CHARACTERIZATION OF PROBABILITY DISTRIBUTIONS USING THE RESIDUAL ENTROPY FUNCTION

Thesis submitted to the
Cochin University of Science and Technology
for the Degree of
DOCTOR OF PHILOSOPHY
Under the Faculty of Science

By

RAJESH G.

Department of Statistics
Cochin University of Science and Technology
Cochin - 682 022

FEBRUARY 2001

CERTIFICATE

Certified that the thesis entitled "CHARACTERIZATION OF PROBABILITY DISTRIBUTIONS USING THE RESIDUAL ENTROPY FUNCTION" is a bonafide record of work done by Sri. Rajesh. G. under my guidance in the Department of Statistics, Cochin University of Science and Technology and that no part of it has been included any where previously for the award of any degree or title.

Cochin University of Science and Technology

February 1, 2001

Dr.K.R. Muraleedharan Nair

Professor of Statistics

CONTENTS

CHAPTER I
INTRODUCTION

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 based on different set of postulates. The works of Fadeev (1956), Khinchin (1957), Tverberg (1958), Chaundy and Mcleod (1960), Renyi (1961), Lee (1964), etc proceed in this direction. The classic monographs by Fisher (1958), Ash (1965), Aczel and Daroczy (1975) and Behra (1990) summarises most of the developments in this area.

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 systems 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, refered 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, ..., 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 $\sum_{i=1}^{n} p_i = 1$,

   $H_n(p_1, p_2, \ldots, p_n) \leq H_n\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right)$.
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 $P(A_iB_j) = p_{ij}$, then

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

where

$$H^i(Q) = -\sum_{j=1}^{m} P(B_j|A_i) \log P(B_j|A_i).$$

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

$$H_f = -\int_a^b f(x) \log f(x)\, dx. \qquad (1.2)$$

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 net work 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 upto 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 upto 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

$$\bar{F}(t) = P(X>t),\ t\geq 0 \qquad (1.3)$$

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

$$h(x) = \frac{f(x)}{\overline{F}(x)}$$

$$= -\frac{d\log \overline{F}(x)}{dx}. \qquad (1.4)$$

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

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

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

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

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

$$\overline{F}(x) = \exp\left\{-\int_0^x h(t)dt\right\}. \qquad (1.6)$$

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

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

characterizes the Exponential distribution specified by

$$\overline{F}(x) = e^{-\lambda x}, \quad x>0, \ \lambda>0 \tag{1.8}$$

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

$$\overline{F}(x) = \alpha^{k} \, (x+\alpha)^{-k}, \quad x>0, \ \alpha>0, \ k>0 \tag{1.9}$$

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

$$\overline{F}(x) = R^{-c} \, (R-x)^{c}, \quad 0<x<R, \ c>0 \tag{1.10}$$

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

$$h(x) = \frac{P(X=x)}{P(X \geq x)}. \tag{1.11}$$

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

$$\overline{F}(x) = \prod_{y=0}^{x-1} [1-h(y)]. \tag{1.12}$$

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

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

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

$$\overline{F}(x) = q^x,\ x = 0, 1, 2, \ldots,\ \ 0<q<1 \tag{1.14}$$

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

$$\overline{F}(x) = \frac{(b)_x}{(a)_x},\ x = 0, 1, 2, \ldots,\ \ a, b>0 \tag{1.15}$$

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

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

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

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

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

$$r(x) = \frac{1}{\overline{F}(x)}\int_x^\infty \overline{F}(t)dt. \tag{1.18}$$

The following relationship between failure rate and the mean residual life function is immediate.

$$h(x) = \frac{1+r'(x)}{r(x)} \qquad (1.19)$$

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

$$\bar{F}(x) = \frac{r(0)}{r(x)} \exp\left\{-\int_0^x \frac{dt}{r(t)}\right\} \qquad (1.20)$$

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

$$m(x) = E(X|X \geq x). \qquad (1.21)$$

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

(iv) $\lim_{x \to 0} m(x) = E(x)$

Moreover

$$m(x) = x + r(x) \tag{1.22}$$

and

$$m'(x) = r(x)h(x). \tag{1.23}$$

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

$$r(x)h(x) = k \tag{1.24}$$

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

$$r(x) = (k\text{-}1) + c, \tag{1.25}$$

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

$$r(x) = ax + b \tag{1.26}$$

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

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

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

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

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

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

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

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

$$m_r(x) = \mathrm{E}[(X-x)^r|X>x)] \tag{1.30}$$

and obtains a reccurence 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

$$V(x) = V(X-x|X\geq x)$$

$$= \mathrm{E}[(X-x)^2|X\geq x] - r^2(x). \tag{1.31}$$

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

$$V(x) = \frac{2}{\bar{F}(x)} \int_x^\infty r(t)\bar{F}(t)dt - r^2(x) \tag{1.32}$$

and

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

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,

$$H(f,\, t) = -\int_t^{\infty} \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)}{\bar{F}(t)}\, dx,\;\; \bar{F}(t)>0. \qquad (1.34)$$

(1.34) can also be written as

$$H(f,\, t) = \log\, \bar{F}(t) - \frac{1}{\bar{F}(t)} \int_t^{\infty} f(x) \log f(x)\, dx. \qquad (1.35)$$

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

$$H(f,\, t) = 1 - \frac{1}{\bar{F}(t)} \int_t^{\infty} f(x) \log h(x)\, dx. \qquad (1.36)$$

$H(f,t)$ 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 $-\infty \leq H(f,\, t) \leq \infty$ and that $H(f,\, 0)$ reduces to Shannon's entropy defined over $(0,\, \infty)$. It is established that $H(f,\, t)$ determines the distribution uniquely. Also

$$H'(f,\, t) = h(t)[\, H(f,\, t) + \log h(t) - 1] \qquad (1.37)$$

and

$$H''(f,\, t) = h'(t)[\, H(f,\, t) + \log h(t)] + H'(f,\, t)\, h(t). \qquad (1.38)$$

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

$$H(f,\, t) \leq 1 + \log r(t). \qquad (1.39)$$

It can be easily verified that the maximum entropy distribution of $(X\text{-}t)$ 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 $v_F(t) = v(t)/r(t)$, is available, Ebrahimi and Kirmani (1996a) has proposed a better bound for $H(f,t)$ as follows. Suppose $E(X^2) < \infty$, then

$$H(f,\, t) \le \tfrac{1}{2}\, \theta_0^2\, r^2(t) + \log\{(2\pi)^{1/2} \left(\frac{1}{\theta_0}\right) r(t)\, \Phi(-\theta_0)\}$$

where θ_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 ϕ and Φ 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

$$h(t) \leq (\geq) \exp\{1 - H(f, t)\},\ t > 0.$$

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,

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

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

$$H(f, t) = \log (a+bt),\ a>0 \qquad (1.40)$$

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

$$H(f, t) = 1+\log r(\mathrm{t}) \qquad (1.41)$$

holds if and only if X follows the exponential distribution.

(iii) A relation of the form

$$H(f, t) = a - \log h(\mathrm{t}) \qquad (1.42)$$

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

$$H(f, t) = cg(t) + d,\ c>0 \qquad (1.43)$$

holds if and only if X follows the Weibull distribution with survival function specified by

$$\bar{F}(x) = e^{at^{b}},\ a>0,\ b>0,\ t>0. \qquad (1.44)$$

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

$$H(f,\ t) = a\ m(t) + b \qquad (1.45)$$

where $m(t) = E(X|X>t)$, characterizes the extreme value distribution with survival function

$$\bar{F}(x) = e^{-pe^{qt}},\ -\infty<t<\infty.$$

1.5 Discrimination between two residual life time 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 $\sum_{i=1}^{n} p_i = 1$, $\sum_{i=1}^{m} q_i = 1$, be the two discrete probability distributions. Then a measure of directed divergence between P and Q is defined as

$$I_n(P,Q) = \sum_i p_i \log \frac{p_i}{q_i} \tag{1.46}$$

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

$$I(P,Q) = \int_{-\infty}^{\infty} f(x) \log \frac{f(x)}{g(x)} dx \tag{1.47}$$

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

$$I(X,Y,t) = I(F,G,t) = \int_t^{\infty} \frac{f(x)}{\overline{F}(t)} \log\left\{\frac{f(x)/\overline{F}(t)}{g(x)/\overline{G}(t)}\right\} dx \qquad (1.49)$$

where $\overline{F}(t) = 1\text{-}F(t)$ and $\overline{G}(t) = 1\text{-}G(t)$.(1.49) can also be written as

$$I(X,Y,t) = H(f,\ t) + \log \overline{G}(t)\ - \int_t^{\infty} \frac{f(x)}{\overline{F}(x)}\ \log g(x)\ dx. \qquad (1.50)$$

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 $\overline{F}(t)$ and $\overline{G}(t)$ are the survival functions of two random variables X and Y then a proportional hazards model for the survival functions exists if the relation.

$$\overline{G}(x) = [\,\overline{F}(x)\,]^{\beta}\ ,\ \beta>0, \text{ holds for all } x.$$

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 in to 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 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.

CHAPTER II

CHARACTERIZATION OF CONTINUOUS PROBABILITY DISTRIBUTIONS

2.1. Introduction

A conventional approach to characterize a life time distribution is by using the failure rate or the mean residual life function. The works of Kotz and Shanbhag (1980), Gupta (1981) and Mukherjee and Roy (1986) proceeds in this direction. As pointed out in Chapter I the residual entropy function, being a measure of the stability of a component, can be advantageously used to describe the physical characteristics of the failure mechanism and so a characterization theorem involving this concept helps one to determine the life time distribution through a knowledge of the form of the residual entropy function. The residual entropy function is evaluated for some standard distributions and is given as Appendix-I.

Some of the results mentioned in this chapter are being published in JISA Vol. 36, pp. 157-166.

2.2 Characterizations using relationship between the residual entropy function and the mean residual life function

Galambos and Kotz (1978) has observed that the three characteristic properties of the exponential distribution namely the lack of memory property, constancy of failure rate and constancy of mean residual life function are equivalent. In addition to the above if we add the property of constancy of the residual entropy function one can see that the four properties are equivalent. This is stated as Theorem 2.1 below

Theorem 2.1

Let X be a non-negative random variable admitting an absolutely continuous distribution function with finite mean. Then the following are equivalent.

(a) $h(t) = c$, where c is a constant

(b) $r(t) = r(0)$

(c) $P(X \geq t+s | X \geq s) = P(X \geq t)$

(d) $H(f, t) = H(f)$.

Proof:

We first show that (d) $\Leftrightarrow$ (a). When (d) holds we have

$$H'(f, t) = 0.$$

Using the expression for $H'(f\ t)$ given in (1.37) namely

$$H'(f,\ t)= h(t)\ [H(f,t)\ -\ 1+\ \log h(t)],$$

we get

$$h(t)\ [H(f,t)\ -\ 1+\ \log h(t)]=0$$

or

$$h(t)\ [H(f)\ -\ 1+\ \log h(t)]=0.$$

This gives either $h(t) = 0$ or $h(t) = c$, where $c = \exp[1-H(f)]$.

This is same as (a).

When (a) holds using the relation

$$H(f,t) = 1 - \frac{1}{\bar{F}(t)}\int_t^{\infty} f(x)\ \log h(x)\ dx,$$

we get

$$H(f,t) = 1- \log c = H(f)$$

which is same as (d). The rest of the proof follows from Galambos and Kotz (1978).

Since properties (a), (b) and (c) are characteristic to the exponential model, in view of Theorem 2.1 it may be observed that the constancy of the residual entropy function is characteristic to the exponential distribution.

Our next result provides a characterization theorem for the exponential distribution using a functional relationship between the residual entropy function and the mean residual life function.

Theorem 2.2

Let X be a non-negative random variable admitting an absolutely continuous distribution function such that $E(X)<\infty$. If $H(f,t)$ be the residual entropy function and $r(t)$ be the mean residual life function, then a relation of the form

$$H(f,t) - r(t) = H(f) - r(0) \tag{2.1}$$

holds for all real $t(\geq 0)$ if and only if X follows the exponential distribution.

Proof:

When (2.1) holds we have

$$H'(f,t) - r'(t) = 0. \tag{2.2}$$

Using (1.19) and (1.37) in (2.2) we get

$$h(t)[H(f,t) - r(t) - 1 + \log h(t)] = -1$$

or

$$h(t)[c-1 + \log h(t)] = -1, \tag{2.3}$$

where $c = H(f) - r(0)$. Differentiating (2.3) with respect to t and rearranging the terms we get

$$h'(t)\{c+\log h(t)\} = 0. \tag{2.4}$$

(2.4) gives either $h'(t) = 0$ or $\log h(t) = -c$. In either case $h(t)$ is a constant. Since the constancy of failure rate is characteristic to the exponential distribution the only if part of the theorem follows.

Conversely when X follows the exponential distribution with survival function

$$\bar{F}(t) = e^{-\lambda t}, \; t \geq 0, \; \lambda > 0,$$

by direct calculations we get

$$r(t) = \frac{1}{\lambda}$$

and

$$H(f,t) = 1 - \log \lambda .$$

(2.1) is immediate from the above expressions.

It may be observed that (2.1) can be written in the form

$$H(f) - H(f,t) = r(0) - r(t). \tag{2.5}$$

In connection with his study relating to memory of distributions, Muth (1980) defines the virtual age, $v(t)$, of a component at time t as

$$v(t) = r(0) - r(t).$$

Also when $v(t)$ =0, $r(t)$ = $r(0)$ and there is no memory. So Theorem 2.2 implies that the excess of entropy resulting from the functioning of the component upto time t is equal to the virtual age of the component if and only if the distribution is exponential.

The following theorem provides a characterization for a family of distributions using a possible relationship between the residual entropy function and the mean residual life function.

Theorem 2.3

For the random variable X considered in Theorem 2.2, the relation

$$H(f,t) - \log r(t) = k \tag{2.6}$$

where $k = H(f) - \log r(0)$, holds for all real t (≥ 0) if and only if X follows any one of the following three distributions

(i) the Exponential distribution with survival function

$$\bar{F}(x) = e^{-\lambda x}, \ x \geq 0, \ \lambda > 0, \tag{2.7}$$

(ii) the Pareto distribution with survival function

$$\bar{F}(x) = \left(\frac{\alpha}{x+\alpha}\right)^{a}, \ x \geq 0, \ a > 1, \ 0 < \alpha < \infty, \tag{2.8}$$

(iii) the Beta distribution with survival function

$$\bar{F}(x) = \left(1 - \frac{x}{R}\right)^{c}, \ 0 < x < R, \ c > 1. \tag{2.9}$$

Proof:

When (2.6) holds, we have

$$H'(f,t) = \frac{r'(t)}{r(t)}$$

Using (1.37) and (2.6) the above equation can be written as

$$h(t)\ r(t)\ [k-1+\log h(t)r(t)] = r'(t).$$

Writing $c(t) = h(t)r(t)$ and using (1.19) we get

$$c(t)\ [k-1+\log c(t)] = c(t) - 1. \tag{2.10}$$

Differentiating (2.10) with respect to t and rearranging the terms we get

$$c'(t)\ [k-2+\log c(t)] = 0 \tag{2.11}$$

(2.11) gives either

$$c'(t) = 0 \text{ or } c(t) = e^{1-k}.$$

In either case

$$h(t)r(t) = p, \tag{2.12}$$

where p is a constant satisfying (2.11).

From Mukherjee and Roy (1986), (2.12) characterizes the Exponential distribution for $p=1$, the Pareto distribution for $p>1$ and the Beta distribution for $p<1$. Hence X follows any one of the three distributions.

The if part of the theorem follows from the expression for $H(f,t)$ and $r(t)$ given below

Distribution	$r(t)$	$H(f,t)$
Exponential	λ^{-1}	$1-\log\lambda$
Pareto	$\dfrac{t+\alpha}{a-1}$	$1+\frac{1}{a}-\log\left(\dfrac{a}{t+\alpha}\right)$
Beta	$\dfrac{R-t}{c+1}$	$1-\frac{1}{c}-\log\left(\dfrac{c}{R-t}\right)$

Observing that (2.6) can be written as

$$r(t) = \mu \exp\{-[H(f) - H(f,t)]\},$$

where $\mu = E(X)$, we notice that for the above class of distributions the expected remaining life increases as the excess of entropy decreases and vice versa.

Recently, in connection with their study on ordering and asymptotic properties of residual income distributions, Belzunce, Candel and Ruiz (1995) consider the random variable

$$X(t) = \frac{X}{t}\big| X > t$$

to define a new class of distributions and used the proportional failure rate defined by

$$p(t) = th(t) \tag{2.13}$$

to derive a model for income distributions. We give below a characterization for the family of distributions considered in Theorem 2.3 using the above concept.

Theorem 2.4

Let X be a non-negative random variable admitting absolutely continuous distribution function such that $E(X)<\infty$. The relationship

$$H(f,t) + \log p(t) = \log kt \tag{2.14}$$

holds for all $t \geq 0$ if and only if X follows

(i) the Exponential distribution with survival function (2.7) if $k=e$,

(ii) the Pareto distribution with survival function (2.8) if $k>e$ and

(iii) the Beta distribution with survival function (2.9) if $k<e$.

