AN EXTENDED PEARSON SYSTEM USEFUL IN RELIABILITY ANALYSIS

THESIS SUBMITTED TO THE

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

UNDER THE FACULTY OF SCIENCE

By

T. K. SINDU

DEPARTMENT OF STATISTICS

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COCHIN-682 022

DECEMBER 2002

CERTIFICATE

Certified that the thesis entitled "AN EXTENDED PEARSON SYSTEM USEFUL IN RELIABILITY ANALYSIS" *is a bonafide record of work done by Smt. Sindu.T.K. under my guidance in the Department of Statistics, Cochin University of Science and Technology, Cochin and that no part of it has been included anywhere previously for the award of any degree or title.*

Cochin University of Science and Technology
December 19, 2002

Dr.P.G.Sankaran
Lecturer in Statistics

UNIVERSITY OF SCIENCE AND TECHNOLOGY
DEPARTMENT OF STATISTICS
COCHIN-22

CONTENTS

CHAPTER I

INTRODUCTION

1.1. Introduction

In the real world, all products and systems are unreliable in the sense that they degrade with age and ultimately fail. Since the process of deterioration leading to failure occurs in a random manner, the concept of reliability requires a probabilistic framework.

The term reliability of a product (system) is the probability that the product (system) will perform its intended function for a specified time period when operating under normal environmental conditions. Even though the above definition of reliability is explained with reference to the behaviour or length of life of a system, it is equally applicable in the analysis of any duration variable that describes a well defined population subject to decrementation due to the operation of forces of attrition overtime. Accordingly the concepts and tools used in reliability analysis have found applications in other areas such as economics, demography, survival analysis, biology, medicine and engineering.

Prior to World War II, the notion of reliability was largely intuitive, subjective and qualitative. The use of actuarial methods to estimate survivorship of rail-road equipments began in the early part of the twentieth century (Nelson, 1982). In the late 1930's extreme value theory was used to model fatigue life of materials and was the forerunner of later probabilistic developments.

A more mathematical and formal approach to reliability grew out of the demands of modern technology and particularly out of the experiences in World War II with complex military systems (Barlow and Proschan, 1975). Barlow (1984) deals with a historical perspective of mathematical reliability theory up to that time. Similar perspectives on reliability engineering in electronic equipment can be found in Coppola (1984), on nuclear power system reliability in Fussel (1984) and on software reliability in Shooman (1984).

1.2 Reliability modeling

Reliability theory deals with the interdisciplinary use of probability, statistics and stochastic modeling combined with engineering insight into the design and the scientific understanding of the failure mechanism. As such it encompasses issues such as

(i) reliability modeling
(ii) reliability analysis and optimization
(iii) reliability engineering
(iv) reliability science
(v) reliability technology
(vi) reliability management

The major endeavor here is to develop new statistical techniques that can be used for modeling the lifetime data. One of the basic problems in reliability modeling when the data on failure times is the only input, is to identify the underlying model that is supposed to generate the observations. Generally it is not easy to isolate all the physical causes that contribute individually or collectively to the failure mechanism and to mathematically account for each and as such the task of identifying the correct model representing the data becomes very difficult. A standard practice adopted in most modeling situation is to ascertain the physical properties of the process generating the observations, express them by means of equations or inequalities and then solve them to obtain the model.

In reliability, analysts have developed certain basic concepts such as failure rate, mean residual life, vitality etc. through which the physical characteristics of the failure mechanism can be adequately described and therefore these concepts form the basis of specifying a probability distribution of failure times. Thus if one can translate the characteristics of the failure mechanism in terms of failure rate, mean residual life or an ageing criteria and if there exists a probability distribution characterized by such a property, the problem of model identification is satisfactorily resolved. As already mentioned, apart from the point of view of reliability theory, the results obtained here are of interest in their own right in distribution theory and also in various applied studies, whose concepts and tools in reliability are used with differing interpretations.

1.3 Basic Reliability concepts

The discussions in the previous section reveals that the difficulties in reliability modeling can be reduced by appealing to certain concepts associated with the failure process, that permit different distributions to be distinguished. In the present section, we discuss these concepts and review the results that will be used on the sequel.

1.3.1 Reliability Function

Let X be a non-negative random variable on a probability space $(\Omega, \mathcal{F}, P)$ with distribution function $F(x) = \mathrm{P}(X \leq x)$. In the reliability context, X generally represents the length of life of a device, measured in units of time and the function,

$$\begin{aligned} R(x) &= 1 - F(x) \\ &= P(X > x) \end{aligned}$$

is called the survival or reliability function. It gives the probability that the device will operate without failure for a mission time x. The probability density function(p.d.f) of X, $f(x)$, is obtained from $R(x)$ by the relationship

$$f(x) = -\frac{d}{dx}R(x). \tag{1.1}$$

1.3.2 Failure Rate

Defining the right extremity L of $F(x)$ by

$$L = \inf\{x: F(x)=1\},$$

the failure rate $h(x)$ of X, when $F(x)$is absolutely continuous with respect to Lebesgue measure with probability density $f(x)$, is defined for $x<L$ by

$$h(x) = \lim_{\Delta x \to 0^+} \frac{P(x < X < x + \Delta x \mid X > x)}{\Delta x}$$

$$= \frac{f(x)}{R(x)}$$

$$= \frac{d}{dx}[-\log R(x)]. \tag{1.2}$$

The distribution of X is uniquely determined by the relationship

$$R(x) = \exp\left[-\int_0^x h(t)dt\right]. \tag{1.3}$$

Accordingly,

$$f(x) = h(x) \exp\left[-\int_0^x h(t)dt\right]. \tag{1.4}$$

In the general case, when X is a random variable on the entire real line, Kotz and Shanbhag (1980) defined the failure rate as the Radon-Nikodym derivative with respect to Lebesgue measure on $\{x: F(x)<1\}$, of the hazard measure,

$$H(B) = \int_B \frac{dF}{1-F(x)}$$

for every Borel set B of $(-\infty, L)$.

Further the distribution of X is uniquely determined by the relationship

$$R(x) = \prod_{u<x}[1 - H(u)]\exp[-H_c(-\infty, c)], \tag{1.5}$$

where H_c is the continuous part of H.

1.3.3 Mean Residual Life Function

The mean residual life (MRL), known in early literature in actuarial studies as expectation of life, was reintroduced in the reliability context by Knight in 1959 (Kupka and Loo, 1989). Later this function was used by Watson and Wells (1961) to study the effect of burn-in on the useful life of articles. MRL represents the average life time remaining to a component which has survived up to time x. When X is defined on the real line with $E(X^{+}) < \infty$, the B-measurable function

$$r(x) = E(X - x \mid X > x) \tag{1.6}$$

for all x such that $P(X>x)>0$ is called the MRL function of X. In the case when X is non-negative with $E(X)<\infty$ and $F(x)$ is absolutely continuous with respect to Lesbesgue measure

$$r(x) = \frac{1}{R(x)} \int_x^{\infty} R(t)dt. \tag{1.7}$$

Further, the MRL function related to hazard function by

$$h(x) = \frac{1 + r'(x)}{r(x)} \tag{1.8}$$

and

$$R(x) = \frac{r(0)}{r(x)} \exp\left\{-\int_0^x \frac{dt}{r(t)}\right\} \tag{1.9}$$

for every x in $(0,L)$, where $r'(x)$ denotes the derivative of $r(x)$ with $r(0) = \mathrm{E}(X)$.

Guerrieri (1965) and Cundy (1966) established that for a non-negative random variable X with finite mean, the MRL function $r(x)=c$, a constant is a characterizing property of the exponential distribution. MRL function had been extensively used in lifetime studies by Hollandar and Proschan (1975), Bryson and Siddique (1969) and Muth (1980). One set of necessary and sufficient condition for a function to be an MRL given by Swartz (1973) is that along with (1.9),

(i) $r(x) \geq 0$

(ii) $r(0) = E(X)$

(iii) $r'(x) \geq -1$ and

(iv) $\int_0^\infty \frac{dx}{r(x)}$ should be divergent.

For more properties and applications, we refer to Muth (1977), Galambos and Kotz (1978) and Kotz and Shanbhag (1980).

1.3.4 Vitality function

The concept of vitality function is closely related to MRL and it is defined as

$$m(x) = E(X|X>x) \tag{1.10}$$

or

$$m(x) = \frac{1}{R(x)} \int_x^\infty t dF(t) \tag{1.11}$$

[See, Kupka and Loo (1989)].

Obviously,

$$m(x) = x + r(x) \tag{1.12}$$

and

$$m'(x) = r(x)h(x)$$

where $m'(x)$ denotes the derivative of $m(x)$.

The vitality function satisfies the following properties

(i) $m(x)$ is non-decreasing and right continuous on $(-\infty, L)$

(ii) $m(x) \geq x$ for all $x<L$

(iii) $\lim_{x \to L} m(x) = L$

(iv) $\lim_{x \to -\infty} m(x) = E(x)$.

Kupka and Loo (1989) proved that if F satisfies increasing vitality function property together with decreasing MRL, then F has increasing failure rate property. Shanbhag (1970) showed that the vitality function $m(x) = x + r(x)$, is a constant characterizes the exponential distribution. Characterizations of probability distributions using vitality function were also given by Osaki and Li (1988), Ahmed (1991), Nair and Sankaran (1991), Ruiz and Navarro (1994) and Navarro et.al. (1998).

1.3.5 Variance Residual Life

The variance residual life (VRL) of a random variable X is defined as

$$V(x) = V(X-x|X\geq x) \tag{1.13}$$

$$= E[(X-x)^2|X\geq x] - r^2(x). \tag{1.14}$$

This concept was introduced by Launer (1984) in order to define certain new classes of life distributions and to provide bounds on the reliability function for certain specified class of distributions. Gupta et. al. (1987) proved that,

$$V(x) = \frac{2}{R(x)} \int_x^\infty r(t)R(t)dt - r^2(x)$$

and

$$\frac{dV(x)}{dx} = h(x)\,[V(x) - r^2(x)].$$

Further they showed that the increasing (decreasing) VRL distribution have close relationship with increasing (decreasing) MRL models, but the former provides a more general class of distributions.

1.3.6 Partial Moments

The r^{th} partial moment of a continuous random variable X about a point t is defined as

$$p_r(t) = \mathrm{E}[(X-t)^+]^r, \quad r = 1, 2, \ldots, t>0 \tag{1.15}$$

where

$$(X-t)^+ = X-t, \quad X \geq t$$
$$= 0, \quad X<t.$$

The random variable $(X-t)^+$ is quite meaningful in the insurance studies. When X represents the income of an individual and t is the tax exemption level, $(X-t)^+$ represents the taxable income. Then $p_1(t)$ gives the average income that exceeds the exemption level. (1.15) can be written as

$$p_r(t) = \int_t^\infty (x-t)^r f(x)dx. \tag{1.16}$$

In reliability, $p_1(t)$ gives the average lifetime that exceeds the current age t.

Also if $\mathrm{E}(X^r)<\infty$, (1.16) is equivalent to

$$p_r(t) = r\int_t^\infty (x-t)^{r-1} R(x)dx.$$

It follows that,

$$p_1(t) = r(t)\, R(t). \tag{1.17}$$

From (1.17), we get

$$h(t) = \frac{r'(t)}{r(t)} - \frac{p_1'(t)}{p_1(t)} = -\frac{p_1''(t)}{p_1'(t)}$$

and

$$r(t) = -\frac{p_1(t)}{p_1'(t)}.$$

Chong (1977) has characterized the exponential distribution by the property

$$E(X-t-s)^+ E(X) = E(X-t)^+ E(X-s)^+.$$

Later Gupta and Gupta (1983) have made an extensive study of partial moments and established that one partial moment is sufficient to determine the parent distribution uniquely. For properties and applications, we refer to Hitha (1991) and Sunoj (2002).

1.4 Families of Distributions

As mentioned earlier, a standard technique adopted in modeling situation is to ascertain the physical properties of the process generating the observations express them by means of equations or inequalities and then solve them to obtain the model. There are however, situations when the system is so complex that the response derived from it may not be amenable to simple mathematical manipulations nor possess such mathematical structures. Also, only very little will be known about the physical characteristics of the system. One method that can be used in such situations is to use a general family of distributions, one member of which could be a possible model. The main reason to prefer this procedure is the desire to find the best possible approximation in a complex situation that generated the data rather than any reasonable evidence to the effect that the model explains the data generating mechanism. Once the model is found it may perhaps be possible to explain the nature of the observations through the model. When the families of distributions are chosen for modeling, it is desirable that

(a) it contains enough members with different shapes so that there is a member that can correspond to a given data situation
(b) the members of the family should have a sufficient number of parameters to impart flexibility
(c) there should be some simple criterion that distinguishes the various members of the family, so that the choice of a member that fits the data become easy and
(d) efficient methods exists for the estimation of the parameters.

The above discussions clearly reveal that the family of distributions play a pivotal role in reliability modeling. Statistical literature is abundant with families of probability distributions arising from different contexts. Various families of distributions used in reliability modeling are Pearson family, Exponential family, Burr family etc. Pearson family is the oldest among them and it is extensively used in reliability modeling as it contains most of the lifetime distributions such as gamma, beta, normal, Pareto, exponential etc.

1.5 Pearson Family

Pearson Family of probability distribution was introduced by Karl Pearson in 1895. A brief description of the Pearson family is as follows.

Let X be a continuous random variable in the support of H={a, b} where a can be $-\infty$ and b can be $+\infty$ and $f(x)$ represent the probability density function (p.d.f) of X. Assume that $f(x)$ is differentiable with respect to x. Then the distribution of X belong to the Pearson family if $f(x)$ satisfies the differential equation

$$\frac{1}{f(x)}\frac{df(x)}{dx} = \frac{-(x+d)}{b_0 + b_1 x + b_2 x^2} \tag{1.18}$$

where b_0, b_1, b_2 and d are real constants. The shape of the distributions depends on the values of the parameters b_0, b_1, b_2 and d. The form of solution of (1.18) evidently depends on nature of the roots of the equation $b_0+b_1x+b_2x^2=0$ and the various types correspond to the roots of the quadratic equation in the denominator of (1.18). Pearson introduced three main types of the curves,

1. Type I corresponds to both the roots are real and of opposite signs,
2. Type IV occurs when both the roots are imaginary and
3. Type VI occurs when both the roots are real and of same sign.

In the limiting cases when one type changes into another we get simple forms of transition type curves. For various properties and applications of the Pearson family of distributions, we refer to Elderton and Johnson (1969), Ord (1972), Johnson, Kotz and Balakrishnan (1994), Nair and Sankaran (1991), Glanzel (1991), Ruiz and Navarro (1994), Navarro, Franco and Ruiz (1998) and Sankaran and Nair (2000).

1.6 Length Biased Models

The length biased model is a particular case of the well known weighted models. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X: \Omega \to Q$ be a random variable where $Q = (a,b)$ is the subset of the real line with $a \geq 0$ and $b > a$ can be finite or infinite. The distribution function $F(x)$ is assumed to be absolutely continuous with respect to Lebesgue measure with probability density function $f(x)$ and $w(X)$ is a non-negative function of X such that $\mu = E[w(X)] < \infty$. The random variable Y with probability density function

$$g(x) = \frac{w(x)f(x)}{\mu}, \; x>0 \qquad (1.19)$$

is said to have a weighted distribution associated with X.

The concept of weighted distributions was introduced by Rao(1965) in connection with modeling statistical data in situations where the usual practice of employing standard distribution for the purpose was not found appropriate. The basic problem when one uses a weighted distribution as a tool for modeling is the identification of the appropriate weight function that fits the data. When $w(x)=x$, the corresponding observed distribution in (1.19) is termed as length (size) biased distribution. That is, when the weight function depends on the length of the unit of interest, the resulting distribution is called length biased. More generally, when the sampling mechanism selects units with probability proportional to some measure of the unit size, the resulting distribution is called length (size) biased. The statistical interpretation of the length biased distribution was originally identified by Cox (1962) in the context of renewal theory. But the same idea has originally been conceived much before as evidenced from Daniels (1942) who discusses length biased sampling in the analysis of the distribution of fiber lengths in wool. An exhaustive account of the research in this area is available in Patil and Rao (1977) and Gupta and Kirmani (1990).

1.7 Reliability Modeling in Discrete Time

In most of the studies relating to life testing and reliability, lifetime is usually represented by a non-negative continuous random variable and accordingly continuous probability distributions are proposed as models. Xekalaki (1983) pointed out that the discrete models are more appropriate in a variety of applied problems due to the limitations in measuring equipments and to the fact that many continuous life length distributions can be very well approximated by the corresponding discrete counterparts. Gupta (1985) has given an example of discrete random variable that occur naturally, such as the case with the time to failure in the fatigue studies measured in terms of the number of cycles to failure. Cox (1972), Kalbfleisch and Prentice (1980) and Lawless (1982) have provided the basic formulation to the study of discrete life distributions. In the type I censoring, the number of failed units up to a certain time period can be represented by a discrete distribution and this may be used to study the failure process of the system. Therefore, a development of concepts and methods when length of life is treated as discrete random variable appears to be in right place. For more applications of discrete models in reliability and survival analysis, we can refer to Padgett and Spurrier (1985), Ebrahimi (1986), Guess and Park (1988), Nair and Hitha (1989) and Shaked et.al (1994, 1995).

1.8 Reliability Concepts in Discrete Time

Let X denote a discrete random variable in the support of $I^+=\{0, 1, 2, \ldots\}$ denoting time to failure of a component or system. Then the survival function defined as

$$R(x) = P(X \geq x),$$

so that

$$p(x) = R(x) - R(x+1) \tag{1.20}$$

where $p(x)$ is the probability mass function (p.m.f) of X.

The failure rate of X is defined as

$$h(x) = \frac{p(x)}{R(x)}$$

and it is shown that the failure rate determines the life distribution uniquely through the relationship

$$R(x) = \prod_{y=0}^{x-1}[1-h(y)]. \tag{1.21}$$

Xekalaki (1983) proved that if X is a random variable taking values in the set $\{0, 1, 2, \ldots, k\}$, $k \in \{0, 1, 2, \ldots\} \cup \{+\infty\}$, then

$$h(x) = \frac{1}{a+bx}$$

holds iff X has geometric distribution for $b=0$, Waring distribution for $b>0$ and negative hyper geometric distribution for $b<0$. Later Hitha (1990) has shown that the continuous approximation of geometric, Waring and negative hyper geometric distributions are respectively exponential, Pareto II and beta distributions which have the same form for failure rate in continuous time.

The Mean Residual Function (MRL) in the discrete setup is given by

$$r(x) = E(X-x|X>x)$$

which provides

$$r(x) = \frac{\sum_{t=x+1}^{\infty} R(t)}{R(x+1)} \tag{1.22}$$

Like the failure rate, MRL determines the distribution of X as given by

$$R(x) = \prod_{u=1}^{x-1} \frac{r(u-1)-1}{r(u)} [1-p(0)] \tag{1.23}$$

where $p(0)$ is determined such that $\sum p(x) = 1$. (Nair and Hitha, 1989).

The relationship between failure rate and MRL is given by,

$$1 - h(x+1) = \frac{r(x)-1}{r(x+1)}, \; x = 0, 1, 2, \ldots \tag{1.24}$$

Nair (1983) has used the function $r(x)$ to define the notion of memory of life distributions and Salvia (1996) established some simple bounds for residual life when the device has a monotonic hazard rate sequence.

As in the continuous case, the vitality function is defined as

$$m(x) = E(X|X>x).$$

Further, $h(x)$, $r(x)$ and $m(x)$ are related to one another by the following identities (Hitha and Nair, 1989)

$$m(x) = x + r(x)$$

and

$$h(x+1)\, r(x+1) = m(x+1) - m(x). \tag{1.25}$$

The r[th] factorial partial moment can be defined as

$$\alpha_r(t) = E[(X-t)^+]^r,\ X>t+r-1$$

$$= \sum_{t+r}^{\infty}(x-t)^{(r)}p(x),\ r = 1, 2, \ldots, t= 0,1,2\ldots \qquad (1.26)$$

where $x^{(r)} = x(x-1)\ldots(x-r+1)$.

