = (\theta^2 - 2)(1+\theta)(1-\omega)\theta^2$$

Estimation of Parameters

The parameter of the distribution (5.1) can be expressed as

$$\omega = 1 - \frac{(\theta + 2)}{(1+\theta)\theta\mu}, \tag{5.7}$$

from the equation (5.5) of the distribution (5.1)
Similarly the parameter of the distribution (5.1) can be expressed in terms of θ, using the equation (5.6)

$$\omega = \frac{A\theta^2 + B\theta + 2}{2\mu\theta^2}, \tag{5.8}$$

where A= $3\mu + 1 - \frac{\sigma^2}{\mu}$, $B = 2\left(\mu - \frac{\sigma^2}{\mu}\right)$.

Eliminating ω from equations (5.7) and (5.8), the function of θ may be obtained as

$$f(\theta) = \theta^3 + 3\theta^2 + 2\theta - \frac{2}{T}, \tag{5.9}$$

where $T = \mu + 1 - \frac{\delta^2}{\mu}$

Now parameter θ can be estimated by using Newton-Raphson method

$$\theta_{r+1} = \theta_r - \frac{f(\theta_r)}{f'(\theta_r)}, \quad \text{for } r=1,2,--- \tag{5.10}$$

by selecting an appropriate guess values of θ, ie. θ_0.
When θ is known, the parameter w of the distribution (5.1) m[illegible] esti-mated using either mean or mean and variance of the distribu[illegible]
Therefore ω, may be estimated either by

$$\omega = \frac{\{\theta(1+\theta)\mu - (\theta + 2)\}}{(\theta+1)\theta\mu} \tag{5.11}$$

or $$\omega = \frac{\{\theta(2+3\mu+\theta^2-2-\theta(1+\theta)\delta^2\}}{2\mu\theta^2-\theta(1+\theta)\delta^2}.$$

(5.12)

where μ and σ^2denote the mean and variance of the distribution.

b) The pgf of Poisson-Lindley distribution derived from the model (3.2) may be given as

$$G(t) = \frac{\omega(\theta+1)(\theta+1-t)^2}{(1+\omega)(\theta+1)(\theta+1-t)^2-(\theta+2-t)\theta^2}, \omega>0, \theta>0 \tag{5.13}$$

Differentiating (5.13) w.r.t. 't' and equating the coefficients of t^r on both sides the recurrence relation for probabilities may be obtained as

$$P_{r+1} = \frac{\{2(\omega+1)(\theta+1)^2-\theta^2\}P_r-(\omega+1)(\theta+1)P_{r-1}}{(\omega+1)(\theta+1)^3-\theta^2(\theta+2)}, \quad r>1 \tag{5.14}$$

where $$P_0 = \frac{\omega(\theta+1)^3}{(\omega+1)(\theta+1)^3-\theta^2(\theta+2)},$$

$$P_1 = \frac{\{2(\omega+1)(\theta+1)^2-\theta^2\}P_0-2\omega(\theta+1)^2}{(\omega+1)(\theta+1)^3-\theta^2(\theta+2)},$$

and $$P_2 = \frac{\{2(\omega+1)(\theta+1)^2-\theta^2\}P_1-(\omega+1)(\theta+1)P_0+\omega(\theta+1)}{(\omega+1)(\theta+1)^3-\theta^2(\theta+2)}$$

Putting r=2,3 --- respectively in the equation (5.14) we have

$$P_3 = \frac{\{2(\omega+1)(\theta+1)^2-\theta^2\}P_2-(\omega+1)(\theta+1)P_1}{(\omega+1)(\theta+1)^3-\theta^2(\theta+2)},$$

$$P_4 = \frac{\{2(\omega+1)(\theta+1)^2-\theta^2\}P_3-(\omega+1)(\theta+1)P_2}{(\omega+1)(\theta+1)^3-\theta^2(\theta+2)}, \text{ etc.}$$

The fmgf of the distribution (5.13) may be written as

$$m(t) = \frac{\omega(\theta+1)(\theta-t)^2}{(\omega+1)(1+\theta)(\theta-t)^2-\theta^2(1+\theta-t)},$$

(5.15)

The factorial moments recurrence relation may be obtained by differentiating (5.15) w.r.t. 't' and equating the coefficients of $t^r/r!$ on both sides,

$$\mu_{(r+1)} = \frac{(r+1)\{2\theta(\omega+1)(\theta+1)-\theta^2\}\mu_{(r)} - r(\omega+1)(\theta+1)\mu_{(r-1)}}{(\theta+1)\omega\theta^2}, \quad r > 1 \qquad (5.16)$$

where $\mu_{(1)} = \dfrac{(\theta+2)}{\omega\theta(1+\theta)}$,

$$\mu_{(2)} = \frac{2\{2\theta(\omega+1)(\theta+1)-\theta^2\}\mu_{(1)} - 2(\theta+1)}{(\theta+1)\omega\theta^2}$$