Proof:

The proof is immediate, observing that (2.14) can be written as

$$H(f,t) + \log h(t) = \log k. \tag{2.15}$$

Differentiating (2.15) with respect to t we get

$$H'(f,t) = -\frac{h'(t)}{h(t)}.$$

Using (1.37) and (2.15), the above equation can be written as

$$-\frac{h'(t)}{h^2(t)} = \log k - 1. \qquad (2.16)$$

(2.16) gives

$$h(t) = [(\log k - 1)t + c]^{-1} \qquad (2.17)$$

where c is a constant. Inview of (1.7), X follows the Exponential distribution when k=e, the Pareto distribution when $k>e$ and the Beta distribution when $k<e$.

2.3 Characterizations using the residual entropy function when the support of the random variable is the real line

Kotz and Shanbhag (1980) has observed that the concept of failure rate and the mean residual life function can be used with out much difficulty if the support of X is the set of real numbers. They defined the failure rate as the Radon-Nikodym derivative with respect to Lebesgue measure on $\{x: F(x)<1\}$ of the hazard measure

$$H(x) = \int_B dF(x)/[1-F(x)],$$

for every Borel set B and the mean residual life function as a real valued Borel measurable function

$$r(x) = \mathrm{E}(X- x|X>x),$$

for all x such that $\mathrm{P}(X\geq x)>0$. Analogously for a continuous random variable X defined over R, we define the residual entropy function as

$$H(f,\ t) = -\int_{t}^{\infty} \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)}\ dx,\ -\infty<t<\infty$$

$$= 1 - \frac{1}{\overline{F}(t)} \int_{t}^{\infty} f(x) \log h(x)\ dx. \qquad (2.18)$$

The following theorem provides a characterization of the extreme value distribution using (2.18).

Theorem 2.5

Let X be a random variable in the support of R, admitting an absolutely continuous distribution such that $E(X)<\infty$ and let $H(f,t)$ be defined as in (2.18). The relation

$$H(f,t) + r(t) = 1-t \qquad (2.19)$$

holds for all real t if and only if X follows the type I extreme value distribution with survival function

$$\overline{F}(x) = e^{-e^{x}} \quad ,-\infty<x<\infty \qquad (2.20)$$

Proof:

When (2.19) holds we have

$$H(f,t) + \frac{1}{\overline{F}(t)} \int_{t}^{\infty} \overline{F}(x)dx = 1-t$$

or

$$\int_{t}^{\infty} \overline{F}(x)dx = [1-t-H(f,t)]\ \overline{F}(t).$$

Differentiating the above equation w.r.t. t and simplifying we get

$$[1 - t - H(f,t)]\,\overline{F}'(t) - H'(f,t)\,\overline{F}(t) = 0. \qquad (2.21)$$

Using (1.37), (2.21) can be written as

$$\log h(t) = t$$

or

$$h(t) = e^{t}.$$

Now the relationship

$$\overline{F}(x) = \exp\left\{-\int_{-\infty}^{x} h(t)dt\right\} \qquad (2.22)$$

gives

$$\overline{F}(x) = e^{-e^{x}}$$

as claimed.

Conversely when the distribution of X is specified by (2.20), using (2.18), we have

$$H(f,t) = 1 - \frac{1}{\overline{F}(t)} \int_{t}^{\infty} x e^{-e^{x}} . e^{x} dx$$

$$= 1 - t - \frac{1}{\overline{F}(t)} \int_{t}^{\infty} \overline{F}(x) dx$$

$$= 1 - t - r(t).$$

which is same as (2.19).

Gompertz (1825), in connection with his empirical studies on human mortality, consider a failure rate function (force of mortality) of the form

$$h(t) = pq^t, \quad t>0,\ p>0,\ q>1$$

and obtained a truncated form of type I extreme value distribution specified by

$$\overline{F}(t) = e^{-p(q^t-1)/\log q},\ t\geq 0 \tag{2.23}$$

as a model for life time data. In the light of Theorem 2.5 it is immediate that the relation

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

when $m(t) = E(X|X>t) = r(t) + t$, holds for all real $t\geq 0$ if and only if X follows (2.23) with $p=1$.

Theorem 2.6

For the random variable X considered in Theorem 2.5, with $\lim_{x\to\infty} f(x) = 0$, the relation

$$H(f,t) + r'(t) = a, \tag{2.24}$$

where a is a constant and $r'(t)$ denotes the derivative of $r(t)$, holds for all real t if and only if X follows the logistic distribution with survival function

$$\overline{F}(x) = \frac{ke^{-cx}}{1+ke^{-cx}}; \quad -\infty<x<\infty,\ c>0, k>0. \tag{2.25}$$

Proof:

When (2.24) holds we have

$$H'(f,t) = -r''(t) \tag{2.26}$$

Using (1.19) and (1.37), (2.26) can be written as

$$h(t)[H(f,t) + \log h(t) - 1] = -h(t)\, r'(t) - r(t)h'(t).$$

In view of (2.24) the above equation simplifies to

$$h(t)[\log h(t) + a - 1] = -r(t)h'(t)$$

or

$$\frac{1}{r(t)} = \frac{\frac{d}{dt}[\log h(t)]}{\log h(t) + a - 1}.$$

This gives

$$\frac{d}{dt}[\log \int_t^{\infty} \overline{F}(x)dx] = \frac{d}{dt}[\log(\log h(t) + a - 1)]. \tag{2.27}$$

From (2.27), we get

$$\int_t^{\infty} \overline{F}(x)dx = b\,[\log h(t) + a - 1],$$

where $\log b$ is the constant of integration. This above equation gives

$$-\overline{F}(t) = b\frac{h'(t)}{h(t)}$$

or

$$-f(t) = bh'(t).$$

Integrating from $-\infty$ to t and using the condition $\lim_{x \to -\infty} f(x) = 0$ we get

$$F(t) = -b\ h(t). \tag{2.28}$$

Since $F(t)$ and $h(t)$ are non-negative, for (2.28) to be valid we must have $b<0$. Thus (2.28) gives

$$f(t) = \alpha\ F(t)[1 - F(t)]$$

where $\alpha = -\frac{1}{b} > 0$. The rest of the proof follows from Galambos (1992).

Conversely when the distribution of X is specified by (2.25) by direct calculation we get

$$h(t) = \frac{c}{1+ke^{-ct}},$$

$$r(t) = \left(\frac{1+ke^{-ct}}{kce^{-ct}}\right)\log(1+ke^{-ct})$$

$$H(f,t) = 2 - \log c - \frac{\log(1+ke^{-ct})}{ke^{-ct}}$$

and

$$H(f,t) + r'(t) = 1 - \log c,$$

which is a constant, so that the conditions of the theorem holds.

In the sequel, we give characterization theorems for the type I extreme value distribution and the logistic distribution using functional relationships between failure rate, mean residual life function and the residual entropy function.

Theorem 2.7

For the random variable considered in Theorem 2.5, the relation

$$H(f,t) + r(t) = 1 - \log h(t), \tag{2.29}$$

holds for all real t if and only if X follows the type I extreme value distribution with survival function

$$\overline{F}(x) = e^{-\frac{1}{m}e^{x}}, \quad m \geq 0, \ -\infty < x < \infty. \tag{2.30}$$

Proof:

When (2.29) holds, we have

$$H'(f,t) + r'(t) = -\frac{h'(t)}{h(t)}. \tag{2.31}$$

Using (1.19) and (1.37), (2.31) can be written as

$$h(t)[k - r(t) - 1] + h(t)\, r(t) - 1 = -\frac{h'(t)}{h(t)}$$

or

$$-\frac{h'(t)}{h^2(t)} + \frac{1}{h(t)} = 0.$$

If $u(t) = \dfrac{1}{h(t)}$, the above equation takes the form

$$u'(t) + u(t) = 0$$

which is a linear differential equation whose solution is

$$u(t) = me^{-t}$$

or

$$h(t) = \frac{1}{m}\, e^{t},$$

with $m \geq 0$. From (2.22) we have

$$\overline{F}(x) = \exp\left\{\frac{-1}{m} e^{x}\right\}$$

The if part follows from the expressions for $h(t)$ and $r(t)$ namely

$$h(t) = \frac{1}{m}\, e^{t}$$

$$r(t) = e^{\frac{1}{m}e^{t}} \int_{t}^{\infty} e^{-\frac{1}{m}e^{x}}\, dx,$$

so that

$$H(f,\, t) = 1 - t - e^{\frac{1}{m}e^{t}} \int_{t}^{\infty} e^{-\frac{1}{m}e^{x}}\, dx + \log m$$

$$= 1 - \log\left\{\frac{1}{m} e^{t}\right\} - e^{\frac{1}{m}e^{t}} \int_{t}^{\infty} e^{-\frac{1}{m}e^{x}}\, dx$$

$$= 1 - \log h(t) - r(t).$$

which is same as (2.29).

Theorem 2.8

The relation

$$H(f,t) + r(t) = 2 - \log h(t)$$

holds for all real t if and only if X follow the logistic distribution with survival function (2.25).

The theorem follows by proceeding along the same lines as in the proof of Theorem 2.7, and so the proof is omitted.

2.4 Conditional measure of uncertainty

For a non-negative random variable X with probability density function $f(x)$ and survival function $\bar{F}(x)$, Sankaran and Gupta (1999) define a conditional measure of uncertainty as

$$M(f,\, t) = -\,\mathrm{E}(\log f(X) \mid X > t)$$

$$= -\,\frac{1}{\bar{F}(t)} \int_t^{\infty} f(x) \log f(x)\, dx. \qquad (2.32)$$

$M(f,\, t)$ is related to $H(f,\, t)$ through the relation

$$M(f,\, t) = H(f,\, t) - \log\, \bar{F}(t). \qquad (2.33)$$

Observing that

$$\int_0^t h(x)dx = -\,\log\, \bar{F}(t)$$

represents the total failure rate, (2.33) implies that the conditional measure of uncertainty is simply the sum of the residual entropy function and the total failure rate

We give below a characterization theorem for the exponential distribution using the conditional measure of uncertainty defined by (2.32).

Theorem 2.9

For a non-negative random variable X with $E(X)<\infty$, a relation of the form

$$M(f, t) - p(t) = k, \qquad (2.34)$$

where $p(t)$ is the proportional hazard rate defined by (2.13) and k is a constant, holds for all real $t(\geq 0)$ if and only if X follows the exponential distribution with survival function (2.7).

Proof:

When (2.34) holds, we have

$$M'(f, t) = p'(t) \qquad (2.35)$$

Using the relationship

$$M'(f, t) = h(t)\,[M'(f, t) + \log f(t)], \qquad (2.36)$$

(2.35) becomes

$$h(t)\,[M'(f, t) + \log f(t)] = t\,h'(t) + h(t)$$

or

$$h(t)\,[k + t\,h(t) + \log f(t)] = t\,h'(t) + h(t). \qquad (2.37)$$

Since

$$\frac{h'(t)}{h(t)} = \frac{f'(t)}{f(t)} + h(t)$$

(2.37) can be written as

$$\frac{d}{dt}\log f(t) - \frac{1}{t}\log f(t) = \frac{k-1}{t}$$

which is a linear differential equation in $\log f(t)$ whose solution is

$$f(t) = \lambda\, e^{-\lambda t}$$

where $\lambda = -e^{1-k}$ Thus X has the exponential distribution.

Conversely when the distribution of X is specified by (2.7), by direct calculation we get

$$M(f,\, t) = 1 - \log \lambda + \lambda t,$$

$$p(t) = \lambda t$$

and

$$M(f,\, t) - p(t) = 1 - \log \lambda,$$

which is same as in (2.34).

In the next theorem we look into the situation where $M(f,\, t)$ is a linear function of $h(t)$. Here we consider the situation where domain of X is the set of real numbers.

Theorem 2.10

For the random variable X considered in Theorem 2.5, the relation

$$M(f,\, t) = a + \tfrac{1}{2}\, t\, h(t). \tag{2.38}$$

with $a = \log \sqrt{2\pi\, e}$, holds for all real t if and only if X follows the standard normal distribution with probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad -\infty<x<\infty \qquad (2.39)$$

Proof:

When (2.38) holds, we have

$$M'(f, t) = ½ [t\, h'(t) + h(t)]. \qquad (2.40)$$

Using (2.36) and (2.38), (2.40) can be written as

$$a + \frac{t}{2} h(t) + \log f(t) = ½t \left[\frac{f'(t)}{f(t)} + h(t)\right] + ½$$

or

$$\frac{f'(t)}{f(t)} - \frac{2}{t} \log f(t) = (a - ½)\frac{2}{t} \qquad (2.41)$$

(2.41) is a linear differential equation in $\log f(t)$, whose solution is

$$\log f(t) = -(a - ½) + bt^2$$

where b is the constant of integration. Therefore

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}, \ -\infty<x<\infty,$$

observing that $b = -½$ for $f(t)$ to be a probability density function.

Conversely by direct calculations we have

$$h(t) = \frac{e^{-\frac{t^2}{2}}}{\Gamma\left(\frac{1}{2}, \frac{t^2}{2}\right)},$$

$$M(f, t) = \log\sqrt{2\pi} + ½ + \frac{t}{\sqrt{2}} \frac{e^{-\frac{t^2}{2}}}{\Gamma\left(\frac{1}{2}, \frac{t^2}{2}\right)},$$

so that

$$M(f, t) = a + \frac{t}{2} h(t).$$

The following theorem characterizes the standard normal distribution using a functional relationship between $M(f, t)$ and the second order moment of residual life.

Theorem 2.11

For the random variable consider in Theorem 2.5, the relation

$$M(f, t) - \tfrac{1}{2} E(X^2|X>t) = A \qquad (2.42)$$

where A is a specified constant, holds for all real t if and only if X follows the standard normal distribution with density function (2.39).

Proof:

When (2.42) holds, we have

$$\bar{F}(t)\, M(f, t) - \tfrac{1}{2} \int_t^{\infty} x^2 f(x)dx = A\, \bar{F}(t).$$

Differentiating the above equation with respect to t and simplifying we get

$$M'(f, t) - h(t)\, M(f, t) + \tfrac{1}{2}\, t^2\, h(t) = -\, Ah(t). \qquad (2.43)$$

Using (2.36) in (2.43) we get

$$\log f(t) = -\, A - \tfrac{1}{2}\, t^2.$$

Therefore

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

where $e^{c} = \frac{1}{\sqrt{2\pi}}$, since f is to be a probability density function.

Conversely if the distribution of X is specified by the density function (2.39), from (2.32), we have

$$M(f, t) = -\frac{1}{\bar{F}(t)} \int_t^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \log \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) dx$$

$$= \log \sqrt{2\pi} + \frac{1}{\bar{F}(t)} \int_t^{\infty} \left(\frac{x^2}{2} \right) \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

$$= \log \sqrt{2\pi} + \tfrac{1}{2} \, E(X^2 | X > t)$$

which is same as (2.42).

CHAPTER III

RESIDUAL ENTROPY OF CONDITIONAL DISTRIBUTIONS

3.1. Introduction

The fundamental constituents of a multivariate distribution are the marginal and the conditional distributions. It is well known that, except in the case of independence, the marginals does not determine the distribution uniquely. Conditions under which a set of marginal and conditional distributions uniquely specifies the distribution has been investigated by Abrahams and Thomas (1984), Hitha and Nair (1991) and Gelman and Speed (1993). In many real life phenomena information about the conditional densities are easily available. So the question of determining the joint distribution using specified conditionals has received considerable attention in recent times. The works of Arnold (1987), Arnold and Press (1988), Arnold and Strauss (1988) and Geetha and Nair(1997) proceed in this direction.

Some of the results in this chapter have appeared in the Statistical Methods (2000), Vol. 2(1), pp. 72-80.

For two random variables X and Y in the support of set of integers, the entropy of the conditional distribution of X given $Y=y$ is defined as

$$H(X|Y=y) = \sum_{x} p(x|y) \log p(x|y) \qquad (3.1)$$

where $p(x|y)$ is the conditional probability mass function of X given $Y=y$. When the support of X and Y are the set of real numbers, then (3.1) takes the form

$$H(X|Y=y) = -\int_{x} f(x|y) \log f(x|y) dx$$

where $f(x|y)$ is the conditional probability density function of X given Y. The conditional entropy of X given Y is defined as the weighted average of the entropy of the conditional distributions, namely

$$H(X|Y) = E_Y[H(X|Y=y)].$$

Analogous to (1.34) one can define the residual entropy of conditional distributions. The form of the same can be advantageously used to arrive at bivariate distributions which have applications in modelling life time data for systems having more than one component.

