CHARACTERIZATION OF PROBABILITY DISTRIBUTIONS

SOME CHARACTERIZATIONS OF PARETO AND RELATED POPULATIONS

THESIS SUBMITTED TO
THE COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
UNDER THE FACULTY OF SCIENCE

By
N. HITHA

DEPARTMENT OF MATHEMATICS AND STATISTICS
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
KOCHI - 682 022
INDIA

APRIL 1991

CERTIFICATE

Certified that the thesis entitled " Some characterizations of Pareto and related populations" is a bonafide record of work done by Smt. N. Hitha under my guidance in the Department of Mathematics and Statistics, Cochin University of Science and Technology and that no part of it has been included anywhere previously for the award of any degree or title.

Dr. N.Unnikrishnan Nair
Professor of Statistics
Dept. of Mathematics and Statistics
Cochin University of Science and Technology
Kochi 682 022

Kochi 682 022
April 17 , '91

CONTENTS

Cont'd.

****

Chapter I

INTRODUCTION

In most cases of analysing statistical data, a fundamental problem that emerges at the outset is the identification of an appropriate model that can describe the real situation which generated the observations. Once the correct model is recognised the original problem permits analysis with lesser effort, as the properties of the model comes handy to the analyst in drawing inferences and decisions. Owing to the availability of a large number of probability distributions at disposal, very often the selection of a particular one in a specific situation turns out to be difficult, unless one has a reasonable basis or criteria that justifies the choice. A general approach to this problem is to make use of empirical methods such as probability plots or goodness-of-fit tests, while another is, to apply some approximation theorems from probability theory. Although some times such considerations may lead to reasonable models, neither they are of universal applicability nor they guarantee the correct solution all the time. The only tool that enables the determination of a probability model exactly, is a characterization theorem and therefore the study of such theorems has emerged as an important area of mathematical statistics. It is also not uncommon that many such theorems are found useful from theoretical considerations as well.

Suppose that, for a random variable X, there is a family F of distributions such that a distribution L(x) belongs to F implies that X has the property P. The characterization theorem concludes that if X exhibits P then L(x) belongs to F. Obviously, the two ingradients of a characterization problem are the family of distributions F and the property P.

The study of characterization of probability distributions appears to have begun with the work of Gauss in 1807 when he proved under certain conditions that the maximum likelihood estimate of the location parameter of a distribution is the sample mean if and only if, the distribution is normal. Eventhough reckoned from this work a long history can be attributed to the research activities in characterizing probability distributions, a full fledged development of this field as part of mathematical statistics, began taking shape only in the late fiftees of the present century.

Consistent with the emphasis placed on normal distribution in the early stages of development of statistical theory, initially the work on characterization theorems also were concerned primarily with normal models. Although in 1923 Polya characterized the normal law, by the identical distribution of two linear statistics, a real spurt in this direction began only with the conjecture of Levy in 1935, that for independent random variables X and Y, X+Y is normal if and only if X and Y also follow the same law. Cramer (1946) proved Levy's conjecture and in the following year Raikov established a similar result for Poisson variables. Modest activities in the forties due to the world war gave way to rapid growth in the next decade, with the review paper by Lukacs (1956) and the monograph by Lukacs and Laha (1964), which established a new line of thought. The first authoritative book on the tools employed in proving characterizations along with a large collection of results covering most probability distributions was published by Kagan, Linnik and Rao in 1973. This along with the books by Galambos and Kotz (1977), Mathai and Pederzoli (1977), Azlarov and Volodin (1986) and Patil et. al. (1975) contain most of the literature on the subject in recent times.

In the present thesis, we consider certain characterization problems associated with non-negative random variables X for which

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

is of the form a+bx for various values of a > 0 and all real values of b including zero. It can be deduced from the papers of Kotz and Shanbag (1980) and Xekalaki (1983a) reviewed in chapter II, that the Pareto, finite range and exponential models in the continuous case and the Waring, negative hypergeometric and geometric models in the discrete domain are the only probability laws admitting the above property. Thus in this sense, these six distributions can be thought of as a particular class sharing similar properties. The question that arises at this juncture is, whether there exist other characterizations which can bring together these models? Our investigation attempts to provide some answers in this direction. An answer to a search of this kind can be achieved in two ways. The first is, by identifying new results characterizing the above class of distributions and the second is to look at the existing characterizations that have been established for any one member of the above group and then to examine the feasibility of establishing a general property that holds for the entire family. In this connection it is to be observed that in the continuous case, there exist monotone transformations that can convert the exponential model into the Pareto and finite range models and accordingly it is possible in most cases to translate a property of the exponential into corresponding results in the other cases. Occasionally such translations become meaningful also in practical situations. But in the present investigation such results are not included and we present only those which are not directly implied by transformations.

The present work is organised into five chapters. After the introductory chapter currently being unfolded, we present in chapter II a review of some basic properties and characterizations of the afore-mentioned models. Since most characterizations of the exponential distributions based on order statistics, lack of memory property and its variants, provide only spin off results arising out of the monotonic transformations mentioned earlier for the other models, they are kept out of discussion. For the same reason, we have chosen the Pareto models as the pivot from which the other models are to be viewed.

In the subsequent three chapters some new results are presented. Chapter III deals with some characterizations of continuous models by properties of equilibrium distribution in section 3.1 and via certain conditional distributions in section 3.2. The next two sections, 3.3 and 3.4 discuss properties of partial and truncated reciprocal moments that are unique to some of the distributions. The chapter concludes with the conditions under which the continuous models admit characterizations in the category of additive damage models. In chapter IV, several theorems characterizing the Geometric, Waring and negative hyper-geometric distributions that are extensions of the results of the previous chapter to the discrete domain are presented. Also established are certain properties of residual life that are unique to these models. Most results proved in these two chapters have relevance to reliability and life testing. In the voluminuous literature in this area, continuous distributions are generally proposed as models of life lengths, in most studies. Accordingly, results and concepts have been developed in most works in this topic by treating time as continuous. We look at the role of discrete distributions in life-length studies and the possibility of developing several concepts parallel to those in the continuous case, that can describe the pattern of ageing. The study is concluded by presenting some new definitions and results in this direction in chapter V. These discussions permit to impart physical meaning to the various notions and characteristic properties of the models, already established in chapter IV, and thereby point out the applications of the results in the context of reliability analysis.

Chapter II

SURVEY OF LITERATURE

In the present chapter we consider some of the properties of the Pareto type I and type II, finite range and Waring distributions, that are of relevance to the investigations carried out in the succeeding chapters, along with an outline of the important developments in characterizing these models.

2.1 Pareto Type I Distribution

The Pareto type I distribution of a random variable X, denoted in the present investigation by PI(a,k) is specified by the probability density function,

$$f(x;a,k) = ak^{-a}\ x^{-(a+1)},\ a > 0,\ x \geqslant k > 0. \qquad (2.1)$$

The distribution is J-shaped with mode located at k. The mean and variance are

$$E(X) = a(a-1)^{-1}k,\ a > 1 \qquad (2.2)$$

and

$$V(X) = a(a-1)^{-2}(a-2)^{-1}k^2,\ a > 2\ . \qquad (2.3)$$

The r^{th} moment about the origin takes the form,

$$\mu'_r = a(a-r)^{-1}\ k^r,\ a > r,\ r = 0,1,2,\ldots\ . \qquad (2.4)$$

The truncated moment of order r defined as,

$$m_r(x) = E[(X-x)^r | X > x], \; r = 0,1,2,\ldots , \qquad (2.5)$$

for the distribution (2.1) is,

$$m_r(x) = r! \; x^r/(a-1)^{(r)}; \; x \geqslant k, \qquad (2.6)$$

where $a^{(r)}$ is the descending factorial expressed as

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

Specialising for r=1, we get the truncated mean, known more popularly as the mean residual life function (MRLF) in reliability theory. Thus,

$$m_1(x) = x(a-1)^{-1}. \qquad (2.7)$$

We notice that the MRLF is a linear function in x, for all $x \geqslant k$. On the other hand, the partial moment of order r,

$$p_r(x) = E[(X-x)^+]^r, \; r = 0,1,2,\ldots, \qquad (2.8)$$

where,

$$(X-x)^+ = \max(0, X-x),$$

is given by the expression

$$p_r(x) = r! \; k^a x^r/(a-1)^{(r)} \; x^a, \; x \geqslant k. \qquad (2.9)$$

Another type of moments, that is meaningful in connection with the Pareto distribution, is the reciprocal moments,

$$b_r = E(X^{-r}),\ r = 0,1,2,\ldots,$$
$$= a(a+r)^{-1}\ k^{-r}. \qquad (2.10)$$

The truncated version of (2.10) known as the truncated reciprocal moments defined by

$$C_r(x) = (-1)^r\ E[(\frac{1}{X} - \frac{1}{x})^r \mid X > x] \qquad (2.11)$$

has the value,

$$C_r(x) = (-1)^r\ ar!\ /\ (a+1)^{[r]}\ x^r,\ x \geq k, \qquad (2.12)$$

where, $(a+1)^{[r]}$ is the ascending factorial, given by

$$a^{[r]} = a(a+1)\ \ldots\ (a+r-1).$$

The distribution of a sum of Pareto variables that are independent and identically distributed is difficult to obtain. However, for the special case of $P\,I(1,k)$, there is a simple closed form for the distribution of X_1+X_2. This is

$$P(X_1+X_2>x) = 2[x^{-1} + x^{-2}\log(x-1)],\ x > 2. \qquad (2.13)$$

The problem of obtaining the distribution of the product of several Pareto variables can be made relatively simple, if one uses the transformation,

$$U = k_1 k_2 \ldots k_n \exp \left[\sum_{i=1}^{n} (Y_i/a) \right], \qquad (2.14)$$

where, Y_i's are independent standard exponential variables and $U = X_1 X_2 \ldots X_n$ with the X_i's following P I(a,k_i). The density function of U (Malik, 1970) is,

$$f(u) = (\Gamma n)^{-1} [k \log(u/k)]^{n-1} (u/k)^{-a} a u^{-1},$$
$$u > 0,\ k = k_1 k_2 \ldots k_n. \qquad (2.15)$$

Notice that here we have taken the parameter a to be the same for all the variables. When the shape parameter a is taken differently for the variables X_i, the form of the distribution becomes complicated. This aspect is discussed in detail by Pederzoli and Rathie (1980).

For deriving the distribution of the quotient $Z = X_1/X_2$ of independent Pareto variables with parameters (a_1,k_1) and (a_2,k_2) respectively, the method is to take the inverse Mellin transform of

$$E(Z^{s-1}) = a_1 k_1^{\,s-1} a_2 k_2^{\,1-s} / (a_1-s+1)(a_2-1+s).$$

In this way, Pederzoli and Rathie (1980) obtained the density of Z as

$$p(z) = \frac{a_1 a_2}{k_1 k_2^{-1}(a_1+a_2)} \left(z/k_1 k_2^{-1} \right)^{a_2-1} \text{ for } z \leqslant k_1 k_2^{-1},$$
$$= \frac{a_1 a_2}{k_1 k_2^{-1}(a_1+a_2)} \left(z/k_1 k_2^{-1} \right)^{a_1-1}, \text{for } z > k_1 k_2^{-1}. \qquad (2.16)$$

In view of the recent interest generated in the Pareto models as distributions of life lengths, it is desirable to look at some properties of P I(a,k) in this connection. Using the well known definition of failure rate

$$h(x) = f(x) \,/\, (1-F(x)),\ x > 0\ , \qquad (2.17)$$

where f(.) is the density function and F(.) is the distribution function of a non-negative random variable X satisfying the condition F(o) = 0, we find that for (2.1)

$$h(x) = ax^{-1},$$

which is a reciprocal linear function of x.

The life time remaining to an equipment at age x called the residual life time, is also a random variable say, Y_x, whose distribution is expressible in terms of the distribution function of the life length X. One can write the relationship as

$$G_x(y) = P[x < X \leq x+y \mid X > x],\ y > 0,$$
$$= 1-[R(x+y) / R(x)],$$

where,

$$R(x) = 1-F(x) = P[X > x],$$

is the survival function of X. Accordingly the survival function of Y_x is,

$$\bar{G}_x(y) = R(x+y)/R(x),\ y > 0. \qquad (2.18)$$

By direct calculation, for the P I(a,k) model,

$$\bar{G}_x(y) = \left(\frac{x+y}{x}\right)^{-a},\ y > 0, \qquad (2.19)$$

which is again of the P I form. Notice that

$$E(Y_x) = E[X-x|X > x] = r(x), \qquad (2.20)$$

is the MRLF defined earlier and shown to be linear in x in equation (2.7). An interesting property of the distribution is that,

$$r(x)\ h(x) = a(a-1)^{-1}, \qquad (2.21)$$

a constant greater than unity. A physical interpretation of the property and an associated characterization will be taken up in section 2.5. Further it is easy to see that h(x) is a decreasing function and r(x) is an increasing function of x so that P I belongs to the DFR (decreasing failure rate) and IMRL (increasing mean residual life) class of probability distributions. Further classes of life distributions based on different criteria of ageing, to which the Pareto models belong to will be taken up subsequently in chapter III.

2.2 Pareto Type II Distribution

The Pareto II distribution, occasionally referred to as the Lomax distribution also, is represented by the density function,

$$f(x) = a\alpha^{a} (x+\alpha)^{-(a+1)};\ x > 0,\ \alpha > 0,\ a > 0. \qquad (2.22)$$

The survival function corresponding to (2.22) becomes

$$R(x) = \alpha^{a} (x+\alpha)^{-a};\ x > 0,\ \alpha > 0,\ a > 0\,. \qquad (2.23)$$

We shall use the notation P II(a,α) to represent the Pareto type II distribution in (2.22). Dubey (1966) derives the same model as a special case of a compound gamma distribution and calls it exponential gamma distribution. If the conditional distribution of X has the exponential distribution with density function,

$$f(x|b) = be^{-bx}, \ x > 0, \ b > 0,$$

and if the parameter b has a gamma distribution,

$$g(b) = \alpha^k \ b^{k-1} \ e^{-\alpha b}, \ \alpha > 0, \ k > 0, \ b > 0,$$

then, the density of X is P II(k,α). There exists a monotone transformation $X = \alpha(e^{-u/a} - 1)$ that takes the standard exponential variable u to P II(a,α).

The r^{th} moment about the origin of the distribution is,

$$\mu'_r = \alpha^r \ \Gamma(a-r) \ \Gamma(1+r) \ / \ \Gamma(a) \ . \tag{2.24}$$

In particular,

$$E(X) = \alpha(a-1)^{-1}, \ a > 1, \tag{2.25}$$

and

$$V(X) = \alpha^2 \ a(a-1)^{-2} \ (a-2)^{-1}, \ a > 2.$$

The standard Pareto type II distribution with

$$R(x) = (1+x)^{-1},$$

has the special property that X and X^{-1} are identically distributed. Utilizing the result that $X = -\alpha + aZ^{-1/a}$, where Z is a rectangular variate in (0,1), Arnold (1983) has obtained the distribution of the sum of two independent random variables following P II(a,α).

The residual life distribution for P II(a,α) has survival function,

$$\bar{G}_x(y) = [(y+x+\alpha) \;/\; (x+\alpha)]^{-a}, \tag{2.26}$$

where $\bar{G}_x(y)$ is defined in equation (2.18). It is interesting to note that the residual life distribution is of the same form as the parent distribution with only a shift in the parameter from α to (x+α). Therefore, the MRLF is deduced from (2.26) as,

$$r(x) = (x+\alpha)\,(a-1)^{-1}. \tag{2.27}$$

The failure rate function is a reciprocal linear function,

$$h(x) = a(x+\alpha)^{-1}. \tag{2.28}$$

As in the case of P I(a,k), here also r(x) h(x) is a constant greater than unity. In spite of the simple form of the failure rate and MRLF, the potential of the Pareto distributions as useful models of failure times is yet to be fully exploited in life length studies. Since P II(a,α) arises by compounding exponential and gamma distributions as shown earlier, there is scope for the model to be used whenever the exponential distribution provides a satisfactory model in which the uncertainty in the parameter can be described in terms of a gamma distribution. Several examples of problems of this nature are discussed in literature such as Harris (1968) and Lindley and Singpurwalla (1986).

In view of the transformation $y=x+\alpha$, that changes the Pareto type I distribution to the type II distribution, it is easy to translate the properties of the former from that of the latter. Consequently, in the present investigation, the results are mainly obtained for the type II distribution with only occasional references to the type I.

2.3 Finite Range Distribution

The finite range distribution in the interval (0,R) is defined by the density function,

$$f(x) = (c/R)\ (1-x/R)^{c-1},\ 0<x<R,\ c>0, \tag{2.29}$$

and is denoted by FR(c,R). When c=1, the distribution reduces to the uniform distribution in (0,R). The distribution is L-shaped for c > 1, a straight line for c = 2 and J-shaped for c < 1. It is a particular case of the Pearson type I distribution with density function,

$$f(x) = (1/B(p,q))\ \frac{(y-a)^{p-1}\ (b-y)^{q-1}}{(b-a)^{p+q-1}}\ ,\ a<y<b,\ p>0,\ q>0, \tag{2.30}$$

as seen from the fact that when a=0 and p=1 (2.30) reduces to

FR(q,b). When R=1, in (2.29) we obtain the standard form of the model.

The r^{th} moment about the origin is,

$$\mu'_r = cR^r B(r+1,c),\ r=0,1,2,3,\ldots . \tag{2.31}$$

In particular, the mean and variance are

$$E(X) = R(c+1)^{-1},$$

and

$$V(X) = cR^2(c+1)^{-2}(c+2)^{-1}. \tag{2.32}$$

The moment generating function is,

$$M(t) = c\, e^{itR}\, I(c,tR), \tag{2.33}$$

where

$$I(p,q) = \int_0^R e^{-qx}\, x^{p-1}\, dx.$$

The distribution has been found useful in several areas of theoretical and applied statistics. From the reliability context, it is a model of life-lengths that have increasing failure rate. This is evidenced from the failure rate function

$$h(x) = c(R-x)^{-1}. \tag{2.34}$$

The MRLF is

$$r(x) = (R-x)(c+1)^{-1}, \tag{2.35}$$

which is linearly decreasing in x. Thus the FR(c,R) belongs to the IFR and DMRL class of life distributions.

In contrast to the Pareto variable, here $r(x)\ h(x)$ is a constant that is less than unity, for all values of x in (0,R). As already mentioned, the distribution belongs to Pearson family. It is also a member of the exponential family.

If $X_{(1)}, X_{(2)}, \ldots, X_{(N)}$ are order statistics of a random sample from a continuous distribution with density f(x), then

$$C_i = \int_{X_{(i-1)}}^{X_{(i)}} x\, dx \tag{2.36}$$

are called the elementary coverages of the random interval $(X_{(i-1)}, X_{(i)})$. The distribution of the i^{th} coverage C_i, $i = 1,2,\ldots,N$, is FR(N,1). This fact is utilized in non-parametric statistical inference (David, 1970). Apart from these, FR(c,1) inherits various properties and applications by virtue of its status as a translated beta distribution. It forms a special class of distributions along with the exponential and Pareto II models with reference to certain special characterizing and closure properties. These aspects will be investigated in chapter III.

2.4 Waring Distribution

One way of obtaining a discrete probability distribution is to consider a mathematical function admitting expansion as a convergent series of inverse factorials of positive terms and then by multiplying these terms by a suitable constant to render its sum unity. The Waring distribution belongs to the class of discrete models obtained in this manner, and make use of the Waring's expansion.

$$\frac{1}{(x-a)} = \frac{1}{x} + \frac{a}{x(x+1)} + \frac{a(a+1)}{x(x+1)(x+2)} + \dots \qquad (2.37)$$

The probability function of the Waring distribution discussed here is,

$$\begin{aligned} f(x) &= P[X=x], \\ &= (a-b)(b)_x/(a)_{x+1}, \ x=0,1,2,\dots, \end{aligned} \qquad (2.38)$$

where, $(b)_x$ is the Pochammer's symbol, defined as

$$(b)_x = \Gamma(b+x)/\Gamma b \ . \qquad (2.39)$$

The model (2.38) forms a particular case of the generalized Waring distribution introduced in Irwin (1975, a,b,c). It was originally found by Irwin (1963), in an attempt to encounter frequency distributions with very long tails suitable to describe the distribution of the number of philarial worms. It is J-shaped and forms the continuous analogue of

Pearson type VI distribution. Irwin (1968) has also shown that his generalized model has a theoretical basis, as a probability model for the number of accidents.

Looking at the properties of the simple model (2.38), we note that the r^{th} factorial moment is

$$\mu_{(r)} = (b)_r/(a-b-1)^{(r)}. \tag{2.40}$$

In particular,

$$E(X) = b(a-b-1)^{-1}$$

and

$$V(X) = b(a-1)(a-b)(a-b-1)^{-2}(a-b-2)^{-1}, \ a>b+2 . \tag{2.41}$$

The Yule distribution arises as a particular case of (2.38) when a=1.

