1.1. Introduction

The concept of entropy is extensively used in literature as a quantitative measure of uncertainty associated with a random phenomena. The development of the idea of entropy by Shannon (1948) provided the beginning of a separate branch of learning namely the ‘Theory of information’. Historically a glimpse of the concept of entropy is available in an early work by Boltzman (1870) in connection with his studies related to the thermodynamic state of a physical system. Hartley (1928) used the entropy measure to ascertain the transmission of information through communication lines. Even though an axiomatic foundation to this concept was laid down by Shannon, this measure was developed in an independent context by Weiner (1948). Earlier work in connection with Shannon’s entropy was centered around characterizing the same

In the reliability context, if $X$ is a random variable representing the life time of a component or a device, a characteristic of special interest in the residual life distribution which is the distribution of the random variable $(X-t)$ truncated at $t(>0)$. A comparison of the residual life distribution and the parent distribution as well as characterization of distributions based on the form of the residual life time distributions has received a lot of interest among researchers. The works of Gupta and Gupta (1983), Gupta and Kirmani (1990) and Sankaran (1992) focuses attention on this aspect.

It is common knowledge that highly uncertain components or svstems are inherently not reliable. At the stage of designing a system, when there is enough information regarding the deterioration, wear and tear of component parts, factors and levels are prepared based on this information. Concepts such as failure
rate and the mean residual life function comes up as a handy tool in such situations. However in order to have a better system, the stability of the component parts should also be taken into account along with deterioration. Recently Ebrahimi and Pellerey (1995) and Ebrahimi (1996) has used the Shannon’s entropy applied to the residual life, referred to in literature as the residual entropy function, as a measure of stability of a component or a system. Because of the above the residual entropy function can be advantageously used as a useful tool at the stage of design and planning in Reliability Engineering.

The measurement and comparison of income among individuals in a society is a problem that has been attracting the interest of a lot of researchers in Economics and Statistics. In addition to the common measures of income inequality such as variance, coefficient of variation, Lorenz curve, Gini index etc, the Shannon’s entropy has been advantageously used as a handy tool to measure income inequality. The utility of this measure is highlighted in the works of Theil (1967) and Hart (1971). Ord, Patil and Taillie (1983) has used the truncated form of the entropy measure as a measure for examining the inequality of income of persons whose income exceeds a specified limit.
One of the main problems encountered in the analysis of statistical data is that of locating an appropriate model followed by the observations. Empirical methods such as probability plots or goodness of fit procedures fails to provide an exact model. However a characterization theorem enables one to determine the distribution uniquely in the sense that under certain conditions a family $F$ of distributions is the only one possessing a specified property. Accordingly characterization theorems are developed in respect of most of the distributions.

The commonly used life time models in Reliability Theory are exponential distribution, Pareto distribution, Beta distribution, Weibull distribution, and Gamma distribution. Several characterization theorems are obtained for the above models using reliability concepts such as failure rate, mean residual life function, vitality function, variance residual life function etc. Cox (1962), Guerrieri (1965), Reinhardt (1968), Shanbhag (1970), Swartz (1973), Laurent (1974), Vartak (1974), Dallas (1975), Nagaraja (1975), Morrison (1978), Gupta (1981), Gupta and Gupta (1983), Mukherjee and Roy (1986), Osaki and Li (1988) etc provide characterization results for the above distributions using reliability
concepts. An excellent review of works in the area is given in Galambos and Kotz (1978) and Azlarov and Volodin (1986).

Most of the works on characterization of distributions in the reliability context centers around the failure rate or the mean residual life function. However only very little work seems to have been done in using the residual entropy function as the criteria for characterization. Since the residual entropy function determines the distribution uniquely, a characterization theorem involving this concept will enable one to determine the model uniquely through a knowledge of its functional form. Motivated by this fact, the present study focuses attention on characterization of probability distributions based on (1) the form of the residual entropy function and (2) relationships between the residual entropy function and other reliability measures.