Putting r=2,3,---, respectively in equation (5.16) we have

$$\mu_{(3)} = \frac{3\{2\theta(\omega+1)(\theta+1)-\theta^2\}\mu_{(2)} - 6(\omega+1)(\theta+1)\mu_{(1)}}{(\theta+1)\omega\theta^2}.$$

$$\mu_{(4)} = \frac{4\{2\theta(\omega+1)(\theta+1)-\theta^2\}\mu_{(3)} - 12(\omega+1)(\theta+1)\mu_{(2)}}{(\theta+1)\omega\theta^2}, \quad \text{etc.}$$

Hence, the mean and variance of the distribution

$$\mu = \frac{(\theta+2)}{(\theta+1)\theta\omega}, \qquad (5.17)$$

and $\sigma^2 = \dfrac{4\omega\theta\mu(\theta+1)+\theta\mu(\theta+2)+\theta^2-2}{\omega(\theta+1)\theta^2}$, respectively (5.18)

Estimation of Parameters

From equation (5.17) the parameter ω can be expressed as

$$\omega = \frac{\theta+2}{\theta(\theta+1)\mu}, \qquad (5.19)$$

Again from equation (5.18) the parameter can be expressed as

$$\omega = \frac{\theta^2\mu + 2\theta\mu + \theta^2 - 2}{\theta(\theta\sigma^2 - 4\mu)(\theta + 1)}.$$

(5.20)

Eliminating from (5.19) and (5.20), the function of θ may be obtained as

$$f(\theta) = A\theta^2 + B\theta + C, \qquad (5.21)$$

The parameter θ is estimated as

$$\theta = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$, assuming θ to be positive,

where $A = \sigma^2 - \mu^2 - \mu$, $B = 2(\sigma^2 - \mu^2 - 2\mu)$, and $C = -6\mu$.

Hence, w may be estimated either in terms of mean or in terms of mean and variance respectively as

$$\omega = \frac{(\theta + 2)}{\mu\theta(\theta + 1)}, \quad \omega = \frac{\mu\theta(\theta + 2) + \theta^2 - 2}{\theta(\theta + 1)(\theta\delta^2 - 4\mu)}.$$

1.6 Properties Of Negative-Binomial Mixing Inf. Div. Distribution

a) The pgf of Negative Binomial mixing Infinitely divisible distribution derived from equation (3.1) will be given by

$$G(t) = \frac{(1-\omega)(1-a)^n(1-at)^{-n}}{1-\omega(1-a)^n(1-at)^{-n}}, \text{ wher } w > 0, \quad a>0, \qquad (6.1)$$

Probability recurrence relation

$$P_{r+1} = T_1\left\{ {}^{n+r}C_{r+1}a^{r+1}(1 - \omega + \omega P_0) + \omega \sum_{j=1}^{r} {}^{n+r-j}C_{r-j+1}a^{r-j+1}P_j \right\}, r \geq 0 \quad (6.2)$$

may be obtained by expanding (6.1) the pgf of negative binomial mixing distribution and equating the coefficients of tr on both sides,

where $P_0 = (1-\omega)T_1$,

and $$T_1 = \frac{(1-a)^n}{1-\omega(1-a)^r},$$

Putting , r=0, 1, 2, ... respectively, in equation (6.2),we have

$$P_1 = naP_0\left\{1 + \frac{\omega P_0}{(1-\omega)}\right\},$$

$$P_2 = aP_1\left\{\frac{P_1}{P_0} - \frac{(n-1)}{2}\right\}, \text{ etc.}$$

Factorial moments recurrence relation

$$\mu_{(r+1)} = K_1\left\{b^{r+1}({}^{n+r}C_{r+1}) + \omega\sum_{j=1}^{r}({}^{n+r-j}C_{r-j+1})b^{r-j+1}\frac{\mu_j}{j!}\right\}, r \geq 1 \qquad (6.3)$$

where $$K_1 = \frac{(r+1)!}{(1-\omega)}$$

may be derived from fmgf corresponding to the pgf (6.1) ,

where $$b = \frac{a}{(1-a)},$$

Hence the first three factorial moments of the distribution (6.1) may be written as

$$\mu_{(1)} = \frac{nb}{1-\omega}$$

$$\mu_{(2)} = (n+1)b\mu_{(1)} + 2\omega\mu_{(1)}^2$$

$$\mu_{(3)} = b^2(n+1)(n+2)\mu_{(1)} + 6\omega b(n+1)\mu_{(1)}^2 + 6\omega^2\mu_{(1)}^3$$

where Mean $$\mu = \frac{nb}{(1-\omega)} \qquad (6.4)$$

and Variance $$\sigma^2 = (2\omega-1)\mu^2(n+1)b\mu + \mu \qquad (6.5)$$

Estimation of Parameters

Using the first three factorial moments of negative binomial mixing inf. div distribution (6.1) we estimate the parameters n, a and ω as

$$n = \frac{M - 2N^2}{M - N^2},$$

$$b = \frac{M - X^2}{\mu X},$$

$$\omega = 1 - \frac{M - 2N^2}{\mu^2 X},$$

and $$a = \frac{M - N^2}{\mu N + M - N^2}.$$

writing $M = \mu_{(3)}\mu_{(1)} - 6\mu_{(1)}^{2}(\lambda + \mu_{(1)})$,

and $N = \mu_{(2)} - 2\mu^2_{(1)}$,

Where μ is the mean of the distribution.