Since the Pareto II distribution in the continuous case and the Waring distribution in the discrete case are heavy tailed distributions, it is natural to expect that they have similar properties. The limiting form of model (2.38) derived in section 4.2 of Chapter IV confirms this fact. From the point of view of reliability characteristics the resemblence is almost perfect.

The survival function of the distribution is

$$R(x) = \sum_{x}^{\infty} P[X=x],$$

$$= \sum_{x}^{\infty} (a-b)(b)_x/(a)_{x+1},$$

$$= (b)_x/(a)_x. \qquad (2.42)$$

The MRLF is

$$r(x) = E[X-x|X>x],$$

which is equal to (see equation 4.6)

$$r(x) = [R(x+1)]^{-1} \sum_{x+1}^{\infty} R(t),$$

$$= (a+x)\ (a+b-1)^{-1}, \qquad (2.43)$$

which is linear and the failure rate function,

$$h(x) = P[X=x]/P[X\geqslant x],$$

$$= (a+b)(a+x)^{-1}, \qquad (2.44)$$

is reciprocal linear. Xekalaki (1983a) has proposed the Waring distribution (2.38) which will be denoted in the rest of the discussions as W(a,b), as a life-length distribution in the discrete time domain. A generalized version of the same distribution has been used by him in relation to accident theory (Xekalaki, 1983b). There is yet another generalization of W(a,b) by Panaretos and Xekalaki (1986) to what they call as Waring distribution of order k, arising out of certain generalized sampling schemes. Since W(a,b) is only investigated in the present discussion the details of the other models and their properties are not presented here.

We shall also need a special case of the usual negative hypergeometric distribution,

$$P[X=x] = \binom{-a}{x}\binom{-b}{n-x} \Big/ \binom{-a-b}{n}, \quad x = 0,1,\dots,n, \qquad (2.45)$$

denoted as NH(a,b,n). The form of the model and its properties will be explained in connection with the characterization theorems in chapter IV.

2.5 Characterizations

In order to motivate certain characterization problems associated with the various models mentioned in the previous sections of this chapter and also to ascertain the present state of art, in this section we take up an overview of the important results in this connection. In view of the monotone transformations existing between the Pareto and exponential populations it is always possible to translate a characteristic property of the latter to suit the former. Since the literature on characterization of exponential distribution is so rich, several results for the Pareto distributions can be deduced in this manner. The following survey does not include any characterizations of this type.

The first characterization of the Pareto distribution appears to be that of Hangstroem (1925) which states that X is P I if and only if

$$E(X|X>t) = ct, \tag{2.46}$$

for some $c > 1$. It is easy to see that, in general, this property need not be true in the case of Pareto type I model only in view of the characterization of P II(a,α) by Laurent (1974). His result is that a mean residual life function of the form

$$r(x) = a+bx, \; x > 0 \; , \tag{2.47}$$

with $b > 0$, leads to P II$(ab^{-1},(b+1)b^{-1})$. Sullo and Rutherford (1977) observed that the relationship $h(x)r(x)>1$ is characteristic of the Pareto type II distribution. They further proved that a constant coefficient of variation of residual life with the constant greater than unity is a characteristic property of the same distribution. To be able to identify life-time models by the numerical value of $h(x)\ r(x)$ or the coefficient of variation of residual life one has to specify the classes of distributions corresponding to all numerical values these quantities can take, rather than to specialised values. In this sense, the characterization of Sullo and Rutherford (1977) provides only a partial answer by restricting the numerical value to be greater than unity. A complete answer to the problem is given in Mukherjee and Roy (1986) who proved the following result.

(1) If X is a non-negative random variable with finite expectation and $h(x)\ r(x) = k$, a constant, then $k=1$ if and only if X is exponential, $k>1$ if and only if X follows Pearson type XI distribution and $k<1$ if and only if X is FR(c,R).

(2) The coefficient of variation of residual life of X with a finite variance is less than, equal to or greater than one if and only if X is distributed respectively as FR(c,R), exponential and Pearson type XI distributions.

In the above paper a physical interpretation of the quantity $h(x)\ r(x)$ in the context of reliability is given as follows. The distribution of X belongs to a member of the increasing mean residual life or decreasing mean residual life class of distributions according as $h(x)\ r(x) >$ or < 1. We observe that another interpretation of the same is also possible. Kupka and Loo (1989) define

$$V_1(x) = E(X|X>x) \tag{2.48}$$

as the vitality function that represents the average age at failure of a component of life length X. The derivative of V(x) is the rate of vitality or gain in the conditional mean life of the component given that it has survived age x. Since

$$V_1(x) = r(x) + x,$$

and (2.49)

$$h(x) = [1+dr/dx]\,[r(x)]^{-1},$$

$$V_1'(x) = r'(x)+1$$

and (2.50)

$$V_1'(x) = r(x)\ h(x),$$

where the ''' denotes differentiation.

In this case, h(x) r(x) = 1 can be interpreted as a constant rate of vitality or no ageing. Similar interpretation would mean the negative ageing of the component for h(x) r(x) > 1 and positive ageing corresponding to the reverse inequality.

The proposition of linearly increasing mean residual life times can be achieved from other considerations as well. Assuming the distribution of life lengths to be exponential, Morrison (1978) considered the question of the possible mixing distributions for the exponential parameter that can guarantee a linearly increasing mean residual life function. He proved that the gamma distribution is the only absolutely continuous model that meets the above requirement. It may be noticed that the compound distribution arising from the exponential and gamma is Pareto type II and therefore, Morrison's (1978) result only confirms the functional form of the mean residual life proposed in the earlier results. Gupta (1980) generalized Morrison's (1978) results by describing a method based on Laplace transform technique to determine the mixing distributions when the life distribution is exponential. The most general result concerning the form of MRLF appears to be that of Kotz and Shanbag (1980) which can be stated as follows.

Let F be the distribution function of a random variable X such that its restriction to a non-degenerate interval (α,β) is absolutely continuous with respect to Lebesgue measure with $E(X) < \infty$, then the failure rate will be a polynomial or a reciprocal polynomial and the MRL function is a polynomial or a reciprocal polynomial if and only if

(i) $F(\beta-)-F(\alpha) = 0$,

or for $\alpha > -\infty$ and $F(\beta-)-F(\alpha) > 0$ either

(ii) $G(x) = \exp -a(x-\alpha)$,

for some $a > 0$ together with

$$\int_{\beta}^{\infty} G(y)dy = a^{-1} \exp -a(\beta-\alpha),$$

or

(iii) $G(x) = [1+c(x-\alpha)]^n, \; c > 0, \; n \leqslant -2,$

or

(iv) $G(x) = [1+c(x-\alpha)]^r, \; c \geqslant (\alpha-\beta)^{-1}, \; r > 0,$

or negative non-integer satisfying,

$$\int_\beta^\infty G(y)dy = -[c(r+1)]^{-1} \lim_{x \to \beta} [1+c(x-\alpha)]^{r+1},$$

where

$$G(x) = [1-F(x)]/[1-F(\alpha)] .$$

In spite of the fact that the MRL is widely discussed in theory and practice it has several limitations. The impracticability of waiting until all items have failed, information on failure times is available only on a censored basis and the high sensitivity to this function to very large values can be cited as some of the reasons for this. Accordingly one can use the median of the residual life times as an alternate to the MRL. This is defined as

$$M(S|t) = R^{-1}(\tfrac{1}{2} R(t))-t.$$

Schmittlein and Morrison (1981) have shown that

$$M(S|t) = a+bt, \; a > b > 0$$

if and only if X is P II.

A considerable volume of literature is available on characterizations based on truncation invariance of the Pareto variables. In such investigations we look at the properties of the random variable X that remain unaltered when the variable is subjected to the right truncation, $X > x$. Bhatacharya (1963) was the first to find a characterization in this direction. His result is that the Lorenz curve and the Gini index will be independent of the point of truncation if and only if X is distributed as P I. Ord et. al. (1983) provided a rigorous proof to Bhatacharya's (1963) result on the Gini index and also established that the property of measures of inequality derived from the Mellin-transform

$$H_r(c) = [E\{(X/\mu_c)^{r+1}|X \geqslant c\}-1]\ [r(r+1)]^{-1},\ -\infty<r<\infty \quad (2.51)$$

where

$$\mu_c = E(X|X \geqslant c),$$

being invariant under truncation. Their results can be stated as follows.

(1) When the density function f(x) of the income distribution is positive almost everywhere on its range $[C_L,\infty)$, the Gini index is truncation invariant if and only if

$$f(x)=\begin{cases} k\ C_L^k / x^{k+1}, & 0 < C_L \leqslant x,\ k > 1, \\ 0 & , \text{otherwise}. \end{cases} \quad (2.52)$$

(2) When $f(x)$ is positive everywhere on its range, the index H_r in (2.51), $-\infty < r < \infty$, is truncation invariant if and only if X has a probability density function (2.52).

Assuming the existence of the r^{th} moment, Dallas(1976) characterized the Pareto distribution by the condition that the r^{th} truncated moment (see equation (2.5)) is the same as the r^{th} moment of the original distribution suitably scaled. According to him, if Y is a random variable having absolutely continuous distribution function with $E(Y^r) < \infty$, then

$$E(Y^r|Y<c) = E(cY/k)^r, \quad 0<k\leq y, \quad c\geq k \,, \tag{2.53}$$

holds for some $r > 0$ if and only if Y has density (2.1).

Krishnaji (1970a) proved the following two results.

(1) Let X be a random variable having absolutely continuous distribution function and R be a random variable having p.d.f

$$h(r)=\begin{cases} pr^{p-1}, & 0\leq r\leq 1,\ p>0 \,, \\ 0 & , \text{ otherwise} \,, \end{cases} \tag{2.54}$$

such that $P(RX > x_0) > 0$, for some $x_0 > 0$. Then the distribution of RX truncated to the left at x_0 coincides with that of X if and only if X has a Pareto II distribution on (x_0,∞).

(2) Let Z and X be random variables such that

$$E(Z|X=x) = \alpha + \eta x,$$

and X has a continuous marginal density function. Further, let R be a random variable independent of Z and X with a density (2.54), then

$$E(Z|Y=RX=y) = \begin{cases} \alpha+\lambda y, & y>x_o,\ x_o>0,\ \lambda(\lambda-\eta)^{-1} > p, \\ \text{constant}, & y_o \leqslant x_o, \end{cases} \qquad (2.55)$$

if and only if X has a Pareto II distribution on (x_o,∞).

The significance of the last two theorems is that they become quite meaningful once we interpret X as the reported income, Y as the true income and R as the under reporting error. With this interpretation Krishnaji's (1970a) results say that the observed income distribution, truncated to the left is the same as the reported income and a variable having linear regression on true income has a linear regression on observed incomes also, if and only if income distribution is Pareto type II. Although Krishnaji has assumed a power distribution for R, it has been shown later by Lau and Rao (1982) that the result is true for any continuous distribution of R in the range [0,1]. (See Krishnaji (1971) also.) However, we observe that to be able to ascertain that income distribution is in fact Pareto in a real situation, one must verify that the truncated distributions of reported and true incomes are identical. Almost always it is the case that the model of the true incomes, truncated or not, is seldom known exactly.

By assuming an alternative formulation $Y=X-R$, $0 < R < \max(0, x-m)$, where m is the tax-exemption level and

$$E(R \mid x > m) = a+bx, \quad 0<b<1 \;, \tag{2.56}$$

Revankar et. al. (1974) shows that for the relation

$$E(R \mid X>y) = c+dy; \quad d > b > 0 \;, \tag{2.57}$$

to hold it is necessary and sufficient that X has Pareto II distribution with finite mean. The theorem is also proved for Pareto I distribution with $c=a$. The work of the last two authors belongs to the general class of models referred to in literature as damage models. A closer look at the properties of (2.56) and (2.57) and the corresponding characterizations will be attempted in chapter IV.

Talwalker (1980) defines a property called "dullness" analogous to the lack of memory of the exponential distribution by calling X to be totally dull at a point t in its support if

$$P[X > st | X \geq t] = P[X > s] \tag{2.58}$$

for all $s \geq 1$. She proves that the Pareto I (a,1) is the only distribution which is totally dull at all points in its support and further that the dullness of the distribution of true incomes at a single reported income point is sufficient to characterize the distribution as Pareto, provided that the distribution function of X is concave. It is easy to see that the first result is closely related to the well known characterization of the exponential distribution by lack of memory.

Yet another characterization of the Pareto I is based on the maximization of entropy (Ord, Patil and Taillie,1981) which seeks a distribution with support $[c, \infty)$ that maximizes $\int_c^\infty f(x) \log f(x)dx$ on the condition that the geometric mean of the distribution is fixed.

The most recent characterizations of the Pareto distributions appears to be that of Korwar (1989). His results can be summerized as follows.

(1) Let X be a positive random variable on $[a,\infty)$ with density $f(x)$ and distribution function $F(x)$. Let Z and W be random variables with respective densities,

$$k(z) = f(z) / [Z \ E(\tfrac{1}{Z})] , \tag{2.59}$$

and

$$L(\omega) = [1-F(\omega)]/\mu, \ \mu = E(X), \tag{2.60}$$

then Z and W have the same distribution if and only if X is Pareto I.

(2) If Y is distributed with density

$$g(y) = \int_0^y xy^{-2} f(x)dx, \ y > a \tag{2.61}$$

then Y and W have the same distribution if and only if X is Pareto II.

There are a large number of results on characterizing the Pareto distributions based on order statistics. Since order statistics of the distributions so far discussed will not be considered in the present work, a detailed exposition of such characterizations is not attempted here. But for the sake of completion of the survey of characterizations we refer to the papers in this connection as Fisz (1958), Rogers (1959, 1963), Rossberg (1972 a,b), Malik (1970), Govindarajlu (1966), Ahsanullah and Kabir (1973), Ferguson (1967), Dallas (1976), Srivastava (1965), Mosimann (1970), James (1979),Crawford(1966), Beg and Kirmani (1974), Shah and Kabe (1981) and Wang and Srivastava (1980).

Analogous to the characterization of Pareto II distribution due to Krishnaji (1970b) in the context of under-reporting of incomes Xekalaki (1983a) obtains a characteristic property of the Yule distribution. The result states that for a random variable U which is uniform over (0,1), the distribution of [UX] truncated at zero coincides with that of X if and only if X has a Yule distribution where [b] denotes the greatest integer in b. Investigating the conditions under which distributions of two random variables are identical, Korwar (1989) has given several characterizations of the Waring (and hence Yule) distribution and its truncated versions. Let X, Y and Z be positive integer valued random variables with respective probability functions,

$$p_r = P[X=r],\ r=1,2,\ldots, \tag{2.62}$$

$$q_y = P[Y=y],$$

$$= \sum_{x=1}^{y} x\, p_x/y(y+1),\ y=1,2,\ldots, \tag{2.63}$$

and

$$q_r' = P[Z=r],$$

$$= (r+a)\, p_r(\mu+a),\ r=1,2,\ldots, \tag{2.64}$$

where $\mu = E(X) < \infty$ and $a > -1$. Further W is assumed to be non-negative integer valued random variable with probability function,

$$q_r'' = P[W=r] = \sum_{j=r+1}^{\infty} p_j \, \mu^{-1}, \quad r=0,1,2,\ldots . \qquad (2.65)$$

The results are stated as follows:

(1) The random variable W truncated on the left at zero has the same distribution as Z if and only if X has Waring distribution W(λ,C) given by (2.38).

(2) When $E(Y^2) < \infty$, Y and Z have the same distributions if and only if the distribution of X is $W(\mu-1, \mu+1)$.

(3) The necessary and sufficient condition that Y and W truncated at zero to have the same distribution is that X follows $W(\mu-1, \mu+1)$.

He further considers a non-negative integer valued random variable X with probability mass function,

$$p_r' = P[X=r], \quad r=1,2,\ldots,m , \qquad (2.66)$$

and defines a new random variable Y such that,

$$q_y''' = P[Y=y] ,$$
$$= \begin{cases} \sum_{x=1}^{y} x \, p_x' \,/\, y(y+1), & y=1,2,\ldots,m, \\ \sum_{x=1}^{y} x \, p_x' \,/\, y(y+1), & y=m+1, \ldots . \end{cases} \qquad (2.67)$$

It is shown that the distribution of

(a) Y truncated on the right at m and Z have the same distribution if and only if X has the Waring $W(\mu-1,\mu+1)$ distribution truncated on the right at m.

(b) Y truncated on the right at m and W truncated on the left at zero have the same distribution implies and is implies by a Waring $W(\mu-1, \mu+1)$ distribution for X truncated on the right at m.

Xekalaki (1983a) introduces a class of life distributions in the discrete time domain consisting of the geometric, Waring (2.38) and the negative hyper-geometric (2.45) with probability function,

$$P[X=x] = P_x, \; x = 0,1,2,..., \tag{2.68}$$

which shares the property that the hazard function is inversely proportional to some linear function of time. He characterizes the probability law of a random variable in the support of $\{0,1,2,...,m\}$, $m \in \{0,1,...\} \cup \{+\infty\}$ by assuming that for $0 < P[X \leqslant 0] < 1$, the failure rate function is of the form

$$h(x) = (a+bx)^{-1}, \; x = 0,1,2,...,m \; , \tag{2.69}$$

to be geometric for $b = 0$, Waring (2.38) for $b > 0$ and

Negative hypergeometric (2.45) for $b < 0$.

It does not seem that much literature is available on characterizations of discrete models by using reliability concepts. The fact that generally continuous distributions are considered in life length studies by treating time as continuous, may be one of the reasons for the lack of interest in discrete distributions in this field. We will examine this point more closely in chapter IV and present some results relating to reliability concepts in discrete time.

Chapter III

CHARACTERIZATION OF CONTINUOUS MODELS

3.1 Characterization By Properties Of Equilibrium Distribution*

Let X be a random variable admitting absolutely continuous distribution function F(x) with respect to Lebesgue measure in the support of the set of non-negative real numbers. Further, assume that F(o) = 0 and $m = \int_0^\infty x dF < \infty$. Associated with X a random variable Y can be defined with probability density function,

$$g(y) = m^{-1}R(y), \; y > 0 , \tag{3.1}$$

where R(x) = P[X>x] is the survival function of X. The distribution specified by X has special significance in renewal theory. Consider a set of components whose failure times are of interest to us and we start experimenting with a new component at time zero, replace it upon failure by a second component and so on. If the failure times X_i, i = 1,2,3,... of the components are independent and identically distributed random variables, then the sequence (S_n) of points where $S_n = X_1 + X_2 + \ldots + X_n$ constitutes a renewal process. If F(.) is the common distribution function

---

* Some results in this section have appeared in the J.Ind.Statist.Assoc. (Reference 66).

of the X's satisfying the conditions stipulated above and U_y and V_y respectively denote the age and remaining life time (residual life) of the component which is in use at time y, then the limiting distributions of U_y or V_y is shown in Cox (1967) to have density function (3.1). The random variables U_y and V_y are called the backward and forward recurrence times and their common asymptotic distribution (3.1), the equilibrium distribution. In this physical situation, Y represents the residual life of the component whose length of life is X. An alternative derivation of the same model based on length biased sampling is also given in Cox (1967).

Deshpande et. al. (1986), Singh (1989) and Bluementhal (1967) have found several applications of the equilibrium distribution in reliability studies. The probabilistic comparison between the distribution functions of X and Y were utilised to explain the phenomenon of ageing.

In the present section we develop some identities connecting the failure rates and MRLF's of the random variables X and Y. Eventhough we use them in the sequel to characterize only the exponential, Pareto, Power and finite range distributions, the relationships are quite general in character and can be used to characterize the distribution of any continuous non-negative random variable.

3.1.1 Some Identities

Let the failure rates of X and Y be

$$h(x) = f(x)/R(x), \tag{3.2}$$

and

$$k(x) = g(x)/S(x), \tag{3.3}$$

where f(.) is the density of X and S(.) is the survival function of Y. If the MRL functions of X and Y are r(x) and s(x) respectively, we have from equation (2.5) specialised for r=1,

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

$$= [R(x)]^{-1} \int_x^\infty R(t)dt, \tag{3.4}$$

and

$$s(x) = [S(x)]^{-1} \int_x^\infty S(t)dt. \tag{3.5}$$

Using the result (Gupta, 1979),

$$k(x) = [r(x)]^{-1}, \tag{3.6}$$

and the relationships,

$$h(x) = [1+r'(x)]/r(x) ,$$

and

$$k(x) = \frac{1+s'(x)}{s(x)} , \tag{3.7}$$

the following identities connecting the various functions are easily proved.

$$h(x) = k(x) - k'(x)/k(x) \quad , \tag{3.8}$$

$$r(x) = s(x)/[1+s'(x)] \quad . \tag{3.9}$$

The primes in the last two equations and elsewhere in the present study denotes differentiation. In view of the one-to-one relationship between failure rates, MRL function and the corresponding survival function, the distribution of X be inferred from that of Y as either

$$R(x) = \exp\,[-\int_0^x h(t)dt] \quad ,$$

$$= (k(x)/k(o))\,\exp[-\int_0^x k(t)dt], \tag{3.10}$$

in terms of the failure rates or as

$$R(x) = (r(o)/r(x))\,\exp\,[-\int_0^x (r(t))^{-1}\,dt],$$

$$= \frac{1+s'(x)}{1+s'(o)} \left\{\frac{s(o)}{s(x)}\right\}^2 \exp[-\int_0^x (s(t))^{-1}\,dt\,], \tag{3.11}$$

in terms of the MRL function, provided the quantities on the right side exist. The equations from (3.6) through (3.11) are quite useful in distribution theory as we shall see in the next few sections. Further, they can be applied to provide alternative simple proofs to many of the theorems in literature concerning the inter relationships between the various criteria for ageing and different modes of stochastic orderings. Since our main thrust is on characterizations, these aspects are not considered in the present study.