3.2 Definition

Let $X = (X_1, X_2)$ be a non-negative random vector admitting an absolutely continuous distribution with density function $f(x_1, x_2)$, survival function $\bar{F}(x_1, x_2)$, marginal density of X_i, $f_i(x_i)$, $i=1,2$, and conditional density of X_i given $X_j = x_j$, $g_i(x_i \mid x_j)$, $i, j = 1,2$, $i \neq j$. Using (1.34), the residual entropy function of the conditional distribution of X_i given $X_j = t_j$ turns out to be

$$H_i(g_i, t_1, t_2) = -\int_{t_i}^{\infty} \frac{g_i(x_i \mid t_j)}{\bar{G}_i(t_i \mid t_j)} \log \frac{g_i(x_i \mid t_j)}{\bar{G}_i(t_i \mid t_j)} dx_i, \; i,j = 1,2 \; i \neq j \qquad (3.2)$$

where $\bar{G}_i(t_i \mid t_j)$ is the conditional survival function of X_i given $X_j = t_j$.

If X represents the life time of the components in a two component system, $Y_1 = (X_1 - t_1)$ given $X_1 > t_1$, $X_2 = t_2$ corresponds to the residual life of the first component subject to the condition that it has survived up to time t_1 and that the second component has failed at time t_2. The Shannon's entropy corresponding to the distribution of Y_1 simplifies to (3.2) with $i=1$. Similar interpretation can be given for $H_2(g_2, t_1, t_2)$.

Since

$$\bar{F}(t_1, t_2) = \int_{t_1}^{\infty}\int_{t_2}^{\infty} f(x_1, x_2) dx_2 dx_1$$

we have

$$\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2} = -\int_{t_1}^{\infty} f(x_1,t_2)dx_1 .$$

Also

$$\frac{g_1(x_1|t_2)}{\bar{G}_1(t_1|t_2)} = - \frac{f(x_1,t_2)}{\int_{t_1}^{\infty} f(x_1,t_2)dx_1} .$$

Hence (3.2) can also be written as

$$H_1(g_1,t_1,t_2) = \log\left(-\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}\right) + \frac{1}{\left(\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}\right)} \int_{t_1}^{\infty} f(x_1,t_2)\log f(x_1,t_2)dx_1 \tag{3.3}$$

and

$$H_2(g_2,t_1,t_2) = \log\left(-\frac{\partial \bar{F}(t_1,t_2)}{\partial t_1}\right) + \frac{1}{\left(\frac{\partial \bar{F}(t_1,t_2)}{\partial t_1}\right)} \int_{t_1}^{\infty} f(t_1,x_2)\log f(t_1,x_2)dx_2 . \tag{3.4}$$

3.3 Characterization theorems

In this section we discuss characterization theorems associated with some bivariate models based on the functional form of the residual entropy function. Our first result focuses attention on the constancy of the residual entropy of the conditional distributions.

Theorem 3.1

Let $X = (X_1, X_2)$ be a non-negative, non-degenerate random vector admitting an absolutely continuous distribution function with respect to Lebesgue measure. The relation

$$H_i(g_i, t_1, t_2) = c_i,\ i = 1,2 \tag{3.5}$$

where c_i's are constants, holds for all real $t_1, t_2 \geq 0$ if and only if X is distributed as a bivariate exponential with independent (exponential) marginals.

Proof:

When (3.5) holds with $i = 1$, using (3.3) we can write

$$c_1 \frac{\partial \overline{F}(t_1,t_2)}{\partial t_2} = \frac{\partial \overline{F}(t_1,t_2)}{\partial t_2} \log\left(-\frac{\partial \overline{F}(t_1,t_2)}{\partial t_2}\right) + \int_{t_1}^{\infty} f(x_1,t_2) \log f(x_1,t_2)\, dx_1 .$$

Differentiating with respect to t_1 we get

$$c_1 f(t_1,t_2) = f(t_1,t_2) + f(t_1,t_2) \log\left(-\frac{\partial \overline{F}(t_1,t_2)}{\partial t_2}\right) - f(t_1,t_2) \log f(t_1,t_2).$$

Since $f(t_1,t_2) > 0$, the above equation can be written as

$$\log\left(-\frac{\partial \overline{F}(t_1,t_2)}{\partial t_2}\right) - \log f(t_1,t_2) = c_1 - 1$$

or

$$-\frac{\partial \overline{F}(t_1,t_2)}{\partial t_2} = k_1 f(t_1,t_2) \tag{3.6}$$

with $k_1 = e^{c_1 - 1} > 0$. Differentiating (3.6) with respect to t_1 we get

$$\frac{\partial \log f(t_1,t_2)}{\partial t_1} = -\frac{1}{k_1}.$$

Proceeding along similar lines with i=2 in (3.5) we get

$$\frac{\partial \log f(t_1,t_2)}{\partial t_2} = -\frac{1}{k_2},\ k_2>0$$

Using the argument in Galambos and Kotz (1978, p.128) we see that $f(t_1,t_2)$ is proportional to $\exp\left\{-\frac{1}{k_1}t_1 - \frac{1}{k_2}t_2\right\}$. The condition

$$\int_0^\infty\int_0^\infty f(t_1,t_2)dt_2dt_1 = 1$$

gives

$$f(t_1,t_2) = \lambda_1\ \lambda_2\ \exp\{-\lambda_1 t_1 - \lambda_2 t_2\};\quad t_1,t_2 \geq 0, \tag{3.7}$$

with $\lambda_1 = \frac{1}{k_1}$, $\lambda_2 = \frac{1}{k_2}>0$, as claimed.

Conversely when the distribution of X is specified by (3.7), by direct calculations we get

$$H_i(g_i,t_1,t_2) = 1 - \log \lambda_i,\ i = 1,2$$

so that the conditions of the theorem holds.

The following theorem looks into the situation where the residual entropy function of the conditional distribution of X_i given $X_j = t_j$ is log linear in t_i.

Theorem 3.2

Let $X = (X_1,X_2)$ be a non-negative, non-degenerate random vector admitting an absolutely continuous distribution function and $\lim_{x_1,x_2 \to 0} f(x_1,x_2)$ (>0) exists. The relation

$$H_i(g_i, t_1, t_2) = \log (At_i + B_i(t_j)), \quad i,j = 1,2 \; i \neq j \tag{3.8}$$

where $B_i(t_j)$ are non-negative non-increasing functions of t_j holds for all t_1, $t_2>0$ if and only if X follows

(i) the bivariate distribution with exponential conditionals [Arnold and Strauss (1988)] with probability density function

$$f(x_1,x_2) = k \exp\{-\beta_1 x_1 - \beta_2 x_2 - \beta_3 x_1 x_2\},$$

$$k,\ \beta_1,\ \beta_2,\ \beta_3 >0,\ x_1,\ x_2>0 \tag{3.9}$$

if $A = 0$.

(ii) the bivariate distribution with Pareto conditionals [Arnold (1987)] specified by

$$f(x_1,x_2) = k_1 (1+ c_1 x_1+ c_2 x_2+ c_3 x_1x_2)^{-d};$$

$$k_1,\ c_1,\ c_2,\ c_3 >0,\ d>1,\ x_1,\ x_2>0, \tag{3.10}$$

if $A > 0$

(iii) the bivariate distribution with Beta conditionals specified by

$$f(x_1,x_2) = k_2 (1- \alpha_1 x_1- \alpha_2 x_2 + \alpha_3 x_1x_2)^p;\ k_2,\alpha_1,\alpha_2>0,\ p>0,$$

$$1-p\leq\alpha_3/\alpha_1\alpha_2 \leq 1,\ 0<x_1<\frac{1}{\alpha_1},\ 0<x_2< \frac{1-\alpha_1 x_1}{\alpha_2-\alpha_3 x_1} \tag{3.11}$$

if $A <0$.

Proof:

When (3 8) holds with $i=1$, using (3.3) we can write

$$\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}\log\left(\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}\right)+\int_{t_1}^{\infty} f(x_1,t_2)\log f(x_1,t_2)dx_1$$

$$= \frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}\log(At_1+B_1(t_2)).$$

Differentiating the above equation with respect to t_1 we get

$$f(t_1,t_2)+f(t_1,t_2)\log\left(-\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}\right)-f(t_1,t_2)\log f(t_1,t_2)$$

$$= f(t_1,t_2)\log(At_1+B_1(t_2))+A\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}\frac{1}{At_1+B_1(t_2)}. \quad (3.12)$$

Dividing (3.12) throughout by $f(t_1,t_2)$ (>0) and setting

$$z_1(t_1,t_2)=\frac{-f(t_1,t_2)(At_1+B_1(t_2))}{\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}}, \quad (3.13)$$

(3.12) can be written as

$$z_1(t_1,t_2)\,[\log z_1(t_1,t_2)-1]=A. \quad (3.14)$$

Differentiating (3.14) with respect to t_1 and solving the resulting equation we get

$$z_1(t_1,t_2)=c_1(t_2) \quad (3.15)$$

where $c_1(t_2)$ is independent of t_1. Similarly differentiating (3.14) with respect to t_2 and proceeding along the same lines we also get

$$z_1(t_1,t_2) = c_2(t_1) \tag{3.16}$$

where $c_2(t_1)$ is independent of t_2. For (3.15) and (3.16) to hold simultaneously we should have

$$z_1(t_1,t_2) = k_1 \tag{3.17}$$

where k_1 is a constant.

Similarly when (3.8) holds with i = 2 and if

$$z_2(t_1,t_2) = \frac{-f(t_1,t_2)(At_2 + B_2(t_1))}{\frac{\partial \overline{F}(t_1,t_2)}{\partial t_1}}$$

we can also get

$$z_2(t_1,t_2) = k_2 \tag{3.18}$$

where k_2 is a constant. From the monotonicity of x (log x-1), it follows that $k_1 = k_2$. Let $k_1 = k_2 = k$.

From (3.13) and (3.17) we get

$$k \frac{\partial \overline{F}(t_1,t_2)}{\partial t_2} = -f(t_1,t_2)\,[At_1+B_1(t_2)].$$

Differentiating with respect to t_1 and rearranging the terms we get

$$\frac{\partial \log f(t_1,t_2)}{\partial t_1} = -\frac{-(k+A)}{At_1 + B_1(t_2)} \tag{3.19}$$

or

$$f(t_1,t_2) = m_1(t_2)\ \left[At_1 + B_1(t_2)\right]^{-\left(\frac{k+A}{A}\right)},\ A \neq 0, \qquad (3.20)$$

Proceeding on a similar lines we also get

$$f(t_1,t_2) = m_2(t_1)\ \left[At_2 + B_2(t_1)\right]^{-\left(\frac{k+A}{A}\right)} \qquad (3.21)$$

As $t_2 \to 0^+$ (3.21) gives

$$f(t_1,\ 0^+) = m_1(t_2)\left[B_2(t_1)\right]^{-\left(\frac{k+A}{A}\right)},$$

or

$$m_1(t_2) = f(t_1,\ 0^+)/\ \left[B_2(t_1)\right]^{-\left(\frac{k+A}{A}\right)} \qquad (3.22)$$

Similarly (3.20) gives

$$m_1(t_2) = f(0^+,\ t_2)/\ \left[B_1(t_2)\right]^{-\left(\frac{k+A}{A}\right)} \qquad (3.23)$$

Also as $t_2 \to 0^+$ (3.20) we get

$$f(t_1,\ 0^+) = m_1(0)\ \left[At_1 + B_1(0)\right]^{-\left(\frac{k+A}{A}\right)}. \qquad (3.24)$$

Similarly from (3.21)

$$f(0^+,\ t_2) = m_2(0)\ \left[At_2 + B_2(0)\right]^{-\left(\frac{k+A}{A}\right)}. \qquad (3.25)$$

From (3.20), (3.23) and (3.25) we have

$$f(t_1,t_2) = \frac{[At_1 + B_1(t_2)]^{-\left(\frac{k+A}{A}\right)} [At_2 + B_2(0)]^{-\left(\frac{k+A}{A}\right)} m_2(0)}{[B_1(t_2)]^{\left(\frac{k+A}{A}\right)}} \qquad (3.26)$$

similarly from (3.21), (3.22) and (3.24) we have

$$f(t_1,t_2) = \frac{[At_2 + B_2(t_1)]^{-\left(\frac{k+A}{A}\right)} [At_1 + B_1(0)]^{-\left(\frac{k+A}{A}\right)} m_1(0)}{[B_2(t_1)]^{\left(\frac{k+A}{A}\right)}} . \qquad (3.27)$$

Equating the two expressions for $f(t_1,t_2)$ we get

$$\left[1 + \frac{At_1}{B_1(t_2)}\right]^{-\left(\frac{k+A}{A}\right)} [At_2 + B_2(0)]^{-\left(\frac{k+A}{A}\right)} m_2(0)$$

$$= \left[1 + \frac{At_2}{B_2(t_1)}\right]^{\left(\frac{k+A}{A}\right)} [At_1 + B_1(0)]^{\left(\frac{k+A}{A}\right)} m_1(0).$$

That is

$$\left[1 + \frac{At_1}{B_1(t_2)}\right]^{\left(\frac{k+A}{A}\right)} \left[1 + \frac{At_2}{B_2(0)}\right]^{-\left(\frac{k+A}{A}\right)} [B_2(0)]^{-\left(\frac{k+A}{A}\right)} m_2(0)$$

$$= \left[1 + \frac{At_2}{B_2(t_1)}\right]^{-\left(\frac{k+A}{A}\right)} \left[1 + \frac{At_1}{B_1(0)}\right]^{-\left(\frac{k+A}{A}\right)} [B_1(0)]^{-\left(\frac{k+A}{A}\right)} m_1(0). \qquad (3.28)$$

But

$$[B_2(0)]^{-\left(\frac{k+A}{A}\right)} m_2(0) = [B_1(0)]^{-\left(\frac{k+A}{A}\right)} m_1(0) = f(0^+, 0^+) < \infty$$

(3.28) can now be written as

$$\left[1+\frac{At_1}{B_1(t_2)}\right]\left[1+\frac{At_2}{B_2(0)}\right] = \left[1+\frac{At_2}{B_2(t_1)}\right]\left[1+\frac{At_1}{B_1(0)}\right]$$

or

$$\frac{1}{t_1B_2(t_1)} - \frac{1}{t_1B_2(0)} + \frac{A}{B_2(t_1)B_1(0)} = \frac{1}{t_2B_1(t_2)} - \frac{1}{t_2B_1(0)} + \frac{A}{B_1(t_2)B_2(0)} \tag{3.29}$$

Since (3.29) is true for all real t_1, $t_2 \geq 0$ we may take both side of (3.29) equal to θ, where θ is a constant. This gives

$$B_2(t_1) = \frac{[At_1 + B_1(0)]B_2(0)}{[1+\theta B_2(0)t_1]B_1(0)}$$

and

$$B_1(t_2) = \frac{[At_2 + B_2(0)]B_1(0)}{[1+\theta B_1(0)t_2]B_2(0)}.$$

Substituting for $B_1(t_2)$ in (3.26) we get

$$f(t_1,t_2) = m_2(0)\,[B_2(0)]^{-\left(\frac{k+A}{A}\right)}\left\{1+\frac{At_1}{B_1(0)}+\frac{At_2}{B_2(0)}+\theta At_1t_2\right\}^{-\left(\frac{k+A}{A}\right)} \tag{3.30}$$

which is of the form (3 10) with $k_1 = m_2(0)\,[B_2(0)]^{-\left(\frac{k+A}{A}\right)}$, $c_1 = \frac{A}{B_1(0)}$, $c_2 = \frac{A}{B_2(0)}$, $c_3 = A\theta$ and $d = \frac{k+A}{A}$. If $A>0$ since $B_i(t_i)$ are non-negative functions of t_i we have k_1, c_1, c_2, $c_3 > 0$.

If $A<0$, (3.30) takes the form.

$$f(t_1,t_2) = k_2\,(1-\alpha_1 t_1-\alpha_2 t_2+\alpha_3 t_1 t_2)^p$$

with $k_2, \alpha_1, \alpha_2>0$, $p>0$. Further for $f(t_1, t_2)$ to be a probability density function we should have $0<t_1<\frac{1}{\alpha_1}$ and $0<t_2<\frac{1-\alpha_1 t_1}{\alpha_2-\alpha_3 t_1}$.