b) The pgf of Negative Binomial mixing Infinitely divisible distribution derived from equation (3.2) will be given by

$$G(t) = \frac{\omega}{\omega + 1 - (1-a)^n (1-at)^{-n}}, \qquad 0 < w < 1 \tag{6.6}$$

Probability Recurrence relation

$$P_{r+1} = T_2 \left\{ {}^{n+r}C_{r+1} a^{r+1} P_0 + \sum_{j=0}^{r} {}^{n+r-j}C_{r-j+1} a^{r-j+1} P_j \right\}, \tag{6.7}$$

may be obtaind by expanding (6.6) the pgf of negative binomial mixing distribution and equating the coefficients of tr on both sides, where

where $$T_2 = \frac{(1-a)^n}{\omega + 1 - (1-a)^n},$$

and $$P_0 = \frac{\omega}{\omega + 1 - (1-a)^n}.$$

Putting r=0, 1, 2,... respectively in (6.7), the higher order probabilities may be computed as

$$P_1 = na(1-a)^n \frac{P_0^2}{\omega},$$

where $N = \mu_{(2)} - 2\mu^2_{(1)}$,

$O = M\mu_{(1)} - X\mu_{(2)}$,

and $M = \mu_{(3)} - \mu_{(1)}(N + 3\mu_{(2)})$,

Goodness of fit

One of the problems in fitting distributions is that of the parameters. Of all the procedures of estimating the parameters, the method of moments is perhaps the oldest and the simplest. In many cases it leads to tractable operations. The method of maximum likelihood is considered to be more accurate for fitting a probability distribution on given data, but it involves much more computational work than is required by the method of moments. It is mainly from this reason, moment estimators are used.

Therefore, using the method of moments we estimate the parameters and of Poisson mixing inf. div. distribution and in case of the negative-binomial mixing inf. div. distribution we used the first three factorial moments for estimating the parameters a, w and n, for each of the above two forms under (3.1) and (3.2).

In this investigation the parameter q of generalized Poisson-Lindley distribution is estimated by Newton-Raphson method, whereas the parameter w is estimated by the method of moments. The Negative-binomial, the Poisson and the Poisson-Lindley distributions are commonly used in ecological and biological problems. Therefore we consider the numerical data of Haemocytometer Counts of Yeast Cells and the data of counts of European red mites on apple leaves for our comparison.

The distributions (4.1) and (4.4) have been fitted to some available data as shown in Table 1 and 2. It is clear that there is some improvement, however small it may be, in fitting these distributions over the other distributions, considered earlier. The fitting of these distributions as indicated here may be used in other situations also. In table 3 we consider the data on the Pyrausta nubilalis, to which generalized Poisson distribution was fitted by Jain (1975). It is observed that generalized Poisson-Lindley distributions (5.1) and (5.13) give better fit. The Distribution of corn borers is considered in table 4, to which the Negative- Binomial and the Neyman's type A distributions were fitted by Bliss and Fisher (1953). It will be seen that the data agree excellently with generalized Poisson-Lindley fit. The generalized Poisson-Lindley distributions are also fitted in table 5, to some data collected by P. Garman (see Bliss et al,1953) and the fitting is found to be satisfactory.

We consider the numerical data of Table 2, Haemocytometer Counts of Yeast Cells, where the expected frequencies match satisfactorily with those computed by distribution (6.1) than the Hermite distribution (Kemp and Kemp), and in case of the counts of European red mites on apple leaves in table 5, is also good filled by the distribution (6.1).

TABLE 1: OBSERVED AND FITTED DISTRIBUTIONS (4.1) AND (6.1)

(Data from the paper by Beall and Rescia)

No. of Insects	Observed frequency	Fitted Distribution (4.1) $\hat{\lambda}=0.1403$ $\hat{\omega}=0.8089$	Fitted Distribution (6.1) $\hat{n}=1.00058$ $\hat{a}=0.9724$ $\hat{\omega}=0.0213$	Pascal-Poisson $\alpha=1$	Pascal-Poisson $\alpha=2$
0	33	31.00	31.93	32.05	33.45
1	12	14.86	13.73	13.65	11.34
2	6	6.05	5.89	5.87	5.86
3	3	2.69	2.52	2.25	3.34
4	1	0.99	1.08	1.08	0.88
5	1	0.41	0.46	0.83	1.13
Total	56	56.00	56.00	56.00	56.00
χ^2		0.8824	0.6389	1.93	0.08

TABLE 2: OBSERVED AND FITTED DISTRIBUTIONS (4.4) AND (6.1)

(Haemocytometer Counts of Yeast Cells)

