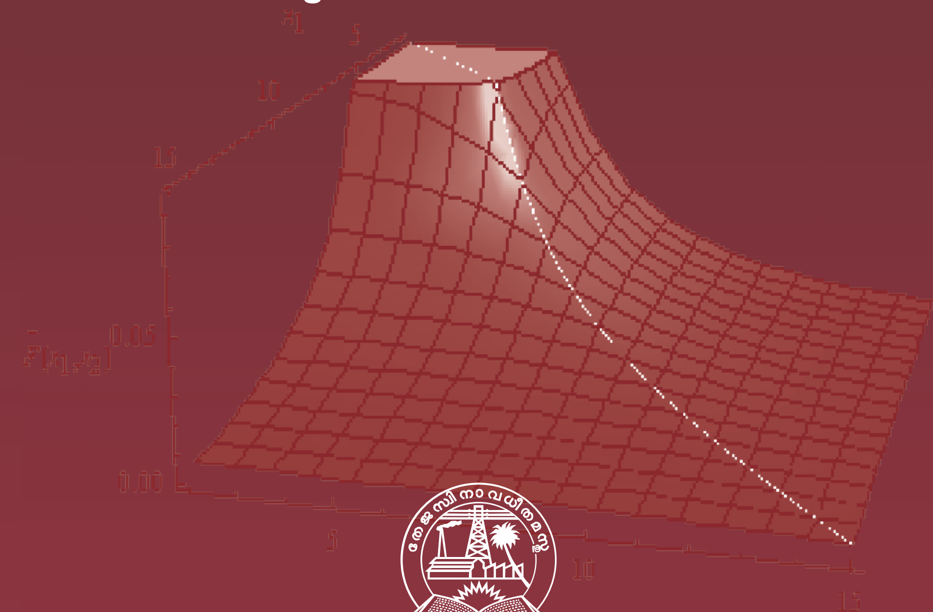
A BIVARIATE PARETO DISTRIBUTION WITH FREUND'S DEPENDENCE STRUCTURE

Thesis submitted to the
Cochin University of Science and Technology
for the Award of Degree of
Doctor of Philosophy
under the Faculty of Science

By
Jagathnath Krishna K.M.

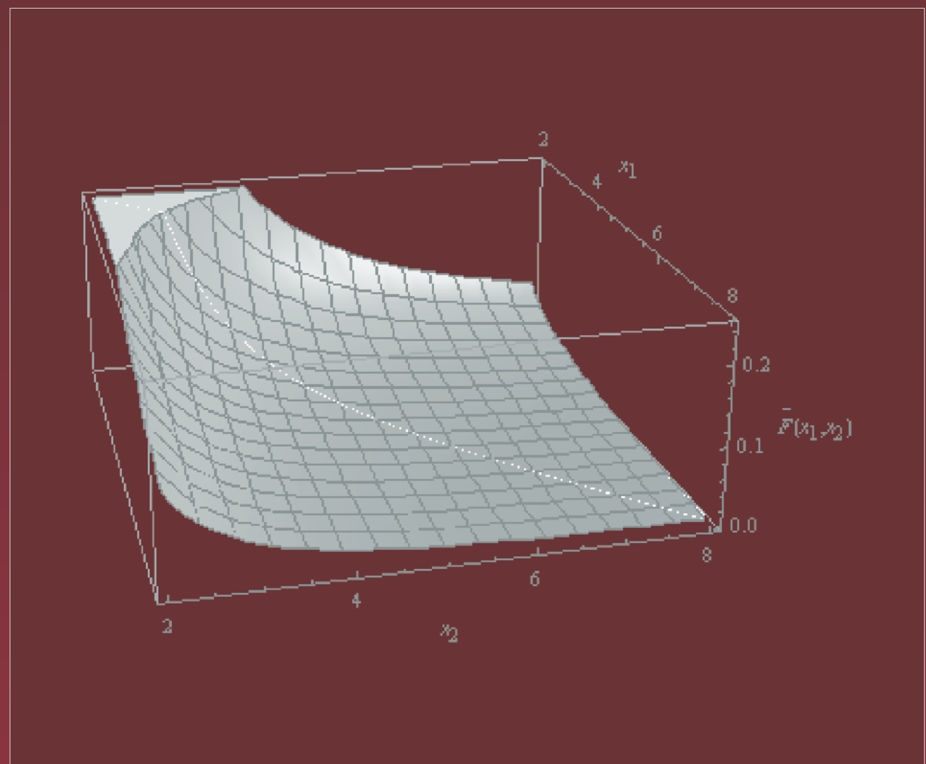


Department of Statistics
Cochin University of Science and Technology
Cochin – 682022

April 2010

A BIVARIATE PARETO DISTRIBUTION WITH
FREUND'S DEPENDENCE STRUCTURE

Jagathnath Krishna K.M.



A BIVARIATE PARETO DISTRIBUTION WITH FREUND'S DEPENDENCE STRUCTURE

Thesis submitted to the
Cochin University of Science and Technology
for the Award of Degree of
Doctor of Philosophy
under the Faculty of Science

By

Jagathnath Krishna K.M.

Department of Statistics
Cochin University of Science and Technology
Cochin – 682022

April 2010

To My Loving Parents

CERTIFICATE

Certified that the thesis entitled 'A Bivariate Pareto Distribution with Freund's Dependence Structure' is a bonafide record of works done by Shri. Jagathnath Krishna K.M. under my guidance in the Department of Statistics, Cochin University of Science and Technology, Cochin-22, Kerala, India and that no part of it has been included anywhere previously for the award of any degree or title.

Cochin-22 Dr. Asha Gopalakrishnan

29 April 2010 (Supervising Guide)

DECLARATION

The thesis entitled 'A Bivariate Pareto Distribution with Freund's Dependence Structure' contains no material which has been accepted for the award of any Degree 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.

Cochin-22 Jagathnath Krishna K.M.

29 April 2010

Acknowledgements

With pleasure, I placed on record my sincere gratitude to my supervising guide Dr. Asha Gopalakrishnan, Reader, Department of Statistics, Cochin University of Science and Technology for the enthusiastic guidance, support and steadfast encouragement for the successful completion of the research work.

I am obliged to Dr. N. Balakrishna, Professor and Head and Dr. K.R. Muraleedharan Nair, Professor and Former Head, Department of Statistics, Cochin University of Science and Technology for the extensive support and encouragement given to me.

I also wish to express my profound gratitude to all the faculty members of the Department of Statistics, Cochin University of Science and Technology for their timely advice and suggestions during the entire period of my research.

I remember with deep gratefulness all my former teachers.

The amity with my fellow researchers and friends was stimulating, inspirational and enriching. I owe each and every one of them, who helped and motivated me whenever it is needed.

I express my earnest gratitude to the non-teaching staffs, Department of Statistics, Cochin University of Science and Technology for the co-operation and help they had rendered.

I am deeply indebted to my family and relatives for their patience, sacrifice, encouragement, prayers and immense support given to me, without which I would not be able to fulfill this work.

I express my heartfelt thanks to the University Grants Commission and Council of Scientific and Industrial Research, Government of India, for the financial support extended by them.

Above all, I bow before the grace of the Almighty.

Jagathnath Krishna K.M.

Contents

	Page No.
Chapter 1 Introduction	1
1.1 The Univariate Pareto Distribution	1
1.2 Basics	4
1.2.1 Survival Function	4
1.2.2 Failure Rate	5
1.2.3 Mean Residual Life Function	6
1.2.4 Vitality Function	7
1.2.5 Geometric Vitality Function	7
1.3 Bivariate Notions	8
1.3.1 Bivariate Failure Rate	9
1.3.2 Bivariate Mean Residual Life Function	11
1.3.3 Bivariate Vitality Function	14
1.3.4 Bivariate Geometric Vitality Function	16
1.4 Information Measures	18
1.4.1 Entropy Function	18
1.4.2 Residual Entropy Function	19
1.4.3 Bivariate Residual Entropy Function	20
1.5 Measures of Inequality	21
1.5.1 The Lorenz Curve	22
1.5.2 The Gini Index	25
1.6 Multivariate Pareto Distributions	28
1.7 Present Study	34

Chapter 2 A Bivariate Pareto Model 38

2.1 Introduction 38

2.2 Model and its Properties 40

2.3 Sums and Ratios for $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ Distribution 53

2.4 Transformed Exponential Variates 55

2.5 Estimation of Parameters 58

2.6 Data Analysis 72

Chapter 3 Characterizations of Bivariate Pareto Distributions 75

3.1 Introduction 75

3.2 Characterizations Using Dullness Property and its Variants 76

3.3 Characterizations Based on Reliability Concepts 84

Chapter 4 Residual Measure of Uncertainty for Bivariate Distributions with Load Sharing Dependence 92

4.1 Introduction 92

4.2 Bivariate Residual Entropy Function 93

4.3 Characterizations Using Bivariate Residual Entropy Function 97

Chapter 5 Inequality Measures for Bivariate Distributions with Load Sharing Dependence 111

5.1 Introduction 111

5.2 Definitions and Relationships 114

5.3 Characterizations 125

Chapter 6 A General Representation 132

6.1 Introduction 132

6.2 The General Representation 136

6.3 Failure Rate of the General Class 140

6.4 Conclusion 147

6.5 Future Work 148

References 150

Chapter 1

Introduction

1.1 The Univariate Pareto Distribution

The Pareto distribution is named after the Italian economist Vilfredo Pareto. Pareto (1897) originally used this distribution to describe the allocation of wealth among individuals since it seemed to show rather well the way that a large portion of wealth of any society is owned by a smaller percentage of the people in that society. The classical Pareto distribution called the Pareto I distribution with survival function $\bar{F}(x) = P\left[X \geq x\right]$ is given by

$$\bar{F}(x) = \left(\frac{x}{\sigma}\right)^{-\alpha} ; x > \sigma, \sigma > 0 \qquad (1.1)$$

where α is the shape parameter and σ is the scale parameter. This model with its heavy tail soon became an accepted model for income. The parameter α is referred to as Pareto's index of inequality.

In much of the literature the standard Pareto distribution is permitted to have an additional location parameter. It is called the Pareto II distribution or Lomax distribution and its survival function is given by

$$\bar{F}(x) = \left(1 + \frac{x-\mu}{\sigma}\right)^{-\alpha} ; x > \mu \tag{1.2}$$

where μ the location parameter, is real, with $\sigma > 0$ and $\alpha > 0$. This model finds application in reliability studies as a model to incorporate environmental influence on a system having exponential lifetime (Marshall (1975)).

An alternative variation on the Pareto theme, which provides tail behaviour similar to (1.2), is provided by the Pareto III family with survival function

$$\bar{F}(x) = \left(1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right)^{-1} ; x > \mu \tag{1.3}$$

where μ is real, $\sigma > 0$ and $\gamma > 0$, is called the inequality parameter.

This distribution is further generalized by introducing a shape parameter, to arrive at the Pareto IV family,

$$\bar{F}(x) = \left(1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right)^{-\alpha} ; x > \mu \tag{1.4}$$

where μ is real, $\sigma > 0$, $\alpha > 0$ and $\gamma > 0$.

Feller (1971, p. 49) defined a Pareto distribution in a some what different manner. Let Y has a beta distribution with parameters γ_1 and γ_2, that is

$$f(y) = \frac{y^{\gamma_1 - 1}(1-y)^{\gamma_2 - 1}}{B(\gamma_1, \gamma_2)}; 0 < y < 1 \tag{1.5}$$

and define $X = Y^{-1} - 1$. Then X has what Feller called a Pareto distribution given by

$$f(x) = \frac{x^{\gamma_2 - 1}(1+x)^{-\gamma_1 - \gamma_2}}{B(\gamma_1, \gamma_2)}; x > 0. \tag{1.6}$$

This family represents a generalization of the Pareto IV family.

Here we restrict our studies to the extensions of Pareto I and Pareto II distributions.

Lot of work has appeared in literature with modification, application, generalization and inference of these models. Krishnaji (1970) has assumed that the reporting errors are multiplicative and obtained characterization results of Pareto I distribution. Ahsanullah and Kabir (1973) studied the model (1.1) using order statistics. Revankar et al. (1974) assumed that the under reporting errors are additive and showed that, for constant $\alpha, \beta > 0$,

$$E\left[U \mid X > y\right] = \alpha + \beta y$$

where X, Y and U denote the actual income, reported income and under reporting error, is necessary and sufficient condition for the random variable X to follow Pareto I distribution. Talwaker (1980) introduced the concept of dullness of distribution defined as follows.

An income distribution of X is said to be dull at a point t, i.e. incapable of utilizing the information about the reported income t, whenever

$$P\left[X \geq st \mid X \geq t\right] = P\left[X \geq s\right] \text{ for all } s \geq 1 \text{ and given } t \geq 1. \quad (1.7)$$

The distribution of X will be called totally dull, if equation (1.7) holds true for all $s \geq 1$ and all $t \geq m > 1$ where m is some fixed real number. Talwaker also obtained characterization for the model (1.1) using this property. Cuadras et al. (2006) expanded Pareto distribution as a series of principal components and made a comparison with exponential distribution. Also he obtained the asymptotic distribution. Nadarajah (2005) has given an exponential Pareto model. Nadarajah and Gupta (2008) obtained a product Pareto distribution and discussed its properties. The generalized Pareto distribution was introduced by Pickands (1975). Davison and Smith (1990) pointed out that the generalized Pareto might form the basis of a broad modeling approach to high level exceedances.

Estimation of parameters of the models has been undertaken by several researchers. One of the earlier works related to estimation of a Pareto model is by Quandt (1966). In his paper 'old and new method of estimation of Pareto distribution', the traditional method of estimation is discussed in detail and he formulated a new method of estimation by minimization of the criterion function. Later on lots of work has appeared in these lines. Recently, Somesh Kumar and Bandhyopadhyay (2005) obtained UMVUE for the scale parameter of the model when the shape parameters are assumed to be equal. Srivastava and Gupta (2005) obtained a modified Pitman estimator for the ordered scale parameters of two Pareto distributions and showed that they improve up on Pitman estimators using simulation study. Garren et al. (2007) obtained improved estimate of the Pareto's location parameter under the squared error loss function. Several variants and properties of the Pareto distribution are discussed in Arnold (1983) and Johnson et al. (1994).

These distributions have been extended to the bivariate and multivariate set up. The approach of extending to multivariate set up has either been that of generalizing univariate distributional properties, univariate notions and obtaining models for which univariate marginals belong to that family. Before looking into some popular generalization of Pareto I and Pareto II distributions, we overview some basic concepts and notions that are been used directly or indirectly in developing these multivariate models. We also include the concepts that will be used through out this thesis.

1.2 Basics

Let X be a non-negative random variable defined on a probability space (Ω, A, P) with distribution function $F(x) = P[X < x]$. The random variable X could represent the income from a source or the length of life of a device, measured in units of time.

1.2.1 Survival Function

The function

$$\bar{F}(x) = P[X \geq x] \tag{1.8}$$

is called survival function or reliability function. $\bar{F}(x)$ is a non-increasing continuous function with $\bar{F}(0)=1$ and $\lim_{x\to\infty}\bar{F}(x)=0$. For an absolutely continuous $\bar{F}(x)$, the probability density function of X is

$$f(x)=-\frac{d\bar{F}(x)}{dx}.$$

1.2.2 Failure Rate

Lifetime distributions are usually characterized using the concept of failure rate $\lambda(x)$, defined as

$$\lambda(x)=\lim_{\vartriangle x\to 0}\frac{P\left[x\le X<x+\vartriangle x\mid X>x\right]}{\vartriangle x}. \tag{1.9}$$

When $f(x)$ is the probability density function of X, (1.9) can be equivalently written as

$$\begin{aligned}\lambda(x)&=\frac{f(x)}{\bar{F}(x)}\\ &=\frac{d}{dx}\left[-\log\bar{F}(x)\right].\end{aligned} \tag{1.10}$$

The failure rate $\lambda(x)$, measures the instantaneous rate of failure or death at time x, given that an individual survives up to time x. The failure rate is also known as conditional failure rate in reliability, the hazard rate in survival analysis, the force of mortality in demography, the age-specific failure rate in epidemiology. In extreme-value theory, it is known as the intensity rate and its reciprocal is termed as Mill's ratio in economics.

When X is non-negative and has a distribution function absolutely continuous with respect to the Lebesgue measure, (1.10) provides

$$\bar{F}(x)=\exp\left(-\int_0^x\lambda(t)dt\right). \tag{1.11}$$

Equation (1.11) indicates that $\lambda(x)$ is a non-negative function with $\int_0^x\lambda(t)dt<\infty$, for some $x>0$ and $\int_0^\infty\lambda(t)dt=\infty$. From equation (1.11), it can be noted that $\lambda(x)$ uniquely determines the distribution. It is shown that constancy of $\lambda(x)$ is the characteristic property of the exponential distribution (Galambos and Kotz (1978)).

1.2.3 Mean Residual Life Function (MRLF)

The mean residual life function $m(x)$, for a random variable X defined on the real line with $E\left[X\right]<\infty$, is given by (Swartz (1973))

$$m(x)=E\left[X-x \mid X\geq x\right] \tag{1.12}$$

for all x. The mean residual life function $m(x)$, represents the average lifetime remaining for a component, which has already survived up to time x. When $F(x)$ is absolutely continuous with respect to Lebesgue measure, (1.12) becomes

$$m(x)=\frac{1}{\overline{F}(x)}\int_{x}^{\infty}\overline{F}(t)dt. \tag{1.13}$$

A function $m(x)$ is a mean residual life function of some random variable with an absolutely continuous distribution function if only if $m(x)$ satisfies the following properties.

(i) $0\leq m(x)<\infty, x\geq 0.$

(ii) $m(0)>0.$

(iii) $m(x)$ is continuous in $x.$

(iv) $m(x)+x$ is increasing on R^{+}, where $R^{+}=\left\{x \mid x\in\left[0,\infty\right]\right\}$.

(v) When there exist an x_0 such that $m(x_0)=0$ then $m(x)=0$ for $x\geq x_0$ otherwise, there does not exist such an x_0 with $m(x_0)=0$, then $\int_0^{\infty}m^{-1}(x)dx=\infty$.

Further $m(x)$ uniquely determine the underlying distribution through the expression,

$$\overline{F}(x)=\frac{m(0)}{m(x)}\exp\left(-\int_{0}^{x}\frac{dt}{m(t)}\right). \tag{1.14}$$

Also the MRLF is related to the failure rate by

$$\lambda(x)=\frac{1+\frac{d}{dx}m(x)}{m(x)}. \tag{1.15}$$

1.2.4 Vitality Function

The vitality function $v(x)$, of a random variable X admitting an absolutely continuous distribution function $F(x)$, with respect to Lebesgue – Stieljes measure on real line is given by (Kupka and Loo (1989))

$$\begin{aligned} v(x) &= E\left[X \mid X \geq x\right] \\ &= \frac{1}{\bar{F}(x)}\int_x^\infty t dF(t). \end{aligned} \tag{1.16}$$

The vitality function satisfies the following properties.

(i) $v(x)$ is non-decreasing and left continuous on $(-\infty, L)$, where $L=\inf\{x : F(x)=1\}$.

(ii) $v(x) > x$ for all $x < L$.

