A STUDY ON THE REVERSED LACK OF MEMORY PROPERTY AND ITS GENERALIZATIONS

Thesis Submitted to the
Cochin University of Science and Technology
for the Award of the Degree of

Doctor of Philosophy
under the Faculty of Science

by

REJEESH C JOHN

Department of Statistics

Cochin University of Science and Technology

Kochi-682022

FEBRUARY 2011

To My Loving Parents

CERTIFICATE

Certified that the thesis entitled **"A Study on the Reversed Lack of Memory Property and its Generalizations"** is a bonafide record of work done by Shri. Rejeesh C. John under my guidance in the Department of Statistics, Cochin University of Science and Technology and that no part of it has been included anywhere previously for the award of any degree or title.

Kochi- 22

23 February 2011

Dr. Asha Gopalakrishnan
Associate Professor,
Department of Statistics,
Cochin University of
Science and Technology.

DECLARATION

This thesis contains no material which has been accepted for the award of any other Degree or Diploma in any University and to the best of my knowledge and belief, it contains no material previously published by any other person, except where due references are made in the text of the thesis.

Kochi- 22 **Rejeesh C. John**

23 February 2011

Acknowledgements

I wish to express my deep sense of respect and gratitude to my supervising guide Dr. Asha Gopalakrishnan, Associate Professor, Department of Statistics, Cochin University of Science and Technology(CUSAT), who has been a constant source of inspiration during the course of my Ph.D work. She has always been patient towards my shortcomings and kept encouraging me to work in a better way. Without her help and support, perhaps, I would not have been able to write this thesis.

I am obliged to Dr. N. Balakrishna, Professor and Head, Department of Statistics, CUSAT for his extensive support.

I take this opportunity to record my sincere respect and gratitude to the Management and Staff, Nirmalagiri College, Kuthuparamba who had kindly agreed to go on with my Ph.D work. Without their kind cooperation, it would have been almost impossible for me to continue the research work.

I wish to express my sincere thanks to all the faculty members of the Department of Statistics, CUSAT for their valuable comments and encouragement. I offer my regards to all non-teaching staff, Department of Statistics, CUSAT for their kind cooperation.

I express my gratitude to all my family members who have provided me with their helping hands in each and every walk of my life.

I owe my appreciation and thanks to University Grants Commision(UGC), Government of India, for backing me with the Faculty Devolopment Programme(FDP) to complete my Ph.D.

I am greatly indebted to all my teachers, research scholars and friends for their untiring support and constant encouragement on which I thrived to the end of my work.

Above all , to the Almighty for the blessings showered upon me during the times of difficulties. I offer my salutations to Him for giving me strength to bring this thesis into fruition.

Rejeesh C. John

Contents

Chapter 1

Preliminaries

1.1 Introduction

In the majority cases of analyzing statistical information, a fundamental crisis that emerges at the beginning is the identification of a suitable model that can explain the real condition which generated the observations. Once the exact model is known the original problem permits analysis with minor effort, as the properties of the model comes helpful to the analyst in drawing inferences and decisions. Due to the accessibility of a huge number of probability distributions at disposal, very frequently the choice of exacting one in an unambiguous situation turns out to be hard, unless one has a reasonable basis or criteria that give good reason for the choice. A universal move toward to this problem is to make use of experimental methods such as probability plots or goodness-of-fit tests while another is, to apply some approximation theorems from probability theory. Even though sometimes such considerations may guide to reasonable models, neither are they of general applicability nor do they promise the correct result all the time. The only means that enables the determination of a probability model precisely is a characterization theorem and as a result the study of such theorems has emerged as an essential area of mathematical statistics. It is also general that many such theorems are found useful from theoretical considerations as well.

One such characterization that has been popular is the lack of memory property of the exponential distribution. This property along with its generalizations and modifications was a subject of study for many researchers. This property manifests in many properties of the exponential distribution. For instance, the constancy of important reliability concepts like hazard rate and mean residual life function. The importance of hazard rate and mean residual life in reliability added to the popularity of this characterization and thus evolved several multivariate generalizations of the lack of memory property.

If X is a nonnegative random variable possessing absolutely continuous distribution with respect to Lebesgue measure, we say that the random variable X or its distribution has lack of memory property if for all x, $y \geq 0$ such that $P(X \geq y) > 0$,

$$P\left(X \geq x + y | X \geq y\right) = P\left(X \geq x\right) \tag{1.1}$$

In expressions of the survival function of the random variable,

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

(1.1) can be restated as

$$R(x + y) = R(x)R(y). \tag{1.2}$$

For an absolutely continuous survival function $R(x)$, its hazard rate $h(x)$ is defined as

$$h(x) = \frac{-d \log R(x)}{dx}. \tag{1.3}$$

The lack of memory property is equivalent to the statement,

$$h(x) = c, \text{ a constant.}$$

Further, the truncated mean or mean residual life is defined as

$$r(x) = E\left(X - x | X \geq x\right) \\ = \frac{1}{R(x)} \int_x^\infty R(t)dt, \tag{1.4}$$

often interpreted as the average lifetime remaining to a component at age x, is connected to the hazard rate through the equation

$$h(x) = \frac{1}{r(x)} \left\{1 + \frac{dr(x)}{dx}\right\}. \tag{1.5}$$

It is given in Cox (1962) that for the exponential distribution,

$r(x)$= k, a constant.

Galambos and Kotz (1978) established the similarity of lack of memory property, constancy of the hazard rate and constancy of the mean residual life.

The hazard rate is found valuable in the investigation of right censored data. But in a number of situations, we come across the left censored data. Left censoring occurs in life test applications when a unit has failed at the time of its first assessment; we know only that the unit failed earlier than the assessment time. In other situations, left censored observations occur when the precise value of a response has not been observed and we have, instead, an upper bound on that response. Consider, for example, a measuring device that lacks the sensitivity required to determine the observations below a known threshold. When the dimension is taken, if the signal is below the instrument threshold, we know only that the measurement is less than the threshold.

In such situations, the reversed hazard rate was found to be more adequate than the hazard rate (Block et al. (1998), Andersen et al. (1993), Gupta and Hann (2001)). This paved way of studying many reliability concepts in the reversed time scale. These measures gained attention as they were not just ”duals” of the existing probability and reliability measures but they found use and applications in the field of actuaries, biometry, maintenance theory, economics etc. in their own right. Motivated by this, in the thesis we subject the lack of memory property to the reversed time scale to develop a new property. We call this property the reversed lack of memory property and show that it is radically different from the lack of memory property. Before getting into the details we first consider the basic concepts required in the ensuing discussions.

1.2 Basics

Consider a random variable X with an absolutely continuous cumulative distribution function $F(x) = P(X \leq x)$, survival function $R(x) = 1 - F(x)$ and probability density function $f(x)$. Let $a = \inf \{x|F(x) > 0\}$ and $b = \sup \{x|F(x) < 1\}$. Then (a, b), $-\infty \leq a < b < \infty$ is the interval of support of X. The distribution function $F(x)$ is a non-decreasing continuous function with $F(a) = 0$ and $\lim_{x \to b} F(x) = 1$. The probability density function of X may be represented as

$$f(x) = \frac{dF(x)}{dx}.$$

1.2.1 Hazard Rate

An essential function that characterizes lifetime distributions is the hazard rate $h(x)$, defined as

$$h(x) = \lim_{\Delta x \to 0} \frac{P(x \leq X < x + \Delta x|X \geq x)}{\Delta x}.$$

The hazard rate specifies the instantaneous rate of death or failure at x, given that the individual survives up to x. Thus $h(x)\Delta x$ is the approximate probability of death in the interval $[x, x + \Delta x)$, given survival up to x. The hazard rate is also known as conditional failure rate in reliability, the force of mortality in demography, the intensity function in stochastic processes, the specific failure rate in epidemiology, the inverse of Mill's ratio in economics or simply the hazard function. When the probability density function of X, $f(x)$ exists, then the hazard rate is expressed as

$$\begin{aligned} h(x) &= \frac{f(x)}{R(x)} \\ &= -\frac{d\log R(x)}{dx}. \end{aligned}$$

The hazard rate completely specifies the distribution of X and determines the survival function. Integrating the above equation with respect to x and using $R(a) = 1$, we obtain

$$R(x) = \exp\left\{-\int_a^x h(u)du\right\}.$$

Thus, the probability density function of X can be obtained as

$$f(x) = h(x)\exp\left\{-\int_a^t h(u)du\right\}.$$

A related function is cumulative hazard rate $H(x)$, defined as

$$H(x) = \int_a^t h(u)du.$$

Then, $R(x)$ can be represented in terms of $H(x)$ as

$$R(x) = \exp\left\{-H(x)\right\}.$$

1.2.2 Reversed Hazard Rate

Recently, another concept, that is valuable in the survival studies, is developed which is referred as the reversed hazard rate. The reversed hazard rate of X is defined for $x > a$ as

$$\lambda(x) = \lim_{\Delta x \to 0} \frac{P\left(x - \Delta x < X \leq x | X \leq x\right)}{\Delta x}. \tag{1.6}$$

That is, in a small interval, $\lambda(x)\Delta x$ is the approximate probability of failure in the interval $(x - \Delta x, x]$, given failure before the end of the interval. Reversed hazard rate was proposed as a dual to the hazard rate by Barlow et al. (1963). When the probability density function of X, $f(x)$ exists, (1.6) can be expressed as

$$\begin{aligned}\lambda(x) &= \frac{f(x)}{F(x)} \\ &= \frac{d\log F(x)}{dx}.\end{aligned} \tag{1.7}$$

The reversed hazard rate, $\lambda(x)$ determines the distribution function uniquely by the relationship

$$F(x) = \exp\left\{-\int_x^b \lambda(u)du\right\},$$

which was given by Keilson and Sumita (1982). The probability density function of X can be obtained from $\lambda(x)$ using the relationship

$$f(x) = \lambda(x)\exp\left\{-\int_x^b \lambda(u)du\right\}.$$

The cumulative reversed hazard rate $\Lambda(x)$ is defined as

$$\Lambda(x) = \int_x^b \lambda(u)du.$$

Then, $F(x)$ can be represented in terms of $\Lambda(x)$ as

$$F(x) = \exp\left\{-\Lambda(x)\right\}.$$

Further, it is seen that the reversed hazard rate of $-X$ is same as the hazard rate $h(-x)$ for $x \in (-b, -a)$. Also, if $E(X) < \infty$ and $\nu(x) = E(X|X > x)$ is the vitality function of X and $\eta(x) = E(X|X \leq x)$ is the conditional expectation, then (Nair et al. (2005))

$$\frac{\nu(x) - E(X)}{h(x)} = \frac{E(X) - \eta(x)}{\lambda(x)}$$

for all $x \in (a, b)$.

The above relation enables one to translate characterizations in terms of $h(x)$ and $\nu(x)$ (Ahmed (1991), Osaki and Li (1988), Ruiz and Navarro (1994), Nair and Sankaran (1991)) to that between $\lambda(x)$ and $\eta(x)$. Thus, one would expect dual results to exist for the reversed hazard rates.

Ware and DeMets (1976) advocated reversed hazard rate for the estimation of the distribution function in the presence of left censored observations. Shaked and Shanthikumar (1994) presented a number of results based on reversed hazard rate ordering and characterization of lifetime distributions based on reversed hazard rate. Block et al. (1998) pointed out that there is no non-negative random variable having an increasing reversed hazard rate distribution and observed that increasing hazard rate distributions like Weibull, gamma and lognormal distributions are decreasing reversed hazard rate distributions. Block et al. (1998) characterized properties for k out of n systems in terms of reversed hazard rate. Kijima (1998) proved that if an irreducible Markov chain in continuous time is monotone in the sense of reversed hazard rate ordering then it must be skip-free to the left. A birth-death process is then characterized as a continuous time Markov chain that is monotone in the sense of reversed hazard rate orderings. Bloch-Mercier (2001) applied the reversed hazard rate orderings in reliability. Chandra and Roy (2001) have considered different implicative relationships with respect to the monotonic behavior of reversed hazard rate. Finkelstein (2002) expressed the relation between hazard rate $h(x)$ and reversed hazard rate $\lambda(x)$ as

$$\begin{aligned}\lambda(x) &= \frac{h(x)R(x)}{F(x)} \\ &= \frac{h(x)}{\exp\left\{\int_a^x h(u)du\right\} - 1}.\end{aligned}$$

Lawless (2003) developed nonparametric estimators of $R(x)$ for the right truncated observations using reversed hazard rates. Reversed hazard rate is useful in forensic science and actuarial science, as the time elapsed since failure is a quantity of interest in order to predict the actual time of failure (Nanda et al. (2003)). Gupta et al. (2006)

studied the monotonicity of the reversed hazard rate of the maximum for two well known bivariate distributions viz the Farlie – Gumbel – Morgenstern (FGM) and Sarmanov family. Increasing hazard rate and decreasing reversed hazard rate properties of the minimum and maximum of multivariate distributions with log-concave densities are studied by Hu and Li (2007). Li et al. (2010) dealt with the reversed hazard rate of general mixture models. They also studied the dependence and monotone properties of the reversed hazard rate. For more properties and applications of reversed hazard rate function, one could refer to Kalbfleisch and Lawless (1989), Gupta and Nanda (2001), Gupta and Wu (2001), Nair and Asha (2004), Chandra and Roy (2005), Nair et al. (2005), Bartoszewicz and Skolimowska (2006) and Sankaran and Gleeja (2007).

1.2.3 Proportional Hazards Model

Cox (1972) defined proportional hazards model as

$$h(x) = \phi h_0(x),$$

where $h_0(x)$ is an arbitrary baseline hazard rate and ϕ is some positive real constant of proportionality and is a measure of relative risk.

Here, the survival functions can be related as

$$R(x) = [R_0(x)]^{\phi},$$

where $R_0(x)$ is the baseline survival function. The class of models provided by this process is sometime referred to as the Lehmann class (Lehmann (1953)). For a comprehensive review on this topic, one can refer Kalbfleisch and Prentice (2002) and Lawless (2003).

As an example, for some positive integer value of ϕ, if $X_1, X_2, ..., X_\phi$ are independently and identically distributed random variables with survival function $R_0(x)$ representing the lifetime of components, in a ϕ-component series system, then the lifetime of the system is given by $X = \min(X_1, X_2, ..., X_\phi)$ with survival function $R(x)$ given by $R(x) = [R_0(x)]^\phi$.

1.2.4 Proportional Reversed Hazards Model

Gupta et al. (1998) proposed a dual model called proportional reversed hazards model, which is expressed as

$$\lambda(x) = \theta\lambda_0(x), \tag{1.8}$$

where $\theta > 0$ and λ_0 is the baseline reversed hazard rate. Then the relation between distribution functions can be expressed as

$$F(x) = [F_0(x)]^\theta,$$

where $F_0(x)$ is the baseline distribution function.

As an example, for some positive integer value of θ, if $X_1, X_2, ..., X_\theta$ are independent and identically distributed random variables with distribution function $F_0(x)$ representing the lifetime of components, in a θ-component parallel system, then the lifetime of the system is given by $X = \max(X_1, X_2, ..., X_\theta)$ with distribution function $F(x)$ given by $F(x) = [F_0(x)]^\theta$.

The proportional reversed hazards model has strong resemblance with the proportional hazards model, but is appropriate in situations where proportional hazards model becomes unsuitable. For example, the model (1.8) is helpful in the analysis of left censored or right truncated data. Gupta et al. (1998) and Gupta and Gupta (2007) studied the monotonicity of hazard rate and reversed hazard rate of the model (1.8). The properties based on stochastic comparisons and results related to ageing notions of random lifetimes are given in Di Crescenzo (2000). Chen et al. (2004) employed the proportional reversed hazard rate models to study the longitudinal pattern of recurrent gap times. Further, Chen et al. (2004) introduced the concept of frailties in proportional reversed hazard rate models. The applications and methods of inference of the model (1.8) are examined in Sengupta et al. (1998) and Gupta and Gupta (2007). Sankaran and Gleeja (2008) derived a class of bivariate distributions having marginal proportional reversed hazard rates. Further, Sankaran and Gleeja (2008) introduced a class of proportional reversed hazard rate frailty models and propose a multivariate correlated gamma frailty model. Li and Li (2008) investigated the properties of mixture model of proportional reversed hazard rate. Li and Da (2010) studied multivariate mixed proportional reversed hazard rate model having dependent mixing variables.

1.2.5 Reversed Mean Residual Life Function

Reversed hazard rate is very much related to another important notion known as the reversed mean residual life function. The reversed mean residual life function of an item failed in an interval $[a, x]$ is defined as

$$m(x) = E(x - X|X \leq x) = \frac{1}{F(x)} \int_a^x F(u)du. \tag{1.9}$$

In reliability studies the reversed mean residual life is also known as mean waiting time, expected inactivity time or mean past lifetime. Assuming $m(x)$ as differentiable the reversed mean residual life is connected to $\lambda(x)$ through the relationship

$$\lambda(x) = \frac{1 - m'(x)}{m(x)},$$

where $m'(x) = \frac{d}{dx}m(x)$. The distribution function can be uniquely determined from the relation (Chandra and Roy (2001))

$$F(x) = \exp\left\{-\int_x^b \frac{1 - m'(u)}{m(u)}du\right\}.$$

The corresponding density function is given by

$$f(x) = \exp\left[-\int_x^b \frac{1 - m'(t)}{m(t)}dt\right]\left\{\frac{1 - m'(x)}{m(x)}\right\}.$$

Let X be a continuous random variable with finite mean μ. Then $F(x)$ is uniquely determined by $\omega(x) = \lambda(x)m(x)$, $x \in (a, b)$ through the relation

$$F(x) = \exp\left[-\int_x^b \frac{\omega(t)}{t - \mu + I(t)}dt\right],$$

where $I(t) = \int_t^b \omega(x)dx$.

Chandra and Roy (2001) studied a range of properties of reversed mean residual life with respect to reversed hazard rate. Finkelstein (2002) focused the importance of reversed mean residual life in defining reversed hazard rate and studied its properties. Li and Lu (2003) established some stochastic comparisons on reversed mean residual life and residual life of series and parallel systems and presented some applications based on these comparisons. Reliability properties of reversed mean residual life and the definition of a new ordering based on reversed mean residual life are discussed in Nanda et al. (2003). Asadi (2006) studied properties of reversed mean residual life for components of parallel system. Kayid (2006) introduced and studied multivariate notions of reversed mean residual life. The properties of mean time to failure in an age replacement model is presented by examining the relationship it has with reversed hazard rate and reversed mean residual life were studied by Asha and Nair (2010). For further properties of reversed mean residual life, one could refer to Kayid and Ahmad (2004), Kayid (2006), Nanda et al. (2006), Sadegh (2008), Tavangar and Asadi (2008) and Kundu and Nanda (2010).

1.2.6 Past Entropy

Study of uncertainty is a subject of interest common to reliability, survival analysis, actuary, economics, business and many other fields. A classical measure of uncertainty for a random variable X having probability density function$f(x)$, cumulative distribution function $F(x)$ and the survival function $R(x) = 1 - F(x)$, is the differential entropy, also known as the Shannon information measure (Shannon (1948)), defined as

$$H(f) = -E(\log f(X)) = -\int_{-\infty}^{\infty} f(x)\log f(x)dx \tag{1.10}$$

where $\log$ denotes the natural logarithm.

The role of differential entropy as a measure of uncertainty in residual lifetime distributions has attracted increasing attention in recent years. According to Ebrahimi (1996), the residual entropy at time t of a random lifetime X is defined as the differential entropy of $(X|X > t)$. Formally, for all $t > 0$ the residual entropy of X is given by

$$\begin{aligned} H(f;t) &= -\int_t^{\infty} \frac{f(x)}{R(t)} \log \frac{f(x)}{R(t)} dx \\ &= \log R(t) - \frac{1}{R(t)} \int_t^{\infty} f(x)\log f(x)dx \\ &= 1 - \frac{1}{R(t)} \int_t^{\infty} f(x)\log r(x)dx. \end{aligned} \tag{1.11}$$

where $r(t) = \frac{f(t)}{F(t)}$ is the hazard function or failure rate of X. Given that an item has survived up to time t, $H(f;t)$measures the uncertainty about its remaining life. Various results concerning $H(f;t)$ have been obtained in recent years by Ebrahimi (1996, 1997, 2000), Ebrahimi and Pellery (1995), Ebrahimi and Kirmani (1996), Asadi and Ebrahimi (2000) and Navarro et al. (2002).

One of the main drawbacks of $H(f)$ specified in (1.10) is that for some probability distribution, it may be negative and then it is no longer an uncertainty measure. Khinchin (1957) generalized (1.10) by choosing a convex function ϕ such that $\phi(1) = 0$ and defined the measure

$$H^{\phi}(f) = \int f(x)\phi(f(x))\,dx. \tag{1.12}$$

For two particular choices of ϕ, (1.12) becomes, for some fixed $\beta > 0$ and $\beta \neq 1$,

$$H_{\beta}(f) = \frac{1}{\beta - 1}\left[1 - \int_0^{\infty} f^{\beta}(x)dx\right] \tag{1.13}$$

and for some $\alpha > 0$ and $\alpha \neq 1$,

$$H_\alpha(f) = \frac{1}{1-\alpha} \log \int_0^\infty f^\alpha(x)dx. \tag{1.14}$$

As $\beta \to 1$ and $\alpha \to 1$ in (1.13) and (1.14) respectively, they reduce to $H(f)$given in (1.10). It may be noted that although (1.10) may be negative for some distribution, but $H_\beta(f)$ and $H_\alpha(f)$ can always be made nonnegative by choosing appropriate value of β and α. As argued by Ebrahimi (1996), equations (1.13) and (1.14) can be redefined for a unit surviving up to an age t as (Nanda and Paul (2006c))

$$H_\beta(f;t) = \frac{1}{\beta-1} \left[1 - \int_t^\infty \left(\frac{f(x)}{R(t)} \right)^\beta dx \right] \tag{1.15}$$

and

$$H_\alpha(f;t) = \frac{1}{1-\alpha} \log \int_t^\infty \left(\frac{f(x)}{R(t)} \right)^\alpha dx \tag{1.16}$$

respectively.

It can be noted that $\beta \to 1$ and $\alpha \to 1$ in (1.15) and (1.16) respectively, they reduce to (1.11). $H_\beta(f;t)$ and $H_\alpha(f;t)$ can be called residual entropy of order β and α respectively. For a detailed survey on entropy function one may refer to Ullah (1996).

However, in many realistic situations, uncertainty is not necessarily related to the future but can also refer to the past (Maiti and Nanda (2009)). For instance if at time t, a system which is observed only at certain pre-assigned inspection times, is found to be down, then the uncertainty of the system's life relies on the past, that is, at which instant in $(0, t)$ the system has failed. To be more specific, in a periodic replacement policy where the system is observed at times $T, 2T, 3T, ...$ for some pre-assigned time T, it is possible that at time $(n-1)T$ the system is functioning, but at time nT the system is found to be down, where n is a positive integer. Then, if X is the failure time of the system, the variable of interest is $[nT - X|X \leq nT]$.

By writing $nT = t$, we have the random variable $X_t = (t - X|X \leq t)$, known as the inactivity time. This is because, once at time X the system fails, and at time t it is observed to be in a failure state, the random time for which the system was down is X_t. Based on this idea, Di Crescenzo and Longobardi (2002, 2004) have studied measures of entropy and discrimination based on past entropy over $(0, t)$.

Let us consider an absolutely continuous random variable X probability density function $f(x)$ and distribution function $F(x)$. Let the support of the random variable X be (a, b) where $a = \inf\{x|F(x) > 0\}$ and $b = \sup\{x|F(x) < 1\}$ with $-\infty \leq a < b < \infty$. Kundu et al. (2010) defined the measure of uncertainty for inactivity time or past time distribution, called past entropy as

$$\begin{aligned} \overline{H}(f;t) &= -\int_a^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} dx \\ &= 1 - \frac{1}{F(t)} \int_a^t f(x) \log \lambda(x) dx, \end{aligned} \tag{1.17}$$

where $\lambda(x) = \frac{f(t)}{F(t)}$ is the reversed hazard rate. Note that, as $t \to b$, $\overline{H}(f;t)$ becomes the well known Shannon entropy given by

$$H(f) = -\int_a^b f(x) \log f(x) dx.$$

For more properties and applications of $\overline{H}(f;t)$ one may refer to Di Crescenzo and Longobardi (2002, 2004), Nanda and Paul (2006a), Maiti and Nanda (2009) and the references therein.

1.2.7 Cumulative Entropy

Recently, Rao et al. (2004) introduced an alternative measure of uncertainty called cumulative residual entropy (CRE). This measure is based on the cumulative distribution function F and is defined in the univariate case and for non-negative random variables as follows.

$$\xi(X) = -\int_0^{\infty} R(x) \log R(x) dx.$$

They have obtained several properties of this measure and provided some applications of it in reliability engineering and computer vision. Rao (2005) developed some more mathematical properties of cumulative residual entropy and gave an alternative formula for it.

Asadi and Zohrevand (2007) studied the relation between cumulative residual entropy and mean residual life function. They proved that if X be a non-negative continuous random variable with mean residual life function $r(x)$ and cumulative residual entropy $\xi(X)$such that $\xi(X) < \infty$. Then

$$\xi(X) = E\left(r(X)\right).$$

Capturing effects of the age t of an individual or an item under study on the information about the residual lifetime is important in many applications. For example, in reliability when a component or a system of components is working at time t, one is interested in the study of the lifetime of component or system beyond t. In such case, the set of interest is the residual lifetime $S_t = \{x : x > t\}$.

Hence the distribution of interest for computing uncertainty and information is the residual distribution with survival function

$$R_t(x) = \begin{cases} \frac{R(x)}{R(t)}, x \in S_t \\ 1, otherwise \end{cases}$$

where R denotes the survival function of X.

Based on these concepts, Asadi and Zohrevand (2007) introduced the cumulative residual entropy for the residual lifetime distribution with survival function $R_t(x)$ is

$$\xi(X;t) = -\int_t^{\infty} R_t(x) \log R_t(x) dx$$

$$
\begin{aligned}
&= -\int_0^\infty \frac{R(x)}{R(t)} \log \frac{R(x)}{R(t)} dx \\
&= -\frac{1}{R(t)} \int_t^\infty R(x) \log R(x) dx + r(t) \log R(t).
\end{aligned}
$$

They called this measure as dynamic cumulative residual entropy (DCRE). It worth noting that $\xi(X;t)$ provides a dynamic information measure for measuring the information of the residual life distribution. It is clear that $\xi(X;0) = \xi(X)$. They also proved that that if F be an absolutely continuous distribution function with mean residual life function $\varepsilon(t)$ and dynamic cumulative residual entropy $\xi(X;t)$ such that $\xi(X;t) < \infty$ for all $t \geq 0$. Then

$$
\xi(X;t) = E\left(r(X) | X \geq t\right).
$$

Recently, Di Crescenzo and Longobardi (2009) introduced a new measure of information, that will be called cumulative entropy is suitable to measure information when uncertainty is related to the past, a dual concept of the cumulative residual entropy which relates to uncertainty of the future lifetime of a system. Moreover, similarly to the cumulative residual entropy, it is defined as

$$
C\xi(X) = -\int_0^\infty F(x) \log F(x) dx.
$$

This measure recalls the differential entropy (1.10), the significant difference being that now the argument of the logarithm is a probability. This implies $C\xi(X) \geq 0$, which does not hold for (1.10). If a system that begins to work at time 0 is observed only at deterministic inspection times, and is found to be ‘down’ at time t, then the uncertainty relies on which instant in $(0,t)$ it has failed. Di Crescenzo and Longobardi (2009) thus introduced the following new dynamic information measure.

$$
C\xi(X;t) = -\int_0^t \frac{F(x)}{F(t)} \log \frac{F(x)}{F(t)} dx, t \geq 0.
$$

The above measure is called as dynamic cumulative entropy which identifies the cumulative entropy $[X|X \leq t]$.

1.3 Bivariate Notions

In many practical situations, one may have paired lifetime data. For example, the time to deterioration level or time to reaction of a treatment may be of interest in pairs of lungs, kidneys, eyes or ears of humans.

Let $X = (X_1, X_2)$ be a random vector in the two-dimensional space admitting an absolute continuous distribution function

$$F(x_1, x_2) = P\left(X_1 \leq x_1, X_2 \leq x_2\right)$$

in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ where $a_i = \inf\{x_i|F_i(x_i) > 0\}$ and $b_i = \sup\{x_i|F_i(x_i) < 1\}$ with $F(x_1, b_2) = F_1(x_1)$ and $F(b_1, x_2) = F_2(x_2)$ as the marginals of X_i, $i = 1, 2$. The joint probability density function of X may be represented as

$$f(x_1, x_2) = \frac{\partial^2 F(x_1, x_2)}{\partial x_1 \partial x_2}.$$

1.3.1 Bivariate Hazard Rate

In the bivariate case we can define the hazard rate in more than one way. The first definition of bivariate hazard rate was given by Basu (1971) which is defined as

$$b(x_1, x_2) = \frac{f(x_1, x_2)}{R(x_1, x_2)},$$

where $R(x_1, x_2) = P\left(X_1 > x_1, X_2 > x_2\right)$ is the survival function of (X_1, X_2). As in the univariate case, $b(x_1, x_2)$, in general, does not determine the bivariate distribution uniquely.