Nair et. al. (2000) explored the properties of partial moments of discrete random variables and pointed out their applications in distribution theory and reliability analysis. For other properties and applications of partial moments to reliability analysis, we can refer to Hitha (1991), Priya et. al. (2000) and Priya (2001).

In the discrete setup also, the family of distributions posseses a vital role in reliability modeling. The families of discrete distributions used in this connection are power series family by Kosambi (1949) and Noack (1950), Katz family (Katz, 1945), Ord family (Ord, 1972) and Kemp family (Kemp, 1968). Among them Ord family of distributions is important because it includes most of the discrete distributions like binomial, Poisson, negative binomial, hyper geometric, Waring etc. Ord family of distributions is defined by the difference equation

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{-(x+u)}{k_0+k_1x+k_2x^2} \qquad (1.27)$$

where k_0, k_1, k_2 and u are real constants.

The nature of the roots of the equation $k_0+ k_1x+k_2x^2=0$ determines different types of distributions. For $k_2 \neq 0$, the denominator of (1.27) has two roots with the posibilities

Type I: one root zero, the other non zero and range finite

Type IV: both the roots are imaginary

Type VI: one root zero, the other negative and infinite range.

For different properties and applications of the Ord family of distributions, we can refer to Ord (1972), Nair and Sankaran (1991), Glanzel (1991)etc.

1.9 Present Study

The present thesis is organized into five chapters. In the introductory chapter we discuss the relevance and the scope of the study along with review of literature on reliability modeling.

We present in Chapter II, an extended version of the Pearson family of distributions. Various properties of the family are discussed. An identity connecting conditional moments and failure rate is developed that enables the determination of the particular model in a practical situation. A characterization result that relates the conditional means is also established which generalizes the result given by Glanzel (1991) to the Pearson family.

In reliability, ageing behaviour of the system is usually studied by the failure rate (hazard rate) function. The increasing (decreasing) failure rates (IFR/DFR) property is the characteristic of the system that consistently deteriorate (improved) with age. This brings the relevance and the need of classification of distributions based on failure rate function which provides information about the system reliability. In Chapter III, we discuss a procedure to identify an IFR/DFR model from the generalized Pearson family. We also derive necessary and sufficient conditions under which the members of the generalized Pearson family are form-invariant (that is having the same form for the density) with respect to the formation of their length biased distributions.

In reliability, there are situations where discrete distribution naturally arises like the number of cycles to failure or the number of failures in a given time interval. Motivated by the relevance and the usefulness of discrete models, we propose to develop some results that have applications in the modeling and analysis of lifetime data in the discrete time domain. Chapter IV, deals with an extended class of Ord family (generalized Ord family) and provides some characterizations using conditional means.

The present thesis concludes with Chapter V by providing a method to identify an IFR/DFR model in the generalized Ord family. Further we derive the conditions under which the members of the generalized Ord family are form-invariant with respect to the formation of their length biased distributions.

CHAPTER II

A GENERALISED PEARSON FAMILY

2.1. Introduction

The normal distribution played a vital role in the statistical analysis till the end of the nineteenth century and the developments in the statistical theory took place on the assumption that the population is normal or at least approximately so. However, there are practical situations where the samples from many populations show characteristics that are different from normal. By the end of the nineteenth century non-normal curves became popular and efforts were underway to generate systems of curves which include the normal only as a particular case. Accordingly, Pearson (1895) introduced a system of distributions represented by the differential equation (1.18) that includes normal as a special case. The family (1.18) is used widely in reliability modeling as it contains many other important probability models such as exponential, gamma, beta, Pareto etc. For various properties and applications of the family (1.18) we refer to Nair and Sankaran (1991), Glanzel (1991), Johnson, Kotz and Balakrishnan (1994) and Navarro, Franco and Ruiz (1998).

Some parts of the work in this chapter is due to appear in Sankaran and Sindu (2003).

There are distributions that does not belong to the family (1.18), but are widely employed in reliability modeling. For example, the inverse Gaussian distribution with probability density function

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{\frac{-\lambda(x-\mu)^2}{2x\mu^2}\right\}, \; x, \; \lambda, \; \mu>0$$

is not a member of the family (1.18). Motivated by this, we study an extended class of distributions that we termed as generalized Pearson family. The proposed class of distributions include the Pearson family as a particular case.

Earlier several extensions to the Pearson family have been proposed (Ord, 1972) including the use of polynomials of general order fitted directly to histogram estimates of $\frac{d \log f}{dx}$ (see Dunning and Hanson, 1977). A multimodal generalization of the Pearson family is available in Stuart and Ord (1994).

In the following section we discuss an extended version of the Pearson family in the context of reliability.

2.2. A Generalized Pearson Family

Let X be a random variable having absolutely continuous distribution function $F(x)$ in the support of (a,b) where $a<b$, a can be $-\infty$ and b can be $+\infty$. Let $f(x)$ denote the probability density function of X. The distribution of X belongs to the generalized Pearson family if $f(x)$ is differentiable and satisfies the differential equation

$$\frac{d\log f(x)}{dx}=\frac{a_0+a_1x+a_2x^2}{b_0+b_1x+b_2x^2} \tag{2.1}$$

where a_0, a_1, a_2, b_0, b_1 and b_2 are real constants. When $a_2=0$, the family (2.1) reduces to the Pearson family (1.18). The family (2.1) is a special case of the multimodel generalization of the Pearson family given in Stuart and Ord (1994). Like the Pearson family of distributions, the generalized version can be classified into a number of types and different types are based on the nature of the roots of the quadratic expression in the denominator of (2.1). Therefore the procedure for finding different types of curves for the generalized Pearson family is same as that of Pearson type distributions.

2.3 Members of the Generalized Pearson Family

All members of the Pearson family are also the members of the family (2.1). In the following we discuss the different types of the curves of the generalized Pearson family and important distributions belonging to these types.

Type I

The family (2.1) can be written as

$$\frac{f'(x)}{f(x)} = c + \frac{x+d}{p_0+p_1x+p_2x^2} \tag{2.2}$$

where

$$c=\frac{a_2}{b_2},\ d=\frac{a_0b_2-a_2b_0}{a_1b_2-a_2b_1}$$

and

$$p_i = \frac{b_i b_2}{a_1 b_2 - a_2 b_1}, \quad a_1 b_2 \neq a_2 b_1,\ i=0,1,2.$$

Type I distribution occurs when the roots of the equation $p_0+p_1x+p_2x^2=0$, are real and of opposite signs.

Let the roots be α and $-\beta$, with $\alpha,\beta > 0$, then

$$\frac{x+d}{p_0+p_1x+p_2x^2} = \left[\frac{A}{x-\alpha}+\frac{B}{x+\beta}\right]\frac{1}{p_2}$$

$$= \frac{A(x+\beta)+B(x-\alpha)}{p_2(x+\beta)(x-\alpha)} \tag{2.3}$$

Now by partial fractions, we get

$$A = \frac{\alpha+d}{\alpha+\beta} \text{ and } B = \frac{\beta-d}{\alpha+\beta}$$

Hence (2.3) becomes,

$$\frac{x+d}{p_0+p_1x+p_2x^2} = \left[\frac{\alpha+d}{(\alpha+\beta)(x-\alpha)}+\frac{\beta-d}{(\alpha+\beta)(x+\beta)}\right]\frac{1}{p_2} \tag{2.4}$$

Integrating (2.2), using (2.4), we obtain

$$\log f(x) = cx + \frac{1}{p_2}\left(\frac{\alpha+d}{\alpha+\beta}\right)\log(x-\alpha) + \frac{1}{p_2}\left(\frac{\beta-d}{\alpha+\beta}\right)\log(x+\beta) + \text{constant}.$$

Therefore,

$$f(x) = Y_0\, e^{cx}\, (x-\alpha)^{m_1}\,(x+\beta)^{m_2},$$

where

$$m_1 = \frac{1}{p_2}\left(\frac{\alpha+d}{\alpha+\beta}\right)$$

and

$$m_2 = \frac{1}{p_2}\left(\frac{\beta-d}{\alpha+\beta}\right).$$

Assume now that the density vanishes at both ends of the distribution. Therefore

$f(x) = 0$, when $x = \alpha$, $x=-\beta$ and $x = \infty$ provided $c<0$.

Hence

$f(x) = Y_0\, e^{-vx}\, (x-\alpha)^{m_1}\, (x+\beta)^{m_2}$, where $c = -v$, $\alpha, \beta, v, m_1, m_2 \geq 0$, $\alpha<x<\infty$

(2.5)

and Y_0 is the normalizing constant.

When $m_2 = 0$ and $\alpha = 0$ in (2.5), we get

$$f(x) = Y_0\, e^{-vx}\, x^{m_1},\ 0<x<\infty$$

which is gamma distribution.

When $v = 0$, (2.5) reduces to

$$f(x) = Y_0\, (x-\alpha)^{m_1}\, (x+\beta)^{m_2},\ -\beta<x<\alpha$$

where

$Y_0 = e^{v\beta}\,(-\alpha-\beta)^{m_1}\,(\alpha+\beta)^{m_2+1}\mathrm{B}(m_1+1, m_2+1)\ {}_1F_1(1+m_2, 2+m_1+m_2, -v(\alpha+\beta))$

[if $\alpha+\beta>0$, Re $v>0$, Re $m_1>-1$, Re $m_2>-1$]

with

$$\mathrm{B}(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}\,dx$$

and

$${}_1F_1(a,\ b,\ z) = \frac{\Gamma(b)}{\Gamma(b-a)\Gamma(a)} \int_0^1 e^{zt} t^{a-1}(1-t)^{b-a-1}\,dt.$$

Then,

$$f(x) = \frac{(x-\alpha)^{m_1}(x+\beta)^{m_2}}{B(m_1+1, m_2+1)e^{\gamma\beta}(-1)^{m_1}(\alpha+\beta)^{m_1+m_2+1}\ {}_1F_1(1+m_2, 2+m_1+m_2, -\gamma(\alpha+\beta))},$$

$-\beta<x<\alpha$

and hence

$$f(x) = \frac{(x-\alpha)^{m_1}(x+\beta)^{m_2}}{B(m_1+1, m_2+1)(-1)^{m_1}(\alpha+\beta)^{m_1+m_2+1}},\ -\beta<x<\alpha. \qquad (2.6)$$

Putting, $\alpha=1$ and $\beta=0$ in (2.6), we get

$$f(x) = \frac{(x-1)^{m_1} x^{m_2}}{B(m_1+1, m_2+1)(-1)^{m_1}}$$

$$= \frac{(1-x)^{m_1+1-1} x^{m_2+1-1}}{B(m_1+1, m_2+1)}, \; 0<x<1.$$

which is beta distribution of first kind.

Type IV

Type IV curve occurs when both roots of $p_0+p_1x+p_2x^2=0$ are imaginary.

Consider the equation (2.2). This can be written as

$$\frac{f'(x)}{f(x)} = c + \frac{y+k}{p_2(y^2+A^2)} \tag{2.7}$$

where $y = x + \frac{p_1}{2p_2}$, $k = d - \frac{p_1}{2p_2}$ and $A^2 = \frac{p_0}{p_2} - \frac{p_1^2}{4p_2^2}$.

Integrating (2.7), we obtain

$$\log f(x) = cx + \int \frac{y}{p_2(y^2+A^2)}dy + k\int \frac{1}{p_2(y^2+A^2)}dy + \text{constant}$$

and hence

$$f(x) = Y_0 \left(1+\frac{x^2}{a^2}\right)^{-m} \exp\left\{vx - p\tan^{-1}\left(\frac{x}{a}\right)\right\}, \; -a<x<a$$

where $c = -v$, $p = \frac{-k}{Ap_2}$, $m = \frac{-1}{2p_2}$, $a = A$ and Y_0 is a normalizing constant obtained from $\int_{-a}^{a} f(x)dx = 1$.

Type VI

This type occurs when the roots are real and of same sign. Let the roots be α and β then,

$$\frac{f'(x)}{f(x)} = c + \left[\frac{A}{x-\alpha} + \frac{B}{x+\beta}\right]\frac{1}{p_2}$$

or

$$\frac{f'(x)}{f(x)} = c + \left[\frac{d-\alpha}{(\beta-\alpha)(x+\alpha)} + \frac{\beta-d}{(\beta-\alpha)(x+\beta)}\right]\frac{1}{p_2} \tag{2.8}$$

Integrating (2.8) , we get

$$\log f(x)=cx+\frac{1}{p_2}\left(\frac{d-\alpha}{\beta-\alpha}\right)\log(x+\alpha)+\frac{1}{p_2}\left(\frac{\beta-d}{\beta-\alpha}\right)\log(x+\beta) + \text{constant.}$$

Therefore,

$$f(x) = Y_0\, e^{-\nu x}\,(x+\alpha)^{m_1}\,(x+\beta)^{m_2},\ \min(-\alpha,-\beta)<x<\infty.$$

where

$$m_1 = \frac{1}{p_2}\left(\frac{d-\alpha}{\beta-\alpha}\right)$$

$$m_2 = \frac{1}{p_2}\left(\frac{\beta-d}{\beta-\alpha}\right)$$

$$c = -\nu$$

and Y_0 is the normalizing constant.

Similarly we can derive the other transition type curves and they are as follows.

Type II

$$f(x) = Y_0\, e^{-\nu x}\left(1-\frac{x^2}{a^2}\right)^m \exp\left\{-p\tanh^{-1}\left(\frac{x}{a}\right)\right\},\ -a\le x\le a.$$

Type III

$$f(x) = Y_0\, e^{-(\nu+p)x}\left(1+\frac{x}{a}\right)^m,\ -a\le x<\infty. \tag{2.9}$$

Type V

$$f(x) = Y_0\, e^{-\nu x}\, x^{-p}\, e^{-q/x},\ 0<x<\infty. \qquad (2.10)$$

Type VII

$$f(x) = Y_0 \left(1+\frac{x^2}{a^2}\right)^m \exp\left\{\nu x - \delta \tan^{-1}\left(\frac{x}{a}\right)\right\},\ -\infty<x<\infty.$$

Type VIII

$$f(x) = Y_0\, e^{-\nu x} \left(1+\frac{x}{a}\right)^{-m} x^{-p},\ -a\le x<\infty.$$

This is a special case of (2.9) with $p=0$.

Type X

$$f(x) = Y_0\, e^{-\nu x}\, x^{-p},\ 0<x<\infty.$$

This is a particular case of (2.10) with $q=0$.

Type XI

$$f(x) = Y_0\, e^{-qx-\nu/x}\, x^{-p},\ 0<x<\infty.$$

Next we consider some important probability distributions belonging to the generalized Pearson family, which are not members of the Pearson family.

1. Inverse Gaussian Distribution

The probability density function of inverse Gaussian distribution is given by

$$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{\frac{-\lambda(x-\mu)^2}{2x\mu^2}\right\},\ x,\ \lambda>0 \qquad (2.11)$$

which gives,

$$\frac{f'(x)}{f(x)} = \frac{-\lambda x^2 - 3x\mu^2 + \lambda\mu^2}{2x^2\mu^2}$$

so that $a_0 = \lambda\mu^2$, $a_1 = -3\mu^2$, $a_2 = -\lambda$, $b_0 = b_1 = 0$ and $b_2 = 2\mu^2$. This distribution belongs to type V curve.

When early occurrences such as product failures or repairs are dominant in a lifetime distribution, failure rate is expected to be non-monotonic, first increasing and later decreasing. In such situations the inverse Gaussian distribution provides a suitable choice as a lifetime model. For details, we refer to Chhikara and Folks (1977) and Seshadri (1999).

2. Random Walk Distribution

The probability density function of random walk distribution is given by

$$f(y) = \sqrt{\frac{\lambda}{2\pi y}} \exp\left\{\frac{-\lambda(y\mu-1)^2}{2y\mu^2}\right\}, \; y,\ \lambda,\mu>0 \qquad (2.12)$$

Random walk distribution is known as the inverse distribution of the inverse Gaussian distribution. That is, if X follows inverse Gaussian distribution with p.d.f (2.11), then the transformation $Y=\frac{1}{X}$ follows random walk distribution with p.d.f (2.12). This distribution also belongs to type V curve of the (2.1).

3. Rayleigh Distribution

The probability density function of Rayleigh distribution is given by

$$f(x) = 2\lambda x\ e^{-\lambda x^2},\ 0<x<\infty,\ \ \lambda>0 \qquad (2.13)$$

and hence

$$\frac{f'(x)}{f(x)} = \frac{1-2\lambda x^2}{x}$$

which is of the form (2.1) with a_0 = 1, a_1 =0, a_2 = -2λ, b_0 = b_2 =0 and b_1 = 1. Rayleigh distribution is a special case of the Weibull distribution (see, Johnson, Kotz and Balakrishnan, 1994).

4. Maxwell Distribution

For the Maxwell distribution with p.d.f

$$f(x) = 4\sqrt{\frac{\lambda^3}{\pi}}\, x^2\, e^{-\lambda x^2},\ 0<x<\infty,\ \ \lambda>0, \tag{2.14}$$

and hence

$$\frac{f'(x)}{f(x)} = \frac{2-2\lambda x^2}{x}$$

which is of the form (2.1) with a_0 = 2, a_1 =0, a_2 = -2λ, b_0 = b_2 =0 and b_1 = 1.

This distribution arises as the distribution of the magnitude of a gas in a closed container under the assumption that the gas is not flowing and that the pressure in the gas is the same in all directions (see, Johnson, Kotz and Balakrishnan, 1994).

2.4 Properties of the Generalized Pearson Family

In this section we discuss some important properties of the generalized Pearson family.

2.4.1 Recurrence Relationship among Moments

Recurrence relationship between moments is useful to find higher order moments from the mean and the variance. For the generalized Pearson family, we obtain the following recurrence relationship among the raw moments.

Consider the differential equation (2.1), then

$$f(x)\,[a_0 + a_1 x + a_2 x^2] = [b_0 + b_1 x + b_2 x^2]\, f'(x) \qquad (2.15)$$

Multiplying both sides of (2.15) by x^n and applying integration with the assumption that $x^r f(x) \to 0$ as $x \to \pm\infty$, we obtain

$$n b_0 \mu'_{n-1} + [(n+1)b_1 + a_0]\,\mu'_n + [(n+2)b_2 + a_0]\,\mu'_{n+1} + a_2\,\mu'_{n+1} = 0 \qquad (2.16)$$

where $\mu'_n = \int x^n f(x)dx$. If we put $n = 0, 1, \ldots 5$ respectively, we get 6 equations those enable us to find the constants a_0, a_1, a_2, b_0, b_1 and b_2 in terms of the moments. Thus

$$(b_1 + a_0)\,\mu'_0 + (2b_2 + a_0)\,\mu'_1 + a_2\,\mu'_2 = 0 \qquad (2.17)$$

$$b_0\,\mu'_0 + (2b_1 + a_0)\,\mu'_1 + (3b_2 + a_0)\,\mu'_2 + a_2\,\mu'_3 = 0 \qquad (2.18)$$

$$2b_0\,\mu'_1 + (3b_1 + a_0)\,\mu'_2 + (4b_2 + a_0)\,\mu'_3 + a_2\,\mu'_4 = 0 \qquad (2.19)$$

$$3b_0\,\mu'_2 + (4b_1 + a_0)\,\mu'_3 + (5b_2 + a_0)\,\mu'_4 + a_2\,\mu'_5 = 0 \qquad (2.20)$$

$$4b_0\,\mu'_3 + (5b_1 + a_0)\,\mu'_4 + (6b_2 + a_0)\,\mu'_5 + a_2\,\mu'_6 = 0 \qquad (2.21)$$

$$5b_0\,\mu'_4 + (6b_1 + a_0)\,\mu'_5 + (7b_2 + a_0)\,\mu'_6 + a_2\,\mu'_7 = 0 \qquad (2.22)$$

As in the Pearson set up, we obtain the parameters of the family in terms of moments by solving the above six equations.

2.4.2 Relationship using Characteristic Function

A simple solution of an extremely wide range of problems of probability theory, especially those associated with the summation of independent random variables is obtainable by means of characteristic function. The characteristic function also plays an important role in the determination of probability distribution. In the following we prove a result in this direction.