When $A=0$ (3.12) reads as

$$f(t_1,t_2)\log\left(-\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}\right)+f(t_1,t_2)-f(t_1,t_2)\log f(t_1,t_2)=f(t_1,t_2)\log B_1(t_2)$$

or

$$\log\left\{\frac{-\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2}}{f(t_1,t_2)}\right\} = \log B_1(t_2) - 1.$$

This gives

$$-\frac{\partial \bar{F}(t_1,t_2)}{\partial t_2} = \frac{1}{e} B_1(t_2)\, f(t_1, t_2).$$

Differentiating with respect to t_1 and rearranging the terms we get

$$\frac{\partial \log f(t_1,t_2)}{\partial t_1} = \frac{-e}{B_1(t_2)}$$

or

$$\log f(t_1,t_2) = \frac{-e}{B_1(t_2)} t_1 + m_1(t_2). \qquad (3.31)$$

Proceeding on similar lines we also get

$$\log f(t_1,t_2) = \frac{-e}{B_2(t_1)}\, t_2 + m_2(t_1). \tag{3.32}$$

When $t_1 \to 0^+$ in (3.31) we get

$$\log f(0^+, t_2) = m_1(t_2).$$

Also

$$\log f(t_1, 0^+) = m_2(t_1).$$

Further

$$\log f(t_1, 0^+) = \frac{-e}{B_1(0)}\, t_1 + m_1(0)$$

and

$$\log f(0^+, t_2) = \frac{-e}{B_2(0)}\, t_2 + m_2(0).$$

Hence from (3.31) we get

$$\log f(t_1,t_2) = \frac{-e}{B_1(t_2)}\, t_1 + \frac{-e}{B_2(0)}\, t_2 + m_2(0) \tag{3.33}$$

and from (3.32) we get

$$\log f(t_1,t_2) = \frac{-e}{B_2(t_1)}\, t_2 + \frac{-e}{B_1(0)}\, t_1 + m_1(0). \tag{3.34}$$

Equating the two expression for $f(t_1,t_2)$ we get

$$\frac{-e}{B_1(t_2)}\, t_1 + \frac{-e}{B_2(0)}\, t_2 + m_2(0) = \frac{-e}{B_2(t_1)}\, t_2 + \frac{-e}{B_1(0)}\, t_1 + m_1(0) \tag{3.35}$$

But since $m_1(0) = m_2(0) = f(0^+, 0^+)$, (3.35) takes the form

$$\frac{-1}{t_2 B_1(t_2)} + \frac{1}{t_2 B_2(0)} = \frac{-1}{t_1 B_2(t_1)} + \frac{1}{t_1 B_1(0)}$$

Since this is true for all real t_1, $t_2 \geq 0$, we should have

$$\frac{-1}{t_2 B_1(t_2)} + \frac{1}{t_2 B_2(0)} = \frac{-1}{t_1 B_2(t_1)} + \frac{1}{t_1 B_1(0)} = \theta$$

where θ is a constant. This gives

$$B_2(t_1) = \frac{B_2(0)}{1-\theta\, t_1 B_2(0)}$$

and

$$B_1(t_2) = \frac{B_1(0)}{1-\theta\, t_2 B_1(0)}.$$

From (3.33) we get

$$\log f(t_1,t_2) = \frac{-e}{B_1(0)}\, t_1 + \frac{-e}{B_2(0)}\, t_2 + e\theta\, t_1\, t_2 + m_2(0)$$

so that $f(t_1,t_2)$ has the form (3.9) with $k = \log m_2(0)$, $\beta_1 = \frac{-e}{B_1(0)}$,

$\beta_2 = \frac{-e}{B_2(0)}$ and $\beta_3 = e\theta$, which are non-negative.

The if part of Theorem 3.2 follows from the expressions for the residual entropy function of the conditional distribution of X_1 given $X_2 = t_2$ when the distributions are specified by (3.9), (3.10) and (3.11) respectively given by $1-\log(\beta_1+\beta_3 t_2)$,

$$\log\left\{\frac{e^{d/d-1}}{d-1}t_1 + \frac{e^{d/d-1}}{d-1}\left(\frac{1+c_2t_2}{c_1+c_3t_2}\right)\right\} \quad \text{and} \quad \log\left\{-\frac{e^{d/d+1}}{d+1}t_1 + \frac{e^{d/d+1}}{d+1}\left(\frac{1-\alpha_2t_2}{\alpha_1+\alpha_3t_2}\right)\right\}$$

with similar expression for $H_2(g_i,t_1,t_2)$.

If $B_i(t_j)$ in (3.8) are linear functions of t_j, say $a_i + b_j t_j$, we have the following theorem.

Theorem 3.3

For the random vector $X = (X_1, X_2)$ considered in Theorem 3.2, the relation.

$$H_i\,(g_i, t_1, t_2) = \log\,(a_i + b_i\, t_i + b_j\, t_j) \quad i, j = 1, 2,\ i \neq j$$

holds for all real t_1, $t_2 > 0$ if and only if X follows

(i) the bivariate exponential distribution with probability density function

$$f(t_1, t_2) = c_1 \exp\left\{-\lambda_1 t_1 - \lambda_2 t_2\right\},\ c_1,\ \lambda_1,\ \lambda_2 > 0$$

if $b_i = 0$, $i = 1,2$

(ii) the bivariate Pareto distribution specified by

$$f(t_1,t_2) = c_2\ \left(1 + k_1 t_1 + k_2 t_2\right)^{k_3},\ c_2,\ k_1,\ k_2,\ t_1,\ t_2 > 0,\ k_3 > 2$$

if $b_i > 0$, $i = 1, 2$

and

(iii) the bivariate Beta distribution specified by

$$f(t_1,t_2) = c_3\ \left(1 - a_1 t_1 - a_2 t_2\right)^{a_3},\ a_1,\ a_2,\ a_3,\ c_3 > 0,\ 0 < t_1 < \frac{1}{a_1},\ 0 < t_1 < \frac{1 - a_1 t_1}{a_2},$$

if $b_i < 0$, $i = 1, 2$

The proof is analogous to that of Theorem 3.2 and hence omitted.

For a random vector $X = (X_1, X_2)$ admitting an absolutely continuous distribution function in the support of R_2^+, it is special interest to consider another type of conditional distribution namely the conditional distribution of X_i given $X_j > t_j$, $i, j = 1, 2$, $i \neq j$. In life testing experiment if (X_1, X_2) represents the life time of the components in a two component system the above conditional distribution focuses attention on the distribution of the i^{th} component subject to the condition that the other has survived up to time t_j. The residual entropy of the conditional distribution of X_1 given $X_2 > x_2$ is

$$H_1^*(f_1, t_1, t_2) = -\int_{t_1}^{\infty} \frac{f(x_1|X_2 > t_2)}{\bar{F}(t_1|X_2 > t_2)} \log \frac{f(x_1|X_2 > t_2)}{\bar{F}(t_1|X_2 > t_2)} dx_1$$

$$= \log \bar{F}(t_1|X_2 > t_2) - \frac{1}{\bar{F}(t_1|X_2 > t_2)}$$

$$\int_{t_1}^{\infty} f(x_1|X_2 > t_2) \log f(x_1|X_2 > t_2) dx_1 \quad (3.36)$$

Since $\bar{F}(t_1, t_2) = \bar{F}(t_1|X_2 > t_2)\, \bar{F}_2(t_2)$, where $\bar{F}_i(t_i) = P(X_i > t_i)$, $i = 1, 2$, we have

$$\frac{\partial \bar{F}(t_1, t_2)}{\partial t_1} = -f(t_1|X_2 > t_2)\, \bar{F}_2(t_2)$$

so that

$$f(t_1|X_2 > t_2) = \frac{-1}{\bar{F}_2(t_2)} \frac{\partial \bar{F}(t_1, t_2)}{\partial t_1}.$$

Also

$$\bar{F}(t_1|X_2>t_2) = \frac{\bar{F}(t_1,t_2)}{\bar{F}_2(t_2)}.$$

Hence the residual entropy of the conditional distribution of X_1 given $X_2>t_2$ can also be written as

$$H_1^*(f_1, t_1, t_2) = 1+\frac{1}{\bar{F}(t_1,t_2)}\int_{t_1}^{\infty}\frac{\partial\bar{F}(x_1,t_2)}{\partial x_1}\log h_1(x_1, t_2)\, dx_1 \quad (3.37)$$

where $h=(h_1, h_2)$ is the vector valued failure rate considered by Johnson and Kotz (1975), namely $h=(h_1,h_2)$ with

$$h_j(t_1,t_2)=-\frac{\partial\log\bar{F}(t_1,t_2)}{\partial t_j}.$$

Similarly

$$H_2^*(f_1, t_1, t_2) = 1+\frac{1}{\bar{F}(t_1,t_2)}\int_{t_2}^{\infty}\frac{\partial\bar{F}(t_1,x_2)}{\partial x_2}\log h_2(t_1, x_2)\, dx_2 \quad (3.38)$$

The following theorems aims at characterizations of certain bivariate distributionsbased on the forms of $H_i^*(f_i, t_1, t_2)$ $i=1, 2$.

Theorem 3.4

Let $X = (X_1, X_2)$ be a non-negative non-degenerate random vector admitting an absolutely continuous distribution function with respect to a Lebsegue measure. The relation

$$H_i^*(f_i, t_1, t_2) = p_i,\ i=1, 2. \quad (3.39)$$

where p_i's are constants, holds for all real t_1, $t_2 \geq 0$ if and only if X is distributed as a bivariate exponential with independent (exponential) marginals.

Proof:

When (3.39) holds with $i = 1$, using (3.37) we have

$$(p_1-1)\ \bar{F}(t_1,t_2) = \int_{t_1}^{\infty} \frac{\partial \bar{F}(x_1,t_2)}{\partial x_1} \log h_1(x_1, t_2)\, dx_1 \qquad (3.40)$$

Differentiating (3.40) with respect to t_1 we get

$$(p_1-1)\ \frac{\partial \bar{F}(t_1,t_2)}{\partial t_1} = \frac{\partial \bar{F}(t_1,t_2)}{\partial t_1} \log h_1(t_1, t_2)$$

Since $\frac{\partial \bar{F}(t_1,t_2)}{\partial t_1} > 0$, the above equation can be written as

$$\log h_1(t_1, t_2) = (p_1-1)$$

or

$$h_1(t_1, t_2) = e^{(p_1-1)} = \lambda_1\ (>0)\ (\text{say})$$

Proceeding along similar lines with $i = 2$ in (3.39) we get

$$h_2(t_1, t_2) = e^{(p_2-1)} = \lambda_2\ (>0).\ (\text{say})$$

From Galambos and Kotz (1978) we have

$$\bar{F}(t_1,t_2) = \exp\{-\lambda_1 t_1 - \lambda_2 t_2\},\ t_1, t_2 \geq 0,$$

as claimed.

Conversely when X is specified by (3.7), by direct calculation we get

$$H_i^*(f_i,\, t_1,\, t_2) = 1 - \log \lambda_i,\; i=1,\, 2.$$

so that the condition of the theorem holds.

Theorem 3.5

For the random vector X, considered in the Theorem 3.1, the relation

$$H_i^*(f_i,\, t_1,\, t_2) = \log(at_i + b_i(t_j)),\; i,\, j = 1,\, 2,\; i \neq j \qquad (3.41)$$

where $b_i(t_j)$ are non negative function of t_j (>0) holds if and only if X is distributed as

1. the Gumbel's bivariate exponential distribution with survival function

   $$\bar{F}(t_1,t_2) = \exp\{-\alpha_1 t_1 - \alpha_2 t_2 - \theta t_1 t_2\},\; \alpha_1,\, \alpha_2,\, t_1,\, t_2 > 0,\; 0 < \theta < \alpha_1 \alpha_2 \qquad (3.42)$$

   if $a=0$.

2. the bivarite Pareto distribution specified by

   $$\bar{F}(t_1,t_2) = (1 + a_1 t_1 + a_2 t_2 + b t_1 t_2)^{-c},\; a_1, a_2, c, t_1, t_2 > 0,\; 0 < b \leq (c+1) a_1 a_2 \qquad (3.43)$$

   if $a>0$ and

3. the bivarite Beta distribution specified by

$$\bar{F}(t_1,t_2) = (1-p_1t_1-p_2t_2+qt_1t_2)^d,\ p_1,p_2,d>0, 0<t_1<\frac{1}{p_1},$$

$$0<t_2<\frac{1-p_1t_1}{p_2-qt_1}, 1\text{-}d\leq q\, p_1^{\ 1}\, p_2^{\ 1}\leq 1 \quad (3.44)$$

if $a<0$.

Proof:

When (3.41) hold with $i=1$, we have

$$\bar{F}(t_1,t_2)\log\,(a_1t_1+b_1(t_2)) = \bar{F}(t_1,t_2)+\int_{t_1}^{\infty}\frac{\partial\bar{F}(x_1,t_2)}{\partial x_1}\log h_1(x_1,\, t_2)\, dx_1$$

Differentiating with respect to t_1 and rearranging the terms we get

$$h(t_1,\, t_2)\ [\log\,(a_1t_1+b_1(t_2))+\log h(t_1,\, t_2)\ \text{-}1] = \frac{a}{at_1+b_1(t_2)}$$

Denoting by

$$c_1(t_1,\, t_2) = h(t_1,\, t_2)[\ at_1 + b_1(t_2)],$$

the above equation takes the form

$$c_1(t_1,\, t_2)[\log c_1(t_1,\, t_2)\ \ \text{-}1] = a.$$

Proceeding on similar lines as in the proof of Theorem 3.2, we get

$$c_1(t_1,\, t_2) = k,$$

where k is a constant. This gives

$$h_1(t_1,\, t_2) = \frac{k}{at_1+b_1(t_2)}.$$

Similarly we can also have

$$h_2(t_1, t_2) = \frac{k}{at_2 + b_2(t_1)}.$$

The rest of the proof follows from Roy (1989).

The if part of the theorem follows from the expression for given by $H_1^*(f_1, t_1, t_2)$ given by $1-\log(\alpha_1+\theta t_2)$,

$$\log\left(\frac{e^{1+\frac{1}{c}}}{c}t_1 + \frac{e^{1+\frac{1}{c}}}{c}\frac{1+a_2t_2}{c(a_1+bt_2)}\right) \text{ and } \log\left(\frac{-e^{1-\frac{1}{d}}}{d}t_1 + \frac{e^{1-\frac{1}{d}}}{d}\frac{1-p_2t_2}{(p_2-qt_2)}\right)$$

respectively for distributions specified by (3.42), (3.43) and (3.44) with the similar expression for $H_2^*(f_2, t_1, t_2)$. Hence the condition of the theorem holds.

Theorem 3.6

For the random vector X, considered in Theorem 3.4, a relation of the form

$$H_i^*(f_i, t_1, t_2) = k - \log h_i(t_1, t_2),\ i=1, 2, \tag{3.45}$$

where $h_i(t_1, t_2)$'s are the components of the bivariate failure rate holds for all real $t_1, t_2 \geq 0$ if and only if X follows any one of the three distributions specified by (3.42), (3.43) and (3.44) respectively according as $k=1$, $k>1$ and $k<1$.

Proof:

When (3.45) holds using (3.37) we can write

$$[k - \log h_1(t_1, t_2)]\ \bar{F}(t_1,t_2) = \bar{F}(t_1,t_2) + \int_{t_1}^{\infty} \frac{\partial \bar{F}(x_1,t_2)}{\partial x_1} \log h_1(x_1, t_2)\, dx_1$$

Differentiating with respect to t_1 and simplifying we get

$$\frac{-1}{h_1^2(t_1,t_2)} \frac{\partial h_1(t_1,t_2)}{\partial t_1} = (k\text{-}1).$$

If $u(t_1, t_2) = \dfrac{1}{h_1(t_1,t_2)}$ the above equation turn out to be

$$\frac{\partial u(t_1,t_2)}{\partial t_1} = (k\text{-}1),$$

whose solution is

$$u(t_1, t_2) = (k\text{-}1)\ t_1 + c_1$$

where c_1 is a constant. This gives

$$h_1(t_1, t_2) = [(k\text{-}1)\ t_1 + c_1]^{-1}.$$

Proceeding along similar lines one can also get

$$h_2(t_1, t_2) = [(k\text{-}1)\ t_2 + c_2]^{-1}.$$

where c_2 is a constant. This shows that the components of the vector valued failure rate are reciprocal linear. The rest of the proof is analogous to that of Theorem 3.5.

The if part follows from the expression for $H_1^*(f_1, t_1, t_2)$ given by $1-\log(\alpha_1+\theta t_2)$, $1+\frac{1}{c}+\log\left(\frac{1+a_1t_1+a_2t_2+bt_1t_2}{c(a_1+bt_2)}\right)$ and $1-\frac{1}{d}+\log\left(\frac{1-p_1t_1-p_2t_2+qt_1t_2}{d(p_1-qt_2)}\right)$ and that of $h_1(t_1,t_2)$ given by $\alpha_1+\theta t_2$, $\frac{c(a_1+bt_2)}{1+a_1t_1+a_2t_2+bt_1t_2}$ and $\frac{d(p_1-qt_2)}{1-p_1t_1-p_2t_2+qt_1t_2}$ when the distribution is specified by (3.42), (3.43) and (3.44) respectively with similar expression for $H_2(f_2, t_1, t_2)$ and $h_2(t_1,t_2)$.

CHAPTER IV

GEOMETRIC VITALITY FUNCTION

4.1. Introduction

The vitality function, extensively studied by Kupka and Loo (1989) in connection with their studies on ageing process, provides a useful tool in modelling life time data. Kotz and Shanbhag (1980) has used this concept, without specifying the name, to obtain several characterizations for life time distributions. Where as the hazard rate reflects the risk of sudden death with in a life span the vitality function provides a more direct measure of the failure pattern in the sense that it is expressed in terms of increased average life span. As mentioned in Section 1.2, the vitality function defined by

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

Some of the results in this chapter have appeared in the IAPQR Transaction (2000), Vol. 25(1), pp. 1-8.

measures the average life span of components whose age exceeds x. In the present chapter we define a new measure based on the geometric mean of the residual life time of the components and examine its properties.