1.2 Review of literature

In this section we give a brief outline of the basic concepts in Information Theory and Reliability Theory that are of use in the investigations that are carried out in the succeeding chapters.
The Shannon's entropy

As pointed out in the introduction the Shannon's entropy have been extensively used as a quantitative measure of uncertainty. Consider a random experiment having \( n \) mutually exclusive events \( A_k, k = 1, 2, \ldots, n \) with respective probabilities \( p_k, k = 1, 2, \ldots, n \) satisfying the conditions \( p_k \geq 0 \) and \( \sum_{k=1}^{n} p_k = 1 \). One can represent such a probability space by a complete finite scheme (CFS),

\[
\begin{pmatrix}
A \\
p
\end{pmatrix} = \begin{pmatrix}
A_1 & A_2 & \cdots & A_n \\
p_1 & p_2 & \cdots & p_n
\end{pmatrix}.
\]

A CFS contains an amount of uncertainty about the particular outcome which will occur when the experiment is performed. As the probability associated with an event, \( A_k \), increases the uncertainty associated with that event decreases and so the amount of information conveyed by the occurrence of the event decreases. In a CFS there are different events and so different amount of information corresponding to these events. Hence the average amount of information can be taken as a measure of uncertainty associated with a CFS. Based on the notion, Shannon (1948) used the quantity,

\[
H_n(p) = - \sum_{i=1}^{n} p_i \log p_i \tag{1.1}
\]
as a quantitative measure of uncertainty associated with a CFS. As a convention 0 log 0 is taken as as zero. If we consider a random experiment with $n$ possible outcomes having probabilities $p_1, p_2, \ldots, p_n$, then (1.1) measures the uncertainty concerning the outcome of experiment. On the other hand, if we consider (1.1) after the experiment has been carried out then it measures the amount of information conveyed by the complete finite scheme.

The Shannon's entropy defined by (1.1) satisfies the following properties [Guiasu (1977)].

1. $H_n(P) \geq 0$, $P = (p_1, p_2, \ldots, p_n)$

2. $H_n(p_1, p_2, \ldots, p_n)$ is a continuous function of $p_1, p_2, \ldots, p_n$.

3. $H_n(p_1, p_2, \ldots, p_n)$ is a symmetric function of $p_1, p_2, \ldots, p_n$.

4. If $p_{i_0} = 1$ and $p_i = 0$ ($1 \leq i \leq n$, $i \neq i_0$) then $H_n(p_1, p_2, \ldots, p_n) = 0$

5. $H_{n+1}(p_1, p_2, \ldots, p_n, 0) = H_n(p_1, p_2, \ldots, p_n)$.

6. For any probability distribution with $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$,

   $$H_n(p_1, p_2, \ldots, p_n) \leq H_n\left(\frac{1}{n}, \ldots, \frac{1}{n}\right).$$

7. For any two independent probability distributions $P = (p_1, p_2, \ldots, p_n)$, $Q = (q_1, q_2, \ldots, q_m)$ where $\sum_{i=1}^{n} p_i = 1$, $\sum_{j=1}^{m} q_j = 1$,

   $$H_{n \cdot m}(P \cup Q) = H_n(P) + H_m(Q)$$
8. If the two schemes are not independent and \( P(A_i, B_j) = p_{ij} \), then

\[
H_{n+m}(P \cup Q) = H_n(P) + \frac{1}{m} \sum_{i=1}^{m} p_i H'(Q),
\]

where

\[
H'(Q) = -\sum_{j=1}^{n} P(B_j|A_i) \log P(B_j|A_i).
\]

In the continuous set up if \( f(.) \) denotes the probability density function associated with a random variable \( X \) defined in the interval \([a, b]\), then the continuous analogue of (1.1) turns out to be the Boltzman's \( H \) function is given by

\[
H_f = - \int_{a}^{b} f(x) \log f(x) \, dx. \tag{1.2}
\]

It may be noted that (1.2) is not the limit of the finite discrete entropies corresponding to a sequence of finer partition of the interval \([a, b]\) when the norms tend to zero.

Another important aspect of interest in the study of entropy is that of locating distributions for which the Shannon's entropy is maximum subject to certain restrictions on the underlying random variable. Depending on the conditions imposed, several maximum entropy distribution are derived. For instance, for a random variable in the support of non-negative real numbers, the maximum entropy probability distribution under the condition that the
arithmetic mean is fixed is the exponential distribution. The rationale behind the study of maximum entropy principle is that the probability distributions desired has maximum uncertainty subject to some explicitly stated known information. The books by Kapur (1989, 1994) gives a review of the various maximum entropy models.

