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 $P(a,k)$ is specified by the probability density function,

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

The distribution is J-shaped with mode located at $k$.

The mean and variance are

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

and

$$V(X) = a(a-1)^{-2}(a-2)^{-1}k^2, \quad a > 2 \hspace{1cm} (2.3)$$

The $r$th moment about the origin takes the form,

$$\mu_r = a(a-r)^{-1}k^r, \quad a > r, \quad r = 0, 1, 2, \ldots \hspace{1cm} (2.4)$$
The truncated moment of order \( r \) defined as,
\[
m_r(x) = E[(X-x)^r | X > x], \quad r = 0, 1, 2, \ldots,
\] (2.5)
for the distribution (2.1) is,
\[
m_r(x) = r! \frac{x^r}{(a-1)(r)}; \quad x > k,
\] (2.6)
where \( a(r) \) is the descending factorial expressed as
\[
a(r) = a(a) \ldots (a-r+1).
\] (2.11)

Specialising for \( r=1 \), we get the truncated mean, known more popularly as the mean residual life function (MRLF) in reliability theory. Thus,
\[
m_1(x) = x(a-1)^{-1}.
\] (2.7)

We notice that the MRLF is a linear function in \( x \), for all \( x > k \). On the other hand, the partial moment of order \( r \),
\[
p_r(x) = E[(X-x)^+]^r, \quad r = 0, 1, 2, \ldots,
\] (2.8)
where, \( (X-x)^+ = \max(0, X-x) \), is given by the expression
\[
p_r(x) = r! \frac{k^a x^r}{(a-1)(r)} x^a; \quad x > k.
\] (2.9)
Another type of moments, that is meaningful in connection with the Pareto distribution, is the reciprocal moments,

\[ b_r = E(X^{-r}), \ r = 0, 1, 2, \ldots, \]

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

(2.10)

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

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

(2.11)

has the value,

\[ C_r(x) = (-1)^r \frac{a r!}{(a+1)^{[r]} x^r}, \ x \geq k, \]  

(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 \(P(1,k)\), there is a simple closed form for the distribution of \(X_1 + X_2\). This is

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

(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,

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

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

\[ f(u) = (\prod_{i=1}^{n})^{-1} \left[ k \log \left( \frac{u}{k} \right) \right]^{n-1} \left( \frac{u}{k} \right)^{a u^{-1}}. \quad u > 0, k = k_1 k_2 \ldots k_n. \quad (2.15) \]

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

\[
p(z) = \frac{a_1a_2}{k_1k_2(a_1+a_2)} \left( \frac{z}{k_1k_2} \right)^{a_2-1} \text{ for } z \leq k_1k_2^{-1},
\]

\[
= \frac{a_1a_2}{k_1k_2(a_1+a_2)} \left( \frac{z}{k_1k_2} \right)^{a_1-1} \text{ for } z > k_1k_2^{-1}.
\]

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

\[
h(x) = \frac{f(x)}{1-F(x)}, \quad x > 0,
\]

where \( f(.) \) is the density function and \( F(.) \) is the distribution function of a non-negative random variable \( X \) satisfying the condition \( F(0) = 0 \), 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

$$G_x(y) = P[x < X \leq x+y | X > x], \ y > 0,$$

$$= 1 - \left[ \frac{R(x+y)}{R(x)} \right],$$

where,

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

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

$$\overline{G}_x(y) = \frac{R(x+y)}{R(x)}, \ y > 0. \quad (2.18)$$

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

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

which is again of the P I form. Notice that

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

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

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

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

The survival function corresponding to (2.22) becomes

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

We shall use the notation $P_{II}(a,\alpha)$ 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,

\[ f(x \mid b) = be^{-bx}, \quad x > 0, \quad b > 0, \]

and if the parameter b has a gamma distribution,

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

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

The \( r \)th moment about the origin of the distribution is,

\[ \mu_r' = a^r \frac{\Gamma(a-r) \Gamma(1+r)}{\Gamma(a)}. \quad (2.24) \]

In particular,

\[ E(X) = \alpha(a-1)^{-1}, \quad a > 1, \quad (2.25) \]

and

\[ V(X) = a^2 a(a-1)^{-2} (a-2)^{-1}, \quad a > 2. \]

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 \( P_{II}(a, \alpha) \).