(iii) $\lim\limits_{x\to L^-} v(x) = L.$

(iv) $\lim\limits_{x\to -\infty} v(x) = E\left[X\right].$

Moreover, $v(x)$ is related to $m(x)$ through the relationship

$$v(x)=m(x)+x \tag{1.17}$$

and

$$\frac{d}{dx}v(x)=m(x)\lambda(x). \tag{1.18}$$

1.2.5 Geometric Vitality Function

Let X be a non-negative random variable admitting an absolutely continuous distribution function $F(x)$ on $(0, L)$, where

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

with $E[X] < \infty$. The geometric vitality function $GV(t)$, for $t > 0$ is defined as (Nair and Rajesh (2000))

$$\log GV(t) = E[\log X \mid X > t]$$

$$= \frac{1}{\bar{F}(t)} \int_t^\infty \log x \, f(x) dx. \tag{1.19}$$

i.e.

$$\log GV(t) = \frac{1}{\bar{F}(t)} \int_t^\infty \frac{\bar{F}(x)}{x} dx. \tag{1.20}$$

In the reliability context, if X represents the life length of a component, $GV(t)$ represents the geometric mean of lifetime of the components which has survived up to time t.

The geometric vitality function satisfies

(i) $\log GV(t)$ is non-decreasing in t.

(ii) $\lim_{t \to 0} \log GV(t) = E[\log X]$.

(iii) $v(t) \geq \log GV(t)$, for all $t > 0$.

(iv) If $\lambda(t) = \frac{f(t)}{\bar{F}(t)}$ is the failure rate of T, then $\lambda(t) = \dfrac{\frac{d}{dt} \log GV(t)}{\log\left[\frac{GV(t)}{t}\right]}$.

1.3 Bivariate Notions

Let (X_1, X_2) be a non-negative random vector on $R_2^+ = (0, \infty) \times (0, \infty)$ with a bivariate distribution function $F(x_1, x_2)$. Then the bivariate survival function of (X_1, X_2), denoted by $\bar{F}(x_1, x_2)$ is defined as

$$\bar{F}(x_1, x_2) = P[X_1 > x_1, X_2 > x_2],$$

which is related to $F(x_1, x_2)$ as

$$\bar{F}(x_1, x_2) = 1 - F(x_1, \infty) - F(\infty, x_2) + F(x_1, x_2).$$

If $F(x_1, x_2)$ is absolutely continuous and if the second order derivative exists, then

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

1.3.1 Bivariate Failure Rate

In the bivariate case, the failure rate can be defined in more than one way. One definition of bivariate failure rate was given by Basu (1971) as

$$r\left(x_1,x_2\right)=\frac{f\left(x_1,x_2\right)}{\bar{F}\left(x_1,x_2\right)}. \tag{1.21}$$

Unlike the univariate case, $r\left(x_1,x_2\right)$, in general does not determine the bivariate distribution uniquely.

Another definition proposed is the bivariate failure rate by Cox (1972) defined as a vector

$$\underline{\lambda}\left(\underline{x}\right)=\left(\lambda\left(x\right),\lambda_{12}\left(x_1 \mid x_2\right),\lambda_{21}\left(x_2 \mid x_1\right)\right) \tag{1.22}$$

where

$$\lambda\left(x\right)=\lambda_{10}\left(x\right)+\lambda_{20}\left(x\right),$$

$$\lambda_{i0}\left(x\right)=\lim_{\vartriangle x\to 0^+}\frac{P\left[x\le X_i<x+\vartriangle x \mid x\le X_1,x\le X_2\right]}{\vartriangle x}, i=1,2,$$

$$\lambda_{12}\left(x_1 \mid x_2\right)=\lim_{\vartriangle x_1\to 0^+}\frac{P\left[x_1\le X_1<x_1+\vartriangle x_1 \mid x_1\le X_1,X_2=x_2\right]}{\vartriangle x_1}, x_2<x_1$$

and

$$\lambda_{21}\left(x_2 \mid x_1\right)=\lim_{\vartriangle x_2\to 0^+}\frac{P\left[x_2\le X_2<x_2+\vartriangle x_2 \mid x_2\le X_2,X_1=x_1\right]}{\vartriangle x_2}, x_1<x_2.$$

Note that if $\bar{F}\left(x_1,x_2\right)$ admits a density function then,

$$\lambda\left(x\right)=\frac{f_Z(x)}{\bar{F}_z(x)},\ \text{where } Z=min\left(X_1,X_2\right).$$

Also,

$$\lambda_{i0}\left(x\right)=p_i\lambda\left(x\right),\ \text{where } p_i=P\left[X_i<X_{3-i}\right],\ i=1,2,$$

$$\lambda_{12}\left(x_1 \mid x_2\right)=\frac{\left[\dfrac{\partial^2 \bar{F}\left(x_1, x_2\right)}{\partial x_1 \partial x_2}\right]}{\left[-\dfrac{\partial \bar{F}\left(x_1, x_2\right)}{\partial x_2}\right]}, x_2<x_1$$

and

$$\lambda_{21}\left(x_2 \mid x_1\right)=\frac{\left[\dfrac{\partial^2 \bar{F}\left(x_1, x_2\right)}{\partial x_1 \partial x_2}\right]}{\left[-\dfrac{\partial \bar{F}\left(x_1, x_2\right)}{\partial x_1}\right]}, x_1<x_2 .$$

The probability density function $f(x_1, x_2)$ in terms of $\lambda(x)$ is given by (Cox (1972))

$$f\left(x_1, x_2\right)=\begin{cases}\exp\left(-\int_0^{x_1-0} \lambda(u) du-\int_{x_1+0}^{x_2-0} \lambda_{21}\left(u \mid x_1\right) du\right) \lambda_{10}\left(x_1\right) \lambda_{21}\left(x_2 \mid x_1\right) \\ \qquad ; x_1<x_2 \\ \exp\left(-\int_0^{x_2-0} \lambda(u) du-\int_{x_2+0}^{x_1-0} \lambda_{12}\left(u \mid x_2\right) du\right) \lambda_{20}\left(x_1\right) \lambda_{12}\left(x_1 \mid x_2\right) \\ \qquad ; x_2<x_1\end{cases} . \tag{1.23}$$

Johnson and Kotz (1975) defined bivariate failure rate as a vector given by

$$h\left(x_1, x_2\right)=\left(h_1\left(x_1, x_2\right), h_2\left(x_1, x_2\right)\right) \tag{1.24}$$

where

$$h_i\left(x_1, x_2\right)=-\frac{\partial \log \bar{F}(x_1, x_2)}{\partial x_i}, i=1,2$$

is the instantaneous failure of X_i at time x_i given that X_i was alive at time x_i and that X_{3-i} survived beyond time x_{3-i}. The Johnson and Kotz (1975) vector failure rate uniquely determine the distribution through the expression.

$$\bar{F}\left(x_1, x_2\right)=\exp\left[-\int_0^{x_1} h_1(u, 0) du-\int_0^{x_2} h_2\left(x_1, u\right) du\right] \tag{1.25}$$

or

$$\bar{F}(x_1,x_2)=\exp\left[-\int_0^{x_1} h_1(u,x_2)\,du-\int_0^{x_2} h_2(0,u)\,du\right]. \tag{1.26}$$

Marshall (1975), Shaked and Shanthikumar (1987), Basu and Sun (1997), Finkelstein (2003) have also discussed different versions of failure rate in bivariate setup. Shaked and Shanthikumar (1987) has given the total hazard accumulation of the random variable for a bivariate random variable and extended it to the multivariate case. Finkelstein (2003) considered two conditional hazards associated with $F(x_1,x_2)$ and exponential representations for the survival function in terms of the conditional hazard rates were also obtained.

1.3.2 Bivariate Mean Residual Life Function

Buchanan and Singpurwalla (1977) defined the bivariate mean residual life function (BMRLF) as a direct extension of the univariate case as

$$m(x_1,x_2)=\frac{\int_0^{\infty}\int_0^{\infty} P\left[X_1>x_1+t_1,X_2>x_2+t_2\right]}{\bar{F}(x_1,x_2)},\ x_i>0,\ i=1,2. \tag{1.27}$$

Although $m(x_1,x_2)$ is a direct extension, it does not uniquely determines the underlying distribution.

Another definition for the bivariate mean residual life function is provided independently by Shanbag and Kotz (1987) and Arnold and Zahedi (1988). For a bivariate random vector defined on R_2^{+} with joint distribution function $F(x_1,x_2)$, $L=(L_1,L_2)$ be a vector of extended real numbers such that $L_i=\inf\{x\mid F_i(x_i)=1\}$ where $F_i(x_i)$ is the distribution function of $X_i, i=1,2.$ Further let $E[X_i]<\infty, i=1,2.$ The vector valued Borel-measurable function $m(x_1,x_2)$ on R_2^{+} is defined as

$$\begin{aligned} m(x_1,x_2)&=E\left[X-x\mid X\geq x\right]\\ &=\left(m_1(x_1,x_2),m_2(x_1,x_2)\right) \end{aligned} \tag{1.28}$$

for all $x=(x_1,x_2)\in R_2^+, x_i<L_i, i=1,2,$ such that $P[X>x]>0$ and $X\geq x$ implies $X_i\geq x_i, i=1,2$ is called the bivariate mean residual life function. When (X_1,X_2) is continuous and non-negative the components of bivariate mean residual life function is given by

$$m_1(x_1,x_2)=E[X_1-x_1\mid X_1\geq x_1,X_2\geq x_2]$$
$$=\frac{1}{\overline{F}(x_1,x_2)}\int_{x_1}^{\infty}\overline{F}(t,x_2)\,dt$$

and

$$m_2(x_1,x_2)=E[X_2-x_2\mid X_1\geq x_1,X_2\geq x_2]$$
$$=\frac{1}{\overline{F}(x_1,x_2)}\int_{x_2}^{\infty}\overline{F}(x_1,t)\,dt.$$

It is established that $m(x_1,x_2)$ determine the distribution of $X=(X_1,X_2)$ uniquely. The unique expression of the survival function in terms of $m(x_1,x_2)$ is provided by Nair and Nair (1988) as

$$\overline{F}(x_1,x_2)=\frac{m_1(0,0)m_2(x_1,0)}{m_1(x_1,0)m_2(x_1,x_2)}\exp\left[-\int_0^{x_1}\frac{dt}{m_1(t,0)}-\int_0^{x_2}\frac{dt}{m_2(x_1,t)}\right] \quad (1.29)$$

or alternatively

$$\overline{F}(x_1,x_2)=\frac{m_1(0,x_2)m_2(0,0)}{m_1(x_1,x_2)m_2(0,x_2)}\exp\left[-\int_0^{x_2}\frac{dt}{m_2(0,t)}-\int_0^{x_1}\frac{dt}{m_1(t,x_2)}\right]. \quad (1.30)$$

The bivariate mean residual life function in (1.28) is related to the bivariate failure rate in (1.24) through the relationship

$$h_i(x_1,x_2)=\frac{1+\frac{\partial}{\partial x_i}m_i(x_1,x_2)}{m_i(x_1,x_2)}, i=1,2. \quad (1.31)$$

Shaked and Shanthikumar (1991) has defined the bivariate conditional mean residual life function corresponding to the Cox's failure rate as

$$m_i(x)=E[X_i-x\mid X_1>x,X_2>x], i=1,2, x\geq 0$$
$$m_1(x\mid x_2)=E[X_1-x\mid X_1>x,X_2=x], x\geq x_2\geq 0 \quad (1.32)$$

and

$$m_2(x \mid x_1) = E[X_2 - x \mid X_1 = x, X_2 > x], \ x \geq x_1 \geq 0.$$

The function defined in (1.32) predict the remaining life of the surviving components by the appropriate expectations, conditioned on the observed past up to time t. Shaked and Shanthikumar (1991) also extended the mean residual life function to the multivariate case.

Another definition closely related to above bivariate mean residual life function proposed is (Asha and Jagathnath (2008))

$$\underline{m}(\underline{x}) = \left(m(x), m_{12}(x_1 \mid x_2), m_{21}(x_2 \mid x_1)\right) \tag{1.33}$$

where

$$m(x) = \frac{1}{\bar{F}_Z(x)} \int_x^{\infty} (t - x) f_Z(t) dt, \ x > 0,$$

$$m_{i0}(x) = p_i m(x), \ p_i = P[X_i < X_{3-i}], i = 1, 2,$$

$$m_{12}(x_1 \mid x_2) = \frac{\int_{x_1}^{\infty} (t - x_1) f(t, x_2) \, dt}{\int_{x_1}^{\infty} f(t, x_2) \, dt}, \ x_1 > x_2$$

and

$$m_{21}(x_2 \mid x_1) = \frac{\int_{x_2}^{\infty} (t - x_2) f(x_1, t) \, dt}{\int_{x_2}^{\infty} f(x_1, t) \, dt}, \ x_1 < x_2.$$

The unique expression for the joint density function in terms of mean residual life function is given by

$$f\left(x_1,x_2\right)=\begin{cases}\dfrac{m(0)m_{21}(x_1\mid x_1)}{m(x_1)m_{21}(x_2\mid x_1)}\left(\dfrac{p_1+\dfrac{d}{dx_1}m(x_1)}{m(x_1)}\right)\left(\dfrac{1+\dfrac{\partial}{\partial x_2}m_{21}(x_2\mid x_1)}{m(x_1)}\right) \\ \qquad \exp\left(-\displaystyle\int_0^{x_1}\frac{1}{m(u)}du-\int_{x_1}^{x_2}\frac{1}{m_{21}\left(u\mid x_1\right)}du\right);x_1<x_2 \\ \dfrac{m(0)m_{12}(x_2\mid x_2)}{m(x_2)m_{12}(x_1\mid x_2)}\left(\dfrac{p_2+\dfrac{d}{dx_2}m(x_2)}{m(x_2)}\right)\left(\dfrac{1+\dfrac{\partial}{\partial x_1}m_{12}(x_1\mid x_2)}{m(x_2)}\right) \\ \qquad \exp\left(-\displaystyle\int_0^{x_2}\frac{1}{m(u)}du-\int_{x_2}^{x_1}\frac{1}{m_{12}\left(u\mid x_2\right)}du\right);x_2<x_1\end{cases} \quad .(1.34)$$

The bivariate mean residual life function is related to Cox's failure rate by the relationships

$$\lambda\left(x\right)=\frac{1+\dfrac{d}{dx}m\left(x\right)}{m\left(x\right)},\ x>0\,,$$

$$\lambda_{12}(x_1\mid x_2)=\frac{1+\dfrac{\partial}{\partial x_1}m_{12}(x_1\mid x_2)}{m_{12}(x_1\mid x_2)},\ x_1>x_2$$

and

$$\lambda_{21}(x_2\mid x_1)=\frac{1+\dfrac{\partial}{\partial x_2}m_{21}(x_2\mid x_1)}{m_{21}(x_2\mid x_1)},\ x_2>x_1. \qquad (1.35)$$

1.3.3 Bivariate Vitality Function

Sankaran and Nair (1991) defined the bivariate vitality function of a random variable $\left(X_1,X_2\right)$ defined on R_2^{+} as the vector

$$v\left(x_1,x_2\right)=\left(v_1\left(x_1,x_2\right),v_2\left(x_1,x_2\right)\right) \qquad (1.36)$$

where

$$v_i\left(x_1,x_2\right)=E\left[X_i\mid X_1\geq x_1,X_2\geq x_2\right],\ i=1,2.$$

The bivariate vitality function is related to the mean residual life function by the relationship

$$v_i(x_1, x_2) = x_i + m_i(x_1, x_2),\ i = 1, 2.$$

We propose another extension of vitality function, which is useful to measure the total life span of a two-component parallel system, by

$$\underline{v}(\underline{x}) = \left(v(x), v_{12}(x_1 \mid x_2), v_{21}(x_2 \mid x_1)\right) \tag{1.37}$$

where

$$v(x) = E\left[X \mid X > x\right] = \frac{1}{\overline{F}_Z(x)} \int_x^{\infty} t\, f_Z(t)dt;\ x > 0,$$

$$v_{i0}(x) = p_i v(x),\ p_i = P\left[X_i < X_{3-i}\right], i = 1, 2,$$

$$v_{12}(x_1 \mid x_2) = \frac{\int_{x_1}^{\infty} t\, f(t, x_2)dt}{\int_{x_1}^{\infty} f(t, x_2)dt};\ x_1 > x_2$$

and

$$v_{21}(x_2 \mid x_1) = \frac{\int_{x_2}^{\infty} t\, f(x_1, t)dt}{\int_{x_2}^{\infty} f(x_1, t)dt};\ x_2 > x_1\ .$$

The bivariate vitality $\underline{v}(\underline{x})$ measures the expected life span of the system using the information about the current age of the components. The first element in the vector gives the expected lifetime of the system using the information that both the component has survived beyond 'x'. The second element gives the expected life span of the first component given that it has survived to an age x_1 and the other has failed at x_2. Similar argument holds for the third element.

The bivariate mean residual life function $\underline{m}(\underline{x})$ is related to the bivariate vitality function $\underline{v}(\underline{x})$ through the relationships

$$v(x) = x + m(x),$$

$$v_{12}(x_1 \mid x_2) = x_1 + m_{12}(x_1 \mid x_2)$$

and

$$v_{21}(x_2 \mid x_1) = x_2 + m_{21}(x_2 \mid x_1). \tag{1.38}$$

1.3.4 Bivariate Geometric Vitality Function

Sathar (2002) has extended the concept of geometric vitality function to the bivariate setup. The bivariate geometric vitality function is defined as

$$\log GV(t_1,t_2) = \left(\log GV_1(t_1,t_2), \log GV_2(t_1,t_2)\right) \tag{1.39}$$

where

$$\log GV_1(t_1,t_2) = \frac{1}{\bar{F}\left(t_1 \mid X_2 \geq t_2\right)} \int_{t_1}^{\infty} \log x_1 \; f\left(x_1 \mid X_2 \geq t_2\right) dx_1$$

and

$$\log GV_2(t_1,t_2) = \frac{1}{\bar{F}\left(t_2 \mid X_1 \geq t_1\right)} \int_{t_2}^{\infty} \log x_2 \; f\left(x_2 \mid X_1 \geq t_1\right) dx_2\,, \tag{1.40}$$

which can be equivalently written as

$$\log\left(\frac{GV_1(t_1,t_2)}{t_1}\right) = \frac{1}{\bar{F}\left(t_1 \mid X_2 \geq t_2\right)} \int_{t_1}^{\infty} \frac{\bar{F}\left(x_1 \mid X_2 \geq t_2\right)}{x_1} dx_1$$

and

$$\log GV_2(t_1,t_2) = \frac{1}{\bar{F}\left(t_2 \mid X_1 \geq t_1\right)} \int_{t_2}^{\infty} \frac{\bar{F}\left(x_2 \mid X_1 \geq t_1\right)}{x_2} dx_2\,. \tag{1.41}$$

Corresponding to the bivariate failure rate in equation (1.22) we propose an extension of the geometric vitality function for $F\left(x_1,x_2\right)$ with $E\left[\log X_i\right] < \infty, i = 1,2$ is given by

$$\log GV(t) = \frac{1}{\bar{F}_Z\left(t\right)} \int_{t}^{\infty} \log x \; f_Z\left(x\right) dx, \; t > 0\,,$$

$$\log GV_{12}(t_1 \mid t_2) = \frac{1}{\left[\frac{\partial}{\partial t_2}\bar{F}\left(t_1,t_2\right)\right]} \int_{t_1}^{\infty} \log x_1 \left[\frac{\partial^2}{\partial x_1 \partial t_2}\bar{F}\left(x_1,t_2\right)\right] dx_1, \; t_1 > t_2$$

and

$$\log GV_{21}(t_2 \mid t_1) = \frac{1}{\left[\frac{\partial}{\partial t_1}\bar{F}(t_1, t_2)\right]} \int_{t_2}^{\infty} \log x_2 \left[\frac{\partial^2}{\partial t_1 \partial x_2}\bar{F}(t_1, x_2)\right] dx_2,\ t_2 > t_1. \qquad (1.42)$$

This can be equivalently written as

$$\log\left[\frac{GV(t)}{t}\right] = \frac{1}{\bar{F}_Z(t)} \int_t^{\infty} \frac{\bar{F}_Z(x)}{x} dx,\ t > 0,$$

$$\log\left[\frac{GV_{12}(t_1 \mid t_2)}{t_1}\right] = \frac{1}{\left[\frac{\partial}{\partial t_2}\bar{F}(t_1, t_2)\right]} \int_{t_1}^{\infty} \frac{\left[\frac{\partial}{\partial t_2}\bar{F}(x_1, t_2)\right]}{x_1} dx_1,\ t_1 > t_2$$

and

$$\log\left[\frac{GV_{21}(t_2 \mid t_1)}{t_2}\right] = \frac{1}{\left[\frac{\partial}{\partial t_1}\bar{F}(t_1, t_2)\right]} \int_{t_2}^{\infty} \frac{\left[\frac{\partial}{\partial t_1}\bar{F}(t_1, x_2)\right]}{x_2} dx_2,\ t_2 > t_1. \qquad (1.43)$$

The bivariate geometric vitality function is related to the Cox's failure rate through the relation,

$$\lambda(t) = \frac{\frac{d}{dt}\left(\log GV(t)\right)}{\log\left[\frac{GV(t)}{t}\right]},$$

$$\lambda_{12}(t_1 \mid t_2) = \frac{\frac{\partial}{\partial t_1}\left[\log GV_{12}(t_1 \mid t_2)\right]}{\log\left[\frac{GV_{12}(t_1 \mid t_2)}{t_1}\right]}$$

and

$$\lambda_{21}(t_2 \mid t_1) = \frac{\frac{\partial}{\partial t_2}\left[\log GV_{21}(t_2 \mid t_1)\right]}{\log\left[\frac{GV_{21}(t_2 \mid t_1)}{t_2}\right]}. \qquad (1.44)$$

1.4 Information Measures

The development idea of entropy by Shannon (1948) provided the beginning of a separate branch of learning namely the 'Theory of Information'. Even though an axiomatic foundation to this concept was laid down by Shannon, this measure was independently developed by Wiener (1948). Initial works related to Shannon's entropy was centered on characterization of different models.