3.1.2 Characterizations

Since our main objective is to explore characterizations of the Pareto, exponential and finite range distributions, attention is confined only to MRL functions that are linear. However, the techniques employed can take care of functions of a quite general nature and therefore, be used in other cases as well.

Theorem 3.1.

The MRL function of X is linear if and only if the MRL function of Y is linear.

Proof.

Suppose that Y has linear MRL function.

Then

$$s(x) = \ell x + m.$$

For s(.) to be an MRL function it is necessary and sufficient that $s(o) = E(Y)$, $s'(x) \geq -1$ and

$$\lim_{x \longrightarrow o} s(x)/(x \log x) = 0,$$

(Muth, 1980). Hence the constants ℓ and m must satisfy the conditions $m > 0$ and $\ell \geq -1$.

Using the above form of s(x) in equation (3.9), we find

$$r(x) = Lx + M,$$

where

$$L = \ell/(1+\ell) \text{ and } M = m/(1+\ell) \tag{3.12}$$

For $\ell \geq -\frac{1}{2}$, $L \geq -1$ and $M > 0$.

Conversely when r(x) has the above form, from (3.9),

$$s'(x) + [s(x)-1]\,(Lx+M)^{-1} = 0.$$

This is a linear differential equation with integrating factor $(Lx+M)^{-1/L}$. Accordingly the general solution is

$$s(x) = \begin{cases} (1-L)^{-1}\,(Lx+M)+K(Lx+M)^{1/L}, & L \neq 1 \\ -(x+M)\,\log(x+M)+K_1(x+M), & L = 1 \end{cases}$$

K and K_1 being the constants of integration. In view of the conditions on s(x), the second solution is inadmissible and in the first solution K must be zero. Thus,

$$s(x) = \ell x + m,$$

where

$$\ell = L/(1-L) \text{ and } m = M/(1-L) \tag{3.13}$$

in conformity with equation (3.12).

Theorem 3.1 will now be utilised to establish a closure property, for the distribution of X with respect to the formation of the distribution of Y, in the sense that X and Y have the same form of distribution. For this purpose, we denote by E(b), the exponential distribution with density function,

$$f(x) = b \exp[-bx] . \tag{3.14}$$

Theorem 3.2.

The distribution of X is E(b), (P II(a,α), FR(c,R)) if and only if Y is E(b), (P II(a-1,α), FR(c+1,R)).

Proof.

When X is distributed as one of the above forms, we have from chapter II the following table of values of MRL, failure rate and means.

Table 1

Failure rates, MRLs and means.

Model	MRL	failure rate	mean
$E(b)$	$1/b$	b	$1/b$
$P\ II(a,\alpha)$	$(x+\alpha)(a-1)^{-1}$	$a(x+\alpha)^{-1}$	$\alpha(a-1)^{-1}$
$FR(c,R)$	$(R-x)(c+1)^{-1}$	$c(R-x)^{-1}$	$R(c+1)^{-1}$
Uniform(o,R)	$(R-x)2^{-1}$	$(R-x)^{-1}$	$R/2$

In the notations of Theorem 3.1, the values of L for $E(b)$, $P\ II(a,\alpha)$ and $FR(c,R)$ are respectively $0,(a-1)^{-1}$ and $-(c+1)^{-1}$ while corresponding values of M are b^{-1}, $\alpha(a-1)^{-1}$ and $R(c+1)^{-1}$. Direct calculations from equation (3.13) give $\ell = 0$, $M = b^{-1}$ when X is exponential, $\ell = (a-2)$, $M = \alpha(a-2)^{-1}$ in the Pareto case, and $\ell = -(c+2)^{-1}$, $M=R(c+2)^{-1}$ for the finite range law. By theorem 3.1 Y has linear form and since $\ell = 0$ (> 0, < 0) Y is $E(b)$,($P\ II(a-1,\alpha)$, $FR(c+1,R)$). The conditions on the parameters of three distributions in the order in which they appear in the theorem for the result to be true are $b > 0$, $\alpha > 0$, $a > 1$, $R > 0$ and $c > -1$.

This theorem enables one to assert the distribution of Y in terms of the distribution of X and vice versa. It may be noticed that the closure property of the exponential distribution proved above is well known and has been properly exploited in literature. That the property holds for a class of distributions, points out to the possibility of a unified treatment of the tail behaviour of the different models involved therein. The present study being oriented more towards applications in reliability theory, some equivalent conditions in terms of the failure rates or MRL functions are proved in the following two theorems. In life length studies, models are often postulated by specifying the behaviour of the MRL's or failure rates and in such situations these theorems may help in identifying the appropriate models.

Theorem 3.3.

The relationship $r(x) = ps(x)$ is satisfied for all $x > 0$ and positive constant p if and only if X is E(b) for p=1, P II(a,α) for $0 < p < 1$ and FR(c,R) for $p > 1$.

Proof:

When X follows (3.14), Y has an identical distribution with the same parameter. Hence $r(x)=s(x)$. For the Pareto case also Y has the same distributional form P II$(a-1,\alpha)$ and therefore from table 1, for a > 2, (condition for E(Y) to exist)

$$r(x)/s(x) = (a-2)/(a-1) < 1.$$

Likewise, for the finite range distribution,

$$r(x)/s(x) = (c+2)/(c+1) > 1.$$

This proves the necessity of the condition.

To prove the only if part, we note that from equations (3.6) and (3.7),

$$r(x) = s(x)/(1+s'(x)). \qquad (3.15)$$

When $p=1$, it follows that $s'(x)=0$ or $s(x)$ is a positive constant. Hence X follows (3.14) for some $b > 0$. When $p<1$, equation (3.15) gives,

$$s(x)/[1+s'(x)] = ps(x),$$

or

$$s'(x) = p^{-1}-1.$$

Thus,

$$s(x) = (p^{-1}-1)x+d,$$

where, $d = E(Y)$. Accordingly the survival function of X is

$$R(x) = \exp -\left[\int_0^x (1+r'(t))dt/r(t) \right],$$

$$= \exp -\left[\int_0^x (2-p)dt/[(1-p)t+dp] \right],$$

$$= \alpha^a (x+\alpha)^{-a},$$

with,

$$\alpha = dp(1-p)^{-1} > 0 \text{ and } a = (2-p)(1-p)^{-1} > 1.$$

This condition on the parameter a is necessary for the existence of the distribution of Y. Lastly for $p > 1$, the same type of calculations imply that X follows FR(c,R), with

$$R = dp(p-1)^{-1} \quad \text{and} \quad c = (p-2)(p-1)^{-1}.$$

This completes the proof.

Theorem 3.4.

$h(x) = pk(x)$ for all $x > 0$ and a positive constant p if and only if X is E(**b**) for $p = 1$, P II(a,α) for $p > 1$ and FR(c,R) for $0 < p < 1$.

Proof:

The proof is on similar lines as that of theorem 3.3 once we write

$$k(x) = 1/r(x) \text{ and } h(x) = [1+r'(x)]/r(x) \cdot$$

Corollary 3.1.

The conditions of Theorems 3.3 and 3.4 characterize the distribution of Y as either exponential or Pareto type II or finite range for values of p mentioned herein. The proof follows from Theorem 3.2.

Remark:

When the support of X is taken as (a,b), b>a>0, the condition $r(x) = ps(x)$ or $k(x) = ph(x)$ for $p > 1$ characterizes the distribution with survival function,

$$R(x) = [(b-x)/(b-a)]^c, \quad c > 0,$$

and the same relationships for $p < 1$ is a unique property of the model specified by,

$$R(x) = [(a+d)/(x+d)]^c, \quad x > a,$$

which includes the Pareto type I as a particular case, when $d = 0$.

The characterization theorems presented so far demands the required properties to hold for all points in the support of the random variable X. One problem that will be of interest is to inquire whether it is possible to identify a weaker set of criterion that can uniquely determine the above mentioned distributions. The following theorem answers the question in the affirmative, with the limitation that the models have to come from a subclass of the exponential family. This is in contrast with the earlier results where one had the freedom of choice from the class of all absolutely continuous distributions of non-negative random variables.

Theorem 3.5.

In the one parameter exponential family specified by,

$$f(x) = u(\theta)\ v(x)\ \exp[\theta\ \omega(x)], \tag{3.16}$$

where θ lies in an open interval on the positive part of the real line, the condition

$$E(X) = k\ E(Y),\ k > 0, \tag{3.17}$$

is satisfied, if and only if X is distributed as

(i) exponential for k=1 and $\omega(x) = x$,

(ii) Pareto type I for $k > 1$ and $\omega(x) = -\log x$ (k, known) and

(iii) Power with density,

$$f(x) = (\theta+1)R^{-(\theta+1)}\, x^{\theta},\quad 0 < x < R,$$

for $0 < k < 1$ and $\omega(x) = \log x$.

Proof:

The proof that k=1 characterizes E(b) is given in Gupta (1979). To prove the other two cases we note that from equation (3.17) and from the definition of the random variable Y in (3.1),

$$\begin{aligned} E(Y) &= m^{-1} \int_0^\infty y\, R(y)dy, \\ &= (2m)^{-1}\, E(X^2), \qquad (3.18) \end{aligned}$$

by the usual integration by parts. Further in the support D of the family,

$$\int_D u(\theta)\, v(x)\, e^{\theta\omega(x)}\, dx = 1.$$

Changing θ to $\theta+1$ in the last expression,

$$\int u(\theta+1)\, v(x)\, e^{(\theta+1)\omega(x)}\, dx = 1,$$

and hence,

$$u(\Theta+1)/u(\Theta) \int u(\Theta)\ v(x)\ e^{(\Theta+1)\omega(x)}dx = 1. \qquad (3.19)$$

When $\omega(x) = \log x$, (3.19) reduces to

$$E(X) = u(\Theta)/u(\Theta+1), \qquad (3.20)$$

and similar calculations after replacing Θ by $\Theta+2$ provides,

$$E(X^2) = u(\Theta)/u(\Theta+2). \qquad (3.21)$$

Now, equation (3.17) along with (3.18) read,

$$2m\ E(X) = k\ E(X^2),$$

or

$$2[E(X)]^2 = k\ E(X^2). \qquad (3.22)$$

Introducing equations (3.20) and (3.21) into (3.22),

$$u(\Theta)/u(\Theta-1) = p\ u(\Theta-1)/u(\Theta-2),$$

which is identical to

$$b(\Theta) = p\ b(\Theta-1),\ \Theta \geq 1, \qquad (3.23)$$

where ,

$$b(\Theta) = u(\Theta)/u(\Theta-1) \text{ and } p = k/2. \qquad (3.24)$$

The next stage of the proof is the solution of the functional equation (3.23). Setting $\theta=1$ there $p=b(1)/b(o)$. Further transformation $c(\theta)=b(\theta)/b(o)$ yields

$$c(\theta) = c(1)\ c(\theta-1) . \qquad (3.25)$$

By successive application of (3.25), for non-negative integer n,

$$c(\theta+n) = c(\theta)\ g(n),$$

in which

$$g(n) = [c(1)]^n.$$

Since $c(o)=1$, $c(n)=g(n)$ and hence,

$$c(\theta+n) = c(\theta)\ c(n),$$

which is the Cauchy functional equation with solution

$$c(\theta) = p^\theta,\ p = c(1).$$

Thus,

$$b(\theta) = b(o)\ p^\theta$$

$$= qp^\theta,\ b(o) = q > 0.$$

Substituting into (3.24) ,

$$u(\theta) = qp^{\theta} u(\theta-1).$$

Recurrsively,

$$u(\theta+n) = q^n p^{n\theta+n(n+1)/2} u(\theta),$$

and in particular,

$$u(n) = q^n p^{n(n+1)/2} u(o).$$

This gives,

$$u(\theta+n) = p^{n\theta} u(n) u(\theta) u(o), \quad (3.26)$$

and

$$u(o) = 1.$$

Taking $g(\theta) = p^{\theta^2/2}/u(\theta)$, once again we have the Cauchy functional equation mentioned earlier with solution,

$$g(\theta) = [g(1)]^{\theta},$$

$$= [p^{1/2}/u(1)]^{\theta} = (qp^{1/2})^{-\theta},$$

and whence,

$$u(\theta) = p^{\theta(\theta+1)/2} q^{\theta}.$$

The density function of X now becomes,

$$f(x) = p^{(1/2)\theta(\theta+1)}\; q^{\theta}\; v(x)\; x^{-\theta}.$$

Assuming the support of X to be (q,∞), one must have,

$$p^{(1/2)\theta(\theta+1)}\; q^{\theta} \int_q^{\infty} v(x)\; x^{-\theta}\, dx = 1,$$

or equivalently,

$$p^{(1/2)\theta(\theta+1)} \int_q^{\infty} v(x)\; x^{-\theta}\, dx \;=\; q^{-\theta}.$$

Differentiating both sides with respect to q,

$$v(q) \;=\; \theta/qp^{(1/2)\theta(\theta+1)} \quad . \tag{3.27}$$

Using the expression for v(.) in equation (3.16),

$$f(x) \;=\; \theta q^{\theta}\; x^{-(\theta+1)},\; x > q > 0,$$

so that X is P I(θ,q). The proof for $k < 1$ is on similar lines except for the choice of $\omega(x) = \log x$ and our assertion follows.

Among the three sub-classes of the exponential family considered in Theorem 3.5 has the following implications.

(i) Since the coefficient of variation of X is $C = (\frac{2}{k} - 1)^{1/2}$, the values C=1, C>1 and 0 < C < 1 characterize respectively E(b), P I(a,k) and the power distribution.

(ii) Muth (1980) has introduced the concept of memory of a distribution at a point x in its support as $-r'(x)$, which we can see to be zero for E(b), negative constant for P I(a,k) and positive constant for power distribution at each point of its support. Our previous discussions show that these are the only continuous distributions possessing this property.

(iii) For arbitrary distributions of non-negative random variables Muth (1980) considers a weighted average of the memories at various points to arrive at a measure for global memory of a distribution as $M = 1-C^2$ and classify the distributions as possessing no memory, positive memory or negative memory according as M = 0, M > 0 or M < 0. Since the weights he has chosen in the calculations are positive, it is evident that if a distribution has a particular type of memory at every point in the support it also has the same type of memory globally. But the converse need not be true as seen from the expression for M and also from the fact that mixed types of memory at various points can produce a negative or a positive quantity as the average. Theorem 3.5 shows that the exponential, Pareto and Power models are the only distributions in the families considered above are characterized by a global lack of memory, negative memory and positive memory.

Another possible application of Theorem 3.5 lies in the measurement of income inequality. A random variable Z has the first moment distribution corresponding to X, if it has distribution function,

$$F(z) = m^{-1} \int_{0}^{z} xf(x)dx \; . \tag{3.28}$$

When X is the income of a unit, F(z) is interpreted as the proportional share of total income of units having income upto z, and is extensively used to define and interpret the Lorenz curve and Gini index. (See Kakwani, 1980). We now prove

Theorem 3.6.

$F_Z(x) = F_Y(x)$ for all $x > \theta$ if and only if X is Pareto type I.

Proof:

The necessary part is easily verified. To prove the sufficiency part we observe as follows. Let X ranges from θ to ∞. In order to make the span of the distribution function 1, we change E(X) to E(X)$-\theta$, the mean of the distribution being $(m-\theta)$. Thus if f(.) is the density function of Y,

$$f_Y(x) = (m-\theta)^{-1} R_X(x), \; x > \theta > 0, \tag{3.29}$$

then,

$$F_Z(x) = m^{-1} \int_0^x t\, f(t)dt,$$

$$= m^{-1} \int_0^x t(-dR(t)).$$

Since X ranges from Θ to ∞, we have,

$$F_Z(x) = -m^{-1} \int_\Theta^x t\, dR(t),$$

$$= -m^{-1} \int_\Theta^x t(m-\Theta)\, f'(t)dt, \qquad (3.30)$$

using (3.29).

When Z and Y have the same density function, (3.30) leads to,

$$f(x) = -m^{-1}\, x(m-\Theta)\, f'(x),$$

so that

$$\log f(x) = \log C\, x^{-(m/(m-\Theta))},$$

where C is a constant, thus giving,

$$f(x) = Cx^{-a},$$

with $a = m(m-\Theta)^{-1} > 0$ and $C = (a-1)\,\Theta^{(a-1)}$.

When income inequality is measured in terms of the index E(Z)/E(X) (See Arnold 1983) Theorem 3.6 offers a characterization in that direction as well.

3.2 Characterization By Properties Of Bivariate Models.

There have been several investigations concerning bivariate distributions that are determined uniquely by conditional densities of specified forms. The papers by Abrahams and Thomas (1984), Arnold (1987), Arnold and Strauss (1988) and Nair (1989) discuss this problem extensively. It is also possible to think in terms of characterizing univariate models by assuming particular forms of conditional distributions derived from certain bivariate models. Seshadri and Patil (1964) has shown that if the conditional density of X_1 given $X_2 = x_2$ is of the form,

$$f(x_1|x_2) = [(1+\theta x_1)(1+\theta x_2)-\theta] \exp[-x_1-\theta x_1 x_2],$$

then the distribution of X_1 is exponential if and only if that of X_2 is also exponential. An analogous result identifying the bivariate density in the discrete domain that characterizes the univariate geometric law is available in Nair and Nair (1990). In the present section our objective is to find out suitable conditional densities that guarantee unique Pareto and finite range densities for the component variates. The following results along with characterizations of a bivariate Pareto distribution has been reported in Hitha and Nair (1990).

Theorem 3.7.

Let (X_1,X_2) be a vector of non-negative random variables with joint density function $f(x_1,x_2)$ and marginal densities $f_1(x_1)$ and $f_2(x_2)$. If the conditional distribution of X_1 given $X_2=x_2$ is P II with density,

$$f(x_1|x_2) = (c+1)(d(x_2))^{c+1}\ (x_1+d(x_2))^{-(c+2)}, \qquad (3.31)$$

where

$$d(x_2) = a_1^{-1}\ (b+a_2x_2),\ x_1,x_2 > 0,\ a_1,a_2,b > 0,\ c \geqslant -1,$$

then the necessary and sufficient condition for X_1 to follow P II$(c, b/a_1)$ is that X_2 has the same type of distribution, P II$(c, b/a_2)$.

Proof:

When X_1 given $X_2=x_2$ is Pareto type II as in (3.31) and X_2 is P II$(c,a_2/b)$, the relationship

$$f_1(x_1) = \int_0^\infty f(x_1|x_2)\ f_2(x_2)dx_2, \qquad (3.32)$$

gives,

$$f_1(x_1) = \int_0^\infty (c+1)(b+a_2x_2)^{c+1}\, a_1^{-(c+1)}[x_1+a_1^{-1}(b+a_2x_2)]^{-(c+2)}\, a_2cb^c(a_2x_2+b)^{-(c+1)}dx_2,$$

$$= \int_0^\infty a_1a_2(c+1)(c+2)b^{c+1}(a_1x_1+a_2x_2+b)^{-(c+3)}\, dx_2,$$

$$= a_1cb^c(a_1x_1+b)^{-(c+1)},$$

which establishes the necessary part.