A following approach in defining bivariate hazard rate is given by Johnson and

Kotz (1975). They defined bivariate hazard rate as a vector given by

$$h(x_1, x_2) = (h_1(x_1, x_2), h_2(x_1, x_2)),$$

where $h_i(x_1, x_2) = -\frac{\partial \log R(x_1, x_2)}{\partial x_i}$, $i = 1, 2$.

Note that $h_1(x_1, x_2)$ is the instantaneous rate of failure of X_1 at time x_1 given that X_1 was alive at $X_1 = x_1-$ and that X_2 survived beyond time $X_2 = x_2$. The meaning of $h_2(x_1, x_2)$ is similar.

The vector $h_1(x_1, x_2)$ uniquely determines the distribution of X through

$$R(x_1, x_2) = \exp\left\{-\int_{a_1}^{x_1} h_1(u, a_2)du - \int_{a_2}^{x_2} h_2(x_1, v)dv\right\}$$

or

$$R(x_1, x_2) = \exp\left\{-\int_{a_1}^{x_1} h_1(u, x_2)du - \int_{a_2}^{x_2} h_2(b_1, v)dv\right\}.$$

1.3.2 Bivariate Reversed Hazard Rate

Unlike the univariate set up discussed above, there is more than one definition for reversed hazard rate in the multivariate set up.

Gurler (1996) defined the bivariate reversed hazard rate as a three component vector given by

$$\Lambda(x_1, x_2) = (\Lambda_{12}(x_1, x_2), \Lambda_1(x_1, x_2), \Lambda_2(x_1, x_2))$$

where $\Lambda_{12}(x_1, x_2) = \frac{F(dx_1, dx_2)}{F(x_1, x_2)}$, $\Lambda_1(x_1, x_2) = \frac{F(dx_1, x_2)}{F(x_1, x_2)}$ and $\Lambda_2(x_1, x_2) = \frac{F(x_1, dx_2)}{F(x_1, x_2)}$.

The vector $\Lambda(x_1, x_2)$ is used for the estimation of $F(x_1, x_2)$ when the lifetime data is right truncated.

Roy (2002a) defined reversed hazard rate as a two component vector given by

$$\lambda(x_1, x_2) = (\lambda_1(x_1, x_2), \lambda_2(x_1, x_2)) \tag{1.18}$$

where $\lambda_i(x_1,x_2) = \lim\limits_{\Delta x_i \to 0} \frac{P(x_i-\Delta x_i<X_i\leq x_i|X_1\leq x_1, X_2\leq x_2)}{\Delta x_i} = \frac{\partial \log F(x_1,x_2)}{\partial x_i}$, $i=1,2$.

For $i = 1$, $\lambda_1(x_1,x_2)\Delta x_1$ is the probability of failure of the first component in the interval $(x_1-\Delta x_1, x_1]$ given that it has failed before x_1 and the second has failed before x_2. The interpretation for $\lambda_2(x_1,x_2)$ is similar. From Roy (2002a), it follows that $\lambda_i(x_1,x_2)$, $i=1,2$ determine $F(x_1,x_2)$ uniquely by the relationship

$$F(x_1,x_2) = \exp\left\{-\int_{x_1}^{b_1}\lambda_1(u,b_2)du - \int_{x_2}^{b_2}\lambda_2(x_1,v)dv\right\}$$

or

$$F(x_1,x_2) = \exp\left\{-\int_{x_1}^{b_1}\lambda_1(u,x_2)du - \int_{x_2}^{b_2}\lambda_2(b_1,v)dv\right\} \tag{1.19}$$

where $\lambda_1(x_1,b_2)=\lambda_1(x_1)$ and $\lambda_2(b_1,x_2)=\lambda_2(x_2)$ are the marginal reversed hazard rates of X_1 and X_2.

Recently, Bismi (2005) defined bivariate scalar reversed hazard rate as

$$r(x_1,x_2) = \frac{f(x_1,x_2)}{F(x_1,x_2)}. \tag{1.20}$$

It can be easily seen that (1.20) is a natural extension of the univariate reversed hazard rate given in (1.7). The term $r(x_1,x_2)\Delta x_1\Delta x_2 + o(\Delta x_1,\Delta x_2)$ can be interpreted as the probability of failure of components 1 and 2 in intervals $(x_1-\Delta x_1,x_1]$ and $(x_2-\Delta x_2,x_2]$ respectively, given that they failed before (x_1,x_2).

It can be seen that $\Lambda_{12}(x_1,x_2) = r(x_1,x_2)dx_1dx_2$, $\Lambda_1(dx_1,x_2) = \lambda_1(x_1,x_2)dx_1$ and $\Lambda_2(x_1,dx_2) = \lambda_2(x_1,x_2)dx_2$.

Recently, another definition of reversed hazard rate has given in Sankaran and Gleeja (2006) which is defined as

$$k(x_1,x_2) = \left(k_1(x_1,x_2), k_2(x_1,x_2)\right), \tag{1.21}$$

where $k_i(x_1,x_2) = \lim\limits_{\Delta x_i\to 0}\frac{P(x_i-\Delta x_i<X_i\leq x_i|X_i\leq x_i, X_j=x_j)}{\Delta x_i} = \frac{f(x_i|X_j=x_j)}{F(x_i|X_j=x_j)}$ with $f(x_i|X_j = x_j)$ as the conditional density function of X_i given $X_j=x_j$ and $F(x_i|X_j=x_j)$ as the

conditional distribution function of X_i given $X_j = x_j$, $i, j = 1, 2$; $i \neq j$.

Further, Sankaran and Gleeja (2006) gave a unique representation for $F(x_1, x_2)$ in terms of bivariate reversed hazard rate given in (1.18) and (1.20) as

$$F(x_1, x_2) = \exp\left\{-\int_{x_1}^{b_1} r_1(u)du\right\} \exp\left\{-\int_{x_2}^{b_2} r_2(v)dv\right\} \exp\left\{\int_{x_1}^{b_1}\int_{x_2}^{b_2} \left(r(u,v) - \lambda_1(u,v)\lambda_2(u,v)\right) dvdu\right\}.$$

They also gave a local dependence measure using bivariate reversed hazard rates and studied its properties.

1.3.3 Bivariate Reversed Mean Residual Life Function

The bivariate reversed mean residual life function of a random vector $X = (X_1, X_2)$ is defined as (Nair and Asha (2008))

$$m(x_1, x_2) = (m_1(x_1, x_2), m_2(x_1, x_2)) \tag{1.22}$$

where $m_i(x_1, x_2) = E\left(x_i - X_i | X_1 \leq x_1, X_2 \leq x_2\right) = \frac{1}{F(x_1,x_2)} \int_{a_i}^{x_i} F(x_i, x_j)dx_i, i, j = 1, 2; i \neq j$.

The bivariate reversed mean residual life function uniquely determines the distribution function through the relation,

$$F(x_1, x_2) = \frac{m_i(b_1, b_2)m_j(x_i, b_j)}{m_i(x_i, b_j)m_j(x_1, x_2)} \exp\left\{-\int_{x_i}^{b_i} \frac{dt}{m_i(b_j, t)} - \int_{x_j}^{b_j} \frac{dt}{m_j(x_i, t)}\right\}, \tag{1.23}$$

for $i, j = 1, 2$; $i \neq j$.

Further, Nair and Asha (2008) extended the definition of bivariate reversed mean residual life function to the multivariate case. They also discussed the models based on proportional reversed mean residual life and their properties. For more discussions of bivariate reversed mean residual life function, we may refer Kayid (2006).

1.4 Discrete Notions

There is an abundance of literature on continuous life distributions used in modeling failure data. In reliability theory and survival analysis, time is assumed to be continuous. But discrete failure data arise in various common situations where the system lifetimes cannot be measured with calendar time. Consider the following examples:

1. A device is monitored only once per day (or month etc.). Then the random variable of interest may be the successful number of periods completed prior to the failure of the device.

2. A piece of equipment operates in cycles. In this case the random variable of interest is the successful number of cycles before the failure. For instance, the number of flashes in a car flasher prior to failure of the device.

3. In some situations the experimenter groups or discretizes the continuous observations.

Also Actuaries and Biostatisticians are interested in the lifetimes of persons or organisms, measured in months, weeks, or days (Kemp (2004)). Since there is a limit on the precision of any measurement, it can be arguably said that samples from a continuous distribution exist only in theory (Nanda and Sengupta (2005)). A discrete life distribution is a natural choice where failure occurs only due to incoming shocks. For example, in weapons reliability, the number of rounds fired until failure is more important than age at failure. Also discrete distributions have important applications in reliability theory. For example, they can be used for modeling discrete lifetimes of nonrepairable systems.

For the use of discrete models in reliability theory and characterization of probability distributions, one may refer to Xekalaki (1983), Nair and Hitha (1989), Adams and

Watson (1989), Roy and Gupta (1992), Shaked et al. (1994), Sengupta et al. (1995), Bracquemond and Gaudoin (2003), Xekalaki and Dimaki (2005) and the references therein.

Let X denote a discrete random variable taking values on $I_m = \{n, n+1, n+2, ..., m\}$ where the integer n could be $-\infty$, but m is finite and positive. Denote f and F respectively the probability mass function (p.m.f.) and the distribution function of X. That is,

$$f(x) = P(X = x)$$

and

$$F(x) = P(X \leq x) = \sum_{j=n}^{x} f(j), x \in I_m.$$

The reversed hazard rate of X is a useful tool in the analysis of left censored data, which is defined as

$$\lambda(x) = P(X = x | X \leq x) = \frac{f(x)}{F(x)}, x \in I_m. \tag{1.24}$$

Dewan and Sudheesh (2009) proposed a new definition for reversed hazard rate as

$$\delta(x) = \ln \frac{F(x)}{F(x-1)}, x \in I_m. \tag{1.25}$$

The rationale behind this later definition is as follows. In the continuous case, the reversed hazard rate is defined as

$$\lambda(x) = \frac{F'(x)}{F(x)} = \frac{d \ln F(x)}{dx}.$$

Instead of taking $[F(x) - F(x-1)]$ for $F'(x)$ which leads to the expression (1.7), we could use $[\ln F(x) - \ln F(x-1)]$ for $\frac{d}{dx} \ln F(x)$ so that (1.25) follows. Note that $\delta(x)$ is not bounded by unity and is additive for parallel system as in the continuous case.

The function $\delta(x)$ determines the distribution of X uniquely by the relation

$$F(x) = \exp\left(-\sum_{y=x+1}^{m} \delta(y)\right). \tag{1.26}$$

Hence, the cumulative reversed hazard rate is given by

$$\Lambda(x) = \sum_{y=x+1}^{m} \delta(y) = -\ln F(x).$$

Similar result holds when the lifetimes are continuous random variables. Since $\lambda(x)$ and $\delta(x)$ are related through

$$\lambda(x) = 1 - e^{-\delta(x)}, \tag{1.27}$$

both $\lambda(x)$ and $\delta(x)$ have same monotonic properties. That means, $\lambda(x)$ is increasing/decreasing if and only $\delta(x)$ is increasing/decreasing in$x$.

Further,

$$f(x) = F(x)\left(1 - e^{-\delta(x)}\right), x \in I_m.$$

The hazard rate defined by

$$h(x) = \frac{f(x)}{1 - F(x-1)}$$

can be written in terms of $\delta(x)$ as

$$\frac{h(x+1)(1-h(x))}{h(x)} = \frac{e^{\delta(x+1)} - 1}{1 - e^{-\delta(x)}}.$$

For the hazard rate $r(x)$ defined in Xie et al. (2002), we see that

$$\frac{e^{\delta(x+1)} - 1}{1 - e^{-\delta(x)}} = \frac{1 - e^{-r(x+1)}}{e^{r(x)} - 1}.$$

The functional form of $\delta(x)$ enables the characterization of the distribution of X.

Another measure of interest is the mean past lifetime (MPL). Consider the conditional random variable $X_x = x - X | X \leq x$, where $x \in I_m$. Then $E(X_x)$, which we denote by $k(x)$, is the mean past lifetime of X. To show why the mean past lifetime function may be important, we give an example here. Consider a new car which has been used for some time and undergoes for the first complete check up. Assume that the technician has found that a unit of the engine system of the car, with lifetime X, has already failed. As such systems are not monitored continuously, the technician might be interested in investigating the history of the system, e.g. when the unit has failed. In this case the random variable of interest is $X_x = x - X|X \leq x$, where x is the day of the check up. The expected value of X_x is called the mean past life, denoted as $k(x)$ and is defined as (Goliforushani and Asadi (2008))

$$\begin{aligned} k(x) &= E\left(x - X|X \leq x\right) \\ &= \frac{1}{F(x)} \sum_{t=n}^{x} F(t). \end{aligned} \tag{1.28}$$

Also, several properties of $k(x)$ for lifetime random variable are studied in Goliforushani and Asadi (2008). An inversion formula to retrieve distribution function from $k(x)$ may be obtained in Ruiz and Navarro (1995). They have given a more general inversion formula for doubly truncated distributions.

Recently, Kundu et al. (2010) discussed about past entropy for a discrete random variable X. The discrete past entropy is defined as

$$\overline{H}(X;j) = -\sum_{k=n}^{j} \frac{f(k)}{F(j)} \ln \frac{f(k)}{F(j)}. \tag{1.29}$$

Note that as $j \to m$, $\overline{H}(X;j)$ becomes the well known Shannon entropy given by

$$\overline{H}(X) = -\sum_{k} f(k) \ln f(k).$$

The discrete generalized past entropy is defined as

$$\overline{H}^{\beta}(X;j) = \frac{1}{1-\beta} \ln \left[\sum_{k=n}^{j} \left(\frac{f(k)}{F(j)} \right)^{\beta} \right]. \tag{1.30}$$

It is to be noted that as $\beta \to 1$, (1.30) reduces to (1.29).

1.5 Present Study

The present study is largely based on the concepts defined above and is reported in the coming five chapters.

In the second chapter, we introduce the reversed lack of memory property and characterize the distributions by this property and its variants. The implications along with characterizations of the reversed lack of memory property is taken up for study. This chapter also considers the bivariate extension of this property and describes distributions that satisfy the bivariate reversed lack of memory property.

In the third chapter, we generalize this property which involves operations different than the "addition". In particular an associative, binary operator "$*$" is considered. The univariate reversed lack of memory property is generalized using the binary operator and a class of probability distributions which include Type 3 extreme value, power function, reflected Weibull and negative Pareto distributions are characterized (Asha and Rejeesh (2009)). We also define the almost reversed lack of memory property and considered the distributions with reversed periodic hazard rate under the binary operation. Further, we give a bivariate extension of the generalized reversed lack of memory property and characterize a class of bivariate distributions which include the characterized extension (CE) model of Roy (2002a) apart from the bivariate reflected Weibull and power function distributions. We proved the equality of local proportionality of the reversed hazard rate and generalized reversed lack of memory property. A few characterizations of the model are also discussed in this chapter.

In the fourth chapter, we generalize the bivariate reversed lack of memory property discussed in the previous chapter using two different binary associative operators and we derive bivariate distributions with non-identical marginals that are characterized by this property.

Study of uncertainty is a subject of interest common to reliability, survival analysis, actuary, economics, business and many other fields. However, in many realistic situations, uncertainty is not necessarily related to the future but can also refer to the past. Recently, Di Crescenzo and Longobardi (2009) introduced a new measure of information called dynamic cumulative entropy. Dynamic cumulative entropy is suitable to measure information when uncertainty is related to the past, a dual concept of the cumulative residual entropy which relates to uncertainty of the future lifetime of a system. We redefine this measure in the whole real line and study its properties. We also discuss the implications of generalized reversed lack of memory property on dynamic cumulative entropy and past entropy. These results are reported in the fifth chapter.

In the last chapter, we extend the idea of reversed lack of memory property to the discrete set up. Here we investigate the discrete class of distributions characterized by the discrete reversed lack of memory property. The concept is extended to the bivariate case and bivariate distributions characterized by this property are also presented. The implication of this property on discrete reversed hazard rate (Dewan and Sudheesh (2009)), mean past life (Goliforushani and Asadi (2008)) and discrete past entropy (Kundu et al. (2010)) are also investigated. The chapter along with the thesis concludes with a small discussion on future work.

Chapter 2

The Reversed Lack of Memory Property

2.1 Introduction

Here we introduce the reversed lack of memory property and derive the distribution characterized by it. The study of characterization of probability distribution appears to have begun with the work of Gauss in 1807 when he proved under certain conditions that the maximum likelihood estimate of the location parameter of distribution is the sample mean if and only if, the distribution is normal. Even though reckoned from this work a long history can be accredited to the research activities in characterizing probability distributions, a full fledged expansion of this field as part of mathematical statistics, began taking shape only in the late fifties of the last century. The first authoritative book on the tools employed in proving characterizations along with a large compilation of results covering most probability distributions was published by Kagan, Linnik and Rao in 1973. This along with the books by Patil et al. (1975), Galambos and Kotz (1978), Mathai and Pederzoli (1977) and Azlarov and Volodin (1986) include most of the literature on characterization results.

In this Chapter, we study the reversed lack of memory property in detail and characterize the distributions by this property and its variants. If X satisfies the lack of memory property, it follows that $-X$ satisfies the reversed lack of memory property. However the converse is not true indicating that the reversed lack of memory property is radically different from lack of memory property and so there is a scope for separate study. In Section 2.2, we define the reversed lack of memory property and characterize the distribution by the same. Further, we extend this to bivariate set up which are included in Section 2.3. Some distributional properties of the above property are also studied.

2.2 The Reversed Lack of Memory Property

Consider a random variable X with an absolutely continuous cumulative distribution function $F(x) = P(X \leq x)$, survival function $R(x) = 1 - F(x)$ and probability density function $f(x)$. Let $a = \inf\{x|F(x) > 0\}$ and $b = \sup\{x|F(x) < 1\}$. Then (a, b), $-\infty \leq a < b < \infty$ is the interval of support of X. We define the reversed lack of memory property by the following (Asha and Rejeesh (2007)).

Definition 2.1. *Let X be a random variable having distribution function $F(x)$ and survival function $R(x)$ with support (a, b) where $a < 0$ and $b \geq 0$. Then X is said to have the reversed lack of memory property (RLMP) if*

$$P(X \leq x|X \leq x+t) = P(X \leq 0|X \leq t) \tag{2.1}$$

for all $a < x \leq x+t \leq b$.

In terms of distribution function (2.1) can be written as

$$F(x)F(t) = F(x+t)\,F(0) \tag{2.2}$$

for all $a < x \leq x+t \leq b$.

Note that for $a = 0$ we require $F(0) \neq 0$ for (2.1) or (2.2) to hold.

A physical interpretation for the reversed lack of memory property can be that given a component has maximum lifetime utmost $x + t$ then the failure of the component at any instant before $x + t$, say x, depends only on the residual time t left, rather than x.

It follows very directly that if X satisfies lack of memory property or X is exponential, then $-X$ should satisfy the reversed lack of memory property or $-X$ is negative exponential with $b = 0$. We now investigate the converse. Let X satisfy the reversed lack of memory property. Does $-X$ satisfy lack of memory property? We indicate that this is not necessarily true by the following.

Let X satisfies the reversed lack of memory property specified by (2.1). Then we have

$$P\left(-X \geq -x\right) P\left(-X \geq -t\right) = P\left(-X \geq -x - t\right) P\left(-X \geq 0\right)$$

or

$$R_{-X}(-x)R_{-X}(-t) = R_{-X}(-x - t)R_{-X}(0)$$

which does not imply the lack of memory property, since $R_{-X}(0) < 1$ by the fact that $-b \leq -X < \infty$. Thus the class of distributions satisfying the reversed lack of memory property is a larger class than the negative exponential. Hence the reversed lack of memory property is radically different from the lack of memory property.

Now, by taking logarithm on both sides of (2.2), we get

$$\ln F(x) + \ln F(t) = \ln F(x + t) + \ln F(0) \tag{2.3}$$

for all $a < x \leq x + t \leq b$.

For an absolutely continuous random variable X differentiating (2.3) with respect to x, yields

$$\frac{f(x)}{F(x)} = \frac{f(x + t)}{F(x + t)}$$

which implies

$$\lambda(x) = \lambda(x+t) \tag{2.4}$$

for all $a < x \leq x + t \leq b$.

Thus $\lambda(x) = c$, a constant when the reversed lack of memory property is satisfied. Since $\lambda(x)$ uniquely determines the underlying distribution, it follows from Block et al. (1998) that $a = -\infty$ and $b < \infty$ and the cumulative distribution function of X is given by

$$F(x) = \begin{cases} \exp\left[c\left(x-b\right)\right], x < b \\ 1, x \geq b, c > 0 \end{cases} \tag{2.5}$$

Hence we have the following characterization for an absolutely continuous random variable X.

Theorem 2.1. *An absolutely continuous random variable X in the support of $(-\infty, b)$, $b \geq 0$ has the reversed lack of memory property if and only if any of the following equivalent conditions hold.*

1. $F(x)F(t) = F(x+t)F(0)$ *for all* $-\infty < x \leq x + t \leq b < \infty$.

2. $\lambda(x) = c$ *where* $c > 0$ *is a constant for all* $-\infty < x \leq b$.

3. X *is distributed as a subclass of Type 3 extreme value distribution defined on* $(-\infty, b)$*,* $b \geq 0$ *specified by*

$$F(x) = \begin{cases} \exp\left[c\left(x-b\right)\right], x < b \\ 1, x \geq b, c > 0 \end{cases}$$

This Type 3 extreme value distribution also belongs to the reversed generalized Pareto distribution developed by Castillo and Hadi (1995) as a fatigue model that satisfy certain compatibility conditions arising out of physical and statistical conditions in fatigue studies.

In the nonnegative support $(0, b)$, $b < \infty$,

$$F(x) = \begin{cases} \exp\left[-cb\right], x = 0 \\ \exp\left[c\left(x - b\right)\right], x < b \\ 1, x \geq b, c > 0 \end{cases}$$

Thus X ceases to be an absolutely continuous random variable. Nevertheless it is characterized by the reversed lack of memory property.

In the next section we attempt a bivariate extension of this property and call it the bivariate reversed lack of memory property (BRLMP) and investigate distributions characterized by the same. It should be noted that multivariate extension is merely an extension of the bivariate case and derivations are quite straight forward.

2.3 The Bivariate Reversed Lack of Memory Property

Consider a random vector $X = (X_1, X_2)$ in the two dimensional space with joint distribution function $F(x_1, x_2)$ in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ where

$$a_i = \inf\{x_i | F_i(x_i) > 0\} \text{ and } b_i = \sup\{x_i | F_i(x_i) < 1\},$$

with $F(x_1, b_2) = F_1(x_1)$ and $F(b_1, x_2) = F_2(x_2)$ as the marginals of X_i, $i = 1, 2$.

Now we define the bivariate reversed lack of memory property as follows

Definition 2.2. *A random vector $X = (X_1, X_2)$ in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ with $a_i < 0$ and $b_i \geq 0$, $i = 1, 2$ is said to have the bivariate reversed lack of memory property (BRLMP) if*

$$P\left[X_1 \leq x_1, X_2 \leq x_2 | X_1 \leq x_1 + t_1, X_2 \leq x_2 + t_2\right] = P\left[X_1 \leq 0, X_2 \leq 0 | X_1 \leq t_1, X_2 \leq t_2\right] \quad (2.6)$$

for all x_i and t_i such that $a_i < x_i \leq x_i + t_i \leq b_i < \infty$, $i = 1, 2$.

In terms of distribution function (2.6) can be written as

$$F(x_1, x_2)F(t_1, t_2) = F(x_1 + t_1, x_2 + t_2)F(0, 0) \tag{2.7}$$

for all x_i and t_i such that $a_i < x_i \leq x_i + t_i \leq b_i < \infty$, $i = 1, 2$.

As in the univariate case, here also we assume $F(0, 0) \neq 0$ when $a_i = 0$, $i = 1, 2$ for (2.7) to hold. From (2.6), by putting $t_1 = 0$ and $x_1 = b_1$ it follows that

$$P\left[X_2 \leq x_2 | X_2 \leq x_2 + t_2\right] = P\left[X_1 \leq 0, X_2 \leq 0 | X_1 \leq 0, X_2 \leq t_2\right] \tag{2.8}$$

for all $a_2 < x_2 \leq x_2 + t_2 \leq b_2$.

For $x_2 = 0$ it further follows that

$$P\left[X_2 \leq 0 | X_2 \leq t_2\right] = P\left[X_1 \leq 0, X_2 \leq 0 | X_1 \leq 0, X_2 \leq t_2\right].$$

Hence (2.8) reduces to

$$P\left[X_2 \leq x_2 | X_2 \leq x_2 + t_2\right] = P\left[X_2 \leq 0 | X_2 \leq t_2\right]$$

for all $a_2 < x_2 \leq x_2 + t_2 \leq b_2$.

In a similar manner it can be shown that

$$P\left[X_1 \leq x_1 | X_1 \leq x_1 + t_1\right] = P\left[X_1 \leq 0 | X_1 \leq t_1\right]$$

for all $a_1 < x_1 \leq x_1 + t_1 \leq b_1$.

This implies that if (X_1, X_2) has bivariate reversed lack of memory property, then the marginals satisfy reversed lack of memory property.

In the coming results we investigate the bivariate distributions characterized by the bivariate reversed lack of memory property.

Theorem 2.2. *Let (X_1, X_2) be a random vector in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ with $a_i < 0$ and $b_i \geq 0$, $i = 1, 2$; then (X_1, X_2) satisfies the bivariate reversed lack of memory property if and only if X_1 and X_2 are independently distributed and the marginals having the reversed lack of memory property.*

Proof. If X_1 and X_2 are independently distributed, making use of the fact that X_1 and X_2 satisfy reversed lack of memory property, the bivariate reversed lack of memory property can be verified directly.

To prove the converse, putting $t_2 = 0$ and $x_2 = b_2$ in (2.7) we have

$$F(x_1, b_2)F(t_1, 0) = F(x_1 + t_1, b_2)F(0, 0)$$

for all x_1 and t_1 such that $a_1 < x_1 \leq x_1 + t_1 \leq b_1$.

Further, for $x_1 = x_2 = 0$ gives that

$$\frac{F(t_1, 0)}{F(0, 0)} = \frac{F_1(t_1)}{F_1(0)}, \tag{2.9}$$

where $F_i(t_i)$ denotes the marginal of X_i, $i = 1, 2$.

Similarly, we can obtain

$$\frac{F(0, t_2)}{F(0, 0)} = \frac{F_2(t_2)}{F_2(0)}. \tag{2.10}$$

Putting $t_i = b_i$, $i = 1, 2$ in either (2.9) or (2.10) we get

$$F(0, 0) = F_1(0)F_2(0).$$

Hence it now follows from (2.9) and (2.10) that

$$F(t_1, 0) = F_1(t_1)F_2(0)$$

and

$$F(0, t_2) = F_1(0)F_2(t_2) \tag{2.11}$$

For $x_2 = 0$ in (2.7) we get

$$F(x_1, 0)F(t_1, t_2) = F(x_1 + t_1, t_2)F(0, 0).$$

Once again by substituting $t_1 = 0$ we get

$$F(x_1, 0)F(0, t_2) = F(x_1, t_2)F(0, 0).$$

Using (2.11), the above equation becomes

$$F(x_1, 0)F_1(0)F_2(t_2) = F(x_1, t_2)F(0, 0).$$

Or,

$$F(x_1, t_2) = \frac{F(x_1, 0)}{F(0, 0)}F_1(0)F_2(t_2).$$

By applying (2.9) we have

$$F(x_1, t_2) = F_1(x_1)F_2(t_2). \tag{2.12}$$

Similarly,

$$F(t_1, x_2) = F_1(t_1)F_2(x_2). \tag{2.13}$$

To see that independence holds for any $a_i < x_i \leq b_i$, put $x_i = b_i - t_i$, $i = 1, 2$ in (2.7) to get

$$F(b_1 - t_1, b_2 - t_2)F(t_1, t_2) = F(0, 0), a_i < b_i - t_i \leq b_i, i = 1, 2.$$

Using (2.12) and (2.13),

$$F(b_1 - t_1, b_2 - t_2) = \frac{F_1(0)}{F_1(t_1)}\frac{F_2(0)}{F_2(t_2)}, a_i < b_i - t_i \leq b_i, i = 1, 2. \tag{2.14}$$

Since bivariate reversed lack of memory property implies reversed lack of memory property it follows that X_1 and X_2 has the reversed lack of memory property hence

$$F_1(b_1 - t_1)F_1(t_1) = F_1(0)\, F_1(b_1)$$

and

$$F_2(b_2 - t_2)F_2(t_2) = F_2(0)F_2(b_2).$$

Substituting in (2.14) we get

$$F(b_1 - t_1, b_2 - t_2) = F_1(b_1 - t_1)F_2(b_2 - t_2), a_i < b_i - t_i \leq b_i, i = 1, 2.$$

It now follows that

$$F(x_1, x_2) = F_1(x_1)F_2(x_2), a_i < x_i \leq b_i, i = 1, 2.$$

Hence the theorem. □

For a random vector(X_1, X_2) with an absolutely continuous distribution function satisfying (2.7) we can have

$$\ln F(x_1, x_2) + \ln F(t_1, t_2) = \ln F(x_1 + t_1, x_2 + t_2) + \ln F(0, 0) \tag{2.15}$$

Differentiating (2.15) with respect to x_i, $i = 1, 2$, we have

$$\lambda(x_1, x_2) = \lambda(x_1 + t_1, x_2 + t_2) \tag{2.16}$$

where $\lambda(x_1, x_2) = (\lambda_1(x_1, x_2), \lambda_2(x_1, x_2))$ is the bivariate reversed hazard gradient (Roy (2002a)) with

$$\lambda_i(x_1, x_2) = \frac{\partial \log F(x_1, x_2)}{\partial x_i}, i = 1, 2.$$

If X_i's are independent it then follows that $\lambda_i(x_1, x_2)$ is $\lambda_i(x_i)$, the univariate reversed hazard rate of X_i, $i = 1, 2$ defined in Section 1.2.2. Thus, under bivariate reversed lack of memory property

$$\lambda(x_1, x_2) = \lambda(x_1 + t_1, x_2 + t_2) = (\lambda_1(x_1), \lambda_2(x_2)) = (c_1, c_2)$$

for some $c_i > 0$, $i = 1, 2$. It now follows from Block et al. (1998) that X_i has a support $a_i = -\infty$, $b_i < \infty$, $i = 1, 2$ and is distributed as (2.5).