Theorem 2.1

Let $\phi(\theta)= E(e^{\theta x})$, where $\theta = it$. If $\phi'(\theta) = \dfrac{d\phi}{d\theta}$ and $\phi''(\theta)=\dfrac{d^2\phi}{d\theta^2}$ exist for some θ in an interval $\theta_L < \theta < \theta_U$ and assuming high order contact at the ends of the range, the characteristic function of a distribution in the generalized Pearson family satisfies the equation

$$(a_2+b_2\theta)\ \phi''(\theta) + (a_1+2b_2+b_1\theta)\ \phi'(\theta) + (a_0+b_1+b_0\theta)\ \phi(\theta)=0. \tag{2.23}$$

Proof

Consider the family (2.1). Then

$$f(x)\ [a_0+ a_1x +a_2\ x^2] = [b_0+ b_1x +b_2\ x^2]\ f'(x). \tag{2.24}$$

Multiplying both sides of (2.24) by $e^{\theta x}$ ($\theta = it$) and integrating with respect to x, we obtain

$$\int_{-\infty}^{\infty} e^{\theta x}\ [a_0+ a_1x +a_2\ x^2]\ f(x)dx = \int_{-\infty}^{\infty} e^{\theta x}\ [b_0+ b_1x +b_2\ x^2]\ f'(x)\ dx$$

or

$$\left[e^{\theta x}(b_0+b_1x+b_2x^2)f(x)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} [\theta e^{\theta x}(b_0+b_1x+b_2x^2)+e^{\theta x}\ (b_1+2b_2x)]f(x)dx$$

$$= a_0 \int_{-\infty}^{\infty} e^{\theta x} f(x)dx + a_1 \int_{-\infty}^{\infty} x\ e^{\theta x} f(x)dx + a_2 \int_{-\infty}^{\infty} x^2 e^{\theta x} f(x)dx$$

which gives (2.23).

Remark 2.1

When $a_2 = 0$, the above result reduces to the Pearson family of distributions.

2.4.3 Relationship using Partial Moments

The properties of partial moments can be used to characterize probability distributions and it is shown that the partial moments determine the distribution completely. In the following, we provide a recurrence relationship for the generalized Pearson family using partial moments.

Theorem 2.2

Let X be a random variable in the support of real line with $E(X^r)<\infty$. Suppose that the distribution of X belongs to the generalized Pearson family. Then

$$p_{r-1}(t)\ [r(b_0+ b_1 t +b_2\, t^2)] + p_r(t)\ [(r+1)\ (b_1 +2b_2\, t) + (a_0+ a_1 t +a_2\, t^2)] + p_{r+1}(t)\ [(r+2)b_2+ a_1 +2a_2\, t)]+ a_2\, p_{r+2}(t) = 0. \quad (2.25)$$

Proof

When the distribution of X belongs to the generalized Pearson family, we have

$$f(x)\ [a_0+ a_1 x +a_2\, x^2] = [b_0+ b_1 x +b_2\, x^2]\, f'(x) \qquad (2.26)$$

Multiplying both sides of (2.26) by $(x-t)^r$ and integrating from t to b, we obtain

$$\int_t^b (x-t)^r\, [a_0+ a_1 x +a_2\, x^2]\, f(x)\, dx = \int_t^b (x-t)^r\ [b_0+ b_1 x +b_2\, x^2]\, f'(x)\, dx. \qquad (2.27)$$

Putting, $x^2 = (x-t)^2 + 2t(x-t) + t^2$ in (2.27), we get

$$-b_1\, p_r(t) - 2b_2\, p_{r+1}(t) - 2b_2\, t\, p_r(t) - rb_0\, p_{r-1}(t) - rb_1\, p_r(t) - \mathrm{rtb}_1\, p_{r-1}(t) - rb_2\, p_{r+1}(t) -2rtb_2\, p_r(t) - rb_2\, t^2\, p_{r-1}(t)$$
$$= a_0\, p_r(t) + a_1\, p_{r+1}(t) + a_1 t\, p_r(t) + a_2\, p_{r+2}(t) + 2ta_2\, p_{r+1}(t) + a_2\, t^2\, p_r(t) \qquad (2.28)$$

Rearranging (2.28), we have (2.25).

Corollary 2.1

When $a_2 = 0$, the relationship (2.25) reduces to the result of Sunoj (2002) for the Pearson family of distributions.

2.5 Characterizations

Characterizations of distributions in statistics are of great interest and are widely used for the modeling of data (See, Kagan, Linnik and Rao (1973) and Galamboz and Kotz (1978)). In particular, several characterizations of distributions using the basic concepts such as failure rate, mean residual life and vitality function have been extensively discussed in reliability. Shanbhag (1970) who first provided characterization for the exponential distribution using the vitality function. Characterizations of Pearson family of distributions using reliability concepts were given by, Nair and Sankaran (1991), Glänzel (1991), Sankaran and Nair (1993), Ruiz and Navarro (1994), Navarro, Franco and Ruiz (1998), and Sankaran and Nair (2000).

In the following, we prove a characterization theorem for the generalized Pearson family, using a relationship between the failure rate and the vitality function.

Theorem 2.3

A necessary and sufficient condition for the distribution of X belongs to the family (2.1) under the regularity condition $\lim_{x \to b} x^r f(x) = 0$ for $r=0,1,2$ is that

$$a_2\, E(X^2|X>x) + (a_1+2b_2)E(X|X>x) + a_0+b_1 + (b_0+b_1x+b_2x^2)h(x)=0 \tag{2.29}$$

where $h(x)$ is the failure rate of X.

Proof

When the distribution of X belongs to the family (2.1), we have

$$f'(x)\ (b_0+b_1x+b_2x^2) = f(x)(a_0+a_1x+a_2x^2)$$

which gives

$$\int_x^b f'(t)(b_0+b_1t+b_2t^2)\ dt = \int_x^b f(t)(a_0+a_1t+a_2t^2)\ dt. \tag{2.30}$$

Using integration by parts and applying the regularity condition, (2.30) provides that

$$-(b_0+b_1x+b_2x^2)\ f(x)- \int_x^b f(t)(b_1+2b_2t)dt = \int_x^b f(t)(a_0+a_1t+a_2t^2)dt. \tag{2.31}$$

Dividing (2.31) by $R(x)$ and rearranging the terms, we obtain (2.29). Conversely, when (2.29) holds, we have

$$a_2 \int_x^b t^2 f(t)\ dt + (a_1+2b_2)\int_x^b tf(t)dt + (a_0+b_1)R(x) +(b_0+b_1x+b_2x^2)\ f(x) = 0. \tag{2.32}$$

Differentiating (2.32) with respect to x and applying the regularity condition, we get

$$-a_2\, x^2\, f(x)- (a_1+2b_2)\, x\, f(x)- (a_0+b_1)f(x) + (b_0+b_1x+b_2x^2)\, f'(x)$$
$$+ f(x)\ (b_1+2b_2x) \quad =0$$

and finally we obtain (2.1). This completes the proof.

Corollary 2.2

When $a_2 = 0$, Theorem 2.3 reduces to the result of Nair and Sankaran (1991).

Corollary 2.3

The distribution of X is inverse Gaussian with p.d.f (2.11) if and only if

$$\lambda\, E(X^2|X>x) = \mu^2\, E(X|X>x) + \lambda\mu^2 + 2\mu^2 x^2 h(x).$$

Corollary 2.4

The distribution of X is Rayleigh distribution with p.d.f (2.13) if and only if

$$E(X^2|X>x) = \frac{1}{\lambda} + \frac{x}{2\lambda} h(x).$$

Corollary 2.5

The relationship

$$E(X^2|X>x) = \frac{3}{2\lambda} + \frac{x}{2\lambda} h(x)$$

holds if and only if X has Maxwell distribution with p.d.f (2.14).

Corollary 2.6

The relationship

$$\lambda\, E(X^2|X>x) = 4\, E(X|X>x) + \lambda/\mu^2 + 2\, x^2 h(x).$$

holds if and only if X has random walk distribution with p.d.f (2.12).

2.6 Characterization using Conditional Moments

Glanzel (1991) proved that the distribution of the continuous random variable X belongs to the Pearson family if and only if

$$E(X^2|X>x) = P(x)\, E(X|X>x) + Q(x)$$

where $P(x)$ and $Q(x)$ are polynomials of degree one atmost, with real coefficients. In the following we prove a theorem that generalizes the result of Glanzel (1991).

Theorem 2.4

Let X be a continuous random variable as defined in section 2.2. Assume that $E(X^3) < \infty$ and $E(X^3|X>x)$, $E(X^2|X>x)$ and $E(X|X>x)$ are differentiable. Then the distribution of X belongs to the generalized Pearson family (2.1) if

$$a_2 E(X^3|X>x) = A(x)E(X^2|X>x) + B(x)E(X|X>x) + C(x) \qquad (2.33)$$

holds, where $A(x) = a_2x+q$, with q as a real constant and $B(x)$ and $C(x)$ are polynomials of degree one atmost with real coefficients. The reverse statement holds if $(b_0+b_1x+b_2x^2)xf(x)\to 0$ as $x\to b$ if $b=+\infty$ and if $(b_0+b_1x+b_2x^2)f(x)\to 0$ as $x\to b$ if $b<+\infty$.

Proof

When $A(x) = a_2x+q$, $B(x) = rx+s$ and $C(x) = tx+u$, (2.33) becomes,

$$a_2E(X^3|X>x)=(a_2x+q)E(X^2|X>x)+(rx+s)E(X|X>x)+ tx+u. \qquad (2.34)$$

Differentiating (2.34) twice and rearranging the terms, we obtain,

$$f'(x)\,[(q+r)x^2+(s+t)x+u] = f(x)\,[a_2x^2 + (2q+3r)x+s+2t]$$

or

$$\frac{f'(x)}{f(x)} = \frac{a_2x^2+(2q+3r)+s+2t}{(q+r)x^2+(s+t)x+u}.$$

Thus the distribution of X is a member of the generalized Pearson family (2.1).

Assume now that the distribution of X belongs to the generalized Pearson family. Then we have,

$$f(x)\,[a_0+ a_1x +a_2\, x^2] = [b_0+ b_1x +b_2\, x^2]\, f'(x). \qquad (2.35)$$

Integrating (2.35) from x to b, applying the regularity condition and dividing both sides by $R(x)$, we get

$$-[b_0+b_1x+b_2x^2]\frac{f(x)}{R(x)}=a_0+b_1+(2b_2+a_1)E(X|X>x)+a_2E(X^2|X>x) \qquad (2.36)$$

Multiplying (2.36) by x and carrying out the previous steps, we obtain

$$-x[b_0+b_1x+b_2x^2]\frac{f(x)}{R(x)} = a_2\, E(X^3|X>x)+(3b_2+a_1)E(X^2|X>x)$$

$$+(2b_1+a_0)E(X|X>x) +b_0. \qquad (2.37)$$

Substituting (2.36) in (2.37), we obtain

$$x[b_1+2b_2\, E(X|X>x) + a_2\, E(X^2|X>x)+a_1E(X|X>x) +a_0]$$

$$=a_2\, E(X^3|X>x)+(3b_2+a_1)E(X^2|X>x) +(2b_1+a_0)E(X|X>x) +b_0$$

which gives,

$$a_2\, E(X^3|X>x) = [a_2x-(a_1+3b_2)]E(X^2|X>x)$$

$$+[(a_1+2b_2)x-(a_0+2b_1)]E(X|X>x) + (a_0+b_1)x -b_0.$$

This is of the form (2.33) with $A(x) = a_2x-(a_1+3b_2)$, $B(x) =(a_1+2b_2)x-(a_0+2b_1)$ and $C(x) = (a_0+b_1)x -b_0$.

Corollary 2.7

When a_2=0 in (2.33), we obtain the result given by Glanzel (1991).

Corollary 2.8

The relationship

$$\lambda\, E(X^3|X>x) = (-\lambda x+3\mu^2)E(X^2|X>x) +(\mu^2x -\lambda\mu^2)E(X|X>x) + \mu^2\lambda x$$

holds if and only if X has the p.d.f (2.11).

Corollary 2.9

The distribution of X is Rayleigh with p.d.f (2.13) if and only if

$$2\lambda\ E(X^3|X>x) = 2\lambda x E(X^2|X>x) + 3E(X|X>x) - 2x$$

In this case, $A(x) = 2\lambda x$, $B(x) = 3$, $C(x) = -2x$, and $p=2\lambda$.

Corollary 2.10

The distribution of X is Maxwell with p.d.f (2.14) if and only if

$$2\lambda\ E(X^3|X>x) = 2\lambda x E(X^2|X>x) + 4E(X|X>x) - 3x.$$

CHAPTER III

AGEING PROPERTIES OF THE GENERALIZED PEARSON FAMILY

3.1 Introduction

In reliability theory the concept of ageing plays a central role as it helps to classify the lifetime models. Earlier works in reliability theory was centered around the problem of estimating the reliability function based on observed data. Recently a lot of interest have been evolved to modeling the lifetime data and to classify the life distributions based on certain ageing properties. Accordingly, large number of research papers have been published which examine the behavior of the life distributions based on certain criteria for ageing [See, Bryson and Siddique (1969), Rolski (1975), Klefsjo (1980), Basu and Ebrahami (1986), Singh and Deshpande (1985), Abouammoh (1988) and Jinhua, Cao and Wang (1991)].

One of the methods of describing the failure mechanism is to expose the manner in which its life length is affected by the advancement of age. Usually by ageing we mean that an older component has a shorter remaining lifetime than a newer or younger one. No ageing is equivalent to saying that, the age of a system has no effect on the distribution of the residual lifetime. Positive ageing implies that the age has an adverse effect on the residual lifetime. That is the residual lifetime tends to be smaller in some probabilistic sense with increasing age. Negative ageing describes that the age has a beneficial effect on the residual lifetime. If the same type of ageing persists throughout the entire lifetime of a unit, the system is said to have monotonic ageing.

The phenomena of ageing had been first extensively studied by Bryson and Siddique (1969) and they had postulated a set of seven criteria for describing the ageing behaviour. Later, Basu and Ebrahami (1986) described, how ageing or wear out have been used to study lifetimes of systems and components. Abouammoh (1988) introduced a new criteria of ageing in terms of the conditional mean remaining life. The phenomenon of ageing can be described by using different reliability concepts such as failure rate, reliability function, MRL and VRL. In the present work, we discuss the ageing behaviour of the lifetime models belonging to the generalized Pearson system using failure rate and mean residual life function.

In reliability the ageing behaviour of the system is usually studied either by the failure rate function or by the mean residual life function. The increasing (decreasing) failure rate (IFR/DFR) property is a characteristic of the system that consistently deteriorate (improved) with age. This brings the relevance and need for classification of distributions based on failure rate function which provides information about the system reliability.

Definition 3.1

The distribution of X possess the increasing (decreasing) failure rate property if $h(x)$ is an increasing (decreasing) function of X.

In the following we discuss a method to identify an IFR(DFR) model in the generalized Pearson system (2.1)using

$$\beta(x) = \frac{-f'(x)}{f(x)}.$$

The function $\beta(x)$ was used earlier by Glaser (1980) for the analysis of bathtub models. Later Mukherjee and Roy (1989) used $\beta(x)$ to characterize certain lifetime models. An important feature of the procedure is that the method can be applicable to the most of the models used in the lifetime data analysis.

Lemma 3.1

Suppose that the distribution of X belongs to the generalized Pearson system (2.1). Let $\beta'(x)$ denote the derivative of $\beta(x)$ with respect to x. Then for $b_2 \neq 0$ in (2.1)

(A) $\beta'(x)>0$ if (i) $p_2>0$ and either

(a) $\Delta=0$ and $x \neq -d$ or

(b) $\Delta<0$ or

(c) $\Delta>0$ and $x \notin (\alpha,\beta)$ or

(ii) $p_2<0$, $\Delta>0$ and $x \in (\alpha,\beta)$ and

(B) $\beta'(x)<0$ if (i) $p_2<0$ and either

(a) $\Delta=0$ and $x \neq -d$ or

(b) $\Delta<0$ or

(c) $\Delta>0$ and $x \notin (\alpha,\beta)$ or

(ii) $p_2>0$, $\Delta>0$ and $x \in (\alpha,\beta)$

where α and β are the roots of the equation

$$p_2x^2 + 2p_2dx + p_1d - p_0 = 0 \tag{3.1}$$

with

$$p_i = \frac{b_i b_2}{a_1 b_2 - a_2 b_1}, \text{ i= } 0,1,2, \quad b_2 a_1 \neq a_2 b_1 \tag{3.2}$$

$$d = \frac{a_0 b_2 - a_2 b_0}{a_1 b_2 - a_2 b_1} \tag{3.3}$$

and

$$\Delta = 4p_2^2 d^2 - 4p_2(p_1 d - p_0).$$

Proof

The generalized Pearson system (2.1) can be written as

$$\frac{d\log f}{dx} = c + \frac{x+d}{p_0 + p_1 x + p_2 x^2} \tag{3.4}$$

where

$c = \frac{a_2}{b_2}$ and p_i's and d are given in (3.2) and (3.3).

From (3.4), we have

$$\beta(x) = \frac{-f'(x)}{f(x)} = -\left[c + \frac{x+d}{p_0 + p_1 x + p_2 x^2}\right]$$

and hence

$$\beta'(x) = \frac{p_2 x^2 + 2p_2 dx + p_1 d - p_0}{(p_0 + p_1 x + p_2 x^2)^2}. \tag{3.5}$$

Thus from (3.5), it is obvious that the sign of $\beta'(x)$ is determined by the sign of equation (3.1). Then from the elementary algebra we have the following results.

If $\Delta = b^2 - 4ac$ is the discriminant of the expression $ax^2 + bx + c = 0$, we have

(a) if $\Delta=0$, ax^2+bx+c has the same sign as that of 'a' for all $x \neq \frac{-b}{2a}$ and $ax^2+bx+c=0$ when $x = \frac{-b}{2a}$

(b) if $\Delta<0$, ax^2+bx+c has the same sign as 'a' for all real x.

(c) if $\Delta>0$ and the roots of $ax^2+bx+c=0$ are α and β with $\alpha>\beta$, then

(i) ax^2+bx+c has the same sign as that of 'a' whenever $x>\alpha$ or $x<\beta$ and

(ii) ax^2+bx+c has the sign opposite that of 'a' whenever $\beta<x<\alpha$.

Thus $\beta'(x)>0$, when (A) holds and $\beta'(x)<0$, when (B) holds. This completes the proof.

Theorem 3.1

A distribution belonging to the generalized Pearson family (2.1) has IFR property in a region if condition (A) of Lemma 3.1 holds in that region and has DFR property in a region if condition (B) of Lemma (3.1) holds in that region.

Proof

The proof directly follows from Lemma (3.1) and theorem given in Glaser (1980 p,667).

Remark 3.1

When $b_2=0$ in (2.1), (3.5) becomes

$$\beta'(x)=\frac{a_0b_1-a_1b_0-2a_2b_0x-a_2b_1x^2}{(b_0+b_1x)^2}. \tag{3.6}$$

Thus $\beta'(x)>(<)\ 0$ according as $a_0b_1 - a_1b_0 - 2a_2b_0x - a_2b_1x^2 >(<)0$. In this case, we have

$$p_1d - p_0 = a_0b_1 - a_1b_0,\ 2p_2d = -2a_2b_0 \text{ and } p_2 = -a_2b_1.$$

Corollary 3.1

When $a_2 = 0$, the above result reduces to the Pearson family given by Sankaran and Sindu (2001).

For the verification of the theorem, consider inverse Gaussian distribution with p.d.f (2.11). Then we have

$$\beta(x) = \frac{-\lambda\mu^2 + \lambda x^2 + 3x\mu^2}{2x^2\mu^2}$$

and hence

$$\beta'(x) = \frac{\lambda}{x^3} - \frac{3}{2x^2}.$$

Thus the distribution is IFR if

$$\frac{\lambda}{x^3} - \frac{3}{2x^2} > 0$$

or

$$0 < x < \frac{2\lambda}{3}$$

and DFR if

$$x > \frac{2\lambda}{3}.$$

Table 3.1 gives the region where the distribution possesses the IFR (DFR) property based on $\beta(x)$ for some popular models belonging to the family (2.1).