No. of Insects	Observed frequency	Fitted Distribution (4.4) $\hat{\lambda}=0.4932$ $\hat{\omega}=0.2774$	Hermite Distribution (Kemp and Kemp)	Fitted Distribution (6.1) $\hat{n}=2.7463$ $\hat{a}=0.2202$ $\hat{\omega}=0.8799$
0	213	212.51	213.12	214.77
1	128	126.18	122.91	121.52
2	37	43.81	46.71	45.39
3	18	12.63	13.31	13.74
4	3	3.50	3.16	3.58
5	1	0.97	0.64	0.81
6	0	0.27	0.15	0.05
Total	400	400.00	400.00	400.00
χ^2		2.3201	3.8825	2.7136

Table-3: OBSERVED AND FITTED POISSON-LINDLEY MIXING DISTRIBUTIONS (5.1) AND (5.13)

Pyrausta nubilslis in 1937 (data by Beall and Rescia 1940)

No.of insects	Observed	GPL1(5.1) $\hat{\omega}=.5276$ $\hat{\theta}=3.4556$	GPL2 (5.13) $\hat{\omega}=0.1123$ $\hat{\theta}=12.74$	Poisson-Lindley (By Sankaran, 1970) $\hat{\theta}=1.8082$	Generalized Poisson (by Jain, 1975)
0	33	31.86	33.10	31.49	32.46
1	12	13.84	12.49	14.16	13.45
2	6	5.92	5.50	6.09	5.60
3	3	2.52	2.49	2.54	2.42
4	1	1.06	1.13	1.04	1.08
> 5	1	0.80	1.29	0.42	0.97
Total	56	56.00	56.00	56.00	56.00
χ^2		0.3743	0.0667	0.6532	0.25

Table-4: OBSERVED AND FITTED POISSON-LINDLEY MIXING DISTRIBUTIONS (5.1) AND (5.13)

Corn Borers (data of Beall 1940)

No of Insects	Observed	GPL1(5 1) θ^=3 4556 ω^=0 7611	GPL2(5 13) θ^=6 64 ω^=0 1148	Poisson-lindley Sankaran, 1970 θ^=1 0096	Bliss et al (1953) Negative Binomial	Neyman-type-A
0	43	48 05	49 11	45 36	44 3	49 8
1	35	29 04	24 83	30 07	31 1	23 3
2	17	17 36	15 30	18 70	19 1	18 9
3	11	10 35	9 84	11 16	11 2	12 3
4	5	6 16	6 38	6 48	6 4	7 3
5	4	3 67	4 14	3 68	3 6	4 1
6	1	2 18	2 69	2 06	2 0	2 2
7	2	1 30	1 75	1 14	1 1	1 1
8	2	1 89	1 14	62	1 2	1 1
Total	120	120 00	120 00	120 00	120 00	120 00
χ^2		1 7289	5 6031	1 7251		

Table-5: OBSERVED AND FITTED POISSON-LINDLEY MIXING DISTRIBUTIONS (5.1) AND (5.13)

Count of the number of European red mites on apple leaves (data of Bliss et al 1953)

No of mites Per leaf	Leaves observed	GPL1(5 1) θ^=1 392 ω^=0 1117	GPL2(5 13) θ^=8 58 ω^=0 1022	Poisson-lindley (By Sankaran 1970) θ^=1 258	Gen Negative Binomial(by Jain and Consul, (1971)	fitted dist (6 1) a^= 8902 w^= 1385 n^=1 0219
0	70	67 62	70 89	67 [illegible]	71 48	68 67
1	38	[illegible] 68	33 35	38 [illegible]	33 98	37 83
2	17	21 04	18 70	21 [illegible]	19 80	20 47
3	10	11 09	10 84	11 [illegible]	11 59	10 97
4	9	5 73	6 73	5 [illegible]	6 57	5 83
5	3	2 92	3 69	2 90	3 55	3 07
6	2	1 47	2 15	1 44	1 80	1 52
7	1	74	1 26	71	84	79
8	0	71	73	34	39	41
Total	150 00	150 00	150 00	150 00	150 00	150 00
χ^2		2 8491	2 4433	3 0136	2 07	2 4318

RFERENCES

Chang, D K. (1989) On infinitely divisible discrete distributions, Utilitas Mathematica, 36, 215- 217.

Feller, W. (1957). An Introduction to Probability Theory and Its Applica tions, Vol. I, (third Edition), New York, Wiley

Fisz, M. (1962) : Infinitely divisible distributions : recent results and ap plications. Ann. Math. Statist., 33, 68- 84.

Godambe, A.V. and Patil, G.P. (1975) Some characterizations involving additivity and infinite divisibility and their applications to Poisson mixtures and Poisson sums, Statistical Distributions in Scientific Work, 3, Characterizations and Applications, G.P. Patil, S. Kotz, and J.K. Ord (editors), 339- 351. Dordrecht: Reidel.

Goldie Charles (1967) : A class of infinitely divisible random variables, Proc Camb. Phil. Soc.

Gnedenko, B.V. and Kolmogorov, A.N (1968) : Limit distributions for sums of independent random variables, revised edition. Addison-Wesley Publishing Co., Reading Massachusetts.