From the above discussions we have the following theorem.

Theorem 2.3. *An absolutely continuous random vector (X_1, X_2) with distribution function $F(x_1, x_2)$ with support in $(a_1, b_1) \times (a_2, b_2) \in R^2$, $b_i \geq 0$, $i = 1, 2$ satisfies the bivariate reversed lack of memory property*

$$F(x_1, x_2)F(t_1, t_2) = F(x_1 + t_1, x_2 + t_2)F(0, 0) \tag{2.17}$$

for all x_i and t_i such that $a_i < x_i \leq x_i + t_i \leq b_i < \infty$, $i = 1, 2$, if and only if the following equivalent conditions hold.

1. *X_i's are independently distributed as*

$$F(x_i) = \begin{cases} exp\left[c_i\left(x_i - b_i\right)\right], x_i < b_i \\ 1, x_i \geq b_i, c_i > 0, i = 1, 2. \end{cases} \tag{2.18}$$

2. *The bivariate reversed hazard gradient*

$$\lambda(x_1, x_2) = (\lambda_1(x_1, x_2), \lambda_2(x_1, x_2)) = (c_1, c_2)$$

where $\lambda_i(x_1, x_2) = \frac{\partial \log F(x_1, x_2)}{\partial x_i}$, $i = 1, 2$ for any constants $c_1, c_2 > 0$.

Proof. From Theorem 2.2 it follows that (2.17) implies X_1 and X_2 are independently distributed and the marginals satisfy the reversed lack of memory property. Hence from Theorem 2.1 it follows that X_i's are distributed as in (2.18). Since X_i's are independent, the bivariate reversed hazard gradient $\lambda(x_1, x_2) = (\lambda_1(x_1), \lambda_2(x_2))$ where $\lambda_i(x_i)$, $i = 1, 2$ is the univariate reversed hazard rate. That (2.18) is characterized by constancy of reversed hazard rate follows from Block et al. (1998). Hence the theorem. □

Thus the class of distributions characterized by the bivariate reversed lack of memory property is the class of distributions possessing constant bivariate reversed hazard gradient.

The bivariate reversed lack of memory property is too strong generalization of the reversed lack of memory property as the distribution function $F(x_1, x_2)$ that satisfies (2.17) is the subclass of trivial bivariate extreme value Type 3 distribution which is the product of its marginals. A meaningful relaxation of (2.17) is

$$P\left[X_1 \leq x_1, X_2 \leq x_2 | X_1 \leq x_1 + t, X_2 \leq x_2 + t\right] = P\left[X_1 \leq 0, X_2 \leq 0 | X_1 \leq t, X_2 \leq t\right] \quad (2.19)$$

for all x_i and t such that $a_i < x_i \leq x_i + t \leq b_i < \infty$, $i = 1, 2$.

In terms of the distribution function this can be written as

$$F(x_1, x_2)F(t, t) = F(x_1 + t, x_2 + t)F(0, 0) \quad (2.20)$$

for all x_i and t such that $a_i < x_i \leq x_i + t \leq b_i < \infty$, $i = 1, 2$.

We shall term this property as BRLMP(1). This is the dual of the bivariate lack of memory property (BLMP) (Galambos and Kotz (1978)).

In particular, for $x_1 \geq x_2$, put $t = b_1 - x_1$ in (2.20) to get

$$F(x_1, x_2) = \frac{F(0, 0)F_2\left(b_1 - (x_1 - x_2)\right)}{F(b_1 - x_1, b_1 - x_1)}. \quad (2.21)$$

Similarly for $x_1 \leq x_2$ putting $t = b_2 - x_2$, we get

$$F(x_1, x_2) = \frac{F(0,0)F_1\left(b_2 - (x_2 - x_1)\right)}{F(b_2 - x_2, b_2 - x_2)}. \tag{2.22}$$

Further for $x_1 = x_2$, the equation (2.20) reduces to

$$F(x,x)F(t,t) = F(x+t, x+t)F(0,0), a < x \leq x+t \leq b$$

which reduces to the Cauchy functional equation

$$H(x)H(t) = H(x+t), a < x \leq x+t \leq b \tag{2.23}$$

where

$$H(x) = \frac{F(x,x)}{F(0,0)}.$$

Solving for (2.23) (Aczel (1966)), we have

$$H(x) = \exp\left(cx\right), a < x \leq b$$

for some constant c.

Since $F(x,x)$ is a distribution function, $F(b,b) = 1$ and hence it follows that

$$F(x,x) = \exp\left[c\left(x - b\right)\right], a < x \leq b, c > 0. \tag{2.24}$$

Thus, substituting (2.24) in (2.21) and (2.22) the general form of distribution satisfying BRLMP(1) is

$$F(x_1, x_2) = \begin{cases} e^{c(x_2 - b_2)} F_1\left(b_2 - (x_2 - x_1)\right), x_1 \leq x_2 \\ e^{c(x_1 - b_1)} F_2\left(b_1 - (x_1 - x_2)\right), x_1 \geq x_2, \end{cases} \tag{2.25}$$

for $a_i < x_i \leq b_i$, $c > 0$, $i = 1, 2$.

Observe that (2.25) is the distribution of $Z = \max(X_1, X_2)$ and Z has an absolutely continuous distribution function only when $a_i = -\infty$, $i = 1, 2$.

The next theorem discusses a bivariate distribution with absolutely continuous marginals characterized by the BRLMP(1).

Theorem 2.4. *Let (X_1, X_2) be a random vector with marginals specified by*

$$F_i(x_i) = \exp\left[c_i\left(x_i - b_i\right)\right], -\infty < x_i \leq b_i, b_i > 0, c_i > 0, i = 1, 2, \tag{2.26}$$

then (X_1, X_2) satisfy the BRLMP(1) if and only if the distribution of (X_1, X_2) is given by

$$F(x_1, x_2) = \exp\left[c_1\left(x_1 - b_1\right) + c_2\left(x_2 - b_2\right) + c_{12}\max\left(x_1 - b_1, x_2 - b_2\right)\right], \tag{2.27}$$

$-\infty < x_i \leq b_i$, $c_i > 0$, $i = 1, 2$, $c_{12} \geq 0$.

Proof. Let the BRLMP(1) be satisfied, then it follows that

$$F(x_1, x_2) = \begin{cases} e^{c(x_2 - b_2)} F_1\left(b_2 - (x_2 - x_1)\right), x_1 \leq x_2 \\ e^{c(x_1 - b_1)} F_2\left(b_1 - (x_1 - x_2)\right), x_1 \geq x_2 \end{cases}$$

Given that the marginals are specified as in (2.26) we get

$$F(x_1, x_2) = \begin{cases} \exp\left[c_1\left(x_1 - b_1\right) + c_2\left(x_2 - b_2\right) + \left(c - c_1 - c_2\right)\left(x_2 - b_2\right)\right], x_1 \leq x_2 \\ \exp\left[c_1\left(x_1 - b_1\right) + c_2\left(x_2 - b_2\right) + \left(c - c_1 - c_2\right)\left(x_1 - b_1\right)\right], x_1 \geq x_2 \end{cases}$$

Writing$(c - c_1 - c_2) = c_{12}$, we retrieve the form (2.27). The converse is quite straight forward. □

Remark 2.1. *Even though marginals are absolutely continuous, the bivariate distribution* (2.27) *is not the same. Also when (X_1, X_2) is having a nonnegative support* *with marginal specified as* (2.26)*, then*

$$F(x_1, x_2) = F_d + F_c + F_{ac},$$

where

$$F_d = \begin{cases} e^{-c_i b_i}, x_i = 0, x_{3-i} = b_i, i = 1, 2 \\ e^{-c_1 b_1 - c_2 b_2 + c_{12} \max(-b_1, -b_2)}, x_i = 0, i = 1, 2 \end{cases},$$

$$F_c = e^{c_1(x-b_1)+c_2(x-b_2)+c_{12}\max(x-b_1,x-b_2)}, x_1 = x_2 = x,$$

and

$$F_{ac} = \exp\left[c_1 (x_1 - b_1) + c_2 (x_2 - b_2) + c_{12} \max (x_1 - b_1, x_2 - b_2)\right],$$

for all $0 < x_i \leq b_i$*,* $c_i > 0$*,* $i = 1, 2$*,* $c_{12} \geq 0$*.*

Thus we observe that the boundary class of distributions of the increasing bivariate reversed hazard rate class and decreasing bivariate reversed hazard rate class does not consist of absolutely continuous distributions.

Theorem 2.5. *A random vector* (X_1, X_2) *with distribution function* $F(x_1, x_2)$*with support in* $(a_1, b_1) \times (a_2, b_2) \in R^2$*,* $b_i \geq 0$*,* $i = 1, 2$ *satisfies the BRLMP(1) if and only if the following conditions are satisfied.*

1. $Z = \max(X_1, X_2)$ *is distributed as*

$$P[Z \leq z] = \exp[c(z - b)], -\infty < z \leq b, b \geq 0. \tag{2.28}$$

2. Z *and* $T = |X_2 - X_1|$ *are mutually independent.*

Proof. From equation (2.25) it follows that if BRLMP(1) is satisfied Z and T are independent. Conversely, let Z and T be independent with Z distributed as (2.28). We then have for $-\infty < x_1 \le x_2 \le b_2$,

$$\begin{aligned} F(x_1, x_2) &= F(x_1, x_1) + P\left[X_1 \le x_1, x_1 \le X_2 \le x_2\right] \\ &= F(x_1, x_1) + \int_{x_1}^{x_2} P\left[T > z - x_1\right] \frac{dP\left[Z \le z\right]}{dz} \\ &= F(x_1, x_1) + \int_{x_1}^{x_2} P\left[T > z - x_1\right] \frac{de^{c(z-b)}}{dz}. \end{aligned}$$

We further have that for a fixed x_1,

$$\frac{\partial}{\partial x_2} F(x_1, x_2) = P\left[T > x_2 - x_1\right] ce^{c(x_2 - b_2)} \tag{2.29}$$

Also,

$$\begin{aligned} \frac{\partial}{\partial x_2} F(x_1 + t, x_2 + t) &= P\left[T > x_2 - x_1\right] ce^{c(x_2 + t - b_2)} \\ &= P\left[T > x_2 - x_1\right] ce^{c(x_2 - b_2)} e^{ct}. \end{aligned}$$

From (2.28) and (2.29) this can be equivalently written as

$$\frac{\partial}{\partial x_2} F(x_1 + t, x_2 + t) = \frac{F(t,t)}{F(0,0)} \frac{\partial}{\partial x_2} F(x_1, x_2).$$

Integrating with respect to x_2,

$$F(x_1 + t, x_2 + t) F(0,0) = F(x_1, x_2) F(t,t), -\infty < x_1 \le x_2 \le b_2.$$

Similar computations holds for $-\infty < x_2 \le x_1 \le b_1$.

Hence the result. □

The following characterization is now evident.

Corollary 2.1. *A random vector (X_1, X_2) with distribution function with support in $(a_1, b_1) \times (a_2, b_2) \in R^2$, $b_i \ge 0$, $i = 1, 2$ specified in* (2.27) *if and only if the following* *conditions hold.*

1. *The marginal distributions of* X_i, $i = 1, 2$ *are specified by* (2.26).

2. $Z = \max(X_1, X_2)$ *is distributed as* (2.28).

3. Z *and* $T = |X_2 - X_1|$ *are mutually independent.*

Chapter 3

Generalized Reversed Lack of Memory Property

3.1 Introduction

In this Chapter we generalize the reversed lack of memory property which involves operations different than the addition. In particular we shall consider an associative, binary operator $*$.

A binary operator $*$ over real numbers is said to be associative if

$$(x * y) * z = x * (y * z) \tag{3.1}$$

for all real numbers x, y, z. The binary operation $*$ is said to be reducible if $x*y = x*z$ if and only if $y = z$ and if $y * w = z * w$ if and only if $y = z$. The general reducible continuous solution of the functional equation (3.1) is (Aczel (1966)),

$$x * y = g^{-1}(g(x) + g(y)) \tag{3.2}$$

where $g(.)$ is a continuous and strictly monotone function provided x, y, $x * y$ belong to a fixed interval A in the real line. The function $g(.)$ in (3.2) is determined up to a multiplicative constant; $g_1^{-1}(g_1(x)+g_1(y)) = g_2^{-1}(g_2(x)+g_2(y))$ for all x, y in a fixed interval A, implies $g_2 = \alpha g_1$ for all x in that interval, for $\alpha \neq 0$. We assume hereafter that the binary operation is reducible and associative with the function $g(.)$ continuous and strictly increasing. Further more assume that there exists an identity element $\widetilde{e} \in \widetilde{R}$ such that

$$x * \widetilde{e} = x, x \in A \tag{3.3}$$

Further more every continuous, reducible and associative operation defined on an interval A in the real line is commutative (Aczel (1966)). Let X be a random variable with distribution function $F(x)$ having support A. Define

$$\Phi_X^*(s) = \int_A \exp\{isg(x)\}dF(x), -\infty < s < \infty,$$

Note that the above function is the characteristic function of the random variable $g(X)$ and hence determine the distribution function of the random variable $g(X)$ uniquely.

Characterization of distributions through binary operations is given in Muliere and Scarcini (1987) and Muliere and Prakasa Rao (2003). In Prakasa Rao (2004), the bivariate lack of memory property (Roy (2002b)) is generalized and classes of bivariate probability distributions which include bivariate exponential, Weibull, Pareto distributions are characterized under binary associative operations. Also implication of these characterizations on hazard rate is considered in Prakasa Rao (2004).

In this chapter, we generalize the reversed lack of memory property and characterize probability distributions using this property. In Section 3.2, the univariate reversed lack of memory property is generalized using the binary operation and a class of probability distributions which include Type 3 extreme value, power function, reflected Weibull, negative Pareto and truncated extreme value distributions are characterized. In Section 3.3, we studied the almost reversed lack of memory property and considered the distributions with reversed periodic hazard rate under the binary operation. Unlike the situation for the univariate case, there is more than one way in which a model can be extended to the multivariate case. In Section 3.4, we give a bivariate extension of the generalized reversed lack of memory property and characterize a class of bivariate distributions which include the characterized extension (CE) model of Roy (2002a) apart from the bivariate reflected Weibull and power function distributions. We also prove the equality of local proportionality of the reversed hazard rate and generalized reversed lack of memory property. A few characterizations of the model are also discussed.

3.2 Univariate Characterizations

Here we shall extend the reversed lack of memory property discussed in the previous Chapter using a binary associative operator $*$.

Definition 3.1. *Let X be a random variable having distribution function $F(x)$ and survival function $R(x)$ with support (a, b) where $a < 0$ and $b \geq 0$. Then X is said to have the generalized reversed lack of memory property (GRLMP) if*

$$P(X \leq x | X \leq x * t) = P(X \leq e | X \leq t) \tag{3.4}$$

for all $a < x < x * t \leq b < \infty$, $a \leq e$, where the binary operator $*$ is associative specified by (3.1) *and its continuous solution is given by* (3.2).

In terms of the distribution function, (3.4) can be written as

$$F(x)F(t) = F(x * t)F(e), \tag{3.5}$$

for all $a < x < x * t \leq b < \infty, a \leq e$.

A class of distributions characterized by the generalized reversed lack of memory property is given in the following theorem which show that the continuous solution of (3.5) are generalized proportional reversed hazards (PRH) models.

Theorem 3.1. *The continuous solution of* (3.5) *is*

$$F(x) = \exp[c(g(x) - g(b))] \tag{3.6}$$

with $c > 0$ *and* $g^{-1}(-\infty) < x < b < \infty$ *where* g *is a continuous and strictly monotone function.*

Proof. Combining (3.2) and (3.5) we have

$$F(x)F(t) = F\left[g^{-1}(g(x) + g(t))\right] F(e) \tag{3.7}$$

Writing $s = g(x)$, $u = g(t)$ and $F \circ g^{-1} = H$, (3.7) can be rewritten as

$$H(s)H(u) = H(s + u)H(g(e))$$

for all $-\infty < s < g(b)$ and $g(e) \leq u \leq g(b)$ which implies

$$G(s)G(u) = G(s + u),$$

where $G(s) = \frac{H(s)}{H(g(e))}$ for all $-\infty < s < g(b)$ and $g(e) \leq u \leq g(b)$.

The solution to the above Cauchy functional equation is (Aczel (1966))

$$G(s) = e^{cs}, -\infty < s < g(b), c > 0.$$

Thus we have

$$F(x) = e^{cg(x)}F(e), x \in \left(g^{-1}(-\infty), b\right). \tag{3.8}$$

Now taking $x = b$ we get,

$$F(e) = e^{-cg(b)}.$$

Thus (3.8) reduces to

$$F(x) = \begin{cases} \exp\left[c\left(g(x) - g(b)\right)\right], x \in (g^{-1}(-\infty), b) \\ 1, x \geq b, c > 0, b < \infty \end{cases}$$

Hence the theorem. □

Remark 3.1. *If we particularize the operation $*$ we obtain different types of distributions. We now present a few members of this class.*

Example 3.1. *Type 3 extreme value distribution (Gumbel (1958), Castillo and Hadi (1995)).*

For $x * y$=$x + y$, we get $g(x) = x$, $x \in (-\infty, b)$ with $a = -\infty$, $b < \infty$ and $e = 0$. The distribution function is now given by

$$F(x) = \begin{cases} \exp\left[c\left(x - b\right)\right], x < b \\ 1, x \geq b, c > 0, b < \infty \end{cases}$$

The lack of memory property for exponential distribution can be reduced from reversed lack of memory property as given below.

We have the reversed lack of memory property is given by

$$P\left[X \leq x\right] P\left[X \leq t\right] = P\left[X \leq x + t\right] P\left[X \leq 0\right]$$

for all $-\infty < x \leq x + t \leq b < \infty$ which implies

$$P\left[-X \geq -x\right] P\left[-X \geq -t\right] = P\left[-X \geq -(x + t)\right] P\left[-X \geq 0\right]$$

or,

$$P\left[b - X \geq b - x\right] P\left[b - X \geq b - t\right] = P\left[b - X \geq b - (x + t)\right] P\left[b - X \geq b\right].$$

Taking $Y = b - X$*, the above equation can be expressed as*

$$P\left[Y \geq b - x\right] P\left[Y \geq b - t\right] = P\left[Y \geq b - (x + t)\right] P\left[Y \geq b\right]. \tag{3.9}$$

From Johnson et al. (1995), it follows that if X *has a Type 3 extreme value distribution in the support of* $(-\infty, b)$*,* $b < \infty$*, then* $Y = b - X$ *follows an exponential distribution in the interval* $(0, \infty)$*.*

Hence

$$\begin{aligned} P\left[Y \geq b - (x+t)\right] P\left[Y \geq b\right] &= e^{-c[b-(x+t)]} e^{-cb} \\ &= e^{-c[(b-x)+(b-t)]} \\ &= P\left[Y \geq (b-x) + (b-t)\right]. \end{aligned}$$

Thus (3.9) *reduces to*

$$P\left[Y \geq b - x\right] P\left[Y \geq b - t\right] = P\left[Y \geq (b-x) + (b-t)\right]$$

which implies

$$P\left[Y \geq x'\right] P\left[Y \geq t'\right] = P\left[Y \geq x' + t'\right]$$

or

$$R(x')R(t') = R(x' + t')$$

where $x' = b - x$*,* $t' = b - t$ *and* $R(x') = 1 - F(x')$*, the survival function. Hence the result.*

Example 3.2. *Power function distribution.*

1. *For* $x * y$*=*xy*, we get* $g(x) = \log x$*,* $x \in (0, b)$ *with* $b < \infty$ *and* $e = 1$*. In this case the distribution function is given by*

$$F(x) = \begin{cases} \left(\frac{x}{b}\right)^c, 0 \leq x < b \\ 1, x \geq b, c > 0 \end{cases}$$

which is the distribution function for the power function distribution.

The dullness property for Pareto distribution can be reduced from the generalized reversed lack of memory property as follows.

Here the reversed lack of memory property becomes

$$P\left[X \leq x\right] P\left[X \leq t\right] = P\left[X \leq xt\right] P\left[X \leq 1\right]$$

for all $0 \leq x \leq b$, $1 \leq x \leq b$ *and* $xt \leq b$ *which implies*

$$P\left[\frac{b}{X} \geq \frac{b}{x}\right] P\left[\frac{b}{X} \geq \frac{b}{t}\right] = P\left[\frac{b}{X} \geq \frac{b}{xt}\right] P\left[\frac{b}{X} \geq b\right].$$

Taking $Y = \frac{b}{X}$, *the above equation can be written as*

$$P\left[Y \geq \frac{b}{x}\right] P\left[Y \geq \frac{b}{t}\right] = P\left[Y \geq \frac{b}{xt}\right] P\left[Y \geq b\right]. \tag{3.10}$$

From Johnson et al. (1995), it follows that if X *has a power function distribution in the support of* $(0, b)$, $b < \infty$, *then* $Y = \frac{b}{X}$ *follows the Pareto distribution in the interval* $(1, \infty)$.

Hence

$$\begin{aligned} P\left[Y \geq \frac{b}{xt}\right] P\left[Y \geq b\right] &= \left(\frac{b}{xt}\right)^{-(c+1)} b^{-(c+1)} \\ &= \left(\frac{b}{x}\frac{b}{t}\right)^{-(c+1)} \\ &= P\left(Y \geq \frac{b}{x}\frac{b}{t}\right). \end{aligned}$$

Thus (3.10) *reduces to*

$$P\left(Y \geq \frac{b}{x}\right) P\left(Y \geq \frac{b}{t}\right) = P\left(Y \geq \frac{b}{x}\frac{b}{t}\right).$$

Or,

$$R(x')R(t') = R(x't')$$

where $x' = \frac{b}{x}$, $t' = \frac{b}{t}$ *and* $R(x') = 1 - F(x')$, *the survival function. Hence the result.*

2. *For* $x * y$=$x + y + xy$, *we get* $g(x) = \log(x+1)$, $x \in (-1, b)$ *with* $b < \infty$ *and* $e = 0$. *In this case the distribution function is given by*

$$F(x) = \begin{cases} \left(\frac{x+1}{b+1}\right)^c, -1 \leq x < b \\ 1, x \geq b, c > 0 \end{cases}$$

which is the distribution function for the power function distribution in the support of $(-1, b)$ *with* $b < \infty$.

Example 3.3. *Reflected Weibull distribution (Lai and Xie (2005)).*

If we take $x * y = \sqrt{x^2 + y^2}$, *we get* $g(x) = -x^2$, $x \in (-\infty, 0)$ *with* $e = 0$. *The corresponding distribution function is given by*

$$F(x) = \begin{cases} \exp\left[-cx^2\right], x < 0 \\ 1, x \geq 0, c > 0, \end{cases}$$

which is the distribution function for a reflected Weibull distribution.

Example 3.4. *Negative Pareto distribution (Malinowska and Szynal (2008)).*

For $x * y = x + y - xy$, *we get* $g(x) = -\log(1 - x)$, $x \in (-\infty, 0)$ *with* $e = 0$. *In this case the distribution function is given by*

$$F(x) = \begin{cases} (1-x)^{-c}, x < 0 \\ 1, x \geq 0, c > 1 \end{cases}$$

Remark 3.2. *Malinowska and Szynal (2008) characterized a class of continuous distributions by the conditional expectation of the k^{th} lower record values and this class contains the distributions discussed above. Thus the above class can be considered as a subclass of Malinowska and Szynal (2008).*

Remark 3.3. *If X is a random variable in the support of $(g^{-1}(e), b)$ with $F(e) \neq 0$, then X satisfies the generalized reversed lack of memory property if and only if*

$$F(x) = \begin{cases} \exp\left[c\left(g(x) - g(b)\right)\right], x \in (g^{-1}(-\infty), b) \\ 1, x \geq b, c > 0, b < \infty \end{cases}$$

which has a probability mass at $x = g^{-1}(e)$. Thus there exist no absolutely continuous distribution satisfying generalized reversed lack of memory in the interval $(g^{-1}(e), b)$.

Remark 3.4. *The distribution specified in* (3.6) *can be considered as a proportional reversed hazard (PRH) model, since $F(x)$ can be written as $F(x) = [F^*(x)]^c$, where $F^*(x) = \exp[g(x) - g(b)]$, $g^{-1}(-\infty) < x < b$ and $c > 0$. Then the reversed hazard rate can be expressed as $\lambda(x) = c\lambda^*(x)$, $c > 0$ where $\lambda^*(x)$ is the reversed hazard rate of $F^*(x)$. Hence, all the distributions discussed above can be considered as proportional reversed hazard models.*

In the next section we studied about the almost reversed lack of memory property.

3.3 Almost Reversed Lack of Memory Property

A random variable X is said to have the almost reversed lack of memory property (ARLMP) if the equation (3.4) holds for a sequence $a < x_n < b$, $n \geq 1$ for all $0 \leq t \leq b$.

Suppose that $*$ is a binary operation with an identity element $e \in \tilde{R}$ and further suppose that the equation (3.5) holds for a random variable X, with a continuous distribution function F with support $(g^{-1}(-\infty), b)$, for a sequence $g^{-1}(-\infty) < x_n < b$,

$n \geq 1$ for all $e \leq t \leq b$. Here $g(.)$ is the continuous strictly increasing function corresponding to the binary associative operation $*$. Equation (3.5) implies that

$$F(x_n * t)F(e) = F(x_n)F(t), n \geq 1 \tag{3.11}$$

A random variable X satisfying the equation (3.11) is said to have the almost reversed lack of memory property under the binary associative operation $*$.

We now characterize the class of all such distributions. Equation (3.11) shows that

$$F\left[g^{-1}\left(g\left(x_n\right)+g\left(t\right)\right)\right]F(e) = F(x_n)F(t), n \geq 1 \tag{3.12}$$

Let $g(x_n) = u_n$, $g(t) = d$ and $F \circ g^{-1} = H$, then (3.12) becomes

$$H(u_n + d)H(g(e)) = H(u_n)H(d), n \geq 1.$$

Or,

$G(u_n + d) = G(u_n)G(d)$, where $G(u_n) = \frac{H(u_n)}{H(g(e))}$.

Applying the results from Lau and Rao (1984), we get

$$G(u) = p_1(u)e^{c_1 u} + p_2(u)e^{c_2 u} \tag{3.13}$$

for some $c_1, c_2 \in R$ (c_1 may be equal to c_2) and the function $p_i(u)$ has period d such that $0 < p_i(x + d) = p_i(x)$, $i = 1, 2$.

Let us take $c_1 = c_2$. Then (3.13) becomes

$$\frac{H(u_n)}{H(g(e))} = [p_1(u) + p_2(u)]e^{cu}$$

which implies

$$F\left(g^{-1}(g(x))\right) = [p_1(g(x)) + p_2(g(x))]e^{cg(x)}F(e) \tag{3.14}$$

Since $F(b) = 1$, we have

$$F\left(e\right) = \left[p_1(g(b)) + p_2(g(b))\right]^{-1} e^{-cg(b)}.$$

Hence,

$$F(x) = \frac{\left[p_1(g(x)) + p_2(g(x))\right]}{\left[p_1(g(b)) + p_2(g(b))\right]} \exp\left(c[g(x) - g(b)]\right), \tag{3.15}$$

for all $x \in (g^{-1}(-\infty), b)$, $c > 0$.