4.2 Definition and Properties

Let X be a random variable admitting an absolutely continuous distribution function $F(x)$, with respect to Lebesgue measure on $(0, L)$, where

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

with $E(\log X) < \infty$. We define the geometric vitality function $G(t)$, for $t > 0$, as

$$\log G(t) = E(\log X \mid X > t)$$

$$= \frac{1}{\bar{F}(t)} \int_t^\infty \log x \, f(x) dx. \tag{4.1}$$

In the reliability context, if X represents the life length of a component, $G(t)$ represents the geometric mean of life time of the components which has survived upto time t. (4.1) can also be written as

$$\log\left(\frac{G(t)}{t}\right) = \frac{1}{\bar{F}(t)} \int_t^\infty \frac{\bar{F}(x)}{x} dx. \tag{4.2}$$

The following properties are immediate from the definition.

(a) log $G(t)$ is non-decreasing

(b) $\lim_{t \to 0} \log G(t) = E[\log X]$

(c) $m(t) \geq \log G(t)$ for all $t>0$

(d) if $h(t) = \dfrac{f(t)}{\bar{F}(t)}$ is the failure rate of X,

$$h(t) = \frac{\frac{d}{dt}\log G(t)}{\log\left(\frac{G(t)}{t}\right)} \tag{4.3}$$

(a) follows from the expression for the derivative of log $G(t)$ obtained from (4.1) namely

$$\frac{d\log G(t)}{dt} = \log\left(\frac{G(t)}{t}\right) h(t) \tag{4.4}$$

in which $h(t) \geq 0$ and $\log\left(\frac{G(t)}{t}\right) \geq 0$, in view of (4.2). (b) is straight forward. Using the fact the $x > \log x$ for all $x>0$ we have

$$E(X|X>t) > E(\log X|X>t)$$

which is same as (c). (d) is immediate from (4.4).

Theorem 4.1

The geometric vitality function determines the distribution uniquely.

Proof:

Let $f_1(t)$ and $f_2(t)$ be two probability density functions with geometric vitality functions $G_1(t)$ and $G_2(t)$ and that

$$G_1(t) = G_2(t).$$

Since

$$\log G_1(t) = \log G_2(t),$$

we have

$$\log\left(\frac{G_1(t)}{t}\right) = \log\left(\frac{G_2(t)}{t}\right),\ t>0. \tag{4.5}$$

From (4.3) and (4.5) we have

$$h_1(t) = h_2(t)$$

where $h_1(t)$ and $h_2(t)$ are the failure rates corresponding to $f_1(t)$ and $f_2(t)$ respectively. Since the failure rate determines the distribution uniquely we have

$$f_1(t) = f_2(t).$$

Further if

$$f_1(t) \neq f_2(t),$$

then

$$\log G_1(t) \neq \log G_2(t), \text{ for all } t.$$

Hence the geometric vitality function determines the distribution uniquely.

4.3 Characterization theorems

In this section we look into the problem of characterizing some well known life time models by the form of the geometric vitality function.

Theorem 4.2

Let X be a random variable in the support of $[x_0, \infty)$, with $x_0>0$, admitting an absolutely continuous distribution function and with geometric vitality function $G(t)$. The relation

$$\log\left(\frac{G(t)}{t}\right) = a, \tag{4.6}$$

where $a(>0)$ is a constant, holds for all t (>0) if and only if X follows the Pareto type I distribution specified by

$$\overline{F}(x) = \left(\frac{x_0}{x}\right)^{\alpha}, \quad x\geq x_0>0,\ \alpha>0 \tag{4.7}$$

Proof:

When (4.6) holds, using (4.2) we can write

$$\int_t^{\infty}\frac{\overline{F}(x)}{x}dx = a\,\overline{F}(t).$$

Differentiating the above equation with respect to t and rearranging the terms we get the expression for the failure rate of X as

$$h(t) = \frac{\alpha}{t},$$

where $\alpha = \frac{1}{a} > 0$. Using the relation

$$\overline{F}(x) = \exp\left\{-\int_{x_0}^{x} h(t)dt\right\} \tag{4.8}$$

we get (4.7) as claimed.

Conversely when the distribution of X is specified by (4.7), by direct calculations using (4.2) we get

$$\log\left(\frac{G(t)}{t}\right) = \frac{1}{\alpha},$$

so that (4.6) holds with $a = \frac{1}{\alpha}$.

It may be noted that (4.6) can also be written in the form

$$G(t) = kt,$$

where k is a constant. Hence Theorem 4.2 provides a characterization for the Pareto distribution when the geometric vitality function is proportional to the age

The following theorem provides a characterization for a family of distributions using a possible relationship between the geometric vitality function and the first order reciprocal moment of X.

Theorem 4.3

Let X be a non-negative random variable admitting an absolutely continuous distribution with respect to Lebesgue measure on $(0, \infty)$. A relation of the form

$$\log\left(\frac{G(t)}{t}\right) = a + bR(t),\ b>0 \tag{4.9}$$

where

$$R(\mathrm{t}) = \mathrm{E}(X^{-1}|X>t)$$

holds for all real t (>0) if and only if X follows

1. the Exponential distribution with survival function (2.7) for $a=0$
2. the Pareto distribution with survival function (2.8) for $a>0$ and
3. the Beta distribution with survival function (2.9) for $a<0$.

Proof:

When (4.9) holds, in the light of (4.2), we have

$$\frac{1}{\overline{F}(t)}\int_t^\infty \frac{\overline{F}(x)}{x}dx = a + \frac{b}{\overline{F}(t)}\int_t^\infty \frac{f(x)}{x}dx$$

or

$$\int_t^\infty \frac{\overline{F}(x)}{x}dx = a\,\overline{F}(t) + b\int_t^\infty \frac{f(x)}{x}dx \tag{4.10}$$

Differentiating (4.10) with respect to t and rearranging the terms we get

$$h(t) = [a+bt]^{-1}. \tag{4.11}$$

The rest of the proof follows from Mukherjee and Roy (1986).

Conversely when the distribution of X is specified by (2.7), (2.8) and (2.9) respectively by direct calculations, using (4.2), we get the expression for $\log\left(\frac{G(t)}{t}\right)$ as λ^{-1}, $p^{-1} + \alpha\, p^{-1} R(t)$ and $-c^{-1} + Rc^{-1} R(t)$ respectively so that the conditions of the theorem holds.

Corollary 4.1

For the random variable X considered in Theorem 4.3, the relation

$$\log\left(\frac{G(t)}{t}\right) = \frac{R_{\alpha}(t)}{\alpha},\ \alpha>0$$

where

$$R_{\alpha}(t) = \mathrm{E}(X^{\alpha}|X>t)$$

holds for all t (>0) if and only if X follows the Weibull distribution with survival function

$$\bar{F}(x) = e^{-x^{\alpha}},\ \ x>0,\ \alpha>0. \tag{4.12}$$

This result can be established proceeding along the same lines as in the proof of Theorem 4.3. The special case with $\alpha=2$ in Corrolary 4.1 provides a characterization for the Rayleigh distribution

Now we look into the problem of characterizing certain distributions using possible relationships between the residual entropy function and the geometric vitality function.

Theorem 4.4

Let X be a random variable admitting an absolutely continuous distribution in the support of (x_0, ∞), with geometric vitality function $G(t)$ and residual entropy function $H(f,t)$. The relation

$$H(f,t) - \log G(t) = c, \qquad (4.13)$$

where c is a constant, holds for all real t (>0) if and only if X follows the Pareto type I distribution specified by (4.7).

Proof:

When (4.13) holds, using (1.36) and (4.1), we get

$$(c-1)\,\overline{F}(t) = -\int_t^{\infty} f(x)\log x \, dx - \int_t^{\infty} f(x)\log h(x)\, dx. \qquad (4.14)$$

Differentiating (4.14) with respect to t we get

$$\log\,(th(t)) = 1 - c,$$

or

$$h(t) = \frac{x_0}{t}$$

with $x_0 = e^{1-c} > 0$. The relation (4.8) gives the form of $\bar{F}(t)$ as (4.7).

Conversely when the distribution of X is specified by (4.7) by direct calculations we get

$$\log G(t) = \frac{1}{\alpha} + \log t$$

and

$$H(f,t) = 1 + \frac{1}{\alpha} + \log\left(\frac{t}{\alpha}\right),$$

so that (4.13) holds with $c=1-\log \alpha$.

Singh and Maddala (1976) has obtained a model for income distributions using an increasing and bounded proportional failure rate. The model considered by them is specified by the distribution function

$$F(x) = 1-(1+ax^b)^{-c},\ x>0,\ a>0,\ b>0,\ c>0. \tag{4.15}$$

(4.15) is also known as the Burr type XII distribution. The following theorem examines how the residual entropy function and the geometric vitality function can be related so as to provide a characterization for (4.15).

Theorem 4.5

For the random variable X considered in Theorem 4.3, the relation

$$H(f,\ t) + (b\text{-}1)\log G(t) = \log(A+Bt^b) \qquad (4.16)$$

where $b>1$, A, $B>0$, holds for all real t (>0) if and only if X follows the Burr type XII distribution with distribution function (4.15).

Proof:

When (4.16) holds, using (1.36) and (4.1), we get

$$\overline{F}(t)\ \log(A+Bt^b) = \overline{F}(t) - \int_t^\infty f(x)\log h(x)\,dx + (b\text{-}1)\int_t^\infty f(x)\log x\,dx$$

Differentiating the above equation with respect to t we get

$$\frac{Bbt^{b-1}}{A+Bt^b}\ \frac{1}{h(t)} = -\log\left[\frac{t^{b-1}}{A+Bt^b}\ \frac{1}{h(t)}\right] - 1. \qquad (4.17)$$

Setting

$$\phi(t) = \frac{t^{b-1}}{A+Bt^b}\ \frac{1}{h(t)}, \qquad (4.18)$$

(4.17) becomes

$$Bb\phi(t) = -\log\phi(t) - 1 \qquad (4.19)$$

Differentiating (4.19) with respect to t and simplifying we get

$$\phi'(t)\{[\phi(t)]^{-1}+Bb\}=0. \qquad (4.20)$$

(4.20) give either $\phi'(t) = 0$ or $[\phi(t)]^{-1} = -Bb$. In either case $\phi(t)$ is a constant. Let

$$\phi(t) = k.$$

Using (4.18) we get

$$h(t) = \frac{t^{b-1}}{k(A+Bt^b)},\ k>0.$$

From the relation

$$\overline{F}(x) = \exp\left\{-\int_0^x h(t)dt\right\}, \qquad (4.21)$$

we get

$$\overline{F}(x) = (1+ax^b)^{-c},$$

with $a = \frac{B}{A}>0$ and $c = \frac{1}{kBb}>0$.

Conversely when the distribution is specified by (4.15), using (1.36) we have

$$H(f,\ t) = 1 - \frac{1}{\overline{F}(t)}\int_t^\infty f(x)\log h(x)\ dx$$

$$= 1-\frac{1}{\overline{F}(t)}\int_t^\infty cabx^{b-1}(1+ax^b)^{-c-1}\log\left(\frac{cabx^{b-1}}{1+ax^b}\right)dx$$

$$= 1 - \log(cab) - (b-1)\ \mathrm{E}(\log X|X>t)$$

$$+\frac{1}{\overline{F}(t)}\int_t^\infty cabx^{b-1}(1+ax^b)^{-c-1}\log(1+ax^b)dx$$

or

$$H(f,\ t) + (b-1)\ \mathrm{E}(\log X|X>t) = \log(A+Bt^b)$$

with $A = \frac{e^{1+\frac{1}{c}}}{cab}$ and $B = \frac{e^{1+\frac{1}{c}}}{cb}$ so that (4.16) holds.

Theorem 4.6

For the random variable X considered in Theorem 4.3, a relation of the form

$$H(f, t) + (c-1) \log G(t) = 1 - \log c, \qquad (4.22)$$

where $c(>0)$ is a constant, holds for all real $t(>0)$ if and only if X follows the Weibull distribution with survival function (4.12)

Proof:

When (4.22) holds using (4.1) we get

$$H(f, t)\ \overline{F}(t) + (c-1) \int_t^\infty f(x) \log x\, dx = (1 - \log c)\, \overline{F}(t). \qquad (4.23)$$

Differentiating (4.23) with respect to t and using (1.37) we get

$$h(t) = ct^{c-1}$$

Using (4.21) we get the form of $\overline{F}(x)$ as (4.12).

Conversely when X has the distribution with survival function (4.12), we have from (1.36)

$$H(f, t) = 1 - \log c - (c-1) \frac{1}{\overline{F}(t)} \int_t^\infty f(x) \log x\, dx$$

which is same as (4.22), as claimed

It may be noted that when $c=1$, (4.22) speaks about the constancy of residual entropy function which is characteristic to the exponential model. This result has already been given in Ebrahimi (1996). When $c=2$, (4.22) reduces to

$$H(f,\, t) + \log G(t) = 1 - \log 2$$

which is a characteristic property of Rayleigh distribution.

Belzunce, Candel and Ruiz (1995) defines a new class of distributions by using the mean left proportional residual income namely

$$e(t) = \mathrm{E}\left(\frac{X}{t}|X > t\right).$$

In a similar way we define the geometric mean left proportional residual life, $S(t)$, through the relationship

$$\log S(t) = \mathrm{E}\left(\log\left(\frac{X}{t}\right)|X > t\right).$$

$$= \frac{1}{\overline{F}(t)}\int_t^{\infty}\frac{\overline{F}(x)}{x}dx. \tag{4.24}$$

It may be noticed that $G(t)$ and $S(t)$ are related by

$$G(t) = t\, S(t),\ t>0. \tag{4.25}$$

The following theorem focuses attention on the monotonic behaviour of $S(t)$.

Theorem 4.8

If log $S(t)$ is increasing and log concave (decreasing and log convex) then the failure rate is decreasing (increasing).

Proof:

From (4.3) we have

$$h(t) = \frac{\frac{d}{dt}\log S(t)}{\log S(t)} + \frac{1}{t}\frac{1}{\log S(t)}$$

$$= \frac{d}{dt}\log[\log S(t)] + \frac{1}{t}\frac{1}{\log S(t)}. \qquad (4.26)$$

Differentiating (4.26) with respect to t we get

$$h'(t) = \frac{d^2}{dt^2}\log[\log S(t)] + \frac{d}{dt}\left[\frac{1}{t\log S(t)}\right]. \qquad (4.27)$$

Suppose log $S(t)$ is increasing. Then for $t_1<t_2$

$$\log S(t_1) \le \log S(t_2)$$

or

$$\frac{1}{t_1\log S(t_1)} \ge \frac{1}{t_2\log S(t_2)}.$$

This implies $\frac{1}{t\log S(t)}$ is decreasing and hence

$$\frac{d}{dt}\left[\frac{1}{t\log S(t)}\right] \le 0.$$

Also if log $S(t)$ is concave, then

$$\frac{d^2}{dt^2}\log[\log S(t)] \leq 0.$$

Hence from (4.26) $h'(t) \leq 0$. This implies that the failure rate is decreasing.

4.4 Geometric vitality function in discrete time

For a random variable X in the support of the set of non-negative integers, we define the geometric vitality function $G(t)$, for t=0, 1, 2, ..., as

$$\log G(t) = E(\log X \mid X > t)$$

$$= \frac{1}{\bar{F}(t+1)} \sum_{x=t+1}^{\infty} f(x) \log x \qquad (4.28)$$

where $f(x)$ = P(X=x)and $\bar{F}(x)$ = P($X \geq x$) are the probability mass function and the survival function of X respectively. Writing $f(x) = \bar{F}(x) - \bar{F}(x+1)$, (4.28) can also be written as

$$\log\left(\frac{G(t)}{t}\right) = \frac{1}{\bar{F}(t+1)} \sum_{x=t+1}^{\infty} \bar{F}(x) \log\left(\frac{x}{x-1}\right). \qquad (4.29)$$

Analogous to the continuous case the geometric vitality function satisfies the property (a), (b) and (c) mentioned in section 4.3. Also if $h(t) = f(t)/\bar{F}(t)$ is the failure rate of X, we have

$$h(t) = \frac{\log\left(\frac{G(t)}{G(t-1)}\right)}{\log\left(\frac{G(t)}{t}\right)}. \tag{4.30}$$

Further, in view of (4.30), the geometric vitality function determines the distribution uniquely. The following theorem characterizes a family of discrete distribution based on a relationship between the geometric vitality function and the reciprocal moments.