1.4.1 Entropy Function

The concept of entropy is extensively used in literature as a quantitative measure of uncertainty associated with random phenomena. Some of the commonly used measures to characterize or to compare the aging process of the units are the failure rate and mean residual life function. Various characteristic properties of these functions can be seen in Swartz (1973), Esary and Marshall (1974), Buchanan and Singpurwalla (1977), Mukherjee and Roy (1986), Guess and Proschan (1988), Ruiz and Navarro (1994), Tan et al. (1999), Asadi (1999), Lin (2003), Gupta and Kirmani (2004). But highly uncertain components or systems are inherently not reliable. One measure of this uncertainty is the Shannon's (1948) information measure defined as

$$H(f) = -\int_0^\infty f(x)\log f(x)dx = -E\left[\log f(X)\right]. \tag{1.45}$$

Low entropy distributions are more concentrated and hence more informative than high entropy distributions. The concept of entropy has been extended to bivariate and multivariate case by several authors (see Cover and Thomas (1991), Darbellay and Vajda (2000), Nadarajah and Zografos (2005), Zografos and Nadarajah (2005)).

The Shannon's entropy finds applications in diverse fields. In communication theory an aspect of interest is the flow of transmission in some network where information is carried from a transmitter to receiver. This may be sending of messages by telegraph, flow of electricity, and visual communications from artist to viewers etc.

1.4.2 Residual Entropy Function

In many reliability and survival analysis problems the current age of an item under study must be taken in to account by information measures of the lifetime distribution. It is common knowledge that highly uncertain components or systems are inherently not reliable. At the stage of designing a system, when there is enough information regarding the deterioration, wear of a component parts, factors and levels are prepared based on this information. This type of information was usually obtained through hazard rate function or mean residual life function. However, in order to have a better design the stability of the component parts should also be taken into account together with deterioration. Capturing effects of the age t of an individual or an item under study on the information about the remaining lifetime is important in many applications.

As an example consider the case where X is the age at death of an insured person who purchases the policy at age t. The length of time between X and t, together with the age at which insurance is purchased, is crucial for pricing life insurance products for individuals in various age groups.

Ebrahimi and Pellery (1995) and Ebrahimi (1996) has modified the Shannon's (1948) entropy function by taking the age into account, which measures the expected uncertainty contained in the conditional density of $X-t$ given $X>t$ about the predictability of the remaining lifetime of the component is defined as

$$H(f,t)=-\int_{t}^{\infty}\frac{f(x)}{\bar{F}(t)}\log\left(\frac{f(x)}{\bar{F}(t)}\right)dx. \tag{1.46}$$

Which is equivalent to

$$\int_t^{\infty} f(x)\log f(x)dx = [1-F(t)]\log[1-F(t)] - [1-F(t)]H(f,t). \qquad (1.47)$$

Differentiating with respect to t, we obtain

$$f(t)\log f(t) = f(t)\left[1-H(f,t)\right] + \log[1-F(t)] + [1-F(t)]\frac{d}{dt}H(f,t).$$

The failure rate (1.10) verifies

$$\lambda(t)\left[H(f,t)-1+\log\lambda(t)\right] = \frac{d}{dt}H(f,t). \qquad (1.48)$$

Under the assumption that $H(f,t)$ is a non-decreasing function, Belzunce et al. (2004) showed that $\lambda(t)$ is unique positive solution of the equation

$$g(y) = y\left[H(f,t)-1+\log\lambda(t)\right] - \frac{d}{dt}H(f,t) = 0. \qquad (1.49)$$

Thus a non-decreasing $H(f,t)$ uniquely determines the underlying distribution.

In particular if $\frac{d}{dt}H(f,t)=0$ then solving (1.49), we have $\lambda(t)=e^{1-H(f,t)}$ which characterizes the exponential distribution. For further characteristic properties of $H(f,t)$, we refer to Ebrahimi (1996), Nair and Rajesh (1998), Sankaran and Gupta (1999), Asadi and Ebrahimi (2000), Belzunce et al. (2004). A dynamic generalized information measure is given in Asadi et al. (2005).

1.4.3 Bivariate Residual Entropy Function

Now if (X_1, X_2) represents the lifetime of the components or system, then the joint residual lifetime distribution at ages $t_1, t_2 \ge 0$ is the conditional (truncated) distribution denoted by

$$F(x_1, x_2, t_1, t_2) = P\left[X_1 \le x_1, X_2 \le x_2 \mid X_1 > t_1, X_2 > t_2\right]. \qquad (1.50)$$

The residual density function will be denoted by

$$f(x_1, x_2, t_1, t_2) = \frac{f(x_1, x_2)}{\bar{F}(t_1, t_2)} \text{ for } x_1 > t_1, x_2 > t_2. \qquad (1.51)$$

The residual entropy function has been extended to the bivariate case by Ebrahimi et al. (2007) for an absolutely continuous distribution function as

$$H(X_1, X_2, t_1, t_2) = H\left[f(x_1, x_2, t_1, t_2)\right]$$

$$= -\int_{t_2}^{\infty}\int_{t_1}^{\infty} f(x_1, x_2, t_1, t_2)\log\left[f(x_1, x_2, t_1, t_2)\right]dx_1 dx_2 \qquad (1.52)$$

$$= \log \bar{F}(t_1, t_2) - \frac{1}{\bar{F}(t_1, t_2)}\int_{t_2}^{\infty}\int_{t_1}^{\infty} f(x_1, x_2)\log f(x_1, x_2)dx_1 dx_2 .$$

The residual entropy (1.52) measures the uncertainty of the remaining lifetime when the ages of components are t_1 and t_2 respectively.

A representation of the residual entropy corresponding to the equation (1.52) in terms of marginal and conditional entropies is obtained as

$$H(X_1, X_2, t_1, t_2) = H(X_i, t_1, t_2) + H\left(X_j \mid X_i, t_1, t_2\right), i \neq j = 1, 2,$$

where $H(X_i, t_1, t_2)$ is the marginal entropy which measures the uncertainty of the marginal residual density of X_i given $X_1 > t_1$, $X_2 > t_2$ and $H\left(X_j \mid X_i, t_1, t_2\right)$ is the conditional residual life entropy which quantifies the uncertainty about X_j on average when we know $X_i, i \neq j.$

If $r(x_1, x_2) = \dfrac{f(x_1, x_2)}{\bar{F}(x_1, x_2)}$, denote the Basu's (1971) bivariate failure rate, then

$$H(X_1, X_2, t_1, t_2) = -\frac{1}{\bar{F}(t_1, t_2)}\int_{t_2}^{\infty}\int_{t_1}^{\infty} f(x_1, x_2)\log r(x_1, x_2)dx_1 dx_2 \qquad (1.53)$$

which doesn't uniquely determine the distribution unless the conditional entropies are given.

1.5. Measures of Inequality

As is customary in most statistical analysis, extend of variation in income is represented in terms of certain summary measures. A measure of income inequality is designed to provide an index that can abridge the variations prevailing among the units in a population. The population measures are Lorenz function and Gini index. The concepts and ideas from reliability theory have been extensively used to study measures of inequality. Chandra and Singpurwalla (1981) pointed out a few relationships between some notions that are common to reliability theory and economics in the context of measuring inequality. These aspects were further carried out by Klefsjo (1984).

1.5.1. The Lorenz Curve

Although there had been many attempts to provide measures of income inequality in the nineteenth century, the first major development in this area can be attributed to the work of M.O. Lorenz in 1905. The Lorenz curve is an important tool for the measurement of income inequality. To compare the distribution of income of a country at different periods of time or of different countries at the same time, the Lorenz curve, takes into account the changes in income and population.

Let X be a non-negative random variable admitting an absolutely continuous distribution function $F(x)$, with finite mean μ. The Lorenz curve $L(p)$ of X is defined in terms of two parametric equations in x namely

$$p = F(x) = \int_0^x f(t)dt \tag{1.54}$$

and

$$L(p) = F_1(x) = \frac{1}{\mu}\int_0^x t\ f(t)dt\,. \tag{1.55}$$

$L(p)$ determined by (1.55) is called ‘standard Lorenz curve’. p can be interpreted as the proportion of individuals having income less than or equal to x. It follows from (1.55) that the Lorenz curve is the first moment distribution function of p. It may be noticed that both p and $L(p)$ lies between zero and one, and the Lorenz curve being the graphical representation of incomes by plotting a curve with co-ordinates $\left(p, L(p)\right)$ in the unit square. $L(p)$ can be interpreted as the proportion of the total wealth owned by the poorest p^{th} fraction of the population. One can easily verify the following properties.

(i) $L(0)=0, L(1)=1, L(p)$ is continuous and strictly increasing in $(0,1)$, as $L'(p)=\frac{1}{\mu}x$, which is greater than zero.

(ii) $L(p)$ is twice differentiable and is strictly convex on $(0,1)$ as $L''(p)=\frac{1}{\mu f(x)}>0$.

Gastwirth (1971) relaxed the assumption that the distribution function $F(x)$ is absolutely continuous and defined the Lorenz curve $L(p)$ by

$$L(p)=\frac{1}{\mu}\int_0^p Q(t)dt, 0 \le p \le 1 \qquad (1.56)$$

where

$$Q(x)=\inf\{x : F(x) \ge x\}$$

is the quantile function. When $F(x)$ is absolutely continuous, $Q(x)$ is the inverse function of $F(x)$ and (1.56) is the solution for $L(p)$ obtained from (1.54) and (1.55).

For statistical or administrative reasons, many surveys of income are truncated at the lower end of the income range. Much of the data on income comes from income tax returns and most countries have a threshold below which no tax is levied. Someone known or suggested to have a low income is much unlikely to file a tax return than a person with high earnings. Hence the importance of studying inequality measures of truncated distribution upon the various measures of income inequality had been a theme of interest among researchers. Bhattacharya (1963) showed that the Lorenz curve of a left truncated distribution is Pareto. The right truncated case was studied by Moothathu (1986), who showed that Lorenz curve is independent of the point of truncation if and only if the distribution is a power function distribution. Ord et al. (1983) examined the effects of truncation upon some derived measures of inequality and it is shown that only for the Pareto distribution, the measures are invariant with respect to truncation.

The truncation form of the Lorenz curve for a non-negative random variable T admitting an absolutely continuous distribution function with $E\left[X\right]<\infty$ is defined by

$$p_x(t)=\frac{1}{\overline{F}(x)}\int_x^t dF(y) \tag{1.57}$$

and

$$L(p_x(t))=\frac{\int_x^t y\,dF(y)}{\int_x^\infty y\,dF(y)},\ x>t \tag{1.58}$$

where x is the truncation point. Then the plot of $\left(p_x(t),L(p_x(t))\right)$ gives the graphical representation of incomes beyond the truncation point x in the unit square.

Many authors have extended the concept of Lorenz curve to higher dimensions. Taguchi (1972 a) defined the concentration surface of a two dimensional random vector $\left(X_1,X_2\right)$ having a continuous density function $f(x_1,x_2)$ and having non-zero finite mean values μ_1 and μ_2 for X_1 and X_2 respectively, by the following implicit function.

$$L\left(p_1,p_2,p_3\right)=0 \tag{1.59}$$

where

$$p_1=\int_{-\infty}^{x_2}\int_{-\infty}^{x_1} f(u,v)dudv\,,$$

$$p_2=\frac{1}{\mu_1}\int_{-\infty}^{x_2}\int_{-\infty}^{x_1} u\ f(u,v)dudv$$

and

$$p_3=\frac{1}{\mu_2}\int_{-\infty}^{x_2}\int_{-\infty}^{x_1} v\ f(u,v)dudv\,. \tag{1.60}$$

He proved that the transformation (1.60) provides a one-to-one correspondence between (X_1, X_2) and (p_1, p_2, p_3). Hence the concentration surface defined by (1.59) can always be expressed as a single valued explicit function.

$$p_3 = L(p_1, p_2). \tag{1.61}$$

Taguchi (1972 b) extended the notions of concentration surface to complete surface, which he called as the Lorenz manifold. In Arnold (1983) a parametric representation of the Lorenz curve is given as

$$L(u,v) = \frac{\int_0^{x_1}\int_0^{x_2} \xi\eta f_{12}(\xi,\eta) d\xi d\eta}{E[X_1 X_2]} \tag{1.62}$$

where f_{12} denotes the joint income density and $f_i, i = 1,2$ denotes the marginals corresponding to the non-negative random variables X_1 and X_2 respectively. Here $U = \int_0^{x_1} f_1(\xi) d\xi$ and $V = \int_0^{x_2} f_2(\eta) d\eta$. Koshevoy (1995) provides a definition in higher dimensions in terms of the Lorenz zonoids and the inequality measures for multivariate distributions are given in Koshevoy and Mosler (1996), Arnold (2005).

1.5.2. The Gini Index

The Gini index is another popular inequality measure defined in terms of geometric features of the Lorenz curve. For a non-negative random variable with distribution function $F(x)$ and a finite mean μ, the Gini index (Gini (1912)) is defined in terms of mean difference as

$$G = \frac{1}{2\mu} \int\int |x - y| dF(x) dF(y). \tag{1.63}$$

As a function of Lorenz curve it can also be defined as (Frosini (1988)) twice the area between the Lorenz curve and the diagonal segment joining the points $(0,0)$ and $(1,1)$. That is

$$G = 1 - 2\int_0^1 L(p)dp \tag{1.64}$$

or

$$G = 1 - 2\int_0^\infty F(x)dF_1(x)\,. \tag{1.65}$$

The line segment joining the points $(0,0)$ and $(1,1)$ is known as line of equal distribution or egalitarian line. The value of G lies between 0 and 1, with $G = 0$ representing perfect equality and $G = 1$ representing perfect inequality. The Gini index is also referred to literature under the name coefficient of concentration, Lorenz concentration ratio and Gini coefficient.

Chakrabarthy (1982) points out that the analysis and criticism of Gini-index and Lorenz curve constitute a major part of the growing literature on inequality. He also stated that Lorenz curve and Gini index has remained the most powerful tool in the analysis of size distribution of income, both empirical and theoretical.

The truncation form of the Gini index was considered by several authors. For a non-negative continuous random variable X admitting an absolutely continuous distribution function with $E\left[X\right] < \infty$, Ord et al. (1983) considered the truncated form of Gini index defined by

$$G(t) = 2\int_t^\infty F(x,t)dF_1(x,t) - 1 \tag{1.66}$$

where $F(x,t)$ is the distribution function of $X_1(t) = X \mid X > t$ and $F_1(x,t)$ is the first moment distribution given by

$$F_1(x,t) = \frac{\int_t^x \frac{y\,f(y)}{\bar{F}(t)}dy}{\int_t^\infty \frac{y\,f(y)}{\bar{F}(t)}dy}\,.$$

It is established that (1.66) is truncation invariant if and only if X follows the Pareto type I distribution. The Gini index has also been extended to higher dimensions. Mosler (2002) extended it as the volume of the Lorenz zoniod and call it the Gini zonoid index. Weymark (2004) describes parameterized family of multivariate generalized Gini indeces. Multivariate extension of Gini index can also be seen in Koshevoy and Mosler (1997). A Gini index for truncated bivariate distribution was proposed by Sathar et al. (2007) which was consistent with that of Ord et al. (1983) for the univariate case and is defined as follows.

For a non-negative random vector (X_1, X_2) admitting an absolutely continuous distribution function, the bivariate Gini index for the truncated distribution is defined as the vector

$$G(t_1,t_2)=\left(G_1(t_1,t_2),G_2(t_1,t_2)\right) \tag{1.67}$$

where

$$G_1(t_1,t_2)=2\int_{t_1}^{\infty} F(x_1,t_1,t_2)\,dF_1(x_1,t_1,t_2)-1 \tag{1.68}$$

and

$$G_2(t_1,t_2)=2\int_{t_2}^{\infty} F(x_2,t_1,t_2)\,dF_2(x_2,t_1,t_2)-1. \tag{1.69}$$

where

$$F(x_1,t_1,t_2)=\int_{t_1}^{x_1}\frac{f(y_1 \mid X_2>t_2)}{\bar{F}(t_1 \mid X_2>t_2)}dy_1,$$

$$F(x_2,t_1,t_2)=\int_{t_2}^{x_2}\frac{f(y_2 \mid X_1>t_1)}{\bar{F}(t_2 \mid X_1>t_1)}dy_2,$$

$$F_1(x_1,t_1,t_2)=\frac{\int_{t_1}^{x_1} y_1\frac{f(y_1 \mid X_2>t_2)}{\bar{F}(t_1 \mid X_2>t_2)}dy_1}{\int_{t_1}^{\infty} y_1\frac{f(y_1 \mid X_2>t_2)}{\bar{F}(t_1 \mid X_2>t_2)}dy_1}$$

and

$$F_2\left(x_2,t_1,t_2\right)=\frac{\int_{t_2}^{x_2} y_2\frac{f\left(y_2\mid X_1>t_1\right)}{\bar{F}\left(t_2\mid X_1>t_1\right)}dy_2}{\int_{t_2}^{\infty} y_2\frac{f\left(y_2\mid X_1>t_1\right)}{\bar{F}\left(t_2\mid X_1>t_1\right)}dy_2}.$$

Recently, in connection with their study on ordering and asymptotic properties of residual income distribution, Belzunce et al. (1998) introduced a measure of income gap ratio among the rich, defined by

$$\begin{aligned}\beta(t)&=1-\frac{t}{E\left[X\mid X>t\right]}\\&=1-\frac{t}{v(t)}.\end{aligned}\tag{1.70}$$

or

$$\beta(t)=\frac{m(t)}{t+m(t)}.\tag{1.71}$$

where $m(t)$ and $v(t)$ are defined in (1.13) and (1.16) respectively.