Thus, we have the following result.

Theorem 3.2. *A random variable X with a continuous distribution function has the almost reversed lack of memory property under a binary operation $*$ as described above if and only if its distribution function F is of the form* (3.15)*, where $g(.)$ is the continuous strictly increasing function corresponding to the binary associative operation $*$ and $p_i(.)$ is a periodic function with period d for some constant $d > 0$ such that $x_n = nd$, $n \geq 1$.*

Remark 3.5. *It is easy to see that*

$$p\left(g(x) + d\right) = p\left(g(x)\right)$$

which implies that

$$p\left(g(x) + g(g^{-1}(d))\right) = p\left(g(x)\right)$$

or,

$$p\left(g(x * g^{-1}(d))\right) = p\left(g(x)\right), x \in \left(g^{-1}(-\infty), b\right)$$

or,

$$\left(p \circ g\right)\left(x * \rho\right) = \left(p \circ g\right)\left(x\right),$$

for all $x \in (g^{-1}(-\infty), b)$ and for some constant $\rho > g^{-1}(-\infty)$ where $\left(p \circ g\right)(x) = p(g(x))$.

In other words the function $(p \circ g)(.)$ *is periodic under the operation* $*$ *with period* ρ.

Remark 3.6. *For different forms for the binary operator* $*$*, we get different members of the family* (3.15) *which are given in Table 3.1.*

3.3.1 Distributions with reversed periodic failure rate under the binary operation $*$.

Consider a binary operation $*$ with an identity e as described earlier. Let $g(.)$ be the corresponding continuous strictly increasing function such that

$$x * y = g^{-1}\left[g(x) + g(y)\right].$$

Let X be random variable with a continuous distribution function of the form

$$F(x) = \frac{[p_1(g(x)) + p_2(g(x))]}{[p_1(g(b)) + p_2(g(b))]} \exp\left(c[g(x) - g(b)]\right),$$

for all $x \in (g^{-1}(-\infty), b)$, $c > 0$ and $(p_i \circ g)(.)$, $i = 1, 2$ is periodic under the operation $*$ with period $\rho > g^{-1}(-\infty)$.

Suppose the function $p(g(.))$ is differentiable with respect to x. Then the probability density function of X is given by

$$f(x) = \tfrac{1}{[p_1(g(b)) + [p_2(g(b))]} \{c[p_1(g(x)) + p_2(g(x))] \exp(c[g(x) - g(b)])g'(x) \\ + \exp(c[g(x) - g(b)])[p_1'(g(x)) + p_2'(g(x))]\}$$

for all $x \in (g^{-1}(-\infty), b)$.

Then the reversed failure rate is given by,

$$\lambda(x) = \frac{f(x)}{F(x)} = cg'(x) + \frac{[p_1'(g(x)) + p_2'(g(x))]}{[p_1(g(b)) + p_2(g(b))]},$$

for all $x \in (g^{-1}(-\infty), b)$.

It is easy to see that

$$(p \circ g)(x * \rho) = (p \circ g)(x)$$

and

$$(p' \circ g)(x * \rho)\, g'(x * \rho) = (p' \circ g)(x)\, g'(x)$$

for $x \in (g^{-1}(-\infty), b)$ from the periodicity of the function $(p \circ g)(.)$ under the binary operation $*$ with period ρ.

Further more,

$$g(x * \rho) = g(x) + g(\rho), x \in \left(g^{-1}(-\infty), b\right)$$

and hence

$$g'(x * \rho) = g'(x), x \in \left(g^{-1}(-\infty), b\right).$$

Therefore,

$\lambda(x * \rho) = \lambda(x)$, $x \in (g^{-1}(-\infty), b)$ with $F(x * \rho) < 1$.

This shows that the distribution function F has reversed periodic failure rate with period ρ under the binary operation $*$.

3.4 Bivariate Extension

In this section we evolve the concept of generalized reversed lack of memory property to higher dimensions. In Roy (2002a) a distribution free characterization of models retaining the local proportionality, that is $\lambda_i(x_1, x_2)$ must be locally proportional to $\lambda_i(x_i)$ where the constant of proportionality, $c_i(x_{3-i})$, using dependence only on x_{3-i}, $i = 1, 2$, is given. The above model is referred to as the characterized extension model. Here we extend the generalized reversed lack of memory property so that it characterizes the characterized extension model and since the local proportionality of the reversed hazard rate is maintained we term it the generalized reversed local lack of memory property (GRLLMP).

Definition 3.2. *A random vector $X = (X_1, X_2)$ in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ with $a_i < 0$ and $b_i \geq 0$, $i = 1, 2$ is said to have the generalized reversed local lack of memory property (GRLLMP) if*

$$P\left(X_i \leq x_i | X_i \leq x_i * t_i, X_j \leq x_j\right) = P\left(X_i \leq e | X_i \leq t_i, X_j \leq x_j\right) \tag{3.16}$$

for all x_i and t_i such that $a_i < x_i \leq x_i * t_i \leq b_i$, $a_i \leq e$, $e \leq t_i \leq b_i$; $i, j = 1, 2$; $i \neq j$, $F(e) \neq 0$.

For $i = 1$, (3.16) can be expressed in terms of distribution function as

$$F\left(x_1, x_2\right) F\left(t_1, x_2\right) = F\left(x_1 * t_1, x_2\right) F\left(e, x_2\right) \tag{3.17}$$

for all x_1 and t_1 such that $a_1 < x_1 \leq x_1 * t_1 \leq b_1 < \infty$, $a_1 \leq e$, $e \leq t_1 \leq b_1$.

As $x_{3-i} \to b_i$ we get

$$F_i\left(x_i\right) F_i\left(t_i\right) = F_i\left(x_i * t_i\right) F_i\left(e\right),$$

for all x_i and t_i such that $a < x_i \leq x_i * t_i \leq b_i < \infty$, $a_i \leq e$, $i = 1, 2$.

It hence follows from Theorem 3.1 that

$$F_i(x_i) = \begin{cases} \exp\left[c_i\left(g(x_i) - g(b_i)\right)\right], g^{-1}(-\infty) < x_i \leq b_i \\ 1, x_i \geq b_i, c_i > 0, i = 1, 2 \end{cases} \tag{3.18}$$

Now combining (3.17) and (3.2) we get

$$F\left(g^{-1}\left(g(x_1) + g(t_1)\right), x_2\right) F(e, x_2) = F(x_1, x_2) F(t_1, x_2).$$

By taking $s_i = g(x_i)$, $i = 1, 2$, $u_1 = g(t_1)$, $F \circ g^{-1} = H$, we can write

$$H(s_1 + u_1, s_2)H(g(e), s_2) = H(s_1, s_2)H(u_1, s_2)$$

for all $g(a_i) < s_i \le g(b_i)$, $i = 1, 2$ and $g(e) \le u_1 \le g(b_1)$, which gives

$$G(s_1 + u_1, s_2) = G(s_1, s_2)G(u_1, s_2) \tag{3.19}$$

for all $g(a_i) < s_i \le g(b_i)$, $i = 1, 2$ and $g(e) \le u_1 \le g(b_1)$ where $G(s_1, s_2) = \frac{H(s_1,s_2)}{H(g(e),s_2)}$.

In a similar manner, we get

$$G(s_1, s_2 + u_2) = G(s_1, s_2)G(s_1, u_2) \tag{3.20}$$

for all $g(a_i) < s_i \le g(b_i)$, $i = 1, 2$ and $g(e) \le u_2 \le g(b_2)$ where $G(s_1, s_2) = \frac{H(s_1,s_2)}{H(s_1,g(e))}$.

In (3.19), s_2 can be interpreted as a fixed parameter. Hence (3.19) reduces directly to the functional equation as

$$f_{s_2}(s_1 + u_1) = f_{s_2}(s_1)f_{s_2}(u_1).$$

Because of the assumed positivity of $G(s_1, s_2)$, it follows that (Aczel (1966))

$$f_{s_2}(s_1) = \exp\left[c(s_2)s_1\right]$$

where the constant c naturally still depends on the parameter s_2.

Thus,

$$G(s_1, s_2) = \exp\left[c(s_2)s_1\right].$$

If we substitute this into (3.20), then

$$c(s_2 + u_2) = c(s_2)c(u_2)$$

Again appealing to the positivity of $G(s_1, s_2)$, we have

$$c(s_2) = cs_2,$$

so that

$$G(s_1, s_2) = \exp\left[cs_1s_2\right], g(a_i) < s_i \leq g(b_i), i = 1, 2; c > 0$$

which implies

$$\frac{H(s_1, s_2)}{H(g(e), s_2)} = \exp\left[cs_1s_2\right], g(a_i) < s_i \leq g(b_i), i = 1, 2; c > 0$$

or

$$F(x_1, x_2) = \exp\left[cg(x_1)g(x_2)\right] F(e, x_2),$$

for all $g^{-1}(-\infty) < x_i \leq b_i$, $i = 1, 2$ and $c > 0$.

As $x_1 \to b_1$, we get

$$F_2(x_2) = \exp\left[cg(b_1)g(x_2)\right] F(e, x_2), g^{-1}(a_2) < x_2 \leq b_2.$$

Thus, it follows that

$$F(x_1, x_2) = \exp\left[cg(x_2)\left\{g(x_1) - g(b_1)\right\} + c_2\left\{g(x_2) - g(b_2)\right\}\right], \tag{3.21}$$

for all $g^{-1}(-\infty) < x_i \leq b_i$, $i = 1, 2$ and $c, c_2 > 0$.

Proceeding similarly and using (3.18) we have

$$F(x_1, x_2) = \exp\left[cg(x_1)\left\{g(x_2) - g(b_2)\right\} + c_1\left\{g(x_1) - g(b_1)\right\}\right], \tag{3.22}$$

for all $g^{-1}(-\infty) < x_i \leq b_i$, $i = 1, 2$ and $c, c_1 > 0$.

Identifying (3.21) and (3.22) we have

$$cg(x_2)\left[g(x_1)-g(b_1)\right]+c_2\left[g(x_2)-g(b_2)\right] = cg(x_1)\left[g(x_2)-g(b_2)\right]+c_1\left[g(x_1)-g(b_1)\right]$$

or,

$$\frac{cg(x_2)-c_1}{g(x_2)-g(b_2)} = \frac{cg(x_1)-c_2}{g(x_1)-g(b_1)}$$

which implies

$$\frac{cg(x_i)-c_j}{g(x_i)-g(b_i)} = \theta,$$

a constant independent of x_i and x_j with $c_1c_2 > \theta \geq 0$, $i, j = 1, 2; i \neq j.$

Thus,

$$cg(x_i) = c_j + \theta\left[g(x_i)-g(b_i)\right], i, j = 1, 2; i \neq j. \tag{3.23}$$

Substituting (3.23) in (3.21) or (3.22), we get

$$F(x_1, x_2) = \exp\left\{c_1\left[g(x_1)-g(b_1)\right] + c_2\left[g(x_2)-g(b_2)\right] + \theta\left[g(x_1)-g(b_1)\right]\left[g(x_2)-g(b_2)\right]\right. \tag{3.24}$$

for all $g^{-1}(-\infty) < x_i \leq b_i < \infty$, $c_i > 0$, $i = 1, 2$ and $c_1c_2 > \theta \geq 0$.

Remark 3.7. *The variables X_1 and X_2 are independent if and only if $\theta = 0$.*

Remark 3.8. *For different forms for the binary operator $*$, we get different members of this family which are given in Table 3.2.*

3.4.1 Characterizations

In Roy (2002a) a distribution free characterization of models retaining the local proportionality, that is $\lambda_i(x_1, x_2)$ must be locally proportional to $\lambda_i(x_i)$ where the constant of proportionality, $c_i(x_{3-i})$, using dependence only on x_{3-i}, $i = 1, 2$, is given. The above property uniquely determines the model referred to as the characterized extension model (Roy (2002a)) which is given by

$$F(x_1, x_2) = F_1(x_1)F_2(x_2)\exp\{-\gamma(\log F_1(x_1))(\log F_2(x_2))\}$$

for some γ.

In this section we establish the equality of the reversed local lack of memory property and local proportionality of the bivariate reversed hazard rate, thus giving a more explicit form to the characterized extension model of (Roy (2002a)). The results are summarized in the following theorem.

Theorem 3.3. *A continuous random vector $X = (X_1, X_2)$ in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ with $a_i < 0$ and $b_i \geq 0$, $i = 1, 2$ has a distribution specified by* (3.24) *if and only if the vector valued reversed hazard rate function is of the form*

$$\lambda(x_1, x_2) = (k_1(x_2)g'(x_1), k_2(x_1)g'(x_2))$$

where $k_i(x_j)$ is a function of x_j alone with $k_i(b_j) = c_i$, $i, j = 1, 2$; $i \neq j$ and $g(.)$ is a continuous, monotone increasing function.

Proof. Assume that $\lambda(x_1, x_2)$ has the above form. Then from Roy (2002a), we have

$$\begin{aligned} F(x_1, x_2) &= \exp\left\{-\int_{x_1}^{b_1} r_1(u, b_2)du - \int_{x_2}^{b_2} r_2(x_1, v)dv\right\} \\ &= \exp\left[c_1\left[g(x_1) - g(b_1)\right] + k_2(x_1)\ \left[g(x_2) - g(b_2)\right]\right] \end{aligned} \tag{3.25}$$

and

$$F(x_1, x_2) = \exp\left[k_1(x_2)\left[g(x_1) - g(b_1)\right] + c_2\left[g(x_2) - g(b_2)\right]\right]. \tag{3.26}$$

Comparing (3.25) and (3.26),

$$c_1\left[g(x_1) - g(b_1)\right] + k_2(x_1)\left[g(x_2) - g(b_2)\right] = k_1(x_2)\left[g(x_1) - g(b_1)\right] + c_2\left[g(x_2) - g(b_2)\right]$$

or,

$$\frac{k_1(x_2) - c_1}{g(x_2) - g(b_2)} = \frac{k_2(x_1) - c_2}{g(x_1) - g(b_1)}$$

holding for all x_1, x_2 which is in fact true if and only if,

$$k_i(x_j) = c_i + \theta \left[g(x_j) - g(b_j)\right], i, j = 1, 2; i \neq j$$

where θ is independent of x_1 and x_2. Substituting this in (3.25) or (3.26) we have the distribution function of the required form. The converse is straightforward. □

For the distribution (3.18), the reversed hazard rate $\lambda_i(x_i)$, $i = 1, 2$ is given by

$$\lambda_i(x_i), = c_i g'(x_i), i = 1, 2.$$

Hence the above theorem restated as follows.

Theorem 3.4. *A continuous random vector* $X = (X_1, X_2)$ *in the support of* $(a_1, b_1) \times (a_2, b_2) \in R^2$ *with* $a_i < 0$ *and* $b_i \geq 0$, $i = 1, 2$ *has a distribution specified by* (3.24) *if and only if the vector valued reversed hazard rate function is of the form*

$$\lambda\left(x_1, x_2\right) = \left(k_1(x_2)\lambda_1(x_1), k_2(x_1)\lambda_2(x_2)\right)$$

where $\lambda_i(x_i)$ *is the univariate reversed hazard rate and* $k_i(x_j)$ *is a function of* x_j *alone with* $k_i(b_j) = 1$, $i, j = 1, 2$; $i \neq j$.

Events of the form $X_j \leq x_j$, where X_j represents a continuous random variable are of specific interest in various fields of applied resrearch. In the bivariate set up it is possible to have access on information about the behaviour of the variable X_i given that $X_j \leq x_j$, $i, j = 1, 2$; $i \neq j$ and therefore there is some apparent interest in characterizing the joint distribution of (X_1, X_2) given the forms of the conditional distribution of X_i given that $X_j \leq x_j$.

If the distribution function of $X = (X_1, X_2)$ is of the form (3.24), we have

$$P(X_i \leq x_i | X_j \leq x_j) = \exp c_i[g(x_i) - g(b_i)] + \theta[g(x_i) - g(b_i)][g(x_j) - g(b_j)]$$

so that differentiation with respect to x_i yields

$$\begin{aligned} f(x_i|X_j \leq x_j) = [c_i + \theta[g(x_j) - g(b_j)]]g'(x_i) \exp c_i[g(x_i) - g(b_i)] \\ + \theta[g(x_i) - g(b_i)][g(x_j) - g(b_j)] \end{aligned}$$

which is in the exponentiated form.

Conversely, assuming $f(x_i|X_j \leq x_j)$ are of exponentiated form, it should be of the form

$$f(x_i|X_j \leq x_j) = c_i(x_j) \exp\left\{c_i(x_j)\left[g(x_i) - g(b_i)\right]\right\} g'(x_i)$$

and therefore

$$P\left(X_i \leq x_i | X_j \leq x_j\right) = \exp\left\{c_i(x_j)\left[g(x_i) - g(b_i)\right]\right\} \tag{3.27}$$

Putting $x_j = b_j$ in (3.27) we have

$P\left(X_i \leq x_i\right) = \exp\left\{c_i\left[g(x_i) - g(b_i)\right]\right\}$, where$c_i = c_i(b_j)$.

Thus, the joint distribution function can be written either as

$$\begin{aligned} F(x_1, x_2) &= P\left(X_1 \leq x_1 | X_2 \leq x_2\right) P\left(X_2 \leq x_2\right) \\ &= \exp\left\{c_1(x_2)\left[g(x_1) - g(b_1)\right] + c_2\left[g(x_2) - g(b_2)\right]\right\} \end{aligned} \tag{3.28}$$

or

$$\begin{aligned} F(x_1, x_2) &= P\left(X_2 \leq x_2 | X_1 \leq x_1\right) P\left(X_1 \leq x_1\right) \\ &= \exp\left\{c_1\left[g(x_1) - g(b_1)\right] + c_2(x_1)\left[g(x_2) - g(b_2)\right]\right\}. \end{aligned} \tag{3.29}$$

Equating (3.28) and (3.29)

$$c_1(x_2)\left[g(x_1)-g(b_1)\right]+c_2\left[g(x_2)-g(b_2)\right]=c_1\left[g(x_1)-g(b_1)\right]+c_2(x_1)\left[g(x_2)-g(b_2)\right]$$

which after rearrangement yields,

$$\frac{c_1(x_2)-c_1}{g(x_2)-g(b_2)}=\frac{c_2(x_1)-c_2}{g(x_1)-g(b_1)}$$

for all $a_i < x_i \leq b_i$, $i = 1, 2$.

Using similar arguments, we find that

$$c_i(x_j)=c_i+\theta\left[g(x_j)-g(b_j)\right], i,j=1,2; i\neq j \tag{3.30}$$

where θ is a constant. Substituting (3.30) in (3.28) or (3.29), we arrive at (3.24) and this completes the proof.

Thus, we have the following result.

Theorem 3.5. *Let $X = (X_1, X_2)$ be a bivariate random vector in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ with $a_i < 0$ and $b_i \geq 0$, $i = 1, 2$ admitting an absolutely continuous distribution with respect to a Lebesgue measure. The distribution function of X is of the form* (3.24) *with continuous and strictly increasing $g(.)$ if and only if for every $a_i < x_i \leq b_i$, the conditional densities $f(x_i|X_j \leq x_j)$, $i, j = 1, 2$; $i \neq j$ are in the exponentiated form.*

Summarizing all the above, we have the following.

Theorem 3.6. *For an absolutely continuous random vector $X = (X_1, X_2)$ in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ with $a_i < 0$ and $b_i \geq 0$, $i = 1, 2$, the following statements are equivalent.*

1. *$X = (X_1, X_2)$ satisfies the GRLLMP specified by* (3.16).

2. *The distribution function of $X = (X_1, X_2)$ is* (3.24).

3. *The bivariate reversed hazard rate of* $X = (X_1, X_2)$ *is of the form*

$$\lambda(x_1, x_2) = (k_1(x_2)\lambda_1(x_1), k_2(x_1)\lambda_2(x_2)),$$

   where $\lambda_i(x_i)$ *is the univariate reversed hazard rate and* $k_i(x_j)$ *is a function of* x_j *alone with* $k_i(b_j) = 1$, $i, j = 1, 2$; $i \neq j$.

4. *The conditional densities* $f(x_i|X_j \leq x_j)$, $i, j = 1, 2$; $i \neq j$ *are in the exponentiated form*

Now we consider the reversed mean residual life function of a bivariate random vector.

Let $X = (X_1, X_2)$ be an absolutely continuous random vector. We define the bivariate reversed mean residual life function (BRMRLF) of X as

$$m(x_1, x_2) = (m_1(x_1, x_2), m_2(x_1, x_2)) \tag{3.31}$$

where $m_i(x_1, x_2) = E(x_i - X_i|X_1 \leq x_1, X_2 \leq x_2)$, $i = 1, 2$.

The BRMRLF uniquely determines the distribution function by the relation (Nair and Asha (2008)),

$$F(x_1, x_2) = \frac{m_i(b_1, b_2)}{m_i(x_j, b_i)} \frac{m_j(x_j, b_i)}{m_j(x_1, x_2)} \exp\left[-\int_{x_i}^{b_i} \frac{dt}{m_i(b_j, t)} - \int_{x_j}^{b_j} \frac{dt}{m_j(x_i, t)}\right], \tag{3.32}$$

for $i, j = 1, 2$; $i \neq j$.

In general for the family specified by (3.24), the BRMRLF defined by (3.31) does not have a closed form expression and therefore, there does not exist characterization of the entire family in terms of simple functional forms of the reversed mean residual life. However, it is worthwhile to investigate the existence of characterization theorems comprising certain subclass of the above family. We now present a result in this direction.

Theorem 3.7. *If $g(x_i)$ is differentiable and $[g'(b_i)]^{-1} < \infty$, $i = 1, 2$, then a necessary and sufficient condition that a bivariate density have a reversed mean residual life vector of the form*

$$\left(\frac{a_1(x_2)}{g'(x_1)}, \frac{a_2(x_1)}{g'(x_2)}\right)$$

where $a_2(x_1)$ and $a_1(x_2)$ are nonnegative continuous functions, is that the corresponding distribution function is

$$F(x_1, x_2) = \frac{g'(x_1)g'(x_2)}{g'(b_1)\,g'(b_2)} \exp\{c_1[g(x_1) - g(b_1)] + c_2[g(x_2) - g(b_2)] + \theta[g(x_1) - g(b_1)][g(x_2) - g(b_2)] \tag{3.33}$$

Proof. Given that

$$m_i(x_1, x_2) = \frac{a_i(x_j)}{g'(x_i)}, i, j = 1, 2; i \neq j.$$

From relation (3.32),

$$\begin{aligned} F(x_1, x_2) &= \frac{\frac{a_1(b_2)}{g'(b_1)}\frac{a_2(x_1)}{g'(b_2)}}{\frac{a_1(b_2)}{g'(x_1)}\frac{a_2(x_1)}{g'(x_2)}} \exp\left[-\int_{x_1}^{b_1} \frac{g'(t)}{a_1(b_2)}dt - \int_{x_2}^{b_2} \frac{g'(t)}{a_2(x_1)}dt\right] \\ &= \frac{g'(x_1)g'(x_2)}{g'(b_1)\,g'(b_2)} \exp\left[\frac{g(x_1) - g(b_1)}{a_1(b_2)} + \frac{g(x_2) - g(b_2)}{a_2(x_1)}\right] \end{aligned}$$

and

$$F(x_1, x_2) = \frac{g'(x_1)g'(x_2)}{g'(b_1)\,g'(b_2)} \exp\left[\frac{g(x_1) - g(b_1)}{a_1(x_2)} + \frac{g(x_2) - g(b_2)}{a_2(b_1)}\right].$$

Equating the two forms,

$$(g(x_1) - g(b_1))\left[\frac{1}{a_1(x_2)} - \frac{1}{a_1(b_2)}\right] = (g(x_2) - g(b_2))\left[\frac{1}{a_2(x_1)} - \frac{1}{a_2(b_1)}\right].$$

By the usual arguments, this yields

$$a_i(x_j) = [c_i + \theta[g(x_j) - g(b_j)]]^{-1}, i, j = 1, 2; i \neq j$$

where $c_i = [a_i(b_j)]^{-1}$.

Substituting this value of $[a_i(x_j)]^{-1}$ in $F(x_1, x_2)$ we get the distribution specified by (3.33). The converse is obvious by taking the relation

$$m_i(x_1, x_2) = \frac{\int_a^{x_i} F(x_i, x_j) dx_i}{F(x_1, x_2)}, i, j = 1, 2; i \neq j.$$

□

Remark 3.9. *Observe that for $x * y = x + y$, we get the distribution function*

$$F(x_1, x_2) = \exp\{c_1(x_1 - b_1) + c_2(x_2 - b_2) + \theta(x_1 - b_1)(x_2 - b_2)\} \tag{3.34}$$

for all $-\infty < x_i \leq b_i < \infty$, $c_i > 0$, $i = 1, 2$ and $c_1 c_2 > \theta \geq 0$, which characterizes that the bivariate reversed mean residual lives $m(x_1, x_2)$ are locally constants (Nair and Asha (2008)).

Theorem 3.8. *The only absolutely continuous distribution of the random vector $X = (X_1, X_2)$ defined on $(-\infty, b_1) \times (-\infty, b_2)$, $b_i < \infty$, $i = 1, 2$ satisfying*

$$r_i(x_1, x_2) m_i(x_1, x_2) = 1, i = 1, 2 \tag{3.35}$$

is the bivariate Type 3 extreme value distribution defined in (3.34).

Proof. The bivariate Type 3 extreme value distribution (3.34) verifies

$$r_i(x_1, x_2) = c_i + \theta(x_j - b_j), i, j = 1, 2; i \neq j$$

and

$$m_i(x_1, x_2) = [c_i + \theta(x_j - b_j)]^{-1}, i, j = 1, 2; i \neq j$$

so tha $r_i(x_1, x_2) m_i(x_1, x_2) = 1$, $i = 1, 2$.

Conversely if (3.35) is assumed, using the relation (Nair and Asha (2008)),

$$r_i(x_1,x_2) = [m_i(x_1,x_2)]^{-1}\left[1-\frac{\partial}{\partial x_i}m_i(x_1,x_2)\right], i=1,2$$

we get

$$1-\frac{\partial}{\partial x_i}m_i(x_1,x_2)=1$$

which on integration gives,

$$m_i(x_1,x_2)=k_i(x_j)$$

where $k_i(x_j)$, $i,j = 1,2$; $i \neq j$ is a constant of integration. Thus, we see that $m_i(x_1,x_2)$, $i = 1,2$ are locally constants which characterizes the bivariate Type 3 extreme value distribution (3.34). □

Theorem 3.9. *The only absolutely continuous distribution of the random vector* $X = (X_1, X_2)$ *defined on* $(0,b_1)\times(0,b_2)$, $b_i < \infty$, $i = 1,2$ *satisfying*

$$r_i(x_1,x_2)m_i(x_1,x_2) = A_i(x_j), i,j=1,2; i\neq j$$

is the bivariate power function distribution defined by

$$F(x_1,x_2) = \left(\frac{x_1}{b_1}\right)^{c_1}\left(\frac{x_2}{b_2}\right)^{c_2+\theta\log\left(\frac{x_1}{b_1}\right)},$$

for all $x_i \in (0,b_i)$*;* $c_i > 0$, $i = 1,2$*;* $\theta \geq 0$, *where* $A_i(x_j)$ *is a positive function independent of* x_i.

Table 3.1: **Distributions characterized by almost reversed lack of memory property**

$x * y$	e	$g(x)$	Distribution Function
$x + y$	0	$g(x) = x$, $x \in (-\infty, b)$.	$F(x) = \left[\frac{p_1(x)+p_2(x)}{p_1(b)+p_2(b)}\right] \exp\left(c(x-b)\right)$, $x \in (-\infty, b)$; $c > 0$.
xy	1	$g(x) = \log x$, $x \in (0, b)$.	$F(x) = \left[\frac{p_1(\log x)+p_2(\log x)}{p_1(\log b)+p_2(\log b)}\right] \left(\frac{x}{b}\right)^c$, $x \in (0, b)$; $c > 0$.
$\sqrt{x^2 + y^2}$	0	$g(x) = -x^2, .$ $x \in (-\infty, 0)$.	$F(x) = \left[\frac{p_1(-x^2)+p_2(-x^2)}{p_1(0)+p_2(0)}\right] \exp\left(-cx^2\right)$, $x \in (-\infty, 0)$, $c > 0$.
$x + y - xy$	0	$g(x) = -\log(1-x)$, $x \in (-\infty, 0)$.	$F(x) = \left[\frac{p_1(-\log(1-x))+p_2(-\log(1-x))}{p_1(0)+p_2(0)}\right] (1-x)^{-c}$, $x \in (-\infty, 0)$; $c > 0$.