Katti, S.K. (1967) Infinite divisibility of integer valued random variables, Annals of Mathematical Statistics, 38, 1306- 1308.

Klebanov, L.B., Maniya, G.M. and Melamed, I.A. (1984) A problem of zolotarev and analogs of infinitely divisible and stable distributions in a scheme of summing a random number of random variables. Theor. prob. Applicn. 4, 791- 794.

Levy, P. (1937) : Theorie de l' aaaition des variables aleatoires, Gauthier-Villars, Paris.

Loeve, M. (1960) : Probability theory, Second Ed., Van Nostrand, Princeton, N.J.

Lukacs,E. (1970) Characteristics Functions (Second edition), London: Griffin

Maceda, E.C. (1948) On the compound and generalized Poisson distribu tions, Annals of Mathematical Statistics, 19, 414- 416.

Pillai, R.N. (1985) Semi a- Laplace distributions. omm. Stat. Theory and Methods 14, 991- 1000.

Pillai, R.N. (1990) Harmonic mixtures and geometric infinite divisibility. J. Indian Stat. Assoc. 28, 87- 98.

Pillai, R.N. and Sandhya, E. (1990) Distributions with complete mono tone derivative and geometric infinite divisibility. Adv. Appl. Prob. 22, 751- 754.

Pillai, R.N. and Jose, K.K. (1994) Geometric infinite divisibility and autoregressive time series modelling, Proceeding of the third Ramanujan Symposium on Stochastic Processes and their Applications, University of Madras, Madras 17-19 January, 1994, 81-87.

Sandhya, E. (1991) : Geometric Infinite Divisibility and Applications, Ph. D. thesis (unpublished) submitted to the University of Kerala, January, 1991

Steutel, F.W. (1970) Preservation of infinite divisibility under mixing, Mathematical Centre Tract, 33, Amsterdam.

Steutel, F.W. (1973). Some recent results in infinite divisibility, Stochastic Pocesses Appl., Vol. I, 125-143.

Steutel, F.W. (1983) Infinite divisibility, Encyclopedia of Statistical Sciences, 4, S. Kotz, N.L. Johnson and C.B. Read (editors), 114-116, New York, Wiley.

Steutel, F.W. (1990) The set of geometrically infinitely divisible distributions, Technische Universiteit Eindhoven Memorandum COSOR.

Warde, W.D. and Katti, S.K. (1971) Infinite divisibility of discrete distributions II, Annals of Mathematical Statistics, 42, 1088-1090.

Proceedings of the Annual Technical Session,
Assam Science Society, 1999

SOME PROPERTIES OF POISSON MIXING INFINITELY DIVISIBLE DISTRIBUTIONS

M. Borah and R. A. Begum[1]

Abstract

Elementary infinitely divisible distributions which are formulated on the basis of simple models seem to be inadequate to describe the situations which may occur in a number of phenomena. In the last few years a number of various infinitely divisible distributions have been derived. In this paper three forms of infinitely divisible discrete distributions have been studied. The recurrence relations for their probabilities and factorial moments are investigated. Further, the recurrence relations for partial derivatives of the probabilities with respect to its parameters are also investigated which may facilitate the calculation of the asymptotic covariance matrix of maximum likelihood estimators. As the method of maximum likelihood will be very cumbersome, some other adhoc methods have been also used to estimate the parameters. A few sets of reported data have been considered for the fitting of the distributions, and the fits are compared with that obtained with other distributions.

Key words : Infinitely divisible distribution, Recurrence relation, Covariance matrix, Factorial moments.

1. INTRODUCTION

The theory of infinitely divisible distributions developed primarily during the period from 1920 to 1950, has played a very important role in the solution of limit problems for sums of independent random variables. A full account of this theory and its applications as it had been developed by the late 40's were presented in the monographs of Levy (1937), Gnedenko and Kolmogorov (1968), and Loeve (1960). In the last few years research in this field has been carried out along many lines. Numerous new results have been obtained and entirely new applications have been found.

A random variable is said to be infinitely divisible if it has a characteristic function (cf). $\phi(t)$ that can be represented for every positive integer n as the nth power of some cf $\phi_n(t)$

$$\phi(t) = \{\phi_n(t)\}^n \tag{1.1}$$

Department of Mathematical Sciences, Tezpur University, Tezpur-784001, Assam
[1]Department of Statistics., Darrang College, Tezpur, Assam

It is well known that all infinitely divisible distributions are generated by stochastic processes, more specifically by processes with stationary independent increments. There is a number of methods to construct new infinitely divisible distributions from given ones. The best known are convolution and compounding. A general compounding theorem is due to Feller (1957) There are some applied processes however, which give rise to classes of infinitely divisible distributions. Certain families of probability distribution functions maintain their infinite divisibility under repeated mixing and convolution.