The Shannon's entropy finds applications in several branches of learning. In communication theory an aspect of interest is the flow of information in some network where information is carried from a transmitter to receiver. This may be sending of messages by telegraph, flow of electricity, visual communications from artist to viewers etc. Things which tends to make errors in the transmission is called noise and in general message cannot be transmitted with complete reliability because of the effect of noise. In a source with a finite number of messages, \( \{x_k\}, k = 1, 2, \ldots, n \), the source selects each of the messages at random with probabilities \( p(x_k) \) and the amount of information associated with the transmission of \( x_k \) is \( - \log p(x_k) \). The average information per message for the source is

\[
I = - \sum_{k=1}^{n} p(x_k) \log p(x_k).
\]
This is referred to as the entropy of the source. This aspect in communication theory was studied by several researchers such as Fadeev (1956), Ash (1957), Reza (1971) etc.

Another field of application of Shannon's entropy is Economics, in connection with measurement of income inequality. If there are \( N \) individuals in a society, there are \( N \) non-negative amounts of individual income which adds up to the total income. Each of the individual earns non-negative fractions \( y_1, y_2, \ldots, y_N \) of total income where \( y_i \)'s are non-negative numbers which add up to 1. When there is equality of income \( y_1 = y_2 = \ldots = y_N = 1/N \) and in the case of complete inequality \( y_i = 1 \) for some \( i \) and zero for each \( i \neq j \). The quantity

\[
H(y) = \sum_{i=1}^{n} y_i \log \left( \frac{1}{y_i} \right)
\]

is the entropy of income shares. When there is complete equality \( H(y) \) is maximum with value \( \log N \). A measure of income inequality due to Theil (1967), is

\[
\log N - H(y) = \sum_{i=1}^{n} y_i \log(Ny_i).
\]
Ord, Patil and Taillie (1983) points out that the main draw back of the above measure is that it is scale dependent and location invariant.

Tilanus and Theil (1965) and Theil (1967) discusses how the entropy concept can be used to forecast input output structures. Cozzolino and Zaheer (1973) have used the principle of maximum entropy for the prediction of future market price of a stock. Golan, Judge and Miller(1996) give a new set of generalized entropy techniques designed to recover information about economic systems by extending the maximum entropy principle.

1.3 Some basic concepts in Reliability

The basic concepts in Reliability Theory, which are extensively studied, are (1) the reliability function (2) the failure rate and (3) the mean residual life function. If \( X \) is a random variable representing the life time of a device, the reliability function (survival function) of \( X \), defined by

\[
\bar{F}(t) = P(X > t), \quad t \geq 0
\]

represents the probability of failure free operation of the device at time \( t(\geq 0) \). Also

\[
\bar{F}(t) = 1 - F(t).
\]
where $F(t)$ is the distribution function of the random variable $X$.

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

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

for $x < L$, the failure rate (hazard rate) is defined as

$$h(x) = \frac{f(x)}{F(x)} = -\frac{d\log F(x)}{dx}. \quad (1.4)$$

In the general case, for a random variable $X$ with support $-\infty < X < \infty$, Kotz and Shanbhag (1980) defines the failure rate as the Radon-Nikodym derivative with respect to Lebesgue measure on $\{x : F(x) < 1\}$, of the hazard measure

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

for every Borel set $B$ of $(-\infty, L)$. Further the distribution of $X$ is uniquely determined through the relationship

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

where $H_c$ is the continuous part of $H$. When $X$ is a non-negative random variable admitting an absolutely continuous distribution function, then (1.5) reduces to

$$F(x) = \exp \left\{-\int_0^x h(t)dt \right\}. \quad (1.6)$$
It is well known that $h(x)$ determines the distribution uniquely and that the constancy of $h(x)$ is characteristic to the exponential model [Galambos and Kotz (1978)]. Further, for a random variable $X$ in the support of non-negative real numbers, a failure rate function of the form

$$h(x) = (ax+b)^{-1}$$  \hspace{1cm} (1.7)

characterizes the Exponential distribution specified by

$$\bar{F}(x) = e^{-\lambda x}, \ x>0, \ \lambda>0$$  \hspace{1cm} (1.8)

if $a=0$, the Pareto distribution specified by

$$\bar{F}(x) = \alpha^k (x+\alpha)^{-k}, \ x>0, \ \alpha>0, \ k>0$$  \hspace{1cm} (1.9)

if $a>0$, and the Beta distribution specified by

$$\bar{F}(x) = R^{-e} (R-x)^{e}, \ 0<x<R, \ e>0$$  \hspace{1cm} (1.10)

if $a<0$.