1.6 Multivariate Pareto Distributions

Multivariate extensions of the Pareto models were basically approached so that important characteristics of the univariate Pareto, such as appropriate density shapes, univariate marginals, appropriate dependence structure and characteristic properties are extended to higher dimensions.

It is also well known that multivariate generalizations of univariate distributions may lead to various functional forms for the survival function. The multivariate extension of the univariate Pareto distribution first appeared in Mardia (1962). In this paper he has given the mathematical formulation of the bivariate Pareto distribution. Mardia’s bivariate Pareto I and II distribution are given by

$$f(x_1, x_2) = a(a+1)(\theta_1\theta_2)^{a+1}(\theta_2 x_1 + \theta_1 x_2 - \theta_1\theta_2)^{-(a+2)}; \qquad x_1 \geq \theta_1 > 0,\ x_2 \geq \theta_2 > 0,\ a > 0 \tag{1.72}$$

and

$$f(x_1, x_2) = \frac{a_1 a_2}{(1-\rho^2) x_1 x_2} \left\{ \left(\frac{\theta_1}{x_1}\right)^{a_1}, \left(\frac{\theta_2}{x_2}\right)^{a_2} \right\}^{1/(1-\rho^2)} I_0\left(\frac{2\rho\sqrt{a_1 a_2 \log\left(\frac{x_1}{\theta_1}\right)\log\left(\frac{x_2}{\theta_2}\right)}}{(1-\rho^2)} \right); x_1 \geq \theta_1,\ x_2 \geq \theta_2. \tag{1.73}$$

These models have been extensively used. Jupp and Mardia (1982) obtained characterization results for the bivariate Pareto model and showed that every multivariate distributions whose mean exists is determined by its mean residual life time. Krishnan (1985) has used the bivariate Pareto distribution given in (1.72) to model the crude birth rate and crude death/infant mortality rate, thus revealing its usefulness in demographic studies. Characteristic properties of the models can be seen in Mardia (1962), Malik and Trudel (1985), Xekalaki and Dimaki (2004).

Arnold (1983) has pointed three basic methods of generating a bivariate Pareto distribution of the fourth kind. The first one is from the mixture of Weibull and gamma distribution. The second method is by using the transformation of exponential distribution and the third is the method of trivariate reduction. The survival functions of the bivariate Pareto distribution corresponding to Arnold (1983) are as follows.

$$\bar{F}(x_1,x_2)=\left[1+\left(\frac{x_1-\mu_1}{\sigma_1}\right)^{1/\gamma_1}+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^{1/\gamma_2}+\right.$$
$$\left.\lambda\max\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^{1/\gamma_1},\left(\frac{x_2-\mu_2}{\sigma_2}\right)^{1/\gamma_2}\right\}\right]^{-\alpha};x_1\geq\mu_1,x_2\geq\mu_2. \tag{1.74}$$

and

$$\bar{F}(x_1,x_2)=\left\{1+\left(\frac{\max(x_1,x_2)-\mu}{\sigma}\right)^{1/\gamma}\right\}^{-\alpha_3}\left\{\left(\frac{x_1-\mu}{\sigma}\right)^{1/\gamma}\right\}^{-\alpha_1}$$
$$\left\{\left(\frac{x_2-\mu}{\sigma}\right)^{1/\gamma}\right\}^{-\alpha_2};x_1\geq\mu,x_2\geq\mu. \tag{1.75}$$

Arnold (1983) further studied the properties of these models and Yeh (1994) obtained characterization results for the bivariate Pareto IV distributions.

Now if the conditional densities $f(x_1 \mid x_2)$ and $f(x_2 \mid x_1)$ be members of Pareto II family of distributions, then the joint density of the conditionally specified bivariate Pareto distribution (Arnold (1987, 1989)) is

$$f(x_1,x_2)\propto(\lambda_0+\lambda_1x_1+\lambda_2x_2+\lambda_3x_1x_2)^{-(\alpha+1)};x_1>0,x_2>0, \tag{1.76}$$
$$\lambda_0>0,\lambda_1,\lambda_2>0 \text{ and } \lambda_3\geq0.$$

The case $\lambda_3=0$ leads to Mardia's bivariate Pareto distribution of second kind with $\alpha>1$. Clearly (1.76) represents a general density for which both conditional densities are Pareto II. Also Arnold et al. (1992) have further generalized the conditionally specified bivariate Pareto distribution of second kind in (1.76) by specifying that both $f(x_1 \mid x_2)$ and $f(x_2 \mid x_1)$ be beta density of second kind leading to the bivariate density function,

$$f(x_1,x_2)\propto\frac{x_1^{a-1}x_2^{a-1}}{(\lambda_0+\lambda_1x_1+\lambda_2x_2+\lambda_3x_1x_2)^{(a+b)}};x_1,x_2>0, \tag{1.77}$$
$$a,b>0,\lambda_0,\lambda_1,\lambda_2>0,\lambda_3\geq0.$$

Arnold et al. (1993) have provided a three dimensional plots and contour plots for the joint probability density function. The characteristic properties of this model are discussed in Arnold et al. (1993), Wesolowski (1994) and Arnold (1995).

Lindley and Singpurwalla (1986) considered a two component system where for a given environment, the component lifetimes X_1 and X_2 are independently distributed as exponential. Assuming the environment effect to be modeled by a gamma distribution, they obtained a bivariate Pareto distribution as

$$f\left(x_1,x_2\right)=\lambda_1\lambda_2(a+1)(a+2)b^{(a+1)}\left(\lambda_1x_1+\lambda_2x_2+b\right)^{a+3}; \qquad (1.78)$$
$$x_1,x_2>0,a,b>0,\lambda_1,\lambda_2>0.$$

Nayak (1987) generalized Lindley and Singpurwalla's (1986) model to the multivariate case, studied their properties and has shown its usefulness in reliability theory. Bandhyopadhyay and Basu (1990) have considered a two component system which operates in a test environment consisting of shocks that leads to Marshall-Olkin type dependent bivariate Pareto model,

$$f\left(x_1,x_2\right)=\begin{cases}\dfrac{a(a+1)\theta_i\theta_j^{*}}{\left[1+\theta_ix_i+\theta_j^{*}x_j\right]^{(a+2)}}; 0<x_i<x_j\\[2ex] \dfrac{a\theta_3}{\left[1+\theta x\right]^{(a+1)}} \;; 0<x_1=x_2=x\end{cases}$$

$$\theta_i=\frac{\lambda_i}{b}, i=1,2,3;\ \theta_j^{*}=\theta_j+\theta_3,\ j=1,2;\ \theta=\theta_1+\theta_2+\theta_3.$$

Muliere and Scarcini (1987) proposed a bivariate Pareto distribution with joint survival function

$$\bar{F}\left(x_1,x_2\right)=\left(\frac{x_1}{\beta}\right)^{-\lambda_1}\left(\frac{x_2}{\beta}\right)^{-\lambda_2}\left\{\max\left(\frac{x_1}{\beta},\frac{x_2}{\beta}\right)\right\}^{-\lambda_0}; \beta\le\min\left(x_1,x_2\right)<\infty. \qquad (1.79)$$

This model shows similarity with Marshall-Olkin (1967) bivariate exponential distribution. Also characterization results for the model (1.79) can be seen in Veenus and Nair (1994).

In constructing (1.78), Lindley and Singpurwalla (1986) assumed independent exponential distribution. But in most of the real life situation, the independence assumption is not valid, because in many systems the component life length has a well-defined dependence structure. Sankaran and Nair (1993) has given a bivariate Pareto model by assuming the Gumbel (1960) bivariate exponential distribution instead of independent exponential as,

$$f\left(x_1, x_2\right) = p\left(p\left(a_1 + bx_2\right)\left(a_2 + bx_1\right) + a_1 a_2 - b\right) \left(1 + a_1 x_1 + a_2 x_2 + b x_1 x_2\right)^{-p-2}; \, x_1, x_2 > 0,$$

where the parameters satisfy the conditions $p, a_1, a_2 > 0$ and $0 \le b \le (p+1)a_1 a_2$. They also obtained characterization results and showed the application of the model in reliability studies. Hanagal (1996) has given characterization results for the bivariate Pareto model (1.79), also extended the concept of dullness property to the bivariate case. He obtained the maximum likelihood estimates of the parameters and their asymptotic multivariate normal distributions. Yeh (2004 a) and Yeh (2004 b) has developed multivariate generalized Pareto distributions corresponding to Marshall-Olkin (1967) exponential distribution. The survival function of the generalized Marshall-Olkin type multivariate Pareto IV (Yeh (2004 a)) is given by

$$\bar{F}_{\underline{x}}(\underline{x}) = \prod_{i=1}^{m} \left\{1 + \left(\frac{x_i - \mu_i}{\sigma_i}\right)^{1/r_i}\right\}^{-\alpha_i} \prod_{i=1<}^{m-1} \prod_{j=2}^{m} \left\{\max\left(1 + \left(\frac{x_i - \mu_i}{\sigma_i}\right)^{1/r_i}\right)^{-\alpha_i}\right\} \left(1 + \left(\frac{x_i - \mu_i}{\sigma_i}\right)^{1/r_i}\right)^{-\alpha_i} \cdots \left\{\max_{1 \le i \le m}\left(1 + \left(\frac{x_i - \mu_i}{\sigma_i}\right)^{1/r_i}\right)^{-\alpha_i}\right\} \quad (1.80)$$

for $\underline{x} > \underline{\mu}$, $0 < \underline{\sigma} \le \underline{\mu} < \infty$, where $\sigma_i, r_i, \alpha_i > 0$, $\underline{X} = (X_1, X_2, ..., X_m)$, $\underline{\sigma} = (\sigma_1, \sigma_2, ..., \sigma_m)$ $x_i \ge \sigma_i$, $\underline{\mu} = (\mu_1, \mu_2 ..., \mu_m)$. The Marshall – Olkin type I multivariate Pareto, $GMOP^{(m)}(I)$ is given as

$$\bar{F}_{\underline{X}}(\underline{x}) = \prod_{i=1}^{m}\left(\frac{x_i}{\sigma_i}\right)^{-\alpha_i} \prod_{1=i<j<2}^{m-1} \prod^{m} \left\{ \max\left(\left(\frac{x_i}{\sigma_i}\right)^{-\alpha_i}, \left(\frac{x_j}{\sigma_j}\right)^{-\alpha_j}\right)\right\} \cdots \max_{1\le i\le m}\left(\frac{x_i}{\sigma_i}\right)^{-\alpha_i}; \underline{x} > \underline{\sigma}. \tag{1.81}$$

These distributions are termed as generalized Marshall-Olkin type Pareto distribution since they are obtained from transformed multivariate Marshall-Olkin (1967) exponential distribution. The bivariate and multivariate Pareto distributions proposed by Muliere and Scarcini (1987), Veenus and Nair (1994), Hanagal (1996) are special cases of (1.80). In Yeh (2004 a, b), Asha and Jagathnath (2006) several characterization of the generalized Marshall-Olkin type Pareto are obtained.

Nadarajah (2008) has developed a bivariate Pareto distribution to model the drought durations and drought intensity. He also derived explicit expressions for the probability density function, cumulative density functions and the moments for the sums, product and ratios of the bivariate random variables, which are highly applicable in drought modeling. Navarro et al. (2008) modeled the life lengths of the units in a system as a symmetric bivariate Pareto II distribution. The survival function of the model is given by

$$\bar{F}(x_1, x_2) = \left(1 + ax_1 + ax_2 + ba^2 x_1 x_2\right)^{-c}; x_1, x_2 \ge 0, a, c > 0, 0 \le b \le c+1. \tag{1.82}$$

The basic reliability properties for the series and parallel systems for the model (1.82) are carried out. Chiragiev and Landsman (2009) introduced two multivariate models whose marginals have different shape parameters and a more flexible dependence structure. They also discussed regression and a measure of dependence for their models.

Even though the Pareto distribution was developed as a model of income distribution, the literature review itself reveals the significance of the model in other areas. The application of the model is wide spread. It is identified as a useful distribution in modeling social, economic, financial, actuarial, demographic, survival, reliability and hydrological data. A more detailed review on Pareto distributions can be seen in Arnold (1983) and Kotz et al. (2000).

1.7 Present Study

Usually in statistical modeling, one comes up with a model taking into consideration the physical aspects of the situation. In the present study we consider a physical situation similar to the one mentioned in Freund (1961). In Freund (1961) a load sharing dependence in considered. Though load sharing dependence is popular in reliability studies this dependence is common in socio-economic situations also.

The earliest work on load sharing models is due to Daniels (1945) and Rosen (1964). They observed that yarns and cables in a bundle fail only when the last fibre (wire) in the bundle breaks. A bundle of fibres can be considered as a parallel system subject to a constant tensile load. After one fibre breaks yarn bundles or the remaining unbroken fibres gets extra load. This is the equal load share rule under which the load of the failed component is distributed equally among the remaining working components.

Apart from textile industry such models arises in manufacturing where a part can be considered failed only when the entire set of welded joints that holds the part together fails. However, the failure of one or two joints can increase the stress on remaining joints.

Freund (1961) has designed a bivariate exponential model for the life testing of a two component parallel system which incorporates the above said load-sharing dependence. Lindley and Singpurwalla (1986) proposed a bivariate Pareto for two component system with independent exponential lifetime that could accommodate changes in the common operating environment. However if the operating environment affects both the components differently, then assuming a gamma distribution to model the environment, the above system reduces to assuming two independent Pareto II distribution instead of exponential distribution. Taking into consideration these assumptions we develop a new model which is also same as a transformed Freund bivariate exponential distribution.

After the introductory chapter, where we have pointed out the relevance and scope of the study along with review of literature, the remaining chapters are devoted to some new results. The present work is organized into six chapters. In Chapter two, a bivariate Pareto model is derived which is applicable for a two component parallel system that could accommodate changes in the test environment. The genesis and properties of the model is discussed. The estimation of the parameters in the model is obtained using three methods namely, maximum likelihood estimation, principle of maximum entropy, and method of moments. A simulation study is carried out and comparison of the three methods is done. The model is fitted for a real data set using bivariate Kolmogorov-Smirnov test.

Characterization results for the bivariate Pareto model introduced in the previous chapter are obtained and are discussed in Chapter three. A general class of bivariate Pareto minima is introduced. Characterizations using the bivariate dullness property and its variants, truncation equivalent to rescaling and truncation invariant property, which are meaningful in income distribution context, are obtained. However when considering the reliability characteristics in the bivariate case we need to understand that they are extended to the bivariate or multivariate cases in more than one way. Thus there is a need to choose an appropriate extension which reflects the ageing characteristics of the bivariate distribution. This choice depends on the dependency enjoyed by the distribution. In this study the bivariate failure rate of Cox (1972) is an apt choice as it reflects the ageing characteristics of the distribution by taking into consideration the load sharing features present. The mean residual life function associated with the Cox failure rate is the dynamic mean residual life function (Shaked and Shanthikumar

(1991)). This chapter concludes with the section which presents characterizations of bivariate Pareto II $\left(BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)\right)$ and bivariate Pareto I $\left(BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)\right)$ distributions based on these measures.

In the previous chapter we advocated the use of Cox's failure rate for model's exhibiting a load sharing dependence. Similar arguments form a reasonable motivation to formulate a definition of information measure for the residual life distribution corresponding to a bivariate distribution with load sharing dependence. Accordingly in the fourth chapter we propose a new measure of bivariate residual entropy function. The properties of the bivariate residual entropy function are discussed. How the bivariate dullness property and its variant, the bivariate lack of memory property manifest on bivariate residual entropy function is explored. Characterizations of lifetime models using this measure are carried out.

In Chapter Five, bivariate inequality measures such as Lorenz function and Gini index are introduced. The properties of the measures and their importance in income related studies are established. Chandra and Singpurwalla (1981) have given an interpretation of Lorenz curve and Gini index for lifetime data, which extended the application of these measures in reliability studies. Motivated by this, in this Chapter, the relationship of the Lorenz function and Gini index with bivariate mean residual life function is also established. Further characterization results for bivariate lifetime models are also obtained. Illustration of Lorenz curve is made by using the real data set given in Kim and Kvam (2004).

In Chapter six, we have given a uniform representation of the Freund bivariate exponential distribution (1961) and its transformation. We showed how this representation can be used to infer on the total failure rate of each distribution having this representation, once we know their uniform translates. We also characterize this uniform representation by what we define as general dullness property. It is further shown that this property implies the bivariate lack of memory property (Marshall and Olkin (1967)) for the Freund's bivariate exponential distribution. It implies the bivariate dullness property (Veenus and Nair (1994), Hanagal (1996), Yeh (2004 a,b)) for the $BP\,I\left(1,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution. It also implies a characterization for the bivariate Weibull distribution. This thesis concludes after discussing few open problems for the future course of work.

Chapter 2

A Bivariate Pareto Model

2.1 Introduction

In the previous chapter, several bivariate Pareto distributions were discussed. A more appealing and realistic form of dependence is exhibited when performance of a functioning component is affected by how the other components within the system are operating. Real examples of such dependent system include software and hardware systems, power plants, automobiles and material subject to failure due to crack growth (Hollander and Pena (1995), Kvam and Pena (2005)). Such systems studied in engineering and the physical sciences are typically based on load-share models. Load-share models dictate that component failure rates are defined on the operating status of the other system components and the effective system structure function. Daniels (1945) originally adopted this model to describe how the strain on yarn fibres increases as individual fibres within a bundle break. Freund (1961) formalized the probability theory for a bivariate exponential load-share model. Drummond et al. (2000) carried out a study in a vertebrate species showing that selective deaths due to

Some results in this chapter is published in Asha and Jagathnath (2008).

food shortage result in surviving offspring receiving an increased share of an undiminished food supply. They observed that littermates of domestic rabbit after individual pups died, the total daily milk weight obtained by the litter continued to be the same.

An important element of the load-share model is the rule that governs how failure rate changes after some component in the system fail. An equal load-share rule implies the existence of a constant system load distributed equally among working components. In the local load-sharing rule a failed components load is transferred to adjacent components. A general monotone load-sharing rule assumes only that load or any individual component is non-decreasing as other items fail. Lynch (1999) characterized some relationship between failure rates and the load-share rule based on monotone load-share. Apart from physical sciences and engineering sciences, this load-sharing dependence is found in economic data too.