Table 3.2: **Distributions characterized by GRLLMP**

No.	$x * y$	e	$g(x)$	$r(x_1, x_2)$	Distribution
1	$x + y$	0	$g(x) = x$, $x \in (-\infty, b)$, $b < \infty$.	$(c_1 + \theta(x_2 - b_2),$ $c_2 + \theta(x_1 - b_1))$	Bivariate Type 3 extreme value distribution (Nair and Asha(2008)). $F(x_1, x_2) = \exp\{c_1(x_1 - b_1) + c_2(x_2 - b_2)$ $+\theta(x_1 - b_1)(x_2 - b_2)\}$ $x_i \in (-\infty, b_i)$; $c_i > 0$, $i = 1, 2$; $\theta \geq 0$.
2	xy	1	$g(x) = \log x$, $x \in (0, b), b < \infty$.	$\left(\left[c_1 + \theta \log\left(\frac{x_2}{b_2}\right)\right] \frac{1}{x_1},\right.$ $\left.\left[c_2 + \theta \log\left(\frac{x_1}{b_1}\right)\right] \frac{1}{x_2}\right)$	Bivariate power function distribution (Nair and Asha(2008)). $F(x_1, x_2) = \left(\frac{x_1}{b_1}\right)^{c_1} \left(\frac{x_2}{b_2}\right)^{c_2 + \theta \log\left(\frac{x_1}{b_1}\right)}$, $x_i \in (0, b_i)$; $c_i > 0$, $i = 1, 2$; $\theta \geq 0$.
3	$\sqrt{x^2 + y^2}$	0	$g(x) = -x^2$, $x \in (-\infty, 0)$.	$((\theta x_2^2 - c_1)2x_1,$ $(\theta x_1^2 - c_2)2x_2)$	Bivariate reflected Weibull distribution $F(x_1, x_2) = \exp\{-c_1 x_1^2 - c_2 x_2^2 + \theta x_1^2 x_2^2\}$, $x_i \in (-\infty, b_i)$; $c_i > 0$, $i = 1, 2$; $\theta \geq 0$.
4	$x + y - xy$	0	$g(x) = -\log(1 - x)$, $x \in (-\infty, 0)$	$\left([c_1 - \theta \log(1 - x_2)] \frac{1}{(1-x_1)},\right.$ $\left.[c_2 - \theta \log(1 - x_2)] \frac{1}{(1-x_2)}\right)$	Bivariate negative Paretol distribution $F(x_1, x_2) = (1 - x_1)^{-c_1} (1 - x_2)^{-c_2 + \theta \log(1 - x_1)}$ $x_i \in (-\infty, b_i)$; $c_i > 0$, $i = 1, 2$; $\theta \geq 0$.

Chapter 4

Bivariate Generalized Reversed Lack of Memory Property

4.1 Introduction

Now we try to evolve the concepts of generalized reversed lack of memory property to the higher dimensions. One of the main problems associated with such an attempt is that there is no unique way of evolution. We here consider the definition of bivariate reversed lack of memory property as specified in Chapter 2 to extend the concept of univariate generalized reversed lack of memory property to the bivariate case and derive bivariate models characterized by the respective property. Since multivariate derivations are rather straight forward extensions, it is excluded.

4.2 Bivariate Extensions

Consider a random vector $X = (X_1, X_2)$ in the two-dimensional space with joint distribution function $F(x_1, x_2) = P[X_1 \leq x_1, X_2 \leq x_2]$ in the support of $(a_1, b_1) \times (a_2, b_2) \in R^2$ where

$$a_i = \inf(x_i | F_i(x_i) > 0) \text{ and } b_i = \sup(x_i | F_i(x_i) < 1),$$

with $F(x_1, b_2)$ and $F(b_1, x_2)$ as the marginals of X_i, $i = 1, 2$.

Then a direct extension of RLMP is

$$F(x_1 + t_1, x_2 + t_2)F(0, 0) = F(x_1, x_2)F(t_1, t_2) \tag{4.1}$$

for all x_i and t_i such that $a_i < x_i \leq x_i + t_i \leq b_i$, $a_i < 0$, $i = 1, 2$.

The only solution for (4.1) is (Aczel (1966))

$$F(x_1, x_2) = \exp\left[c_1\left(x_1 - b_1\right) \ + c_2\left(x_2 - b_2\right)\right],$$

for all $-\infty < x_i < b_i$, $c_i > 0$, $i = 1, 2$.

A more fruitful way of extending the reversed lack of memory property to the bivariate case is to investigate the equation

$$F(x_1 + t, x_2 + t).F(0, 0) = F(x_1, x_2).F(t, t), \tag{4.2}$$

for all $a_i < x_i \leq x_i + t \leq b_i$, $a_i < 0$ $i = 1, 2$.

Then the unique solution of (4.2) among probability distribution functions is

$$F(x_1, x_2) = \exp\left[c_1\left(x_1 - b_1\right) \ + \ c_2\left(x_2 - b_2\right) \ + \ c_{12} \ \max\left(x_1 - b_1, \ x_2 - b_2\right)\right],$$

where $-\infty < x_i < b_i$, $c_i > 0$, $i = 1, 2.$; $c_{12} \geq 0$.

We consider analogous equation of (4.1) for the associative binary operators $*$ and $\circ$ given by the following proposition.

Proposition 4.1. *Let*

$$F\left(x_1 * t_1, \ x_2 \circ t_2\right) \ F\left(e_1, \ e_2\right) \ = \ F\left(x_1, \ x_2\right) \ F\left(t_1, \ t_2\right) \tag{4.3}$$

for all $a_1 < x_1 \leq \; x_1 * t_1 \leq b_1$, $a_2 < x_2 \leq x_2 \circ t_2 \leq b_2$; $a_i < e_i, \; i = 1, 2$. *Then*

$$F\left(x_1, x_2\right) = \exp\left[c_1\left[g\left(x_1\right) - g\left(b_1\right)\right] + c_2\left[h\left(x_2\right) - h\left(b_2\right)\right]\right],$$

for all $g^{-1}(-\infty) < x_1 < b_1, \; h^{-1}(-\infty) < x_2 < b_2, \; c_i > 0, \; i = 1, 2$, *where* g *and* h *are continuous and strictly increasing functions.*

Proof. We have

$$F\left(x_1 * t_1, x_2 \circ t_2\right).F\left(e_1, e_2\right) = F\left(x_1, x_2\right) F\left(t_1, t_2\right).$$

Consider the representations $x*y = g^{-1}\left(g(x) + g(y)\right)$ and $x \circ y = h^{-1}\left(h\ (x) + h\ (y)\right)$, then we get

$$\begin{aligned} F\left[g^{-1}\left(g\left(x_1\right) + g\left(t_1\right)\right), h^{-1}\left(h\left(x_2\right) + h\left(t_2\right)\right)\right] F\left[e_1, e_2\right] = \\ F\left[g^{-1}\left(g\left(x_1\right)\right), h^{-1}\left(h\left(x_2\right)\right)\right]\ F\left[g^{-1}\left(g\left(t_1\right)\right), h^{-1}\left(h\left(t_2\right)\right)\right] \end{aligned}$$

which implies

$$\begin{aligned} H\left(g\left(x_1\right) + g\left(t_1\right), h\left(x_2\right) + h\left(t_2\right)\right) H\left(g(e_1), h(e_2)\right) = \\ H\left(g\left(x_1\right), h\left(x_2\right)\right)\ H\left(g\left(t_1\right), h\left(t_2\right)\right) \end{aligned}$$

where $H(.,.) \;=\; F\left(g^{-1}(.), \; h^{-1}(.)\right)$.

Taking $g\left(x_1\right) = u_1, \; h\left(x_2\right) = u_2, \; g\left(t_1\right) = v_1$ and $h(t_2) = u_2$, the above equation becomes,

$$H\left(u_1 + v_1, u_2 + v_2\right) H\left(g(e_1), h(e_2)\right) = \; H\left(u_1, u_2\right) H\left(v_1, v_2\right).$$

which implies

$$G\left(u_1 + v_1, u_2 + v_2\right) = \; G\left(u_1, u_2\right) G\left(v_1, v_2\right)$$

where $G\left(u_1, u_2\right) = \; \frac{H(u_1, u_2)}{H(g(e_1), h(e_2))}$.

The solution to the above Cauchy functional equation is (Aczel (1966)),

$$G\left(u_1, u_2\right) = \ e^{c_1 u_1 + c_2 u_2}, c_1, c_2 > 0$$

which means,

$$\frac{H\left(u_1, u_2\right)}{H\left(g(e_1), h(e_2)\right)} = \ e^{c_1 u_1 + c_2 u_2}$$

or

$$H\left(u_1, u_2\right) = \ e^{c_1 u_1 + c_2 u_2} H\left(g(e_1), h(e_2)\right).$$

That is,

$$F\left[g^{-1}\left(g\left(x_1\right)\right), h^{-1}\left(h\left(x_2\right)\right)\right] = e^{c_1 g(x_1) + c_2 h(x_2)} F(e_1, e_2)$$

which implies,

$$F\left(x_1, x_2\right) = e^{c_1 g(x_1) + c_2 h(x_2)} F(e_1, e_2)$$

Since $F(b_1, b_2) = 1$, we have $F(e_1, e_2) = e^{-c_1 g(b_1) - c_2 h(b_2)}$

Hence,

$$F\left(x_1, x_2\right) = exp\left[c_1\left(g(x_1) - g(b_1)\right) + c_2\left(h(x_2) - h(b_2)\right)\right],$$

for all $g^{-1}(-\infty) < x_i < b_i;\ c_i > 0,\ i = 1, 2.$

Thus, X_1 and X_2 are independent with marginal distribution functions specified by $F_{X_1}(x_1) = \exp[c_1(g(x_1) - g(b_1))]$ and $F_{X_2}(x_2) = \exp[c_2(g(x_2) - g(b_2))]$ where g and h are continuous and strictly increasing functions. □

Now we investigate the possibility of generalizing the bivariate reversed lack of memory property (4.2) using two different associative operations. In order to solve this problem, we need the following.

Lemma 4.1. *(a) Let* $F\left(x_1 * t, \varphi\left(x_2 * t\right)\right) F\left(e_1, \varphi(e_2)\right) \ = \ F\left(x_1, \varphi\left(x_2\right)\right) F\left(t, \varphi\left(t\right)\right)$ *with* φ *continuous and strictly increasing.*

(b) $F_1(x_1 * t) F_1(e_1) = F_1(x_1) F_1(t)$ *with* $F_1(x_1) = F_1(x_1, \varphi(b_2))$.

(c) $F_2(\varphi(x_2 * t)) F_2(\varphi(e_2)) = F_2(\varphi(x_2)) F_2(\varphi(t))$ *with* $F_2(x_2) = F_2(b_1, x_2)$.

Then,

$$F(x_1, x_2) = exp\left[c_1 (g(x_1) - g(b_1)) + c_2 (g(\varphi^{-1}(x_2)) - g(b_2)) + c_{12}\, max\, (g(x_1) - g(b_1), g(\varphi^{-1}(x_2)) - g(b_2))\right]$$

for all $g^{-1}(-\infty) < x_i < b_i;\ c_i > 0,\ i = 1, 2;\ c_{12} \geq 0$.

Proof. From (a), we get

$$F\left[g^{-1}(g(x_1) + g(t)), \varphi\left(g^{-1}(g(x_2) + g(t))\right)\right] F(e_1, \varphi(e_2)) = F(x_1, \varphi(x_2)) F(t, \varphi(t))$$

Setting $g(x_1) = v, g(t) = u, g(x_2) = w$, we have

$$F\left[g^{-1}(v + u), \varphi(g^{-1}(w + u))\right] F\left[g^{-1}(g(e_1)), \varphi(g^{-1}(g(e_2)))\right] = F\left[g^{-1}(v), \varphi(g^{-1}(w))\right] F\left[g^{-1}(u), \varphi(g^{-1}(u))\right] \quad (4.4)$$

When $v = w$, we obtain

$$F\left[g^{-1}(v + u), \varphi(g^{-1}(v + u))\right] F\left[g^{-1}(g(e_1)), \varphi(g^{-1}(g(e_2)))\right] = F\left[g^{-1}(v), \varphi(g^{-1}(v))\right] F\left[g^{-1}(u), \varphi(g^{-1}(u))\right]$$

which implies,

$$G\left[g^{-1}(v + u), \varphi(g^{-1}(v + u))\right] = G\left[g^{-1}(v), \varphi(g^{-1}(v))\right] G\left[g^{-1}(u), \varphi(g^{-1}(u))\right]$$

where $G\left[g^{-1}(v), \varphi(g^{-1}(v))\right] = \frac{F\left[g^{-1}(v), \varphi\left(g^{-1}(v)\right)\right]}{F\left[g^{-1}(g(e_1)), \varphi(g^{-1}(g(e_2)))\right]}$.

This is a Cauchy equation, whose solution is

$$G\left[g^{-1}(v), \varphi\left(g^{-1}(v)\right)\right] = e^{\delta v}, \delta > 0$$

which implies

$$F\left[x_1, \varphi(x_1)\right] = e^{\delta g(x_1)} F\left(e_1, \varphi(e_2)\right),$$

which means

$$F\left[e_1, \varphi(e_2)\right] = e^{-\delta g(b_1)}.$$

Hence,

$$F\left[g^{-1}(v), \varphi\left(g^{-1}(v)\right)\right] = e^{\delta(v - g(b_1))}, \delta > 0 \tag{4.5}$$

Therefore,

$$F\left(x_1, \varphi(x_1)\right) = e^{\delta(g(x_1) - g(b_1))}.$$

From (4.4), we get

$$F\left(x_1 * t, \varphi(x_2 * t)\right) = \frac{F\left(g^{-1}(v), \varphi\left(g^{-1}(w)\right)\right) F\left(g^{-1}(u), \varphi\left(g^{-1}(u)\right)\right)}{F\left(e_1, \varphi(e_2)\right)}$$

When $x_2 = b_2$,

$$\begin{aligned}
F(x_1 * t, \varphi(b_2 * t)) &= \frac{F(g^{-1}(v), \varphi(g^{-1}(b_2)))F(g^{-1}(u), \varphi(g^{-1}(u)))}{F(e_1, \varphi(e_2))} \\
&= \frac{F_1(g^{-1}(v))F(g^{-1}(u), \varphi(g^{-1}(u)))}{F(e_1, \varphi(e_2))} \quad \text{using (b)} \\
&= \frac{F_1(x_1)F(g^{-1}(u), \varphi(g^{-1}(u)))}{F(e_1, \varphi(e_2))} \\
&= \frac{e^{c_1[g(x_1) - g(b_1)]}e^{\delta[g(t) - g(b_1)]}}{e^{\delta g(b_i)}} \\
&= e^{c_1[g(x_1) - g(b_1)] + \delta g(t)}.
\end{aligned}$$

Now let $x_1 * t = s_1$ and $b_2 * t = s_2$ then $g(x_1) = g(s_1) - g(t)$ and $g(b_2) = g(s_2) - g(t)$.

Thus,

$$F(s_1,\varphi(s_2)) = e^{c_1[g(s_1)-g(b_1)-g(s_2)+g(b_2)]+\delta[g(s_2)-g(b_2)]};\ s_2 \geq s_1.$$

Arguing similarly we can prove that

$$F(s_1,\varphi(s_2)) = e^{c_2[g(s_2)-g(b_2)-g(s_1)+g(b_1)]+\delta[g(s_1)-g(b_1)]};\ s_2 \leq s_1.$$

Writing $c_{12} = \delta - c_1 - c_2$ and rearranging we obtain,

$$\begin{aligned} F(s_1,\varphi(s_2)) = \exp\{&c_1(g(s_1)-g(b_1)) + c_2(g(s_2)-g(b_2)) \\ &+ c_{12}\max(g(s_1)-g(b_1), g(s_2)-g(b_2))\}, \end{aligned}$$

or

$$\begin{aligned} F(x_1,\varphi(x_2)) = \exp\{&c_1(g(x_1)-g(b_1)) + c_2(g(x_2)-g(b_2)) \\ &+ c_{12}\max(g(x_1)-g(b_1), g(x_2)-g(b_2))\}. \end{aligned}$$

Hence,

$$\begin{aligned} F(x_1,x_2) = \exp\{&c_1(g(x_1)-g(b_1)) + c_2(g(\varphi^{-1}(x_2))-g(b_2)) \\ &+ c_{12}\max(g(x_1)-g(b_1), g(\varphi^{-1}(x_2))-g(b_2))\} \quad (4.6) \end{aligned}$$

where $g^{-1}(-\infty) < x_i < b_i$; $c_i > 0$, $i = 1, 2$, $c_{12} \geq 0$. □

The next theorem characterizes the distribution characterized by bivariate generalized reversed lack of memory property.

Theorem 4.1. *(a) Let*

$$F(x_1 * t, x_2 \circ z)\, F(e_1, e_2) = F(x_1, x_2)\, F(t, z),$$

for all $x_i \in (a_i, b_i)$, $i = 1, 2$, $t \in (e_1, b_1)$, $z \in (e_2, b_2)$ *with* $x*t = g^{-1}(g(x) + g(t))$, $y \circ z = h^{-1}(h(y) + h(z)),\ h(z) = g(t)$.

(b) $F_1(x_1 * t) F_1(e) = F_1(x_1) F_1(t)$ *with* $F_1(x_1) = F(x_1, b_2)$.

(c) $F_2(x_2 \circ z) F_2(e) = F_2(x_2) F_2(z)$ *with* $F_2(x_2) = F(b_1, x_2)$.

Then,

$$F(x_1, x_2) = \exp\{c_1(g(x_1) - g(b_1)) + c_2(h(x_2) - h(b_2)) + c_{12}\max(g(x_1) - g(b_1), h(x_2) - h(b_2))\} \tag{4.7}$$

for $x_1 \in (g^{-1}(-\infty),\ b_1)$, $x_2 \in (h^{-1}(-\infty),\ b_2)$, $c_i > 0$, $i = 1, 2$; $c_{12} \geq 0$.

Proof. Put $x_2 = h^{-1}(g(v)) = \varphi(v)$. Then (a) becomes

$$F\left(x_1 * t, \varphi\left(g^{-1}(g(v) + g(t))\right)\right) F(e, e) = F(x_1, \varphi(v)) F(t, \varphi(t))$$

This is just (a) of the above lemma. It is very easy to verify that conditions (b) and (c) of this theorem are equivalent to (b) and (c) of above lemma. The equivalence of (4.6) and (4.7) is immediate. □

Remark 4.1. *If we take the same operators* $*$ *and* $*$ *instead of* $*$ *and* $\circ$*, then we get the class of bivariate distributions with identical marginals as*

$$F(x_1, x_2) = \exp\{c_1(g(x_1) - g(b)) + c_2(g(x_2) - g(b)) + c_{12}\max(g(x_1) - g(b), g(x_2) - g(b))\} \tag{4.8}$$

for all $g^{-1}(-\infty) < x_i < b$, $c_i > 0$, $i = 1, 2$; $c_{12} > 0$.

Remark 4.2. *The members of the class of distributions* (4.7) *and* (4.8) *are given in Table 4.1 and Table 4.2 respectively.*

Remark 4.3. *It can be seen that for* (4.7) *and* (4.8)*, a necessary and sufficient condition for* X_1 *and* X_2 *to be independent is* $c_{12} = 0$.

Remark 4.4. *If* (X_1, X_2) *is distributed as* (4.8)*, then the distribution function of* $\max(X_1, X_2)$ *has the following form:*

$$\begin{aligned} P[\max(X_1, X_2) < s] &= P[X_1 < s, X_2 < s] \\ &= F(s, s) \\ &= e^{(c_1+c_2+c_{12})[g(s)-g(b)]} \end{aligned}$$

Hence, the distribution of $\max(X_1, X_2)$ *has the identical form as the marginal distributions of* X_i *,* $i = 1, 2$.

Theorem 4.2. *Let* (X_1, X_2) *be a bivariate random variable with joint distribution function* $F(x_1, x_2)$*. Then* F *is as* (4.8) *if and only if there exist independent and identically distributed random variables* U*,* V*,* W *whose marginal distributions are given by* $F(x) = e^{\lambda[g(x)-g(b)]}$*, such that* $X_1 = \max(U, W)$*,* $X_2 = \max(V, W)$.

Proof.

$$\begin{aligned} F(x_1, x_2) &= P(X_1 < x_1, X_2 < x_2) \\ &= P(U < x_1, W < x_1, V < x_2, W < x_2) \\ &= P(U < x_1).P(V < x_2).P(W < \max(x_1, x_2)) \\ &= e^{\lambda_1[g(x_1)-g(b)]+\lambda_2[g(x_2)-g(b)]+\lambda_{12}\max[g(x_1)-g(b), g(x_2)-g(b)]}. \end{aligned}$$

□

Table 4.1: **Distributions with non-identical marginals**

$x * y$	$x \circ y$	e_1	e_2	$g(x)$	$h(y)$	Distribution
$x+y$	xy	0	1	$g(x)=x$, $x\in(-\infty,b)$.	$h(y)=\log y$, $y\in(0,b)$.	$F(x_1,x_2) = (\frac{x_2}{b_2})^{c_2}\exp\{c_1(x_1-b_1)+c_{12}\max[x_1-b_1,\log(x_2/b_2)]\}$, $x_1\in(-\infty,b_1)$, $x_2\in(0,b_2)$; $c_1,c_2>0$, $c_{12}\geq 0$.
$x+y$	$\sqrt{x^2+y^2}$	0	0	$g(x)=x$, $x\in(-\infty,b)$.	$h(y)=-y^2$, $y\in(-\infty,0)$.	$F(x_1,x_2) = \exp\{c_1(x_1-b_1)-c_2x_2^2+c_{12}\max[x_1-b_1,-x_2^2]\}$, $x_1\in(-\infty,b_1)$, $x_2\in(-\infty,0)$; $c_1,c_2>0$, $c_{12}\geq 0$.
$x+y$	$x+y-xy$	0	0	$g(x)=x$, $x\in(-\infty,b)$.	$h(y)=-\log(1-y)$, $y\in(-\infty,0)$.	$F(x_1,x_2) = (1-x_2)^{-c_2}\exp\{c_1(x_1-b_1)+c_{12}\max[x_1-b_1,-\log(1-x_2)]\}$, $x_1\in(-\infty,b_1)$, $x_2\in(-\infty,0)$; $c_1,c_2>0$, $c_{12}\geq 0$.
$x+y$	$\sqrt{x^2+y^2}$	0	0	$g(x)=\log x$, $x\in(0,b)$.	$h(y)=-y^2$, $y\in(-\infty,0)$.	$F(x_1,x_2)=(\frac{x_1}{b_1})^{c_1}\exp\{-c_2x_2^2+c_{12}\max[\log(x_1/b_1),-x_2^2]\}$, $x_1\in(0,b_1)$, $x_2\in(-\infty,0)$; $c_1,c_2>0$, $c_{12}\geq 0$.
xy	$x+y-xy$	1	0	$g(x)=\log x$, $x\in(0,b)$.	$h(y)=-\log(1-y)$, $y\in(-\infty,0)$.	$F(x_1,x_2)=(\frac{x_1}{b_1})^{c_1}(1-x_2)^{-c_2}\exp\{c_{12}\max[\log(x_1/b_1),-\log(1-x_2)]\}$, $x_1\in(0,b_1)$, $x_2\in(-\infty,0)$; $c_1,c_2>0$, $c_{12}\geq 0$.
$\sqrt{x^2+y^2}$	$x+y-xy$	0	0	$g(x)=-x^2$, $x\in(-\infty,0)$.	$h(y)=-\log(1-y)$, $y\in(-\infty,0)$.	$F(x_1,x_2)=(1-x_2)^{-c_2}\exp\{-c_1x_1^2+c_{12}\max[-x_1^2,-\log(1-x_2)]\}$, $x_1\in(0,b_1)$, $x_2\in(-\infty,0)$; $c_1,c_2>0$, $c_{12}\geq 0$.

Table 4.2: **Distributions with identical marginals**

No.	$x * y$	e	$g(x)$	Distribution
1	$x + y$	0	$g(x) = x, x \in (-\infty, b)$.	Bivariate type 3 extreme value distribution $F(x_1, x_2) = \exp\{c_1(x_1 - b_1) + c_2(x_2 - b_2) + c_{12} \max[x_1 - b_1, x_2 - b_2]\}$, $x_i \in (-\infty, b_i)$; $c_i > 0$, $i = 1, 2$; $c_{12} \geq 0$.
2	xy	1	$g(x) = \log x, x \in (0, b)$.	Bivariate power function distribution $F(x_1, x_2) = (\frac{x_1}{b_1})^{c_1} (\frac{x_2}{b_2})^{c_2} \exp\{c_{12} \max[\log(x_1/b_1), \log(x_2/b_2)]\}$, $x_i \in (0, b_i)$; $c_i > 0$, $i = 1, 2$; $c_{12} \geq 0$.
3	$\sqrt{x^2 + y^2}$	0	$g(x) = -x^2, x \in (-\infty, 0)$.	Bivariate reflected Weibull distribution $F(x_1, x_2) = \exp\{-c_1 x_1^2 - c_2 x_2^2 + c_{12} \max[-x_1^2, -x_2^2]\}$, $x_i \in (-\infty, 0)$; $c_i > 0$, $i = 1, 2$; $c_{12} \geq 0$.
4	$x + y - xy$	0	$g(x) = -\log(1 - x)$, $x \in (-\infty, 0)$.	Bivariate negative Pareto distribution. $F(x_1, x_2) = (1 - x_1)^{-c_1} (1 - x_1)^{-c_2} \exp\{c_{12} \max[-\log(1 - x_1), -\log(1 - x_2)]\}$, $x_i \in (-\infty, 0)$; $c_i > 0$, $i = 1, 2$; $c_{12} \geq 0$.

Chapter 5

Characterizations of the GRLMP Class Based on Past Entropy Measures

5.1 Introduction

Information coding and transmission play a central role in understanding and describing the behavior of biological and engineering systems. Entropy as a baseline concept in the field of information theory was introduced by Shannon (1948) and Wiener (1948), and it is, for instance, also invoked to deal with information in the context of theoretical neurobiology (Johnson and Glantz (2004)).

As is well-known, the classical approach to the description of information related to an absolutely continuous random variable X is based on the differential entropy of X, or Shannon information measure, defined by

$$H(X) = -E\left(\log f(X)\right) = -\int_{-\infty}^{\infty} f(x) \log f(x) dx, \tag{5.1}$$

where 'log' refers to the natural logarithm and $f(x)$ is the probability density function of X. The integrand function on the right-hand side of (5.1) depends on x only via $f(x)$, thus making $H(X)$ shift-independent. In other terms, $H(X)$ is position-free, in the sense that X possesses the same differential entropy as $X + b$, for any $b \in R$. The differential entropy (5.1) presents various deficiencies when it is used as a continuous counterpart of the classical Shannon entropy for discrete random variables. One among these is that, this measure is not always nonnegative. Various attempts have been made in order to define possible alternative information measures. One alternative is due to Schroeder (2004), who proposed a measure which, unlike the classical Shannon's entropy for discrete variables, can be easily and consistently extended to the continuous probability distributions, and unlike differential entropy is always positive and invariant with respect to linear transformation of coordinates.

Rao et al. (2004) defined the cumulative residual entropy (CRE) as

$$\xi(X) = -\int_0^\infty R(x)\log R(x)dx, \tag{5.2}$$

where $R(x) = P\,(X > x)$ is the cumulative residual function or survival function of X. The key idea in the definition (5.2) is to use the cumulative distribution in place of the density function in Shannon's information measure. The distribution function is more regular because it is defined in an integral form unlike the density function, which is computed as the derivative of the distribution. Moreover, in practice what is of interest and/or measurable is the distribution function. For example, if the random variable describes the life span of a light bulb, then the event of interest is not whether the life span equals t, but whether it exceeds t. The definition (5.2) also preserves the well established principle that the logarithm of the probability of an event should represent the information content in the event. Applications of $\xi(X)$, to image alignment and to measurements of similarity between images can be found in Wang and Vemuri (2007) and Wang et al. (2003a, b). The cumulative residual entropy is also suitable to describe the information in problems related to ageing properties of reliability theory based on the mean residual life function (Asadi and Zohrevand (2007)). In addition, in Asadi and Zohrevand (2007) a dynamic version of (5.2) is proposed in order to pinpoint the age effect of a system. The definition of $\xi(X)$ has been extended to the case of distribution with support in $R = (-\infty, \infty)$ has been presented and studied in Drissi et al. (2008) and Zografos and Nadarajah (2005) as

$$\xi(X) = -\int_{-\infty}^{\infty} R(x) \log R(x) dx.$$

However, it is reasonable to presume that in many realistic situations uncertainty is not necessarily related to the future but can also refer to the past. For instance, consider a system whose state is observed only at certain preassigned inspection times. If at time t the system is inspected for the first time and it is found to be ‘down’, then the uncertainty relies on the past, i.e. on which instant in $(0,\, t)$ it has failed. Based on this idea, Di Crescenzo and Longobardi (2009) introduced a new information measure similar to $\xi(X)$, that turns out to be particularly useful to measure information on the inactivity time of a system. The inactivity time is a rather recent concept in reliability that is suitable to describe the time elapsing between the failure of a system and the time when it is found to be ‘down’. In other words, the measure introduced by them called cumulative entropy is suitable to measure information when uncertainty is related to the past, a dual concept of the cumulative residual entropy which relates to uncertainty on the future lifetime of a system. The cumulative entropy for a non-negative random variable X is defined as (Di Crescenzo and Longobardi (2009))

$$C\xi(X) = -\int_{0}^{\infty} F(x) \log F(x) dx, \tag{5.3}$$

where $F(x) = P\,(X \leq x)$ is the distribution function of X. Since the argument of the logarithm is a probability, it easily follows from (5.3) that $C\xi(X)$ takes values in $[0, +\infty]$. In particular, $C\xi(X) = 0$ if and only if X is a constant.