The residual life distribution for \( P_{II}(a, \alpha) \) has survival function,

\[
G_x(y) = \frac{(y+x+\alpha)}{(x+\alpha)} - a, \quad (2.26)
\]

where \( G_x(y) \) 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 \( \alpha \) to \( (x+\alpha) \). Therefore, the MRLF is deduced from (2.26) as,

\[
r(x) = (x+\alpha) (a-1)^{-1} \quad (2.27)
\]

The failure rate function is a reciprocal linear function,

\[
h(x) = a(x+\alpha)^{-1}. \quad (2.28)
\]

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 \( P_{II}(a, \alpha) \) 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 Singpurwala (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,

$$f(x) = \frac{c}{R} \left(1 - \frac{x}{R}\right)^{c-1}, \quad 0 < x < R, \quad c > 0, \quad (2.29)$$

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,

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

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,

$$\mu_r = cR^r B(r+1,c), \quad r=0,1,2,3,...$$  \hspace{1cm} (2.31)

In particular, the mean and variance are

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

and

$$\text{Var}(X) = cR^2 (c+1)^{-2} (c+2)^{-1}. \hspace{1cm} (2.32)$$

The moment generating function is,

$$M(t) = c e^{itR} I(c,tR), \hspace{1cm} (2.33)$$

where

$$I(p,q) = \int_0^R e^{-qx} x^{p-1} dx.$$  \hspace{1cm} (2.35)

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

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

The MRLF is

$$r(x) = (R-x)(c+1)^{-1}. \quad (2.35)$$
which is linearly decreasing in $x$. Thus the $\text{FR}(c,R)$ 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

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

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 $\text{FR}(N,1)$. This fact is utilized in non-parametric statistical inference (David, 1970). Apart from these, $\text{FR}(c,1)$ 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.

\[ \frac{1}{(x-a)} = \frac{1}{x} + \frac{a}{x(x+1)} + \frac{a(a+1)}{x(x+1)(x+2)} + \ldots \] (2.37)

The probability function of the Waring distribution discussed here is,

\[ f(x) = P[X=x], \]

\[ = \frac{(a-b)(b)_x}{(a)_{x+1}}, \quad x=0,1,2,\ldots \] (2.38)

where, \((b)_x\) is the Pochhammer's symbol, defined as

\[ (b)_x = \Gamma(b+x)/\Gamma(b). \] (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

\[
\mu_r = \frac{b^r}{(a-b-1)^r}. \tag{2.40}
\]

In particular,

\[
E(X) = b(a-b-1)^{-1}
\]

and

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

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 resemblance is almost perfect.

The survival function of the distribution is
\[ R(x) = \sum_{x} P[X=x], \]
\[ = \sum_{x} \frac{(a-b)(b)}{(a)^{x+1}}, \]
\[ = \frac{(b)^{x}}{(a)^{x}}. \]

The MRLF is
\[ r(x) = E[X-x|X>x], \]
which is equal to (see equation 4.6)
\[ P[X=x] = \left( \begin{array}{c} -a \\ -b \end{array} \right) / \left( \begin{array}{c} -a-b \end{array} \right), \quad x = 0, 1, \ldots, n, \] (2.45)
\[ r(x) = [R(x+1)]^{-1} \sum_{t=x+1}^{\infty} R(t), \]
\[ = (a+x) (a+b-1)^{-1}, \] (2.43)
which is linear and the failure rate function,
\[ h(x) = P[X=x]/P[X>x], \]
\[ = (a+b)(a+x)^{-1}, \] (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,

$$P[X = x] = \binom{-a}{x} \binom{-b}{n-x} / \binom{-a-b}{n}, \quad x = 0, 1, \ldots, n,$$  \hspace{1cm} (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

$$E(X|X>t) = ct, \quad (2.46)$$

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($\alpha, \alpha$) by Laurent (1974). His result is that a mean residual life function of the form

$$r(x) = a+bx, \quad x > 0, \quad (2.47)$$

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
\[ V_1(x) = E(X|X>x) \]  
(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
\[ h(x) = \left[ 1 + dr/dx \right] \left[ r(x) \right]^{-1}, \]
(2.49)
\[ V_1'(x) = r'(x) + 1 \]
and
\[ V_1'(x) = r(x) h(x), \]
(2.50)
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 $(a, \beta)$ 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