In the present study we propose a bivariate Pareto distribution which is designed in particular for load-sharing dependence. No particular load-sharing rule is assumed here. The proposed model is different from all the models existing in the literature in the sense that they do not have Pareto marginal but have mixture of Pareto distribution as marginals. However, this distribution enjoys the bivariate extensions of several properties of the univariate Pareto that justifies it being called a bivariate Pareto. Unlike the other bivariate Paretos, a multivariate extension of the proposed distribution is not straightforward. Every extension would need an explicit definition of dependency among the components. In the section that follows the mathematical formulation of the model is proposed. The model obtained is same as the one obtained by transforming the Freund (1961) bivariate exponential distribution. In the third section, the distributional properties of the proposed distribution are considered.

In the fourth section its relationship with some well known distributions is discussed. In the fifth section, the estimation of parameters of the model is considered. In the last section we illustrate the model by using a data given in Kim and Kvam (2004).

2.2 Model and its Properties

Mathematically the model can be formulated as follows. Let T_1 and T_2 be two independent random variables with density function conditioned on η, λ_i given as

$$f_i(t_i \mid \eta, \lambda_i) = \eta\lambda_i e^{-\eta\lambda_i t_i},\ t_i > 0, \eta, \lambda_i > 0, i = 1, 2. \tag{2.1}$$

Typically one can think of T_i's to be component lifetimes of a parallel system comprising of two components. Here η is the effect of the operating environment. Assume that this system operates in an environment, which does not change over time but may be different from the test environment. Also assume that the environment of the system influences both the components independently and when one of the components fails the other work with a change in parameter.

Let the effect of the operating environment be described by the distribution function $G(\eta)$, then

$$f_i(t_i \mid \lambda_i) = \int \eta\lambda_i e^{-\eta\lambda_i t_i} g(\eta) d\eta, t_i > 0, \lambda_i > 0, i = 1, 2, \tag{2.2}$$

where $g(\eta) = \dfrac{dG(\eta)}{d\eta}$.

An easily flexible and analytically tractable model for $G(\eta)$ is the gamma distribution with density

$$g(\eta) = \frac{m^p}{\Gamma p} e^{-m\eta} \eta^{p-1}; m > 0,\ p > 0. \tag{2.3}$$

It is well known and requires only simple calculations to reveal that

$$f_i(t_i) = \int_0^\infty \frac{m^p \eta \lambda_i}{\Gamma p} e^{-\eta\lambda_i t_i} e^{-m\eta} \eta^{p-1} d\eta$$

$$= \frac{p\lambda_i}{m}\left(1+\frac{\lambda_i t_i}{m}\right)^{-(p+1)} ; \; m>0, \; p>0, \lambda_i>0, t_i>0 , \quad (2.4)$$

is the Pareto II distribution. This distribution was used by Lomax (1954) to fit data in business failure. This distribution has been identified by Pickands (1975) as one of the distributions that can approximate residual life distributions.

So now let T_1 and T_2 be two independent Pareto type II random variables with survival function $\bar{F}_i(t_i)$ specified by

$$\bar{F}_i(t_i) = \left(1+\frac{t_i-\mu}{\sigma}\right)^{-\alpha_i} ; \; t_i > \mu, \; i = 1,2$$

where μ, the location parameter, is real and σ is positive.

Assume that the first component fails at $T_1 = t_1$, then the time until the subsequent failure of the second component is given by the random variable T_2' which now follows a Pareto model with renewed parameter α_2' for some $\alpha_2' > 0$ with $T_2' > t_1$. Similarly if the second component fails first, then the random variable T_1' represents the remaining life of the first component. The random variable T_1' now follows a Pareto model with renewed parameter α_1' for some $\alpha_1' > 0$ with $T_1' > t_2$. If (X_1, X_2) denote the component lifetime of the system, then

$$X_1 = T_1, \; X_2 = T_1' + T_2 \;\text{ if } T_1 < T_2$$

$$X_1 = T_1 + T_2', \; X_2 = T_2 \;\text{ if } T_1 > T_2 .$$

Then the joint density of (X_1, X_2) is now derived from

$$f(x_1, x_2) = \begin{cases} f_Z(x_1) P[X_1 < X_2] f(x_2 \mid Z = x_1, X_2 > X_1); \; 0 < x_1 < x_2 \\ f_Z(x_2) P[X_1 > X_2] f(x_1 \mid Z = x_2, X_1 > X_2); \; 0 < x_2 < x_1 \end{cases} \quad (2.5)$$

where $f_Z(.)$ is the density of $Z = min(X_1, X_2)$.

Here,

$$f_Z(x) = -\frac{d}{dx}\left(1+\frac{x-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2)}$$

$$f_Z(x) = \frac{(\alpha_1+\alpha_2)}{\sigma}\left(1+\frac{x-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2+1)}.$$

Also,

$$\begin{aligned} P[X_1 > X_2] &= \int_\mu^\infty P[X_1 > x_2 \mid X_2 = x_2] P[X_2 = x_2] dx_2 \\ &= \int_\mu^\infty \left(1+\frac{x_2-\mu}{\sigma}\right)^{-\alpha_1} \frac{\alpha_2}{\sigma}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-(\alpha_2+1)} dx_2 \\ &= \frac{\alpha_2}{\alpha_1+\alpha_2}. \end{aligned}$$

Similarly,

$$P[X_1 < X_2] = \frac{\alpha_1}{\alpha_1+\alpha_2}.$$

And

$$\begin{aligned} f(x_1 \mid Z = x_2, X_1 > X_2) &= \frac{\frac{\alpha_1'}{\sigma}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-(\alpha_1'+1)}}{\left(1+\frac{x_2-\mu}{\sigma}\right)^{-\alpha_1'}} \\ &= \frac{\alpha_1'}{\sigma}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-(\alpha_1'+1)}\left(1+\frac{x_2-\mu}{\sigma}\right)^{\alpha_1'} ;\ x_2 < x_1. \end{aligned}$$

Similarly,

$$f(x_2 \mid Z = x_1, X_2 > X_1) = \frac{\alpha_2'}{\sigma}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-(\alpha_2'+1)}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-\alpha_2'} ;\ x_1 < x_2.$$

Thus the joint density is obtained as

$$f(x_1, x_2) = \begin{cases} \frac{\alpha_1\alpha_2'}{\sigma^2}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_2'+1)}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-(\alpha_2'+1)} ;\ \mu < x_1 < x_2 \\ \frac{\alpha_2\alpha_1'}{\sigma^2}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_1'+1)}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-(\alpha_1'+1)} ;\ \mu < x_2 < x_1 \end{cases} \tag{2.6}$$

$\sigma > 0,\ \alpha_i, \alpha_i' > 0,\ \alpha_1+\alpha_2 \neq \alpha_i',\ i = 1, 2$ and real μ.

In our discussion we consider $\mu > 0$. Ruling out simultaneous failures, the random variables X_1 and X_2 are now no longer independent, the dependency between X_1 and X_2 is essentially such that the failure of one of the component changes the shape parameter of the other. In the rest of the paper we denote the bivariate Pareto in equation (2.6) as $BP\ II(\mu, \sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2')$. When $\mu = \sigma$, the distribution reduces to

$$f(x_1, x_2) = \begin{cases} \dfrac{\alpha_1 \alpha_2'}{\sigma^{-(\alpha_1+\alpha_2)}} x_1^{-(\alpha_1+\alpha_2-\alpha_2'+1)} x_2^{-(\alpha_2'+1)}; & \sigma < x_1 < x_2 \\ \dfrac{\alpha_2 \alpha_1'}{\sigma^{-(\alpha_1+\alpha_2)}} x_2^{-(\alpha_1+\alpha_2-\alpha_1'+1)} x_1^{-(\alpha_1'+1)}; & \sigma < x_2 < x_1 \end{cases} \quad (2.7)$$

$$\sigma > 0,\ \alpha_i, \alpha_i' > 0, \alpha_1 + \alpha_2 \neq \alpha_i',\ i = 1, 2.$$

The distribution (2.7) will be referred to as bivariate Pareto I $\left(BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)\right)$ distribution here after.

The graphical representation of the model $BP\ II(\mu, \sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2')$ is shown in the following figure.

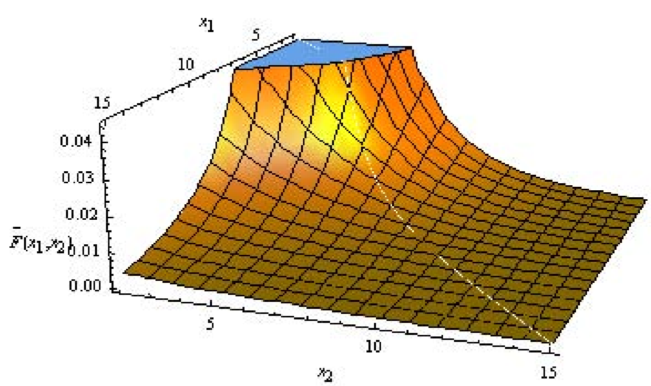


Figure 2.1 Plot of the bivariate survival function of a BP II distribution with $\mu = 2,\ \sigma = 3,\ \alpha_1 = 2,\ \alpha_2 = 2.5,\ \alpha_1' = 4$ and $\alpha_2' = 5$.

Unlike the other bivariate Paretos discussed in Chapter one, these bivariate Paretos exhibits marginals, which are mixture of Pareto distribution. They are obtained as

$$f_i(x_i) = \frac{\alpha_{3-i}\alpha_i'}{(\alpha_1+\alpha_2-\alpha_i')\sigma}\left(1+\frac{x_i-\mu}{\sigma}\right)^{-(\alpha_i'+1)} + \frac{(\alpha_i-\alpha_i')(\alpha_1+\alpha_2)}{(\alpha_1+\alpha_2-\alpha_i')\sigma}\left(1+\frac{x_i-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2+1)};$$

$$0<\mu<x_i,\ \sigma>0,\ \alpha_i,\alpha_i'>0,\ \alpha_1+\alpha_2\neq\alpha_i',\ i=1,2. \tag{2.8}$$

and

$$E[X_i] = \frac{\sigma}{(\alpha_1+\alpha_2-\alpha_i')}\left[\frac{\alpha_{3-i}}{(\alpha_i'-1)}+\frac{(\alpha_i-\alpha_i')}{(\alpha_1+\alpha_2-1)}\right]+\mu \tag{2.9}$$

which is finite whenever $\alpha_i' > 1$ and $\alpha_1+\alpha_2>1, \alpha_1+\alpha_2\neq\alpha_i',\ i=1,2.$

The survival function of (X_1, X_2) is obtained as

$$\bar{F}(x_1,x_2) = \begin{cases} \frac{\alpha_1}{(\alpha_1+\alpha_2-\alpha_2')}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_2')}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-\alpha_2'} \\ \qquad +\frac{(\alpha_2-\alpha_2')}{(\alpha_1+\alpha_2-\alpha_2')}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2)};\ \mu\le x_1\le x_2 \\ \frac{\alpha_2}{(\alpha_1+\alpha_2-\alpha_1')}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_1')}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-\alpha_1'} \\ \qquad +\frac{(\alpha_1-\alpha_1')}{(\alpha_1+\alpha_2-\alpha_1')}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2)};\ \mu\le x_2\le x_1 \end{cases} \tag{2.10}$$

$$\sigma>0,\ \alpha_i,\alpha_i'>0,\ \alpha_1+\alpha_2\neq\alpha_i',\ i=1,2.$$

and the marginal survival function is obtained as

$$\overline{F}_i(x_i)=\frac{\alpha_{3-i}}{\left(\alpha_1+\alpha_2-\alpha_i'\right)}\left(1+\frac{x_i-\mu}{\sigma}\right)^{-\alpha_i'}+\left[1-\frac{\alpha_{3-i}}{\alpha_1+\alpha_2-\alpha_i'}\right]\left(1+\frac{x_i-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2)};$$

$$\mu\le x_i,\sigma>0,\ \alpha_i,\alpha_i'>0,\ \alpha_1+\alpha_2\ne\alpha_i',\ i=1,2.$$

The representation for the product moment $E\left[X_1{}^r X_2{}^s\right]$ is obtained in the following theorem.

Theorem 2.1 If (X_1,X_2) has a joint density function $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ and if $r,\ s<\alpha_1+\alpha_2-\alpha_i'$ and $r,s<\alpha_i',i=1,2$, then

$$E\left[X_1{}^r X_2{}^s\right]=H_1\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_2'\right)-H_2\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_2'\right)+G_1\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1'\right)-G_2\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1'\right)$$

for all $r,\ s\ge1$, where

$$H_1\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_2'\right)=\frac{\mu^{(r+s-\alpha_1-\alpha_2)}\sigma^{-(\alpha_1+\alpha_2+2)}}{\left(\alpha_1+\alpha_2-\alpha_2'-r\right)^2\left(\alpha_2'-s\right)^2}\sum_{j=0}^{\infty}\frac{\left(\alpha_2'+1\right)_j\left(\alpha_2'-s\right)_j}{j!\left(\alpha_2'-s+1\right)_j}\left(\frac{\mu-\sigma}{\mu}\right)^j$$
$$\sum_{j=0}^{\infty}\frac{\left(\alpha_1+\alpha_2-\alpha_2'+1\right)_j\left(\alpha_1+\alpha_2-\alpha_2'-r\right)_j}{j!\left(\alpha_1+\alpha_2-\alpha_2'-r+1\right)_j}\left(\frac{\mu-\sigma}{\mu}\right)^j,$$

$$H_2\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_2'\right)=$$
$$\left(\sigma-\mu\right)^{-(\alpha_2'+1)}\sigma^{-(\alpha_1+\alpha_2-2\alpha_2')}\sum_{j=0}^{\infty}\frac{\left(\alpha_1+\alpha_2-\alpha_2'+1\right)_j\left(\alpha_1+\alpha_2-\alpha_2'-r\right)_j}{j!\left(\alpha_1+\alpha_2-\alpha_2'-r+1\right)_j}$$
$$\int_{\mu}^{\infty}x_2^{(r+s-\alpha_1-\alpha_2-\alpha_2')}\left(1+\frac{x_2}{\sigma-\mu}\right)^{-(\alpha_2'+1)}\left(\frac{x_2-\sigma}{x_2}\right)^j dx_2,$$

$$G_1\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1'\right)=\frac{\mu^{(r+s-\alpha_1-\alpha_2)}\sigma^{-(\alpha_1+\alpha_2+2)}}{\left(\alpha_1+\alpha_2-\alpha_1'-s\right)^2\left(\alpha_1'-r\right)^2}\sum_{j=0}^{\infty}\frac{\left(\alpha_1'+1\right)_j\left(\alpha_1'-r\right)_j}{j!\left(\alpha_1'-r+1\right)_j}\left(\frac{\mu-\sigma}{\mu}\right)^j$$
$$\sum_{j=0}^{\infty}\frac{\left(\alpha_1+\alpha_2-\alpha_1'+1\right)_j\left(\alpha_1+\alpha_2-\alpha_1'-s\right)_j}{j!\left(\alpha_1+\alpha_2-\alpha_1'-s+1\right)_j}\left(\frac{\mu-\sigma}{\mu}\right)^j$$

and

$$G_2\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1'\right)=$$
$$\left(\sigma-\mu\right)^{-\left(\alpha_1'+1\right)}\sigma^{-\left(\alpha_1+\alpha_2-2\alpha_1'\right)}\sum_{j=0}^{\infty}\frac{\left(\alpha_1+\alpha_2-\alpha_1'+1\right)_j\left(\alpha_1+\alpha_2-\alpha_1'-r\right)_j}{j!\left(\alpha_1+\alpha_2-\alpha_1'-r+1\right)_j}$$
$$\int_{\mu}^{\infty}x_1^{\left(r+s-\alpha_1-\alpha_2-\alpha_1'\right)}\left(1+\frac{x_1}{\sigma-\mu}\right)^{-\left(\alpha_1'+1\right)}\left(\frac{x_1-\sigma}{x_1}\right)^{j}dx_1.$$

Proof

From equation (2.6), we have

$$E\left[X_1{}^rX_2{}^s\right]=\int_{\mu}^{\infty}\int_{\mu}^{x_2}x_1^rx_2^s\left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_2'+1\right)}dx_1dx_2+$$
$$\int_{\mu}^{\infty}\int_{\mu}^{x_1}x_1^rx_2^s\left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1'+1\right)}dx_2dx_1$$

$$=\int_{\mu}^{\infty}\left\{\int_{\mu}^{\infty}x_1^rx_2^s\left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_2'+1\right)}dx_1\right.$$
$$\left.-\int_{x_2}^{\infty}x_1^rx_2^s\left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_2'+1\right)}dx_1\right\}dx_2$$
$$+\int_{\mu}^{\infty}\left\{\int_{\mu}^{\infty}x_1^rx_2^s\left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1'+1\right)}dx_2\right.$$
$$\left.-\int_{x_1}^{\infty}x_1^rx_2^s\left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1'+1\right)}dx_2\right\}dx_1$$

$$= \int_{\mu}^{\infty} x_2^s \left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_2'+1\right)} dx_2 \int_{\mu}^{\infty} x_1^r \left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)} dx_1$$

$$-\int_{\mu}^{\infty} x_2^s \left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_2'+1\right)} dx_2 \int_{x_2}^{\infty} x_1^r \left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)} dx_1$$

$$+\int_{\mu}^{\infty} x_1^r \left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1'+1\right)} dx_1 \int_{\mu}^{\infty} x_2^s \left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)} dx_2$$

$$-\int_{\mu}^{\infty} x_1^r \left(1+\frac{x_1-\mu}{\sigma}\right)^{-\left(\alpha_1'+1\right)} dx_1 \int_{x_1}^{\infty} x_2^s \left(1+\frac{x_2-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)} dx_2$$

$$=\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_2'+1\right)} \int_{\mu}^{\infty} x_2^{(s+1)-1}\left(1+\frac{x_2}{\sigma-\mu}\right)^{-\left(\alpha_2'+1\right)} dx_2$$

$$\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)} \int_{\mu}^{\infty} x_1^{(r+1)-1}\left(1+\frac{x_1}{\sigma-\mu}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)} dx_1$$

$$-\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_2'+1\right)} \int_{\mu}^{\infty} x_2^{(s+1)-1}\left(1+\frac{x_2}{\sigma-\mu}\right)^{-\left(\alpha_2'+1\right)} dx_2$$

$$\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)} \int_{x_2}^{\infty} x_1^{(r+1)-1}\left(1+\frac{x_1}{\sigma-\mu}\right)^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)} dx_1$$

$$+\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_1'+1\right)} \int_{\mu}^{\infty} x_1^{(r+1)-1}\left(1+\frac{x_1}{\sigma-\mu}\right)^{-\left(\alpha_1'+1\right)} dx_1$$

$$\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)} \int_{\mu}^{\infty} x_2^{(s+1)-1}\left(1+\frac{x_2}{\sigma-\mu}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)} dx_2$$

$$-\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_1'+1\right)} \int_{\mu}^{\infty} x_1^{(r+1)-1}\left(1+\frac{x_1}{\sigma-\mu}\right)^{-\left(\alpha_1'+1\right)} dx_1$$

$$\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)} \int_{x_1}^{\infty} x_2^{(s+1)-1}\left(1+\frac{x_2}{\sigma-\mu}\right)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)} dx_2$$