In the discrete set up, Xekalaki (1983) defines the failure rate for a random variable $X$ in the support of non-negative integers as

$$h(x) = \frac{P(X=x)}{P(X \geq x)}.$$  \hspace{1cm} (1.11)

It is established that $h(x)$ determines the distribution uniquely through the formula

$$\bar{F}(x) = \prod_{y=0}^{x-1} [1 - h(y)].$$  \hspace{1cm} (1.12)
Further it is shown that if $X$ is a random variable in the support of the set $\{0, 1, 2, \ldots\}$ then a relation of the form

$$h(x) = (px + q)^{-1}$$  \hspace{1cm} (1.13)

holds if and only if $X$ follows the Geometric distribution specified by

$$F(x) = q^x, \quad x = 0, 1, 2, \ldots, \quad 0 < q < 1$$  \hspace{1cm} (1.14)

if $p = 0$, the Waring distribution specified by

$$F(x) = \frac{(b)^x}{(a)^x}, \quad x = 0, 1, 2, \ldots, \quad a, b > 0$$  \hspace{1cm} (1.15)

if $p > 0$, and the Negative hyper geometric distribution specified by

$$F(x) = \binom{-1}{x} \binom{-k}{n-x} \binom{x}{n-k}, \quad x = 0, 1, 2, \ldots, n, \quad k > 0$$  \hspace{1cm} (1.16)

if $p < 0$.

For a continuous random variable $X$ with $E(X) < \infty$, the mean residual life function is defined as the Borel measurable function

$$r(x) = E(X - x | X \geq x),$$  \hspace{1cm} (1.17)

for all $x$ such that $F(x) > 0$. If $X$ is absolutely continuous, $r(x)$ can also be expressed as

$$r(x) = \frac{1}{F(x)} \int_x^\infty \frac{F(t)}{F(x)} dt.$$  \hspace{1cm} (1.18)
The following relationship between failure rate and the mean residual life function is immediate.

\[ h(x) = \frac{1 + r'(x)}{r(x)} \quad (1.19) \]

Also the mean residual life function determines the distribution uniquely through the relationship

\[ \bar{F}(x) = \exp \left\{ -\frac{\int_0^x \frac{dt}{r(t)}}{r(x)} \right\} \quad (1.20) \]

for every \( x \) in \((0, l)\). A set of necessary and sufficient condition for \( r(x) \) to be a mean residual life function, given by Swartz (1973), is that along with (1.20), the following conditions holds

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

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

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

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

Kupka and Loo (1989) defines the vitality function as the Borel measurable function on the real line given by

\[ m(x) = E(X|X\geq x). \quad (1.21) \]

The vitality function satisfies the properties

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

(ii) \( m(x) \geq x \) for all \( x < l \)
(iii) \[ \lim_{x \to \infty} m(x) = L \]

(iv) \[ \lim_{x \to \cdot} m(x) = E(x) \]

Moreover

\[ m(x) = x + r(x) \quad (1.22) \]

and

\[ m'(x) = r(x)h(x). \quad (1.23) \]

Cox (1972) established that the mean residual life function is constant for the exponential distribution. Mukherjee and Roy (1986) observed that a relation of the form

\[ r(x)h(x) = k \quad (1.24) \]

where \( k \) is a constant, holds if and only if \( X \) follows the exponential distribution specified by (1.8) when \( k=1 \), the Pareto distribution specified by (1.9) when \( k>1 \) and the Beta distribution specified by (1.10) when \( k<1 \). The Pareto case is also established in Sullo and Rutherford (1977). In view of (1.19), (1.24) reduces to

\[ r(x) = (k-1) + c, \quad (1.25) \]

where \( c = r(0) = E(X) \). Hence a linear mean residual life function of the form