Note that, in (5.3) the support of the random variable X is R^+, the set of non-negative real numbers. Although the positive case is of great interest for many applications, cumulative entropy entail difficulties when working with random variables with supports that are not restricted to positive values. This motivates us to generalize the measure (5.3) from R^+ to $R = (-\infty, \infty)$. This is discussed in Section 2 of the present chapter. Section 3 focuses the implication of generalized reversed lack of memory property on the dynamic cumulative entropy. In Section 4, we consider the β and α-order generalized past entropy measures and characterize the GRLMP class using these measures. These characterizations generalize many existing characterization results.

5.2 Generalized Cumulative Entropy

Let X be a random variable with absolutely continuous distribution function $F(x)$, then the generalized cumulative entropy is defined as

$$C\xi(X) = -\int_{-\infty}^{\infty} F(x) \log F(x) dx. \tag{5.4}$$

Observe that, $C\xi(X) \geq 0$.

The existence of (5.4) can be established without further assumption upon distribution than the existence of absolute moments of order p, $p > 1$.

Theorem 5.1. *$C\xi(X) < \infty$ if for some $p > 1$, $E\left[|X|^p\right] < \infty$.*

Proof. From the definition of $C\xi(X)$ specified in (5.4), we have

$$C\xi(X) = -\int_{-\infty}^{0} F(x) \log F(x) dx - \int_{0}^{\infty} F(x) \log F(x) dx.$$

We first prove the existence of the first term.

Since $\lim_{x \to 0} x \log x = 0$ and $\log x < 0$, we have the inequality,

$$x \log x \leq x \text{ , for } 0 \leq x \leq 1.$$

Using this inequality, we have

$$F(x) \log F(x) \leq F(x).$$

Thus,

$$\int_{-\infty}^{0} F(x)\log F(x)dx \leq \int_{R} F(x)I_{[-\infty,0]}(x)dx,$$

where $I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ otherwise.

Now,

$$\begin{aligned}
F_X(x)I_{(-\infty,0]}(x) &\leq I_{[-1,0]}(x) + F_X(x)I_{(-\infty,-1)}(x) \\
&= I_{[0,1]}(u) + F_X(-u)I_{(1,\infty)}(u); u = -x \\
&= I_{[0,1]}(u) + P\left(X \leq -u\right) I_{(1,\infty)}(u) \\
&\leq I_{[0,1]}(u) + P\left(|X| \geq u\right) I_{(1,\infty)}(u)
\end{aligned}$$

Now using the Markov inequality

$$P\left(|X| \geq u\right) \leq u^{-p} E\left[|X|^p\right]$$

we have

$$F_X(x)I_{(-\infty,0]}(x) \leq I_{[0,1]}(u) + u^{-p} E\left[|X|^p\right] I_{(1,\infty)}(u)$$

.

Hence,

$$\begin{aligned}
\int_R F_X(x)I_{(-\infty,0]}(x)dx &\leq \int_R \left\{I_{[0,1]}(u) + u^{-p} E\left[|X|^p\right] I_{(1,\infty)}(u)\right\} du \\
&\leq 1 + \int_1^{\infty} u^{-p} E\left[|X|^p\right] du
\end{aligned}$$

Since $E\left[|X|^p\right] < \infty$, we have

$$\int_R F_X(x)I_{(-\infty,0]}(x)dx < \infty$$

.

For proving the existence of the second term, we use the inequality

$$x \log x \leq 1 - x \text{ for } 0 \leq x \leq 1,$$

which gives

$$F(x) \log F(x) \leq 1 - F(x) = R(x).$$

Thus,

$$\int_0^\infty F(x) \log F(x) dx \leq \int_0^\infty R(x) dx < \infty,$$

under the assumption $E\left[|X|^p\right] < \infty$, $p > 1$. Hence the result. □

Remark 5.1. *$C\xi(X)$ need not exist for distributions that do not have finite mean. For example, the Pareto distribution with cumulative distribution function $F(t) = 1 - \frac{1}{t}$, $t \geq 1$. Thus all the results discussed here after, is based on the assumption that $C\xi(X) < \infty$.*

It is easy to check that like Shannon entropy the generalized cumulative entropy in (5.4) remains constant with respect to variable translation. That means,

$$C\xi(X + a) = C\xi(X), \forall a \in R. \tag{5.5}$$

In the same way, it is clear that

$$C\xi(aX) = aC\xi(X), \forall a \in R^+. \tag{5.6}$$

When $a < 0$, we do not have such a nice property. However, let us consider the important particular case where the distribution of X has a symmetry of the form, there exist μ, for all x

$$F(\mu - x) = 1 - F(\mu + x). \tag{5.7}$$

In this case, we get the following result.

Theorem 5.2. *For a random variable X that satisfies symmetry property* (5.7)*, one has*

$$\forall a \in R, C\xi(aX) = |a|C\xi(X).$$

Proof. Since it is clear that for all $a \in R^+$, $C\xi(aX) = aC\xi(X)$, we just have to check that $C\xi(-X) = C\xi(X)$, which can be established as follows:

$$\begin{aligned}
-C\xi(-X) &= \int_{-\infty}^{\infty} F_{-X}(x) \log F_{-X}(x) dx \\
&= \int_{-\infty}^{\infty} F_{-X}(x-\mu) \log F_{-X}(x-\mu) dx \\
&= \int_{-\infty}^{\infty} R_X(-x+\mu) \log R_X(-x+\mu) dx \\
&= \int_{-\infty}^{\infty} (1 - R_X(x+\mu)) \log (1 - R_X(x+\mu))\, dx \\
&= \int_{-\infty}^{\infty} F(x+\mu) \log F(x+\mu) dx \\
&= \int_{-\infty}^{\infty} F(x) \log F(x) dx \\
&= -C\xi(X).
\end{aligned}$$

□

Another measure of uncertainty which has gained much interest is the dynamic information measure (Di Crescenzo and Longobardi (2009)). For a random variable X with distribution function $F(x)$ in support of R^+, the new dynamic information measure is defined as

$$C\xi(X;t) = -\int_0^t \frac{F(x)}{F(t)} \log \frac{F(x)}{F(t)}\, dx, t \in (-\infty, \infty), F(t) > 0. \tag{5.8}$$

$C\xi(X;t)$ identifies the cumulative entropy of the event $[X|X \le t]$ and is useful to measure information on the inactivity time $[t - X|X \le t]$.

For a random variable X with distribution function $F(x)$ with support R, (5.8) can be redefined as

$$\begin{aligned} C\xi(X;t) &= -\int_{-\infty}^{t} \frac{F(x)}{F(t)} \log \frac{F(x)}{F(t)} dx, t \in R, F(t) > 0 \\ &= -\frac{1}{F(t)} \int_{-\infty}^{t} F(x) \log F(x) dx + \log F(t) \frac{1}{F(t)} \int_{-\infty}^{t} F(x) dx \qquad (5.9) \\ &= m(t) \log F(t) - \frac{1}{F(t)} \int_{-\infty}^{t} F(x) \log F(x) dx. \end{aligned}$$

We call the measure (5.9) as generalized dynamic cumulative entropy.

5.3 Implication of GRLMP on Dynamic Cumulative Entropy

In this section we study the implication of GRLMP on the generalized dynamic cumulative entropy.

We have

$$C\xi(X;t) = -\frac{1}{F(t)} \int_{-\infty}^{t} F(x) \log \frac{F(x)}{F(t)} dx,$$

for t $\in (-\infty, \infty), F(t) > 0.$

Let X be as in Definition 3.1. Then,

$$\begin{aligned} C\xi(X;t*t') &= -\frac{1}{F(t*t')} \int_{g^{-1}(-\infty)}^{t*t'} F(x) \log \frac{F(x)}{F(t*t')} dx \\ &= -\frac{F(e)}{F(t)F(t')} \int_{g^{-1}(-\infty)}^{t*t'} F(x) \log \frac{F(x)}{F(t*t')} dx. \end{aligned}$$

Take $x = y * t'$ so that $dx = d(y * t')$.

When $x = g^{-1}(-\infty)$, $y = g^{-1}(-\infty) * (t')^{\oplus}$, where $(t')^{\oplus}$ is the inverse such that $t' * (t')^{\oplus} = e$, the identity element. Similarly, when $x = t * t'$, $y = t$.

It follows that,

$$\begin{aligned} C\xi(X; t*t') &= -\frac{F(e)}{F(t)F(t')}\int_{g^{-1}(-\infty)*(t')^{\oplus}}^{t} F(y*t')\log\frac{F(y*t')}{F(t*t')}d(y*t') \\ &= -\frac{F(e)}{F(t)F(t')}\int_{g^{-1}(-\infty)*(t')^{\oplus}}^{t} \frac{F(y)F(t')}{F(e)}\log\frac{F(y)F(t')}{F(t)F(t')}d(y*t') \\ &= -\frac{1}{F(t)}\int_{g^{-1}(-\infty)*(t')^{\oplus}}^{t} F(y)\log\frac{F(y)}{F(t)}d(y*t'). \end{aligned} \tag{5.10}$$

To study the above relation for particular forms of the operator $*$ or equivalently different forms of $g(x)$, we need the following theorems.

Theorem 5.3. *Let X be an absolutely continuous random variable with distribution function $F(t)$, reversed mean residual life $m(t)$, reversed hazard rate $\lambda(t)$ and generalized dynamic cumulative entropy $C\xi(X;t)$. Then*

$$C\xi(X;t) = km(t), k > 0,$$

if and only if X is distributed as

1. *Type 3 extreme value when $k = 1$,*
2. *Power function when $0 < k < 1$,*
3. *Negative Pareto when $k > 1$.*

Proof. The 'if' part of the theorem is straightforward to prove.

To prove the 'only if' part, let us assume that

$$C\xi(X;t) = km(t), k > 0.$$

That is,

$$m(t)\log F(t) - \frac{1}{F(t)}\int_a^t F(x)\log F(x)dx = km(t).$$

Differentiating both sides of this with respect to t gives

$$\begin{aligned} km'(t) &= m'(t)\log F(t) + m(t)\lambda(t) - \log F(t) + \lambda(t)\frac{1}{F(t)}\int_a^t F(x)\log F(x)dx \\ &= m'(t)\log F(t) + m(t)\lambda(t) - \log F(t) + \lambda(t)\left[m(t)\log F(t) - km(t)\right]. \end{aligned} \tag{5.11}$$

On the other hand we have

$$m'(t) = 1 - m(t)\lambda(t).$$

Substituting $m'(t)$ from this last equation in (5.11) we get

$$\begin{aligned} k\left[1 - m(t)\lambda(t)\right] = \left[1 - m(t)\lambda(t)\right]\log F(t) + m(t)\lambda(t) - \log F(t) \\ + m(t)\lambda(t)\log F(t) - km(t)\lambda(t). \end{aligned}$$

From this we obtain for any $t \in (a, b)$, $-\infty \leq a < t \leq b < \infty$

$$m(t)\lambda(t) = k.$$

Or,

$$m'(t) = 1 - k. \tag{5.12}$$

1. When $k = 1$, (5.12) becomes

   $$m'(t) = 0$$

   implies

   $$m(t) = \text{constant},$$

   which characterizes the Type 3 extreme value distribution.

2. For $0 < k < 1$, the proof follows from Theorem 6.2 of Di Crescenzo and Longobardi (2009).

3. For the negative Pareto distribution

$$F(x) = \begin{cases} (1-x)^{-c}, x < 0 \\ 1, x \geq 0, c > 1 \end{cases},$$

the reversed mean residual life function is given by

$$\begin{aligned} m(t) &= \frac{1-t}{c-1} \\ &= m(0) + \frac{t}{1-c}, \end{aligned}$$

where $m(0) = E(-X) = \frac{1}{c-1}$.

Now, integrating both sides of (5.12) with respect to t on $(0, t)$ yields the following linear form for $m(t)$:

$$\begin{aligned} m(t) &= (1-k)\, t + m(0) \\ &= \frac{t}{(1-c)} + m(0), \end{aligned}$$

by taking $k = \frac{c}{c-1} < 1$, which characterizes the negative Pareto distribution.

□

Theorem 5.4. *Let X be an absolutely continuous random variable. Then for all values of t, $t \in (-\infty, b)$,*

$$C\xi(X;t) = \frac{E\left((t-X)^2 \,|X \leq t\right)}{2m(t)}$$

if and only if X has the Type 3 extreme value distribution.

Proof. Assume that

$$C\xi(X;t) = \frac{E\left((t-X)^2 \,|X \leq t\right)}{2m(t)}.$$

Then using Di Crescenzo and Longobardi (2009), we have

$$2m(t)\int_{-\infty}^{t} m(x)f(x)dx = \int_{-\infty}^{t} (t-x)^2 f(x)dx.$$

Differentiating both sides of this with respect to t and after some simplification, we get

$$m(t)m(t)f(t) + m'(t)\int_{-\infty}^{t} m(x)f(x)dx = \int_{-\infty}^{t} (t-x)^2 f(x)dx.$$

Dividing both sides by $F(t)$ we have

$$m'(t)C\xi(X;t) + m(t)m(t)\lambda(t) = m(t),$$

which gives

$$m'(t)C\xi(X;t) + m(t)\left[m(t)\lambda(t) - 1\right] = 0.$$

By using the relation $m'(t) = 1 - m(t)\lambda(t)$, the above equation becomes

$$m'(t)\left[C\xi(X;t) - m(t)\right] = 0.$$

This means either $m'(t) = 0$ or $C\xi(X;t) = m(t)$. In each case we get the required result. That means, the distribution is Type 3 extreme value. □

We now study the implication of the generalized reversed lack of memory property on $C\xi(X;t)$.

Theorem 5.5. *Let X be a random variable in the interval $(-\infty, b)$, $b < \infty$. Then X has the Type 3 extreme value distribution if and only if*

$$C\xi(X;t+t') = C\xi(X;t); t, t' \in (-\infty, b).$$

Theorem 5.6. *Let X be a random variable in the interval $(0, b)$, $b < \infty$. Then X has* *the power function distribution if and only if*

$$C\xi(X;t+t') = t'C\xi(X;t); t,t' \in (0,b).$$

Theorem 5.7. *Let X be a random variable in the interval $(-\infty, 0)$. Then X has the negative Pareto distribution if and only if*

$$C\xi(X;t+t'-tt') = (1-t')C\xi(X;t); t,t' \in (-\infty,0).$$

5.4 Characterizations Based on Generalized Past Entropy Measures

The Khinchin's (1957) measure of information is given by

$$H^{\phi}(f) = \int f(x)\phi\left(f(x)\right) dx \tag{5.13}$$

where ϕ is a convex function such that $\phi(1) = 0$.

In particular when $\phi(f(x)) = \frac{1}{\beta-1}\left[1 - f^{\beta-1}(x)\right]$, $H^{\phi}(f)$ reduces to

$$\overline{H}_{\beta}(f;t) = \frac{1}{\beta-1}\left[1 - \int_a^t \left(\frac{f(x)}{F(t)}\right)^{\beta} dx\right], \tag{5.14}$$

for $\beta > 0$ and $\beta \neq 1$. This measure has been studied in Nanda and Paul (2006b) and they referred to (5.14) as the generalized past entropy of order β. Another measure known as the generalized past entropy of order α was also introduced in Nanda and Paul (2006b) as

$$\overline{H}_{\alpha}(f;t) = \frac{1}{1-\alpha}\log\int_a^t \left(\frac{f(x)}{F(t)}\right)^{\alpha} dx, \tag{5.15}$$

for $\alpha > 0$ and $\alpha \neq 1$.

Observe that $\beta \to 1$ and $\alpha \to 1$ in (5.14) and (5.15) respectively, they reduce to the past entropy, $\overline{H}(f;t)$ (Kundu et al. (2010)) specified in (1.17).

Gupta and Nanda (2002) studied the properties of generalized past entropies of order β and α for the proportional reversed hazard model. The ordering and ageing properties in terms of generalized past entropies have been defined and their properties have been studied in Nanda and Paul (2006b). In their work, they analyzed how the generalized past entropies behave when the distribution is truncated above by t. One may find a similar kind of result in Belzunce et al. (2004).

In Nanda and Paul (2006b) it is shown that if $\overline{H}_\beta(f;t)$ $\left(\left(\overline{H}_\alpha(f;t)\right)\right)$ is increasing in t, then $\overline{H}_\beta(f;t)$ $\left(\left(\overline{H}_\alpha(f;t)\right)\right)$ uniquely determines the underlying distribution function.

In this section, we investigate the implication of the generalized reversed lack of memory property on these measures. The β and α-order generalized past entropies for the GRLMP class (3.6) are given by

$$\overline{H}_\beta(f;t) = \frac{1}{\beta-1}\left[1-\int_{g^{-1}(-\infty)}^{t}\frac{\left(cg'(x)\right)^\beta e^{\beta c[g(x)-g(b)]}}{e^{\beta c[g(t)-g(b)]}}dx\right] \tag{5.16}$$

and

$$\overline{H}_\alpha(f;t) = \frac{1}{1-\alpha}\log\left\{\int_{g^{-1}(-\infty)}^{t}\frac{\left(cg'(x)\right)^\alpha e^{\alpha c[g(x)-g(b)]}}{e^{\alpha c[g(t)-g(b)]}}dx\right\}. \tag{5.17}$$

The following behavior of the generalized past entropy measures of order α and β are the characterizations for the GRLMP class.

Theorem 5.8. *Let X be an absolutely continuous random variable. Then X has Type 3 extreme value distribution if and only if*

$$\overline{H}_\alpha(f;t+t') = \overline{H}_\alpha(f;t),$$

for t, $t' \in (-\infty, b)$.

Proof. Let X has Type 3 extreme value distribution specified by

$$F(x) = e^{c(t-b)}, t \in (-\infty, b), c > 0$$

. Then, from (5.17) we have

$$\overline{H}_\alpha(f;t) = k(\alpha),$$

where $k(\alpha)$ is a function of α independent of t, which gives

$$\overline{H}_\alpha(f;t+t') = \overline{H}_\alpha(f;t); t, t' \in (-\infty, b)$$

Conversely, let

$$\overline{H}_\alpha(f;t+t') = \overline{H}_\alpha(f;t)$$

which gives

$$e^{-\overline{H}_\alpha(f;t+t')} = e^{-\overline{H}_\alpha(f;t)}.$$

From Nanda and Paul (2006b), the above equation becomes

$$\lambda\left(t+t'\right) = \lambda\left(t\right)$$

which implies

$$\lambda\left(t\right) = \text{constant}.$$

Then, from Asha and Rejeesh (2007), the underlying distribution is Type 3 extreme value. □

A similar result hold for $\overline{H}_\beta(f;t)$. Thus we have

Theorem 5.9. *The generalized past entropy measure of order β satisfies*

$$\overline{H}_\beta(f;t+t') = \overline{H}_\beta(f;t),$$

for t, $t' \in (-\infty, b)$ if and only if X is distributed as Type 3 extreme value distribution (2.5).

Corollary 5.1. *The generalized past entropy measures of order α and β are of the form*

1. $\overline{H}_\alpha(f;t) = k(\alpha)$

2. $\overline{H}_\beta(f;t) = k(\beta)$,

where $k(\alpha)$ and $k(\beta)$ are functions of α and β respectively which are independent of t, if and only if F has a Type 3 extreme value distribution.

Proof. From Asha and Rejeesh (2007) the Type 3 extreme value distribution is characterized by

$$\lambda(t) = c, t \in (-\infty, b)\,.$$

Substituting the last equation in (5.16) and (5.17) we obtain $\overline{H}_\alpha(f;t) = k(\alpha)$ and $\overline{H}_\beta(f;t) = k(\beta)$.
Hence the result. □

Theorem 5.10. *Let X be an absolutely continuous random variable. Then X has the power function distribution if and only if*

$$\overline{H}_\alpha(f;tt') = \overline{H}_\alpha(f;t) + \log t',$$

for t, $t' \in (0,b)$.

Proof. Let us assume that X has the power function distribution given by

$$F(x) = \left(\tfrac{t}{b}\right)^c, t \in (0,b) \text{ and } c > 0.$$

Then, from (5.17), we have

$$\overline{H}_\alpha(f;t) = \frac{1}{1-\alpha}\log\left(\frac{c}{\alpha c - \alpha + 1}\right) - \log\left(\frac{c}{t}\right),$$

which gives

$$\begin{aligned}\overline{H}_{\alpha}(f;tt') &= \frac{1}{1-\alpha}\log\left(\frac{c}{\alpha c-\alpha+1}\right)-\log\left(\frac{c}{tt'}\right)\\ &= \frac{1}{1-\alpha}\log\left(\frac{c}{\alpha c-\alpha+1}\right)-\log\left(\frac{c}{t}\right)+\log t'\\ &= \overline{H}_{\alpha}(f;t)+\log t';t,t'\in(0,b).\end{aligned}$$

Conversely, let

$$\overline{H}_{\alpha}(f;tt') = \overline{H}_{\alpha}(f;t)+\log t';t,t'\in(0,b)$$

Then,

$$-\overline{H}_{\alpha}(f;t) = -\overline{H}_{\alpha}(f;tt')+\log t'$$

which gives,

$$e^{-\overline{H}_{\alpha}(f;t)} = t'e^{-\overline{H}_{\alpha}(f;tt')}.$$

Using Nanda and Paul (2006b), the above equation becomes

$$\lambda\left(t\right) = t'\lambda\left(tt'\right)$$

or,

$$\lambda\left(tt'\right) = \frac{\lambda\left(t\right)}{t'}.$$

The solution to the above equation is $\lambda\left(t\right)=\frac{c}{t}$, $c>0$ which characterizes the power function distribution. □

In the next theorems, we discuss certain characterizations of the GRLMP class using the relationship between the β and α order generalized past entropies and reversed hazard rate $\lambda\left(t\right)$.

Theorem 5.11. *The generalized past entropy of order β for a random variable X is of* *the form*

$$\overline{H}_{\beta}(f;t) = \frac{1}{\beta-1}\left[1-\frac{1}{\beta}\lambda^{\beta-1}(t)\right], \tag{5.18}$$

if and only X has the Type 3 extreme value distribution (2.5).

Proof. Let X have the Type 3 extreme value distribution as in (2.5),

$$F(x) = e^{c(t-b)},\ t \in (-\infty, b) \text{ and } c > 0.$$

Then the reversed hazard rate is

$$\lambda(t) = c.$$

Then using (5.16),

$$\begin{aligned}\overline{H}_{\beta}(f;t) &= \frac{1}{\beta-1}\left[1-\frac{c^{\beta-1}}{\beta}\right]\\ &= \frac{1}{\beta-1}\left[1-\frac{1}{\beta}\lambda^{\beta-1}(t)\right].\end{aligned}$$

Conversely, let (5.18) holds. Then,

$$\frac{\int_{g^{-1}(-\infty)}^{t} f^{\beta}(x)dx}{F^{\beta}(t)} = \frac{1}{\beta}\lambda^{\beta-1}(t).$$

Differentiating the above equation with respect to t on both sides, we get, after simplification

$$\lambda^{\beta}(t) - \beta\lambda(t)\frac{\lambda^{\beta-1}(t)}{\beta} = \frac{1}{\beta}\left(\beta-1\right)\lambda^{\beta-2}(t)\frac{d}{dt}\lambda(t).$$

Or,

$$\lambda^{-2}(t)\frac{d}{dt}\lambda(t) = \lambda^{\beta}(t) - \lambda^{\beta}(t) = 0$$

which implies,

$$\frac{d}{dt}\lambda(t) = 0,$$

which in turn implies $\lambda(t) = c$, a constant. Since $\lambda(t)$ uniquely determines the underlying distribution it follows that X has the Type 3 extreme value distribution (Asha and Rejeesh (2007)). □

We have a similar result for generalized past entropy of order α.

Theorem 5.12. *The generalized past entropy of order α for a random variable X is of the form*

$$\overline{H}_{\alpha}(f;t) = \frac{1}{1-\alpha}\log\left[\frac{1}{\alpha}\lambda^{\alpha-1}(t)\right], \tag{5.19}$$

if and only if X has the Type 3 extreme value distribution (2.5).

From the above theorems also we can arrive at Corollary 5.1

Remark 5.2. *Since the mean inactivity time $m(t) = E(t-X|X \leq t)$ is related to $\lambda(t)$ as*

$$\lambda(t) = \frac{1-m'(t)}{m(t)},$$

for the Type 3 extreme value distribution the mean inactivity time is given by $m(t) = \frac{1}{c}$. Thus, from (i) of the above corollary we have.

$$\overline{H}_{\alpha}(f;t) = k(\alpha) + \log m(t).$$

In the following theorems, we characterize $\overline{H}_{\alpha}(f;t)$ for the power function distribution in terms of reversed hazard rate and mean inactivity time.

Theorem 5.13. *The generalized past entropy measure of order α is of the form*

$$\overline{H}_{\alpha}(f;t) = k(\alpha) - \log\lambda(t)$$

if and only if F has a power function distribution.

Proof. Assume that F has a power function distribution given by

$$F(x) = \left(\tfrac{t}{b}\right)^{c},\ t \in (0,b) \text{ and } c > 0.$$

Then using (5.17), we get

$$\begin{aligned}\overline{H}_\alpha(f;t) &= \frac{1}{1-\alpha}\log\left(\frac{c}{\alpha c-\alpha+1}\right)-\log\left(\frac{c}{t}\right)\\ &= k(\alpha)-\log\lambda(t),\end{aligned}$$

where $k(\alpha)=\frac{1}{1-\alpha}\log\left(\frac{c}{\alpha c-\alpha+1}\right)$.

Conversely, let

$$\overline{H}_\alpha(f;t)=k(\alpha)-\log\lambda(t).$$

Then,

$$\int_{-\infty}^{t}\frac{f^\alpha(x)}{F^\alpha(t)}dx=a\lambda^{\alpha-1}(t), \text{ where } a=e^{k(\alpha)[1-\alpha]}$$

which implies,

$$\int_{-\infty}^{t}f^\alpha(x)dx=aF(t)f^{\alpha-1}(t).$$

Differentiating both sides with respect to t, we get

$$f^\alpha(t)=aF(t)\left(\alpha-1\right)f^{\alpha-2}(t)f'(t)+af^\alpha(t)$$

which gives,

$$1-a=a\left(\alpha-1\right)F(t)\frac{f'(t)}{f^2(t)}.$$

From Kundu et al. (2010), we have

$$\lambda'(t)=\frac{f'(t)}{F(t)}-\lambda^2(t).$$

Thus,

$$\begin{aligned}1-a &= \frac{a\left(\alpha-1\right)F(t)}{f^2(t)}\left[F(t)\lambda'(t)+F(t)\lambda^2(t)\right]\\ &= \frac{a\left(\alpha-1\right)F^2(t)\lambda'(t)}{f^2(t)}+\frac{a\left(\alpha-1\right)F^2(t)\lambda^2(t)}{f^2(t)}\end{aligned}$$

or,

$$1 - a = a\,(\alpha - 1)\,\frac{\lambda'(t)}{\lambda^2(t)} + a\,(\alpha - 1)$$

which in turn implies that

$$\frac{\lambda'(t)}{\lambda^2(t)} = \frac{1 - a\alpha}{a\,(\alpha - 1)}.$$

Solving this differential equation yields,

$$\lambda(t) = \frac{c}{t}$$

where $c = \frac{a(\alpha-1)}{a\alpha-1}$, which is the reversed hazard rate of power function distribution.

This completes the proof. □

Theorem 5.14. *The generalized past entropy measure of order α is of the form*

$$\overline{H}_\alpha(f;t) = b(\alpha) + \log m(t),$$

where $b(\alpha)$ is a function of α if and only if F has a power function distribution.

Proof. It is easily seen that the mean inactivity time of X having a power function distribution is

$$m(t) = \frac{t}{(c+1)}.$$

Assume that F has the power function distribution. Then,

$$\begin{aligned}\overline{H}_\alpha(f;t) &= \frac{1}{1-\alpha}\log\left[\frac{c^\alpha}{(\alpha c - \alpha + 1)}\right] + \frac{1}{1-\alpha}\log\left[\left(\frac{t}{c+1}\right)^{1-\alpha}(c+1)^{1-\alpha}\right] \\ &= \frac{1}{1-\alpha}\log\left(\frac{c^\alpha\,(c+1)^{1-\alpha}}{\alpha c - \alpha + 1}\right) + \log\left(\frac{t}{c+1}\right) \\ &= b(\alpha) + \log m(t),\end{aligned}$$

where $b(\alpha) = \frac{1}{1-\alpha}\log\left(\frac{c^\alpha(c+1)^{1-\alpha}}{\alpha c-\alpha+1}\right)$.