Table 3.1
The region where the distribution possesses the IFR (DFR) property

Sl. No.	Distributions with pdf	$\beta(x)$	Region
1	Gamma $\frac{m^p}{\Gamma(p)} e^{-mx} x^{p-1}$, $x>0$, $p,m>0$	$\frac{mx-(p-1)}{x}$	IFR if $p>1$ DFR if $0<p<1$
2	Pareto $ak^a x^{-(a+1)}$, $a>0$, $x \geq k > 0$	$\frac{a+1}{x}$	DFR
3	Normal $\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\}$, $-\infty<x<\infty$, $\sigma>0$, $-\infty<\mu<\infty$	$\frac{x-\mu}{\sigma^2}$	IFR
4	Finite Range $\frac{d}{R}\left(1-\frac{x}{R}\right)^{d-1}$, $0<x<R$, $d>1$.	$\frac{d-1}{R-x}$	IFR
5	Exponential $\lambda e^{-\lambda x}$, $x>0$, $\lambda>0$	λ	Both IFR and DFR
6	Inverse Gaussian $\sqrt{\frac{\lambda}{2\pi x^3}} \exp\left\{\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right\}$ $x>0$, $\lambda.\mu>0$	$\frac{-\lambda\mu^2+\lambda x^2+3\mu^2 x}{2\mu^2 x^2}$	IFR if $0< x < \frac{2\lambda}{3}$ DFR if $x> \frac{2\lambda}{3}$
7	Maxwell $4\sqrt{\frac{\lambda^3}{\pi}} x^2 e^{-\lambda x^2}$, $x>0$, $\lambda>0$	$\frac{2\lambda x^2-2}{x}$	IFR
8	Rayleigh $2\lambda x e^{-\lambda x^2}$, $x>0$, $\lambda>0$	$\frac{2\lambda x^2-1}{x}$	IFR

3.2 Characterization using Mean Residual Life

As mentioned in Chapter I, mean residual life (MRL) function is extensively used in the analysis of lifetime data. It is shown that an increasing (decreasing) failure rate class of distributions is a subclass of decreasing (increasing) MRL class of distributions. In the following we prove a characterization result for IFR (DFR) class of distributions in the generalized Pearson family using MRL function.

Definition 3.2

Let X be a non-negative continuous random variable with survival function $R(x)$. Then the distribution of X is said to have increasing mean residual life (IMRL) property if

$$r(x) = \frac{1}{R(x)} \int_x^\infty R(t)dt$$

is increasing in $x>0$ and have decreasing mean residual life (DMRL) property if

$$r(x) = \frac{1}{R(x)} \int_x^\infty R(t)dt$$

is decreasing in $x > 0$.

Theorem 3.2

Let the distribution of X belong to the generalized Pearson family (2.1) with$(b_0+b_1x+b_2x^2)\geq 0$. Then X has IFR (DFR) property if and only if

$$a_2\, V(x) + r(x)\, [a_2x+a_1] + m_1(x)[a_2\, r(x)+2b_2] + b_1 \leq (\geq)\ 0 \qquad (3.7)$$

where

$$V(x)=\mathrm{E}[(X-x)^2 \mid X>x],$$

$$r(x)=\mathrm{E}[(X-x) \mid X>x]$$

and

$$m_1(x)=\mathrm{E}[X \mid X>x].$$

Proof

When the distribution of X belongs to the family (2.1), we have (2.29),

$$a_2\, m_2(x) + (a_1+2b_2)\, m_1(x) + a_0+b_1 + (b_0+b_1x+b_2x^2)h(x)=0 \quad (3.8)$$

where

$$m_2(x) = \mathrm{E}[X^2 \,|\, X>x].$$

Using (1.12) and

$$m_2(x) = V(x) + 2x\, m_1(x) - x^2 + r^2(x)$$

(3.8) becomes

$$a_2\, [V(x) + 2x(x+r(x)) - x^2 + r^2(x)] + (a_1+2b_2)(x+r(x)) + a_0+b_1 + (b_0+b_1x+b_2x^2)h(x)=0. \quad (3.9)$$

Differentiating (3.9), we obtain

$$a_2\, [V'(x) + 2(xr'(x)+ r(x)) +2x +2r(x)\, r'(x)] + (a_1+2b_2)[1+ r'(x)] +(b_0+b_1x+b_2x^2)\, h'(x) +(b_1+2b_2x)\, h(x)=0. \quad (3.10)$$

Since $V'(x) = h(x)[\, V(x) - r^2(x)]$ and $h(x) = \dfrac{1+r'(x)}{r(x)}$, (3.10) becomes

$$a_2\, [h(x)[\, V(x) - r^2(x)]+h(x)\, r(x)[2x+2r(x)]] + (a_1+2b_2)\, h(x)\, r(x) +(b_0+b_1x+b_2x^2)\, h'(x) +(b_1+2b_2x)\, h(x)=0$$

which provides,

$$h(x)[a_2\, V(x) - a_2\, r^2(x) + r(x)[2a_2x + 2a_2\, r(x) + a_1+2b_2]+(b_1+2b_2x)] +(b_0+b_1x+b_2x^2)\, h'(x)=0. \quad (3.11)$$

If $(b_0+b_1x+b_2x^2)\geq 0$, then $h'(x) \geq 0$, if and only if

$$a_2\, V(x) + r(x)\, [a_2x+a_1] + m_1(x)[a_2\, r(x)+2b_2] +b_1 \leq (\geq)\ 0.$$

This completes the proof.

For the Pearson family, Theorem 3.2 reduces to a much simpler form, as shown below.

Theorem 3.3

Let the distribution of X belongs to the Pearson family (1.18). Then X has IFR/DFR property if and only if for every x in (a, b)

$$r(x) \geq (\leq)\ c_1 + 2c_2 x$$

where

$$c_i = \frac{b_i}{1-2b_2},\ i = 1, 2.$$

Proof

Since X belongs to the Pearson family (1.18), we have (Nair and Sankaran, 1991)

$$r(x)+x = (c_0+c_1x+c_2x^2)\ h(x)+\mu \tag{3.12}$$

where

$$\mu = \frac{b_1-d}{1-2b_2} \text{ and } c_i = \frac{b_i}{1-2b_2},\ i = 0,\ 1,\ 2.$$

Differentiating (3.12) with respect to x and substituting the relationship between $h(x)$ and $r(x)$,

$$h(x)\ r(x) = 1 + r'(x)$$

we get,

$$h(x)[\ r(x)-c_1-2c_2x] = (c_0+c_1x+c_2x^2)\ h'(x). \tag{3.13}$$

Now, to prove $(c_0+c_1x+c_2x^2) \geq 0$, for all x, we consider

$$m_1(x) - m_1(a) = (c_0+c_1x+c_2x^2)\ h(x) \tag{3.14}$$

Since $m_1(x)$ is non decreasing and $h(x) \geq 0$, (3.14) gives

$$(c_0+c_1x+c_2x^2) \geq 0, \text{ for all } x.$$

Thus from (3.13) $h'(x) \geq (\leq) 0$ if and only if $r(x) \geq (\leq)\ c_1 + 2c_2x$. This completes the proof.

For example, consider

(i) the Lomax distribution with p.d.f

$$f(x) = c\alpha^c\,(x+\alpha)^{-(c+1)},\ x>0,\ \alpha>0,\ c>1$$

then, $c_1 = \dfrac{\alpha}{c-1}$, $c_2 = \dfrac{1}{c-1}$ and $r(x) = \dfrac{x+\alpha}{c-1}$.

Since $r(x) < c_1 + 2c_2x$, the Lomax distribution has DFR property.

(ii) Consider the finite range distribution with p.d.f

$$f(x) = \frac{d}{R}\left(1-\frac{x}{R}\right)^{d-1},\ 0<x<R,\ d>0$$

then $c_1 = \dfrac{R}{d+1}$, $c_2 = \dfrac{-1}{d+1}$ and $r(x) = \dfrac{R-x}{d+1}$.

Since $r(x) > c_1 + 2c_2x$, the finite range distribution has IFR property.

3.3 Form Invariant Length Biased Models from the Generalized Pearson Family

We have discussed basic properties and various applications of length biased models in chapter I. Gupta and Keating (1986) observed that it is worthwhile to investigate the structural relationships between the distributions of X and Y in the context of reliability. Later Jain, Singh and Bagai (1989) extended the Gupta-Keating results for an arbitrary weight function $w(x) > 0$. The major relationship established by Gupta and Keating (1986) are

(i) $G(x) = \dfrac{m(x)}{\mu}R(x)$

(ii) $k(x) = \dfrac{x}{m(x)}h(x)$

(iii) $s(x) = \dfrac{r(x)}{m(x)}\displaystyle\int_x^{\infty}\frac{t+r(t)}{r(t)}\exp\left\{-\int_x^t\frac{du}{r(u)}\right\}dt.$

where $G(x)$, $k(x)$ and $s(x)$ are respectively the survival function, failure rate and MRL of Y.

The above identities along with some characterization theorems cited in Gupta and Kirmani (1990) show how length biased sampling affects the original distribution and how the corresponding reliability characteristics change under such a scheme of sampling. While comparing the distribution under length biased sampling with the parent model, it will be of some definite advantage if the original distribution keeps the same form under length biased sampling, except possibly for a change in the parameters. This will lead to the form invariance property of length biased models and is described as follows.

According to Patil and Ord (1976), the distribution of X with p.d.f $f(x;\theta)$ is said to be form-invariant under length biased sampling of order α if the observed variable Y has the same distribution as X, with a change in parameter. In other words, $f(x;\theta) \equiv f(x;\eta)$. Also they proved that a necessary and sufficient condition for X to be form invariant under size-bias of order α is that its p.d.f belongs to the log-exponential family. Some important members of this family are the log normal, Pareto, gamma, beta first and second kinds and Pearson type V. Motivated by the relevance of form-invariance in characterizing families of distributions ,Sankaran and Nair (1993) derived the condition under which the Pearson family is form-invariant with respect to the length biased sampling. They proved that, the members of the Pearson family satisfying the differential equation (1.18) with $b_2 \neq 1$, have the same type of distribution for Y if and only if $b_0=0$ and the p.d.f of Y satisfies

$$\frac{d\log g(x)}{dx} = \frac{-(x+d_1)}{c_1 x + c_2 x^2}$$

where

$$c_i = \frac{b_i}{1-2b_2} \qquad i = 1,2$$

and

$$d_1 = \frac{d-b_1}{1-b_2}.$$

Now we prove a general theorem in this direction by identifying those distributions of X belonging to the generalized Pearson family (2.1) that retain the same form of the distribution of Y. Note that we restrict our study to distribution of non-negative random variables belonging to the family (2.1).

Theorem 3.4

Among the members of (2.1), X and Y have the same type of distributions if and only if $b_0=0$ and the probability density function of Y satisfies

$$\frac{d\log g(x)}{dx} = \frac{p_o + p_1 x + p_2 x^2}{q_1 x + q_2 x^2} \tag{3.15}$$

where p_0, p_1, p_2, q_1 and q_2 are real constants.

Proof

Suppose that (2.1) holds and that X and Y have the same distributional form. Then from (1.19), we have

$$\frac{d\log g(x)}{dx} = \frac{f'(x)}{f(x)} + \frac{1}{x} \tag{3.16}$$

$$= \frac{a_0 + a_1 x + a_2 x^2}{b_0 + b_1 x + b_2 x^2} + \frac{1}{x}$$

or

$$\frac{d\log g(x)}{dx} = \frac{a_0x + a_1x^2 + a_2x^3 + b_0 + b_1x + b_2x^2}{x(b_0 + b_1x + b_2x^2)} . \tag{3.17}$$

Since Y also must belong to the same family, the equation (3.17) must be of the form,

$$\frac{p_o + p_1x + p_2x^2}{q_0 + q_1x + q_2x^2} = \frac{a_0x + a_1x^2 + a_2x^3 + b_0 + b_1x + b_2x^2}{x(b_0 + b_1x + b_2x^2)} . \tag{3.18}$$

Equating the coefficients of like powers of x in (3.18), we have six equations

$$q_0b_0 = 0 \tag{3.19}$$

$$q_0(a_0 + b_1) + q_1b_0 = p_0b_0 \tag{3.20}$$

$$q_1(a_0 + b_1) + q_2b_0 + q_0(a_1 + b_2) = p_1b_0 + p_0b_1 \tag{3.21}$$

$$q_0a_2 + q_1(a_1 + b_2) + q_2(a_0 + b_1) = p_1b_1 + p_2b_0 + p_0b_2 \tag{3.22}$$

$$(a_1 + b_2)\, q_2 + a_2q_1 = p_1b_2 + p_2b_1 \tag{3.23}$$

$$a_2q_2 = p_2b_2 \tag{3.24}$$

Now consider the equation (3.19), we have three cases,

(i) when $b_0 \neq 0$, $q_0 = 0$ in (3.19), we get

$$\frac{d\log f(x)}{dx} = \frac{a_2}{b_2} + \frac{(a_1b_2 - a_2b_1)x + a_0b_2 - a_2b_0}{b_2(b_0 + b_1x + b_2x^2)}$$

and

$$\frac{d\log g(x)}{dx} = \frac{p_2}{q_2} + \frac{(p_1q_2 - p_2q_1)x + p_0q_2}{q_2(q_1x + q_2x^2)}$$

in which $f(x)$ and $g(x)$ have different forms.

Similarly for case (ii) $b_0 = 0$, $q_0 \neq 0$, $f(x)$ and $g(x)$ have different forms.

On the other hand, for case (iii) when $b_0 = 0$, $q_0 = 0$, we get

$$\frac{d\log f(x)}{dx} = \frac{a_2}{b_2} + \frac{(a_1b_2 - a_2b_1)x + a_0b_2}{b_2(b_1x + b_2x^2)} \tag{3.25}$$

and

$$\frac{d\log g(x)}{dx} = \frac{a_2}{b_2} + \frac{(a_1b_2 - a_2b_1 + b_2^2)x + (a_0 + b_1)b_2}{b_2(b_1x + b_2x^2)}. \qquad (3.26)$$

Since the roots of the quadratic equation in the denominators of equations (3.25) and (3.26) are the same, $f(x)$ and $g(x)$ have same distributional form though with possibly different parameters. Conversely suppose that $b_0 = 0$ in (2.1), then from (3.16), we have

$$\frac{f'(x)}{f(x)} = \frac{g'(x)}{g(x)} - \frac{1}{x}$$

$$= \frac{p_0 + p_1x + p_2x^2}{(q_1x + q_2x^2)} - \frac{1}{x}$$

$$= \frac{(p_0 - q_1) + (p_1 - q_2)x + p_2x^2}{(q_1x + q_2x^2)}$$

which is of the form (2.1) with $a_0 = p_0 - q_1$, $a_1 = p_1 - q_2$, $a_2 = p_2$, $b_1 = q_1$ and $b_2 = q_2$. This completes the proof.

Corollary 3.2

When $a_2 = 0$, Theorem 3.4 reduces to the result of Sankaran and Nair (1993) concerning the original Pearson family.

To verify Theorem 3.4, consider the generalized inverse Gaussian distribution with p.d.f (Johnson, Kotz and Balakrishnan, 1994)

$$f(x) = Kx^{\nu} \exp\left\{-\frac{\lambda(x-\mu)^2}{2x\mu^2}\right\}, \; x>0, \; r, \; \lambda, \; \mu>0 \qquad (3.27)$$

where K is a normalizing constant.

The length biased distribution (LBD) for (3.27) is obtained as

$$g(x) = K' x^{\nu+1} \exp\left\{-\frac{\lambda(x-\mu)^2}{2x\mu^2}\right\}, \; x>0, \; r, \; \lambda, \; \mu>0$$

which has the same form as (3.27), but different parameters.

Sankaran and Nair (1993) derived the conditions under which models belonging to the Pearson family retain the same form for their length biased distribution. But there are situations where both the original and LBD do not have the same form when they belong to Pearson family.

For example

1. Consider the exponential distribution with p.d.f

$$f(x) = \lambda\, e^{-\lambda x} \quad x>0,\ \lambda>0$$

then LBD is obtained as

$$g(x) = \lambda^2\, x\, e^{-\lambda x} \quad x>0,\ \lambda>0$$

which has not the same form as X. In fact Y is gamma.

2. When X is Pareto type II with p.d.f

$$f(x) = c\alpha^c\, (x+\alpha)^{-(c+1)},\ x>0,\ \alpha>0,$$

then LBD of X is

$$g(x) = c(c-1)\alpha^c\, x\, (x+\alpha)^{-(c+1)},\ x>0, c>1.$$

Here also $g(x)$ has different form but both X and Y belong to the Pearson family.

3. When X is half normal with p.d.f

$$f(x) = \sqrt{\frac{2}{\pi}}\,\frac{1}{\sigma}\,\exp\left\{-\frac{x^2}{2\sigma^2}\right\},\quad x\geq 0,\ \sigma>0$$

Then LBD of X is obtained as,

$$g(x) = \frac{x}{\sigma^2}\,\exp\left\{-\frac{x^2}{2\sigma^2}\right\},\quad x\geq 0$$

and hence

$$\frac{g'(x)}{g(x)} = \frac{\sigma^2 - x^2}{\sigma^2 x}$$

which is of the form (2.1). Therefore from example 3, we can infer that, it is not necessary that X is in the Pearson family Y also belongs to that family. In this direction, next we investigate the condition under which the length biased distribution of X belong to the generalized Pearson family when the original belongs to the Pearson family.

Suppose that the distribution of X belongs to the family (1.18) and that of Y belongs to the family (2.1), we have

$$\frac{d\log g(x)}{dx} = \frac{d\log f(x)}{dx} + \frac{1}{x}.$$

Thus

$$\frac{p_0 + p_1 x + p_2 x^2}{q_0 + q_1 x + q_2 x^2} = \frac{-(x+d)}{b_0 + b_1 x + b_2 x^2} + \frac{1}{x}$$

which leads to

$$(p_0 + p_1 x + p_2 x^2)(b_0 + b_1 x + b_2 x^2)x = (q_0 + q_1 x + q_2 x^2)$$
$$(b_0 + (b_1 - d)x + (b_2 - 1)x^2). \quad (3.28)$$

Equating the coefficients of like powers of x in (3.28), we have the following six equations

$$q_0 b_0 = 0 \quad (3.29)$$

$$q_0(b_1 - d) + q_1 b_0 = p_0 b_0 \quad (3.30)$$

$$q_2 b_0 + q_1(b_1 - d) + q_0(b_2 - 1) = p_1 b_0 + p_0 b_1 \quad (3.31)$$

$$q_1(b_2 - 1) + q_2(b_1 - d) = p_1 b_1 + p_2 b_0 + p_0 b_2 \quad (3.32)$$

$$(b_2 - 1)\, q_2 = p_1 b_2 + p_2 b_1 \quad (3.33)$$

$$p_2 b_2 = 0. \quad (3.34)$$

Now we have the following cases arising from (3.29) and (3.34).

Case I

When $b_0 = 0$, $b_2 = 0$, $q_0 \neq 0$ and $p_2 \neq 0$, we obtain

$$b_1 = d = \frac{-q_0}{p_0} = \frac{-q_1}{p_1} = \frac{-q_2}{p_2}.$$

This leads to exponential distribution with parameter $\frac{-1}{d}$ for Y.

Case II

When $b_0 = 0$, $p_2 = 0$, $q_0 \neq 0$ and $b_2 \neq 0$, we get

$$b_1 = d \text{ and } b_2 = 1 + \frac{dp_0}{q_0}.$$

This provides

$$\frac{d \log g(x)}{dx} = \frac{dp_0}{q_0 + (dp_0 + q_0)x}$$

or

$$g(x) = Y_0 \left[q_0 + (dp_0 + q_0)x\right]^{\frac{dp_0}{dp_0 + q_0}}, \; 0 < x < \infty$$

leading to Lomax law, where Y_0 is the normalizing constant.

Case III

When $b_2 = 0$, $q_0 = 0$, we obtain

$$p_0 = q_1, \; b_1 = \frac{-q_2}{p_2} \text{ and } b_0 = \frac{dq_1}{q_2 - p_1} \; (p_1 \neq q_2).$$

This provides

$$\frac{d \log g(x)}{dx} = \frac{q_1 + p_1 x + p_2 x^2}{q_1 x + q_2 x^2}.$$

Therefore the distribution of Y belongs to generalized Pearson family (2.1).