\[ r(x) = ax + b \quad (1.26) \]

is characteristic to the exponential distribution specified by (1.8) if \( a=0 \), the Pareto distribution specified by (1.9) if \( a>0 \) and the Beta distribution specified by (1.10) if \( a<0 \).
For a discrete random variable $X$, in the support of the set of non-negative integers, the mean residual life function is defined as

$$r(x) = E(X-x|X\geq x)$$

$$= [ \tilde{F}(x + 1) ]^{-1} \sum_{y=x+1}^{\infty} \tilde{F}(y). \quad (1.27)$$

The mean residual life function determines the distribution uniquely through the relation

$$\tilde{F}(x) = \prod_{y=1}^{x-1} \frac{r(y-1)-1}{r(y)}[1-f(0)] \quad (1.28)$$

where $f(0)$ is determined such that $\sum f(x) = 1$. Further

$$1- h(x) = \frac{r(x)-1}{r(x+1)}, \quad x = 0, 1, 2, \ldots. \quad (1.29)$$

Nair (1983) discusses the notion of memory of life distributions by using mean residual life function and also classify life time distributions as those possessing no memory, negative memory and positive memory. Salvia and Bollinger (1982), Ebrahimi (1986), Guess and Park (1988), Abouammoh (1990), Hitha (1991), Roy and Gupta (1992), Mi (1993) also discuss the monotone behaviour of discrete reliability characteristics such as failure rate and mean residual life function.

Gupta and Gupta (1983) defines the moments of the residual life distribution through the relation
\[ m_r(x) = E[(X-x)^r | X>x] \] (1.30)

and obtains a recurrence relation satisfied by them. Further it is established that in general one higher moment does not determine a distribution uniquely and that the ratio of two higher moments will be required to do so. As a special case, the variance residual life function is

\[ V(x) = V(X-x | X \geq x) = E[(X-x)^2 | X \geq x] - r^2(x). \] (1.31)

This concept was introduced by Launer (1984) in order to define certain new classes of life distributions and to provide bounds for the reliability function for certain specified class of distributions. Gupta and Kirmani (1987) has established the following relations

\[ V(x) = \frac{2}{F(x)} \int_x^\infty r(t)F(t)dt - r^2(x) \] (1.32)

and

\[ \frac{dV(x)}{dx} = h(x) [V^2(x) - r^2(x)]. \] (1.33)

1.4 The residual entropy function

For a continuous non-negative random variable \( X \), representing the life time of a component, Ebrahimi (1996) defines the residual entropy function as the Shannon’s entropy associated with the random variable \( (X-t) \) truncated at \( t(>0) \), namely,
\[ H(t, t) = - \int_t^\infty f(x) \log \frac{f(x)}{F(t)} \, dx, \quad F(t) > 0. \tag{1.34} \]

(1.34) can also be written as

\[ H(t, t) = \log F(t) - \frac{1}{F(t)} \int_t^\infty f(x) \log f(x) \, dx. \tag{1.35} \]

The residual entropy function can be expressed in terms of the hazard rate through the relation

\[ H(t, t) = 1 - \frac{1}{F(t)} \int_t^\infty f(x) \log h(x) \, dx. \tag{1.36} \]

\( H(t, t) \) measures the expected uncertainty contained in the conditional density of \( (X-t) \) given \( X>t \) about the predictability of remaining life time of the component. It may be noticed that \( -\infty < H(t, t) \leq \infty \) and that \( H(t, 0) \) reduces to Shannon's entropy defined over \((0, \infty)\). It is established that \( H(t, t) \) determines the distribution uniquely. Also

\[ H'(t, t) = h(t)[ H(t, t) + \log h(t) - 1] \tag{1.37} \]

and

\[ H''(t, t) = h'(t)[ H(t, t) + \log h(t)] + H'(t, t) h(t). \tag{1.38} \]