Godambe and Patil (1975) considered a mixture of Poisson distributions where the mixing distributions have non-negative support. The importance of the property of infinite divisibility in modelling was stressed by Steutel (1983); see also the monograph by Steutel (1970). Infinitely divisible discrete distributions with pgf's of the forms

$$G(t) = \frac{(1-\omega)\, g(t)}{1-\omega\, g(t)}, \quad 0 < \omega < 1 \tag{1.2}$$

$$G(t) = \frac{g(t)}{\omega + (1-\omega)\, g(t)}, \quad \omega > 0 \tag{1.3}$$

$$G(t) = \frac{\omega}{\omega + 1 - g(t)}, \quad \omega > 0 \tag{1.4}$$

were studied by Steutel (1990) under the name geometrically infinitely divisible distributions. Conditions for discrete distributions to be infinitely divisible were discussed by Katti (1967), Warde and Katti (1971), and Chang (1989)

Here taking g(t) equal to the p.g.f. of Poisson distributions in the above three forms in the equations (1.2),(1.3), and (1.4) respectively, three infinitely divisible distributions have been derived. Further, the recurrence relations for probabilities and factorial moments of the newly derived distributions are investigated. The problem of estimation and the fitting of the distributions have also been considered.

2. RECURRENCE RELATIONS FOR PROBABILITIES

(a) The pgf of Poisson mixing infinite divisible distribution derived from the first form (1.2) will be given by

$$G(t) = \frac{(1-\omega)\, e^{\lambda(t-1)}}{1-\omega\, e^{\lambda(t-1)}}, \quad \omega > 0,\ \lambda > 0, \tag{2.1}$$

Probability recurrence relation

$$P_{r+1} = e^{-\lambda}/(1-\omega e^{-\lambda})[\lambda^{r+1}(1-\omega+P_0\omega)/(r+1)! + \omega \sum_{j=1}^{r} \lambda^j P_{r-j+1}/(j)!)] \tag{2.2}$$

may be derived by expanding the pgf G(t) of equation (2.1) w.r.t.'t' and equating the coefficients of t^r on both sides, where, $P_0 = (1 - \omega)\, e^{-\lambda}/(1 - \omega\, e^{-\lambda})$,

$P_1 = \lambda P_0\{1+ \omega\, P_0/(1 - \omega)\}$.

$P_2 = \lambda P_1\{1/2 + \omega\, P_0/(1 - \omega)\}$.

$P_3 = \lambda P_2\{P_2/\lambda P_1 - \lambda P_1/12 P_2\}$.

Factorial moments recurrence relation may be derived from the factorial moment generating function (f.m.g.f) corresponding to the p.g.f.(2.1) which will be given by

$$\mu_{(r+1)} = \omega(r+1)/(1-\omega) \sum_{j=1}^{r} \binom{r}{r-j+1} \mu_{(j)} \lambda^{r-j+1}/j + \lambda^{r+1}/(1-\omega), \tag{2.3}$$

where $\mu_{(1)} = \lambda \,/\, (1-\omega)$,

The higher order factorial moments may be obtained by putting r= 1,2,3 etc in (2.3) respectively as

$$\mu_{(2)} = \mu_{(1)}\ \{\lambda + 2\ \omega \mu_{(1)}\},$$

$$\mu_{(3)} = 3\omega \mu_{(1)}\ \{\lambda \mu_{(1)} + \mu_{(2)}\} + \lambda^2 \mu_{(1)},$$

$$\mu_{(4)} = 4\omega \mu_{(1)}\ \{\mu_{(3)} + 3\lambda \mu_{(2)} / 2 + \lambda^2 \mu_{(1)}\} + \lambda^2 \mu_{(1)},$$

where $\mu_{(r)}$ denotes r^{th} factorial moment, and we have

MEAN $= \mu = \lambda/(1-\omega)$, and

VARIANCE $= \mu\{(\lambda+1) + (2\omega - 1)\ \mu\}$.

(b). The pgf of Poisson mixing infinitely divisible distribution derived from the second form (1.3) will be given by

$$G(t) = \frac{e^{\lambda(t-1)}}{\omega + (1-\omega)e^{\lambda(t-1)}}, \quad 0 < \omega < 1, \quad \lambda > 0 \tag{2.4}$$

Probability recurrence relation may be derived from the pgf (2.4) by equating the coefficients of t^r on both sides as

$$P_{r+1} = e^{-\lambda}/\{\omega + (1-\omega)\, e^{-\lambda}\}[\lambda^{r+1}\{1-(1-\omega)P_0\}/(r+1)! - (1-\omega) \sum_{j=1}^{r} (\lambda^j/j!)\ P_{r+j-1}] \tag{2.5}$$

where $P_0 = e^{-\lambda}/\{\omega + (1-\omega)\, e^{-\lambda}\}$

Putting r=0,1,2, respectively in equation (2.5), higher order probabilities may be calculated as

$$P_1 = \lambda P_0\{1-(1-\omega)P_0\}.$$

$$P_2 = \lambda P_1\{1/2-(1-\omega)P_0\}.$$

$$P_3 = \lambda P_2[P_2/\lambda P_1 - \lambda^2 P_1/12P_2\}.$$

Similarly factorial moments recurrence relation may be derived from the f.m g f of the distribution as given by