Case IV

When $p_2 = 0$, $q_0 = 0$, we have,

$$p_0 = q_0,\ b_0 = \frac{dp_0}{q_2 - p_1},\ b_1 = \frac{q_1 + dq_2}{q_2 - p_1} \text{ and } b_2 = \frac{q_2}{q_2 - p_1}$$

and hence

$$\frac{d \log g(x)}{dx} = \frac{q_1 + p_1 x}{q_1 x + q_2 x^2}$$

or

$$g(x) = Y_0\ x\ (q_1 + q_2 x)^{\frac{p_1 - q_2}{q_2}}\ ,\ 0<x<\infty$$

that leads to beta distribution with Y_0 as normalizing constant.

Case V

When $p_2 = 0$, $q_0 = 0$, $b_0 = 0$, $b_2 = 0$, we have,

$$q_2 = 0,\ b_1 = \frac{dq_1}{q_1 - p_0}.$$

This provides

$$\frac{d \log g(x)}{dx} = \frac{(p_0 - q_1)x + d(q_1 + p_0)}{dq_1 x}$$

or

$$g(x) = Y_0\ \exp\left\{\frac{p_0 - q_1}{dq_1} x\right\} x^{\frac{1+p_0}{q_1}}\ ,\ 0<x<\infty$$

which is Pearson type III (Gamma) distribution, where Y_0 as normalizing constant.

The cases other than the above turns out to be special cases of the case V, will not be discussed further. The above discussion leads to the following characterization theorems whose proofs are direct.

Theorem 3.5

Suppose that the distribution of X belongs to the Pearson family (1.18). Then the length biased distribution (LBD) of X is exponential with parameter $\frac{-1}{d}$, if and only if $b_0 = 0$ and $b_2 = 0$.

Theorem 3.6

Suppose that the distribution of X belongs to the Pearson family (1.18) and that of Y belongs to the family (2.1). Then the distribution of Y is

(a) Lomax (Pareto II) if and only if $b_0 = 0$ and $p_2 = 0$

(b) a member of generalized Pearson family if and only if $q_0 = 0$ and $b_2 = 0$

(c) beta if and only if $q_0 = 0$ and $p_2 = 0$

(d) Pearson Type III (Gamma) if and only if $q_0 = 0$, $p_2 = 0$, $b_0 = 0$ and $b_2 = 0$.

3.4 Characterization by Conditional Expectations

When $b_0 = 0$ in (2.1), we obtain a subclass of the generalized Pearson family. This subclass to be denoted by C, contains many distribution of interest in reliability analysis such as gamma, beta, inverted gamma and inverse Gaussian. In the following, we prove a characterization of the class C based on conditional moments.

Theorem 3.7

Let $\lim_{x \to b} (b_1 x + b_2 x^2) f(x) = 0$. Then $f(x)$ belongs to C if and only if

$$a_2 m_3(x) + [(a_1+3b_2)-a_2 x] m_2(x) = m_1(x) \, [(a_1+2b_2)x - (a_0+2b_1)] + x(a_0+b_1) \qquad (3.35)$$

where

$$m_i(x) = \mathrm{E}(X^i|X>x),\ i = 1, 2, 3.$$

Proof

Let $h(x)$ be the failure rate of X. From theorem 2.3, for the class C, we obtain,

$$-(b_1x+b_2x^2)\ h(x) = a_2m_2(x)+(\ a_1+2b_2)m_1(x)+\ a_0+b_1 \tag{3.36}$$

and

$$-m_1(x)(b_1x+b_2x^2)k(x)=a_2m_3(x)+(a_1+3b_2)m_2(x)+(a_0+2b_1)m_1(x) \tag{3.37}$$

where $k(x)$ is the failure rate of Y (LBD).

Using the identity given by Gupta and Kirmani (1990), we have,

$$\frac{k(x)}{h(x)} = \frac{x}{m_1(x)}$$

which gives,

$$a_2m_3(x)+(a_1+3b_2)m_2(x)+(a_0+2b_1)m_1(x)$$
$$= x[a_2m_2(x)+(a_1+2b_2)m_1(x)+a_0+b_1]. \tag{3.38}$$

Rearranging the terms in (3.38), we obtain (3.35).

Conversely suppose that (3.35) holds for all x, then we have,

$$a_2 \int_x^b t^3 f(t)\ dt + (a_1+3b_2-a_2x) \int_x^b t^2 f(t)\ dt$$

$$= xR(x)(a_0+b_1)+[(a_1+2b_2)x-(a_0+2b_1)] \int_x^b t f(t)\ dt. \tag{3.39}$$

Differentiating (3.39) with respect to x, we get

$$\frac{d\log f(x)}{dx} = \frac{a_0+a_1x+a_2x^2}{b_1x+b_2x^2}$$

which has the same form as (2.1) with $b_0 = 0$. This completes the proof.

Corollary 3.3

When $a_2 = 0$, Theorem 3.7 reduces to the result of Sankaran and Nair (1993).

Corollary 3.4

The distribution of X is inverse Gaussian with p.d.f (2.11) holds if and only if

$$2\mu^2\, m_3(x) + (\lambda x+3\mu^2)m_2(x) = m_1(x)\ (\mu^2 x - \lambda\mu^2)+ \lambda\mu^2 x.$$

Corollary 3.5

The relationship

$$-2\lambda\ m_3(x) + 2\lambda m_2(x) = 3m_1(x)+2x.$$

holds if and only if X has Rayleigh distribution with p.d.f (2.13).

Corollary 3.6

The relationship

$$-m\ m_2(x) = m_1(x)\ [1-m-p]+px.$$

holds if and only if X has gamma distribution with p.d.f

$$f(x) = \frac{m^p}{\Gamma(p)}\, e^{-mx}\, x^{p-1},\ x>0,\ m,\ p>0.$$

CHAPTER IV

A GENERALIZED ORD FAMILY OF DISTRIBUTIONS

4.1 Introduction

The Ord family of distributions is the discrete analogue of the Pearson family of continuous distributions. Many distributional properties of the members of this family can be obtained in the same manner as in the Pearson family. The similarity between exponential and geometric distributions, beta and hyper geometric distributions, gamma and negative binomial distributions etc, in the continuous and discrete set ups, makes the investigation of analogous results in the Ord family corresponding to these distributions in the Pearson family worthwhile. In view of the results concerning the generalized Pearson system obtained in the Chapters II and III and their usefulness in reliability modeling, in the present chapter we define an extended Ord family and explore the possibility of obtaining results that have applications when the observation are in the form of integer values.

Let X be a non-negative integer valued random variable with probability mass function (p.m.f) $p(x)$. Then the distribution of X belongs to the Ord family (Ord, 1972) if $p(x)$ satisfies the difference equation

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{-(x+u)}{k_0+k_1x+k_2x^2} \tag{4.1}$$

where k_0, k_1, k_2 and u are real constants.

Ord's classification of distributions is mainly depends on the nature of the roots of quadratic expression in the denominator of (4.1) or rather on the value of

$$k = \frac{k_1^2}{4k_0k_2}.$$

When $k_0 = k_1$ and k_2=0, yields Katz' family of distribution with

$$\frac{p(x+1)}{p(x)} = \frac{\alpha+\beta x}{1+x}, \ \alpha>0 \ \beta<1, \ x = 0, 1, \ldots$$

Ord (1972)labeled these III B, III P, III N for the binomial, Poisson and negative binomial respectively. Also for k_2=0 and $k_0 \neq k_1$ leads to a system of hyper distributions, in particular to the hyper-Poisson distribution. Discrete distributions like Poisson, binomial, negative binomial, hypergeometric, Waring etc that have applications in reliability analysis belonging to the family (4.1). For various properties and applications of (4.1) we refer to Ord (1972) and Johnson, Kotz and Kemp (1992). However there are other distributions that are not members of the family (4.1). For example, the distributions like confluent hypergeometric and Haight are not members of (4.1) [See, Ord (1972)].Motivated by this we consider an extension of (4.1) and study its properties in the context of reliability. Earlier Davies (1934) considered an extension to the Ord family in terms of hyper functions. Later Bowman, Shenton and Kastenbaum (1991) [See, Johnson, Kotz and Kemp (1992)] have studied an extension of Ord's family with

$$p(x) = \left(1+\frac{\alpha - x}{c_0+c_1y+c_2y^2}\right) p(x\text{-}1) \tag{4.2}$$

where $y=x-\mu$ and μ = E(X), the ratio of successive probabilities is here the ratio of two quadratic expressions in x.

In the following section we study an extended version of Ord family, which is a generalization of (4.2), in the context of reliability analysis.

4.2 A Generalized Ord Family of Distributions

Let X be a discrete random variable as stated in Section 4.1. The distribution of X belongs to the generalized Ord family if the p.m.f of X satisfies

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{c_0+c_1x+c_2x^2}{d_0+d_1x+d_2x^2} \tag{4.3}$$

where c_0, c_1, c_2, d_0, d_1 and d_2 are real constants.
Obviously when c_2=0, (4.3) reduces to the form (4.1). The roots of the equation $d_0+d_1x+d_2x^2=0$, appeared in the denominator of (4.3) determines various members of the family (4.3).

The family (4.3) can be written as

$$\frac{p(x+1)-p(x)}{p(x)} = c + \frac{x+u}{k_0+k_1x+k_2x^2} \tag{4.4}$$

where

$$c = \frac{c_2}{d_2}, \quad u = \frac{c_0d_2-c_2d_0}{c_1d_2-c_2d_1}$$

and

$$k_i = \frac{d_id_2}{c_1d_2-c_2d_1}, i = 0,\ 1,\ 2, \quad c_1d_2 \neq c_2d_1.$$

When $c=0$, (4.4) reduces to (4.2). The different types of curves are based on the nature of the roots of the quadratic expression in the denominator of (4.4).

4.3 Members of the Generalized Ord Family

As mentioned in the previous section, all members of the Ord family are the members of (4.3). Some important distributions other than the members of the Ord family belonging to the family (4.3) are given below.

1. Kemp family of distributions [Johnson, Kotz and Kemp(1992)] can be written as

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{(e_1+x)(e_2+x)\lambda}{(g_1+x)(g_2+x)}. \tag{4.5}$$

When, $c_0 = e_1e_2\lambda$, $c_1 = (e_1+e_2)\lambda$, $c_2=\lambda$, $d_0=g_1g_2$, $d_1=(g_1+g_2)d_2$ and $d_2=1$, (4.5) reduces to (4.3).

2. Consider confluent hyper geometric distribution defined by Bhattacharya (1966) with

$$p(x)=\frac{\Gamma(\gamma+x)\Gamma(1+b)}{\Gamma(1+b+x)\Gamma(\gamma)\phi(\gamma;1+b;\theta)}\frac{\theta^x}{x!}\quad, x = 0,1,2, \tag{4.6}$$

where

$$\phi(\gamma;1+b;\theta)=1+\frac{\gamma}{(1+b)1!}\theta+\frac{\gamma(\gamma+1)}{(1+b)(b+2)2!}\theta^2+\ldots$$

$$=\sum_{j=0}^{\infty}\frac{(\gamma)_j\theta^j}{(1+b)_j\,j!}$$

with $(\gamma)_j$ as Pochhammer's symbol. (4.6) can also be written in the form,

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{-x^2+x(\theta-b-2)+\gamma\theta-b-1}{(x+1)(x+b+1)}$$

where

$c_0 = \gamma\theta - b - 1$, $c_1 = \theta - b - 2$, $c_2 = -1$, $d_0 = b+1$, $d_1 = b+2$ and $d_2 = 1$.

Thus (4.6) is a member of the generalized Ord family (4.3).

The confluent hypergeometric distribution has found application in the theory of accident proneness.

When $\gamma = 1$, equation (4.6) reduces to the hyper Poisson distribution with

$$p(x) = Y_0 \frac{\theta^x \Gamma(1+b)}{\Gamma(1+b+x)}, \quad x=0,1\ldots \tag{4.7}$$

where

$$Y_0 = \frac{1}{\phi(1;1+b;\theta)}.$$

This distribution arises in birth-and-death processes (Hall 1956).

3. Consider the Borel-Tanner distribution (Tanner, 1953) with p.m.f

$$p(x) = \frac{e^{-nq}(nq)^{n-x}.x}{(n-x)!n}, \qquad x = 1,2,\ldots \tag{4.8}$$

which can be written as

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{-x^2 + x[n(1-q)-1]+n}{nqx}$$

Thus (4.8) has the same form as (4.3) with $c_0=n$, $c_1=n(1-q)-1$, $c_2=-1$, $d_0=d_2=0$ and $d_1=nq$.

4. Consider the Haight distribution (Haight, 1961) with p.m.f

$$p(x) = \frac{\binom{2n-x-1}{n-1} x\,\alpha^{n-x}}{n(1+\alpha)^{2n-x}}, \quad x = 1,2,\ldots \tag{4.9}$$

(4.9) can be written in the form,

$$\frac{p(x+1)-p(x)}{p(x)}=\frac{-x^2+x[2n-1+\alpha(n-1)]+2n(\alpha+1)}{\alpha x(n-x)}$$

which has the same form as (4.3) with $c_0=2n(\alpha+1)$, $c_1=2n-1+\alpha(n-1)$, $c_2=-1$, $d_0=0$, $d_1=n\alpha$ and $d_2=-\alpha$. The applications of the distributions (4.8) and (4.9) in the queuing theory have already been studied (Ord,1972).

4.4 Properties of the Generalized Ord Family of Distributions

In this section we discuss some important properties of the generalized Ord family.

Property 1

For the generalized Ord family (4.3), the recurrence relationship among the raw moments is obtained as

$$d_0\sum_{j=0}^{r}\binom{r}{j}(-1)^j\mu'_{r-j}+d_1\sum_{j=0}^{r+1}\binom{r+1}{j}(-1)^j\mu'_{r+1-j}+d_2\sum_{j=0}^{r+2}\binom{r+2}{j}(-1)^j\mu'_{r+2-j}$$

$$=\mu'_r(c_0+d_0)+\mu'_{r+1}(c_1+d_1)+\mu'_{r+2}(c_2+d_2)+(-1)^r(d_0-d_1)+d_2)\ p(0) \quad (4.10)$$

where $\mu'_r=E(X^r)$.

Proof

From (4.3), we have

$$(d_0+d_1x+d_2x^2)p(x+1)=[\,(c_0+d_0)+(c_1+d_1)x+(c_2+d_2)x^2]\ p(x). \quad (4.11)$$

Multiplying both sides of (4.11) by x^r and taking summation from 0 to ∞, we get,

$$\sum_{x=0}^{\infty} x^r \ (d_0+d_1x+d_2x^2)p(x+1)$$

$$=\sum_{x=0}^{\infty} x^r \ [(c_0+d_0)+(c_1+d_1)x+(c_2+d_{\partial 2})x^2] \ \ p(x)$$

which gives,

$$\sum_{x=1}^{\infty}(x-1)^r \ (d_0+d_1(x-1)+d_2(x-1)^2)p(x)$$

$$=\sum_{x=0}^{\infty} x^r \ [(c_0+d_0)+(c_1+d_1)x+(c_2+d_2)x^2] \ p(x). \qquad (4.12)$$

Thus (4.12) provides,

$$\sum_{x=0}^{\infty}(x-1)^r \ (d_0+d_1(x-1)+d_2(x-1)^2)p(x)$$

$$=\sum_{x=0}^{\infty} x^r \ [(c_0+d_0)+(c_1+d_1)x+(c_2+d_2)x^2] \ p(x) \ + \ (-1)^r \ (d_0-d_1)+d_2) \ p(0)$$

which reduces to the form (4.10).

Property 2

Let $g(s)$ = $\mathrm{E}(s^x)$. Assume that the derivatives $g'(s)$ and $g''(s)$ with respect to s exists. Then the probability generating function $g(s)$ of the generalized Ord family (4.3) satisfies the relationship

$$g(s)\left[\frac{d_0-d_1+d_2}{s}-(c_0+d_0)\right]+g'(s)\,[d_1-d_2-(c_1+d_1)s-(c_2+d_2)s]$$

$$+\ g''(s)\ \ [d_2s-(c_2+d_2)s^2]+p(0)\left[\frac{d_0-d_1+d_2}{s}\right]=0. \qquad (4.13)$$

Proof

From (4.3), we have

$$\sum_{x=0}^{\infty} s^x \ (d_0+d_1x+d_2x^2)p(x+1) \ = \ \sum_{x=0}^{\infty} s^x \ [(c_0+d_0)+(c_1+d_1)x+(c_2+d_2)x^2] \ p(x). \qquad (4.14)$$

Put $x+1=u$ in the left side of (4.14), we get

$$d_0\sum_{x=1}^{\infty} s^{x-1}p(x)+d_1\sum_{x=1}^{\infty}(x-1)s^{x-1}p(x)+d_2\sum_{x=1}^{\infty}(x-1)^2 s^{x-1}p(x)$$

$$= (c_0+d_0)\sum_{x=1}^{\infty}s^x p(x) + (c_1+d_1)\sum_{x=0}^{\infty}xs^x p(x)+(c_2+d_2)\sum_{x=0}^{\infty}x^2s^x p(x)$$

which provides,

$$p(0)\left[\frac{d_0-d_1+d_2}{s}\right]+\frac{d_0}{s}g(s)+d_1g'(s)-\frac{d_1}{s}g(s)+d_2sg''(s)-d_2g'(s)+\frac{d_2}{s}g(s)$$

$$=(c_0+d_0)g(s)+(c_1+d_1)sg'(s)+(c_2+d_2)s^2g''(s)+(c_2+d_2)sg'(s). \quad (4.15)$$

Thus we obtain (4.13).

Property 3

If a turning point exists, a maximum of $p(x)$ occurs at a value of x for which,

$$\frac{p(x)}{p(x+1)}\geq 1 \qquad \text{and} \qquad \frac{p(x)}{p(x-1)}\geq 1$$

which can be written as,

$$\frac{p(x+1)}{p(x)}\leq 1 \qquad \text{and} \qquad \frac{p(x)}{p(x-1)}\geq 1. \qquad (4.16)$$

Thus from (4.3) and (4.16), we have

$$\frac{c_0+d_0+(c_1+d_1)x+(c_2+d_2)x^2}{d_0+d_1x+d_2x^2}\leq 1$$

and

$$\frac{c_0+d_0+(c_1+d_1)x+(c_2+d_2)x^2}{d_0+d_1x+d_2x^2}\geq 1$$

which provides

$$c_0 + c_1\, x + c_2\, x^2 \leq 0$$

and

$$c_0 + c_1\, x + c_2\, x^2 \geq 0.$$

Thus the maximum of $p(x)$ occurs when $x=\dfrac{c_1^2}{4c_0c_2}$, which is the mode of $p(x)$.

Property 4

An inflexion of $p(x)$ occurs at that value of x for which $\Delta^2 p(x)$ changes sign, where $\Delta\, p(x)= p(x+1)-p(x)$. This provides

$$\begin{aligned}\Delta^2\, p(x) &= \Delta[\Delta\, p(x)]\\ &= \Delta\, p(x+1)-\Delta\, p(x)\\ &= \frac{c_0+c_1(x+1)+c_2(x+1)^2}{d_0+d_1(x+1)+d_2(x+1)^2}\,p(x+1)-\left[\frac{c_0+c_1x+c_2x^2}{d_0+d_1x+d_2x^2}\right]p(x).\end{aligned}$$

When $\Delta^2\, p(x)=0$,

$$\left[c_0+c_1(x+1)+c_2(x+1)^2\right]\left[d_0+d_1x+d_2x^2\right]\,p(x+1)-$$

$$\left[c_0+c_1x+c_2x^2\right]\left[d_0+d_1(x+1)+d_2(x+1)^2\right]\,p(x)=0\,. \tag{4.17}$$

From (4.3), we have

$$p(x+1)=p(x)\left[\frac{c_0+d_0+(c_1+d_1)x+(c_2+d_2)x^2}{d_0+d_1x+d_2x^2}\right]. \tag{4.18}$$

Substituting (4.18) in (4.17), we obtain

$$\begin{aligned}&c_2^2x^4+x^3(2c_2^2+2c_1c_2-4d_2c_2)+x^2[c_2^2+c_1^2+2c_0c_2+3c_1c_2-3c_2d_1-3c_1d_2-2c_2d_2]\\ &\quad+x[c_1^2+2(c_0c_1+c_2c_0-c_2d_0-c_0d_2-c_1d_1)+c_1c_2-c_2d_1-c_1d_2]\\ &\qquad+c_0^2+c_0c_1-c_1d_0+c_2c_0-c_2d_0-c_0d_1-c_0d_2=0\,.\end{aligned} \tag{4.19}$$

The points of inflexion is obtained by solving the equation (4.19).