Conversely, assuming that X_1 is P II$(c,b/a_1)$ so that,

$$f_1(x_1) = a_1cb^c(a_1x_1+b)^{-(c+1)},$$

from equation (3.32)

$$a_1cb^c(a_1x_1+b)^{-(c+1)} = \int_0^\infty \frac{a_1(c+1)(a_2x_2+b)^{c+1}}{(a_1x_1+a_2x_2+b)^{c+2}} f_2(x_2)dx_2.$$

Taking the Mellin transform of both sides with respect to x_1,

$$\int_0^\infty \frac{cb^c\, x_1^{s-1}\, dx_1}{(a_1x_1+b)^{c+1}}\, dx_1 = \int_0^\infty\int_0^\infty \frac{(c+1)(a_2x_2+b)^{c+1}\, x_1^{s-1}}{(a_1x_1+a_2x_2+b)^{c+2}}\, f_2(x_2)\, dx_1\, dx_2.$$

Transforming x_1 to

$$y = a_1x_1(a_1x_1+a_2x_2+b)^{-1}$$

on the right side and

$$z = a_1x_1(a_1x_1+b)^{-1}$$

on the left side leaves the equation,

$$\int_0^\infty (a_2x_2+b)^{s-1} f_2(x_2)dx_2 = cb^{s-1}(c-s+1)^{-1}, \; c>s-1,$$

which is equivalent to

$$a_2^{-1} \int_0^\infty F(y) \; y^{s-1} \; dy = cb^{s-1}(c-s+1)^{-1},$$

where

$$F(y) = H(y-b) \; f_2(\frac{y-b}{a_2}),$$

and H(.) is the Heaviside unit function. Proceeding to the inverse Mellin transform,

$$f_2(x_2) = a_2cb^c(a_2x_2+b)^{-(c+1)}, \; x_2 \geq 0.$$

Theorem 3.7 provides a characterization of the distribution of the random vector $X=(X_1,X_2)$ as well, in terms of the marginal and conditional distributions of the same component X_1(or X_2). To show this, we assume that the distribution of X_1 given $X_2=x_2$ is as in equation (3.31) and further X_1 is P II$(c,b/a_1)$. By Theorem 3.7 X_2 is P II$(c,b/a_2)$ and therefore the joint distribution is computed as,

$$f(x_1,x_2) = a_1a_2(c+1)(c+2)b^{c+1}\,(a_1x_1+a_2x_2+b)^{-(c+3)},$$
$$x_1,x_2 > 0,\ a_1,a_2,\ b > 0,\ c \geq -1,$$

which is the bivariate distribution of Lindley and Singpurwalla (1986).

Theorem 3.8.

If (X_1,X_2) is distributed such that X_1 given $X_2=x_2$ has finite range distribution

$$f(x_1|x_2) = (r-1)a_1(1-a_2x_2)^{-1}[1-a_1x_1/(1-a_2x_2)]^{r-2}, \tag{3.33}$$
$$0<x_1<a_1(1-a_2x_2)^{-1},\ 0<x_2<a_2^{-1},$$
$$a_1,a_2 > 0,\ r > 2,$$

then X_1 has finite range distribution with density,

$$f_1(x_1) = ra_1(1-a_1x_1)^{r-1},\ 0 < x_1 < a_1^{-1}, \tag{3.34}$$

if and only if X_2 is likewise distributed with density,

$$f_2(x_2) = ra_2(1-a_2x_2)^{r-1},\ 0<x_2<a_2^{-1}. \tag{3.35}$$

Proof:

The if part follows from direct calculations as shown below. If X_2 has finite range distribution FR(a_2,r) as in (3.35) then

$$f(x_1,x_2) = r(r-1)\, a_1a_2(1-a_1x_1-a_2x_2)^{r-2}, \qquad (3.36)$$
$$a_1,a_2 > 0,\ a_1x_1 + a_2x_2 \leq 1,\ r > 2.$$

Integrating out x_2, we recover (3.34).

On the other hand, assuming that X_1 has density (3.34), equation (3.32) gives

$$ra_1(1-a_1x_1)^{r-1} = \int_{X_2} f(x_1|X_2=x_2)f_2(x_2)dx_2.$$

That is,

$$r(1-a_1x_1)^{r-1} = \int_0^{(1-a_1x_1)a_2^{-1}} (r-1)(1-a_2x_2)^{-1} [1-a_1x_1/(1-a_2x_2)]^{r-2}\, dx_2,$$

or

$$ra_2y^{r-1} = \int_0^y (r-1)(y-x_2)^{r-2}\, g(x_2)dx_2, \qquad (3.37)$$

where

$$y = a_2^{-1}(1-a_1x_1),$$

and

$$g(x_2) = f_2(x_2)\ (1-a_2x_2)^{1-r}.$$

Equation (3.37) can be written as

$$(r/r-1)\ a_2 y^{r-1}/\Gamma(r-1) \quad = \quad I_g^{r-1},$$

with I_g^{r-1} standing for the Riemann-Liouville fractional integral of order r-1 of the function g(.) (Erdélyi et. al. (1954) p.181) defined as

$$I_g^r \quad = \quad \frac{1}{\Gamma r} \int_0^y (y-x_2)^{r-1}\ g(x_2)dx_2 \quad .$$

The operator I_g^r is connected with differential and integral operators as follows.

$$\frac{d}{dx}\ I^r_{g(x)} \quad = \quad I^{r-1}_{g(x)} \text{ and } \frac{d}{dx}\ I^2_{g(x)} \quad = \quad \int_0^y\ g(t)dt \ .$$

Thus we get

$$\frac{d}{dx_2}\ I^{r-1}_{g(x_2)} \quad = \quad r(r-1)a_2 y^{r-2}/(r-1)\ \Gamma(r-1),$$

$$= \quad r(r-1)\ a_2\ y^{\ r-2}/(r-1)!$$

and

$$\frac{d}{dx_2}\ I^2_{g(x_2)} \quad = \quad \frac{r(r-1)\ldots 2}{(r-1)!}\ a_2 y \quad = \quad \int_0^y\ g(t)dt \ .$$

The unique inverse relation is, therefore,

$$g(x_2) \qquad = \quad ra_2,$$

and thus,

$$f_2(x_2) = ra_2(1-a_2x_2)^{r-1}.$$

This completes the proof of the theorem.

A characterization of the bivariate finite range distribution specified by the density (3.36), with the forms of X_1 and X_1 given $X_2=x_2$ following the finite range distribution is also evident.

3.3 Characterization By Properties Of Partial Moments.

We recall from equation (2.8) that the r^{th} partial moment of a random variable X about a point t is defined as

$$p_r(t) = E[(X-t)^+]^r, \; r = 0,1,2,\ldots ,$$

where, $(X-t)^+ = \max(0, X-t)$.

The properties of partial moments can be used to characterize probability distributions in the same way as truncated moments are employed by many authors in characterizing distributions like the exponential. Chong (1977) has characterized the exponential distribution by the property

$$E[X-t-s]^+ \, E(X) = E(X-t)^+ \, E(X-s)^+$$

of the partial means. In a recent paper 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.

The random variable $(X-t)^+$ used in defining the partial moments are meaningful in the study of personal incomes. If X represents the income of an individual and t the tax-exemption level, $(X-t)^+$ represents the taxable income. Those incomes which fall short of t is of no consequence in the computation of taxes and therefore is as good as treated to be zero. Thus the study of partial moments is useful in analysing the incomes that exceeds the exempt level without truncating the distribution at t.

In the following theorem Hitha (1990) shows that the Pareto distribution is characterized by the property that any partial moment at a given point is proportional to the product of partial moments at two other points which are factors of the original point.

Theorem 3.8.

Let X be a non-negative random variable in the support of $[k, \infty)$, $k > 0$, having absolutely continuous distribution with respect to Lebesgue measure and with

$E(X^{-r}) < \infty$. Then the partial moments satisfy the relation,

$$p_r(t)\ p_r(s) = A(r)\ p_r(ts), \tag{3.38}$$

for all $t,s \geq 1$, $r = 0,1,2,\ldots$, $A(r) > 0$ if and only if X follows the P I(a,k), where A(r) satisfies the equation,

$$a = r[1+ A(r-1)/A(r)],\ r \geq 1,\ a > r\ . \tag{3.39}$$

Proof:

If X is distributed as P I(a,k),

$$p_r(t) = r!\ k^a t^r/t^a (a-1)^{(r)},\ r = 0,1,2,\ldots,$$

so that

$$p_r(t)\ p_r(s) = \frac{r!\ k^a t^r}{(a-1)^{(r)} t^a}\ \frac{r!\ k^a s^r}{(a-1)^{(r)}\ s^a}\ .$$

Now setting $t = s = 1$ in (3.38),

$$A(r) = p_r(1) = \frac{r!k^a}{(a-1)^{(r)}}\ ,$$

and hence,

$$\begin{aligned} p_r(t)\ p_r(s) &= p_r(ts)\ r!k^a/(a-1)^{(r)}, \\ &= A(r)\ p_r(ts). \end{aligned}$$

Conversely, with $p_r(1) > 0$, (3.38) reduces to the functional equation,

$$G(r,t)\, G(r,s) = G(r,ts), \quad t,s \geqslant 1,$$

where

$$G(r,t) = p_r(t) \,/\, p_r(1).$$

Considering the transformations $u = \log t$ and $v = \log s$ and writing $G(r,e^u) = g(u)$ and $G(r,e^v) = g(v)$, we get the Cauchy functional equation,

$$g(u+v) = g(u)\, g(v), \quad u,v \geqslant 0,$$

whose solution is

$$g(v) = \lambda(r)\, e^{\alpha(r)v}.$$

It follows that

$$G(r,t) = \lambda(r)\, t^{\alpha(r)}.$$

From $G(r,1) = 1$, we have $\lambda(r) = 1$ and hence,

$$p_r(t) = p_r(1)\, t^{\alpha(r)} . \tag{3.40}$$

By differentiating the integral form of $p_r(t)$, viz.,

$$p_r(t) = \int_t^{\infty} (x-t)^r \, dF, \tag{3.41}$$

$$p_r'(t) = -r\, p_{r-1}(t), \quad r \geqslant 1 . \tag{3.42}$$

Equations (3.40) and (3.42) mean that

$$-r\, p_{r-1}(t) = \alpha(r)\, p_r(1)\, t^{\alpha(r)-1}, \tag{3.43}$$

and

$$-r\, p_{r-1}(1) = \alpha(r)\, p_r(1) . \tag{3.44}$$

Using (3.39) and $A(r) = p_r(1)$ the conclusions

$$(a-r)\, p_r(1) = r\, p_{r-1}(1), \tag{3.45}$$

and from (3.44),

$$\alpha(r) = -(a-r)$$

follows.

Further,

$$p_r(t) = p_r(1)\, t^{r-a},$$

and from the recurrence relation (3.45),

$$p_r(1) = ar!\, p_o(1)/a^{(r)}.$$

By definition,

$$p_o(1) = \int_1^\infty dF,$$

which is in fact a positive constant lying between 0 and 1 so that it can be written as k^a for some $k > 0$. Thus,

$$p_r(t) = \int_t^\infty (x-t)^r \, dF = r! \; k^a \; t^{r-a}/(a-1)^{(r)} . \qquad (3.46)$$

Differentiating (3.46) successively with respect to t, and using the relation (3.41) we arrive at,

$$1-F(t) = (k/t)^a, \; t \geq k, \; a > 0, \; k > 0,$$

which completes the proof.

Theorem 3.8 provides a series of results relating to other univariate families like exponential, power function, Burr, logistic, etc. in terms of the monotone transformations connecting them. Using the logarithmic transformation to the Pareto variable, the result corresponding to (3.38) for the exponential distribution is

$$E[(X-t-s)^+]^r \, E(X) = E[(X-t)^+]^r \, E[(X-s)^+]^r, \qquad (3.47)$$

$$r \geq 1, \; t,s > 1.$$

Setting r=1 in (3.47) we get the result due to Chong (1977) mentioned at the beginning of this section. There is a similar result that concerns the power distribution cited in Theorem 3.5 which is stated as follows.

Theorem 3.9.

Let X be a non-negative random variable in the support of (0,R), R > 0, having absolutely continuous distribution with respect to Lebesgue measure such that $EX^{-r} < \infty$. Then the partial moments

$$U_r(t) = E[(t-X)^+]^r,$$

satisfy the relation

$$U_r(t)\, U_r(s) = U_r(ts)\, B(r), \tag{3.48}$$

for all $0 < t, s < R$, $r = 0,1,2,\ldots$, $B(r) > 0$ if and only if X follows the power distribution with density,

$$f(x) = (\theta+1)\, x^{\theta} / R^{\theta+1},\ 0 < x < R, \tag{3.49}$$

where B(r) satisfies the relation

$$(r+1) - rB(r-1)/B(r) = \theta,\ r \geq 1. \tag{3.50}$$

Proof:

For the distribution (3.49),

$$U_r(t) = r!\ t^{\theta+r+1}/\ R^{\theta+1}(\theta+2)\ \ldots\ (\theta+r+1),$$

so that from this expression and $B(r) = U_r(1)$, the necessity of the condition follows. The proof of sufficiency part is along the lines of the proof of the previous Theorem and therefore only the important steps are presented. As before $G(r,t) = U_r(t)/U_r(1)$ and subsequently $G(r,e^x) = g(x)$ produces the Cauchy functional equation whose solution turns out to be

$$G(r,t) = t^{\alpha(r)} \quad \text{and} \quad \alpha(r) = \theta + r + 1.$$

Further,

$$U_r(1) = r!\ U_o(1)\ /\ (\theta+2)\ \ldots\ (\theta+r+1)$$

One can take $U_o(1)$ a positive constant to be $R^{\theta+1}$ for some $R > 0$ and this leads to

$$U_r(t) \;=\; r!\ R^{\theta+1}/(\theta+2)\ \ldots\ (\theta+r+1),$$

and **whence** the distribution (3.49).

3.4 Characterization By Truncated Reciprocal Moments

The r^{th} truncated reciprocal moment of a random variable X is given by equation (2.11) provided the expectation on the right side exists. The reciprocal moments exist only for those distributions which have a zero of order r at least at the origin. A characteristic property of the Pareto distribution associated with the reciprocal moments will now be proved.

Theorem 3.10.

The truncated reciprocal moments $C_r(t)$ of the random variable X given by equation (2.11) satisfy the relation

$$C_r(t)\, C_r(s) = M(r)\, C_r(ts), \tag{3.51}$$

for all $t,s \geq 1$, $r = 0,1,2,\ldots$, $C_r(1) > 0$, $M(r) > 0$, if and only if X has P I(a,k) distribution with

$$a = r[M(r-1)/M(r)-1],\ r \geq 1. \tag{3.52}$$

Proof:

If X is P I(a,k), direct calculation yields equation (2.12) and the relationship (3.51) is easily verified. Conversely, let (3.51) be true with the fact $C_r(1) = M(r)$. Proceeding exactly as in Theorem 3.8, we arrive at the equation

$$C_r(t) = M(r)\ t^{\alpha(r)}, \tag{3.53}$$

which is analogous to (3.40).

Now,

$$C_r(t) = (1/R(t))(-1)^r \int_t^\infty ((1/x)-(1/t))^r dF, \tag{3.54}$$

where

$$R(t) = P[X > t].$$

Thus,

$$C_r(t)\ R(t) = (-r)(-1)^r \int_t^\infty ((1/x)-(1/t))^{r-1} (1/x^2)(1-F)dx\ .$$

Differentiating the above equation with respect to t,

$$C_r'(t)\ R(t)+C_r(t)R'(t) = ((-r)/t^2)(r-1) \int_t^\infty ((1/x)-(1/t))^{r-2}(1/x^2)R(x)dx$$

$$= ((-r)/t^2)\,C_{r-1}(t)\ R(t)\ . \tag{3.55}$$

Thus we get

$$[(r/t^2)C_{r-1}(t) + C_r'(t)]/C_r(t) = (-R'(t)/R(t)),$$

and so the recurrence relation

$$[(r/t^2)C_{r-1}(t)+C_r'(t)]/C_r(t)$$
$$= [(r-1)/t^2)C_{r-2}(t) + C_{r-1}'(t)]/ C_{r-1}(t) \cdot$$

Using the relation (3.53) we will arrive at

$$\frac{(r/t^2)\ M(r-1)t^{\alpha(r-1)} + \alpha(r)\ M(r)\ t^{\alpha(r)-1}}{M(r)\ t^{\alpha(r)}}$$

$$= \frac{((r-1)/t^2)\ M(r-2)t^{\alpha(r-2)} + \alpha(r-1)\ M(r-1)\ t^{\alpha(r-1)-1}}{M(r-1)\ t^{\alpha(r-1)}} .$$

Putting t=1,

$$\alpha(r) + rM(r-1)/M(r) = \alpha(r-1)+(r-1)\ M(r-2)/M(r-1). \quad (3.56)$$

Further we have (3.52) to obtain,

$$rM(r-1)/M(r) = a+r ,$$

so that (3.56) will become,

$$\alpha(r) + a + r = \alpha(r-1) + a+r-1.$$

That is,

$$\alpha(r) = \alpha(r-1)-1,$$

or

$$\alpha(r) = -r, \text{ since } \alpha(o) = 1.$$

Thus,

$$C_r(t) = C_r(1)t^{-r}.$$

Now,

$$C_r(1) = (-1)^r a.\ r!\ C_o(1)/\ a^{[r]},$$

where

$$C_o(t) = (1/R(t)) \int_t^\infty dF = 1,$$

for all t.

Thus,

$$C_r(t) = (-1)^r\ ar!\ /(a^{[r]}\ t^r).$$

Further,

$$r!R(t)a(-1)^r/(a^{[r]}t^r) = (-1)^r \int_t^\infty ((1/x)-(1/t))^r\ dF,$$

$$= (-1)^r\ (-r) \int_t^\infty ((1/x)-(1/t))^{r-1}(1/x^2)dF.$$

Differentiating with respect to t, we get

$$(-1)^r\ a.r!/a^{[r]}\left\{\frac{R'(t)}{t^r} - \frac{rR(t)}{t^{r+1}}\right\}$$

$$= (-1)^{r+1}\ \frac{r(r-1)}{t^2} \int_t^\infty ((1/x)-(1/t))^{r-2}\ \frac{R(x)}{x^2}\ dx,$$

$$= \frac{(-r/t^2)\ R(t)(-1)^{r-1} a.(r-1)!}{a^{[r-1]}\ t^{r-1}} \quad .$$

That is,

$$[R'(t) - (r/t)\ R(t)]\ /\ (a+r) = -\ R(t)/t,$$

or

$$R'(t)/R(t) \quad = \quad -a/t.$$

Hence,

$$F(t) = 1-(k/t)^a, \ t \geq k, \ a > 0, \ k > 0,$$

and our result is completely proved.

3.5 Characterization By Additive Damage Model

The concept of damage models introduced by Rao and Rubin (1964) involves a random variable X reduced to another random variable U by some random mechanism represented by the conditional distribution of U given X = x. The quantity Y = X-U is the reduction in X and is called the damaged component in X and the objective in the formulation of such models, is to characterize the distribution of X in terms of the distribution of U. Instead of the additive model one can also have a multiplicative model of the form U=XY. A comprehensive survey of the various results in damage models is available in Patil and Retnaparki (1975) and also in Galambos and Kotz (1977) where the connection between such models and those arrived at by geometric compounding and rarefactions, is also explained.

In the present section, a result due to Revankar et. al. (1974) that characterizes the P II(a,α) model will be extended to cover the finite range and exponential variables.

Theorem 3.11.

If,

$$E(U|X=x) = a+bx, \qquad (3.57)$$

a necessary and sufficient condition that X has either an exponential distribution with $\beta = b>0$ and $\alpha > a > 0$ or a Pareto type II distribution with $\beta>b>0$ and $a>\alpha$ or a finite range distribution with $\beta<b<0$ and $a>\alpha$ is that

$$E(U|X>y) = \alpha+\beta y. \qquad (3.58)$$

Proof:

The case when $\beta>b>0$ for the Pareto distribution is proved in Revankar et.al. (1974).

Assuming X to be FR(c,R),

$$E(U|X>y) = (1/R(y)) \int_y^R (a-bx)(c/R)\ (1-x/R)^{c-1}\, dx,$$

$$= (a-bR/(c+1)) - bcy/(c+1),$$

$$= \alpha + \beta y,$$

where $\beta = bc/(c+1)$ and $a > \alpha$ and $b<\beta<0$.

Conversely, with

$$E(U|X>y) = \alpha+\beta y, \; b < \beta < 0,$$

$$R(y)\;(\alpha+\beta y) = -\int_y^R (a-bx)\, dR(x),$$

$$= (a-by)\, R(y) - b \int_y^R R(x)dx.$$

Differentiating both sides with respect to y,

$$R'(y)/R(y) = -\beta/[(\alpha-a) + (\beta+b)y] \cdot$$

The solution is

$$R(y) = K[(\alpha-a) + (\beta+b)y]^{-\beta/(\beta+b)} \; .$$

Evaluating K, using the condition $R(o)=1$,

$$K = (\alpha-a)^{\beta/(\beta+b)}.$$

Thus,

$$R(y) = (1 - y/R)^{c},$$

where

$$R = (b+\beta)/(a-\alpha) \text{ and } c = -\beta/(\beta+b) > 0,$$

and Y is FR(c,R). The proof of the exponential case is trivial and our theorem is proved.

3.6 Residual Life Distributions

In section 2.5 we have examined the various characterizations of the Pareto and related distributions by properties of residual life time. These properties will appear in a natural way once we look at the entire distribution of the residual life time that is being currently discussed. The distribution function of the residual life Y_x of a non-negative random variable X with F(o) = 0, where F(x) is the distribution function of X, is defined as (Arnold, 1983),

$$G(y;x) = \frac{F(x+y)-F(x)}{1-F(x)}, \quad y > 0 \,. \tag{3.59}$$

The corresponding survival function is

$$S(y;x) = R(x+y)/R(x) \,. \tag{3.60}$$

It is easy to see that the mean of (3.60) is the MRL function r(x) defined in equation (2.20), for,

$$\begin{aligned} E(Y_x) &= \int_0^\infty y\left(\frac{-\partial}{\partial y} S(y;x)\right)dy, \\ &= (1/R(x)) \int_0^\infty y\left(-\frac{\partial}{\partial y} R(x+y)\right)dy, \\ &= (1/R(x)) \int_x^\infty (z-x)\, dF(z) \\ &= E[X-x|X>x], \end{aligned}$$

provided that $E[X] < \infty$.