$$\mu_{(r+1)} = \omega\lambda^{r+1} - (1-\omega)\ )\ (r+1) \sum_{j=1}^{r} \binom{r}{j-1} \lambda^j \mu_{(r-j+1)}/j, \tag{2.6}$$

where, $\mu_{(1)} = \omega\lambda$

The higher order factorial moments may be obtained by putting r=1,2,3 etc. respectively in (2.6) as

$$\mu_{(2)} = \mu_{(1)}\ (2\mu_{(1)} - \lambda)$$

$$\mu_{(3)} = 3\mu_{(1)}{}^2\lambda + 3\mu_{(1)}\ \mu_{(2)} - 3\lambda\mu_{(1)} - 2\lambda^2\mu_{(1)}$$

$$\mu_{(4)} = 3\mu_{(1)}{}^2\lambda + 3\mu_{(1)}\ \mu_{(2)} - 3\lambda\mu_{(1)} - 2\lambda^2\mu_{(1)}$$

and MEAN = $\mu = \omega\lambda$,

VARIANCE = $\mu\ (\mu+1-\lambda)$.

(c) The pgf of Poisson mixing infinitely divisible distribution derived from equation (1.4) will be given by

$$G(t) = \frac{\omega}{\omega + 1 - e^{\lambda(t-1)}}\ , \quad 0 < \omega < 1,\ \ \lambda > 0 \tag{2.7}$$

Probability recurrence relation may be derived from (2 7) by expanding and equating the coefficients of t^r on both sides as

$$P_{r+1} = e^{-\lambda}/\{\omega + 1 - e^{-\lambda}\}[\lambda^{r+1} P_0/(r+1)! - \sum_{j=1}^{r} (\lambda^j/j!)\ P_{r-j+1}], \tag{2.8}$$

The higher order probabilities may be obtained by putting r =1, 2, 3 etc. respectively in (2.7) as

$$P_1 = \lambda e^{-\lambda} P^2{}_0/\omega.$$

$$P_2 = \lambda P_1\{1/2 + e^{-\lambda} P_0/\omega.\}.$$

$$P_3 = P_2\{P_2/P_1 - \lambda P_1/12P_2\}$$

Similarly the factorial moment recurrence relation may be derived from the factorial moment generating function (f.m g.f) corresponding to the p g.f. (2.7) as

$$\mu_{(r+1)} = (r+1)/\omega \, [\sum_{j=1}^{r} \binom{r}{j-1} \lambda^j \mu_{(r-j+1)}/j] + \lambda^{r+1}/\omega, \tag{2 9}$$

where, $\mu_{(1)} = \lambda/\omega$,

$$\mu_{(2)} = \mu_{(1)} \, (2\mu_{(1)}+\lambda)$$

$$\mu_{(3)} = \mu_{(1)} \, \{\lambda^2+3(\mu_{(2)} + \lambda\mu_{(1)})\}$$

$$\mu_{(4)} = 4\mu_{(1)}\{\mu_{(3)}+ (3/2)\lambda\mu_{(2)} + \lambda^2\mu_{(1)}\}+\lambda^3\mu_{(1)}$$

and MEAN $= \mu = \lambda/\omega$, and

VARIANCE $= \mu\,(1+\mu+\lambda)$.

3. ESTIMATION OF PARAMETERS

(a) Poisson mixing distribution (2.1) with parameters (ω, λ)

$$\omega = \frac{\sigma^2 - \mu}{\mu^2}, \quad \text{and} \quad \lambda = \mu + 1 - \sigma^2/\mu$$

where μ and σ^2 denote the mean and the variance of the distribution (2.1)

(b) Poisson mixing distribution (2 4) with parameters (ω, λ)

$$\omega = \frac{\mu^2}{\mu^2 + \mu - \sigma^2}, \quad \text{and} \quad \lambda = 1 + \mu - \sigma^2/\mu$$

where μ and σ^2 denote the mean and the variance of the distribution (2 4)

(c) Poisson mixing distribution (2.7) with parameters (ω, λ)

$$\omega = \frac{\sigma^2 - \mu^2 - \mu}{\mu^2}, \quad \text{and}$$

$$\lambda = \frac{\sigma^2}{\mu} - \mu - 1.$$

where μ and σ^2 denote the mean and the variance of the distribution (2 7)

4. GOODNESS OF FIT

One of the problems in fitting distributions is that of estimation of the parameters Of all the procedures of estimating the parameters, the method of moments is perhaps the oldest and the simplest. In many cases it leads to tractable operations. Whereas the other methods become computationally complicated, it is mainly for this reason that, moment estimators are used. The distributions have been fitted to some available data as shown in Tables 1, 2 and 3. It is clear that there is some improvement, however small it may be, in fitting these distributions over the other distributions, considered earlier. The fitting of these distributions as indicated here many be used in other situations also.