Property 5

If the distribution of X belongs to the generalized Ord family (4.3) with $E(X^{r+2})<\infty$, then

$$c_2\alpha_{r+2}(t)+\alpha_{r+1}(t)[\,c_1+c_2(\alpha t+\alpha r+1)+d_2(r+2)\,]+$$
$$\alpha_r(t)\,\{\,c_0+(t+r)[c_1+c_2(3t+3r+2)]+d_2(r+1)+d_2[r(2t+2r+3)+2t+1\,]\,\}+$$
$$r\alpha_{r-1}(t)\,\{\,d_0+(t+r)[d_1+d_2(3t+3r+2)]\,\}=0 \qquad (4.20)$$

where $\alpha_r(t)$ is defined in (1.26).

Proof

When the distribution of X belongs to generalized Ord family (4.3), we have

$$p(x+1)(\,d_0+d_1x+d_2x^2\,)=p(x)\,[c_0+d_0+(c_1+d_1)x+(c_2+d_2)x^2\,]. \qquad (4.21)$$

Multiplying (4.21) by $(x\text{-}t)^{(r)}$ and taking summation from $t+r$ to ∞, we get

$$\sum_{t+r}^{\infty}(x-t)^{(r)}\ \ (\,d_0+d_1x+d_2x^2\,)p(x+1)$$
$$=\ \sum_{t+r}^{\infty}(x-t)^{(r)}\ [c_0+d_0+(c_1+d_1)x+(c_2+d_2)x^2\,]\ p(x). \qquad (4.22)$$

Putting

$$x=\ (x\text{-}t\text{-}r)\ +\ (t+r)$$

and

$$x^2=(x\text{-}t\text{-}r)\ (x\text{-}t\text{-}r\text{-}1)\ +(x\text{-}t\text{-}r)(2t+2r+1)+(t+r)(3t+3r+2)$$

in (4.22), we obtain

$$d_0\sum_{t+r}^{\infty}(x-t)^{(r)}\ \ p(x+1)+\ d_1\sum_{t+r}^{\infty}(x-t-r)(x-t)^{(r)}\ \ p(x+1)$$
$$+\ d_1(t+r)\sum_{t+\gamma}^{\infty}(x-t)^{(r)}\ \ p(x+1)\ +d_2\sum_{t+r}^{\infty}(x-t-r)(x-t-r-1)(x-t)^{(r)}\ \ p(x+1)$$
$$+d_2(2t+2r\ +1)\ \sum_{t+r}^{\infty}(x-t)^{(r)}\,(x\text{-}t\text{-}\ r)\ p(x+1)$$
$$+d_2(t+\ r)(3t+3r\ +2)\ \sum_{t+r}^{\infty}(x-t)^{(r)}\ \ p(x+1)$$

$$= (c_0+d_0) \sum_{t+r}^{\infty}(x-t)^{(r)}p(x)+ (c_1+d_1) \sum_{t+r}^{\infty}(x-t)^{(r)} (x-t-r)\, p(x)$$

$$+(c_1+d_1)(t+r)\sum_{t+r}^{\infty}(x-t)^{(r)}\ p(x)$$

$$+ (c_2+d_2) \sum_{t+r}^{\infty}(x-t-r)(x-t-r-1)(x-t)^{(r)}p(x)$$

$$+(c_2+d_2)\ (2t+2r+1) \sum_{t+r}^{\infty}(x-t-r)(x-t)^{(r)}p(x)$$

$$+(c_2+d_2)\ (t+r)(3t+3r+2) \sum_{t+r}^{\infty}(x-t)^{(r)}\ p(x) \tag{4.23}$$

Putting $x+1= u$, in the left side of (4.23) and using the relationship

$$\alpha_r(t+1) = \alpha_r(t)-r\alpha_{r-1}(t),\ r\geq 1,\ t\geq 0,$$

we obtain,

$$d_2\alpha_{r+2}(t+1)+ d_2\alpha_{r+2}(t)+ \alpha_{r+1}(t)\ [d_1+d_2(2t+r-1)]+ \alpha_r(t)\ [d_0+d_1(t-1)$$
$$+d_2(3t^2+4tr+r^2-r-1)]\ -r\alpha_{r-1}(t)\ [d_0+d_1(t+r)+d_2(t+r)(3t+3r+2)]$$
$$=\alpha_r(t)\ [c_0+d_0+(c_1+d_1)(t+r)+(c_2+d_2)(t+r)(3t+3r+2)]$$
$$+\alpha_{r+1}(t)\,[c_1+d_1+(c_2+d_2)(2t+2r+1)]+\alpha_{r+2}(t)[c_2+d_2] \tag{4.24}$$

Rearranging (4.24), provides

$$c_2\alpha_{r+2}(t)+\alpha_{r+1}(t)\,[c_1+d_1+(c_2+d_2)(2t+2r+1)-\{d_1+d_2(2t+r-1)\}]$$
$$+\alpha_r(t)\{\,c_0+d_0+(c_1+d_1)(t+r)+(c_2+d_2)(t+r)(3t+3r+2)-d_0-d_1(t-1)$$
$$-d_2[3t^2+4tr+r^2-r-1]\}+r\alpha_{r-1}(t)\,[d_0+d_1(t+r)+d_2(t+r)(3t+3r+2)]=0$$

which leads to (4.20).

Remark 4.1

When c_2=0, the relationship (4.20) reduces to the result of Nair et. al (2000) for the Ord family.

4.5 Characterizations

Several characterizations of discrete distributions using the concepts such as failure rate, mean residual life, and vitality function have been extensively discussed by different researchers like Shanbhag (1970), Xekalaki (1983), Osaki and Li (1988), Ahmed (1991), Nair and Hitha (1989) etc. The characterizations of the Ord family of distributions using different reliability concepts were studied in Nair and Sankaran (1991), Glanzel (1991), Sankaran and Nair (1993), Ruiz and Navarro (1994), Navarro, Franco and Ruiz (1998).

Nair and Sankaran (1991) established the relationship

$$m(x) = \mu + (y_0+y_1x+y_2x^2)\ h(x+1)$$

where $\mu = E(X) = \dfrac{k_1-u}{1-2k_2}$ and $y_i = \dfrac{k_i}{1-2k_2}$, $i=0, 1, 2$, that characterizes the Ord family (4.1).

Now we prove a characterization theorem using the relationship between failure rate and conditional moments, which generalizes the result given by Nair and Sankaran (1991).

Theorem 4.1

A necessary and sufficient condition for the distribution of X belongs to the family (4.3) is that

$$c_2m_2(x)+(c_1+2d_2)m_1(x)+c_0+d_1-d_2+(d_0+d_1x+d_2x^2)h(x+1)=0 \qquad (4.25)$$

where

$$m_i(x)=E[X^i \mid X>x] \qquad i=1, 2.$$

Proof

When the distribution of X belongs to the family (4.3), we have

$$\sum_{x+1}^{\infty}[p(t+1)-p(t)]\,[d_0+d_1t+d_2t^2]= \sum_{x+1}^{\infty}(c_0+c_1t+c_2t^2)p(t)$$

which gives

$$-d_0p(x+1)-d_1(x+1)p(x+1) -d_1R(x+2) -d_2(x+1)^2p(x+1)$$

$$-2d_2\sum_{x+1}^{\infty}tp(t) +2d_2(x+1)p(x+1) +d_2R(x+2)$$

$$= c_0R(x+1) +c_1\sum_{x+1}^{\infty}tp(t) +c_2\sum_{x+1}^{\infty}t^2p(x). \qquad (4.26)$$

Dividing (4.26) by $R(x+1)$ and using the relationship

$$h(x+1) = 1-\frac{R(x+2)}{R(x+1)}$$

we obtain,

$$-d_0h(x+1)-d_1(x+1)h(x+1) -d_1[1-h(x+1)] -d_2(x+1)^2h(x+1)$$

$$-2d_2m_1(x)+2d_2(x+1)h(x+1) +d_2[1-h(x+1)] = c_0+c_1m_1(x)+c_2m_2(x)$$

which leads to (4.25).

Conversely when (4.25) holds, we have

$$c_2\sum_{x+1}^{\infty}t^2p(t) +(c_1+2d_2)\sum_{x+1}^{\infty}t\;p(t) +(c_0+d_1-d_2)\sum_{x+1}^{\infty}p(t)$$

$$+(d_0+d_1x+d_2x^2)p(x+1)=0\cdot \qquad (4.27)$$

Changing x to $(x-1)$ in (4.27) and subtracting the resulting expression from (4.27), we get (4.3). This completes the proof.

Corollary 4.1

When $a_2=0$, Theorem 4.1 reduces to the result of Nair and Sankaran (1991).

Corollary 4.2

The relationship

$$\lambda\, m_2(x) + [\lambda(e_1+e_2) + 2]\, m_1(x) + \lambda e_1 e_2 + g_1 + g_2 - 1 + (g_1 g_2 + (g_1+g_2)x + x^2)h(x+1) = 0$$

holds if and only if X belongs to the Kemp family (4.5).

Corollary 4.3

The distribution of X is confluent hyper geometric with p.m.f (4.6) if and only if

$$-m_2(x) + (\theta - b)m_1(x) + \nu\theta + (b+1+(b+2)x + x^2)h(x+1) = 0.$$

Corollary 4.4

The distribution of X is Borel-Tanner with p.m.f. (4.8) if and only if

$$-m_2(x) + [n(1-q)-1]m_1 x + n(1+q) + (nqx)h(x+1) = 0.$$

Corollary 4.5

The relationship

$$-m_2(x) + m_1(x)[2n-1+\alpha(n-3)] + (3n+1)\alpha + 2n + (\alpha nx - \alpha x^2)h(x+1) = 0$$

holds if and only if X has Haight distribution with p.m.f (4.9).

4.6 Characterization through some Conditional Moments of Generalized Ord Family

Glanzel (1991) proved that the distribution of the discrete random variable X belongs to the Ord family if and only if

$$E(X^2 \mid X \geq x) = P(x)\; E(X \mid X \geq x) + Q(x);$$

where $P(x)$ and $Q(x)$ are polynomials of degree one almost with real coefficients. The following theorem generalizes the result of Glanzel (1991).

Theorem 4.2

Let $(\Omega, \mathrm{F}, \mathrm{P})$ be a probability space and let X be an integer valued random variable. Assume that $E(X^3)<\infty$ and $x^3p(x)\to 0$ as $x\to\pm\infty$. Then the distribution of the random variable X belongs to the generalized Ord family (4.3) if and only if

$$c_2\, E(X^3 \mid X\geq x) = A(x)\, E(X^2 \mid X\geq x) + B(x)\, E(X \mid X\geq x) + C(x) \qquad (4.28)$$

where $A(x) = c_2x+q$ with q as a real constant and $B(x)$ and $C(x)$ are polynomials of degree one at most with real coefficients.

Proof

Let $B(x) = rx+t$ and $C(x) = ux+w$. Now assume that (4.28) holds, then

$$c_2\, E(X^3 \mid X\geq x) = (c_2x+q)\, E(X^2 \mid X\geq x) +(rx+t)\, E(X \mid X\geq x) + (ux+w). \qquad (4.29)$$

From (4.29), we obtain for all $x \in H$,

$$c_2 \sum_{x}^{\infty} t^3 p(t) = (c_2x+q) \sum_{x}^{\infty} t^2 p(t) +(rx+t) \sum_{x}^{\infty} t\, p(t) + (ux+w)R(x). \qquad (4.30)$$

Changing x to $(x+1)$ in (4.30), we obtain

$$c_2 \sum_{x+1}^{\infty} t^3 p(t)=(c_2(x+1)+q) \sum_{x+1}^{\infty} t^2 p(t)+(r(x+1)+t) \sum_{x+1}^{\infty} t\, p(t)$$
$$+(u(x+1)+w)R(x+1). \qquad (4.31)$$

Subtracting (4.31) from (4.30), we have

$$c_2\, x^3\, p(x) = (c_2x+q)\, x^2\, p(x) -p \sum_{x+1}^{\infty} t^2 p(t) +(rx+t)xp(x)-r \sum_{x+1}^{\infty} t\, p(t)$$
$$+ (ux+w)[R(x)- R(x+1)]- u\, R(x+1).$$

Repeating the same procedure as above after changing x to $x+1$, we obtain

$$p(x)\ [(q+r)x^2+(t+u)x+w] = p(x+1)\ [(c_2+q+r)x^2+(2\ c_2+2q+3r+t+u)x$$
$$+(\ c_2+t+w+2u)]$$

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{-[c_2x^2+x(2c_2+2q+3r)+c_2+t+2u]}{(c_2+q+r)x^2+(2q+2c_2+3r+t+u)x+(c_2+t+w+2u)}$$

which is of the form (4.3).

Conversely assume that (4.3) holds, then

$$\frac{p(x+1)}{p(x)} = \frac{c_0+d_0+(c_1+d_1)x+(c_2+d_2)x^2}{d_0+d_1x+d_2x^2} \tag{4.32}$$

or

$$(d_0+d_1x+d_2x^2)\ p(x+1) = p(x)\ [(c_0+d_0)+(c_1+d_1)x+(c_2+d_2)x^2]. \tag{4.33}$$

Taking summation on both sides of (4.33), we get

$$d_0\sum_x^{\infty}p(t)\ -\ d_0\ p(x)\ +\ d_1\ \sum_x^{\infty}t\,p(t)\ -\ d_1\ xp(x)\text{-}\ d_1\sum_x^{\infty}p(t)\ +\ d_1\ p(x)$$
$$+d_2\sum_x^{\infty}t^2\,p(t)\ \ -\ d_2\ x^2\ p(x)\text{-}2d_2\sum_x^{\infty}t\,p(t)\ \ +2d_2\ xp(x)\ +\ d_2\sum_x^{\infty}p(t)\ -\ d_2\ p(x)$$
$$= (c_2+d_2)\ \sum_x^{\infty}t^2\,p(t)\ +(c_1+d_1)\ \sum_x^{\infty}t\,p(t)\ +\ (c_0+d_0)\ \sum_x^{\infty}p(t)$$

which provides,

$$p(x)\ [d_0\text{-}d_1+d_2]\ +\ x\ p(x)\ [d_1\text{-}2d_2]\ +d_2\ x^2\ p(x)$$
$$= [d_2\text{-}d_1\text{-}c_0]\ \sum_x^{\infty}p(t)\ +\ [\text{-}2d_2\text{-}c_1]\ \sum_x^{\infty}t\ \ p(t)\ -\ c_2\ \sum_x^{\infty}t^2\,p(t). \tag{4.34}$$

Now consider,

$$(d_0x+d_1x^2+d_2x^3)\ p(x+1) = p(x)\ [(c_0+d_0)x+(c_1+d_1)x^2+(c_2+d_2)x^3]. \tag{4.35}$$

Taking summation on both sides of (4.35), we obtain

$$d_0 \sum_x^{\infty} t\,p(t) - d_0\, xp(x) + d_0 \sum_x^{\infty} p(t) + d_0\, p(x) + d_1 \sum_x^{\infty} t^2 p(t) - d_1\, x^2 p(x)$$
$$-2d_1 \sum_x^{\infty} t\,p(t) + 2d_1\, xp(x) + d_1 \sum_x^{\infty} p(t) - d_1\, p(x) + d_2 \sum_x^{\infty} t^3 p(t) - d_2\, x^3 p(x)$$
$$-3d_2 \sum_x^{\infty} t^2 p(t) + 3d_2\, x^2 p(x) + 3d_2 \sum_x^{\infty} t\,p(t) - 3d_2\, xp(x) - d_2 \sum_x^{\infty} p(t)$$
$$+ d_2\, p(x) = (c_2+d_2) \sum_x^{\infty} t^3 p(t) + (c_1+d_1) \sum_x^{\infty} t^2 p(t) + (c_0+d_0) \sum_x^{\infty} t\,p(t).$$

which gives,

$$p(x)\,[d_0-d_1+d_2] + x\,p(x)\,[d_1-2d_2] + d_2\, x^2 p(x) - x[p(x)\,(d_0-d_1+d_2)$$
$$+ x\,p(x)\,(d_1-2d_2) + d_2\, x^2 p(x)]$$
$$= [d_0+d_2-c_1] \sum_x^{\infty} p(t) + [c_0+2d_1-3d_2] \sum_x^{\infty} t\,p(t)$$
$$+[c_1+3d_2] \sum_x^{\infty} t^2 p(t) + c_2 \sum_x^{\infty} t^3 p(t). \quad (4.36)$$

Substituting (4.34) in (4.36), we get

$$[d_0+c_0] \sum_x^{\infty} p(t) + [c_0+2d_1+c_1-d_2] \sum_x^{\infty} t\,p(t) + [c_1+c_2+3d_2] \sum_x^{\infty} t^2 p(t)$$
$$+ c_2 \sum_x^{\infty} t^3 p(t) + x\{[d_2-d_1-c_0] \sum_x^{\infty} p(t) + [-2d_2-c_1] \sum_x^{\infty} t\,p(t) - c_2 \sum_x^{\infty} t^2 p(t)\} = 0.$$
$$(4.37)$$

Dividing (4.37) by R(x), we obtain

$$c_2\, \mathrm{E}(X^3 \mid X \geq x) = [(c_0+d_1-d_2)\, x - (c_0+d_0)]$$
$$+ [(c_1+2d_2)\, x + d_2 - c_1 - 2d_1 - c_0]\mathrm{E}(X \mid X \geq x)$$
$$+ [c_2\, x - (c_1+c_2+3d_2)]\ \mathrm{E}(X^2 \mid X \geq x).$$

which has the same form as (4.28) with

$$A(x) = c_2\, x - (c_1+c_2+3d_2)$$
$$B(x) = (c_1+2d_2)\, x + d_2 - c_1 - 2d_1 - c_0$$

and

$$C(x) = (c_0+d_1-d_2)\, x - (c_0+d_0).$$

This completes the proof.

Corollary 4.6

When $c_2=0$, Theorem 4.2 reduces to the result of Glanzel (1991) for Ord family of distributions.

Corollary 4.7:

The relationship

$$\begin{aligned}\lambda\, E(X^3 \mid X\geq x) = {} & \{\lambda x - [3+\lambda(e_1+e_2+1)]\}\, E(X^2 \mid X\geq x) \\ & +\{[\lambda(e_1+e_2)+2]x+1-\lambda[(e_1+e_2)+e_1 e_2] \\ & -2(g_1+g_2)\}\, E(X \mid X\geq x) +[\lambda e_1 e_2 +(g_1+g_2)(x-1)-x]\end{aligned}$$

holds if and only if X belongs to the Kemp family (4.5).

Corollary 4.8

The relationship

$$E(X^3 \mid X\geq x)+\nu\theta(x-1)+E(X^2 \mid X\geq x)(b-\theta-x)+E(X \mid X\geq x)[(\theta-b)x-\theta(1+\nu)]=0$$

holds if and only if the distribution of X is confluent hyper geometric with p.m.f (4.6).

Corollary 4.9

The distribution of X is Borel-Tanner with p.m.f. (4.8) if and only if

$$\begin{aligned}E(X^3 \mid X\geq x)-{} & [n(1-q)+x+2]E(X^2 \mid X\geq x) \\ & +\{[n(1-q)-1]x-n(2+q)+1\}\, E(X \mid X\geq x)+[n(1+q)x-n].\end{aligned}$$

CHAPTER V

CLASSIFICATION OF MODELS IN DISCRETE TIME

5.1 Introduction

The majority of literature on the various criteria for ageing center around continuous life time models. Recently there is some spurt of activity towards reliability analysis in the discrete time domain. As mentioned earlier there are several instances in which the failure time distribution can be modeled by a discrete random variable. The pioneer work in this area is due to Xekalaki (1983), who pointed out that, limitations of measuring devices and the fact that discrete models provide good approximations to their continuous counter parts, necessitate assessment of reliability in discrete time. Accordingly, elaboration of various concepts analogous to those in the continuous case become necessary to distinguish classes of life distributions based on the notions of ageing.