Making use of Gauss hypergeometric function and equation 3.194.2* in Gradshteyn et al. (1994), we get the following

$$=\frac{\sigma^{(\alpha_1+\alpha_2+2)}\mu^{(r+s-\alpha_1-\alpha_2)}}{\left(\alpha_1+\alpha_2-\alpha_2'-r\right)^2\left(\alpha_2'-s\right)^2}\,{}_2F_1\left(\alpha_2'+1,\alpha_2'-s;\alpha_2'-s+1;\frac{\mu-\sigma}{\mu}\right)$$
$${}_2F_1\left(\alpha_1+\alpha_2-\alpha_2'+1,\alpha_1+\alpha_2-\alpha_2'-r;\alpha_1+\alpha_2-\alpha_2'-r+1;\frac{\mu-\sigma}{\mu}\right)$$

$$-\int_{\mu}^{\infty}\sigma^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)}x_2^{\left(r+s-\alpha_1-\alpha_2-\alpha_2'\right)}\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_2'+1\right)}\left(1+\frac{x_2}{\sigma-\mu}\right)^{-\left(\alpha_2'+1\right)}$$
$${}_2F_1\left(\alpha_1+\alpha_2-\alpha_2'+1,\alpha_1+\alpha_2-\alpha_2'-r;\alpha_1+\alpha_2-\alpha_2'-r+1;\frac{x_2-\sigma}{x_2}\right)dx_2$$

$$+\frac{\sigma^{(\alpha_1+\alpha_2+2)}\mu^{(r+s-\alpha_1-\alpha_2)}}{\left(\alpha_1+\alpha_2-\alpha_2'-s\right)^2\left(\alpha_2'-r\right)^2}\,{}_2F_1\left(\alpha_1'+1,\alpha_1'-r;\alpha_1'-r+1;\frac{\mu-\sigma}{\mu}\right)$$
$${}_2F_1\left(\alpha_1+\alpha_2-\alpha_1'+1,\alpha_1+\alpha_2-\alpha_1'-s;\alpha_1+\alpha_2-\alpha_1'-s+1;\frac{\mu-\sigma}{\mu}\right)$$

$$-\int_{\mu}^{\infty}\sigma^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)}x_2^{\left(r+s-\alpha_1-\alpha_2-\alpha_1'\right)}\left(\frac{\sigma-\mu}{\sigma}\right)^{-\left(\alpha_1'+1\right)}\left(1+\frac{x_1}{\sigma-\mu}\right)^{-\left(\alpha_1'+1\right)}$$
$${}_2F_1\left(\alpha_1+\alpha_2-\alpha_1'+1,\alpha_1+\alpha_2-\alpha_1'-s;\alpha_1+\alpha_2-\alpha_2'-s+1;\frac{x_1-\sigma}{x_1}\right)dx_1$$

$$^{*}\int_{u}^{\infty}\frac{x^{\mu-1}dx}{(1+\beta x)^{v}}=\frac{u^{\mu-v}}{\beta^{v}(v-\mu)}\,{}_2F_1\left(v,v-\mu;v-\mu+1;-\frac{1}{\beta u}\right);\left[\operatorname{Re}v>\operatorname{Re}\mu\right]$$

$$= \frac{\mu^{(r+s-\alpha_1-\alpha_2)}\sigma^{-(\alpha_1+\alpha_2+2)}}{(\alpha_1+\alpha_2-\alpha_2'-r)^2(\alpha_2'-s)^2}\sum_{j=0}^{\infty}\frac{(\alpha_2'+1)_j(\alpha_2'-s)_j}{j!(\alpha_2'-s+1)_j}\left(\frac{\mu-\sigma}{\mu}\right)^j$$

$$\sum_{j=0}^{\infty}\frac{(\alpha_1+\alpha_2-\alpha_2'+1)_j(\alpha_1+\alpha_2-\alpha_2'-r)_j}{j!(\alpha_1+\alpha_2-\alpha_2'-r+1)_j}\left(\frac{\mu-\sigma}{\mu}\right)^j$$

$$-(\sigma-\mu)^{-(\alpha_2'+1)}\sigma^{-(\alpha_1+\alpha_2-2\alpha_2')}\sum_{j=0}^{\infty}\frac{(\alpha_1+\alpha_2-\alpha_2'+1)_j(\alpha_1+\alpha_2-\alpha_2'-r)_j}{j!(\alpha_1+\alpha_2-\alpha_2'-r+1)_j}$$

$$\int_{\mu}^{\infty}x_2^{(r+s-\alpha_1-\alpha_2-\alpha_2')}\left(1+\frac{x_2}{\sigma-\mu}\right)^{-(\alpha_2'+1)}\left(\frac{x_2-\sigma}{x_2}\right)^j dx_2$$

$$+\frac{\mu^{(r+s-\alpha_1-\alpha_2)}\sigma^{-(\alpha_1+\alpha_2+2)}}{(\alpha_1+\alpha_2-\alpha_1'-s)^2(\alpha_1'-r)^2}\sum_{j=0}^{\infty}\frac{(\alpha_1'+1)_j(\alpha_1'-r)_j}{j!(\alpha_1'-r+1)_j}\left(\frac{\mu-\sigma}{\mu}\right)^j$$

$$\sum_{j=0}^{\infty}\frac{(\alpha_1+\alpha_2-\alpha_1'+1)_j(\alpha_1+\alpha_2-\alpha_1'-s)_j}{j!(\alpha_1+\alpha_2-\alpha_1'-s+1)_j}\left(\frac{\mu-\sigma}{\mu}\right)^j$$

$$-(\sigma-\mu)^{-(\alpha_1'+1)}\sigma^{-(\alpha_1+\alpha_2-2\alpha_1')}\sum_{j=0}^{\infty}\frac{(\alpha_1+\alpha_2-\alpha_1'+1)_j(\alpha_1+\alpha_2-\alpha_1'-s)_j}{j!(\alpha_1+\alpha_2-\alpha_1'-s+1)_j}$$

$$\int_{\mu}^{\infty}x_1^{(r+s-\alpha_1-\alpha_2-\alpha_1')}\left(1+\frac{x_1}{\sigma-\mu}\right)^{-(\alpha_1'+1)}\left(\frac{x_1-\sigma}{x_1}\right)^j dx_1$$

$$= H_1(\mu,\sigma,\alpha_1,\alpha_2,\alpha_2') - H_2(\mu,\sigma,\alpha_1,\alpha_2,\alpha_2') + G_1(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1') - G_2(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1').$$

Hence the result.

Corollary 2.1 If (X_1, X_2) has a joint density function $BP\,I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$, then

$$E\left[X_1^{\,r}X_2^{\,s}\right] = \frac{\sigma^{r+s}}{(\alpha_1+\alpha_2-s-r)}\left[\frac{\alpha_1\alpha_2'}{(\alpha_2'-s)}+\frac{\alpha_1'\alpha_2}{(\alpha_1'-r)}\right]$$

and $E\left[X_1^{\,r}X_2^{\,s}\right] < \infty$, when $\alpha_1+\alpha_2 > s+r, \alpha_2' > s, \alpha_1' > r.$

Corollary 2.2 If (X_1, X_2) has a joint density function $BP\ II(\mu, \sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2')$ then for $r = s = 1$,

$$E[X_1 X_2] = \mu^2 + \mu\sigma\left[\frac{\alpha_1}{(\alpha_2'-1)(\alpha_1+\alpha_2-\alpha_2')} + \frac{\alpha_2}{(\alpha_1'-1)(\alpha_1+\alpha_2-\alpha_1')}\right] -$$

$$\frac{2\sigma}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2-1)}\left[\mu + \frac{\sigma}{(\alpha_1+\alpha_2-2)}\right]\left[\frac{\alpha_1\alpha_2'}{(\alpha_1+\alpha_2-\alpha_2')} + \frac{\alpha_2\alpha_1'}{(\alpha_1+\alpha_2-\alpha_1')}\right]$$

$$+\frac{\mu\sigma}{(\alpha_1+\alpha_2-1)}\left[\frac{\alpha_1}{(\alpha_1+\alpha_2-\alpha_2')} + \frac{\alpha_2}{(\alpha_1+\alpha_2-\alpha_1')}\right] - \frac{\sigma^2}{(\alpha_1+\alpha_2-1)(\alpha_1+\alpha_2-2)}$$

$$\left[\frac{\alpha_1\alpha_2'}{(\alpha_1+\alpha_2-\alpha_2')(\alpha_1+\alpha_2-\alpha_2'-1)} + \frac{\alpha_2\alpha_1'}{(\alpha_1+\alpha_2-\alpha_1')(\alpha_1+\alpha_2-\alpha_1'-1)}\right] +$$

$$\left[\frac{\sigma^2\alpha_1\alpha_2'}{(\alpha_2'-1)(\alpha_1+\alpha_2-\alpha_2')(\alpha_1+\alpha_2-\alpha_2'-1)} + \right.$$

$$\left.\frac{\sigma^2\alpha_2\alpha_1'}{(\alpha_1'-1)(\alpha_1+\alpha_2-\alpha_1')(\alpha_1+\alpha_2-\alpha_1'-1)}\right]. \qquad (2.11)$$

Corollary 2.3 If (X_1, X_2) has a joint probability density function $BP\ I(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2')$ then for $r = s = 1$,

$$E[X_1 X_2] = \frac{\sigma^2}{(\alpha_1+\alpha_2-2)}\left[\frac{\alpha_1\alpha_2'}{(\alpha_2'-1)} + \frac{\alpha_1'\alpha_2}{(\alpha_1'-1)}\right].$$

Further,

$$V[X_i] = \frac{\sigma^2}{(\alpha_1+\alpha_2-\alpha_i')}\left\{\frac{\alpha_{3-i}}{(\alpha_i'-1)}\left[\frac{2}{(\alpha_i'-2)} - \frac{\alpha_{3-i}}{(\alpha_1+\alpha_2-\alpha_i')(\alpha_i'-1)}\right] + \frac{(\alpha_i-\alpha_i')}{(\alpha_1+\alpha_2-1)}\right.$$

$$\left.\left[\frac{2}{(\alpha_1+\alpha_2-2)} - \frac{2\alpha_{3-i}}{(\alpha_1+\alpha_2-\alpha_i')(\alpha_i'-1)} - \frac{(\alpha_i-\alpha_i')}{(\alpha_1+\alpha_2-\alpha_i')(\alpha_1+\alpha_2-1)}\right]\right\}, i = 1, 2,$$

(2.12)

from which

$$Cov(X_1,X_2)=\frac{\mu\sigma}{(\alpha_1+\alpha_2-1)}\left[\frac{\alpha_1}{\left(\alpha_1+\alpha_2-\alpha_2'\right)}+\frac{\alpha_2}{\left(\alpha_1+\alpha_2-\alpha_1'\right)}\right]-$$

$$\frac{2\sigma}{(\alpha_1+\alpha_2)(\alpha_1+\alpha_2-1)}\left[\mu+\frac{\sigma}{(\alpha_1+\alpha_2-2)}\right]\left[\frac{\alpha_1\alpha_2'}{\left(\alpha_1+\alpha_2-\alpha_2'\right)}+\frac{\alpha_2\alpha_1'}{(\alpha_1+\alpha_2-\alpha_1')}\right]$$

$$-\frac{\sigma^2}{(\alpha_1+\alpha_2-1)(\alpha_1+\alpha_2-2)}\left[\frac{\alpha_1\alpha_2'}{(\alpha_1+\alpha_2-\alpha_2')(\alpha_1+\alpha_2-\alpha_2'-1)}+\right.$$

$$\left.\frac{\alpha_2\alpha_1'}{(\alpha_1+\alpha_2-\alpha_1')(\alpha_1+\alpha_2-\alpha_1'-1)}\right]-\frac{\sigma^2}{(\alpha_1+\alpha_2-\alpha_1')(\alpha_1+\alpha_2-\alpha_2')}$$

$$\left[\frac{\alpha_1}{(\alpha_2'-1)}+\frac{(\alpha_2-\alpha_2')}{(\alpha_1+\alpha_2-1)}\right]\left[\frac{\alpha_2}{(\alpha_1'-1)}+\frac{(\alpha_1-\alpha_1')}{(\alpha_1+\alpha_2-1)}\right]$$

$$-\frac{\mu\sigma}{(\alpha_1+\alpha_2-1)}\left[\frac{(\alpha_1-\alpha_1')}{\left(\alpha_1+\alpha_2-\alpha_2'\right)}+\frac{(\alpha_2-\alpha_2')}{\left(\alpha_1+\alpha_2-\alpha_1'\right)}\right]. \tag{2.13}$$

Evidently as $\alpha_i\to\alpha_i'$, $i=1,2$, $Cov\ (X_1,X_2)\to 0.$

The correlation coefficient of the bivariate Pareto models discussed in chapter one, namely, Mardia's (1962) bivariate Pareto of first kind, Mardia's (1962) bivariate Pareto of second kind, the conditionally specified bivariate Pareto have positive correlation and Arnold et al. (1993) Model II yield non-positive correlation, while the correlation coefficient for the model $BP\ II(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ admits both positive and negative correlations.

When $\alpha_1=\alpha_2=\alpha$ and $\alpha_1'=\alpha_2'=0$ and $\mu=\sigma=1$, the correlation coefficient ρ is given by

$$\rho=-\frac{1}{1-\left[2\alpha(1-\alpha)\right]^{-1}}. \tag{2.14}$$

The expression for $E\left[X_i\mid x_{3-i}\right]$ is generally non-linear and is obtained as

$$E\left[X_i \mid x_{3-i}\right]=\frac{1}{f_{3-i}\left(x_{3-i}\right)}\left\{\frac{\alpha_i\alpha'_{3-i}}{\left(\alpha_1+\alpha_2-\alpha'_{3-i}\right)}\left(1+\frac{x_{3-i}-\mu}{\sigma}\right)^{-(\alpha'_{3-i}+1)}\right.$$
$$\left(\frac{\mu}{\sigma}+\frac{1}{\left(\alpha_1+\alpha_2-\alpha'_{3-i}-1\right)}\right)+\frac{x_{3-i}}{\sigma}\left(1+\frac{x_{3-i}-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2+1)}$$
$$\left(\alpha_{3-i}-\frac{\alpha_i\alpha'_{3-i}}{\left(\alpha_1+\alpha_2-\alpha'_{3-i}\right)}\right)+\left(1+\frac{x_{3-i}-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2)} \tag{2.15}$$
$$\left.\left(\frac{\alpha_{3-i}}{\left(\alpha'_i-1\right)}-\frac{\alpha_i\alpha'_{3-i}}{\left(\alpha_1+\alpha_2-\alpha'_{3-i}\right)\left(\alpha_1+\alpha_2-\alpha'_{3-i}-1\right)}\right)\right\}.$$

Evidently for $\alpha'_i=\alpha_i,\ i=1,2,\ E\left[X_i \mid x_{3-i}\right]=\left(\mu+\dfrac{\sigma}{\alpha_i-1}\right)$.

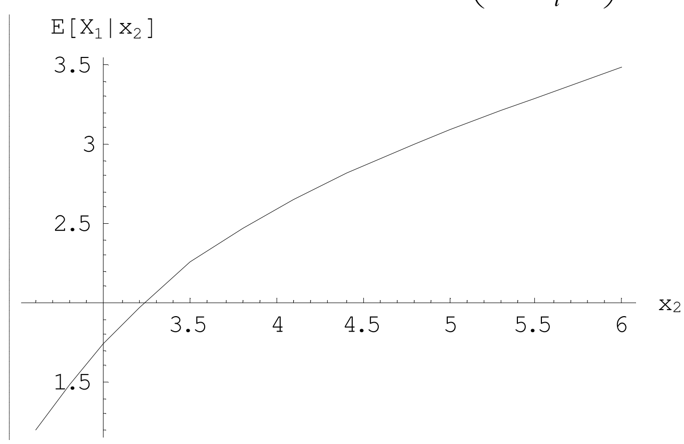


Figure 2.2 Plot of regression line of X_1 given X_2 of a $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha'_1,\alpha'_2\right)$ distribution with $\mu=2.5$, $\sigma=3$, $\alpha_1=5$, $\alpha_2=6$, $\alpha_1'=5.5$ and $\alpha_2'=4$.

Further if (X_1,X_2) has $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha'_1,\alpha'_2\right)$ distribution then $Z=min(X_1,X_2)$ is a univariate Pareto given by

$$f(z)=\frac{\left(\alpha_1+\alpha_2\right)}{\sigma}\left(1+\frac{z-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-1)};\ z>\mu. \tag{2.16}$$

It should be noted that the Marhall-Olkin type bivariate Pareto distribution (Muliere and Scarsini (1987), Veenus and Nair (1994), Hanagal (1996) and Yeh (2004 a,b)) also satisfies this property. In fact we can show that the absolute continuous part of the Marshal-Olkin type bivariate Pareto distribution (1.79) given by

$$\bar{F}_a(x_1,x_2)=\begin{cases}\dfrac{\lambda}{\lambda_1+\lambda_2}\left(\dfrac{x_1}{\sigma}\right)^{-(\lambda_1+\lambda_{12})}\left(\dfrac{x_2}{\sigma}\right)^{-\lambda_2}-\dfrac{\lambda_{12}}{\lambda_1+\lambda_2}\left(\dfrac{x_1}{\sigma}\right)^{-\lambda}; & \sigma<x_2<x_1\\ \dfrac{\lambda}{\lambda_1+\lambda_2}\left(\dfrac{x_2}{\sigma}\right)^{-(\lambda_2+\lambda_{12})}\left(\dfrac{x_1}{\sigma}\right)^{-\lambda_2}-\dfrac{\lambda_{12}}{\lambda_1+\lambda_2}\left(\dfrac{x_2}{\sigma}\right)^{-\lambda}; & \sigma<x_1<x_2\end{cases}\tag{2.17}$$

where $\lambda=\lambda_1+\lambda_2+\lambda_{12}$ is a particular case of the survival function of $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$. To see this observe that when $\lambda_1=\alpha_1+\alpha_2-\alpha_2'$, $\lambda_2=\alpha_1+\alpha_2-\alpha_1'$, $\lambda=\alpha_1+\alpha_2$, $\lambda_1+\lambda_{12}=\alpha_1'$ and $\lambda_2+\lambda_{12}=\alpha_2'$, the absolute continuous part of the survival function specified in (2.17) reduces to that of $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$.