Conversely, assume

$$\overline{H}_\alpha(f;t) = b(\alpha) + \log m(t).$$

Then, $\int_{-\infty}^{t} \frac{f^\alpha(x)}{F^\alpha(t)}dx = km^{1-\alpha}(t)$, where $k = e^{b(\alpha)[1-\alpha]}$

or,

$$\int_{-\infty}^{t} f^\alpha(x)dx = kF^{2\alpha-1}(t)\left[\int_{-\infty}^{t} F(x)dx\right]^{1-\alpha}.$$

Differentiating both sides with respect to t, we get

$$f^\alpha(t) = kF^{2\alpha-1}(t)\,(1-\alpha)\left[\int_{-\infty}^{t} F(x)dx\right]^{-\alpha} F(t) + k\,(2\alpha-1)\,F^{2\alpha-2}(t)f(t)\left[\int_{-\infty}^{t} F(x)dx\right]^{1-\alpha}$$

or,

$$\left[f(t)\int_{-\infty}^{t} F(x)dx\right]^\alpha = k\,(2\alpha-1)\,f(t)F^{2(\alpha-1)}(t)\int_{-\infty}^{t} F(x)dx + (1-\alpha)\,kF^{2\alpha}(t)$$

which gives

$$\left[\frac{f(t)\int_{-\infty}^{t} F(x)dx}{F^2(t)}\right]^\alpha = k\,(2\alpha-1)\,\frac{f(t)}{F(t)}\frac{1}{F(t)}\int_{-\infty}^{t} F(x)dx + (1-\alpha)\,k$$

or,

$$\lambda^\alpha(t)m^\alpha(t) = k\,(2\alpha-1)\,\lambda(t)m(t) + k\,(1-\alpha)\,.$$

By using the relation,

$$\lambda(t) = \frac{1-m'(t)}{m(t)}$$

we get

$$[1-m'(t)]^\alpha = k\,(2\alpha-1)\,[1-m'(t)] + k\,(1-\alpha)\,.$$

If the second derivative of $m(t)$ exists, then we get

$$\alpha\left[1-m'(t)\right]^{\alpha-1}m''(t)=k\left(2\alpha-1\right)m''(t)$$

which gives,

$$m'(t)\text{= constant.}$$

Hence, the mean inactivity time of X is linear which is the required result. □

As $\beta\to1$ and $\alpha\to1$, the β and α- order generalized past entropies reduces to the past entropy given by Kundu et al. (2010) which is defined as follows:

Let X be an absolutely continuous random variable with distribution function F in the support of (a,b), where $a=\inf\{t:F(t)>0\}$ and $b=\sup\{t:F(t)<1\}$, $-\infty\leq a<t\leq b<\infty$. Then the measure of uncertainty for inactivity time or past time distribution called past entropy is defined as

$$\begin{aligned}\overline{H}(X;t)&=-\int_a^t\frac{f(x)}{F(t)}\ln\left(\frac{f(x)}{F(t)}\right)dx\\&=1-\frac{1}{F(t)}\int_a^t f(x)\log\lambda(x)dx.\end{aligned}\tag{5.20}$$

In the following theorem, we characterize the GRLMP class in terms of past entropy and the reversed hazard rate.

Theorem 5.15. *Let X be an absolutely continuous random variable with distribution function $F(t)$ and reversed hazard rate $\lambda(t)$. Then X belongs to the GRLMP class if and only if the following equation holds.*

$$\overline{H}(f;t)+\log\lambda(t)=1+\frac{1}{F(t)}\int_{g^{-1}(-\infty)}^t\left(\frac{g''(x)}{g'(x)}\right)F(x)dx.$$

Proof. Assume that X belongs to the GRLMP class. Then,

$$F(x)=e^{c(g(t)-g(b))},t\in\left(g^{-1}(-\infty),b\right),c>0$$

with

$$\lambda\,(t) = cg'(t).$$

Now, from (5.20), we have

$$\begin{aligned} 1 - \overline{H}(f;t) &= \frac{1}{F(t)} \int_a^t f(x) \log \lambda(x) dx \\ &= \frac{1}{F(t)} \left\{ \log \lambda(t) F(t) - \int_a^t \frac{\lambda'(x)}{\lambda(x)} F(x) dx \right\} \\ &= \log \lambda(t) - \frac{1}{F(t)} \int_a^t \frac{\lambda'(x)}{\lambda(x)} F(x) dx. \end{aligned}$$

Thus,

$$\overline{H}(f;t) + \log \lambda(t) = 1 + \frac{1}{F(t)} \int_a^t \frac{\lambda'(x)}{\lambda(x)} F(x) dx.$$

Since X belongs to the GRLMP class, the above equation becomes

$$\overline{H}(f;t) + \log \lambda(t) = 1 + \frac{1}{F(t)} \int_{g^{-1}(-\infty)}^t \left(\frac{g''(x)}{g'(x)} \right) F(x) dx.$$

Conversely, suppose that

$$\overline{H}(f;t) + \log \lambda(t) = 1 + \frac{1}{F(t)} \int_{g^{-1}(-\infty)}^t \left(\frac{g''(x)}{g'(x)} \right) F(x) dx.$$

Differentiating with respect to t, we get

$$\begin{aligned} \overline{H}'(f;t) + \frac{\lambda'(t)}{\lambda(t)} &= \frac{F(t) \frac{g''(t)}{g'(t)} F(t) - f(t) \int_{g^{-1}(-\infty)}^t \left(\frac{g''(x)}{g'(x)} \right) F(x) dx}{F^2(t)} \\ &= \frac{g''(t)}{g'(t)} - \frac{\lambda(t)}{F(t)} \int_{g^{-1}(-\infty)}^t \left(\frac{g''(x)}{g'(x)} \right) F(x) dx \end{aligned}$$

or,

$$\overline{H}'(f;t) = \frac{g''(t)}{g'(t)} - \frac{\lambda(t)}{F(t)} \int_{g^{-1}(-\infty)}^t \left(\frac{g''(x)}{g'(x)} \right) F(x) dx - \frac{\lambda'(t)}{\lambda(t)}.$$

From Kundu et al. (2010), we have

$$\overline{H}'(f;t) = \lambda(t)\left[1 - \log\lambda(t) - \overline{H}(f;t)\right].$$

Then,

$$\lambda(t) - \lambda(t)\log\lambda(t) - \lambda(t) - \frac{\lambda(t)}{F(t)}\int_{g^{-1}(-\infty)}^{t}\left(\frac{g''(x)}{g'(x)}\right)F(x)dx + \lambda(t)\log\lambda(t)$$
$$= \frac{g''(t)}{g'(t)} - \frac{\lambda(t)}{F(t)}\int_{g^{-1}(-\infty)}^{t}\left(\frac{g''(x)}{g'(x)}\right)F(x)dx - \frac{\lambda'(t)}{\lambda(t)}$$

which gives,

$$\frac{cg''(t)}{cg'(t)} = \frac{\lambda'(t)}{\lambda(t)}$$

or,

$$\lambda(t) = cg'(t).$$

Then, from Block et al. (1998),

$$F(x) = e^{c(g(t)-g(b))}, t \in \left(g^{-1}(-\infty), b\right), c > 0.$$

Hence the theorem. □

Theorem 5.15 enables us to generalize the results obtained in Kundu et al. (2010) which are given in Table 5.1.

It should be noted that the theorems through 5.8 to 5.14 holds for $\overline{H}(X;t)$ as α and $\beta \to 1$ reduces to Theorem 5.15. In particular we have the following.

Corollary 5.2. *Let X be an absolutely continuous random variable. Then X has Type 3 extreme value distribution if and only if*

$$\overline{H}(f;t+t') = \overline{H}(f;t); t, t' \in (-\infty, b)$$

Table 5.1:

Characterizations (Kundu et al.(2010))	$g(t)$	$\lambda(t) = cg'(t)$	Distribution
$\overline{H}(f;t) = 1 - \log \lambda(t)$	t	c	Type 3 extreme value distribution $F(t) = \exp\{c(t-b)\}$, $t \in (-\infty, b)$, $c > 0$.
$\overline{H}(f;t) = \frac{c-1}{c} - \log \lambda(t)$	$\log t$	$\frac{c}{t}$	Power function Distribution $F(t) = \left(\frac{t}{b}\right)^c$, $t \in (0, b)$, $c > 0$.
$\overline{H}(f;t) = \frac{c+1}{c} - \log \lambda(t)$	$-\log(1-t)$	$\log(1-t)^{-c}$	Negative Pareto Distribution $F(t) = (1-t)^{-c}$, $t \in (-\infty, 0)$, $c > 1$.
$\overline{H}(f;t) = 1 - \log \lambda(t) - m(t)$	$-e^{-t}$	ce^{-t}	Truncated extreme value distribution (when $c = 1$) $F(t) = \exp\left\{e^{-b} - e^{-t}\right\}$, $t \in (-\infty, b)$.

Corollary 5.3. *Let X be an absolutely continuous random variable. Then X has the power function distribution if and only if*

$$\overline{H}(f;tt') = \overline{H}(f;t) + \log t'; t, t' \in (0, b).$$

Corollary 5.4. *Let X be an absolutely continuous random variable. Then X has the negative Pareto function distribution if and only if*

$$\overline{H}(f;t + t' - tt') = \overline{H}(f;t) + \log(1 - t'); t, t' \in (-\infty, 0).$$

Chapter 6

Reversed Lack of Memory Property in Discrete Domain

6.1 Introduction

The results in this chapter can be considered as an addendum to those discussed earlier. All the same conceptual definitions in the discrete domains need to be addressed with clarity. They are often overlooked or ignored in literature. However they find application when considering group data, the underlying modeling assumption being that of a continuous time distribution. Discrete models and their study has been conventionally restricted to applications where failure process involves discrete trials. But discrete time models are worth greater consideration. A number of problems which occur with continuous time models are overcome by using a discrete time model. Further discussion on discrete concepts is found in Xekalaki (1983), Nair and Hitha (1989), Adams and Watson (1989), Roy and Gupta (1992), Nanda and Sengupta (2005) and the references therein.

Nanda and Paul (2006a) defined the past entropy for the discrete random variable X with probability mass function $P\left[X=i\right]=p(i)$, $i=n,n+1,...,m$ where n could be $-\infty$ and $0<m<\infty$. Now let $P(k)=P\left[X\leq k\right]=\sum_{i=n}^{k}p(i)$, the distribution function, $\delta(j) = \ln \frac{P(j)}{P(j-1)}$, the reversed hazard rate (Dewan and Sudheesh (2009)) and $m(t) = \frac{1}{P(t)} \sum_{i=n}^{t} P(i)$, the mean inactivity time (Goloforushani and Asadi (2008)) of X. The discrete past entropy of X is now defined as

$$\overline{H}(X;j) = -\sum_{k=n}^{j} \frac{p(k)}{P(j)} \ln \frac{p(k)}{P(j)}.$$

Note that, as $j \to m$, $\overline{H}(X;j)$ becomes the Shannon entropy given by $\overline{H}(X) = -\sum_k p(k) \ln p(k)$.

The discrete generalized past entropy (Nanda and Paul (2006b)) is defined as

$$\overline{H}^{\beta}(X;j) = \frac{1}{1-\beta} \ln \left[\sum_{k=n}^{j} \left(\frac{p(k)}{P(j)} \right)^{\beta} \right].$$

As in the continuous case, when $\beta \to 1$, $\overline{H}^{\beta}(X;j) \to \overline{H}(X;j)$.

However these measures which consider only probabilities and ignore the value of the random variable takes may in some situations not do justice to the literature and practical notions of randomness or information. For any discrete random variable X, the entropy is computed solely using the probabilities $P[X = t]$ and one interprets the entropy as a measure of the randomness in X. However if X denotes the number of components that has failed in a complex system, or time for the next scheduled maintenance of a system or for that matter of any left truncated data, the appropriate probabilities to be considered are $P[X \leq t]$ instead of $P[X = t]$. Accordingly we define the cumulative past entropy as

$$\begin{aligned} H(X;j) &= -\sum_{k=n}^{j} \frac{P(k)}{P(j)} \ln \frac{P(k)}{P(j)} \\ &= -\sum_{k=n}^{j} \sum_{i=k}^{j-1} \left(\prod_{i=k}^{j-1} \frac{P(i)}{P(i+1)} \right) \ln \frac{P(i)}{P(i+1)} \\ &= \sum_{k=n}^{j} \sum_{i=k}^{j-1} \delta(i) e^{-\sum_{i=k}^{j-1} \delta(i)} \end{aligned}$$

Observe that

$$H(X;j) - H(X;j-1) = \sum_{i=k}^{j-1} \delta(i) e^{-\sum_{i=k}^{j-1}\delta(i)}$$
$$= -P(j-1)\ln P(j-1).$$

It hence follows that $H(X;j)$ is increasing in j and also that $H(X;j)$ takes values in $(0, \infty)$. This measure is the discrete analogue of the measure in (5.9) and hence the appropriateness of this measure follow in a similar argument as for (5.9). Here we study the behaviour of these discrete measures when X has the discrete reversed lack of memory property.

In this chapter we discuss the notion of reversed lack of memory property (RLMP) in the discrete set up and derive the discrete RLMP class. Similar to the continuous case the discrete reversed lack of memory property has its implication on many probability and uncertainty measures. It is further shown that the class of distributions defined by Goliforushani and Asadi (2008) and Kundu et al. (2010) is the RLMP class. All these results are discussed in Section 6.2. In Section 6.3 we consider the bivariate analogue of the reversed lack of memory property and its implications on bivariate reliability measures. The multivariate analogue of this property is included in Secton 6.4. Finally the thesis concludes by discussing future work to be carried out.

6.2 Discrete Reversed Lack of Memory Property

Definition 6.1. *Let X be as in Section 6.1. Then X is said to possess the reversed lack of memory property (RLMP), if for all $t = 0, 1, ..., m$ with $x + t \leq m$,*

$$P(X \leq x | X \leq x+t) = P(X \leq 0 | X \leq t) \tag{6.1}$$

To interpret RLMP physically, let X represent the number of cycles of operation of equipment (interpreted as age) before it fails. Then the right side represents the probability that new equipment fails before it completes the first cycle given that it fails within t cycles. On the left side, the probability of the same equipment fails before x cycles for any $x \in \{n, n+1, ..., m\}$ and $t = 0, 1, ..., m$ with $x + t \leq m$ given that it fails within an additional t cycles. Thus the expected time elapsed since failure is independent of the age of the equipment whenever RLMP is satisfied.

The next theorem characterizes models with RLMP.

Theorem 6.1. *Let X be a discrete random variable defined as above with $m < \infty$. Then the following statements are equivalent.*

(i) X has a distribution specified by

$$P(x) = p^{(m-x)}, x \in \{n, n+1, ..., m\}, 0 < p < 1 \tag{6.2}$$

(ii) X has the reversed lack of memory property (6.1) *which is equivalent to*

$$P(x)P(t) = P(x+t)P(0), \tag{6.3}$$

for $x + t \leq m$.

(iii) The reversed hazard rate of X satisfies

$$\delta(x+t) = \delta(x), x \in \{n, n+1, ..., m\}. \tag{6.4}$$

(iv) The discrete generalized past entropy of X is satisfied by

$$\overline{H}^{\beta}(X; j+t) = \overline{H}^{\beta}(X; j).$$

(v) The mean past lifetime of X is satisfied by

$$m(x+t) = m(x).$$

(vi) The discrete cumulative past entropy of X is satisfied by

$$H(X; j+t) = H(X; j).$$

Proof. We first prove the equivalence of (i) and (ii). That of $(i) \Rightarrow (ii)$ is evident by substituting for $P(x)$ in (6.3) from (6.2).

Assuming,

$$P(x)P(t) = P(x+t)P(0),$$

for all $x \in \{n, n+1, ..., m\}$ with $x + t \leq m$, we have

$$P(x) = \frac{P(0)}{P(1)} P(x+1), x \in \{n, n+1, ..., m\}.$$

Or,

$$P(x) = \left[\frac{P(0)}{P(1)}\right]^{m-x}.$$

Hence,

$$P(x) = p^{(m-x)}, \text{ where } 0 < p = \tfrac{P_0}{P_1} < 1.$$

Thus by the definition, $\delta(x)$ is a constant and hence (6.4) holds. Thus $(ii) \Rightarrow (iii)$ and since $\delta(x)$ uniquely determines the distribution function by the relation (1.26) it follows that $(iii) \Rightarrow (i)$.

The equivalence between (i) and (iv) follows by Theorem 4.3 of Kundu et al. (2010).

By assuming (i) and using $m(t) = \frac{1}{P(t)} \sum_{i=n}^{t} P(i)$, we see that the mean past lifetime is a constant. Conversely, by assuming (v) and applying Theorem 2.6 of

Goliforushani and Asadi (2008), we get (i).

Thus, the proof is complete. □

Corollary 6.1. $\overline{H}^{\beta}(X;j) = c$, *where* c *is a constant if and only if* $P(x)$ *is of the form* (6.2).

Corollary 6.2. $H(X;j) = c$, *where* c *is a constant if and only if* $P(x)$ *is of the form* (6.2).

Corollary 6.3. $\delta(x) = c$, *where* c *is a constant if and only if* $P(x)$ *is of the form* (6.2).

Corollary 6.4. $m(x) = c$, *where* c *is a constant if and only if* $P(x)$ *is of the form* (6.2).

From Kundu et al. (2010), we have

$$e^{\delta(j)} = \frac{m(j-1)}{m(j)-1}. \tag{6.5}$$

Hence, by using the results in Nanda and Paul (2006b) we get

$$\overline{H}^{\beta}(X;j-1) = \log\left\{\frac{m(j)-1}{1+m(j-1)-m(j)}\right\}.$$

In fact $\delta(x)$ uniquely determines the distribution through $P(x) = \exp\left(-\sum_{y=x+1}^{m}\delta(y)\right)$, Corollary 6.3 can be generalized to get the following result.

Theorem 6.2. *Let* X *be a random variable defined as above. Then the reversed hazard rate* $\delta(x)$ *is of the form*

$$\delta(x) = lx + k, l, k > 0 \tag{6.6}$$

for all $x \in \{n, n+1, ..., m\}$ *if and only if* X *is distributed as*

$$P(x) = p_1^{(m-x)(m+x+1)} p_2^{(m-x)}, x \in \{n, n+1, ..., m\}$$

for some $0 < p_1, p_2 < 1$, *where* $p_1 = e^{-l/2}$ *and* $p_2 = e^{-k}$.

Proof. Let $\delta(x)$ be as in (6.6). Then the definition

$$\ln \frac{P(x)}{P(x-1)} = lx + k$$

implies

$$\begin{aligned} P(x) &= \exp\left\{-\sum_{y=x+1}^{m}(lx+k)\right\} \\ &= p_1^{\sum_{y=x+1}^{m} y} p_2^{m-x} \\ &= p_1^{(m-x)(m+x+1)} p_2^{(m-x)}. \end{aligned}$$

The constants l and k in (6.6) are determined as $l = -2\log p_1$ and $k = -\log p_2$. The converse is straight forward. □

In the next section we extend the concept of reversed lack of memory property to the bivariate case.

6.3 Bivariate Characterizations

In this section, we extend the idea of discrete reversed lack of memory property to the bivariate case and investigate its implication on bivariate reliability measures like bivariate reversed hazard rate and bivariate mean past lifetime.

Definition 6.2. *Let $X = (X_1, X_2)$ be random vector with distribution function $P(x_1, x_2)$ defined on $I_m \times I_m$, where $I_m = \{n, n+1, n+2, ..., m\}$ where the integer n could be $-\infty$, but m is finite and positive. Then the bivariate reversed hazard rate is defined as*

$$\underline{\delta}(\underline{x}) = (\delta_1(\underline{x}), \delta_2(\underline{x})) \tag{6.7}$$

where $\delta_i(\underline{x}) = \ln \frac{P(\underline{x})}{P(\underline{x}_i)}$, $P(\underline{x}) = P[X_1 \le x_1, X_2 \le x_2]$, $\underline{x}_i = (x_i - 1, x_{3-i})$, $i = 1, 2$ and $\underline{x} = (x_1, x_2) \in I_m \times I_m$.

For $x_i = m$, it follows from (6.7) that $\delta_i(m, x_{3-i}) = \ln \frac{P(x_{3-i})}{P(x_{3-i}-1)}$, the univariate reversed hazard rate (Asha and Nair (2004)).

Also,

$$P(x_1, x_2) = \exp\left[-\sum_{y=x_1+1}^{m} \delta_1(y, x_2) - \sum_{y=x_2+1}^{m} \delta_2(m, y)\right] \tag{6.8}$$

or,

$$P(x_1, x_2) = \exp\left[-\sum_{y=x_2+1}^{m} \delta_2(x_1, y) - \sum_{y=x_1+1}^{m} \delta_1(y, m)\right] \tag{6.9}$$

The bivariate reversed mean residual life function defined by $\underline{\nu}(\underline{x}) = (\nu_1(\underline{x}), \nu_2(\underline{x}))$ where

$$\nu_1(\underline{x}) = E\left(X_1 - x_1 | X_1 \leq x_1, X_2 \leq x_2\right) \tag{6.10}$$

$$= \frac{1}{P(x_1, x_2)} \sum_{t=n}^{x_1} P(t, x_2) \tag{6.11}$$

and

$$\nu_2(\underline{x}) = \frac{1}{P(x_1, x_2)} \sum_{t=n}^{x_2} P(x_1, t) \tag{6.12}$$

is related to $\underline{\delta}(\underline{x})$ as

$$\exp\left[-\delta_1(x_1, x_2)\right] = \frac{\nu_1(x_1, x_2) - 1}{\nu_1(x_1 - 1, x_2)} \tag{6.13}$$

and

$$\exp\left[-\delta_2(x_1, x_2)\right] = \frac{\nu_2(x_1, x_2) - 1}{\nu_2(x_1, x_2 - 1)} \tag{6.14}$$

6.3.1 Characterizations

Characterization of $\delta_i(x_i)$ has been found in Asha and Nair (2004). A problem of natural interest is how far these characterizations can be extended to higher dimensions. In this section we look into few characterizations of $\underline{\delta}(x_1, x_2)$ and their functional form.

Theorem 6.3. *Let $X = (X_1, X_2)$ be a bivariate random vector defined as above with $m < \infty$. Then the following statements are equivalent.*

(i)

$$\underline{\delta}(\underline{x}+\underline{t})=\underline{\delta}(\underline{x}), \tag{6.15}$$

where $\underline{t}=(t_1,t_2)$.

(ii)

$$P(x_1+t_1,x_2+t_2)P(0,0)=P(x_1,x_2)P(t_1,t_2), \tag{6.16}$$

for all x_i *and* t_i *such that* $x_i+t_i\leq m$, $i=1,2$.

(iii) X *has a distribution specified by*

$$P(x_1,x_2)=p_1^{m-x_1}p_2^{m-x_2}; x_1,x_2=n,n+1,n+2,...,m \tag{6.17}$$

where $p_1=\frac{P(0,0)}{P(1,0)}$ *and* $p_2=\frac{P(0,0)}{P(0,1)}$.

Proof. Assume $\underline{\delta}(\underline{x}+\underline{t})=\underline{\delta}(\underline{x})$.

This implies $\underline{\delta}(\underline{x})$= constant. Then by using (6.8) we get (6.17). The converse is straightforward.

To prove the equivalence of (i) and (ii), first we assume that X is distributed as in (6.17). Then

$$P[X_1\leq x_1+t_1,X_2\leq x_2+t_2]=\frac{p_1^{m-x_1}p_2^{m-x_2}p_1^{m-t_1}p_2^{m-t_2}}{p_1^m p_2^m} \tag{6.18}$$

or,

$$P(x_1+t_1,x_2+t_2)=\frac{P(x_1,x_2).P(t_1,t_2)}{P(0,0)}.$$

Thus, (i) implies (ii).

Conversely, let (ii) is satisfied.

Put $x_2=0$ and $t_2=m$ in (6.16), we have

$$P_1(x_1+t_1)P(0,0)=P(x_1,0)P_1(t_1) \tag{6.19}$$

which implies

$$\frac{P(x_1,0)}{P(0,0)} = \frac{P_1(x_1)}{P_1(0)} \tag{6.20}$$

Similarly, $x_1 = 0$ and $t_1 = m$ in (6.16) gives

$$P_2(x_2+t_2)P(0,0) = P(0,x_2)P_2(t_2) \tag{6.21}$$

which implies

$$\frac{P(0,x_2)}{P(0,0)} = \frac{P_2(x_2)}{P_2(0)}. \tag{6.22}$$

Again put $t_2 = 0$ and $x_2 = m$ in (6.16) and using (6.21)

$$P_1(x_1+t_1) = \frac{P_1(t_1)}{P_1(0)}P_1(x_1)$$

which implies

$$p_1^{m-(x_1+t_1)}P(0,0) = P(t_1,0)p_1^{m-x_1}.$$

Thus,

$$p_1 = \frac{P(0,0)}{P(1,0)}. \tag{6.23}$$

Similarly,

$$p_2 = \frac{P(0,0)}{P(0,1)}. \tag{6.24}$$

By putting $x_1 = 0$, $t_1 = x_1$, $t_2 = 0$ in (6.16) and using (6.23) we obtain

$$P(x_1,x_2) = P(0,0)p_1^{-x_1}p_2^{-x_2} \tag{6.25}$$

$x_1 = x_2 = m$ in (6.25) implies

$$P(0,0) = p_1^m p_2^m. \tag{6.26}$$

Thus,

$$P(x_1, x_2) = p_1^{m-x_1} p_2^{m-x_2}.$$

Hence (ii) implies (iii). □

Note that the statement (ii) is the analogue of the lack of memory property (LMP). The LMP plays a pivotal role in modeling lifetime data and also forms the basis for many results in theoretical and applied probability. For left censored data we define a corresponding property, the bivariate reversed lack of memory property (BRLMP) as

$$P\left[X_1 \leq x_1, X_2 \leq x_2 | X_1 \leq x_1 + t_1, X_2 \leq x_2 + t_2\right] = P\left[X_1 \leq 0, X_2 \leq 0 | X_1 \leq t_1, X_2 \leq t_2\right] \quad (6.27)$$

for all x_i and t_i such that $x_i + t_i \leq m$, $i = 1, 2$.

If $\underline{x} + \underline{t} = (x_1 + t_1, x_2 + t_2)$ represent the number of cycles of operations of two components before it fails, then the right hand side represents the probability that a new equipment with two components fails before it completes the first cycle given that the components fail before it completes $\underline{t} = (t_1, t_2)$ cycles. Thus the expected time elapsed since failure is independent of the age of the components whenever the reversed lack of memory property is satisfied. The next characterization discuss the condition that $\underline{t} = (t, t)$.

Theorem 6.4. *Let $X = (X_1, X_2)$ be a bivariate random vector defined as above with $m < \infty$. Then the following statements are equivalent.*

(i)

$$\underline{\delta}(\underline{x} + \underline{t}) = \underline{\delta}(\underline{x}) = \begin{cases} \left(\ln \frac{1}{p_1}, \ln \frac{p_1}{p}\right), x_2 > x_1 \\ \left(\ln \frac{1}{p_2}, \ln \frac{p_2}{p}\right), x_1 > x_2 \\ \left(\ln \frac{1}{p}, \ln \frac{1}{p}\right), x_1 = x_2. \end{cases} \quad (6.28)$$

(ii) X satisfies the reversed lack of memory property defined as

$$P(x_1+t,x_2+t)P(0,0)=P(x_1,x_2)P(t,t), \tag{6.29}$$

for all x_i and t such that $x_i+t\leq m$, $i=1,2$.

(iii) X has a distribution specified by

$$P(x_1,x_2)=\begin{cases} p^{m-x_2}p_1^{x_2-x_1}, x_2>x_1 \\ p^{m-x_1}p_2^{x_1-x_2}, x_1>x_2 \\ p^{m-x}, x_1=x_2 \end{cases} \tag{6.30}$$

where $p=\frac{P(0,0)}{P(1,1)}$, $p_1=\frac{P(0,0)}{P(1,0)}$ and $p_2=\frac{P(0,0)}{P(0,1)}$.

Proof. To prove the equivalence of (i) and (iii), first assume that (i) holds. Then by using (6.8) or (6.9), we have (6.30). The converse is straightforward.

To prove equivalence of (i) and (ii), we assume that (ii) holds.

Put $x_1=x_2=x$ in (6.29) implies

$$\begin{aligned} P(x,x) &= \frac{P(1,1)}{P(0,0)}P(x+1,x+1) \\ &= p^{-x}P(0,0), \end{aligned} \tag{6.31}$$

where $p=\frac{P(0,0)}{P(1,1)}$.