Given \( r(t) \), if the domain is limited to a half line, the maximum entropy occurs for the exponential distribution with mean \( r(t) \). Therefore

\[ H(t, t) \leq 1 + \log r(t). \tag{1.39} \]
It can be easily verified that the maximum entropy distribution of \((X-t)\) truncated at \(t (>0)\) subject to the condition that the arithmetic mean is fixed is the exponential distribution. From (1.42), the finiteness of \(H(f,t)\) is guaranteed whenever \(r(t) < \infty\). It also provide a useful upper bound for \(H(f,t)\) in terms of the mean residual life function \(r(t)\). However, if additional information in terms of the variance residual life function \(V(t)\) or equivalently, in terms of the residual coefficient of variation \(v_{f}(t) = \nu(t)/r(t)\), is available, Ebrahimi and Kirmani (1996a) has proposed a better bound for \(H(f,t)\) as follows. Suppose \(E(X^2) < \infty\), then

\[
H(f, t) \leq \frac{1}{2} \theta_0^2 r^2(t) + \log\left\{2\pi \right\} + \left( \frac{1}{\theta_0} \right) r(t) \Phi(-\theta_0)
\]

where \(\theta_0\) is the solution of the equation

\[
\theta^2 r^2(t) = 1 + \psi(-\theta),
\]

where \(\psi(x) = x \phi(x)/\Phi(x)\), \(\Phi(x) = 1 - \phi(x)\) and \(\phi\) and \(\Phi\) are the density and the distribution function respectively of the standard normal distribution.

Ebrahimi (1996) has also proved the following results

1. If \(\overline{F}\) is an increasing (decreasing) failure rate distribution [IFR (DFR)] then it is also a decreasing uncertainty residual life
(increasing uncertainty residual life) [DURL (IURL)] distribution

2 Let \( \bar{F} \) be a DURL(IURL) then

\[
h(t) \leq \exp\{1-H(f, t)\}, \quad t > 0.
\]

3 Let \( \bar{F} \) be a DURL (IURL) then

\[
H(f, t) \leq (\geq) 1 - \log h(0) = 1 - \log f(0).
\]

4 Let \( \bar{F} \) be a DURL, then

\[
H(f, t) \leq 1 + \log r(0)
\]

and \( \bar{F} \) be a IURL, then

\[
\exp\{H(f, 0) - 1\} \leq r(t).
\]

He has also established there is no relationship between IURL (DURL) class of distributions and the class of increasing failure rate in average (IFRA) distributions. Subsequently Ebrahimi and Kirmani (1996a) has extended 1 to the family of decreasing mean residual life (increasing mean residual life) distributions, DMRL (IMRL). Further, Ebrahimi and Pellerey (1995) used the residual entropy function to introduce a new partial ordering for comparing the uncertainties associated with two non-negative random variables.
Recently Sankaran and Gupta (1999) has proved the following characterization results using the functional form of the residual entropy function.

(i) If $X$ is a non-negative random variable admitting absolutely continuous distribution function, the residual entropy function of the form

$$H(f, t) = \log(a + bt), \quad a > 0$$

characterizes the Exponential distribution with survival function (1.8) if $b = 0$, the Pareto distribution with survival function (1.9) if $b > 0$ and the Beta distribution with survival function (1.10) if $b < 0$.

(ii) A relation of the form

$$H(f, t) = 1 + \log r(t)$$

holds if and only if $X$ follows the exponential distribution.

(iii) A relation of the form

$$H(f, t) = a - \log h(t)$$

holds if and only if $X$ follows the Exponential distribution with survival function (1.8) if $a = 1$, the Pareto distribution with survival function (1.9) if $a > 1$ and the Beta distribution with survival function (1.10) if $a < 1$.

(iv) If $g(t) = \mathbb{E}(-\log X | X > t)$, then a relationship of the form

$$H(f, t) = cg(t) + d, \quad c > 0$$

(1.43)
holds if and only if $X$ follows the Weibull distribution with survival function specified by
\[
\bar{F}(x) = e^{-ae^{bt}}, \quad a>0, \quad b>0, \quad t>0.
\] (1.44)

Further they have extended the concept of residual entropy function to the entire real line and has established the following characterization theorem of extreme value distribution.