2.3 Sums and Ratios for $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ Distribution

The statistics literature has seen many developments in the theory and applications of linear combinations and ratios of random variables (see Gupta and Nadarajah (2006)). In this section, we consider the distribution of $R=X_1+X_2$ and $W=\dfrac{X_1}{X_1+X_2}$ when X_1 and X_2 are Pareto variables with the joint probability density function given by $BP\ I\left(1,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$,

$$f(x_1,x_2)=\begin{cases}\alpha_1\alpha_2'\,x_1^{-(\alpha_1+\alpha_2-\alpha_2'+1)}x_2^{-(\alpha_2'+1)}; & 1<x_1<x_2\\ \alpha_2\alpha_1'\,x_2^{-(\alpha_1+\alpha_2-\alpha_1'+1)}x_1^{-(\alpha_1'+1)}; & 1<x_2<x_1\end{cases}.\tag{2.18}$$

$$\alpha_i,\alpha_i'>0,\ \alpha_1+\alpha_2\neq\alpha_i',\ i=1,2.$$

Theorem 2.2 If (X_1, X_2) are jointly distributed as (2.18) and $R = X_1 + X_2$, $W = \dfrac{X_1}{X_1 + X_2}$ then

$$f_R(r) = \alpha_1 \alpha_2' \, r^{-(\alpha_1+\alpha_2+3)} I_{[0,1/2]}(w, 1-w) + \alpha_2 \alpha_1' \, r^{-(\alpha_1+\alpha_2+3)} I_{[1/2,1]}(1-w, w)$$

where

$$I_{[0,\,1/2]}(w, 1-w) = \int_0^{1/2} w^{-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)} (1-w)^{-\left(\alpha_2'+1\right)} dw,$$

$$I_{[1/2,\,1]}(1-w, w) = \int_{1/2}^{1} (1-w)^{-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)} w^{-\left(\alpha_1'+1\right)} dw$$

and

$$f_W(w) = \begin{cases} \dfrac{\alpha_1 \alpha_2' \, 2^{-(\alpha_1+\alpha_2+2)}}{(\alpha_1+\alpha_2+2)} w^{-(\alpha_1+\alpha_2-\alpha_2'+1)} (1-w)^{-(\alpha_2'+1)}; \ 0 < w < \tfrac{1}{2} \\ \dfrac{\alpha_2 \alpha_1' \, 2^{-(\alpha_1+\alpha_2+2)}}{(\alpha_1+\alpha_2+2)} (1-w)^{-(\alpha_1+\alpha_2-\alpha_1'+1)} w^{-(\alpha_1'+1)}; \ \tfrac{1}{2} < w < 1 \end{cases}.$$

$$\alpha_i, \alpha_i' > 0, \alpha_1 + \alpha_2 \neq \alpha_i', \ i = 1, 2.$$

Proof

Let (X_1, X_2) follows a $BP\ I\left(1, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution specified as in (2.18). Consider the transformations, $R = X_1 + X_2$ and $W = \dfrac{X_1}{X_1 + X_2}$, then the corresponding Jacobian of transformation is obtained as $|J| = \dfrac{1}{r}$.

Now the joint density of (R, W) is obtained as

$$f(r, w) = \begin{cases} \alpha_1 \alpha_2' \, r^{-(\alpha_1+\alpha_2+3)} w^{-(\alpha_1+\alpha_2-\alpha_2'+1)} (1-w)^{-(\alpha_2'+1)}; \ 0 < w < \tfrac{1}{2} \\ \alpha_2 \alpha_1' \, r^{-(\alpha_1+\alpha_2+3)} (1-w)^{-(\alpha_1+\alpha_2-\alpha_1'+1)} w^{-(\alpha_1'+1)}; \ \tfrac{1}{2} < w < 1 \end{cases} \quad (2.19)$$

and $2 < r < \infty$.

The marginal density function for the random variable R is given by

$$f_R(r)=\alpha_1\alpha_2{}'\ r^{-(\alpha_1+\alpha_2+3)}\int_0^{1/2} w^{-\left(\alpha_1+\alpha_2-\alpha_2{}'+1\right)}\left(1-w\right)^{-\left(\alpha_2{}'+1\right)}dw+$$
$$\alpha_2\alpha_1{}'\ r^{-(\alpha_1+\alpha_2+3)}\int_{1/2}^{1}(1-w)^{-\left(\alpha_1+\alpha_2-\alpha_1{}'+1\right)}w^{-\left(\alpha_1{}'+1\right)}dw.$$

i.e.,

$$f_R(r)=\alpha_1\alpha_2{}'\ r^{-(\alpha_1+\alpha_2+3)}I_{[0,1/2]}\left(w,1-w\right)+\alpha_2\alpha_1{}'\ r^{-(\alpha_1+\alpha_2+3)}I_{[1/2,1]}\left(1-w,w\right).$$

Now the marginal density for the random variable W follows from (2.19) by integrating with respect to r,

$$f_W(w)=\begin{cases}\alpha_1\alpha_2' w^{-(\alpha_1+\alpha_2-\alpha_2'+1)}(1-w)^{-(\alpha_2'+1)}\displaystyle\int_2^\infty r^{-(\alpha_1+\alpha_2+3)}dr;\ 0<w<1/2\\ \alpha_2\alpha_1'\,(1-w)^{-(\alpha_1+\alpha_2-\alpha_1'+1)}w^{-(\alpha_1'+1)}\displaystyle\int_2^\infty r^{-(\alpha_1+\alpha_2+3)}dr;\ 1/2<w<1\end{cases}.$$

Thus we obtain the marginal density function of W as

$$f_W(w)=\begin{cases}\dfrac{\alpha_1\alpha_2'\,2^{-(\alpha_1+\alpha_2+2)}}{\left(\alpha_1+\alpha_2+2\right)}w^{-(\alpha_1+\alpha_2-\alpha_2'+1)}(1-w)^{-(\alpha_2'+1)};\ 0<w<1/2\\ \dfrac{\alpha_2\alpha_1'\,2^{-(\alpha_1+\alpha_2+2)}}{\left(\alpha_1+\alpha_2+2\right)}(1-w)^{-(\alpha_1+\alpha_2-\alpha_1'+1)}w^{-(\alpha_1'+1)};\ 1/2<w<1\end{cases}.$$

2.4 Transformed Exponential Variates

It is quite intuitive to recognize the relationship of $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution and the Freund bivariate exponential distribution. Consider the transformation $\left(1+\dfrac{X_i-\mu}{\sigma}\right)=e^{rY_i}$ for $\sigma,\ \mu,\ r>0.$

The corresponding Jacobian of transformation is

$$\left|J\right|=\frac{1}{\sigma^2r^2\left(1+\dfrac{x_1-\mu}{\sigma}\right)\left(1+\dfrac{x_2-\mu}{\sigma}\right)}.$$

Then density of $\left(Y_1,Y_2\right)$ is obtained as

$$f(y_1, y_2) = \begin{cases} \beta_1\beta_2'\ e^{-\beta_2' y_2 - (\beta_1+\beta_2-\beta_2')y_1}; 0 < y_1 < y_2 \\ \beta_2\beta_1'\ e^{-\beta_1' y_1 - (\beta_1+\beta_2-\beta_1')y_2}; 0 < y_2 < y_1 \end{cases};\ \beta_i,\ \beta_i' > 0,\ i = 1,2 \qquad (2.20)$$

which is the Freund bivariate exponential distribution with $\beta_i = r\alpha_i,\ \beta_i' = r\alpha_i',\ i = 1,2.$

Conversely it is quite straightforward to observe that if (Y_1, Y_2) is distributed as (2.20) then (X_1, X_2) has $BP\ II\left(\mu, \sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ with $\alpha_i = \frac{\beta_i}{r},\ \alpha_i' = \frac{\beta_i'}{r},\ i = 1,2.$ Hence, we have the following theorem.

Theorem 2.3 Let (X_1, X_2) be a random vector with $\left(1 + \frac{X_i - \mu}{\sigma}\right) = e^{rY_i}$ for $\sigma, \mu, r > 0.$ Then (X_1, X_2) is distributed as in (2.6) if and only if (Y_1, Y_2) is distributed as Freund bivariate exponential distribution specified by (2.20) with $\alpha_i = \frac{\beta_i}{r},\ \alpha_i' = \frac{\beta_i'}{r},\ i = 1,2.$

Corollary 2.4 Let (X_1, X_2) be a random vector with $\frac{X_i}{\sigma} = e^{rY_i}$. Then (X_1, X_2) is distributed as $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ if and only if (Y_1, Y_2) has Freund bivariate exponential distribution specified by (2.20) with $\alpha_i = \frac{\beta_i}{r},\ \alpha_i' = \frac{\beta_i'}{r},\ i = 1,2.$

Corollary 2.5 Let Y_1 and Y_2 be two independent standard exponential distributions, if

$$X_1 = \begin{cases} \sigma^{\left(\frac{Y_1}{\alpha_1 \ln\sigma} + 1\right)} & ;\ \alpha_2 Y_1 < \alpha_1 Y_2 \\ \sigma^{\left(\frac{Y_1}{\alpha_1' \ln\sigma} - \frac{(\alpha_1 - \alpha_1')Y_2}{\alpha_1'\alpha_2 \ln\sigma} + 1\right)} & ;\ \alpha_1 Y_2 < \alpha_2 Y_1 \end{cases}$$

and

$$X_2 = \begin{cases} \sigma^{\left(\frac{Y_2}{\alpha_2' \ln\sigma} - \frac{(\alpha_2-\alpha_2')Y_1}{\alpha_2'\alpha_1 \ln\sigma} + 1\right)}; & \alpha_2 Y_1 < \alpha_1 Y_2 \\ \sigma^{\left(\frac{Y_2}{\alpha_2 \ln\sigma} + 1\right)} & ; \ \alpha_1 Y_2 < \alpha_2 Y_1 \end{cases} \tag{2.21}$$

then the joint density of (X_1, X_2) is obtained as given in (2.7).

Proof

The joint density function for the independent standard exponential random variables Y_1 and Y_2 is given by

$$f\left(y_1, y_2\right) = e^{-y_1 - y_2}, \ y_1, y_2 > 0\,.$$

Now consider the transformation given in (2.21), and then the corresponding Jacobian of transformation is given by

$$\left|J\right| = \begin{cases} \frac{\alpha_1\alpha_2'}{x_1x_2}; \ x_1 < x_2 \\ \frac{\alpha_2\alpha_1'}{x_1x_2}; \ x_2 < x_1 \end{cases}.$$

Thus the joint density function for the bivariate random variable (X_1, X_2) is now

$$f(x_1, x_2) = \begin{cases} e^{-\alpha_1 \ln\left(\frac{x_1}{\sigma}\right) - \alpha_2' \ln\left(\frac{x_2}{\sigma}\right) - \left(\alpha_2 - \alpha_2'\right)\ln\left(\frac{x_1}{\sigma}\right)} \frac{\alpha_1\alpha_2'}{x_1x_2}; \ \sigma < x_1 < x_2 \\ e^{-\alpha_2 \ln\left(\frac{x_2}{\sigma}\right) - \alpha_1' \ln\left(\frac{x_1}{\sigma}\right) - \left(\alpha_1 - \alpha_1'\right)\ln\left(\frac{x_2}{\sigma}\right)} \frac{\alpha_2\alpha_1'}{x_1x_2}; \ \sigma < x_2 < x_1 \end{cases}.$$

or

$$f(x_1, x_2) = \begin{cases} \frac{\alpha_1\alpha_2'}{\sigma^{-(\alpha_1+\alpha_2)}} x_1^{-(\alpha_1+\alpha_2-\alpha_2'+1)} x_2^{-(\alpha_2'+1)}; \ \sigma < x_1 < x_2 \\ \frac{\alpha_2\alpha_1'}{\sigma^{-(\alpha_1+\alpha_2)}} x_2^{-(\alpha_1+\alpha_2-\alpha_1'+1)} x_1^{-(\alpha_1'+1)}; \ \sigma < x_2 < x_1 \end{cases}$$

the $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution specified as in (2.7).

2.5 Estimation of Parameters

To complete our discussion on the $BP\ I(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ and $BP\ II(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ distributions, we consider briefly the problem of estimation of parameters of the distributions. We confine our study to the method of maximum likelihood. Here we consider a two stage procedure where the scale parameter σ is estimated from the distribution of $min(X_1,X_2)$ which is a univariate Pareto distribution. Making use of this estimate of σ we estimate for $\alpha_1,\alpha_2,\alpha_1'$ and α_2'. Apart from considering the maximum likelihood estimate of σ, we consider alternative methods of estimation for σ namely, method of moments and method based on principle of maximum entropy (POME). It is seen that generally the maximum likelihood estimate is more efficient than the method of moments and method based on POME. This is seen to reflect on the estimates of $\alpha_1,\alpha_2,\alpha_1'$ and α_2'.

2.5.1 Estimation of Parameters of a $BP\ II(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ Distribution

In this section the estimation of the parameters in the bivariate Pareto II $BP\ II(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ model is considered for known σ. The density function is given by

$$f(x_1,x_2)=\begin{cases}\dfrac{\alpha_1\alpha_2'}{\sigma^2}\left(1+\dfrac{x_1-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_2'+1)}\left(1+\dfrac{x_2-\mu}{\sigma}\right)^{-(\alpha_2'+1)} & ;\ \mu<x_1<x_2\\ \dfrac{\alpha_2\alpha_1'}{\sigma^2}\left(1+\dfrac{x_2-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_1'+1)}\left(1+\dfrac{x_1-\mu}{\sigma}\right)^{-(\alpha_1'+1)} & ;\ \mu<x_2<x_1\end{cases}.$$

$\sigma>0,\ \alpha_i,\alpha_i'>0,\ \alpha_1+\alpha_2\neq\alpha_i',\ i=1,2$ and $\mu>0.$

Now consider a random sample of size n from a $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution with density function specified above. Then the likelihood function is given by

$$L = \alpha_1^{n_1} (\alpha_1')^{n_2} \alpha_2^{n_2} (\alpha_2')^{n_1} \sigma^{-2n} \prod_{i=1}^{n_1} \left(1+\frac{x_{1i}-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_2'+1)} \left(1+\frac{x_{2i}-\mu}{\sigma}\right)^{-(\alpha_2'+1)}$$
$$\prod_{i=1}^{n_2} \left(1+\frac{x_{2i}-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_1'+1)} \left(1+\frac{x_{1i}-\mu}{\sigma}\right)^{-(\alpha_1'+1)} \tag{2.22}$$

where n_i = number of $x_i < x_{3-i}$, $i = 1,2$ and $n = n_1 + n_2$.

Observe $Z = min(X_1, X_2)$ has a univariate Pareto II distribution specified by

$$f_Z(x) = \frac{\alpha}{\sigma}\left(1+\frac{x-\mu}{\sigma}\right)^{-\alpha} ; \mu < x, \alpha = \alpha_1 + \alpha_2 > 0, \sigma > 0.$$

Proceeding now as in the univariate case, for a fixed α the maximum likelihood estimate for μ is obtained from the distribution of Z as $\hat{\mu} = \underset{i}{min}\,\underset{j}{min}(X_{ij}),\ i = 1,2,\ j = 1,\ldots,n$. Substituting this $\hat{\mu}$ in (2.22) and for a known σ, the likelihood is now becomes

$$L = \alpha_1^{n_1} (\alpha_1')^{n_2} \alpha_2^{n_2} (\alpha_2')^{n_1} \sigma^{-2n} \prod_{i=1}^{n_1} \left(1+\frac{x_{1i}-\hat{\mu}}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_2'+1)} \left(1+\frac{x_{2i}-\hat{\mu}}{\sigma}\right)^{-(\alpha_2'+1)}$$
$$\prod_{i=1}^{n_2} \left(1+\frac{x_{2i}-\hat{\mu}}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_1'+1)} \left(1+\frac{x_{1i}-\hat{\mu}}{\sigma}\right)^{-(\alpha_1'+1)}$$

The estimates of the shape parameters are now obtained by taking the derivatives of the log-likelihood function with respect to the corresponding parameters and equating to zero. Thus we obtain,

$$\hat{\alpha}_1 = \frac{n_1}{\sum_{i=1}^{n_1} \log\left(1+\frac{x_{1i}-\hat{\mu}}{\sigma}\right) + \sum_{i=1}^{n_2} \log\left(1+\frac{x_{2i}-\hat{\mu}}{\sigma}\right)}, \tag{2.23}$$

$$\hat{\alpha}_2 = \frac{n_2}{\sum_{i=1}^{n_1} \log\left(1+\frac{x_{1i}-\hat{\mu}}{\sigma}\right) + \sum_{i=1}^{n_2} \log\left(1+\frac{x_{2i}-\hat{\mu}}{\sigma}\right)}, \tag{2.24}$$

$$\hat{\alpha}_1' = \frac{n_2}{\sum_{i=1}^{n_2} \log\left(1+\frac{x_{1i}-\hat{\mu}}{\sigma}\right) - \sum_{i=1}^{n_2} \log\left(1+\frac{x_{2i}-\hat{\mu}}{\sigma}\right)} \tag{2.25}$$

and

$$\hat{\alpha}_2' = \frac{n_1}{\sum_{i=1}^{n_2} \log\left(1+\frac{x_{2i}-\hat{\mu}}{\sigma}\right) - \sum_{i=1}^{n_1} \log\left(1+\frac{x_{1i}-\hat{\mu}}{\sigma}\right)}. \tag{2.26}$$

Remark 2.1 For a known μ and σ, the density function given in (2.6) belongs to an exponential family. Hence

$$\left(\sum_{i=1}^{n_1} \log\left(1+\frac{x_{1i}-\mu}{\sigma}\right),\ \sum_{i=1}^{n_1} \log\left(1+\frac{x_{2i}-\mu}{\sigma}\right),\ \sum_{i=1}^{n_2} \log\left(1+\frac{x_{1i}-\mu}{\sigma}\right),\ \sum_{i=1}^{n_2} \log\left(1+\frac{x_{2i}-\mu}{\sigma}\right)\right)$$

are jointly sufficient for $\left(\alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$.

A simulation study is conducted and is given as follows.

2.5.2 Simulation Study for $BP\ II\left(\mu, \sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ Distribution

The estimates of the Bivariate Pareto II distributions are obtained numerically. Random samples of sizes 25, 50 and 100 were generated using Theorem 2.3. The estimate of the parameters obtained using equations (2.23) to (2.26) are given in Table 2.1.

Table 2.1 MLE of a $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ when α_1=2, α_2=1.5, α_1'=1.75, α_2'=1.3, μ=3 and σ=3.5.

Sample Size	Parameters	α_1 (2)	α_2 (1.5)	α_1' (1.75)	α_2' (1.3)	μ (3)
25	Estimated Value	2.15077	1.54877	2.00093	1.41604	3.04427
	Bias	0.15077	0.04877	0.25093	0.11604	0.04427
	Variance	0.41345	0.24613	0.80046	0.16567	0.00140
50	Estimated Value	2.07767	1.56001	1.81442	1.3218	3.02194
	Bias	0.07767	0.06001	0.06442	0.0218	0.02194
	Variance	0.17265	0.09889	0.21106	0.05445	0.00036
100	Estimated Value	2.04010	1.55793	1.78173	1.30272	3.01002
	Bias	-0.04010	0.05793	0.03173	0.00272	0.01002
	Variance	0.07494	0.04592	0.07104	0.03096	8.8E-05

2.5.3 Estimation of Parameters of a $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ Distribution

Consider a random sample of size n from a population having bivariate density function $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$. The likelihood function is

$$L=\alpha_1^{n_1}(\alpha_1')^{n_2}\alpha_2^{n_2}(\alpha_2')^{n_1}\sigma^{n(\alpha_1+\alpha_2)}\prod_{i=1}^{n_1}x_{1i}^{-(\alpha_1+\alpha_2-\alpha_2'+1)}x_{2i}^{-(\alpha_2'+1)}$$
$$\prod_{i=1}^{n_2}x_{2i}^{-(\alpha_1+\alpha_2-\alpha_1'+1)}x_{1i}^{-(\alpha_1'+1)}$$

where n_i = number of $x_i<x_{3-i}$, $i=1,2$ and $n=n_1+n_2$. Then

$$\begin{aligned}\log L=&\,n_1\log\alpha_1+n_2\log(\alpha_1')+n_2\log\alpha_2+n_1\log(\alpha_2')+n(\alpha_1+\alpha_2)\log\sigma\\&-\left(\alpha_1+\alpha_2-\alpha_2'+1\right)\sum_{i=1}^{n_1}\log\left(x_{1i}\right)-\left(\alpha_2'+1\right)\sum_{i=1}^{n_1}\log\left(x_{2i}\right)\\&-\left(\alpha_1+\alpha_2-\alpha_1'+1\right)\sum_{i=1}^{n_2}\log\left(x_{2i}\right)-\left(\alpha_1'+1\right)\sum_{i=1}^{n_2}\log\left(x_{1i}\right).\end{aligned}\qquad(2.27)$$

If σ is known, the procedure to obtain the MLE for $\underline{\alpha}=\left(\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ is straight forward. However if σ is unknown we proceed by estimating σ from the distribution of Z. For fixed $\underline{\alpha}=\left(\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$, the maximum likelihood estimate for σ is

$$\hat{\sigma}=\min_i \min_j (X_{ij}),\ i=1,2,\ j=1,...,n. \tag{2.28}$$

Substituting $\hat{\sigma}$, we get the MLE of $\underline{\alpha}=\left(\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ as

$$\hat{\alpha}_1=\frac{n_1}{\sum_{i=1}^{n_1}\log\left(x_{1i}\right)+\sum_{i=1}^{n_2}\log\left(x_{2i}\right)-n\log\hat{\sigma}}, \tag{2.29}$$

$$\hat{\alpha}_2=\frac{n_2}{\sum_{i=1}^{n_1}\log\left(x_{1i}\right)+\sum_{i=1}^{n_2}\log\left(x_{2i}\right)-n\log\hat{\sigma}}, \tag{2.30}$$

$$\hat{\alpha}_1'=\frac{n_2}{\sum_{i=1}^{n_2}\log\left(x_{1i}\right)-\sum_{i=1}^{n_2}\log\left(x_{2i}\right)} \tag{2.31}$$

and

$$\hat{\alpha}_2'=\frac{n_1}{\sum_{i=1}^{n_2}\log\left(x_{2i}\right)-\sum_{i=1}^{n_1}\log\left(x_{1i}\right)}. \tag{2.32}$$

Remark 2.2 The estimator of α_1' and α_2' are independent of σ.