Theorem 4.7

Let X be a random variable in the support of the set of non-negative integers. The relation

$$\log\left(\frac{G(t)}{t}\right) = \sum_{n=1}^{\infty}\frac{1}{n}[\alpha R_n(t) + \beta R_{n-1}(t)] \tag{4.31}$$

where $R_n(t)$ is the n^{th} order truncated reciprocal moment of X, namely

$$R_n(t) = \mathrm{E}(X^{-n}|X>t), \tag{4.32}$$

holds for all integers t (>0) if and only if X follows

(a) the geometric distribution specified by

$$\overline{F}(x) = q^x,\ x = 0, 1, 2, \ldots, 0<q<1 \tag{4.33}$$

if $\beta=0$.

(b) the Waring distribution specified by

$$\bar{F}(x) = \frac{(b)_x}{(a)_x},\ x = 0,\ 1,\ 2,\ \ldots,\ a>0,\ b>0,\ a>b \tag{4.34}$$

if $\beta>0$ and

(c) the negative hypergeometric distribution specified by

$$\bar{F}(x) = \frac{\binom{k+m-x}{m-x}}{\binom{k+m}{m}},\ x = 0,\ 1,\ 2,\ \ldots,\ m,\ k\geq 0, \tag{4.35}$$

if $\beta<0$.

Proof:

Using the series expansion of the logarithm in (4.29) we obtain

$$\log\left(\frac{G(t)}{t}\right) = \sum_{n=1}^{\infty}\frac{1}{n}\left[\frac{1}{\bar{F}(t+1)}\sum_{x=t+1}^{\infty}\frac{\bar{F}(x)}{x^n}\right] \tag{4.36}$$

when (4.31) holds, using (4.32) and (4.36) we can write

$$\sum_{n=1}^{\infty}\frac{1}{n\bar{F}(t+1)}\left[\sum_{x=t+1}^{\infty}\frac{\bar{F}(x)}{x^n} - \alpha\frac{f(x)}{x^n} - \beta\frac{f(x)}{x^{n-1}}\right] = 0 \tag{4.37}$$

Interchanging the order of summation and using the expansion of $\log(1-x^{-1})$ we get.

$$\frac{1}{\bar{F}(t+1)}\sum_{x=t+1}^{\infty}\log\left(1-\tfrac{1}{x}\right)[\ \bar{F}(x) - \alpha f(x)\ - \beta x f(x)] = 0.$$

This gives

$$\bar{F}(x) - \alpha f(x)\ - \beta x f(x) = 0$$

or

$$h(x) = (\alpha+\beta x)^{-1}.$$

The only if part now follows from Xekalaki (1983).

Conversely when the distribution of X is specified by (4.33), (4.34) and (4.35), (4.31) holds with $\alpha = (1-q)^{-1}$ and $\beta = 0$ for the geometric, $\alpha = \frac{a}{a-b}$, $\beta = \frac{1}{a-b} (>0)$ for the Waring and $\alpha = \frac{k+m}{k}$ and $\beta = -\frac{1}{k} (<0)$ for the negative hypergeometric distribution.

CHAPTER V

AVERAGING OF THE RESIDUAL ENTROPY FUNCTION AND RESIDUAL ENTROPY FUNCTIONS OF HIGHER ORDER

5.1. Introduction

When one is interested in the failure of a device in a finite interval, instead of examining the nature of failure rate at each point in the interval, it will be of more use if the average of the failure rate in the whole interval is used. Roy and Mukherjee(1989) have defined the averages of failure rate and has examined utility of the same in ordering of life distributions.

The arithmetic, geometric and harmonic mean of failure rates for a non-negative random variable X have been defined through the relations

$$A(x) = \frac{1}{x} \int_0^x h(t)dt$$

$$G(x) = \exp\left\{\frac{1}{x}\int_0^x \log h(t)dt\right\}$$

and

$$H(x) = \left\{\frac{1}{x}\int_0^x \frac{1}{h(t)}dt\right\}^{-1} \tag{5.1}$$

The problem of characterizing some well-known life time distributions based on the above concept are also examined by them. Analogously one can define the averages of the residual entropy function. In the sequel we look into the problem of characterizing some life time distributions using the residual entropy function and the averages of failure rates.

5.2 Characterization theorems

Theorem 5.1

Let X be a non-negative random variable admitting an absolutely continuous distribution function with arithmetic, geometric and harmonic mean of failure rates $A(x)$, $G(x)$ and $H(x)$ respectively. Denote by $H(f, t)$, the residual entropy function. The relation

$$A(t) = G(t) = H(t) = \exp\{1 - H(f,t)\} \tag{5.2}$$

holds for all real $t(>0)$ if and only if X follows the exponential distribution.

Proof:

When (5.2) holds, we have

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

Using (5.1), the above equation can be written as

$$t\, H(f,t) + \int_0^t \log h(x)dx = 1. \qquad (5.3)$$

Differentiating with respect to t and using the expression for $H'(f,t)$ given in (1.37), (5.3) simplifies to

$$H(f,t) + \log h(t) = 1. \qquad (5.4)$$

This gives

$$H'(f,t) + \frac{h'(t)}{h(t)} = 0. \qquad (5.5)$$

Using (1.37), (5.5) can be written as

$$h(t)[\, H(f,t) + \log h(t) - 1] = -\frac{h'(t)}{h(t)}.$$

In view of (5.4) the above equation simplifies to

$$h'(t) = 0,$$

so that $h(t) = \lambda$, where λ is a constant. Since the constancy of failure rate is characteristic to the exponential model, X follows the exponential distribution From Roy and Mukherjee (1989) the properties $A(x) = G(x) = H(x)$ is characteristic to the exponential model and so the sufficiency part follows.

Conversely when X follows the exponential distribution with probability density function

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

by direct calculations we get $A(x)=G(x)=H(x)=\lambda$ and $H(f,\ t)=1-\log\lambda$, so that (5.2) holds.

The following theorem provides a characterization for a family of distributions using a relationship between residual entropy function and arithmetic mean failure rate

Theorem 5.2

For the random variable X considered in Theorem 5.1, the relation

$$H(f,t) + ct\, A(t) = k, \tag{5.6}$$

where k is a constant, holds for all real t (>0) if and only if X follows

(i) the Exponential distribution with survival function (2.7) if $c=0$

(ii) the Pareto distribution with survival function (2.8) if $c<0$ and

(iii) the Beta distribution with survival function (2.9) if $c>0$.

Proof:

When (5.6) holds we have

$$H'(f,t) + ct\,A'(t) + c\,A(t) = 0.$$

Using (1.37) the above equation can be written as

$$h(t)[\,H(f,t) + \log h(t) - 1] + c[t\,A'(t) + A(t)] = 0. \tag{5.7}$$

From (5.1) we have

$$t\,A'(t) + A(t) = h(t). \tag{5.8}$$

Using (5.6) and (5.8), (5.7) can be written as

$$h(t)[k - ctA(t) + \log h(t) + c - 1] = 0. \tag{5.9}$$

Assume $h(t) \neq 0$. From (5.9) we get

$$\log h(t) - ctA(t) = 1 - c - k. \tag{5.10}$$

Differentiating (5.10) with respect to t and using (5.8) we get

$$\frac{h'(t)}{h(t)} = c.$$

The above equation gives

$$h(t) = [\alpha - ct]^{-1}, \tag{5.11}$$

where α is the constant of integration. From Mukherjee and Roy (1986), (5.11) is characteristic to the exponential distribution for $c=0$, the Pareto distribution for $c<0$ and the Beta distribution for $c>0$.

The if part of the theorem follows from the expression for $H(f,t)$ and $A(t)$ which are given below.

Distribution	$A(t)$	$H(f,t)$
Exponential	λ	$1-\log\lambda$
Pareto	$\frac{a}{t}\log\left(\frac{t+\alpha}{\alpha}\right)$	$1+\frac{1}{a}-\log\left(\frac{a}{t+\alpha}\right)$
Beta	$\frac{-c}{t}\log\left(\frac{R-t}{R}\right)$	$1-\frac{1}{c}-\log\left(\frac{c}{R-t}\right)$

Instead of using the residual entropy function $H(f,t)$, as such, one can utilize average value of $H(f,t)$, namely

$$\overline{H}(f,t) = \frac{1}{t}\int_0^t H(f,x)dx \qquad (5.12)$$

as a measure of stability of components in a finite interval. The following theorem provides a characterization for the family of distributions considered in Theorem 5.2 using a functional relation between $\overline{H}(f,t)$ and $G(t)$.

Theorem 5.3

For the random variable X considered in Theorem 5.1, the relation

$$\overline{H}(f,t) + \log G(t) = k, \qquad (5.13)$$

holds for all real $t(>0)$ if and only if X follows

(i) the Exponential distribution with survival function (2.7) if $k=0$

(ii) the Pareto distribution with survival function (2.8) if $k>0$ and

(iii) the Beta distribution with survival function (2.9) if $k<0$.

Proof:

Observing that (5.13) can be written as

$$H(f,t) + \log h(t) = k,$$

the proof is immediate in view of theorem 2.4.

It may further be noted that the residual entropy function and its arithmetic average coincides if and only if the distribution is exponential.

5.3 Other measures of residual entropy

Entropies of higher order are defined by several authors and their properties are being examined. The works of Renyi (1961), Havrada and Charvat (1967), Kapur (1968), Behra and Chawla (1974), Sharma and Mittal (1975), proceed in this direction.

Renyi (1961) defines entropies of order α as

$$R_\alpha(f) = \frac{1}{1-\alpha} \log \sum_{x=0}^{t} f^{\alpha}(x), \ \alpha \neq 1, \ \alpha > 0. \qquad (5.14)$$

For a continuous non-negative random variable X admitting an absolutely continuous distribution, (5.14) takes the form

$$R_\alpha(f) = \frac{1}{1-\alpha} \log \int_0^\infty f^\alpha(x)dx. \qquad (5.15)$$

When $\alpha \to 1$, (5.15) reduces to the Shannon's entropy. For the random variable $(X-t)$ truncated at t (>0), (5.15) reads as

$$R_\alpha(f,t) = \frac{1}{1-\alpha} \log \int_t^\infty \left\{\frac{f(x)}{\bar{F}(t)}\right\}^\alpha dx. \qquad (5.16)$$

Further, (5.16) simplifies to the residual entropy function considered in section 1.3, as $\alpha \to 1$.

The following theorem focuses attention on the constancy of $R_\alpha(f,t)$.

Theorem 5.4

Let X be a non-negative continuous random variable admitting an absolutely continuous distribution function with Renyi's entropy measure $R_\alpha(f,t)$ ($<\infty$). The relation

$$R_\alpha(f,t) = c \qquad (5.17)$$

where c is a constant holds for all real $t(>0)$ if and only if X follows the exponential distribution.

Proof:

When (5.17) holds, using (5.16) we have

$$(1-\alpha)\, c = -\, \alpha \log \bar{F}(t) + \log \int_t^{\infty} f^{\alpha}(x)dx \qquad (5.18)$$

Differentiating (5.18) with respect to t we get

$$\alpha\, h(t) - \frac{f^{\alpha}(t)}{\int_t^{\infty} f^{\alpha}(x)dx} = 0$$

or

$$f^{\alpha}(t) = \alpha\, h(t) \int_t^{\infty} f^{\alpha}(x)dx. \qquad (5.19)$$

Differentiating (5.19) with respect to t and simplifying we get

$$\alpha \frac{f'(t)}{f(t)} = -\, \alpha\, h(t) + \frac{h'(t)}{h(t)}. \qquad (5.20)$$

Since

$$\frac{h'(t)}{h(t)} = \frac{f'(t)}{f(t)} + h(t)$$

(5.20) simplifies to

$$\frac{f'(t)}{f(t)} = -\, h(t),\ \alpha \neq 1.$$

This gives

$$\frac{d \log f(t)}{dt} = \frac{d \log \bar{F}(t)}{dt}$$

or

$$f(t) = k\, \bar{F}(t)$$

Hence

$$h(t) = k,$$

where k is the constant of integration. Hence X follows the exponential distribution.

Further for the exponential distribution specified by the survival function (2.7), by direct calculations we find

$$R_\alpha(f,t) = \frac{\log\alpha}{1-\alpha},$$

so that (5.17) is satisfied.

In connection with their studies relating to income inequality, Ord, Patil and Taillie (1983) has proposed the measure defined by

$$e_\gamma(f,\ t) = \int_t^\infty \frac{f(x)}{\bar{F}(t)}\left[1-\left\{\frac{f(x)}{\bar{F}(t)}\right\}^\gamma\right]\frac{dx}{\gamma} \tag{5.21}$$

as a useful measure of income inequality. When $\gamma\neq 0$, (5.21) can be written as

$$1-\gamma\, e_\gamma(f,\ t) = \frac{1}{\{\bar{F}(t)\}^{\gamma+1}}\int_t^\infty f^{\gamma+1}(x)dx. \tag{5.22}$$

Further when t=0, (5.21) takes the form

$$e_\gamma(f) = \int_0^\infty f(x)[1-f^\gamma(x)]\ \frac{dx}{\gamma}. \tag{5.23}$$

Taking the limit of (5.23) as $\gamma\to 0$ we get

$$\lim_{\gamma \to 0} e_\gamma(f) = \lim_{\gamma \to 0} \frac{\int_0^\infty \frac{df(x)[1-f^\gamma(x)]}{d\gamma}dx}{\frac{d\gamma}{d\gamma}}$$

$$= - \int_0^\infty f(x) \log f(x)\, dx$$

which is the Shannon's entropy encountered in Section (1.1). By a similar argument with (5.21), one can verify that $\lim_{\gamma \to 0} e_\gamma(f, t)$ is the residual entropy function defined by (1.34).

The following relationship exists between the $e_r(f, t)$ and the Renyi's entropy measure. We have from (5.16)

$$(1-\alpha)\, R_\alpha(f,t) = \log \frac{1}{\overline{F}^\alpha(t)} \int_t^\infty f^\alpha(x)dx$$

so that

$$\exp\{(1-\alpha)\, R_\alpha(f,t)\} = \frac{1}{\overline{F}^\alpha(t)} \int_t^\infty f^\alpha(x)dx$$

From (5.22) the above takes the form

$$\exp\{(1-\alpha)\, R_\alpha(f,t)\} = 1-(1-\alpha)\, e_{\alpha-1}(f,t). \qquad (5.24)$$

We now establish a recurrence relation satisfied by $e_\gamma(f, t)$, as a consequence of which it is seen that $e_\gamma(f, t)$ determines the distribution uniquely.

Theorem 5.5

Let X be a continuous non-negative random variable with $e_\gamma(f, t) <\infty$. Then $e_\gamma(f, t)$ uniquely determines the distribution.

Proof:

From (5.22) we have

$$[1-\gamma\, e_\gamma(f, t)]\ \bar{F}^{\gamma+1}(t) = \int_t^\infty f^{\gamma+1}(x)dx. \qquad (5.25)$$

Differentiating (5.25) with respect to t and dividing by $\bar{F}^{\gamma+1}(t)$ we get

$$-(\gamma+1)\ [1-\gamma\, e_\gamma(f, t)]h(t)\ -\ \gamma e_\gamma'(f, t) = -\ h^{\gamma+1}(t)$$

or

$$\gamma e_\gamma'(f, t) = h(t)\ [(\gamma+1)\gamma\, e_\gamma(f, t)+h^\gamma(t)\ -\ (\gamma+1)]. \qquad (5.26)$$

Suppose $f_1(.)$ and $f_2(.)$ are density functions with

$$e_{r\gamma}(f_1, t) = e_\gamma(f_2, t),\ t\geq 0.$$

Using (5.26) we get

$$h_1(t)\ [(\gamma+1)\ \gamma\, e_\gamma(f_1,t)+ h_1^\gamma(t)\ -\ (\gamma+1)]$$

$$= h_2(t)\ [(\gamma+1)\gamma\, e_\gamma(f_2, t)+\ h_2^\gamma(t)\ -\ (\gamma+1)]. \qquad (5.27)$$

where $h_1(t)$ and $h_2(t)$ are the failure rates corresponding to $f_1(.)$ and $f_2(.)$ respectively. To prove $\overline{F}_1(t) = \overline{F}_2(t)$ we need to show that $h_1(t) = h_2(t)$ for all $t(\geq 0)$. Suppose

$$h_1(t) > h_2(t).$$

From (5.27) we have

$$(\gamma+1)\gamma\, e_\gamma(f_2, t) + h_2^\gamma(t) - (\gamma+1) > (\gamma+1)\gamma\, e_\gamma(f_1, t) + h_1^\gamma(t) - (\gamma+1).$$

This gives

$$h_1(t) < h_2(t)$$

which is a contradiction. Similarly we can see that $h_1(t) < h_2(t)$ also leads to a contradiction. This implies

$$h_1(t) = h_2(t)$$

and so the proof is complete.

Now we look into the problem of characterizing probability distributions using the functional form of $e_\gamma(f, t)$.