The first problem we investigate is the form of the random variable Y_x when X belongs to the class of models under consideration in this chapter. The answer is provided in the following

Theorem 3.12.

The random variable X follows

(a) E(b) if and only if Y_x is E(b)

(b) P I(a,k) if and only if Y_x is P I(a,k) with the origin shifted from k to k+x.

(c) P II(c,α) if and only if Y_x is P II(c,x+α)

(d) FR(d,R) if and only if Y_x is FR(d, R−x) .

Proof:

When Y_x is Pareto II (c, x+α) it follows from (2.26) that

$$\frac{R(x+y)}{R(x)} = \left(\frac{x+\alpha}{x+y+\alpha}\right)^c .$$

As x tends to zero, $R(y) = \left(\frac{\alpha}{y+\alpha}\right)^c$ and X is P II(c,α). The if part is verified through direct calculations of S(y;x) using (3.60). Proof for the other distributions follow suit.

As an alternative to the mean residual life function the median residual life has also been considered in literature, which stands for the time expected for half of the number of items that operate at time x fail. The new measure enjoys relative superiority over the MRL in situations where the latter (a) becomes unstable in the presence of outliers (b) does not exist (but the median is always finite) (c) is less desirable for fat tailed distributions such as those considered here, and (d) the data is in the form of censored observations with at least half of those remaining have recorded failured times. Moreover, it has simple closed form expressions for many useful failure time models while the Mean Residual life has too complicated a functional form to be of use. We cite for example, the Weibull case where the median residual life is

$$(b^{-1} \log 2 + x^{c})^{1/c} - x,$$

corresponding to the survival function $R(x) = e^{-bx^{c}}$, in contrast with the MRL which at best can be written only in terms of incomplete gamma function and is analytically intractable.

With respect to the residual life distribution the median residual life function is the solution for y of the equation

$$P(Y_x > y) = \frac{1}{2},$$

or (3.61)

$$S(y;x) = R(x+y)/R(x) = \frac{1}{2} ,$$

which is in general a function of x to be denoted by M(x). Thus M(x) satisfies the functional equation

$$R(x+M(x)) = \frac{1}{2} R(x).$$

When x tends to zero, the last expression gives,

$$R(M(o)) = \frac{1}{2} ,$$

so that M(o) = M becomes the median of the random variable X. Then the median residual life is given by

$$R(x+M(x)) = R(x) R(M). \tag{3.62}$$

Theorem 3.13.

Let R(x) be an absolutely continuous survival function with R(o) = 1 then the residual life distribution S(y;x) satisfies the equation,

$$S(g(x)y;x) = R(y), \tag{3.63}$$

where $g(x) = M(x)/M(o)$,

if and only if X is distributed as either exponential or Pareto II or finite range.

Proof:

Lemma:

As a first step we show that the median residual life function is of the form $\ell t+m$, $m>0$ if and only if X is E(b) for $\ell=0$, P II(c,α) for $\ell>0$ and FR(d,R) for $\ell<0$.

Proof:

By solving the equation (3.62) we see that

$$M(x) = \begin{cases} b^{-1}\log 2, & \text{for } E(b), \\ (2^{1/c}-1)(x+\alpha), & \text{for } P\ II(c,\alpha), \\ [1-(1/2)^{1/d}](R-x), & \text{for } FR(d,R), \end{cases}$$

satisfy the conditions of the Theorem. Conversely, when $\ell=0$, M(x) = a constant so that (3.62) becomes the Cauchy functional equation,

$$R(x+M) = R(x)\,R(M),$$

whose only continuous solution that satisfy the probability requirements for R(x) is

$$R(x) = e^{-bx},$$

where $R(M) = e^{-bM} = 1/2$. Then the result is true for the exponential distribution. The Pareto II case is proved in

Schmittlein and Morrison (1981) and the finite range situation follows from the same proof.

In establishing the main result we first assume that (3.62) is true. Then with the help of (3.61) and (3.63) the equation

$$R(g(x)y+x) = R(x)\, R(y) \qquad (3.64)$$

can be reached. Differentiating (3.64) with respect to x,

$$R'[g(x)y+x]\,[g'(x)y+1] = R'(x)\, R(y),$$

and the same operation on (3.64) with respect to y yields,

$$R'[g(x)y+x]g(x) = R(x)\, R'(y).$$

Hence,

$$(g'(x)y+1)/g(x) = R'(x)R(y)/R(x)\, R'(y),$$

and,

$$1/g(x) = R'(x)/R(x)\, R'(y).$$

Combining the last two equations,

$$g'(x)y+1 = R'(o)\, R(y)/R'(y).$$

The rightside being independent of x, so should be the left side also which implies $g'(x)$ is a constant or $g(x)= \ell x+m$.

From the definition of $g(x)$ this reduces to the linearity of $M(x)$ and therefore by lemma, X has one of the distributions stated in the Theorem. In order to establish the converse, we observe that in the exponential case $g(x)=1$ so that relation (3.64) holds. For P II(c,α) we replace y by $yg(x) = y.(x/\alpha)/\alpha$ in

$$S(y;x) = [(x+\alpha)/(x+y+\alpha)]^c,$$

to get

$$S(yg(x);x) = (\alpha/(\alpha+y))^c = S(y).$$

The result for FR(d,R) is established similarly and our Theorem stands proved.

The concept of median of residual life extends itself to the notion of percentile residual life if one wishes to have a finer set of summary measures of location. Such measures are used for inference in reliability studies by Joe and Proschan (1983). From the point of view of characterizations also they are quite handy and produces conclusions similar to that of the median.

The q^{th} percentile residual life time is according to Haines and Singpurwalla (1974) is

$M_q(x) = S^{-1}[qS(x)]-x, \; 0 < q < 1.$

For the family of distributions under investigation,

$$M_q(x) = \begin{cases} (1/b)\log q, & \text{for } E(b)\,, \\ (x+\alpha)\,(q^{1/c}-1), & \text{for } P\,II(c,\alpha)\,, \\ [1-(1/q)^{1/d}]\,[R-x], & \text{for } FR(d,R). \end{cases}$$

Thus the form of the percentile residual life for the three distributions remains identical with that of the median residual life. Therefore the conclusions of the last theorem can be extended to involve the percentile residual life.

Chapter IV

CHARACTERIZATION OF DISCRETE MODELS

4.1 Preliminaries

The aim of the present chapter is to extend to the discrete sample space some of the results that were established earlier to the Pareto, finite range and exponential models. As mentioned in the introduction, the distributions in this context centres around the three discrete distributions, Waring, negative hypergeometric and geometric which exhibit properties analogous to the continuous models just mentioned. Using the notations of section 2.4, we shall denote by W(a,b), the Waring distribution with probability function

$$P[X=x] = (a-b)(b)_x/(a)_{x+1}, \quad x = 0,1,2,\ldots, \quad a > b > 0, \tag{4.1}$$

while the symbol NH(k,n) is reserved for the negative hypergeometric law specified by,

$$P[X=x] = \binom{-1}{x}\binom{-k}{n-x} \Big/ \binom{-1-k}{n}, \quad x=0,1,2,\ldots,n, \quad k > 0, \tag{4.2}$$

and G(p) for the normal geometric model

$$P[X=x] = q^x p, \quad x = 0,1,2,\ldots, \quad 0<p<1, \quad p+q=1. \tag{4.3}$$

Notice that the discrete uniform distribution in the support of (0,1,...,n) appears as a special case of (4.2) when k=1.

4.2 Continuous Approximations

As a prelude to the explorations of characteristic properties of models (4.1), (4.2) and (4.3) in the rest of the chapter, a justification for alluding to them properties similar to the exponential, Pareto and finite range laws, as in order. One way of accomplishing this is by calculating the continuous approximation of the discrete distributions and thereby conclude that they lead to the desired continuous counter parts. The slope-ordinate ratio method propounded by Irwin (1975c) will be involved to realise our objective. Roughly speaking, the method consists in equating the ratio $[f_r-f_{r-1}]/\frac{1}{2}[f_r+f_{r-1}]$ to the logarithmic derivative of f(x) evaluated at $x = r - \frac{1}{2}$ and then solve the resulting differential equation. Notice that f_r is the frequency for r=0,1,... and f(x) is the corresponding continuous density. For the Waring distribution W(a,b) of equation (4.1)

$$f_r = N(a-b)(b)_r/(a)_{r+1}, \quad r=0,1,2,...,$$

so that

$$f_r - f_{r-1} = N(a-b)\ [\ \frac{(b)_r}{(a)_{r+1}} - \frac{(b)_{r-1}}{(a)_r}\ ]\ ,$$

$$= N(a-b)(b-a-1)\ (b)_{r-1}/(a)_{r+1},$$

and

$$f_r + f_{r-1} = N(a-b)\ [\ \frac{(b)_r}{(a)_{r+1}} + \frac{(b)_{r-1}}{(a)_r}\ ],$$

$$= N(a-b)\ (a+b+2r-1)(b)_{r-1}/(a)_{r+1}.$$

Hence ,

$$\frac{f_r - f_{r-1}}{\frac{1}{2}[f_r + f_{r-1}]} = \frac{2(b-a-1)}{(a+b+2r-1)}\ ,$$

and this is called the slope-ordinate ratio at $x = r - \frac{1}{2}$.

We now write

$$\left(\frac{d}{dx} \log y \right)_{x=r-\frac{1}{2}} = \frac{2(b-a-1)}{(a+b+2r-1)}\ ,$$

where $y = f(x)$.

Setting $z = x + \frac{1}{2}$, we have $z=r$ whenever $x = r-\frac{1}{2}$ and we further deduce

$$\frac{d}{dz} \log y = (b-a-1)\ (\frac{a-b-1}{2} + z),\ a>b.$$

The solution of the last differential equation is of the form

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

for some α, $c > 0$, leading to P II(c,α).

Now consider the negative hypergeometric distribution with

$$f_r = N \binom{-1}{r} \binom{-k}{n-r} \Big/ \binom{-1-k}{n},$$

$$= N \binom{k+n-r-1}{n-r} \Big/ \binom{k+n}{n},$$

giving

$$f_r - f_{r-1} = N \binom{k+n-r-2}{n-r-1} \frac{k-1}{n-r} \Big/ \binom{k+n}{n},$$

and

$$\tfrac{1}{2}[f_r + f_{r-1}] = N \binom{k+n-r-2}{n-r-1} \frac{(\frac{k-1}{2})+n-r}{n-r} \Big/ \binom{k+n}{n} .$$

Thus

$$\frac{d}{dz} \log y = (k-1) \Big/ ((\tfrac{k-1}{2})+n-z)$$

or

$$f(z) = \text{const } (1- z/R)^{d-1} ,$$

where d=K and $R = \frac{k}{2} + n - \frac{1}{2}$ are both positive. This shows that the finite range model is the continuous approximation of the negative hypergeometric law.

4.3 Failure Rate and MRL

The distributions that are to follow requires elaborate use of the concepts of failure rates and MRL's as applied to discrete random variables. These concepts have been touched upon in many works such as Kalbfech and Prentice (1980). Many interrelationships and identities in this connection will be investigated now.

Let X be a discrete random variable in the support of $I^+ = (0,1,2,...)$ with probability mass function f(x). We also define

$$R(x) = P[X \geqslant x],$$

so that

$$f(x) = P[X=x] = R(x)-R(x+1) \cdot \qquad (4.4)$$

The failure rate of X is

$$h(x) = f(x)/R(x), \qquad (4.5)$$

and the MRL is

$$r(x) = E[X-x|X>x],$$

$$= \frac{1}{R(x+1)} \sum_{x+1}^{\infty} (y-x)\ f(y).$$

This gives

$$r(x)\ R(x+1) = \sum_{x+1}^{\infty} R(y), \qquad (4.6)$$

and the recurrence relation

$$r(x)\ R(x+1) - r(x+1)R(x+2) = R(x+1),$$

or

$$r(x)\ R(x+1) = [r(x-1)-1]\ R(x),\ x \geq 1. \qquad (4.7)$$

From the equation (4.5) and (4.4)

$$h(x) = [R(x)-R(x+1)]\ /\ R(x),$$

$$= 1-R(x+1)\ /\ R(x).$$

Hence ,

$$1-h(x) = \frac{f(x+1)/h(x+1)}{f(x)/\ h(x)},$$

and

$$f(x+1) = f(x)\ h(x+1)\ (1-h(x))/h(x).$$

Thus we get by iteration on x

$$f(x) = h(x)\ (1-h(x-1))\ \ldots\ (1-h(o)), \qquad (4.8)$$

and

$$R(x) = \prod_{y=o}^{x-1} (1-h(y)). \qquad (4.9)$$

Equations (4.8) and (4.9) show that h(x) determines the distribution of X uniquely.

Combining (4.7) and (4.9) we get the relationship between the failure rate and MRL of X as

$$1-h(x+1) = (r(x)-1)/r(x+1),\ x \geq 0. \qquad (4.10)$$

It follows that MRL function also determines the distribution uniquely through

$$R(x) = \prod_{u=1}^{x-1} \frac{r(u-1)-1}{r(u)}\ (1-f(o)). \qquad (4.11)$$

These interrelationships are useful in lifelength studies when time is treated as discrete as explained in chapter V but their immediate application is restricted to characterizing probability distributions.

4.4 Characterizations By Distribution Based on Partial Sums*

If X is a random variable defined in the previous section with $E(X) = m < \infty$, the variate Y specified by

$$g(y) = P[Y=y] = m^{-1} P[X>y],\ y=0,1,\ldots, \qquad (4.12)$$

is said to have the distribution based on partial sums corresponding to X. The probabilities assumed by the values of Y are proportional to the survival probabilities of X. Some properties of (4.12) are discussed in Johnson and Kotz (1969, p.261). Under certain conditions Gupta (1979) describes Y as the residual life time of a component, in a system where a component of life length X is replaced upon failure by another, having the same life distribution, so that the sequence of life lengths forms a renewal process. He showed that the failure rate of Y is the reciprocal of the MRL of X and that when the renewal distribution belongs to the class of modified power series distribution, the geometric law is the only one satisfying the property E(X)=E(Y). In this section we supplement Gupta's results by extending some of his results to cover the class of discrete distributions under consideration and also explore the possibility of arriving at some new characterizations.

---

* These results have been published in Prob.Statist.Letters (reference 65).

4.4.1 Basic Results

Analogous to equation (4.5), we write the failure rate of Y as

$$k(s) = P[Y=s] \,/\, P[Y>s],$$

$$= R(s+1) \,/\, \sum_{s+1}^{\infty} R(u),$$

$$= [r(s)]^{-1}, \quad s>t \geq 0, \tag{4.13}$$

where r(s) is as in equation (4.6), the MRL of X. Result (4.13) is obtained by Gupta (1979). Further,

$$k(s) = \sum_{r=1}^{\infty} P[Y=s+r] \,/\, \sum_{s+1}^{\infty} R(u),$$

$$= \sum_{r=1}^{\infty} R(s+r)\, h(s+r) \,/\, \sum_{s+1}^{\infty} R(u),$$

$$= \sum_{s+1}^{\infty} h(t) \quad w(t), \tag{4.14}$$

with

$$w(t) = R(t) \,/\, \sum_{s+1}^{\infty} R(s), \; t = s+1, s+2, \ldots .$$

Thus the failure rate of Y is a weighted average of the failure rates of X beyond the time point s. From (4.14), for $G(u) = P[Y \geqslant u]$,

$$a(s) = \sum_{s+1}^{\infty} G(t)/G(s+1), \tag{4.15}$$

so that together with (4.12) giving

$$G(s) = m^{-1} \sum_{s+1}^{\infty} R(t),$$

we can write,

$$a(s)m^{-1} \sum_{s+2}^{\infty} R(t) = \sum_{s+1}^{\infty} G(t),$$

$$= m^{-1} \sum_{r=1}^{\infty} \sum_{s+r}^{\infty} R(t).$$

Substituting

$$r(s)\, R(s+1) = \sum_{s+1}^{\infty} R(t),$$

and (4.16)

$$a(s) = \sum_{s+1}^{\infty} r(t)\, w(t+1),$$

the terms in (4.16) can be simplified to

$$w(t+u+1)\, r(u+t) = \begin{cases} 1 & , u=1, \\ \prod_{j=1}^{u} [1-k(t+1+j)], & u=2,3,\dots . \end{cases}$$

Thus we see that, as in the case of failure rate, the MRL of Y is also a weighted average of the MRL's of X.

Now with the aid of (4.13), equation (4.10) becomes,

$$1-h(s+1) = [(k(s))^{-1}-1]\, k(s+1),$$

or (4.17)

$$h(s) = 1+k(s)[1-(k(s-1))^{-1}],\quad s \geqslant 1,$$

giving the relationship between failure rate of X and Y. On the otherhand, the MRL's of X and Y satisfy

$$r(s+1) = a(s+1)\,[1+a(s+1)-a(s)]^{-1},$$

which follows from (4.17), (4.13) and (4.10). We can also have the equation

$$k(s) = [1+a(s)-a(s-1)]/a(s),$$

connecting the failure rate of Y and the MRL of X.

4.4.2 Characterizations

An immediate consequence of the identities developed in the previous section is in characterizing some discrete distributions. We first prove

Theorem 4.1.

A necessary and sufficient condition for X to be distributed as G(p) (W(a,b); NH(k,n)) is that Y is G(p) (W(a,b+1); NH(k+1, n-1)).

Proof:

From Xekalaki (1983a), it is seen that a failure rate function of the form $h(x) = (L+Mx)^{-1}$ characterizes

G(p) for $M = 0$, $L = p^{-1}$;

W(a,b) for $M = (a-b)^{-1}$, $L = a(a-b)^{-1}$;

and

NH(k,n) for $M = -k^{-1}$, $L = k^{-1}(n+k)$.

Substituting this form of h(x) in (4.17), we find,

$$1+k(x)\ [1-(k(x-1))^{-1}] = (L+Mx)^{-1}.$$

Since, $k(x) = (r(x))^{-1}$, this should read

$$\frac{1-r(x-1)}{r(x)} = (L+Mx)^{-1} - 1.$$

Applying the above result recurssively for ascending values of x,

$$r(x-1) = \sum_{\gamma=0}^{\infty} (\alpha+x)_{(\gamma)} \ / \ (\beta+x)_{(\gamma)}, \qquad (4.18)$$

with

$\alpha = (L-1)/M$ and $\beta = L/M$.

Obviously the last equation represents the hypergeometric series $2F_1(\alpha+x,\ 1,\ \beta+x,\ 1)$ in the usual notations (see Abramowitz and Stegun, 1972). From the well known formula,

$$2F_1(p,q,\gamma,1) = \frac{\Gamma\gamma \ \Gamma(\gamma-p-q)}{\Gamma(\gamma-p) \ \Gamma(\gamma-q)},$$

it is easy to see that

$$2F_1(p,1,\gamma,1) = (\gamma-1)/(\gamma-p-1). \qquad (4.19)$$

Thus

$$r(x-1) = (\beta+x-1) \ / \ (\beta-\alpha-1), \qquad (4.20)$$

or

$$k(x) = (A+Bx)^{-1},$$

where

$$A = L(1-M)^{-1} \text{ and } B = M(1-M)^{-1}. \qquad (4.21)$$

The form of $k(x)$ suggests that the distribution of Y is of the same type as that of X and further (4.20) gives the parameters of each model after inserting the values of L and M for each stated at the beginning of the proof.

Note:

The fact that the above models are closed with respect to the formation of the distribution of Y can also be established by direct calculation using the equation (4.12). Our motivation in using the above method are (i) direct calculation are more extensive, (ii) the form of the failure rate and MRL is handy when the modelling is based on these concepts, and (iii) the relationships used here are important in their own right in other areas of applications as will be shown in chapter V.

Corollary 4.1.

The MRL of X is of the form A+Bx if and only if X is either G(p) for A=0 or W(a,b) for A > 0 or NH(k,n) for A < 0.

Proof:

When X has a distribution that follows one of the models stated in the theorem if and only if the failure rate is of the form $(L+Mx)^{-1}$. In this situation, from equation (4.20) we see that

$$r(x) = (\beta+x) / (\beta-\alpha-1) , \qquad (4.22)$$

and therefore, the MRL is of the required form. Conversely, if MRL is as in (4.22), the failure rate is

$$h(x) = 1 - \left(\frac{r(x-1)-1}{r(x)} \right)$$

$$= (\beta-\alpha) \;/\; (\beta+x) = (L+Mx)^{-1}$$

Accordingly, by the characterization theorem of Xekalaki (1983a), X has the above specified distributional form.

To enable future reference, we note Table 4.1 for the actual expressions for the failure rates and MRLs of the various models.

Table 4.1.