Hence,

$$P(x,x)=p^{m-x} \tag{6.32}$$

with $P(0,0)=p^m$.

Then (6.32) in (6.29) implies

$$P(x_1+t,x_2+t)p^t=P(x_1,x_2) \tag{6.33}$$

Taking$x_2 + t = m$, we get

$$P(x_1, x_2) = p^{m-x_2} p_1^{x_2-x_1}. \tag{6.34}$$

Similarly when $x_1 + t = m$, we have

$$P(x_1, x_2) = p^{m-x_1} p_2^{x_1-x_2}. \tag{6.35}$$

Thus (ii) implies (iii).

The converse is straightforward. □

We can define the bivariate reversed local lack of memory property as

$$P\left[X_1 \leq x_1, X_2 \leq x_2 | X_i \leq x_i + t\right] = P\left[X_j \leq x_j | X_i \leq t\right], \tag{6.36}$$

for all x_i and t such that $x_i + t \leq m$, $i, j = 1, 2$, $i \neq j$

However, if the expected time is a local constant with respect to age we have the following characterization.

Theorem 6.5. *Let* $X = (X_1, X_2)$ *be a random vector defined as above with* $m < \infty$. *Then*

$$\underline{\delta}(\underline{x}) = (lx_2 + k, sx_1 + t) \tag{6.37}$$

if and only if X *is distributed as*

$$P(x_1, x_2) = p^{(m-x_1)(m-x_2)} p_1^{m-x_1} p_2^{m-x_2}, \tag{6.38}$$

for all $x_1, x_2 = n, n+1, n+2, ..., m$ *where* $0 < p = e^{-\theta} < 1$, $\theta > 0$, $0 < p_1 = e^{-A(m)} < 1$, $0 < p_2 = e^{-B(m)} < 1$, $A(m)$ *and* $B(m)$ *are constants and* $\frac{1-p^m p_2}{p^m p_1} \geq 1 - p^{m+1} p_2$.

Proof. Let $\underline{\delta}(\underline{x})$ is of the form (6.37). Then from (6.8) and (6.9), we have

$$\begin{aligned}&\exp\left\{-\left[(lx_2+k)(m-x_1)+B(m)(m-x_2)\right]\right\}=\\&\qquad\qquad\exp\left\{-\left[(sx_1+t)(m-x_2)+A(m)(m-x_1)\right]\right\}\end{aligned} \tag{6.39}$$

which implies

$$\frac{lx_2+k-A(m)}{m-x_2}=\frac{sx_1+t-B(m)}{m-x_1}=\theta, \tag{6.40}$$

where θ is a constant.

Then,

$$\begin{aligned}P(x_1,x_2)&=\exp\left\{-\theta(m-x_1)(m-x_2)-A(m)(m-x_1)-B(m)(m-x_2)\right\}\\&=p^{(m-x_1)(m-x_2)}p_1^{m-x_1}p_2^{m-x_2}\end{aligned}$$

For $p(0,0)\geq 0$, we need $\frac{1-p^m p_2}{p^m p_1}\geq 1-p^{m+1}p_2$.

The converse is straightforward. □

Corollary 6.5. *Consider the vector of probabilities*

$$\underline{\lambda}(\underline{x})=\left(\lambda_1(\underline{x}),\lambda_2(\underline{x})\right),$$

where

$$\begin{aligned}\lambda_1(x_1,x_2)&=P\left[X_1=x_1|X_1\leq x_1,X_2\leq x_2\right]\\&=1-\frac{P(x_1-1,x_2)}{P(x_1,x_2)}\end{aligned} \tag{6.41}$$

and

$$\begin{aligned}\lambda_2(x_1,x_2)&=P\left[X_2=x_2|X_1\leq x_1,X_2\leq x_2\right]\\&=1-\frac{P(x_1,x_2-1)}{P(x_1,x_2)}.\end{aligned} \tag{6.42}$$

Then $\underline{\lambda}(\underline{x})$ can be viewed as discrete analogue of reversed hazard rate (Bismi (2005)).

Since $\underline{\delta}(\underline{x})$ and $\underline{\lambda}(\underline{x})$ are connected by

$$(\lambda_1(\underline{x}), \lambda_2(\underline{x})) = \left(1 - e^{-\delta_1(\underline{x})}, 1 - e^{-\delta_2(\underline{x})}\right) \tag{6.43}$$

the above three characterizations can be proved using $\underline{\lambda}(\underline{x})$ also.

In the next section we discuss the multivariate analogue of the reversed lack of memory property in the previous sections.

6.4 Multivariate Extension

The multivariate reversed lack of memory property is a straightforward extension of the bivariate case. For completeness, we state the reversed lack of memory property in the multidimensional case and derive the corresponding multivariate distribution characterized by it.

Definition 6.3. *Let $X = (X_1, X_2, ..., X_k)$ be random vector with distribution function $P(x_1, x_2, ..., x_k)$ defined on $I_m \times I_m \times ... \times I_m$, where $I_m = \{n, n+1, n+2, ..., m\}$ where the integer n could be $-\infty$, but m is finite and positive. Then the multivariate reversed lack of memory property is defined as*

$$P\left(\underline{x} + \underline{t}\right) P\left(\underline{0}\right) = P\left(\underline{x}\right) P\left(\underline{t}\right), \tag{6.44}$$

such that $x_i + t_i \leq m_i$, $i = 1, 2, ..., k$ where $\underline{x} = (x_1, x_2, ..., x_k)$, $\underline{t} = (t_1, t_2, ..., t_k)$ and $(0, 0, ..., 0)$.

The distribution satisfying (6.44) is given by

$$P\left(\underline{x}\right) = \prod_{i=1}^{k} p_i^{(m-x_i)}, \tag{6.45}$$

for $x_i = n, n+1, n+2, ..., m$, $i = 1, 2, ..., k$ where $p_i = \frac{P(0,0,...,0)}{P(0,0,...,1,...,0)}$.

A meaningful relaxation of (6.44) is by considering $\underline{t} = (t, t, ..., t)$. Then the multivariate reversed lack of memory property becomes

$$P\left(\underline{x}+\underline{t}\right)P\left(\underline{0}\right)=P\left(\underline{x}\right)P\left(\underline{t}\right), \tag{6.46}$$

such that $x_i + t \leq m_i$, $i = 1, 2, ..., k$ where $\underline{x} = (x_1, x_2, ..., x_k)$, $\underline{t} = (t, t, ..., t)$ and $(0, 0, ..., 0)$.

The distribution satisfying (6.46) is given by

$$P\left(\underline{x}\right)=p_{i_1}^{m-x_{i_1}}\left(\frac{p_{i_1i_2}}{p_{i_1}}\right)^{m-x_{i_2}}\left(\frac{p_{i_1i_2i_3}}{p_{i_1i_2}}\right)^{m-x_{i_3}}...\left(\frac{p_{i_1i_2...i_k}}{p_{i_1i_2...i_{k-1}}}\right)^{m-x_{i_k}}, \tag{6.47}$$

where $x_{i_1} \geq x_{i_2} \geq ... \geq x_{i_k}$ for each permutation $(i_1, i_2, ..., i_k)$ of the integers from 1 to k. The parameters are such that

$$0 < p_{i_1i_2...i_k} \leq ... \leq p_{i_1i_2} \leq p_1, p_2, ..., p_k < 1,$$

$$p_{i_1i_2...i_j} = p_{123...j} \text{ for } j = 2, 3, ..., k \text{ and}$$

$$1-\sum_{i=1}^{k}p_i-\sum_{i<j}\sum p_{ij}+...+(-1)^{k-1}p_{12...k}\geq 0.$$

Let $X = (X_1, X_2, ..., X_k)$ be a random vector and $\underline{t} = (t_1, t_2, ..., t_k)$ defined as above. Then the multivariate reversed local lack of memory property is defined as

$$P\left[X_i \leq x_i | X_1 \leq x_1, X_2 \leq x_2, ..., X_i \leq x_i + t_i, ..., X_k \leq x_k\right]$$
$$= P\left[X_i \leq 0 | X_1 \leq x_1, X_2 \leq x_2, ..., X_i \leq t_i, ..., X_k \leq x_k\right] \tag{6.48}$$

for all x_i and t_i such that $x_i + t_i \leq m$, $i = 1, 2, ..., k$.

The distribution satisfying (6.48) is given by

$$P\left(\underline{x}\right)=\left(\prod_{i=1}^{k}p_i^{(m-x_i)}\right)\left(\prod_{i<j}p_{ij}^{(m-x_i)(m-x_j)}\right)...\left(p_{12...k}^{(m-x_i)(m-x_2)...(m-x_k)}\right),$$

where $0 < p_i < 1$, $0 < p_{ij}, p_{ijl}, ..., p_{12...k} < 1$ and

$$1 - \sum_{i=1}^{k} p_i - \sum_{i<j}\sum p_{ij} + ... + (-1)^{k-1} p_{12...k} \geq 0.$$

6.5 Future Work

In this thesis, the concept of reversed lack of memory property and its generalizations is studied. The implications of this property on various statistical measures are also investigated. A model characterized by this property is also presented. A detailed study on distributional properties of the model characterized by the generalized reversed lack of memory property remains to be looked up on. As a future course of work we also intended to look upon the parametric and nonparametric estimation problems associated with the generalized reversed lack of memory property class based on the characterizations discussed here.

Bibliography

[1] Aczel, J. (1966). *Lectures on Functional Equations and Their Applications.* Academic Press, New York.

[2] Adams, G. and Watson, R. (1989). *A discrete time parametric model for the analysis of failure time data*. Australian Journal of Statistics, 31, 365-384.

[3] Ahmed, A. N. (1991). *Characterization of beta, binomial and Poisson distribution*. IEEE Transactions on Reliability, 40, 290-293.

[4] Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (1993). *Statistical models based on counting processes*. Springer Verlag, New York.

[5] Asadi, M. (2006). *On the mean past lifetime of the components of a parallel system*. Journal of Statistical Planning and Inference, 136 (4), 1197-1206.

[6] Asadi, M. and Ebrahimi, N. (2000). *Residual entropy and its characterizations in terms of hazard function and mean residual life function.* Statistics and Probability Letters, 49, 263-269.

[7] Asadi, M. and Zohrevand, Y. (2007). *On the dynamic cumulative residual entropy.* Journal of Statistical Planning and Inference, 137, 1931-1941.

[8] Asha, G. and Nair, N. U. (2010). *Reliability properties of mean time to failure in age replacement models.* International Journal of Reliability, Quality and Safety Engineering, 17 (1), 1-12.

[9] Asha, G. and **Rejeesh, C. J.** (2007). *Models characterized by the reversed lack of memory property.* Calcutta Statistical Association Bulletin, 65 (233-234), 1-14.

[10] Asha, G. and **Rejeesh, C. J.** (2009). *Characterizations using the generalized reversed lack of memory property*. Statistics and Probability Letters, 79, 1480-1487.

[11] Azlarov, T. A. and Volodin, N. A. (1986). *Characterization Problems Associated with the Exponential Distribution*. Springer Verlag, New York.

[12] Barlow, R. E., Marshall, A. W. and Proshan, F. (1963). *Properties of probability distributions with monotone hazard rate*. The Annals of Mathematical Statistics, 34 (2), 375-389.

[13] Bartoszewicz, J. and Skolimowska, M. (2006). *Preservation of classes of life distributions and stochastic orders under weighting.* Statistics and Probability Letters, 76(6), 587-596.

[14] Basu, A. P. (1971). *Bivariate failure rate.* Journal of the American Statistical Association, 66, 103-104.

[15] Belzunce, F., Navarro, J., Ruiz, J. M. and Del Aguila, Y. (2004). *Some results on residual entropy function*. Metrika, 59, 147-161.

[16] Bismi, G. (2005). *Bivariate Burr Distributions.* PhD Thesis, Cochin University of Science and Technology, India.

[17] Bloch-Mercier, S. (2001).*Monotone markov processes with respect to the reversed hazrad rate ordering: An application to reliability.* Journal of Applied probability, 38 (1), 195-208.

[18] Block, H. W., Savits, T. H. and Singh, H. (1998). *The reversed hazard rate function*. Probability in the Engineering and Informational Sciences, 12, 69-90.

[19] Bracquemond, C. and Gaudoin, O. (2003). *A survey on discrete lifetime distributions*. International Journal of Reliability, Quality and Safety Engineering, 10, 69-98.

[20] Castillo, E. and Hadi, A. S. (1995). *Modeling lifetime data with applications to fatigue models.* Journal of the American Statistical Association, 90 (431), 1041-1054.

[21] Chandra, N. K. and Roy, D. (2001). *Some results on reversed hazard rate.* Probability in the Engineering and Informational Sciences, 15 (1), 95-102.

[22] Chandra, N. K. and Roy, D. (2005). *Classification of distributions based on reversed hazard rate.* Calcutta Statistical Association Bulletin, 56, 221-224.

[23] Chen, Y. Q., Wang, M. C. and Huang, Y. (2004). *Semiparametric regression analysis on longitudinal pattern of recurrent gap times.* Biostatistics, 5 (2), 277-290.

[24] Cox, D. R. (1962). *Renewal Theory.* Methuen, London.

[25] Cox, D. R. (1972). *Regression models and life tables.* Journal of the Royal Statistical Society, Series B, 34 (2), 187-220.

[26] Dewan, I. and Sudheesh, K. K. (2009). *On proportional (reversed) hazard model for discrete data.* Technical report, Indian Statistical Institute, New Delhi, India.

[27] Di Crescenzo, A. (2000). *Some results on the proportional reversed hazards model.* Statistics and Probability Letters, 50, 313-321.

[28] Di Crescenzo, A. and Longobardi, M. (2002). *Entropy-based measure of uncertainty in past lifetime distributions.* Journal of Applied Probability, 39, 434-440.

[29] Di Crescenzo, A. and Longobardi, M. (2004). *A measure of discrimination between past lifetime distributions.* Statistics and Probability Letters, 67, 173-182.

[30] Di Crescenzo, A. and Longobardi, M. (2009). *On cumulative entropies.* Journal of Statistical Planning and Inference, 139, 4072-4087.

[31] Drissi, N., Chonavel, T. and Boucher, J. M. (2008). *Generalized cumulative residual entropy for distributions with unrestricted supports.* Research Letters in Signal Processing, Article ID 790607, 5p.

[32] Ebrahimi, N. (1996). *How to measure uncertainty in the residual lifetime distribution*. Sankhya Series A, 58, 48-56.

[33] Ebrahimi, N. (1997). *Testing whether lifetime distribution is decreasing uncertainty.* Journal of Statistical Planning and Inference, 64, 9-19.

[34] Ebrahimi, N. (2000). *The maximum entropy method for lifetime distributions.* Sankhya Series A, 62, 236-243.

[35] Ebrahimi, N. and Kirmani, S. N. U. A. (1996). *A characterization of the proportional hazards model through a measure of discrimination between two residual life distributions.* Biomertika, 83 (1), 233-235.

[36] Ebrahimi, N. and Pellerey, F. (1995). *New partial ordering of survival functions based on notion of uncertainty.* Journal of Applied Probability, 32, 202-211.

[37] Finkelstein, M. S. (2002). *On the reversed hazard rate.* Reliability Engineering and System Safety, 78 (1), 71-75.

[38] Galambos, J. and Kotz, S. (1978). *Characterization of Probability Distributions, Lecture Notes in Mathematics.* Springer-Verlag Berlin Heidelberg, New York.

[39] Gauss, C. F. (1807). *Disquisitiones Arithmeticae.* Leipzig, 1801; French translation by Poullet Delisle.

[40] Goliforushani, S. and Asadi, M. (2008). *On the discrete mean past lifetime.* Metrika, 68, 209-217.

[41] Gumbel, E. J. (1958). *Statistics of Extremes.* Columbia University Press, New York.

[42] Gupta, R. C., Gupta, P. L. and Gupta, R. D. (1998). *Modeling failure time data with Lehmann alternative.* Communications in Statistics Theory and Methods, 27, 887-904.

[43] Gupta, R. C., Gupta, R. D. and Gupta, P. L. (2006). *Monotonicity of the (reversed) hazard rate of the (maximum) minimum in bivariate distributions.* Metrika, 63, 223-241.

[44] Gupta, R. C. and Gupta, R. D. (2007). *Proportional reversed hazard rate model and its applications.* Journal of Statistical Planning and Inference, 137 (11), 3525-3536.

[45] Gupta, R. C. and Hann, W. (2001). *Analysing survival data by proportional reversed hazard model.* International Journal of Reliability and Applications, 2, 203-216.

[46] Gupta, R. D. and Nanda, A. K. (2001). *Some results on reversed hazard rate ordering.* Communications in Statistics Theory and Methods, 30 (11), 2447-2457.

[47] Gupta, R. D. and Nanda, A. K. (2002). *α - and β - entropies and relative entropies of distributions.* Journal of Statistical Theory and Applications, 1 (3), 177-190.

[48] Gupta, R. C. and Wu, H. (2001). *Analysing survival data by proportional reversed hazard model*. International Journal of Reliability and Applications, 2 (1), 1-26.

[49] Gurler, U. (1996). *Bivariate estimation with right-truncated data*. Journal of the American Statistical Association, 91, 1152-1165.

[50] Hu, T. and Li, Y. (2007). *Increasing failure rate and decreasing reversed hazard rate properties of the minimum and maximum of multivariate distributions with log-concave densities.* Metrika, 65, 325-330.

[51] Johnson, D. H. and Glantz, R. M. (2004). *When does interval coding occur?* Neurocomputing, 59-60, 13-18.

[52] Johnson, N. L. and Kotz, S. (1975). *A vector multivariate hazard rate.* Journal of Multivariate Analysis, 5, 53-66.

[53] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions, Volume 2*. John Wiley and Sons, New York.

[54] Kagan A. M., Linnik, Yu. V. and Rao, C. R. (1973). *Characterization Problems in Mathematical Statistics*. USSR Academy of Sciences(English translation), Wiley, New York.

[55] Kalbfleisch, J. D. and Lawless, J. F. (1989). *Inference based on retrospective ascertainment: an analysis of the data based on transfusion-related AIDS*. Journal of the American Statistical Association, 84, 360-372.

[56] Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data.* John Wiley and Sons, New York.

[57] Kayid, M. (2006). *Multivariate mean inactivity time functions with reliability applications.* International Journal of Reliability and Applications, 7 (2), 125-139.

[58] Kayid, M. and Ahmad, I. A. (2004). *On the mean inactivity time ordering with reliability applications*. Probability in the Engineering and Informational Sciences, 18, 395-409.

[59] Keilson, J. and Sumita, U. (1982). *Uniform stochastic ordering and related inequalities.* The Canadian Journal of Statistics, 10 (3), 181-198.

[60] Kemp, A. W. (2004). *Classes of discrete lifetime distributions.* Communications in Statistics Theory and Methods, 33 (12), 3069-3093.

[61] Khinchin, A. J. (1957). *Mathematical Foundation of Information Theory.* Dover, New York.

[62] Kijima, M. (1998). *Hazard rate and reversed hazard rate monotonicities in continuous time markov chains.* Journal of Applied Probability, 35, 545-556.

[63] Kundu, C. and Nanda, A. K. (2010). *Some reliability properties of the inactivity time.* Communications in Statistics Theory and Methods, 39 (5), 899-911.

[64] Kundu, C., Nanda, A. K. and Maiti, S. S. (2010). *Some distributional results through past entropy.* Journal of Statistical Planning and Inference, 140, 1280-1291.

[65] Lai, C. D. and Xie, M. (2005). *Stochastic aging and dependence for reliability.* Springer Verlag, New York.

[66] Lau, K. S. and Rao, C. R. (1984). *Solution to the integrated Cauchy functional equation on the whole line.* Sankhya Series A, 46 (3), 311-318.

[67] Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data.* Wiley, New York.

[68] Lehmann, E. L. (1953). *The power of rank tests.* Annals of Mathematical Statistics, 24, 23-43.

[69] Li, X. and Da, G. (2010). *Stochastic comparisons in multivariate mixed model of proportional reversed hazard rate with applications*. Journal of Multivariate Analysis, 101, 1016-1025.

[70] Li, X., Da, G. and Zhao, P. (2010). *On reversed hazard rate in general mixture models*. Statistics and Probability Letters, 80, 654-661.

[71] Li, X. and Li, Z. (2008). *A mixture model of proportional reversed hazard rate.* Communications in Statistics Theory and Methods, 37 (18), 2953-2963.

[72] Li, X. and Lu, J. (2003). *Stochastic comparisons on residual life and inactivity time of series and parallel systems.* Probability in the Engineering and Informational Sciences, 17 (2), 267-275.

[73] Maiti, S. S. and Nanda, A. K. (2009). *A loglikelihood-based shape measure of past lifetime distribution.* Calcutta Statistical Association Bulletin, 61, 303-320.

[74] Malinowska, I. and Szynal, D. (2008). *On characterization of certain distributions of k^{th} lower (upper) record values.* Applied Mathematics and Computation, 202, 338-347.

[75] Mathai, A. M. and Pederzoli, G. (1977). *Characterizations of the Normal Probability Law.* Wiley Eastern, New Delhi.

[76] Muliere, P. and Prakasa Rao, B. L. S. (2003). *Characterization of probability distributions under binary associative operation.* Sankhya Series A, 65, 799-806.

[77] Muliere, P. and Scarsini, M. (1987). *Characterization of a Marshall-Olkin type distributions.* Annals of Institute of Statistical Mathematics, 39, 429-441.

[78] Nair, N. U. and Asha, G. (2004). *Characterizations using failure and reversed failure rates.* Journal of the Indian Society for Probability and Statistics, 8, 45-56.

[79] Nair, N. U. and Asha, G. (2008). *Some characterizations based on bivariate reversed mean residual life.* ProbStat Forum, 1, 1-14.

[80] Nair, N. U. and Hitha, N. (1989). *Characterization of discrete models by distribution based on their partial sums.* Statistics and Probability Letters, 8, 335-337.

[81] Nair, N. U. and Sankaran, P. G. (1991). *Characterization of Pearson family of distributions.* IEEE Transactions on Reliability, 40 (1), 75-77.

[82] Nair, N. U., Sankaran, P. G. and Asha, G. (2005). *Characterizations of distributions using reliability concepts.* Journal of Applied Statistical Science, 14, 237-241.

[83] Nanda, A. K., Bhattacharjee, S. and Alam, S. S. (2006). *On upshifted reversed mean residual life order.* Communications in Statistics Theory and Methods, 35 (8), 1513-1523.

[84] Nanda, A. K. and Paul, P. (2006a). *Some properties of past entropy and their applications.* Metrika, 64, 47-61.

[85] Nanda, A. K. and Paul, P. (2006b). *Some results on generalized past entropy.* Journal of Statistical Planning and Inference, 136, 3659-3674.

[86] Nanda, A. K. and Paul, P. (2006c). *Some results on generalized residual entropy.* Information Sciences, 176, 27-47.

[87] Nanda, A. K. and Sengupta, D. (2005). *Discrete life distributions with decreasing reversed hazard.* Sankhya, Series A, 67 (1), 106-125.

[88] Nanda, A. K., Singh, H., Misra, N. and Paul, P. (2003). *Reliability properties of reversed residual lifetime.* Communications in Statistics Theory and Methods, 32 (10), 2031-2042.

[89] Navarro, J., Ruiz, J. M. and Sandoval, C. J. (2002). *A characterization of Weibull and Gumbel multivariate exponential distributions.* IAPQR Transactions, 27, 151-158.

[90] Osaki, S. and Li, X. (1988). *Characterizations of the gamma and negative binomial distributions.* IEEE Transactions on Reliability, 37, 379-382.

[91] Patil, G. P., Kotz, S. and Ord, J. K. (1975). *A Modern Course on Statistical Distributions in Scientific Work, Volume 3.* Characterizations and Applications, D. Reidel Publishing Company, Holland.

[92] Prakasa Rao, B. L. S. (2004). *On a bivariate lack of memory property under binary associative operation.* Communications in Statistics - Theory and Methods, 33 (12), 3103-3114.

[93] Rao, M. (2005). *More on a new concept of entropy and information.* Journal of Theoretical Probability, 18, 967-981.

[94] Rao, M., Chen, Y., Vemuri, B. C. and Wang, F. (2004). *Cumulative residual entropy: a new measure of information.* IEEE Transactions on Information Theory, 50 (6), 120-128.

[95] Roy, D. (2002a). *A characterization of model approach for generating bivariate life distributions using reversed hazard rates.* Journal of the Japan Statistical Society, 32 (2), 239-245.

[96] Roy, D. (2002b). *On bivariate lack of memory property and a new definition.* Annals of Institute of Statistical Mathematics, 54, 404-410.

[97] Roy, D. and Gupta, R. P. (1992). *Classifications of discrete lives.* Microelectronics and Reliability, 32 (10), 1459-1473.

[98] Ruiz, J. M. and Navarro, J. (1994). *Characterization of distributions by relationships between failure rate and mean residual life.* IEEE Transactions on Reliability, 43 (4), 640-644.

[99] Ruiz, J. M. and Navarro, J. (1995). *Characterization of discrete distributions using expected values.* Statistical Papers, 36, 237-252.

[100] Sadegh, M. K. (2008). *Mean past and mean residual life functions of a parallel system with nonidentical components.* Communications in Statistics-Theory and Methods, 37 (7), 1134-1145.

[101] Sankaran, P. G. and Gleeja, V. L. (2006). *On bivariate reversed hazard rates.* Journal of the Japan Statistical Society, 36 (2), 213-224.

[102] Sankaran, P. G. and Gleeja, V. L. (2007). *Nonparametric estimation of reversed hazard rate.* Calcutta Statistical Association Bulletin, 59, 55-68.

[103] Sankaran, P. G. and Gleeja, V. L. (2008). *Proportional reversed hazard and frailty models.* Metrika, 68 (3), 333-342.

[104] Schroeder, M. J. (2004). *An alternative to entropy in the measurement of information.* Entropy, 6, 388-412.

[105] Sengupta, D., Chatterjee, A. and Chakraborty, B. (1995). *Reliability bounds and other inequalities for discrete life distributions.* Microelectronics Reliability, 35 (12), 1473-1478.

[106] Sengupta, D., Singh, H. and Nanda, A. K. (1998). *The proportional reversed hazard model.* Technical Report, Applied Statistics Division, Calcutta: Indian Statistical Institute.

[107] Shaked, M. and Shanthikumar, J. G. (1994). *Stochastic Orders and Their Applications.* Academic Press, New York.

[108] Shaked, M., Shanthikumar, J. G. and Valdez-Torres, J. B. (1994). *Discrete probabilistic orderings in reliability theory.* Statistica Sinica, 4, 567-579.

[109] Shannon, C. E. (1948). *A mathematical theory of communications.* Bell System Technical Journal, 27, 379-423.

[110] Tavangar, M. and Asadi, M. (2008). *On a characterization of generalized Pareto distribution based on generalized order statistics.* Communications in Statistics-Theory and Methods, 37 (9), 1347-1352.

[111] Ullah, A. (1996). *Entropy, divergence and distance measures with economic applications.* Journal of Statistical Planning and Inference, 49, 137-162.

[112] Wang, F. and Vemuri, B. C. (2007). *Non-rigid multi-modal image registration using cross-cumulative residual entropy.* International Journal of Computer Vision, 74 (2), 201-215.

[113] Wang, F., Vemuri, B. C., Rao, M. and Chen, Y. (2003a). *A new and robust information theoretic measure and its application to image alignment,* In: Information Processing in Medical Imaging, Lecture Notes in Computer Science, Volume 2732, Springer Heidelberg, 388-400.

[114] Wang, F., Vemuri, B. C., Rao, M. and Chen, Y. (2003b). *Cumulative residual entropy, a new measure of information and its application to image alignment,* In: Proceedings on the Ninth IEEE International Conference on Computer Vision (ICCV'03), Volume 1, IEEE Computer Society Press, Silver Spring, MD, 548-553.

[115] Ware, J. H. and DeMets, D. L. (1976). *Reanalysis of some baboon descent data.* Biometrics, 32 (2), 459-463.

[116] Wiener, N. (1948). *Cybernetics.* second ed. 1961, MIT Press and Wiley, New York.

[117] Xekalaki, E. (1983). *Hazard functions and life distributions in discrete time.* Communications in Statistics - Theory and Methods, 12, 2503-2509.

[118] Xekalaki, E. and Dimaki, C. (2005). *Identifying the Pareto and Yule distributions by properties of their reliability measures.* Journal of Statistical Planning and Inference, 131, 231-252.

[119] Xie, M., Tang, Y. and Goh, T. N. (2002). *A modified Weibull extension with bathtub-shaped failure rate function, Reliability Engineering and System Safety, 76 (3), 279-285.*

[120] Zografos, K. and Nadarajah, S. (2005). *Survival exponential entropies.* IEEE Transactions on Information Theory, 51, 1239-1246.