Theorem 5.6

For the random variable X considered in Theorem 5.5, the relation

$$1-\gamma\, e_\gamma(f, t) = (At+B)^{-\gamma},\ \gamma>-1 \tag{5.28}$$

holds for all real t (>0) if and only if X follows

(i) the Exponential distribution with survival function (2.7) if $A=0$

(ii) the Pareto distribution with survival function (2.8) if $A>0$ and

(iii) the Beta distribution with survival function (2.9) if $A<0$.

Proof:

When (5.28) holds, using (5.22), we can write

$$\int_t^\infty f^{\gamma+1}(x)dx = \bar{F}^{\gamma+1}(t)(At+B)^{-\gamma}.$$

Differentiating the above equation with respect to t we get

$$-f^{\gamma+1}(t) = -A\gamma(At+B)^{-\gamma-1}\bar{F}^{\gamma+1}(t) - (\gamma+1)\bar{F}^{\gamma}(t)f(t)(At+B)^{-\gamma}$$

Dividing throughout by $\bar{F}^{\gamma+1}(t)\ (At+B)^{-\gamma-1}$ we get

$$[h(t)(At+B)]^{\gamma+1} = A\gamma + (\gamma+1)h(t)\ (At+B). \tag{5.29}$$

Denoting by

$$y(t) = h(t)\ (At+B), \tag{5.30}$$

(5.29) takes the form

$$y^{\gamma+1}(t) - (\gamma+1)\ y(t) = -\gamma A$$

Differentiating the above equation with respect to t we get

$$(\gamma+1)\ y^{\gamma}(t)y'(t) - (\gamma+1)\ y'(t) = 0$$

or

$$y'(t)\ [(\gamma+1)\ y^{\gamma}(t) - (\gamma+1)] = 0.$$

In either case

$$y(t) = k,$$

where k is a constant. From (5.30) we get

$$h(t) = \frac{k}{At + B}$$

$$= [at+b]^{-1} \tag{5.31}$$

with $a = \frac{A}{k}$, $b = \frac{B}{k}$. From Roy and Mukherjee and Roy (1986), (5.31) is characteristic to the exponential distribution if a=0, the Pareto distribution if a>0 and the Beta distribution if a<0. The if part follows by translating the result for A.

The only if part follows from the expression for $1-\gamma\, e_\gamma(f_2, t)$ when the distribution are specified by (2.7), (2.8) and (2.9) respectively given by $\frac{\lambda^{\gamma+1}}{\gamma+1}$, $\frac{a^{\gamma+1}(t+\alpha)^{\gamma}}{(a+1)(\gamma+1)-1}$ and $\frac{c^{\gamma+1}(R-t)^{\gamma}}{(c-1)(\gamma+1)+1}$.

Instead of assuming γ>-1 if we restrict the range of γ to the set of non-negative reals we arrive at a more general result which given as Theorem 5.7.

Theorem 5.7

For the random variable X considered in theorem 5.5, the relation

$$1-\gamma\, e_\gamma(f, t) = k\, h^\gamma(t),\ \gamma>0 \tag{5.32}$$

holds for all real $t(>0)$ if and only if X follows any one of the distributions specified by (2.7), (2.8) and (2.9) respectively according as $k = (1+\gamma)^{-1}$, $k < (1+\gamma)^{-1}$ and $k > (1+\gamma)^{-1}$.

Proof:

When (5.32) holds, using (5.22) we can write

$$\int_t^\infty f^{\gamma+1}(x)dx = \bar{F}^{\gamma+1}(t)\, h^{\gamma}(t). \tag{5.33}$$

Differentiating (5.33) with respect to t and rearranging the terms we get

$$-\frac{h'(t)}{h(t)} = \frac{1-k(\gamma+1)}{k\gamma}$$

or

$$\frac{d\left(\dfrac{1}{h(t)}\right)}{dt} = \frac{1-k(\gamma+1)}{k\gamma}.$$

This gives

$$h(t) = \left[\left(\frac{1-k(\gamma+1)}{k\gamma}\right)t + d\right]^{-1} \tag{5.34}$$

where d is the constant of integration. (5.34) takes the form

$$h(t) = (pt+d)^{-1}$$

where $p = \dfrac{1-k(\gamma+1)}{k\gamma}$. The rest of the proof of the sufficiency part is similar to that of Theorem 5.6.

The only if part of the theorem follows from the expression for $e_\gamma(f, t)$ and $h(t)$ given below.

Distribution	$h(t)$	$e_\gamma(f, t)$
Exponential	λ	$\dfrac{\lambda^\gamma}{\gamma+1}$
Pareto	$\dfrac{a}{t+\alpha}$	$\dfrac{a^{\gamma+1}(t+\alpha)^{-\gamma}}{(a+1)(\gamma+1)-1}$
Beta	$\dfrac{c-1}{R-t}$	$\dfrac{c^{\gamma+1}(R-t)^{-\gamma}}{(c-1)(\gamma+1)+1}$

CHAPTER VI

RESIDUAL ENTROPY FUNCTION IN DISCRETE TIME

6.1. Introduction

Most of the works in reliability modelling assumes that the underlying life time model is a continuous distribution. However the limitation of measuring devices and the fact that discrete models provide good approximations for their continuous counterparts necessitate assessment of reliability in discrete time. Xekalaki (1983) provides examples of situations where discrete models are appropriate by citing examples. The works of Gupta and Gupta (1983), Lawless (1982), Hitha and Nair (1989), Roy and Gupta (1992), Shaked, Shanthi Kumar and Torres (1995) aims at characterization of probability distribution using discrete reliability concepts

---

Some of the results in this chapter have appeared in the Far East Journal of Theoretical Statistics (1998), Vol. 2(1), pp. 1-10.

The residual entropy function discussed in Section 1.3 can also be defined in the discrete set up. This enable one to determine the model through a knowledge of the form of the residual entropy function. In the present chapter we define the residual entropy function in the discrete set up examine its properties.

6.1 Definition and Properties

Let X be a random variable in the support of the set of non negative integers, with the probability mass function $f(x)$ and survival function $\overline{F}(x)$. Analogous to the definition of failure rate given in (1.11) and that of mean residual life function given by (1.27), we define the residual entropy function associated with the random variable X as

$$H(f,t) = - \sum_{x=t+1}^{\infty} \frac{f(x)}{\overline{F}(t+1)} \log \frac{f(x)}{\overline{F}(t+1)}. \qquad (6.1)$$

(6.1) can also be written as

$$H(f,t) = \log \overline{F}(t+1) - \frac{1}{\overline{F}(t+1)} \sum_{x=t+1}^{\infty} f(x) \log f(x). \qquad (6.2)$$

Observing that for the random variable $Y = X-t$ truncated at t (≥ 0), the survival function $\overline{G}(y)$ and probability mass function $g(y)$ are respectively

$$\overline{G}(y) = \overline{F}(t+y+1)/\overline{F}(t+1)$$

and

$$g(y) = f(t+y+1)/\overline{F}(t+1)$$

we notice that the Shannon's entropy corresponding to Y, namely

$$H(g,t) = -\sum_{y=0}^{\infty} \frac{f(t+y+1)}{\overline{F}(t+1)} \log \frac{f(t+y+1)}{\overline{F}(t+1)},$$

simplifies to (6.1). Hence the Shannon's entropy corresponding to the residual life is same as the residual entropy function (6.1). It may also be observed that (6.1) serves as a measure of stability of the component at time t when time is measured at discrete points. Also using the relationship $h(x) = \dfrac{f(x)}{\overline{F}(x)}$ and $f(x) = \overline{F}(x) - \overline{F}(x+1)$, (6.2) can be written as

$$H(f,t) = -\log h(t+1) \quad \sum_{x=t+1}^{\infty} \frac{\overline{F}(x+1)}{\overline{F}(t+1)} \log \frac{h(x)}{h(x+1)[1-h(x)]}. \qquad (6.3)$$

In terms of the mean residual life function, considered in section 1.2, $H(f, t)$ can also be written as

$$H(f,t) = -\log \left[\frac{r(t+1)-r(t)+1}{r(t+1)}\right]$$

$$+\sum_{x=t+1}^{\infty}\frac{\bar{F}(x+1)}{\bar{F}(t+1)}\log\left[\frac{r(x+1)[r(x)-r(x-1)+1]}{[r(x-1)-1][r(x+1)-r(x)+1]}\right].$$

We now establish a recurrence relation satisfied by $H(f,t)$.

Theorem 6.1

Let X be a discrete random variable in the support of I^+ with probability mass function $f(x)$, failure rate $h(x)$ and residual entropy function $H(f,t)$. Then $H(f,t)$ satisfies the recurrence relation,

$$H(f,t)=\frac{1}{1-h(t)}\{H(f,t-1)+h(t)\log h(t)+[1-h(t)]\log[1-h(t)]\},$$

$$t=1,2,3,\ldots . \qquad (6.4)$$

Proof:

From (6.2) we have

$$H(f,t)\bar{F}(t+1)=\bar{F}(t+1)\log\bar{F}(t+1)-\sum_{x=t+1}^{\infty}f(x)\log f(x) \qquad (6.5)$$

Subtracting (6.5) from the equation obtained by changing t to $(t+1)$ in (6.5) we get

$$H(f,t+1)\bar{F}(t+2)-H(f,t)\bar{F}(t+1)$$

$$=f(t+1)\log f(t+1)+\bar{F}(t+2)\log\bar{F}(t+2)-\bar{F}(t+1)\log\bar{F}(t+1). \qquad (6.6)$$

Since $f(t+1)=\bar{F}(t+1)-\bar{F}(t+2)$, (6.6) can be written as

$$H(f,t+1) = \frac{\overline{F}(t+1)}{\overline{F}(t+2)}\,[H(f,t)+\log h(t+1)] - \log\frac{f(t+1)}{\overline{F}(t+2)}$$

or

$$H(f,t+1) = \frac{1}{1-h(t+1)}\,[H(f,t)+\log h(t+1)] - \log\frac{h(t+1)}{1-h(t+1)}.$$

Rearranging the terms in the above equation we get

$$H(f,t+1)=\frac{1}{1-h(t+1)}\{H(f,t)+h(t+1)\log h(t+1)$$

$$+ [1-h(t+1)]\log[1-h(t+1)]\}. \qquad (6.7)$$

Taking t in the place of $(t+1)$ in (6.7) we get (6.4).

Theorem 6.2

The $H(f,t)$, considered in theorem 6.1, uniquely determines the distribution.

Proof:

Substituting for $H(f,t-1)$ in the recurrence relation (6.4) we get

$$H(f,t) = \frac{1}{1-h(t)}\,\frac{1}{1-h(t-1)}H(f,t-2)$$

$$+ \frac{h(t-1)\log h(t-1)+[1-h(t-1)]\log[1-h(t-1)]}{[1-h(t)][1-h(t-1)]}$$

$$+ \frac{h(t)\log h(t)+[1-h(t)]\log[1-h(t)]}{1-h(t)}.$$

Proceeding recursively, we get

$$H(f,t) = \frac{H(f,-1)}{[1-h(t)][1-h(t-1)]\ldots[1-h(0)]}$$

$$+ \frac{h(0)\log h(0) + [1-h(0)]\log[1-h(0)]}{[1-h(t)][1-h(t-1)]\ldots[1-h(0)]}$$

$$+ \frac{h(1)\log h(1) + [1-h(1)]\log[1-h(1)]}{[1-h(t)][1-h(t-1)]\ldots[1-h(1)]}$$

$$+ \ldots$$

$$+ \frac{h(t)\log h(t) + [1-h(t)]\log[1-h(t)]}{[1-h(t)]}$$

$$= \frac{H(f)}{\prod_{x=0}^{t}[1-h(x)]} + \frac{h(0)\log h(0) + [1-h(0)]\log[1-h(0)]}{\prod_{x=0}^{t}[1-h(x)]}$$

$$+ \frac{h(1)\log h(1) + [1-h(1)]\log[1-h(1)]}{\prod_{x=1}^{t}[1-h(x)]} + \ldots$$

$$+ \frac{h(t)\log h(t) + [1-h(t)]\log[1-h(t)]}{[1-h(t)]} \tag{6.8}$$

where $H(f)$ is the Shannon's entropy associated with X. Since $h(t)$ determines the distribution uniquely in view of (6.8), $H(f,t)$ also determines the distribution uniquely.

6.3 Characterization Theorem

We now look into the situation where the residual entropy function is constant.

Theorem 6.3

Let X be a discrete random variable in the support of the set of non negative integers with residual entropy function $H(f,t)$. The relation

$$H(f,t)=c \tag{6.9}$$

where c is a constant holds for all integers $t \geq 0$ if and only if X follows the geometric distribution.

Proof:

When X follows the geometric distribution with probability mass function

$$f(x) = q^x p, \quad x = 0,1,2,\ldots, \; 0<p<1, \; p+q=1,$$

direct calculation using (6.3) gives

$$H(f,t) = -\frac{1}{p}[p\log p + (1-p)\log(1-p)] \;, \tag{6.10}$$

so that the conditions of the theorem are satisfied.

Conversely when (6.9) holds, (6.4) takes the form

$$h(x)\log h(x) + [1-h(x)]\log[1-h(x)] + ch(x) = 0, \quad x=1,2,3,\ldots \tag{6.11}$$

Let t_1 and t_2 be two positive integers such that $t_1<t_2$. Denote by

$$A(t) = \frac{t_2 - t}{t_2 - t_1} h(t_1) + \frac{t - t_1}{t_2 - t_1} h(t_2), \quad t_1 \leq t \leq t_2. \tag{6.12}$$

Consider

$$B(t) = A(t)\log A(t) + [1-A(t)]\log[1-A(t)] + cA(t). \qquad (6.13)$$

From (6.11) and (6.12) we have

$$B(t_1) = B(t_2) = 0.$$

By mean value theorem there exists an $x_0 \in (t_1, t_2)$ such that

$$B'(x_0) = \frac{B(t_2) - B(t_1)}{t_2 - t_1} = 0. \qquad (6.14)$$

But from (6.13)

$$B'(x_0) = A'(x_0)\left[\log\frac{A(x_0)}{1 - A(x_0)} + c\right]. \qquad (6.15)$$

Assume that

$$A'(x_0) \neq 0. \qquad (6.16)$$

From (6.14) and (6.15) we get

$$A(x_0) = \frac{e^{-c}}{1 + e^{-c}}. \qquad (6.17)$$

Without loss of generality assume

$$h(t_1) < h(t_2).$$

Since $A(t)$ is the equation of the line segment joining $(t_1, h(t_1))$ and $(t_2, h(t_2))$ we have

$$h(t_1) \leq A(x_0) \leq h(t_2)$$

or

$$h(t_1) \leq \frac{e^{-c}}{1 + e^{-c}} \leq h(t_2). \qquad (6.18)$$

Now from (6.11)

$$e^{-c} = h(x)\,[1-h(x)]^{\frac{1-h(x)}{h(x)}}$$

and

$$\frac{e^{-c}}{1+e^{-c}} = \frac{h(x)[1-h(x)]^{\frac{1-h(x)}{h(x)}}}{1+h(x)[1-h(x)]^{\frac{1-h(x)}{h(x)}}} \le h(x), \text{ for all x.} \qquad (6.19)$$

From (6.18) and (6.19) we get

$$h(t_1) = \frac{e^{-c}}{1+e^{-c}}.$$

From (6.11) with $x = t_1$ we have

$$\frac{e^{-c}}{1+e^{-c}} = \frac{h(t_1)[1-h(t_1)]^{\frac{1-h(t_1)}{h(t_1)}}}{1+h(t_1)[1-h(t_1)]^{\frac{1-h(t_1)}{h(t_1)}}}.$$

This gives

$$h(t_1) = \frac{h(t_1)[1-h(t_1)]^{\frac{1-h(t_1)}{h(t_1)}}}{1+h(t_1)[1-h(t_1)]^{\frac{1-h(t_1)}{h(t_1)}}}$$

which is not true since $h(t_1) \neq 0$.

Therefore, in (6.15), $\left[\log\frac{A(x_0)}{1-A(x_0)}+c\right]$ cannot be zero.

Hence

$$A'(x_0)=0.$$

That is

$$\frac{h(t_2)-h(t_1)}{t_2-t_1} = 0.$$

This implies $h(t_1) = h(t_2)$ for all $t_1>t_2$. Proceeding on similar lines with $t_1<t_2$ we can observe that $h(t_1) = h(t_2)$ for all $t_1<t_2$. This implies that $h(t)$ is a constant.

Since the constancy of failure rate is characteristic to the geometric distribution, X follows the geometric distribution.

For a random variable $(X-t)$ truncated at $t(>0)$ in the support of non-negative integers, the maximum entropy probability distribution under the condition that the arithmetic mean is fixed is the geometric distribution. An upperbound for the residual entropy function in terms of the mean residual life function can be obtained from the above.

Theorem 6.4

Given that $r(t)<\infty$, then

$$H(f, t) \le r(t) \log r(t) - [r(t)-1] \log[r(t)-1].$$

Proof:

From Kapur (1989) for a random variable X in the support of non-negative integers, the maximum entropy probability distribution subject to the condition that the arithmetic mean is fixed is the Geometric distribution. Hence given $r(t)$, if the domain is restricted to the set of non-negative integers, the maximum entropy occurs when the underlying distribution is geometric with mean $r(t)$.