As in the continuous set up, the ageing behaviour of the system or component usually studied by failure rate function or by MRL function. Various authors have studied classes of life distributions based on different concept of ageing. Langberg et.al(1980) discussed properties of discrete models with decreasing failure rates. Ebrahami (1986) provided two parametric families of discrete distributions which are suitable for fitting decreasing and increasing mean residual life models to life test data in discrete time. Guess and Park (1988) developed a general approach to modeling discrete bathtub shaped MRL function. Salvia and Bollinger (1982) have established simple bounds for residual life when the device has a monotonic hazard rate sequence.

As is well known, the monotonicity of failure rate of a life distribution plays a very important role in modeling failure time data. Therefore, the identification of the increasing failure rate (IFR) or decreasing failure rate (DFR) distributions and their properties have been extensively discussed in the literature for the continuous case. However, for the discrete case, the determination of the IFR and DFR models is not straightforward because of the complexity of the failure rate. In this direction, Gupta et.al (1997) developed techniques for the determination of IFR and DFR models for a wide class of discrete distributions.

In the following section, we provide a new method to identify an IFR/DFR model in the generalized Ord family.

Definition 5.1

The distribution of X is said to have discrete IFR (DFR) property if $h(x) \leq (\geq)\ h(x+1)$ for every $x=0,1,2,\ldots$

For the classification of discrete lifetime models through failure rate function, we refer to Abouammoh (1990) and Roy and Gupta (1992). However for many distributions $h(x)$ is not in a simple form. Now we suggest a method, using $\beta(x)$ to identify an IFR(DFR) model in the generalized Ord family, where $\beta(x)$ is defined as

$$\beta(x) = \frac{p(x) - p(x+1)}{p(x)}.$$

Theorem 5.1:

If the inequality

$$\beta(x) \leq (\geq) \beta(x+1)$$

holds for every non-negative integer x, then X has IFR (DFR) property.

Proof

From the definition of $h(x)$, we have

$$\frac{1}{h(x)} = 1 + \frac{p(x+1)}{p(x)} + \frac{p(x+2)}{p(x)} + \ldots \tag{5.1}$$

If the inequality $\beta(x) \leq \beta(x+1)$ holds, (5.1) becomes $h(x) \leq h(x+1)$. Thus X has IFR property. The proof when $\beta(x) \geq \beta(x+1)$ is similar.

Now we use $\beta(x)$ for the classification of distributions belonging to the generalized Ord family. For the generalized Ord family (4.3),

$$\beta(x+1) - \beta(x) = \frac{kx^2 + mx + n}{(d_0 + d_1 x + d_2 x^2)(d_0 + d_1(x+1) + d_2(x+1)^2)} \tag{5.2}$$

where

$$k = c_1 d_2 - c_2 d_1$$

$$m = c_1 d_2 - c_2 d_1 + 2(c_0 d_2 - c_2 d_0)$$

and

$$n = c_0(d_2 + d_1) - d_0(c_2 + c_1).$$

When $d_0+d_1x+d_2x^2 > 0$ for all $x = 0, 1, 2, \ldots$

$$\beta(x+1) - \beta(x) >(<) 0$$

according as $kx^2 + mx + n > (<) 0$. Thus, the roots of the equation $kx^2 + mx + n = 0$ determines the sign of $\beta(x+1) - \beta(x)$.
If $\Delta = m^2 - 4kn$ is the discriminant of the expression $kx^2 + mx + n = 0$, then from the elementary algebra we have the following theorems.

Theorem 5.2

Suppose that the p.m.f of X belongs to the family of (4.3) with $d_0+d_1x+d_2x^2>0$, then

(A) $\beta(x+1) - \beta(x) > 0$ if (i) $k > 0$ and either

(a) $\Delta = 0$ and $x \neq \dfrac{-m}{2k}$ or

(b) $\Delta < 0$ or

(c) $\Delta > 0$ and $x \notin (\alpha, \beta)$ or

(ii) $k < 0$, $\Delta > 0$ and $x \in (\alpha, \beta)$.

(B) $\beta(x+1) - \beta(x) < 0$ if (i) $k < 0$ and either

(a) $\Delta = 0$ and $x \neq \dfrac{-m}{2k}$ or

(b) $\Delta < 0$ or

(c) $\Delta > 0$ and $x \notin (\alpha, \beta)$ or

(ii) $k > 0$, $\Delta > 0$ and $x \in (\alpha, \beta)$.

Theorem 5.3

A distribution belonging to the generalized Ord family (4.3) has IFR property in a region if condition (A) of Theorem 5.2 holds in that region and has DFR property in a region if condition (B) of Theorem 5.2 holds in that region.

Remark 5.1

When k=0, $\beta(x+1)$-$\beta(x)$>(<)0 according as $mx+n$>(<)0.

Corollary 5.1

When c_2=0 in (4.3), Theorem 5.2 reduces to the result of Ord family of distributions given by Sankaran and Sindu (2001).

Corollary 5.2

For the verification of the theorem, we consider the Borel-Tanner distribution with p.m.f. (4.8), then we have

$$\beta(x) = \frac{x^2 + x(nq - n + 1) - n}{nqx}.$$

Since

$$\beta(x+1) - \beta(x) = \frac{x^2 + x + n}{nqx + nqx^2} > 0$$

the distribution (4.8) is IFR.

Remark 5.2

Table 5.1 gives the region where the distribution possesses the IFR (DFR) property based on $\beta(x)$ for some popular models belonging to the family (4.3).

Table 5.1

The region where the distribution possesses the IFR (DFR) property

Sl. No.	Distributions with p.m.f	$\beta(x)$	Region
1	Binomial $\binom{n}{x} p^x (1-p)^{n-x}$, x=0,1,...	$\frac{x-[np-(1-p)]}{(1-p)(x+1)}$	IFR
2	Poisson $\frac{e^{-\lambda}\lambda^x}{x!}$, x=0,1,...	$\frac{x+1-\lambda}{x+1}$	IFR
3	Negative Binomial $\binom{x+y-1}{x-1} p^x (1-p)^y$, x=0,1,...	$\frac{x-\frac{py}{1-p}}{x/(1-p)}$	IFR
4	Hypergeometric $\frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}}$, Max(0,D-N+n)$\le x \le$Min(D,n)	$\frac{x+n(D+1)-[N+1+D]}{(x+1)(N-D-n+x+1)}$	IFR
5	Waring $\frac{(a-b)(b)_x}{(a)_x}$, x=0,1,... $(b)_x = \frac{\Gamma(b+x)}{\Gamma(b)}$,	$\frac{x(1+a-b)}{x(1+a+x)}$	DFR
6	Beta Pascal $\frac{\frac{A}{A+K}\binom{K+x-1}{x}\binom{A+B-1}{A}}{\binom{K+A+B+x-1}{K+A}}$, x=1,2,...	$\frac{x(A+1)-[KB-(K+A+B)]}{x^2+(K+A+B+1)x+K+A+B}$	DFR
7	Borel-Tanner $\frac{e^{-nq}(nq)^{n-x}x}{n(n-x)!}$, x = 1,2....	$\frac{x^2+x(nq-n+1)-n}{nqx}$	IFR

5.2 Characterization using Mean Residual Life

Mean Residual Life (MRL) function is widely used in the analysis of lifetime data. Muth (1977) pointed out that the MRL function is superior to the failure rate function in many practical situations. It is well known that an increasing (decreasing) failure rate class of distributions is a subclass of a decreasing (increasing) MRL class of distributions. In the following, we establish a characterization result for IFR(DFR) class of distributions in the generalized Ord family, using MRL function.

Definition 5.2

A non-negative random variable X has increasing mean residual life (IMRL) if

$$r(x) = \frac{1}{R(x+1)} \sum_{x+1}^{\infty} R(t)$$

is non-decreasing in x=0,1,2... and decreasing mean residual life (DMRL) if

$$r(x) = \frac{1}{R(x+1)} \sum_{x+1}^{\infty} R(t)$$

is non-increasing in x=0,1,2...

Theorem 5.4

Let the distribution of X belongs to the generalized Ord family (4.3). If $d_0+d_1x+d_2x^2>0$ and $c_2[V(x)-V(x-1)]>0$, then X is said to have discrete IFR(DFR) property if and only if

$$d_1 + 2d_2x - d_2 + r(x)[\, c_1 + 2d_2 + c_2(r(x)+r(x-1)) + c_2(2x-1)\,] \leq (\geq) 0. \quad (5.3)$$

Proof

When the distribution of X belongs to the generalized Ord family (4.3), we have the identity (4.25).

From(1.14),

$$m_2(x) = V(x) + 2xm_1(x) - x^2 + r^2(x)$$

where $V(x)$ is the variance residual life,(4.25) can be written as

$$c_2[V(x)+x^2+2xr(x)+r^2(x)] + c_1+2d_2(x+r(x)) +(c_0+d_1-d_2) + (d_0+d_1x+d_2x^2)\, h(x+1) =0 \cdot \quad (5.4)$$

Changing x to x-1 in (5.4) and subtracting the resulting expression from (5.4), we get

$$c_2\, [(V(x)- V(x-1))+2x(r(x)-r(x-1)) + (r^2(x)- r^2(x-1))+2x-1+2r(x-1)] + (c_1+2d_2)(r(x)- r(x-1)+1) +(d_1+2d_2x-d_2)h(x) + (d_0+d_1x+d_2x^2)\, [h(x+1)- h(x)] =0. \quad (5.5)$$

Substituting the relationship between the failure rate and the MRL in the discrete case given by

$$1- h(x) = \frac{r(x-1)-1}{r(x)} \quad (5.6)$$

(5.5) provides,

$$h(x)[\, d_1+2d_2x-d_2+(c_1+2d_2)r(x) +c_2\, r(x)[\, r(x)+ r(x-1)+2x-1] +c_2[V(x)- V(x-1)] +(d_0+d_1x+d_2x^2)\, [h(x+1)- h(x)] =0. \quad (5.7)$$

If $d_0+d_1x+d_2x^2>0$ and $c_2[V(x)-V(x-1)]>0$, then from (5.7),

$$h(x+1)-h(x)\geq(\leq)0$$

if and only if,

$$d_1+2d_2x-d_2+r(x)[\, c_1+2d_2+c_2(r(x)+r(x-1)) +c_2(2x-1)\,]\leq(\geq)0$$

which completes the proof.

Corollary 5.3

When $c_2=0$, Theorem 5.4 reduces to the simple form for Ord family of distributions as shown below.

Theorem 5.5

Let the distribution of X belong to the Ord family (4.1). Then X is said to have discrete IFR(DFR) property if and only if

$$r(x)\geq(\leq)\ p_1-p_2+2p_2x \tag{5.8}$$

for all $x=0,1,2\ldots$ where $p_i=\dfrac{k_i}{1-2k_2}$, $i=1, 2$.

Proof

When the distribution of X belong to the Ord family (4.1), we have (Nair and Sankaran (1991))

$$r(x)+x = (p_0+p_1x+p_2x^2)h(x+1)+\mu \tag{5.9}$$

where $\mu = E(x)$.

Changing x to $(x-1)$ in (5.9) and subtracting the resulting expression from (5.9), we get

$$r(x)-r(x-1)+1=(p_0+p_1x+p_2x^2)[h(x+1)-h(x)]+h(x)\ (p_1-p_2+p_2x). \tag{5.10}$$

Substituting the relationship (5.6), (5.10) becomes,

$$h(x)[r(x)-(p_1-p_2+2p_2x)]=(p_0+p_1x+p_2x^2)[h(x+1)-h(x)]. \tag{5.11}$$

It is easy to verify that $(p_0+p_1x+p_2x^2)\geq 0$. Thus from (5.11),

$$[h(x+1)-h(x)]\geq(\leq)\,0$$

if and only if

$$r(x)\ \geq(\leq)\,(p_1-p_2+2p_2x).$$

This completes the proof.

Remark 5.3

The p_i's in (5.8) are directly related to the moments of the distributions. To apply the result in a practical situation one need to take the sample moments and sample MRL function as estimators.

For the verification of Theorem 5.5, consider the Waring distribution with

$$p(x)= \frac{(a-b)(b)_x}{(a)_x}, \; x=0,1,2\ldots \quad a>b>0$$

where $(b)x= b(b+1)\ldots.(b+x-1)$.

By direct computation we get,

$$r(x)=\frac{a+x}{a-b-1}, \; p_1=\frac{a+1}{a-b-1} \text{ and } p_2=\frac{1}{a-b-1}$$

Since $r(x)< p_1-p_2+2p_2x$ for any $x=0,1,2,\ldots$. Waring distribution has DFR property.

5.3. Length Biased Models

In this section we discuss the form-invariant length biased models from generalized Ord family.

Analogous to the continuous case, the length biased distribution of a discrete random variable X with the set of non-negative integers as the support is defined as (Gupta 1979),

$$g(x) = \frac{xp(x)}{\mu} \quad x= 1,2\ldots \tag{5.12}$$

where μ= E(X)<∞. Clearly the above random variable Y will have no zero in its support. Applying a displacement of Y to the left, by taking $Z=Y-1$, Z would be realized by length biased sampling on X with the above displacement and the support becomes the set of non negative integers (See Patil and Ord 1976). The resulting probability mass function of Z is

$$p(x)=g(x+1) \quad \text{for } x=0,1,2\ldots$$

For the application of (5.12) to reliability we can refer to Patil and Rao (1977), and Gupta and Kirmani (1990).

5.3.1 Form Invariance

The distribution of X with p.m.f. $p(x)$ is said to be form-invariant under length biased sampling if observed variable Z has the same distribution as X, with a change in parameter. The major relationships between the survival function, failure rate and MRL of the original distribution and its corresponding length biased version is given as

$$G(x)= \frac{m(x)R(x+1)}{\mu} \tag{5.13}$$

$$k(x+1)= \frac{(x+1)h(x+1)}{m(x)} \tag{5.14}$$

$$e(x-1)= \frac{\sum_{x}^{\infty} R(t+1)m(t)}{R(x+1)m(x)} \tag{5.15}$$

where $G(x)$, $k(x)$ and $e(x)$ are respectively the survival function, failure rate and MRL of Y. The above identities connecting reliability characteristics of X and Y can be employed in the characterization of the distribution of X. Sankaran and Nair (1993) derived conditions under which models belonging to the Ord family retain the same form for their length biased distributions.

In reliability the ageing patterns of system can be studied by comparing the structural properties of their life lengths with those from the corresponding length biased distributions. In the following we derive the conditions under which the members of the generalized Ord family (4.2) are form invariant with respect to the formation of their length biased distributions.

Theorem 5.6

Among the members of family (4.3), X and Y have the same type of distribution if and only if $a_0 + b_0 = 0$, and the p.m.f of Y satisfies

$$\frac{g(x+1)-g(x)}{g(x)} = \frac{p_0+p_1x+p_2x^2}{q_0+q_1x+q_2x^2} \tag{5.16}$$

where p_0, p_1, p_2, q_0, q_1 and q_2 are real constants.

Proof

Suppose that (4.3) holds and X and Y have the same distributional form. Then from (5.12), we have

$$\frac{g(x+1)-g(x)}{g(x)} = \frac{x+1}{x}\frac{f(x+1)}{f(x)} - 1 \tag{5.17}$$

which gives,

$$\frac{g(x+1)-g(x)}{g(x)} = \frac{x+1}{x}\left[1+\frac{c_0+c_1x+c_2x^2}{d_0+d_1x+d_2x^2}\right] - 1. \tag{5.18}$$

Since Y also must belong to the family (4.3), (5.18) must be of the form,

$$\frac{p_0+p_1x+p_2x^2}{q_0+q_1x+q_2x^2} = \frac{c_2x^3+(c_1+c_2+d_2)x^2+(c_0+c_1+d_1)x+c_0+d_0}{(d_0x+d_1x^2+d_2x^3)}$$

or

$$(d_0x+d_1x^2+d_2x^3)(p_0+p_1x+p_2x^2)= [c_2x^3 + (c_1+ c_2+ d_2)x^2+(c_0+ c_2+ d_1)x + (c_0+ d_0)](q_0+q_1x+q_2x^2) \quad (5.19)$$

Equating the coefficients of like powers of x in (5.18), we have six equations,

$$q_0 (c_0 + d_0) = 0 \quad (5.20)$$

$$p_0d_0 = q_0 (c_0 +c_1+d_1)+ q_1 (c_0 + d_0) \quad (5.21)$$

$$p_0d_1+ p_1d_0= q_0 (c_1 +c_2+d_2)+ q_1 (c_0 +c_1+d_1)+ q_2(c_0 + d_0) \quad (5.22)$$

$$p_2d_0+ p_1d_1+ p_0d_2= q_1 (c_1 +c_2+d_2)+ q_2 (c_0 +c_1+d_1)+q_0c_2 \quad (5.23)$$

$$p_2d_1+ p_1d_2= q_2 (c_1 +c_2+d_2)+q_1c_2 \quad (5.24)$$

$$p_2d_2 = q_2c_2. \quad (5.25)$$

From (5.20), we have the following cases,

(1) $q_0 \neq 0$ and $c_0 + d_0=0$, which leads to

$$\frac{g(x+1)-g(x)}{g(x)} = \frac{c_2x^2+(c_1+c_2+d_2)x+c_0+c_1+d_1}{-c_0+d_1x+d_2x^2}. \quad (5.26)$$

It is easy to see that, $p(x)$ and $g(x)$ have same distributional form though with possibly different parameters.

(ii) When $q_0=0$ and $c_0 + d_0\neq0$,

$$\frac{g(x+1)-g(x)}{g(x)} = \frac{p_2}{q_2} + \frac{p_0q_2+(p_1q_2-p_2q_1)x}{q_2+(q_1x+q_2x^2)}$$

which has not the same form as $p(x)$.

(iii) When $q_0=0$ and $c_0 + d_0=0$, we obtain an equation

$$p_0\, d_0=0$$

which leads to three different cases,

(a) $p_0=0,\ d_0\neq0$

(b) $p_0\neq0,\ d_0=0$

(c) $p_0=0,\ d_0=0$.

The discussions based on above three cases lead to the situations parallel to those we have already mentioned with $c_0+d_0=0$.

Conversely when $c_0 + d_0 = 0$, (4.3) and (5.18) provide that X and Y have the same type of distributions. This completes the proof.

Corollary 5.4

When c_2=0, Theorem 5.6 reduces to the result of Sankaran and Nair (1993) for the Ord family of distributions.

To verify Theorem 5.6, consider the confluent hyper geometric distribution with p.m.f (4.6), then the LBD can be obtained as

$$g(x) = K\frac{\Gamma v+x}{\Gamma 1+b+x\ (x-1)!}\theta^{x-1}\ ,\ x = 1, 2, \ldots\ldots$$

where K is the normalizing constant. This has the same form as parent distribution with different parameters.

Next we prove the condition under which the length biased distribution of X belongs to the generalized Ord family when the original belongs to Ord family.

Suppose that the distribution of X belongs to the Ord family (4.1) and that of Y belong to (4.3). Then we have,

$$\frac{g(x+1)-g(x)}{g(x)} = \frac{x+1}{x}\frac{p(x+1)}{p(x)} - 1$$

or

$$\frac{c_0 + c_1 x + c_2 x^2}{d_0 + d_1 x + d_2 x^2} = \frac{k_0 - u + (k_1 - d_1 - 1)x + (k_2 - 1)x^2}{k_0 x + k_1 x^2 + k_2 x^3} \qquad (5.27)$$

which provides,

$$(k_0x+k_1x^2+k_2x^3)(c_0+c_1x+c_2x^2)= (d_0+d_1x+d_2x^2)$$
$$[k_0 - u + (k_1 - d_1 - 1)x + (k_2 - 1)x^2] \cdot \qquad (5.28)$$

Equating the coefficients of like powers of x in (5.28), we obtain the following equations

$$d_0[k_0-u]=0 \tag{5.29}$$

$$c_0k_0 = d_0(k_1-u-1)+d_1(k_0-u) \tag{5.30}$$

$$c_1k_0+ c_0k_1= d_0(k_2-1)+ d_1(k_1-u-1)+ d_2(k_0-d_K) \tag{5.31}$$

$$c_2k_0+ c_0k_2+ c_1k_1= d_1(k_2-1)+ d_2(k_1-u-1) \tag{5.32}$$

$$c_1k_2+ c_2k_1= d_2(k_2-1) \tag{5.33}$$

$$c_2k_2=0. \tag{5.34}$$

Now from (5.29) and (5.34), we have the following cases.