If $X$ is a random variable defined over the real line then the residual entropy function of the form
\[
H(f, t) = a m(t) + b
\] (1.45)
where $m(t) = E(X|X>t)$, characterizes the extreme value distribution with survival function
\[
\bar{F}(x) = e^{-pe^{qt}}, \quad -\infty < t < \infty.
\]

1.5 Discrimination between two residual lifetime distributions

Kullback and Leibler (1951) has extensively studied the concept of directed divergence which aims at discrimination between two populations. An axiomatic foundation to this concept was laid down by Aczel and Daroczy (1975). Kannappan and Rathie (1973) has obtained some characterization results based on the directed divergence. The concept of generalized directed divergence is discussed by Kapur (1968) and Rathie (1971).
Let \( P = (p_1, p_2, \ldots, p_n) \), \( Q = (q_1, q_2, \ldots, q_m) \) where \( \sum_{i=1}^{n} p_i = 1 \), \( \sum_{j=1}^{m} q_j = 1 \). be the two discrete probability distributions. Then a measure of directed divergence between \( P \) and \( Q \) is defined as

\[
I_\varepsilon(P, Q) = \sum_{i} p_i \log \frac{p_i}{q_i} \tag{1.46}
\]

If \( p_i = q_i \), then (1.46) reduces to zero. The continuous analogue to (1.46) turns out to be

\[
I(P, Q) = \int_{\mathbb{R}} f(x) \log \frac{f(x)}{g(x)} \, dx \tag{1.47}
\]

where \( f(x) \) and \( g(x) \) be the probability density functions corresponding to the probability measures \( P \) and \( Q \).

Let \( X \) and \( Y \) be non-negative random variables admitting absolutely continuous distribution functions \( F(x) \) and \( G(x) \) respectively, then (1.47) takes the form

\[
I(X, Y) = I(F, G) = \int_{0}^{\infty} f(x) \log \frac{f(x)}{g(x)} \, dx. \tag{1.48}
\]

Recently Ebrahimi and Kirmani (1996a) proposed a measure of discrimination between two residual life distributions based on (1.48) given by
\[ I(X, Y, t) = I(F, G, t) = \int \frac{f(x)}{F(t)} \log \left( \frac{f(x)}{F(t)} \right) dx = f' \log \{ f(x) / F(t) \} dx \]  

(1.49)

where \( F(t) = 1 - f(t) \) and \( G(t) = 1 - G(t) \). (1.49) can also be written as

\[ I(X, Y, t) = H(f, t) + \log G(t) - \int \frac{f(x)}{F(x)} \log g(x) dx. \]  

(1.50)

Further they have studied the properties of \( I(X, Y, t) \) and their implications.

According to the minimum discrimination information (MDI) principle, among the probability distributions satisfying the given constraints, one should choose that one for which directed divergence from a given prior distribution is minimum. Ebrahimi and Kirmani (1996a) has established that MDI principle when applied to modelling survival functions leads to the proportional hazard model, given in Cox (1972). If \( F(t) \) and \( G(t) \) are the survival functions of two random variables \( X \) and \( Y \) then a proportional hazards model for the survival functions exists if the relation

\[ G(x) = [F(x)]^\beta, \beta > 0, \]  

holds for all \( x \).

Ebrahimi and Kirmani (1996b) has further proved that the constancy of \( I(F, G, t) \) with respect to \( t \) is a characteristic property of the proportional hazard model.
The present thesis is organised into six chapters. After the present chapter which includes a brief review of literature on the topic, we look into the problem of characterizing probability distributions based on the form of the residual entropy function in Chapter II. Accordingly characterization theorems are established in respect of the Exponential distribution, Pareto distribution, Beta distribution and the Extreme value distribution. We devote Chapter III to the study of the residual entropy function of conditional distributions. Certain bivariate life time models such as bivariate exponential distribution with independent exponential marginals, Gumbel's bivariate Exponential distribution, bivariate Pareto distribution and bivariate Beta distribution are being characterized using this concept.

In Chapter IV we define the geometric vitality function and examine its properties. It is established that the geometric vitality function determines the distribution uniquely. Further characterization theorems in respect of some standard life time models are also obtained. The problem of averaging the residual entropy function is examined in Chapter V. Also the truncated form version of entropies of higher order are defined. Further we look into the problem of characterizing probability distributions using
the above concepts. Chapter VI is devoted to the study of the residual entropy function in the discrete time domain. It is established that in this case also the residual entropy function determines the distribution uniquely and that the constancy of the same is characteristic to the geometric distribution.