Failure rates and MRL's of Discrete Models

Model	Failure Rate	MRL
$G(p)$	p	p^{-1}
$W(a,b)$	$(a-b)(a+x)^{-1}$	$(a+x)(a-b-1)^{-1}$
$NH(k,n)$	$k(k+n-x)^{-1}$	$(k+n-x)(k+1)^{-1}$
(Uniform in [o,n])	$(1+n-x)^{-1}$	$\frac{1}{2}\,(1+n-x)$

Theorem 4.2

The property $h(x) = C.k(x)$ for all integers $x \geq 0$ and a constant C, characterizes G(p) for C=1, W(a,b) for $C > 1$ and NH(k,n) for $0 < C < 1$.

Proof:

For the geometric law X and Y have identical distributions with the same parameters and therefore, in this case $h(x) = k(x)$. When X is W(a,b), Y is W(a,b+1) so that from table 4.1,

$$h(x)/k(x) = (a-b)/(a-b-1) > 1.$$

In the negative hypergeometric case

$$h(x)/k(x) = k/(k+1) < 1.$$

Conversely, $h(x) = Ck(x)$ is equivalent to $h(x)\ r(x) = C$ or to

$$r(x)-r(x-1)+1 = C.$$

The solution of this equation is

$$r(x) = (C-1)x + r(o),$$

which is of the form A+Bx and therefore by Corollary 4.1 our Theorem is proved.

Corollary 4.2.

The MRL and failure rate of X is such that $r(x)\ h(x) = C$ for all integers $x \geq 0$ and a constant $C > 0$ if and only if X is G(p) for C=1, W(a,b) for $C > 1$ and NH(k,n) for $0 < C < 1$.

This follows from Theorem 4.2 and the relationship $k(x) = [r(x)]^{-1}$.

Corollary 4.3.

The relationship $r(x) = Ka(x)$ is satisfied for all integers $x \geq 0$ and a constant K if and only if X is geometric (for K=1) or Waring (for K<1) or negative hypergeometric (for K>1).

Proof:

The result follows from the relationships between a(x) and k(x), equation (4.17) and Theorem 4.2.

The utility of these results and the physical interpretation of the properties enjoyed by the models in the last theorem in the context of ageing will be taken up in chapter V. The advantages arising out of these Theorems in the context of evaluating the memory of discrete distributions will now be explored.

A measure of memory at any point x in the support (0,1,2,...) of a discrete random variable X is defined in terms of MRL (Nair, 1983) as

$$m(x) = r(x-1) - r(x), \quad x \geqslant 1.$$

The distribution of X is said to have no memory, negative memory or positive memory according as m(x) is zero, negative or positive. Since a distribution can have different types of memory at the various points of its support, a consolidated measure of memory for the entire support was obtained as a weighted average of the measures at various points. The proposed measure was

$$M = \frac{2\,E^2(X) + E(X) - E(X^2)}{E(X^2) + E(X)}\,. \tag{4.23}$$

The distribution itself has lack of memory, negative memory or positive memory according as M is zero, negative or positive. The following Theorems follow from the above definition.

Theorem 4.3.

The geometric, **Waring** and negative hypergeometric laws in that order are the only discrete distributions that possess lack of memory, constant negative memory and constant positive memory at each point of its support.

(This deduction is easily seen from the fact that $r(x) = A+Bx$ for these models).

Theorem 4.4.

If X has a particular type of memory, at a given point, then Y also has the same type of memory.

Proof:

Suppose X has positive memory at the point x. This implies that

$$r(x-1) > r(x) \quad \text{or} \quad k(x) > k(x-1).$$

The last inequality, however, is equivalent to

$$g(x)\, G(x-1)-g(x-1)\, G(x) > 0,$$

or to

$$G(x)/G(x-1) \; > \; G(x+1)/G(x).$$

Hence $H(x) = G(x)/G(x-1)$ is decreasing in x.

Accordingly,

$$a(x) \; = \; H(x) + H(x)\, H(x+1)+H(x)\, H(x+1)\, H(x+2)+ \dots ,$$

and

$$a(x)-a(x+1) = [H(x)-H(x+1)]+H(x+1)[H(x)-H(x+2)]+ \dots$$

$$> 0,$$

which means that Y has positive memory at x. The proof is similar when X has lack of memory or has negative memory.

From the expression for M it is clear that, in general, M=0 does not imply that X is geometric (although for this law M=0). This means that we have to restrict the family of distributions to be considered, in order to characterize the geometric law by the property M=0. Suppose that X belongs to the modified power series family

$$P[X=x] = a(x)\ [g(\theta)]^x/f(\theta),\ x \in B,$$

where B is a subset of the set of non-negative integers, $a(x) > 0$, $g(\theta)$ and $f(\theta)$ are positive, finite and differentiable. Then from Gupta (1979), we have $E(Y) = E(X)$.

If we denote the generating function of $\langle f(x) \rangle$ and $\langle R(x) \rangle$ by

$$A(t) = \sum_x f(x)\ t^x \text{ and } B(t) = \sum_x R(x) t^x,$$

it is easy to see that

$$A'(1) = E(X) \text{ and } A''(1) = EX(X-1). \tag{4.24}$$

Further ,

$$(1-t)\,B(t) \;=\; 1-tA(t). \qquad (4.25)$$

Differentiating (4.25) twice, with respect to t and setting t=1

$$B'(1) \;=\; \sum_{x=1}^{\infty} xR(x),$$

and

$$2B'(1) \;=\; A''(1) + 2E(X),$$

$$= \; EX(X-1) + 2E(X).$$

Accordingly ,

$$E(X) \;=\; E(Y) \;=\; m^{-1}\,\Sigma\; yP[X>y],$$

$$= \; B'(1)-E(X).$$

Thus,

$$E(X) \;=\; \tfrac{1}{2}\, m^{-1}[E(X^2)-E(X)]\;,$$

$$= \; [E(X^2)-E(X)]/2E(X),$$

or

$$E(X^2) \;=\; 2E^2(X) + E(X),$$

and hence M=0. We have, therefore, established the following result.

Theorem 4.5.

Among the modified power series family, geometric law is the only one for which M=0.

4.5 Residual Life Distribution

Let X be a random variable representing the life time of a component or device in the support of the set of non-negative integers with survival function R(x). When the life times are expressed only in completed units of time, the domain of X will be restricted to the support just mentioned. In such a situation, the residual life distribution (RLD) of X at the elapse of x units of time is specified by the distribution function

$$F(y;x) = P[x < X \leq x+y | X > x],$$

$$= \frac{F(x+y) - F(x)}{1-F(x)},$$

$$= [R(x+1)-R(x+y+1)]/R(x+1).$$

The corresponding survival function is

$$R(y;x) = R(x+y+1)/R(x+1), \; y \geq 0 \cdot \qquad (4.26)$$

For convenience, let Y_x denote the random variable with survival function (4.26). Then the MRL of X is from (4.6)

$$r(x) = \frac{1}{R(x+1)} \sum_{x+1}^{\infty} (z-x)\, f(z),$$

$$= \frac{1}{R(x+1)} \sum_{x+1}^{\infty} R(z),$$

where, f(.) is the probability mass function of X. This can be written as

$$r(x) = \frac{1}{R(x+1)} \sum_{y=o}^{\infty} R(x+y+1) = \sum_{y=o}^{\infty} R(y;x),$$

$$= E(Y_x),$$

so that the definition (4.26) is consistent with the notion of MRL and residual life given in section 4.4.

Our first concern is the form of the RLD when X follows the class of models considered so far. This is vindicated through the following theorem.

Theorem 4.6.

Y_x is G(p) (or W(a+x+1, b+x+1) or NH(k,n-x-1)) if and only if X is G(p) (or W(a,b) or NH(k,n)) and conversely.

Proof:

When X follows G(p), $R(x) = q^x$ and hence ,

$$R(y;x) = q^{x+y+1} / q^{x+1} = q^y = R(y) \quad . \tag{4.27}$$

Then the RLD is also geometric with the same parameter as X. On the other hand, X is W(a,b) implies,

$$R(x) = (a-b) \sum_{x}^{\infty} [(b)_t/(a)_{t+1}],$$

$$= (a-b) \frac{(b)_x}{(a)_{x+1}} \sum_{r=o}^{\infty} [(b+x)_r/(a+x+1)_r],$$

$$= (a-b) \frac{(b)_x}{(a)_{x+1}} \frac{a+x}{a-b} \text{ ,(from formula (4.19))}$$

$$= (b)_x/(a)_x. \tag{4.28}$$

Accordingly,

$$R(y;x) = \frac{(b)_{x+y+1} \; (a)_{x+1}}{(a)_{x+y+1} \; (b)_{x+1}} \; ,$$

$$= \frac{(b+x+1)_y}{(a+x+1)_y} \; , \; y = 0,1,2,\ldots \; , \tag{4.29}$$

and therefore Y_x is W(a+x+1, b+x+1). Lastly when X is NH(k,n)

$$R(x) = \sum_{x}^{\infty} \binom{-1}{t} \binom{-k}{n-t} / \binom{-1-k}{n},$$

$$= \sum_{x}^{n} \binom{k+n-t-1}{n-t} / \binom{k+n}{n},$$

on using $\binom{-p}{k} = (-1)^k \binom{p+k-1}{k}$.

Thus,

$$R(x) = \sum_{o}^{n-x} \binom{k-t-1}{t} / \binom{k+n}{n}$$

The last sum is, however, reduced to the following by virtue of the combinatorial identity (Riordan, 1968)

$$\sum_{x=o}^{n} \binom{a+n-x-1}{n-x} = \binom{a+n}{n}, \tag{4.30}$$

so that,

$$R(x) = \binom{k+n-x}{n-x} / \binom{k+n}{n},$$

and

$$R(y;x) = \binom{k+n-x-y-1}{n-x-y-1} / \binom{k+n-x-1}{n-x-1},$$

$$= \sum_{t=y}^{n-x-1} \left[\binom{-1}{t} \binom{-k}{n-x-t-1} / \binom{-1-k}{n-x-1} \right],$$

showing that Y_x is NH(k,n-x-1).

Let us look at the converse proposition. Here we are given that, for example, in the Waring case,

$$R(y;x) = (b+x+1)_y \,/\, (a+x+1)_y,$$

and therefore from (4.26),

$$\frac{R(x+y+1)}{R(x+1)} = (b+x+1)_y \,/\, (a+x+1)_y \;.$$

Setting x to zero ,

$$\frac{R(y+1)}{R(1)} = \frac{(b+1)_y}{(a+1)_y} \;,$$

and hence ,

$$R(y) = (b)_y \,/\, (a)_y.$$

The proof for other models are on similar lines and the truth of the theorem is established.

4.6 Characterization By Properties Of Residual Life*

In continuation with the characterization of the models by properties of mean residual life renewed in

* The results in this section have appeared in Cal. Statist. Assoc. Bull. (Reference 33)

chapter 2, we now present certain properties based on the form of the variance of residual life. The applications of the Theorem in reliability analysis is discussed in chapter V.

Theorem 4.7.

$$\text{If } b(x) = V(Y(x))/EY(x)\, E[Y(x)-1] = C,$$

a constant, then a necessary and sufficient condition that X follows

(i) geometric distribution is $C=1$

(ii) negative hypergeometric distribution, $NH(k,n)$, is $C < 1$, and

(iii) Waring distribution, $W(a,b)$, is $C > 1$.

Note that the symbol $V(X)$ stands for the variance of the random variable X.

Proof:

First we prove that the condition is necessary. By definition,

$$V[Y(x)] = Cr(x)\,(r(x)-1),$$

where,

$$V[Y(x)] = \frac{1}{R(x+1)} \sum_{x+1}^{\infty} (y-x)^2\, f(y) - r^2(x),$$

or,

$$r(x)\ R(x+1)[Cr(x)-C+r(x)] = \sum_{n=1}^{\infty} n^2 f(x+n)\ ,$$

$$= \sum_{n=1}^{\infty} n^2[R(x+n)-R(x+n+1)]\ ,$$

$$= R(x+1)+\sum_{n=1}^{\infty} [(n+1)^2-n^2]\ R(x+n+1).$$

The right hand side may be simplified as,

$$R(x+1)+2\sum_{n=1}^{\infty} nR(x+n+1)+\sum_{n=1}^{\infty} R(x+n+1)=\sum_{n=1}^{\infty} R(x+n)+2\sum_{n=1}^{\infty} nR(x+n+1),$$

$$= r(x)\ R(x+1) + 2\sum_{n=1}^{\infty} nR(x+n+1),$$

giving,

$$2\sum_{n=1}^{\infty} nR(x+n+1) = r(x)R(x+1)(C+1)(r(x)-1). \qquad (4.31)$$

Changing x to x+1 in (4.31) and subtracting the resulting expression from (4.31), we get

$$2[R(x+2)-R(x+3)+2R(x+3)-2R(x+4)+\ \dots\ ]$$

$$= (C+1)[r(x)\ R(x+1)(r(x)-1)-r(x+1)R(x+2)(r(x+1)-1)].$$

On using the recurrence relation,

$$[r(x)-1)]\ R(x+1) = r(x+1)\ R(x+2),$$

we find,

$$2\sum_{n=1}^{\infty} R(x+n+1) = (C+1)\ [r(x)-r(x+1)+1]\ r(x+1)\ R(x+2),$$

which is the same as

$$2r(x+1)R(x+2) = (C+1)r(x+1)R(x+2)[r(x)-r(x+1)+1].$$

This, however, reduces to,

$$r(x) = r(x+1) + (1-C)/(1+C). \qquad (4.32)$$

Notice that when C=1,

$$r(x) = r(x+1), \text{for all } x \geq 0\ ,$$

and this implies that r(x) = K, a constant. From the definition of r(x), K is greater than unity so that there exists a p, satisfying $0<p<1$ such that $K=p^{-1}$. Hence from (4.11),

$$R(x) = (1-f(o))\ q^{x-1},\ q=1-p.$$

Determining f(o) such that R(o)=1, we get,

$$f(x) = pq^{x},\ x = 0,1,2,\ \ldots,$$

and X has geometric distribution as claimed.

Now taking C<1, from equation (4.32), we find

$$r(x+1) = r(x) + m,$$

so that r(x) is of the form $\ell + mx$ where $\ell > 1$ and $m < 0$ and therefore X is W(a,b).

Lastly, when C>1, $r(x) = \ell + mx$ with $\ell > 1$, but $m > 0$ so that by applying corollary 4.1 we conclude that X is NH(k,n).

It remains to prove the sufficiency of the conditions of the Theorem. We use the formula (4.31) to get,

$$(C+1)r(x)[r(x)-1] = 2[R(x+1)]^{-1} \sum_{n=1}^{\infty} nR(x+n+1),$$

$$Cr(x)\ (r(x)-1) = 2\ s(x)-r(x)\ [r(x)-1],$$

and

$$V(Y(x)) = 2s(x)-r(x)(r(x)-1),$$

where,

$$s(x) = (R(x+1))^{-1} \sum_{n=1}^{\infty} nR(x+n+1),$$

$$b(x) = 2s(x)\ [r(x)\ (r(x)-1)]^{-1} -1. \tag{4.33}$$

When X is geometric, $r(x)=p^{-1}$, $s(x) = qp^{-2}$, so that $b(x)=1$. For the Waring distribution in (4.1),

$$R(x+1)s(x) = \sum_{i=2}^{\infty} \sum_{j=i}^{\infty} (b)_{x+j} \,/\, (a)_{x+j}, \tag{4.34}$$

$$= \sum_{i=2}^{\infty} \frac{(b)_{x+i}}{(a)_{x+i-1}} \left\{ \frac{1}{(a+x+i-1)} + \frac{(b+x+i)}{(a+x+i-1)(a+x+i)} + \cdots \right\},$$

$$= \sum_{i=2}^{\infty} \frac{b(b+1)_{x+i-1}}{(a)_{x+i-1}} \left\{ \frac{1}{a-b-1} \right\}, \quad a>b+1,$$

$$= \frac{b}{a-b-1} \left\{ \frac{(b+1)_{x+1}}{(a)_{x+1}} + \frac{(b+1)_{x+2}}{(a)_{x+2}} + \cdots \right\},$$

$$= \frac{b}{a-b-1} \frac{(b+1)_{x+1}}{(a)_x} \left\{ \frac{1}{(a+x)} + \frac{(b+x+2)}{(a+x)(a+x+1)} + \cdots \right\},$$

$$= \frac{b}{a-b-1} \frac{(b+1)_{x+1}}{(a)_x} \frac{1}{a-b-2}, \quad a > b+2.$$

Thus,

$$s(x) = \frac{b}{(a-b-1)(a-b-2)} \frac{(b+1)_{x+1}(a)_{x+1}}{(a)_x \, (b)_{x+1}},$$

$$= \frac{b}{(a-b-1)(a-b-2)} \frac{\Gamma(b+x+2)\,\Gamma a\,\Gamma b\,\Gamma(a+x+1)}{\Gamma(b+1)\,\Gamma(a+x)\,\Gamma(b+x+1)\,\Gamma a},$$

$$= \frac{(b+x+1)(a+x)}{(a-b-1)(a-b-2)}, \quad a>b+2 .$$

Also, from table 4.1,

$$r(x) = (a+x)/(a-b-1).$$

Finally, we write b(x) as

$$b(x) = \frac{2(b+x+1)(a+x)(a-b-1)^2}{(a-b-1)(a-b-2)(a+x)(b+x+1)} - 1,$$

$$= -1 + 2(a-b-1)/(a-b-2) > 1 .$$

To compute the expression (4.32) for the negative hypergeometric distribution, observe that its probability mass function (4.2) can be converted into the form

$$f(x) = \binom{k+n-x-1}{n-x} / \binom{k+n}{n},$$

resulting in the survival function,

$$R(x) = \binom{k+n-x}{n-x} / \binom{k+n}{n}.$$

Accordingly from (4.34),

$$\binom{k+n-x-1}{n-x-1} s(x) = \sum_{i=2}^{n} \sum_{y=o}^{n-i} \binom{k+m-y-x}{m-y-x}, \quad m=n-i ,$$

$$= \sum_{i=2}^{n} \sum_{y=o}^{n-i} \binom{k+1+m-y-x-1}{m-y-x},$$

$$= \sum_{i=2}^{n} \sum_{y=o}^{n-i} \binom{k+1+n-i-y-x-1}{n-i-y-x},$$

$$= \sum_{i=2}^{n} \binom{k+1+n-i-x}{n-i-x},$$

on using the combinatorial identity (4.30).

Thus,

$$\binom{k+n-x-1}{n-x-1} s(x) = \sum_{y=o}^{n-2} \binom{k+1+n-y-2-x}{n-y-2-x} = \binom{k+n-x}{n-x-2},$$

or

$$s(x) = (k+n-x)(n-x-1)/(a+1)(a+2).$$

Again, we have

$$r(x) = \frac{1}{R(x+1)} \sum_{y=x+1}^{n} R(y),$$

$$= \binom{k+n-x-1}{n-x-1}^{-1} \sum_{y=x+1}^{n} \binom{a+n-y}{n-y},$$

$$= \binom{k+n-x-1}{n-x-1}^{-1} \sum_{y=o}^{n-x-1} \binom{k+n-y-x-1}{n-y-x-1},$$

$$= \binom{k+n-x-1}{n-x-1}^{-1} \binom{k+n-x}{n-x-1},$$

$$= \frac{(k+n-x)!\ (n-x-1)!\ \ k!}{(n-x-1)!\ (k+1)!\ (k+n-x-1)!},$$

$$= \quad (k+n-x)/(k+1),$$

and hence,

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

$$= \frac{2(k+1)}{(k+2)} - 1 = \frac{k}{k+2} < 1 .$$

This completes the proof.

4.7. Characterization By Additive Damage Model

Let X,Y and U denote random variables in the support of the set of non-negative integers such that

$$Y = X-U,\ 0 < U < \max\ (0,\ X-t)$$

for some positive integer t. We further assume that for $x > t$, $t > 0$,

$$E[U|X=x] = \ell+mx,\ \ell > 0,\ m \neq 0.$$

Thus, Y becomes the damaged component of X where, the random mechanism that reduces X to Y is represented by the regression function U on X. If we assume that the conditional mean of U given X>x is a different linear function of the form

$$E[U|X>x] = \alpha+\beta x,$$

then it is possible to arrive at a characterization of our discrete models.

Theorem 4.8.

Given that

$$E[U|X=x] = \ell+mx, \tag{4.35}$$

it is necessary and sufficient that

$$E[U|X>y] = \alpha+\beta y,$$

for X to be distributed as geometric for $\beta = m>0$ and $\alpha > \ell > 0$; Waring for $\beta > m > 0$ and $\alpha > \ell$; negative hypergeometric for $\beta < m < 0$ and $\alpha > \ell$.