(i) $d_0=0$, $k_2=0$, $k_0-u\neq0$, $c_2\neq0$, then

$$\frac{g(x+1)-g(x)}{g(x)}=\frac{k_0-u+(k_1-d_1-1)x+x_2}{k_0x+k_1x^2}$$

and

$$\frac{p(x+1)-p(x)}{p(x)}=\frac{-(x+u)}{k_0+k_1x}$$

clearly $p(x)$ and $g(x)$ have different forms.

(ii) $d_0=0$, $c_2=0$, $k_2\neq0$, $k_0-u\neq0$, then

$$\frac{p(x+1)-p(x)}{p(x)}=\frac{-(x+u)}{k_0+k_1x+k_2x^2}$$

and

$$\frac{g(x+1)-g(x)}{g(x)}=\frac{k_0-u+(k_1-d_1-1)x+k_2x^2}{(k_0+k_1x+k_2x^2)x}$$

here also, $p(x)$ and $g(x)$ have different forms.

(iii) $d_0\neq0$, $c_2=0$, $k_2\neq0$, $k_0-u=0$, then

$$\frac{g(x+1)-g(x)}{g(x)}=\frac{(k_1-d_1-1)+(k_2-1)x}{k_0+k_1x+k_2x^2}$$

and

$$\frac{p(x+1)-p(x)}{p(x)}=\frac{-(x+u)}{k_0+k_1x+k_2x^2}.$$

Thus $p(x)$ and $g(x)$ have same distributional forms but with different parameters.

(iv) $d_0 \neq 0$, $k_2=0$, $c_2 \neq 0$, $k_0-u=0$, then

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{-(x+u)}{k_1 x+u}$$

and

$$\frac{g(x+1)-g(x)}{g(x)} = \frac{k_1-d_1-1-x}{k_1 x+u}.$$

Thus $p(x)$ and $g(x)$ have same distributional forms.

(v) $k_0-u = 0$, $d_0 = 0$, $k_2 = 0$ and $c_2 = 0$, we obtain $d_2=0$, then

$$\frac{p(x+1)-p(x)}{p(x)} = \frac{-(x+u)}{k_1 x+u}$$

and

$$\frac{g(x+1)-g(x)}{g(x)} = \frac{k_1-d_1-1-x}{k_1 x+u}$$

thus $p(x)$ and $g(x)$ have same distributional forms.

When $d_0 = -c_0$ in (4.3), we obtain a sub class of the generalized Ord family. This sub-class, to be denoted by D, contains many distributions of interest in reliability analysis.

Now we prove a characterization of the class D based on conditional moments.

Theorem 5.7

The distribution of X belongs to the class D if and only if

$$c_2\, m_3(x) + [(c_1+3d_2) - c_2 x]\, m_2(x) = [(c_1+2d_2)\, x+3d_2 - c_0 - 2d_1]m_1(x) + (d_1 + c_0 - d_2)(x+1) \quad (5.35)$$

where $m_i(x) = E[X^i|X>x]$, $i=1,2,3$.

Proof

Since $d_0 = -c_0$,(4.25) leads to

$$(c_0 - d_1 x - d_2 x^2)\, h(x+1) = c_2\, m_2(x) + (c_1+2d_2)m_1(x) + c_0 + d_1 - d_2. \tag{5.36}$$

On similar lines, for the random variable Y, (5.26) gives,

$$(c_0 - d_1 x - d_2 x^2)k(x+1) = c_2\, V_2(x) + (c_1+c_2+3d_2)V_1(x) + c_0 + c_1 + 2d_1 - d_2 \tag{5.37}$$

where $V_i(x) = E[Y^i|Y>x]$, $i = 1, 2$. and $k(x)$ is the failure rate of Y. From (5.14), we have

$$k(x+1) = \frac{(x+1)h(x+1)}{m_1(x)} \tag{5.38}$$

and

$$V_i(x) = \frac{m_{i+1}(x)}{m_1(x)},\ i=1, 2. \tag{5.39}$$

Dividing (5.37) by (5.36) and substituting the relationships (5.38) and (5.39) in the resulting equations, we obtain (5.35).

Conversely suppose that (5.35) holds. Then we have

$$c_2 \sum_{x+1}^{\infty} t^3 p(t) + [c_1+3d_2-c_2x] \sum_{x+1}^{\infty} t^2 p(t) = [(c_1+2c_2)x+3d_2-c_0-2d_1] \sum_{x+1}^{\infty} t p(t)$$
$$+(c_0+d_1-d_2)(x+1)R(x+1)\cdot \tag{5.40}$$

Changing x to $(x-1)$ in (5.40) and subtracting (5.40) from the resulting equation, we get

$$d_2 x^2\, p(x) + c_2 \sum_{x}^{\infty} t^2 p(t) = (2d_2-d_1)\, xp(x) - (c_1+2d_2) \sum_{x+1}^{\infty} t p(t)$$
$$+(c_0+d_1-d_2)R(x+1)\cdot \tag{5.41}$$

Now changing the variable x to $(x+1)$ in (5.41) and subtracting (5.41) from the resulting expression, we obtain

$$p(x+1)\ [d_0 + d_1 x + d_2 x^2] = p(x)\ [(c_1 + d_1)x + (c_2 + d_2)x^2]$$

which is of the form (4.3) with $d_0 = -c_0$.

This completes the proof.

Corollary 5.5

When $c_2=0$, Theorem 5.7 reduces to the result of Sankaran and Nair (1993).

Corollary 5.6

The distribution of X is confluent hyper geometric with p.m.f (4.6) holds if and only if

$$-m_3(x) + (x+\theta-b+1)\ m_2(x) = [(\theta-b)x-(\nu\ \theta+3b+2)]m_1(x) + \nu\theta\ (x+1).$$

Corollary 5.7

The relationship

$$-m_3(x) + [x+(n-3)\alpha+n-1)m_2(x) = \{[n(1-\alpha)-2\alpha-1]x+4-5n\alpha-3\alpha-n\}m_1(x) + (3n\alpha+n)\ (x+1)$$

holds if and only if X has Haight distribution with p.m.f (4.9).

REFERENCES

1. Abouammoh, A.M. (1988): On the criteria of the mean remaining life, *Statistics and Probability Letters*, 6, 205-211.

2. Ahmed, A. N. (1991): Characterization of beta, binomial and Poisson distribution, *IEEE Transactions on Reliability*, 40, 290-295.

3. Barlow, R.E. (1984): Theory of reliability, A historical perspective, *IEEE Transactions on Reliability*, 33, 16-20.

4. Barlow, R.E. and Porschan, F. (1975): *Statistical Theory of Reliability and Life Testing*, Probability Models, Holt, Rinehart and Winston, New York.

5. Basu, A.P. and Ebrahimi, N (1986): HNBUE and HNWUE distributions-A Survey; *Reliability and Quality control*, edited by A.P Basu (North Holland)

6. Bhattacharya, S.K. (1966): Confluent hyper geometric distributions of discrete and continuous type with applications to accident proneness, *Calcutta Statistical Association Bulletin*, 60, 1060-1066.

7. Bowman,K.O., Shenton,L.R. and Kastenbaum, M.A. (1991): Discrete Pearson distributions, *Oak Ridge National Labroratry Technical Report*, TM-11899, Oak Ridge.

8. Bryson, M.C. and Siddiqui, M.M. (1969): Some criteria for ageing, *Journal of the American Statistical Association*, 64, 1472-1483.

9. Chhikara, R.S. and Folks, J.L. (1977): The inverse Gaussian distribution as a lifetime model, *Technometrics*, 19, 461-468.

10. Chong, K. (1977): On Characterizing exponential and geometric Distribution by Expectation, *Journal of the American Statistical Association*, 72,1 960-1961.

11. Coppola, A (1984):Reliability Engineering of Electronic Equipment; A historical perspective, *IEEE Transactions on Reliability*, 33, 29-35.

12. Cox, D.R. (1962): *Renewal Theory*, Methuen, London.

13. Cox, D.R(1972):"Regression models and life tables", *Journal of Royal Statistical Society*, 34,187-202.

14. Cundy,A (1966): Birds and atoms, *Mathematical Gazette*, 50,294-295.

15. Daniels, H.E. (1942): A new technique for the analysis of fiber length distribution in wool, *Journal of Textile Institute*, 33, 137-150.

16. Davies,O.L.(1934): On asymptotic formula for the hypergeometric series, *Biometrika*, 25, 295-322.

17. Dunning,K.A and Hanson,J.N (1977): Generalized Pearson distributions and non-linear programming, *Journal of Statistical Computation and Simulation*, 6, 115.

18. Ebrahimi, N(1986), Classes of discrete decreasing and increasing mean residual life distributions, *IEEE Transactions on Reliability*, 35, 403-405.

19. Elderton,W.P. and Johnson,N.L. (1969): *Systems of Frequency Curves*, Cambridge University Press, London.

20. Fussel,J.B (1984): Nuclear power system reliability; A historical perspective, *IEEE Transactions on Reliability*, 33, 41-47.

21. Galambos, J and Kotz, S. (1978): *Characterizations of probability distributions*, Springer-Verlag, New York.

22. Glanzel,W (1991): Characterization through some conditional moments of Pearson-type distributions and discrete analogues, *Sankhya*, B, 53, 17-24.

23. Glaser, R.E. (1980): Bathtub and related failure rate characterizations, *Journal of the American Statistical Association*, 75, 667-672.

24. Guerrieri, G (1965): Some characteristic properties of the exponential distribution, G. *Economists*, 24, 427-437.

G18598

25. Guess, F.M. and Park, D.H. (1988): Modeling discrete bathtub and upside-down bathtub mean residual life functions, *IEEE Transactions on Reliability*, 37, 545-549.

26. Gupta, P.L. (1985): Some characterizations of distributions by truncated moments, *Statistics and Probability Letters*, 16, 465-473.

27. Gupta, P.L. and Gupta, R.C (1983): On the moments of residual life in reliability and some characterization results, *Communications in Statistics, Theory and Methods*, 12(4), 449-461

28. Gupta,P.L., Gupta,R.C and Tripathi,R.C (1997):On the monotonic properties of discrete failure rates, *Journal of Statistical Planning and Inference*,65,255-268.

29. Gupta, R.C. (1979): On the characterization of survival distributions in reliability by properties of their renewal densities, *Communications in Statistics, Theory and Methods*, 8(7), 685-697.

30. Gupta, R.C and Kirmani, S.N.U.A. (1990): The role of weighted distributions on stochastic modeling, *Communications in Statistics, Theory and Methods*, 19, 3147-3162.

31. Gupta, R.C., Kirmani, S.N.U.A. and Launer,R.L. (1987): On life distributions having monotone residual variance, *Prob. Engg. Inf. Sci.*, 1, 299-307.

32. Gupta,R.C and Keating,J.P (1986): Relations for reliability measures under length biased sampling, *Scandinavian Journal of Statistics*,13,49-56.

33. Haight, F.A. (1961): A distribution analogous to the Borel-Tanner, *Biomterika*, 48, 167-73.

34. Hall, W.J. (1966): Some hypergeometric series distribution occurring in birth-and-death processes at equilibrium, *Annals of mathematical Statistics*, 27, 221.

35. Hitha, N (1990): Two characterization of Pareto distribution, *Proceedings of 77th session of the Indian Science Congress Association*, 13.

36. Hitha, N (1991): *Some Characterizations of Pareto and Related Distributions*, Unpublished Ph.D thesis, Cochin University of Science and Technology, Cochin.

37. Hitha, N and Nair, N.U. (1989): Characterizations of some discrete models by properties of residual life function, *Calcutta Statistical Association Bulletin*,39,219-223.

38. Hollander, W and Proschan, F (1975): Test for mean residual life, *Biometrika*, 62, 585-593.

39. Jain,K., Singh,H and Bagai,I (1989): Relations for reliability measures of weighted distributions, *Communications in Statistics-Theory and methods*, 18(12), 4393-4412

40. Jinhua, C and Wang, Y (1991): NBUC and NWUC classes of life distributions; *Journal of Applied Probability*, 28, 473-479.

41. Johnson,N.L, Kotz, S and Balakrishnan, N. (1994): *Continuous Univariate Distributions*, Vol II, John Wiley, New York.

42. Johnson, N.L., Kotz, S and Kemp, A.W. (1992): *Univariate Discrete Distributions*, John Wiley and Sons, New York.

43. Kagan, A. M., Linnik, Y.V., and Rao, C.R.(1973): *Characterization Problems in Mathematical Statistics*, USSR, Academy of Sciences.

44. Kalbfleisch, J.D. and Prentice, R.L. (1980): *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, Inc. New York.

45. Pearson, K (1895): Contributions to the mathematical theory of evolution I, skew distribution in homogeneous material, *Philosophical Transactions of the Royal Society of London*, A, 186, 343-414.

46. Katz, L (1945): *Characteristics of frequency functions defined by first order difference equations*, Dissertation, Ann Arbor, MI; University of Michigan

47.Kemp, A.W. (1968) A wide class of discrete distributions and the associated differential equations, *Sankhya, A*, 30, 401-410.

48.Klefsjo, B(1982): The HNBUE and HNWUE classes of life distributions, *Naval Research Logistics Quarterly*, 29, 331-344.

49.Kosambi, D.D. (1949): Characteristic properties of series distributions, *Proceedings of the National Institute of Science*, India, 15, 109-113.

50.Kotz, S and Shanbhag, D.N. (1980): Some new approaches to probability distributions, *Advances in Applied Probability*, 12, 903-921.

51.Kupka,J and Loo,S (1989): The hazard and vitality measures of ageing, *Journal of Applied Probability*, 26, 532-542.

52.Langberg, N.A., Leon, R.V., Lynch,J. and Proschan, F. (1980): Extreme points of the class of discrete decreasing failure rate life distributions, *Mathematics of Operations Research*, 5, 35-42.

53.Launer, R.L. (1984): Inequalities for NBUE and NWUE life distributions, *Operations Research*, 32, 660-667.

54.Lawless,J.F (1982):*Statistical Models and Methods for Lifetime Data*, John Wiley and Sons, New York.

55.Mukherjee,S.P. and Roy,D. (1989): Properties of classes of probability distributions based on the concept of reciprocal co-ordinate sub tangent, *Calcutta Statistical Association Bulletin*, 38, 169-180.

56.Muth,E.J. (1977): Reliability models with positive memory derived from the mean residual life function, In T.P.Tsokos and T.N.Shimi (eds) *Theory and Applications of Reliability*, Academic Press.

57.Muth,E.J. (1980): Memory as a property of probability distributions, *IEEE Transactions on Reliability*,29, 160-165.

58.Nair, N.U. (1983): A measure of memory for some discrete distributions, *Journal of Indian Statistical Association*, 21, 141-147.

59.Nair, N.U. and Hitha, N (1989): Characterizations of discrete models by distributions based on their partial sums, *Statistics and Probability Letters*, 8, 335-337.

60.Nair,N.U., Priya,P. and Geetha, K.G (2000) : On Partial moments of discrete distributions and their applications, *Far East Journal of Theoretical Statistics*,4(1),153-164.

61.Nair,N.U. and Sankaran,P.G. (1991): Characterization of the Pearson family of distributions, *IEEE Transactions on Reliability*, 40, 75-77.

62.Navarro,J., Franco,M. and Ruiz,J.M. (1998): Characterization through moments of the residual life and conditional spacing, *Sankhya* A, 60, 36-48.

63.Nelson, W (1982): *Applied Life Data Analysis*, John Wiley and Sons, New York.

64.Noack, A (1950): A class of random variables with discrete distributions, *Annals of Mathematical Statistics*, 21, 127-132.

65.Ord,J.K. (1972): *Families of Frequency Distributions*, Charles Griffin and Co. Ltd., London.

66.Osaki, S and Li, X. (1988): Characterizations of gamma and negative binomial distributions, *IEEE Transactions on Reliability*, 37, 379-382.

67.Padgett, W.J. and Spurrier, J.D. (1985): On discrete failure models, *IEEE Transactions on Reliability*, 31, 458-459.

68.Patil, G.P. and Ord, J.K. (1976): On size-biased sampling and elated form-invariant weighted distributions, *Sankhya B*, 38, 48-61.

69.Patil, G. P. and Rao, C.R. (1977), The Weighted distributions: A survey of their applications, in *Applications of Statistics*, P.R.Krishniah (Ed.), North Holland Publishing Co., 383-405.

70. Priya,P.(2001): *Some Concepts and Models Useful in the Analysis of Discrete Life Time Data*, Unpublished Ph.D thesis, Cochin University of Science and Technology, Cochin.

71. Priya, P. Menon, Sankaran,P.G and Nair, N.U (2000): Factorial partial moments and their properties, *Journal of the Indian Statistical Association*, 38, 45-55.

72. Rao, C.R. (1965): On discrete distributions arising out of methods of ascertainments. In *Classical and Contagious Discrete Distributions*, G.P.Patil (ed.), Pergamon Press and Statistical Publishing Society, Calcutta, 320-332. Also reprinted in *Sankhya* A, 27, 311-324.

73. Rolski,T (1975): Mean residual life, *Bulletin of the International Statistical Institute*, 46,266-270.

74. Roy, D and Gupta, R.P. (1992): Classification of discrete lives, *Micro electron reliability*, 31, 1459-1473.

75. Ruiz, J.M. and Navarro, J (1994): Characterizations of distributions by relationship between failure rate and mean residual life, *IEEE Transactions on Reliability*, 43, 640-644.

76. Salvia, A.A. (1996): Some results on discrete mean residual life, *IEEE Transactions on Reliability*, 45, 359-361.

77. Salvia, A.A and Bollinger,R.C. (1982): On discrete hazard functions, *IEEE Transactions on Reliability*, 31, 458-459.

78. Sankaran, P.G. and Nair, N.U. (1993): On form-invariant length biased models from Pearson family, *Journal of the Indian Statistical Association*, 31, 85-89.

79. Sankaran, P.G. and Nair, N.U. (2000):On Some reliability aspects of Pearson family of distributions, *Statistical Papers*, 41, 109-117.

80. Sankaran, P.G. and Sindu, T.K. (2001): On IFR(DFR) classes of distributions, *IAPQR Transactions* , 26, 53-61.

81. Sankaran, P.G. and Sindu, T.K. (2003): A generalized Pearson system useful in reliability analysis, *Statistical Papers* (to appear).

G78598

82. Seshadri, V (1999): The inverse Gaussian distribution, *Statistical Theory and Applications*, Springer Verlage, New York.

83. Shaked, M., Shanthikumar,G and Torres,J, B.V(1994): Discrete probability orderings in reliability theory, *Statistica Sinica*, 4(2), 567-579.

84. Shaked,M, Shanthikumar,G and Torres,J.B.V (1995): Discrete hazard functions, *Comp. Operation Research*, 22(4), 391-402.

85. Shanbhag, D.N. (1970): Characterizations for exponential and geometric distributions, *Journal of the Indian Statistical Association*, 65, 1256-1259.

86. Shooman,M.L (1984): Software Reliability; A historical perspective, *IEEE Transactions on Reliability*, 33, 48-55.

87. Singh, H and Despande, J.V. (1985):On some new ageing properties, *Scandinavian Journal of Statistics*, 12, 213-220.

88. Stuart,A and Ord,J.K (1994): *Kendall's Advanced Theory of Statistics* ,Volume 1,John Wiley and Sons, New York.

89. Sunoj,S.M(2002): Some characterizations of continuous distributions using partial moments (Communicated).

90. Swartz, G.B. (1973): The mean residual lifetime function, *IEEE Transactions on Reliability*, 22, 108-109.

91. Tanner, J.C. (1953): A problem of interference between two queues, *Biometrika*, 40, 58-69.

92. Watson, G.S. and Wells, W.T. (1961): On the possibility of improving the mean useful life of items by eliminating these with short lives, *Technometrics*, 3, 281-298.

93. Xekalaki, E. (1983): Hazard functions and life distributions in discrete time, *Communications in Statistics, Theory and Methods*, 12, 2503-2509.