TABLE 1: OBSERVED AND FITTED POISSON MIXING DISTRIBUTION (2 1)

(Data from the paper by Beall and Rescia)

No of Insects	observations	Fitted Dist. $\hat{\lambda} = 0.1403$ and $\hat{\omega} = 0.8089$	Pascal-Poisson Distribution	
			$\alpha=1$	$\alpha=2$
0	33	31.00	32.05	33.45
1	12	14.86	13.65	11.34
2	6	6.05	5.87	5.86
3	3	2.69	2.25	3.34
4	1	0.99	1.08	.88
5	1	0.41	0.83	1.13
Total	56	56.00	56.00	56.00
χ^2		0.8824	1.93	0.08

TABLE 2: OBSERVED AND FITTED POISSON MIXING DISTRIBUTION (2.4)

(Counts of the number of European Red Mites on Apple Leaves)

No of Obs.	Observations	Fitted Dist. $\hat{\lambda} = 0.1638$, and $\hat{\omega} = 0.8571$	Negative Binomial distribution
0	70	66.79	67.49
1	38	40.17	39.03
2	17	20.87	20.86
3	10	10.75	10.97
4	9	5.76	5.66
5	3	2.95	2.90
6	2	1.57	1.48
7	1	0.75	0.75
8	0	0.39	0.76
Total	150	150 00	150.00
χ^2		2.8640	2.8936

TABLE 3: OBSERVED AND FITTED POISSON MIXING DISTRIBUTION (2.7)
(Haemocytometer Counts of Yeast Cells)

No of Obs.	Observations	Fitted Dist. $\hat{\lambda} = 0.4932$ $\hat{\omega} = 0.2774$	Hermite distribution (Kemp and Kemp)
0	213	212.51	213.12
1	128	126.18	122.91
2	37	43.81	46.71
3	18	12.63	13.31
4	3	3.50	5.16
5	1	0.97	0.64
6	0	0.27	0.15
Total	400	400.00	400.00
χ^2		2.2301	2.8825

RFERENCES

Chang, D.K (1989) On infinitely divisible discrete distributions, *Mathematica*, 36, 217.

Feller, W. (1957) An introduction to probability theory and its applications, Vol. I, (third edition), New York, Wiley

Fisz, M. (1962): infinitely divisible distributions: recent results and applications. *Ann. Math. Stats*, 33,68-84.

Godambe, A.V. and Patil, G.P. (1975) Some characterizations involving additivity and infinite divisibility and their applications to Poisson sums, *Statistical Distributions in Scientific Work*, 3, Characterization and Applications, G.P. Patil, S. Kotz, and J.K. Ord (editors), 339-351. Dordrecht: Reidel.

Goldie Charles (1967) : A class of infinitely divisible random variables, *Proc. Camb. Phil. Soc.*

Gnedenko, B.V. and Kolmogorov, A.N. (1968) : Limit distributios for sums of independent random variables, Revised edition. Addison- Wesley Publishing Co., Reading, Massachusetts.

Katti, S.K. (1967) Infinite divisibility of integer valued random variables *Annals of Mathematical Statistics,* 38, 1906-1308.

Klebanov, L.B., Maniya, G.M. and Melamed, I.A. (1984) A problem of Zolotarev and analogs of infinitely divisible and stable distributions in a scheme of summing a random number of random variables. *Theor. Prob. Applicn.* 4, 791-794

Steutel, F.W. (1970) Preservation of infinite divisibility under mixing, *Mathematical Centre Tract,* 33, Amsterdam.

1

Pillai, R.N. (1985) Semi α- Laplace Distributions. *Comm. Stat. Theory and Methods* 14,991-1000.

Pillai, R.N. (1990) Harmonic Mixtures and Geometric infinite divisibility. *J. Indian Stat. Assoc.* 28,87-98.

Pillai, R.N. and Sandhya, E. (1990) Distributions with complete monotone derivative and geometric infinite divisibility, *Adv. Appl. Prob.* 22, 751-754.

Pillai, R.N. and Jose, K.K (1994) Geometric infinite divisibility and autoregressive time series modelling, *Proceedings of the Third Ramanujan Symposium on Stochastic Processes and Their* Applications, University of Madras, Madras 17-19 January, 1994,81-87.

Samdhya, E. (1991) : Geometric Infinite Divisibility and Applications, Ph. D. thesis (unpublished) submitted to the University of Kerala, January, 19

Steutel, F.W. (1973) : Some recent results in infinite divisibility, *Stochastic Processes Appl.*, Vol. 1, 125-143.

Steutel, F.W. (1983) Infinite divisibility, *Encyclopedia of Statistical Sciences,* 4, S. Kotz, N.L. Johnson and C.B. Read (editors), 114-116, New York, Wiley.

Steutel, F.W. (1990) The set of geometrically infinitely divisible distributions, *Technische Universiteit Eindhoven Memorandum COSOR.*

Warde, W.D. and Katti, S.K. (1971) Infinite divisibility of discrete distributions II, *Annals of Mathematical Statistics,* 42, 1088-1090.

Levy, P. (1937) : Theorie de l' aeraition des variables aleatoires, Gauthier-Villars, Paris.

Loeve, M. (1960): Probability theory, Second Ed., Van Nostrand, Princeton, N.J.