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

or

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

or negative non-integer satisfying,

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

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

\[M(S|t) = R^{-1}(\frac{1}{2} R(t)) - t.\]

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 $P_{1}$. 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

$$H_{x}(c) = [E\left\{ \left( \frac{X}{\mu_{c}} \right)^{r+1} | X > c \right\} - 1] \left[ r(r+1) \right]^{-1}, \quad -\infty < r < \infty (2.51)$$

where

$$\mu_{c} = E(X | X > c),$$

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

$$f(x) = \begin{cases} \frac{k C_{L}^{k}}{x^{k+1}}, & 0 < C_{L} < x, \ k > 1, \\ 0, & \text{otherwise}. \end{cases} (2.52)$$
(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

\[
E(Y^r | Y < c) = E(cY/k)^r, \quad 0 < k < c, \quad (2.53)
\]

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

\[
h(r) = \begin{cases} \frac{p^r}{p-1}, & 0 < r < 1, \quad p > 0, \\ 0, & \text{otherwise} \end{cases} \quad (2.54)
\]

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 \((x_0, \infty)\).
(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

$$E(Z|Y=RX=y) = \begin{cases} 
\alpha + \lambda y, & y > x_0, \ x_0 > 0, \ \lambda(\lambda-\eta)^{-1} > p, \\
\text{constant,} & y_0 \leq x_0,
\end{cases} \quad (2.55)$$

if and only if $X$ has a Pareto II distribution on $(x_0, \infty)$. 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 underreporting 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 \( Y = X - R \), \( 0 < R < \max (O, x-m) \), where \( m \) is the tax-exemption level and

\[
E(R | x > m) = a + bx, \quad 0 < b < 1 \quad (2.56)
\]

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

\[
E(R | X > y) = c + dy; \quad d > b > 0 \quad (2.57)
\]

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
\[ P[X > s | X > t] = P[X > s] \] (2.58)
for all \( s > 1 \). 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 \( c \) 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,
\begin{equation}
k(z) = \frac{f(z)}{[Z E(\frac{1}{Z})]}, \tag{2.59}
\end{equation}

and
\begin{equation}
L(\omega) = \frac{[1-F(\omega)]}{\mu}, \mu = E(X), \tag{2.60}
\end{equation}

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

(2) If Y is distributed with density
\begin{equation}
g(y) = \int_{a}^{y} x^{y-2} f(x) dx, \quad y > a \tag{2.61}
\end{equation}

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), Govindarajalu (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,

$$p_r = P[X=r], \quad r=1,2,\ldots, (2.62)$$

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

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

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

$$= \frac{\sum x p_x / y(y+1)}{\Sigma x p_x / y(y+1), \quad y=1,2,\ldots,}$$

$$= (r+a) p_r(\mu+a), \quad r=1,2,\ldots, (2.64)$$

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

$$p_x' = P[X=r], \quad r=1,2,...,m,$$

and defines a new random variable $Y$ such that,

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

$$q_y''' = \sum_{x=1}^{y} \frac{x p_x'}{y(y+1)}, \quad y=1,2,...,m,$$

$$h(x) = (1/q_x) - 1, \quad x=0,1,2,...,m,$$

$$h(x) = \begin{cases} \sum_{x=1}^{\gamma} \frac{x p_x'}{y(y+1)}, & y=m+1, \\ \sum_{x=1}^{\gamma} \frac{x p_x'}{y(y+1)}, & y=1,2,...,m \end{cases}$$

for $\gamma > 0$. Waring (2.38) for $b > 0$ and
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 distribu-
tions in the discrete time domain consisting of the geometric,
Waring (2.38) and the negative hyper-geometric (2.45) with
probability function,

\[
P[X=x] = P_x, \quad x = 0, 1, 2, \ldots
\]

(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 \([0, 1, 2, \ldots, m]\), \( m \in \{0, 1, \ldots\} \cup \{+\infty\} \) by
assuming that for \( 0 < P[X < 0] < 1 \), the failure rate
function is of the form

\[
h(x) = (a+bx)^{-1}, \quad x = 0, 1, 2, \ldots, m
\]

(2.69)
to be geometric for \( b = 0 \), Waring (2.38) for \( b > 0 \) and
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.

\[ g(y) = \frac{1}{m} R(y), \quad y > 0, \quad (3,1) \]

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 \( X_1, i = 1, 2, 3, \ldots \) 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).