From (6.10) we have,

$$H(f, t) \le -\frac{1}{p}[p\log p + (1-p)\log(1-p)]$$

$$= -r(t)\left[\frac{1}{r(t)}\log\frac{1}{r(t)} + \left(1-\frac{1}{r(t)}\right)\log\left(1-\frac{1}{r(t)}\right)\right]$$

or

$$H(f, t) \le r(t) \log r(t) - [r(t)-1] \log[r(t)-1].$$

It may be noticed that $H(f, t)<\infty$ whenever $r(t)<\infty$.

Plan for future Study

Several problems have opened out during the present investigation. The problem of extending the concept of the residual entropy function to higher dimensions is yet to be examined. Characterizations of some bivariate distributions based on the functional form of the bivariate residual entropy function can be obtained, analogous to that of bivariate failure rate. The problem of obtaining distributions which maximizes the residual entropy function under different set of constraints is to be studied in detail. Recently Ebrahimi and Kirmani (1996a) has studied the truncated version of the Kullback-Leibler measure of directed divergence information measure. Characterization of certain bivariate models using the above also seems to be in order. The problem of estimating the residual entropy function using standard procedures and their comparisons is yet another problem to be examined. These works are proposed to be under taken in a future study.

REFERENCES

1. Abouammoh, A.M. (1990), Partial ordering of discrete life distributions based on ageing, *Pakistan J. of Statist.*, 6, 25-45

2. Abrahams, J. and Thomas, J.B. (1984), A note on the characterization of bivariate densities by conditional densities, *Commun. Statist. Theory and Methods*, 13(3), 395-400

3. Arnold, B.C. (1987), Bivariate distributions with Pareto conditionals, *Statist. and Prob. Lett.,* 5, 263-266.

4. Arnold, B.C. and Press, S.J. (1988), Compatible conditional distributions, *J. Amer. Statist. Assoc.*, 84, 152-156.

5. Arnold, B.C. and Strauss, D. (1988), Bivariate distributions with exponential conditionals, *J. Amer. Statist. Assoc.*, 83, 522-527.

6. Aczel, J. and Daroczy,Z. (1975), **On measure of information and their characterizations**, Academic Press, New York

7. Ash, R.B. (1965), **Information theory**, Wiley, New York

8. Azlarov T.A. Volodin, N.A. (1986), **Characterization problems associated with the Exponential distribution**, Springer Verlag, New York.

9. Belzunce, F, Candel, J and Ruiz, J.M. (1995), Ordering of truncated distributions through concentration curves, *Sankhya,* Ser. A, 57, 375-383.

10. Behra, M and Chawla, J.S. (1974), Generalized gamma entropy, In Kapur (1989).

11. Behra, M. (1990). **Additive and non- additive measures of entropy**, Wiley Eastern, New York.

12. Chaundy, T. W and Macleod, J.B (1960), On a functional equation, *Proc. Edinburgh Math. Soc. Notes*, 43, 7-8.

13. Cox, D.R.(1972), Regression models and life tables, *J. Roy. Statist. Soc.*, Ser. B,34,187-220.

14.Cozzolino, J,.M. and Zahner, M.J. (1973), The maximum entropy distribution of future market price of a stock, *Oper. Res.*, 21 (6), 1200-1211.

15.Dallas, A.C. (1975), On a characterization by conditional variance, In Galambos and Kotz (1978).

16.Ebrahimi, N. (1986), Classes of discrete decreasing and increasing mean residual life distributions, *IEEE Trans. Rel.*, 35, 403-405.

17.Ebrahimi , N. and Kirmani, S.N.U.A. (1996a), Some results on ordering of survival function through uncertainty, *Statist. Probab. Lett.*, 29, 167-176.

18.Ebrahimi, N. and Kirmani, S.N.U.A (1996b), A measure of discrimination between two residual life time distributions and its applications, *Ann. Math. Statist.*, 48, 257-265.

19.Ebrahimi, N. (1996), How to measure uncertainty in the residual life time distribution, *Sankhya* A, 58, 48-56.

20.Ebrahimi, N.. and Pellerey F. (1995), New partial ordering of survival functions based on the notion of uncertainty, *J. App. Prob.*, 32, 202-211.

21.Fadeev, D.K. (1956), On the concept of entropy of a finite probabilistic scheme, Uspeki. Mat. Nauk, 11(1), 227-231.

22.Fisher, A. (1958), **Foundations of information theory**, McGraw Hill, New Delhi.

23.Galambos J. (1992), **In Handbook of the Logistic Distribution**, edited by N. Balakrishnan, Marcel Dekker, New York.

24.Galambos, J. and Kotz, S. (1978), **Characterizations of probability distributions**, Springer- Verlag, Berlin.

25.Gelman, A. and Speed, T.P. (1993), Characterizing a joint probability distribution by conditionals, *J. Roy. Statist. Soc.*, Ser. B, 55(1), 185 –188.

26.Geetha, K.G. and Nair , V.K.R. (1997), A class of bivariate distributions with specified conditionals, *Far East J. Theo. Statist.*, 1(1), 63-68.

27.Golan, A. Judge, G. and Miller, D.M. (1996), Maximum Entropy Econometrics: Robust Estimation with Limited Data, John Wiley and Sons, New York.

28.Gompertz, B. (1825), On the nature of the function expressive of the law of human mortality, In Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995), **Continuous univariate distributions**, Vol.2, John Wiley and Sons, New York.

29.Guerrieri, G. (1965), Some characteristic properties of the exponential distribution, *G. Economist*, 24,427-437.

30.Guess, F.M. and Park, D.H. (1988), Modelling discrete bathtub and upside down bath tub mean residual life functions, *IEEE Tran. Rel.*, 35,545-549.

31.Guiasu, S. (1977), **Information theory with applications**, McGraw Hill International Book Company, New York.

32.Gupta, R.C. (1981), On the mean residual life function in survival studies. In Taillie, C. Patil, G.P. and Baldessari, B.A. (Eds), **Statistical distributions in scientific work**, Vol.5, Reidel, Dordrecht – Boston, 327-334.

33.Gupta, P.L and Gupta, R.C. (1983), On the moments of residual life in reliability and some characterization results, *Commu. Statist. Theory and Methods*, 12, 449-461.

34.Gupta, R.C. and Kirmani, S.N.U.A. (1987), On order relation between reliability measures, *Commu. Statist. Stochastic Models*, 3, 149-156.

35.Gupta, R.C. and Kirmani, S.N.U.A. (1990), The role of weighted distribution on stochastic modelling, *Commu. Statist. Theory and Methods*, 19, 3147-3162.

36.Hart, P.E. (1971), Entropy and other measures of concentration, *J. Roy. Statist. Soc., Ser* A, 134, 73-85.

37.Hartley, R.V.L. (1928), Transmission of information, *Bell System Tech. J.*, 7, 535-563.

38.Havrada, J. and Charvat, P. (1967), Quantification method of classification processes concept of structural entropy, *Kybernetika*, 3, 30-35.

39.Hitha, N. (1991), **Some characterizations of Pareto and related populations**, PhD Thesis, Cochin University of science and Technology, India.

40.Hitha, N. and Nair, N.U. (1989), Characterizations of some discrete models of properties of residual life function, *Cal. Statist. Assoc. Bull.*, 39, 219-225.

41.Hitha, N. and Nair, N.U. (1991), Characterization of bivariate Lomax and finite range distributions, *Cal. Statist. Assoc. Bull.*, 163-167.

42.Johnson, N.L. and Kotz,S. (1975), A vector valued multivariate hazard rate , *J. Multiv. Anal.*, 5,53-66.

43.Kannappan, P.I. and Rathie, P.N. (1973), On characterization of directed divergence, *Information and Control*, 22, 163-171.

44.Kapur J.N. (1968), On information of order α and β, *Proc. Ind. Acad. Sci.*, 48(A), 65-75.

45.Kapur, J.N. (1989), **Maximum entropy models in science and engineering**, Wiley Eastern Limited.

46.Kapur. J.N. (1994), **Measures of information and their application**, Wiley Eastern Limited.

47.Khinchin, A.J. (1953), The concept on entropy in the theory of probability, *Uspeki. Math. Nauk.*, 8, 3-20.

48.Kotz, S. and Shanbhag, D.N. (1980), Some new approaches to probability distributions, *Adv. Appl. Prob.*, 12, 903-921.

49.Kullback, S. and Leibler, R.A. (1951) On information and sufficiency, *Ann. Math. Statist.*, 22, 79-86.

50.Kupka, J. and Loo, S. (1989), The hazard and vitality measures of ageing, *J. Appl. Prob.*, 26, 532-592.

51.Launer, R.L. (1984), Inequalities for NBUE and NWUE life distribution, *Oper. Res.*, 32, 660-667.

52.Laurent, A.G. (1974), On characterization of some distribution by truncation, *J. Amer. Statist. Assoc.*, 69, 823-827.

53.Lawless, J. F. (1982), **Statistical models and methods for life time data**, John Wiley, New York.

54.Lee, P.M. (1964), On the axioms of information theory, *Ann. Math. Statist.*, 35, 415-418.

55.Mi, J. (1993), Discrete bath tub failure rate and upside down bathtub mean residual life, *Naval Res. Log. Qrt.*, 40, 361-371.

56.Morrison, D.G. (1978), On linearly increasing mean residual life times, *J. Appl. Prob.*, 15, 617- 620.

57.Mukherjee, S.P. and Roy, D. (1986), Some characterization of the exponential and related life distributions, *Cal. Statist. Assoc. Bull.*, 35, 189-197.

58.Muth, E.J. (1980), Memory as a property of probability distribution, *IEEE Trans. Rel.*, 29, 160-165.

59.Nagaraja, H.N. (1975), Characterization of some distributions by conditional moments, *J. Ind. Statist. Assoc.*, 13, 57-61.

60.Nair, N.U. (1983), A measure of memory for some discrete distributions, *J. Ind. Statist. Assoc.*, 21, 141-147.

61.Nair, K.R.M. and Rajesh, G. (1998) Characterization of probability distributions using the residual entropy function, *J. Ind. Satist. Assoc.*, 36, 157-166.

62.Nair, K.R.M. and Rajesh, G. (2000), Geometric vitality function and its application to reliability, *IAPQR Tran.*, 25(1) 1-8.

63.Ord, J.K. , Patil, G.P. and Taillie, C. (1983), Truncated distribution and measures of income inequality, *Sankhyā*, Ser B, 45, 413-430.

64.Osaki, S. and Li, X. (1988), Characterization of gamma and negative binomial distributions, *IEEE Trans. Rel.*, 37, 379-382

65.Rajesh, G. and Nair, K.R.M (1998), Residual entropy function in discrete time, *Far East. J. Theor. Statist.*, 2(1), 1-10.

66.Rajesh, G. and Nair, K.R.M. (2000), Residual entropy of conditional distributions, *Statist. Meth.*, 2 (1), 72-80.

67.Reinhardt, H.E. (1968), Characterizing the exponential distribution, *Biometrics*, 24, 437-438.

68.Renyi, A. (1961), On measures of entropy and information, *Proc. Fourth Berkely Symp. Math.*, Statistics and Probability. 1960. University of California Press, Vol. 1, 547-561.

69.Reza, F.M. (1961), **An introduction to information theory**, McGraw Hill Book company Inc.

70.Roy, D. (1989), A characterization of Gumbel's bivariate exponential and Lindley and Singpurwalla's bivarate Lomax distribution, *J. Appl. Prob.*, 27, 886-891.

71.Roy, D and Gupta, R.P., (1992), Classification of discrete lives, *Micro Electron Reliability*, 31, 1459-1473.

72.Roy D, and Mukherjee, S.P. (1989), Characterization based on arithmetic, geometric and harmonic means of failure rates, *Contribution to Stochastic*, 178-185.

73.Salvia, A.A. and Bollinger, R.C. (1982), On discrete hazard functions, *IEEE Trans. Rel.*, 31, 458-459.

74.Sankaran, P.G. (1992), Characterization by properties of probability distributions by reliability concepts, Ph. D. thesis, Cochin University of Science and Technology.

75.Sankaran, P.G. and Gupta, R.P. (1999), Characterization of life time distribution using measure of uncertainty, *Cal. Statist. Assoc. Bull.*, 49,159-166.

76.Shaked, M., Shanthikumar, J.G. and Torres, J.B.V (1995), Discrete Hazard rate functions, pre print (to appear in *Comp. Oper. Res.*).

77.Shanbhag, D.N. (1970), Characterization for exponential and geometric distributions, *J. Amer. Statist. Assoc.*, 65, 1256-1259.

78.Shannon, C.E. (1948), A Mathematical theory of communication, *Bell system Tech. J.*, 279-423, 623-656.

79.Sharma, B.D. and Mittal, D.P. (1975), New non-additive measures of entropy for discrete probability distributions, *J. Math. Sciences*, 10, 28-40.

80.Singh, S.K. and Maddala, G.S. (1976), A function for size distribution of incomes, *Econometrica*, 44, 963-970.

81.Sullo, P. and Rutherford, D.K. (1977), Characterization of the Power distribution by conditional exceedences, *The American Association Proceeding of the Business and Economics*, 787-792.

82.Swartz, G.B. (1973), The mean residual life time function, *IEEE Tran. Rel.*, 22, 108-109.

83.Theil, H. (1967), **Economics and information theory**, North Holland, Amsterdam.

84.Thilanas, C. B. and Theil, H. (1965), Information approaches to the valuation of input output forecasts, *Econometrica*, 33, 847-862.

85.Tverberg, H. (1958), A new derivation of the information function, Math. Scand., 6, 297-298.

86.Vartak, M. N. (1974), Characterization of certain classes of probability distributions,. *J. Ind. Statist. Assoc.*, 12, 63-74.

87.Wiener, N. (1948), *Cybernetics*, John Wiley, New York.

88.Xekalaki, E. (1983), Hazard functions and life distributions in discrete time, *Commun. Statist. Theory and Methods*, 12, 2503-2509.

Appendix

Distribution	Probability density function $f(x)$	Residual entropy function $H(f, t)$
Exponential distribution	$\lambda \exp\{-\lambda x\}$, $x>0$, $\lambda>0$	$1-\log \lambda$
Pareto type I distribution	$a\, \beta^a\, x^{-(a+1)}$, $x \geq a$, $\beta>0$, $a>0$	$1+\frac{1}{a} - \log\left(\frac{a}{t}\right)$
Pareto type II distribution	$a\, c^a\, (x+c)^{-(a+1)}$, $x>0$, $c>0$, $a>0$	$1+\frac{1}{a} - \log\left(\frac{a}{c+t}\right)$
Beta distribution	$cR^{-c}\, (R-x)^{+c}$, $0<x<R$, $c>0$	$1 - \frac{1}{c} - \log\left(\frac{c}{R-t}\right)$
Logistic distribution	$\frac{\exp\{-(x-a)/b\}}{b[\exp\{-(x-a)/b\}]^2}$, $-\infty<x<\infty$, $b>0$	$\log b - \log[1+\exp\{-((t-a)/b)\}]\, \exp\{-((t-a)/b)\} + 2$
Uniform distribution	$(b-a)^{-1}$, $a \leq x \leq b$	$2 - \log(b-t)$
Power function distribution	$\frac{cx^{c-1}}{b^c}$, $b>0$, $a \leq x \leq b$	$\frac{c-1}{c^2} - \log(b^c - t^c) - [b^c \log cb^{c-1} - t^c \log(ct^{c-1})]/(b^c - t^c)$
Extreme value type I distribution	$\frac{1}{b}\exp\left\{\frac{x-a}{b}\right\} \exp\left[-\exp\left\{\frac{x-a}{b}\right\}\right]$, $-\infty<x<\infty$, $b>0$	$1-\exp\left[-\exp\left\{\frac{x-a}{b}\right\}\right] \left\{\log\left[1-\exp\left[-\exp\left\{\frac{t-a}{b}\right\}\right]\right]\right\}$ $+\Gamma\left[0, \exp\left\{\frac{t-a}{b}\right\}\right] + \left\{\log\left[b \exp\left[-\exp\left\{\frac{t-a}{b}\right\} - 1\right]\right]\right\}$ $+\frac{t-a}{b} \exp\left[-\exp\left\{\frac{t-a}{b}\right\}\right]$
Weibull distribution	$\frac{cx^{c-1}}{b^c} \exp\left\{-\left(\frac{x}{b}\right)^c\right\}$, $x>0$, $b, c>0$	$1-\exp\left(\frac{t}{b}\right)^c \Gamma\left(0, \left(\frac{t}{b}\right)^c\right) + \log\left(\frac{b^c}{ct^c - 1}\right)$
Geometric distribution	$p\,(1-p)^x$, $x = 0, 1, 2, \ldots$, $0<p<1$	$-\frac{1}{p}[p \log p + (1-p)\log(1-p)]$