Proof:

We have

$$E[U|X>y] = \frac{1}{R(y+1)} \sum_{y+1}^{\infty} (\ell+mx)\ f(x),$$

$$= \ell + m\ E(X|X>y),$$

$$= \ell + m[y+E(X-y|X>y)],$$

$$= \ell + my + mr(y). \qquad (4.37)$$

From table 4.1, we read the values of r(y), to find,

$$E[U|X>y] = \begin{cases} (\ell+mp^{-1})+my, & \text{for } G(p), \\ \ell + am(a-b-1)^{-1} + \frac{m(a-b)y}{(a-b-1)}, & \text{for } W(a,b), \\ \ell + \frac{(k+n)m}{(k+1)} + \frac{mk}{(k+1)}y, & \text{for } NH(k,n), \end{cases}$$

which is of the form $\alpha+\beta y$. The conditions on the parameters α and β can easily verified to be as stated in the Theorem.

Conversely, if $E(U|X>y)$ is of the form in (4.36), we have from (4.37),

$$\alpha+\beta y = \ell+my+mr(y)$$

so that

$$r(y) = Ay+B$$

with

$$A = \frac{\beta-m}{m} \quad \text{and} \quad B = \frac{\alpha-\ell}{m} .$$

The form of the distribution of X follows from corollary 4.1 to Theorem 4.1.

Chapter V

RELIABILITY CONCEPTS IN DISCRETE TIME

5.1 Preliminaries

The focal theme of the last two chapters has been the characterization of discrete and continuous distributions which share the common property that their mean residual life is of linear form. Apart from this property that has immense value to reliability modelling, most other features relating to these models have also some significant implications in the context of reliability analysis. In the present chapter, we look more closely at some concepts that are frequently used in life length studies vis a vis their relationships, with reference to various notions discussed earlier along with their inherent properties.

A failure time distribution represents an attempt to describe mathematically the length of life of a component or device. However, our inability to isolate the vast body of causes, that individually or collectively are responsible for the failure of the device at a particular instant, often renders, the identification of the failure distribution very difficult. Available at the disposal of the analyst is only some actual observations on the time to failure and these have to be made use of to explore a plausible model. When the data set indicates the desirability of a skew distribution, the problem becomes even more difficult, as assymetric models differ markedly at the tails and the actual observations at the right tail are sparse on account of limited sample size. This has led to several concepts that enable differentiation between various models based upon physical considerations that governed the failure phenomenon. Some of these, such as failure rate, mean (median) residual life function, equilibrium distribution along with the essential conditions under which their properties provide specific models have already been discussed. Yet another and perhaps more versatile way of describing the failure mechanism is to expose the manner in which its life length is affected by the advancement of age. In other words, one can check whether the life length of the device is increasing, decreasing or remaining steady together with an assessment of the manner in which these improvements or deterioration in the effectiveness of the device takes place with regard to its age. The various concepts designed for this purpose are called criteria for ageing. The vast majority of literature on the various criteria for ageing treats life time as continuous with only occasional references to the discrete. Recently there is some spurt of activity towards reliability analysis in the discrete time domain.

Xekalaki (1983a), points 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. Further discrete models do occur in a natural way as in fatigue studies the time to failure is measured in terms of the number of cycles to failure which is obviously integer valued. Accordingly elaboration of various concepts analogous to those in the continuous cases become necessary to distinguish classes of life distributions based on the notions of ageing. The definitions of the various classes, their characterizations and some implications among them are discussed in the following sections. It may be noticed that Klefsjö (1982) have touched upon the definitions of the various classes introduced below.

5.2 Increasing (Decreasing) Failure Rate

We first present the oldest and perhaps the simplest concept of ageing based on the monotone character of the failure rate.

Definition 5.1.

A discrete random variable X or the corresponding survival function $R(x) = P[X>x]$ belongs to the increasing failure rate or IFR (decreasing failure rate or DFR) class if

$$h(x) = P[X=x] \;/\; R(x)$$

is an increasing (decreasing) function of x, for all x in I, where I is the set of non-negative integers.

From the point of view of elucidating the above definition and also in proving some other results, two characterizations of the IFR class will be established. Throughout the sequel, the proof will be limited to the IFR class, it being understood that by reversing the monotonicity, results for the dual DFR class follow atonce.

5.2.1 Characterizations.

Theorem 5.1.

X is IFR (DFR) if and only if $R(x+y)/R(x)$ is a decreasing (increasing) function of x for all y in I.

Proof:

When $R(x+y)/R(x)$ is an increasing function of x for all y we can write,

$$\frac{R(x+1)}{R(x+y+1)} - \frac{R(x)}{R(x+y)} \geq 0, \text{ for all y in I.}$$

Hence,

$$\frac{R(x+1)}{R(x)} \geq \frac{R(x+y+1)}{R(x+y)} ,$$

and

$$1- \frac{R(x+1)}{R(x)} \leq 1- \frac{R(x+y+1)}{R(x+y)} ,$$

since, for all x the ratio $R(x+1)/R(x)$ does not exceed unity. It now follows that $h(x) \leq h(x+y)$ and therefore $h(x)$ is an increasing function of x. Thus X is IFR. The converse is obtained by retracing the above proof from end to the beginning.

Theorem 5.2.

X is IFR (DFR) if and only if $H(x,y)$ is an increasing (decreasing) function of x for all y in I, where $H(x,y)$ is the cumulated failure rate in the interval $[x,x+y-1]$ defined by,

$$H(x,y) = \sum_{t=x}^{x+y-1} h(t), \quad x \geq 0. \tag{5.1}$$

Proof:

First we suppose that X is IFR. Then for all y in I,

$$H(x+1,y)-H(x,y) = \frac{1}{x+1}\sum_{t=y}^{x+y} h(t) - \frac{1}{x}\sum_{t=y}^{x+y-1} h(t), \qquad (5.2)$$

$$= \left(\frac{1}{x+1} - \frac{1}{x}\right)\sum_{t=y}^{x+y-1} h(t) + \frac{h(x+y)}{x+1}.$$

$$= \frac{1}{x+1}\left[h(x+y) - \frac{1}{x}\sum_{t=y}^{x+y-1} h(t)\right].$$

Since h(x) is increasing in x,

$$\frac{1}{x}\sum_{t=y}^{x+y-1} h(t) \leqslant h(x+y-1),$$

and therefore,

$$H(x+1,y)-H(x,y) \geqslant \left(\frac{1}{x+1}\right)[h(x+y)-h(x+y-1)],$$

$$\geqslant 0.$$

Thus H(x,y) is increasing with x. Conversely, relation (5.2) gives,

$$x\sum_{x}^{x+y} h(t) \geqslant (x+1)\sum_{x}^{x+y-1} h(t),$$

or

$$xh(x+y) \geqslant \sum_{x}^{x+y-1} h(t).$$

Hence,

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

and

$$h(x+y+1) \geq \frac{x+1}{x} h(x+y), \text{ for all } x > 0$$

and y in I which implies

$$h(x+y+1) \geq h(x+y),$$

and the required conclusions.

In section 4.4 properties were discussed of the distribution based on the partial sums of a discrete model. If X denotes the life of a component with survival function R(x), and whenever the component fails it is replaced by another new unit which acts independently of the first. When the renewal of the system is continued indefinitely, Feller (1968) has shown that the asymptotic distribution of the residual life Y of the unit under observation at time t has the distribution (4.12). Deshpande et. al. (1986) considers the comparison between R(x) and G(x), the survival functions of X and Y in the continuous case as meaningful. Their point of view is that the life distribution of a unit which ages more rapidly will come off worse in such a comparison. In this sense, the following result is meaningful.

Theorem 5.3.

If X is IFR then $h(x) \leq k(x)$, for all x in I. (See section 4.4.1 for the definitions of h(.) and k(.).)

Proof:

When X is IFR, from Gupta (1979), Y is also IFR. This means that

$$k(x+1) \geq k(x),$$

or

$$1 - \frac{k(x+1)}{k(x)} \leq 0.$$

Therefore,

$$k(x+1)+1 - \frac{k(x+1)}{k(x)} \leq k(x+1),$$

and

$$1 + k(x+1)\left[1- \frac{1}{k(x)}\right] \leq k(x+1).$$

The left hand side is h(x+1) from equation (4.17) and our result is proved.

5.3 Increasing (Decreasing) Failure Rate Average

The class of distributions distinguished by increasing failure rate average or IFRA (decreasing failure rate average or DFRA) property was introduced for continuous random variables by Birnbaum et.al.(1966) in an attempt to find a new class of life distributions that reflect the phenomenon of wear-out. They have shown that this class (1) limiting case of no wear; that is, all exponential distributions, (2) preserves the wearing out phenomenon for a system in which the components also have the same behaviour, (3) is the smallest one with properties (1) and (2). Klefsjö (1982) has considered the discrete IFRA class, prefering to define it in terms of the behaviour of $[\bar{F}(x)]^{1/x}$ where $\bar{F}(x)=R(x+1)=P[X>x]$, as in the continuous case. While in the continuous case

$$-\frac{1}{x}\log\bar{F}(x) = \frac{1}{x}\int_{o}^{x} h(t)dt$$

provides the average failure rate in $[o,x]$ no such meaning can be given to $\log\bar{F}(x)$ in the discrete case. In order to retain the notion of averaging the failure rate, we adopt the following definition of the IFRA class.

Definition 5.2.

A discrete random variable X or its survival function belongs to the IFRA class if

$$\frac{1}{x+1}\sum_{t=o}^{x} h(t) \geq \frac{1}{x}\sum_{t=o}^{x-1} h(t),$$

or equivalently, $H(x,o)$ is an increasing function of x, for every x in I. The DFRA class is defined by reversing the above inequality.

5.3.1 Properties

Directly from the discussions in the previous section we conclude

Theorem 5.4.

1. The IFRA (DRFA) class contains all IFR(DFR) distributions.

2. If $R(x+y)/R(x)$ is a decreasing (increasing) function of x, for all $y \geqslant 0$, then X is IFRA (DFRA).

3. The function $H(x,y)$ is increasing (decreasing) in x for all y implies that X is IFRA (DFRA).

Theorem 5.5.

If X is IFRA then

$$R(x)\, R(y) \geqslant R(x+y),$$

for all $y \geqslant 0$.

Proof:

By definition 5.2, X belongs to the IFRA class if

$$\frac{1}{x+1} \sum_{t=o}^{x} h(t) - \frac{1}{x} \sum_{t=o}^{x-1} h(t) \geqslant 0,$$

or

$$\frac{1}{x+1} \sum_{t=o}^{x} \left[1- \frac{R(t+1)}{R(t)}\right] - \frac{1}{x} \sum_{t=o}^{x-1} \left[1- \frac{R(t+1)}{R(t)}\right] \geqslant 0,$$

or

$$\frac{1}{x} \sum_{t=o}^{x-1} \frac{R(t+1)}{R(t)} \geqslant \frac{1}{x+1} \sum_{t=o}^{x} \frac{R(t+1)}{R(t)},$$

or

$$\sum_{t=o}^{x-1} \frac{R(t+1)}{R(t)} \geqslant \frac{x}{x+1} \sum_{t=o}^{x} \frac{R(t+1)}{R(t)} . \qquad (5.3)$$

Now, suppose that the condition $R(x)\, R(y) \geqslant R(x+y)$ is violated for the IFRA class for atleast one y, say, y=1. Then one should have

$$\sum_{t=o}^{x-1} \frac{R(t+1)}{R(t)} > x\, R(1).$$

This implies

$$\frac{1}{x} \sum_{t=o}^{x-1} \frac{R(t+1)}{R(t)} > \frac{R(1)}{R(0)} ,$$

and

$$1- \frac{1}{x} \sum_{t=o}^{x-1} \frac{R(t+1)}{R(t)} < 1 - \frac{R(1)}{R(0)} .$$

Since,

$$\frac{R(t+1)}{R(t)} = 1-h(t),$$

this should mean that

$$\frac{1}{x} \sum_{t=o}^{x-1} h(t) < h(0),$$

which contradicts our hypothesis and therefore the theorem is proved.

Theorem 5.6.

If X belongs to the IFRA class, then $r(x) \leqslant r(o)$, for all $x \geqslant 0$.

Proof:

The IFRA nature of X means that by

Theorem 5.5,

$$R(x)\ R(y) \geqslant R(x+y), \text{ for all } x,y \geqslant 0.$$

Then,

$$R(x+1) \sum_{y=1}^{\infty} R(y) \geqslant \sum_{y=1}^{\infty} R(x+y+1),$$

for all $x \geqslant 0$.

That is,

$$R(x+1)\ r(o)\ R(1) \geqslant \sum_{y+1}^{\infty} R(t),$$

or

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

this gives

$$r(o) \geqslant r(x).$$

The next property concerns a characterization of the IFRA nature of the equilibrium distribution.

Theorem 5.7.

Y is IFRA if and only if $\frac{1}{x+1} \sum_{o}^{x} \frac{1}{r(t)}$ is an increasing function of x, for all $x > 0$.

Proof:

The proof follows from definition 5.2 and equation (4.13).

5.4. Decreasing (Increasing) Mean Residual Life

A chronological review of the development of reliability concepts reveals that the notion of failure rate was pursued to the more fundamental than the mean residual life. The potential of the mean residual life in describing various laws of failure had received only limited attentions from earlier researchers. In fact, the independence- dependence relationship between $h(x)$, $R(x)$ and $r(x)$ can be reversed and it is possible to look at the failure patterns based on the behaviour of MRL. At times, the MRL appears to be a better concept than the failure rate once we closely examine the definitions of the two. The failure rate take into account, the behaviour of the survival probabilities at times x and x+1, while $r(x)$ utilizes the entire information about the survival probabilities from age x onwards till failure. Thus, if an equipment does not fail in the near future, its failure rate may be zero and at the same time, the MRL may be decreasing as failure is bound to occur once the period for which $h(x) = 0$ is surpassed. This difference in the behaviour of the two functions has resulted in ageing concepts based on MRL. We first introduce the decreasing mean residual life or DMRL (increasing mean residual life or IMRL) class.

Definition 5.3.

A discrete random variable X or its distribution belongs to the DMRL class if $r(x) \geq r(x+1)$ and belongs to the dual IMRL class if $r(x) \leq r(x+1)$ for every x in I.

5.4.1. Properties.

Theorem 5.7.

A sufficient condition for X to have DMRL(IMRL) is that $R(x+y)/R(x)$ is a decreasing (increasing) function of x, for every y in I.

Proof:

By hypothesis, $R(x+y+1)/R(x+1)$ is a decreasing function of x. Summation over y yields,

$$\frac{1}{R(x+1)} \sum_{y=o}^{\infty} R(x+y+1) \quad \text{or} \quad \frac{1}{R(x+1)} \sum_{x+1}^{\infty} R(y)$$

is decreasing in x. This is the same as saying that R(x) decreases with x for all y.

Corollary 5.1.

If X is IFR(DFR), then X is DMRL (IMRL).

Corollary 5.2.

If $H(x+1,y) \geq H(x,y)$, then X is DMRL.

Notice that these corollaries are direct consequences of Theorems 5.1 and 5.2.

It remains to demonstrate that as in the continuous case, DMRL does not imply IFR; otherwise we would not be introducing a new notion of ageing since the IFR class implies the DMRL class. For the purpose we present the following example.

X	0	1	2	3	4
f(x)	0.21	0.15	0.22	0.22	0.20
R(x)	1.00	0.79	0.64	0.42	0.20
h(x)	0.21	0.19	0.34	0.52	1.00
r(x)	2.60	2.00	1.50	1.00	0

It is clear that r(x) steadly decreases with x and h(x) is not always increasing.

We have postponed the physical interpretation available to some of the characteristic properties of the discrete models considered in Theorems 4.2 and 4.5 of the previous chapter for later consideration. It will now be established that they are in fact closely associated with the notion of being discussed in this section.

Theorem 5.8.

A necessary and sufficient condition for X to be DMRL is that

$$V(Y_x) \leqslant r(x)\,[r(x)-1].$$

The condition for IMRL class is obtained by reversing the inequalities.

Proof:

We have

$$EY_x^{\,2} = \frac{1}{R(x+1)} \sum_{y=x+1}^{\infty} (y-x)^2 \, f(x),$$

$$= \frac{1}{R(x+1)} \sum_{1}^{\infty} n^2 f(x+n),$$

$$= \frac{1}{R(x+1)} \sum_{1}^{\infty} n^2 [R(x+n)-R(x+n+1)],$$

$$= 1+ \frac{1}{R(x+1)} \sum_{1}^{\infty} (2n+1) \, R(x+n+1). \qquad (5.4)$$

Also,

$$EY_x = \frac{1}{R(x+1)} \sum_{n=1}^{\infty} R(x+n),$$

$$= 1+ \frac{1}{R(x+1)} \sum_{n=1}^{\infty} R(x+n+1) \qquad (5.5)$$

Subtracting (5.5) from (5.4)

$$EY_x^2 - r(x) = \frac{2}{R(x+1)} \sum_{n=1}^{\infty} n\, R(x+n+1),$$

$$= \frac{2}{R(x+1)} \left[\sum_{n=2}^{\infty} R(x+n) + \sum_{n=3}^{\infty} R(x+n) + \ \dots \ \right],$$

$$= \frac{2}{R(x+1)} \sum_{x+1}^{\infty} R(t+1)\ r(t),$$

$$= \frac{2}{R(x+1)} \sum_{x}^{\infty} R(t+1)\ r(t) - 2r(x),$$

$$= \frac{2}{R(x+1)} \sum_{x}^{\infty} R(t+1)\ [r(t)-r(x)] + 2r^2(x) - 2r(x).$$

Hence,

$$EY_x^2 + EY_x - 2[EY_x]^2 = \frac{2}{R(x+1)} \sum_{x}^{\infty} R(t+1)[r(t)-r(x)].$$

If X has DMRL,

$$V(Y_x) - r(x)\ [r(x)-1] \leq 0.$$

Conversely, we have shown in Theorem 4.5 that for every $x > 0$,

$$V(Y_x) = c\, r(x)\ [r(x)-1]$$

is equivalent to the statement

$$r(x)-r(x+1) = \frac{1-C}{1+C},$$

where C is some positive constant. Taking $0<C\leq 1$, it follows that

$$V(Y_x) \leq r(x)\,[r(x)-1],$$

implies

$$r(x)\,[r(x)-1] \geq 0.$$

Accordingly r(x) is a decreasing function of x and the theorem is completely proved.

Theorem 5.9.

X is DMRL (IMRL) if r(x) h(x) is not less than (not greater than) unity.

Proof:

From equation (4.10),

$$1-h(x+1) = [r(x)-1] \,/\, r(x+1),$$

and therefore,

$$h(x+1)\, r(x+1) = r(x+1)-r(x)+1.$$

Thus, $h(x+1)\ r(x+1) \lessgtr 1$ implies $r(x+1)-r(x) \lessgtr 0$ and consequently X is DMRL.

The results of the last two theorems can be used to characterize the Waring and negative hypergeometric distributions among the class of distributions for which $h(x)\ r(x)$ is a constant or $b(x) = V(Y_x)/r(x)[r(x)-1]$ is a constant. We give the proof in one case only as the other follows by using the same argument.

Theorem 5.10.

Among the class of distributions with strictly increasing (decreasing) MRL, Waring (negative hypergeometric) is the only member for which $b(x)$ is a constant.

Proof:

The constancy of $b(x)$ gives rise to three cases, $b(x) \gtreqless 1$, of which $b(x)=1$ does not provide a strictly increasing or decreasing MRL as it corresponds to the geometric law with constant MRL. Let us take $b(x)$ to be greater than unity. From Theorem 4.5 we find that $b(x)=C$ implies

$$r(x) \quad = \quad r(x+1) + (1-C)/(1+C).$$

For $C>1$, we thus have $r(x) < r(x+1)$ and hence MRL is an increasing function of x. Similarly, when $C<1$, $r(x)$ is decreasing in x and by Theorem 4.5 these values of C are characteristic of the Waring and negative hypergeometric distributions. This completes the proof.

5.5. Criteria Used in Maintenance Policies

Another category of ageing concepts considered in literature are those that help the study of maintenance policy which are followed to reduce the incidence of system failure or to return a failed system to the operating state. We consider some classes of distributions that are specially designed for application in this context. The continuous version of these classes are discussed in Marshall and Proschan (1972).

Definition 5.4.

A discrete random variable X, with positive integer values as its support, or its distribution is new better than used or NBU (new worse than used or NWU) if

$$R(x+y+1) \leqslant (\geqslant)\ R(x+1)\, R(y+1), \tag{5.6}$$

for $x \geqslant 0$, $y \geqslant 0$.

Definition 5.5.

X is new better than used in expectation or NBUE (new worse than used in expectation or NWUE) if

(a) X has finite (finite or infinite) mean m,

(b) $r(x) \leqslant (\geqslant)$ m, for $x \geqslant 0$.

The quantity R(x+y+1)/R(y+1) represents the survival function of a unit of age y or the conditional probability that a unit of age y will survive for an addition x unit of time. At y=0, [R(x+y+1)/R(y+1)] = R(x+1) is the survival function of a new unit and accordingly the ageing of the device can be studied by comparing R(x+y+1)/R(y+1) and R(x+1). Thus, $R(x+y+1) \leqslant R(x+1)\, R(y+1)$ if and only if the older system has aged is that it has no better chance of surviving for a duration of x than does a new system. In other words, the new unit is better than the used one or NBU. A similar interpretation can be given to the NWU which exhibits the benefit of ageing. On the other hand the condition $r(x) \leqslant r(o) = m$ states that the expected remaining life of a unit surviving age x is not larger than the expected life of a new unit so that the new unit fairs better than the unit of age x in terms of the expected life length. This explains the terminology NBUE. It is easy to see that the boundary of both the classes is the geometric law,

$$P[X=x] = q^{x-1}p, \quad x = 1,2,\ldots, \quad 0 < p < 1, \; p+q = 1.$$

Definition 5.6.

The distribution of a positive integer valued random variable belongs to the HNBUE (harmonic new better than used in expectation) class if

$$\sum_{t=x}^{\infty} R(t+1) \leqslant \sum_{t=x}^{\infty} G(t+1) = m(1-\frac{1}{m})^{x},$$

where G(x) is the survival function of a geometric random variable with mean $m = \sum_{0}^{\infty} R(x)$. The class of HNBUE distributions was introduced in the continuous case by Rolski (1975) in the above definition is as in Klefsjö (1982). It is obvious from the definition that the NBUE class is contained in the HNBUE class. We further have the following properties.

1. If X is NBU then X is NBUE.

This is obtained by summation of (5.6) with respect to y from 0 to ∞.

2. X is NBU if and only if $k(x) \geqslant k(o)$. (We use the definition of NBUE and the relationship $k(x) = \frac{1}{r(x)}$.)

3. X is HNBUE if and only if $\frac{1}{x}\sum_{t=1}^{x} k(t) \geq k(o)$ for $x=1,2,\ldots$

4. Y is NBUE if and only if $h(o) \leq h(x)$.

5. Y is HNBUE if and only if $h(o) \leq \frac{1}{x}\sum_{t=1}^{x} h(t)$.

We have also,

Theorem 5.11.

For a non-decreasing function g(.), the random variable X is HNBUE if and only if

$$Eg(Z) \geq Eg(Y), \qquad (5.7)$$

where Z is distributed as geometric with mean m and Y is the random variable in section 4.4.

Proof:

Since g is monotonic, the condition (5.7) is satisfied if and only if $Z \geq Y$. This means that

$$[1-\frac{1}{m}]^{x-1} \geq \sum_{x=o}^{\infty} \frac{R(x+1)}{m},$$

or

$$m[1-\frac{1}{m}]^{x-1} \geq \sum_{o}^{\infty} R(x+1).$$

Accordingly X is HNBUE.

As a consequence of Theorem 5.11 we can obtain a characterization of the geometric distribution as in the following theorem.

Theorem 5.12.

Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables having $E(X) = m < \infty$ with common survival function $R(x)$. Then among the class of HNBUE laws the geometric distribution,

$$P[X=x] = pq^{x-1}, \quad x=1,2,\ldots$$

is the only law for which

$$\frac{E(X_{(1)})-1}{E(X_{(1)})} = \left(\frac{m-1}{m}\right)^n,$$

where $X_{(1)} = \min_{1 \leq i \leq n} X_i$.

Proof:

Assume X_i to be geometric with mean m. Then,

$$R(x) = \left(1-\frac{1}{m}\right)^{x-1}.$$

Now,
$$P[X_{(1)} \geq x] = [R(x)]^n$$
$$= \left[\left(1-\frac{1}{m}\right)^n\right]^{x-1},$$

so that, $X_{(1)}$ is geometric with mean

$$E(X_{(1)}) = \frac{1}{1-(1-\frac{1}{m})^n} .$$

Therefore,

$$\frac{E(X_{(1)})-1}{E(X_{(1)})} = \left(\frac{m-1}{m}\right)^n ,$$

as stated.

Conversely, assume that X_i is HNBUE. Then we can write,

$$\begin{aligned} E(X_{(1)}) &= \sum_{1}^{\infty} [R(x)]^n, \\ &= \sum_{1}^{\infty} R(x)\, [R(x)]^{n-1}, \\ &= \sum_{1}^{\infty} R(x)\, [\, \ell(x) - \ell(x-1)], \end{aligned}$$

where,

$$\ell(x) = \sum_{t=o}^{x} [R(t)]^{n-1}.$$

Therefore,

$$E(X_{(1)}) = R(1)\ [\ \ell(1)-\ell(0)]+R(2)[\ell(2)-\ell(1)]+\ \ldots,$$

$$= \ell(1)\ [R(1)-R(2)]+\ell(2)\ [R(2)-R(3)]+\ \ldots,$$

$$= \sum_{x=1}^{\infty} \ell(x)\ f(x),$$

where,

$$f(x) = P[X_i=x],$$

$$= R(x)-R(x+1).$$

Defining,

$$h(x) = \sum_{t=o}^{x} (1-\frac{1}{m})^{(n-1)t},$$

we have from the HNBUE property of X_i,

$$\sum_{x=t+1}^{\infty} R(x) \leqslant m(1-\frac{1}{m})^{t},$$

$$= \sum_{t+1}^{\infty} (1-\frac{1}{m})^{x},$$

or

$$\sum_{x=o}^{\infty} R(x) \geqslant \sum_{x=1}^{t} (1-\frac{1}{m})^{x-1}.$$

Also,

$$\sum_{x=1}^{t}[R(x)]^2 - \sum_{x=1}^{t}(1-\frac{1}{m})^{2x} = \sum_{x=1}^{t}\ [R(x)-(1-\frac{1}{m})^{x-1}]R(x)\ +$$

$$\sum_{x=1}^{t}\ [R(x)-(1-\frac{1}{m})^{x-1}](1-\frac{1}{m})^{x-1} \geq 0.$$

Proceeding similarly we have,

$$\sum_{x=1}^{t}[R(x)]^{n-1} \geq \sum_{x=1}^{t}(1-\frac{1}{m})^{(n-1)(x-1)}.$$

Therefore,

$$\ell(t) \geq h(t).$$

Then,

$$E(X_{(1)}) = \sum_{x=1}^{\infty}\ell(x)f(x) \geq \sum_{x=1}^{\infty} h(x)f(x),$$

$$= \sum_{x=1}^{\infty}[\ \sum_{t=1}^{x}(1-\frac{1}{m})^{(n-1)(x-1)}]f(x),$$

$$= \sum_{x=1}^{\infty}[m(1-\frac{1}{m})^{(n-1)(x-1)}]\ \frac{R(x)}{m},$$

$$= E\phi(Y),$$

where,

$$\phi(Y) = m(1-\frac{1}{m})^{(n-1)(x-1)}.$$

By Theorem 5.11,

$$E\phi(Y) \geq E\phi(Z),$$

$$= \sum_{x=1}^{\infty} m(1-\frac{1}{m})^{(n-1)(x-1)} \frac{1}{m}(1-\frac{1}{m})^{x-1},$$

$$= \frac{1}{1-(1-\frac{1}{m})^{n}} .$$

In the above inequality, the equality sign will hold good if and only if $\ell(x) = h(x)$, in which case,

$$R(x) = (1-\frac{1}{m})^{x-1},$$

as required in the Theorem.

5.6 Conclusion

The various results established in the last three chapters constitute the properties of the class of discrete and continuous distributions that are characterized by linear mean residual life. The discrete version of the ageing concepts that parallel with their continuous counterparts have also been introduced. An elaborate study of these ageing concepts in the discrete time domain with respect to the preservation of these properties in relation to convolutions, mixing and coherent structures remains an open problem. The behaviour of the tail distributions point out to several applications in the analysis of income and reliability. Also, further characterizations of these models are to be investigated. The discussion in the present study is confined to univariate models only. It will be interesting to identify the multivariate models possessing appropriate multivariate analogues of the univariate properties discussed here. The answers to the various questions are being investigated and will be presented elsewhere.

REFERENCES

1. Abrahams, J. and Thomas, J.B. (1984) A note on characterization of bivariate densities by conditional densities, Commun. Statist. - Theory Meth., 13(3), 395-400.

2. Abramowitz, M. and Stegun, I.A. (1972) Hand Book of Mathematical Functions, Dover Publications, Inc. New York.

3. Arnold, B.C. (1983) Pareto Distributions, International Co-operative Publishing House, Maryland, U.S.A.

4. Arnold, B.C. (1987) Bivariate distributions with Pareto conditionals, Statist. Prob. Letters, 5, 263-266.

5. Arnold, B.C. and Strauss, D. (1988) Bivariate distributions with exponential conditionals, J.Amer.Statist.Assoc., 83, 522-527.

6. Ahsanullah, M. and Kabir, A.B.M.L. (1973) A characterization of the Pareto distribution, Canadian J.Statist., 1, 109-112.

7. Azlarov, T.A. and Volodin, N.A. (1986) Characterization Problems Associated With The Exponential Distribution, Springer-Verlag, New York.

8. Beg, M.I. and Kirmani, S.N.U.A. (1974) On a characterization of exponential and related distributions, Aust.J. Statist ., 16, 163-166.

9. Bhatacharya, N. (1963) A property of the Pareto distribution, Sankhya, B, 25, 195-196.

10. Birnbaum, Z.W., Esary, J.D., and Marshall, A.M. (1966) A stochastic characterization of wear-out for components and systems, Ann.Math.Statist., 37(4), 816-825.

11. Bluementhal, S. (1967) Proportional sampling in life length studies, Technometrics, 9, 205-218.

12. Chong, K.M. (1977) On characterizations of exponential and geometric distributions by expectation, J.Amer.Statist. Assoc., 72, 160-161.

13. Cox, D.R. (1967) Renewal Theory, Methuen, London.

14. Cramer, H. (1946) Mathematical Methods of Statistics, Princeton University Press.

15. Crawford, C.B. (1966) Characterization of geometric and exponential distributions, Ann.Math.Statist., 37, 1790-1795.

16. Dallas, A.C. (1976) Characterizing the Pareto and power distributions, Ann.Inst.Statist.Math., 28, 491-497.

17. David, H.A. (1970) Order Statistics, Wiley, New York.

18. Deshpande, J.V., Kochar, S.C. and Singh, H. (1986) Aspects of positive ageing, J.Appl.Prob., 23, 748-758.

19. Dubey, S.D. (1966) Characterization theorems for several distributions and their applications, J.Indust. Math.Soc., 16, 1-22.

20. Erdélyi, A., Magnus, W., Oberhettinger, F. and Tricomi, F.G. (1954) Tables of Integral Transforms, Vol.2, Mc Graw Hill Book Co., Inc.

21. Feller, W. (1968) An Introduction to Probability Theory and Its Applications, Vol.1, Third Edition, John Wiley and Sons, Inc., New York.

22. Ferguson, T.S. (1967) On characterizing distributions by properties of order statistics, Sankhya, A, 29, 265-278.

23. Fisz, M. (1958) Characterization of some probability distributions, Skand. Aktuar., 1-2, 65-70.

24. Galambos, J. and Kotz, S. (1977) Characterizations of Probability Distributions, Springer-Verlag, Berlin.

25. Govindarajlu, Z. (1966) Characterizations of the exponential and power distributions, Skand. Aktuar., 49,132-136.

26. Gupta, P.L. and Gupta, R.C. (1983) On the moments of residual life in reliability and some characterization results, Commun.Statist.- Theory Meth., 12(4), 449-461.

27. Gupta, R.C. (1979) On the characterization of survival distributions in reliability by properties of their renewal densities, Commun.Statist.- Theory Meth., 8(7), 685-697.

28. Gupta, R.C. (1980) On the mean residual life function of a mixture in reliability studies, Commun. Statist.- Theory Meth., 9(13), 1365-1369.

29. Haines, A. and Singpurwalla, N.D. (1974) Some contributions to stochastic characterization of wear, Reliability and Biometry, F.Proschar and R.J.Serfling, Eds., Soc. for Indus.Appl.Math., Philadelphia.

30. Hangstroem, K.G. (1925) La loi de Pareto et la reassurance, Skand.Aktuar., 8, 65-88.

1. Harris, C.M. (1968) The Pareto distribution as a queue service discipline, Operat.Res., 16, 307-313.

2. Hitha, N. (1990) Two characterizations of Pareto distribution, Proc. 77th Session of the Indian Science Congress Association, 13.

33. Hitha, N. and Nair, N.U. (1989) Characterization of some discrete models by properties of residual life function, Cal. Statist.Assoc. Bull., 38(151-152), 219-223.

34. Hitha, N. and Nair, N.U. (1990) Characterizations of bivariate Lomax and finite range distributions (communicated).

35. Irwin, J.O. (1963) The place of Mathematics in medical and Biological Statistics, J.Roy.Statist.Soc., A, 126, 1-44.

36. Irwin, J.O. (1968) The generalized Waring distribution applied to accident theory, J.Roy.Statist.Soc.A, 131, 205-225.

37. Irwin, J.O. (1975 a,b,c) The generalized Waring distribution, Parts I, II, III, J.Roy. Statist.Soc., A, 138, 18-31, 204-227, 374-384.

38. James, I.R. (1979) Characterization of a family of distributions by independence of size and shape variables, Ann. Statist., 7, 869-881.

39. Joe, H. and Proschan, F. (1983) Tests for properties of the percentile residual life function, Commun. Statist.-Theory Meth., 12(10), 1087-1119.

40. Johnson, N.L. and Kotz, S. (1969) Distributions in Statistics: Discrete Distributions, John Wiley and Sons, Inc., New York.

41. Kagan, A.M., Linnik, Yu. V., and Rao, C.R. (1973) Characterization Problems in Mathematical Statistics, U.S.S.R. Academy of Sciences (English translation, Wiley, New York).

42. Kakwani, C.N. (1980) Income Inequality and Poverty, Oxford University Press.

43. Kalbfleisch, J.D. and Prentice, R.L. (1980) The Statistical Analysis of Failure Time Data, John Wiley and Sons, Inc. New York.

44. Klefsjö, B. (1982) The HNBUE and HNWUE classes of life distributions, Naval Res. Logist. Quart., 29, 331-344.

45. Korwar, R.M. (1989) On a new system of discrete distributions and characterizations of several discrete distributions by equality of distributions, Ann. Inst. Statist. Math., 41(2), 305-321.

46. Kotz,S. and Shanbag,D.N. (1980) Some new approaches to probability distributions, Adv.Appl.Prob., 12, 903-921.

47. Krishnaji, N. (1970a) Characterization of the Pareto distribution through a model of under-reported incomes, Econometrika, 38, 251-255.

48. Krishnaji, N. (1970b) A characteristic property of the Yule distribution, Sankhya, A, 32, 343-346.

49. Krishnaji, N. (1971) Note on a characterizing property of the exponential distribution, Ann. Math. Statist., 42, 361-362.

50. Kupka, J. and Loo, S. (1989) The hazard and vitality measures of ageing, J.Appl.Prob., 26, 532-542.

51. Lau, Ka-Sing and Rao, C.R. (1982) Integrated Cauchy functional equation and characterizations of the exponential law, Sankhya, A, 44, 72-90.

52. Laurent, A.G. (1974) On characterization of some distributions by truncation properties, J. Amer. Statist. Assoc., 69, 823-827.

53. Lindley, D.V. and Singpurwalla, N.D. (1986) Multivariate distributions for life lengths of a system sharing a common environment, J.Appl.Prob., 23, 418-431.

54. Lukacs, E. (1956) Characterization of populations by properties of suitable statistics, Proce. Third Berkeley Sympos. Math.Statist. and Prob., Vol.2, 195-214.

55. Lukacs, E. and Laha, R.G. (1964) Applications of Characteristic Functions, Hafner Publishing Co., New York.

56. Marshall, A.W. and Proschan, F. (1972) Classes of distributions applicabl in replacement with renewal theory implications, Proce. Sixth Berkeley Sympos. Math. Statist. and Prob., University of California Press, Berkeley,395-415

57. Malik, H.J. (1970) Distribution of product statistics from a Pareto population, Metrika, 15, 19-22.

58. Mathai, A.M. and Pederzoli, G. (1977) Characterizations of the Normal Probability Law, Wiley Eastern, New Delhi.

59. Morrison, D.G. (1978) On linearly increasing mean residual life times, J.Appl. Prob., 15, 617-620.

60. Mosimann, J.E. (1970) Size allometry: Size and shape variables with characterizations of the lognormal and generalized gamma distributions, J.Amer. Statist. Assoc., 65, 930-945.

61. Mukherjee, S.P. and Roy, D. (1986) Some characterizations of the exponential and related distributions, Cal.Statist.Assoc. Bull., 35, 189-197.

62. Muth, E.J. (1980) Memory as a property of probability distributions, I.E.E.E. Trans.Rel., 29(2), 160-164.

63. Nair, N.U. (1983) A measure of memory for some discrete distributions, J.Ind. Statist.Assoc., 21, 141-147.

64. Nair, N.U. (1989) Characterization of some multivariate distributions, J.Korean Statist.Soc., 18(1), 72-79.

65. Nair, N.U. and Hitha, N. (1989) Characterization of discrete models by distribution based on their partial sums, Statist. Prob.Letters, 8, 335-337.

66. Nair, N.U. and Hitha, N. (1990) Characterizations of Pareto and related distributions, J.Ind. Statist.Assoc., 28, 75-79.

67. Nair, N.U. and Nair, K.R.M. (1990) Characterizations of a bivariate geometric distribution, Statistica, 46(1), 247-253.

68. Ord, J.K., Patil, G.P. and Taillie, C. (1981) The choice of a distribution to describe personal incomes, Statistical Distributions in Scientific Work, Vol.6, C.Taillie, G.P.Patil and B.Baldessari, Eds., D.Reidel, Holland, 193-202.

69. Ord, J.K., Patil, G.P. and Taillie, C. (1983) Truncated distributions and measure of income inequality, Sankhya, B,45, 413-430.

70. Panaretos, J. and Xekalaki, E. (1986) On some distributions arising from certain generalized sampling schemes, Commun. Statist.- Theory Meth., 15(3), 873-891.

71. Patil, G.P. and Ratnaparkhi, M.V. (1975) Problems of damaged random variables and related characterizations, Statistical Distributions in Scientific Work, Vol.3, G.P. Patil, S. Kotz and J.K.Ord, Eds., D. Reidel, Holland.

72. Patil, G.P., Kotz,S. and Ord, J.K. (1975) A modern course on Statistical Distributions in Scientific Work, Vol.3 - Characterizations and Applications, D. Reidel Publishing Co., Holland.

73. Pederzoli, G. and Rathie, P.N. (1980) Distribution of product and quotient of Pareto variates, Metrika, 27, 165-169.

74. Rao, C.R. and Rubin, H. (1964) On a characterization of the Poisson distribution, Sankhya, A, 26, 294-298.

75. Revankar, N.S., Hartley, M.J. and Pagano, M. (1974) A characterization of the Pareto distribution, Ann. Statist., 2, 599-601.

76. Riordan, J. (1968) <u>Combinatorial Identities</u>, John Wiley and Sons, Inc., New York.

77. Rogers, G.S. (1959) A note on stochastic independence of functions of order statistics, Ann.Math.Statist., 30, 1263-1264.

78. Rogers, G.S. (1963) An alternative proof of the characterization of the density Ax^B, Amer.Math.Monthly, 70,857-858

79. Rolski, T. (1975) Mean residual life, Bull. Int. Statist. Inst., 46, 266-270.

80. Rossberg, H-J. (1972a) Characterization of the exponentia and Pareto distributions by means of some properties of the distribu tions which the differences and quotients of order statistics are subject to, Math. Operationsfors U. Statist., 3, 207-216.

81. Rossberg, H-J. (1972b) Characterization of distribution functions by the independence of certain functions of order statistics, Sankhya, A, 34, 111-12

82. Schmittlein, D.C. and Morrison, D.G. (1981) The mean residual life time: A characterization theorem and an application, Operat. Res., 29, 392-399.

83. Seshadri, V. and Patil, G.P. (1964) A characterization of bivariate distributions by the marginal and conditional distribution of the same component, Ann. Inst. Statist. Math., 15, 215-221.

84. Shah, S.M. and Kabe, D.G. (1981) Characterizations of exponential, Pareto, power function, Burr and logistic distributions by order statistics, Biom.J., 23(2), 141-146.

85. Singh, H. (1989) On partial orderings of life distributions, Naval Res. Logist. Quart., 36, 103-110.

86. Srivastava, M.S. (1965) A characterization of Pareto's distribution and $(k+1)x^k/\theta^{k+1}$ (Abstract), Ann. Math. Statist., 36, 361-362.

87. Sullo, P. and Rutherford, D. (1977) Characterizations of the power distribution by conditional exceedence, Research Report 37-77-P6, Renesselaer Poly. Inst., Troy, New York.

88. Talwalker, S. (1980) On the property of dullness of Pareto distribution, Metrika, 27, 115-119.

89. Wang, Y.H. and Srivastava, R.C. (1980) A characterization of the exponential and related distributions by linear regression, Ann. Statist., 8, 217-220.

90. Xekalaki, E. (1983a) Hazard functions and life distributions in discrete time, Commun. Statist.- Theory Meth., 12(21), 2503-2509.

91. Xekalaki, E. (1983b) The univariate generalized Waring distribution in relation to accident theory: Proneness, spells or contagion?, Biometrics, 39, 887-895.