Theorem 2.4 For a given σ, $\sqrt{n}\left(\left(\hat{\alpha}_1,\hat{\alpha}_2,\hat{\alpha}_1',\hat{\alpha}_2'\right)-\left(\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)\right)\xrightarrow{D}$ $N(0,\Sigma)$ as $n\to\infty$, where $\Sigma=\frac{I_\theta^{-1}}{n}$, I_θ is the information matrix given by

$$I_\theta = \begin{bmatrix} \frac{n_1}{\alpha_1^2} & 0 & 0 & 0 \\ 0 & \frac{n_2}{\alpha_2^2} & 0 & 0 \\ 0 & 0 & \frac{n_2}{\alpha_1'^2} & 0 \\ 0 & 0 & 0 & \frac{n_1}{\alpha_2'^2} \end{bmatrix}$$

The proof follows directly from Serfling (1980 p. 145).

Corollary 2.6 For a given σ, the estimates of shape parameters $\underline{\alpha} = \left(\alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ are independent of each other.

We now consider alternative approaches to estimate σ. If $Z = min(X_1, X_2)$, then

$$\begin{aligned} \bar{F}_Z(x) &= P\left[X_1 > x, X_2 > x\right] \\ &= \left(\frac{x}{\sigma}\right)^{-\alpha}; x > \sigma,\ \alpha = \alpha_1 + \alpha_2 > 0,\ \sigma > 0 \end{aligned}$$

a univariate Pareto distribution given by

$$f_Z(x) = \frac{\alpha}{\sigma}\left(\frac{x}{\sigma}\right)^{-(\alpha+1)}; x > \sigma,\ \alpha = \alpha_1 + \alpha_2,\ \sigma > 0. \tag{2.33}$$

In the above approach we estimated σ using the maximum likelihood method. Also as observed in Remark 2.2, the estimates of α_1' and α_2' does not depend on σ. We consider alternative approaches to estimate σ from (2.33) and investigate if it gives better estimates in terms of bias and mean square error than in (2.29) and (2.30).

The two alternative approaches considered to estimate σ are

(i) Method of moments and

(ii) Method using principle of maximum entropy (POME)

(i) The method of moments:

The moment estimates $(\tilde{\sigma}, \tilde{\alpha})$ for the parameters σ and α are obtained by solving

$$\frac{1}{n}\sum_{i=1}^{n} X_i = \frac{\tilde{\alpha}\tilde{\sigma}}{(\tilde{\alpha}-1)} \tag{2.34}$$

and

$$\frac{1}{n}\sum_{i=1}^{n} {X_i}^2 = \frac{\tilde{\alpha}\tilde{\sigma}^2}{(\tilde{\alpha}-2)}. \tag{2.35}$$

(ii) Method using principle of maximum entropy (POME)

The principle of maximum entropy is employed to estimate σ from the distribution of $Z = min(X_1, X_2)$ given in (2.33). This method of estimation has been discussed in Singh and Guo (1995), Singh and Guo (1997), Singh (1998), Singh and Ahmad (2004) and Hao and Singh (2009). In Singh and Guo (1997), POME method is used for estimating the parameters of a generalized Pareto distribution. They showed that the parameter estimates yielded by POME were comparable or better than the maximum likelihood estimates obtained by method of moments within certain ranges of sample sizes. This method is detailed as follows.

Shannon (1948) defined entropy as a numerical measure of uncertainty or conversely the information content associated with a probability distribution, say $f(x,\theta)$ where θ represents the parameter vector used to describe a random variable X. The Shannon's entropy function $H(f)$ for a continuous random variable X is given in (1.45). Accordingly to Jayness (1961), the minimally biased distribution of X is the one which maximizes the entropy function subject to given information or which satisfies the principle of maximum entropy. Therefore the parameters of the distribution are obtained by achieving maximum $H(f)$. POME is a logical and rational criterion for choosing some specific

$f(x)$ that maximizes $H(f)$ and satisfies the given information expressed as constraints. In other words, for given information (mean, geometric mean, variance, skewness, lower limit, upper limit etc.) the distribution derived by POME would best represent X, implicitly this distribution would best represent the sample from which the information was derived. Inversely if it is desired to fit a particular probability distribution to a sample of data, then POME can uniquely specify the constraints (or the information) needed to derive the distribution. The distribution parameters are then related to these constraints.

Given m linearly independent constraints $C_i, i = 1, 2, .., m,$ in the form

$$C_i = \int g_i(x) f(x) dx, i = 1, 2, ..., m \tag{2.36}$$

where $g_i(x)$ are some functions whose averages over $f(x)$ are specified, then the maximum of $H(f)$ subject to equation (2.36) is given by the distribution (Tribus (1969))

$$f(x) = \exp\left[-\lambda_0 - \sum_{i=1}^{m} \lambda_i g_i(x)\right] \tag{2.37}$$

where $\lambda_i, i = 1, 2, ..., m,$ are the Lagrangian multipliers, and can be determined from (2.36) and (2.37). Inserting equation (2.37) in (1.45) yields the entropy of $f(x)$ in terms of the constraints and Lagrangian multipliers:

$$H(f) = \lambda_0 + \sum_{i=1}^{m} \lambda_i C_i. \tag{2.38}$$

Maximization of $H(f)$ then establishes the relation between constraints and Lagrangian multipliers. Thus, to derive a method using POME for the estimation of the parameters α and σ in (2.33), three steps are involved: (a) specification of the appropriate constraints; (b) derivation of the entropy of the distribution; and (c) derivation of the relations between Lagrangian multipliers and constraints.

(a) Specification of the constraints

The entropy function of Pareto I distribution can be derived by inserting equation (2.33) into equation (1.45) to get

$$H(f) = -\log\left(\frac{\alpha}{\sigma}\right)\int_{\sigma}^{\infty} f(x)dx + (\alpha+1)\int_{\sigma}^{\infty}\log\left(\frac{x}{\sigma}\right)f(x)dx. \tag{2.39}$$

Comparing equation (2.39) with (2.38), the constraints appropriate for equation (2.33) can be written (Kapur (1989)) as

$$\int_{\sigma}^{\infty} f(x)dx = 1 \tag{2.40}$$

and

$$\int_{\sigma}^{\infty}\log\left(\frac{x}{\sigma}\right)f(x)dx = E\left[\log\left(\frac{X}{\sigma}\right)\right] \tag{2.41}$$

in which $E[.]$ denotes the expectation of the bracketed quantity. The first constraint specifies the total probability. The second constraint specifies the geometric mean of $\frac{X}{\sigma}$. The parameters of the distribution are related to these constraints.

(b) Construction of entropy function

The probability density function of the Pareto I distribution corresponding to POME and consistent with equation (2.40) and (2.41) takes the form

$$f(x) = \exp\left(-\lambda_0 - \lambda_1\log\left(\frac{x}{\sigma}\right)\right) \tag{2.42}$$

where λ_0 and λ_1 are Lagrangian multipliers and $g(x) = \log\left(\frac{x}{\sigma}\right)$.

By applying equation (2.42) to the total probability condition in equation (2.40), one obtains

$$e^{\lambda_0} = \int_{\sigma}^{\infty} e^{\left(-\lambda_1\log\left(\frac{x}{\sigma}\right)\right)}dx, \tag{2.43}$$

which yields,

$$e^{\lambda_0} = \frac{\sigma}{(\lambda_1 - 1)}. \tag{2.44}$$

The zeroth Lagrangian multiplier is now given by

$$\lambda_0 = \log\left(\frac{\sigma}{\lambda_1 - 1}\right). \tag{2.45}$$

Inserting (2.44) in (2.42) yields,

$$f(x) = \frac{\lambda_1 - 1}{\sigma}\left(\frac{x}{\sigma}\right)^{-\lambda_1}. \tag{2.46}$$

A comparison of equation (2.46) with (2.33), we get

$$\lambda_1 = \alpha + 1. \tag{2.47}$$

Taking logarithm of equation (2.46) gives

$$\log f(x) = \log(\lambda_1 - 1) - \log\sigma - \lambda_1 \log\left(\frac{x}{\sigma}\right). \tag{2.48}$$

Therefore, the entropy $H(f)$ of the Pareto I distribution follows,

$$H(f) = -\log(\lambda_1 - 1) + \log\sigma + \lambda_1 E\left(\log\left(\frac{X}{\sigma}\right)\right). \tag{2.49}$$

(c) Relation between distribution parameters and constraints

According to Singh and Rajagopal (1986), the relation between distribution parameters and constraints is obtained by taking partial derivatives of the entropy function $H(f)$ with respect to the Lagrangian multipliers (other than zeroth) as well as distribution parameter, and then equating the derivatives to zero, and making use of the constraints. To that end, taking partial derivatives of equation (2.49) with respect to λ_1 and equating to zero yields

$$\frac{\partial}{\partial\lambda_1} H(f) = -\frac{1}{(\lambda_1 - 1)} + E\left(\log\left(\frac{X}{\sigma}\right)\right) = 0$$

$$E\left[\log\left(\frac{X}{\sigma}\right)\right] = \frac{1}{\lambda_1 - 1}. \tag{2.50}$$

Alternatively, the estimation equation (2.50) can also be obtained by differentiating the zeroth Lagrangian multiplier with respect to the Lagrangian multiplier λ_1 and equating the derivative to zero. Equation (2.43) is written as

$$\lambda_0 = \log \int_{\sigma}^{\infty} e^{-\lambda_1 \log\left(\frac{x}{\sigma}\right)} dx . \tag{2.51}$$

Differentiating (2.51) with respect to λ_1, we obtain

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{-\int_{\sigma}^{\infty} \log\left(\frac{x}{\sigma}\right) e^{-\lambda_1 \log\left(\frac{x}{\sigma}\right)} dx}{\int_{\sigma}^{\infty} e^{-\lambda_1 \log\left(\frac{x}{\sigma}\right)} dx} . \tag{2.52}$$

Using (2.43), equation (2.52) becomes

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \int_{\sigma}^{\infty} e^{-\lambda_0 - \lambda_1 \log\left(\frac{x}{\sigma}\right)} \log\left(\frac{x}{\sigma}\right) dx . \tag{2.53}$$

Now using (2.42), we get

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -E\left[\log\left(\frac{X}{\sigma}\right)\right]. \tag{2.54}$$

Once again differentiating (2.52) with respect to λ_1 and simplifying, we get

$$\frac{\partial^2 \lambda_0}{\partial {\lambda_1}^2} = \int_{\sigma}^{\infty} e^{-\lambda_0 - \lambda_1 \log\left(\frac{x}{\sigma}\right)} \left(\log\left(\frac{x}{\sigma}\right)\right)^2 dx - \left(\int_{\sigma}^{\infty} e^{-\lambda_0 - \lambda_1 \log\left(\frac{x}{\sigma}\right)} \log\left(\frac{x}{\sigma}\right) dx\right)^2$$

i.e.

$$\frac{\partial^2 \lambda_0}{\partial {\lambda_1}^2} = E\left[\log\left(\frac{X}{\sigma}\right)\right]^2 - \left(E\left[\log\left(\frac{X}{\sigma}\right)\right]\right)^2,$$

$$\frac{\partial^2 \lambda_0}{\partial {\lambda_1}^2} = V\left[\log\left(\frac{X}{\sigma}\right)\right]. \tag{2.55}$$

From equation (2.44),

$$\lambda_0 = \log \sigma - \log(\lambda_1 - 1) . \tag{2.56}$$

Differentiating (2.56) with respect to λ_1,

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{1}{\lambda_1 - 1} \tag{2.57}$$

and

$$\frac{\partial^2 \lambda_0}{\partial \lambda_1^{\,2}} = \frac{1}{(\lambda_1 - 1)^2} . \tag{2.58}$$

Equating (2.57) to (2.54) and (2.58) to (2.55) leads to

$$E\left[\log\left(\frac{X}{\sigma}\right)\right] = \frac{1}{(\lambda_1 - 1)} \tag{2.59}$$

and

$$V\left[\log\left(\frac{X}{\sigma}\right)\right] = \frac{1}{(\lambda_1 - 1)^2} . \tag{2.60}$$

Equations (2.59) and (2.50) are equivalent. Hence the parametric estimation of POME consists of two equations (2.59) and (2.60). Inserting $\lambda_1 = \alpha + 1$ from (2.47) into (2.59) and (2.60), we get

$$E\left[\log\left(\frac{X}{\sigma}\right)\right] = \frac{1}{\alpha} \tag{2.61}$$

and

$$V\left[\log\left(\frac{X}{\sigma}\right)\right] = \frac{1}{\alpha^2} . \tag{2.62}$$

Equations (2.61) and (2.62) are the POME based estimation equations. The POME estimates of the parameters σ and α can be obtained by solving the equations (2.61) and (2.62) numerically. We denote them as $\sigma*$ and $\alpha*$. Replacing $\hat{\sigma}$ by $\sigma*$ in (2.29) to (2.32), we get another set of estimates as $(\sigma*, \alpha_1*, \alpha_2*, \alpha_1'*, \alpha_2'*)$. Observe that $\alpha_1'*$ and $\alpha_2'*$ remain the same as $\hat{\alpha}_1'$ and $\hat{\alpha}_2'$ since their estimates are independent of σ.

On substitution of the moment estimate of σ, $\tilde{\sigma}$ from the solution of (2.34) and (2.35), we get the moment estimates corresponding to σ and $\underline{\alpha} = \left(\alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$.

2.5.4 Simulation for $BP\,I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ Distribution

The random samples from $BP\,I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution are generated using Corollary 2.4. For each population, 100 random samples of sizes 25, 50 and100 were generated and the parameters are estimated in two stages as discussed in Section 2.5.3. The bias, variance and efficiency of the estimates obtained are given in Table 2.2.

Table 2.2 Estimates of parameters $\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'$ of a $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution.

N	Parameters		POME	Moment	MLE	Efficiency POME *MLE	Efficiency Moment *MLE	Efficiency Moment *POME
25	σ (2)	Bias	0.1682	0.8358	0.0581	0.1327	0.00012	0.0009
		Variance	0.0167	19.2961	0.0022			
	α_1 (1.5)	Bias	0.4659	7.0022	0.1326	0.5103	0.00006	0.00013
		Variance	0.4747	3745.352	0.2422			
	α_2 (1.25)	Bias	0.4755	5.7079	0.1833	0.5475	0.00006	0.0001
		Variance	0.3859	3683.895	0.2113			
	α_1' (1.8)	Bias	0.1394	0.1394	0.1394	1	1	1
		Variance	0.4367	0.4367	0.4367			
	α_2' (1.6)	Bias	0.0981	0.0981	0.0981	1	1	1
		Variance	0.7166	0.7166	0.7166			
50	σ (2)	Bias	0.1235	0.3242	0.0301	0.0747	0.0511	0.6841
		Variance	0.0106	0.0155	0.0008			
	α_1 (1.5)	Bias	0.2954	1.1383	0.0532	0.6245	0.1579	0.2530
		Variance	0.1293	0.5111	0.0807			
	α_2 (1.25)	Bias	0.3545	1.1223	0.1355	0.6385	0.6385	0.2759
		Variance	0.1868	0.6770	0.1193			
	α_1' (1.8)	Bias	0.0982	0.0982	0.0982	1	1	1
		Variance	0.1343	0.1343	0.1343			
	α_2' (1.6)	Bias	-0.0167	-0.0167	-0.0167	1	1	1
		Variance	0.5033	0.5033	0.5033			
100	σ (2)	Bias	0.0997	0.2790	0.0125	0.0266	0.0161	0.6037
		Variance	0.0050	0.0083	0.0001			
	α_1 (1.5)	Bias	0.2570	0.8971	0.0451	0.6442	0.2128	0.3153
		Variance	0.0671	0.2128	0.0432			
	α_2 (1.25)	Bias	0.1849	0.7078	0.0119	0.6581	0.2149	0.3267
		Variance	0.0516	0.1578	0.0339			
	α_1' (1.8)	Bias	0.0114	0.0114	0.0114	1	1	1
		Variance	0.0873	0.0873	0.0873			
	α_2' (1.6)	Bias	-0.2741	-0.2741	-0.2741	1	1	1
		Variance	0.1054	0.1054	0.1054			

2.6. Data Analysis

For the illustration of the bivariate Pareto model, a real data set given in Kim and Kvam (2004) is considered. The data set consists of the failure times of 20 sample units from a system consists of three components. Exponential transformation of the data is taken for the illustration the model. The transformed observations are given in the Table 2.3.

Table 2.3 Transformed data set

Sl. No.	X	Y	**Sl. No.**	X	Y
1	7.3046	1124.393	11	2.4102	0.818
2	0.9153	0.9935	12	17.3809	4.0224
3	8.9994	11.0694	13	5405.896	1.5002
4	7.0784	7.4593	14	453753.5	2.9878
5	2.0737	1579.856	15	94.7863	37.5636
6	1.9326	6034.524	16	36.077	11.5354
7	2.2236	171.2747	17	3.54E+08	6.9313
8	1872.675	12.0204	18	159.6718	4.781
9	4126.667	109.0827	19	32380.1	6034.524
10	4.1107	0.8536	20	9.7781	3.3759

The Kolmogorov-Smirnov test for the bivariate random variable is made as follows (Justel et al. (1997)).

(1) Compute the maximum distance in the observed points, $D_n{}^1 = \max\limits_{i=1,\dots,n} D_n{}^+(u_i)$.

(2) Compute the maximum and minimum distances in the intersection points,

$$D_n{}^2 = \max_{i,j=1,\dots,n} \left\{ D_n{}^+(x_j, y_i) \mid x_j > x_i, y_j < y_i \right\}$$

and

$$D_n{}^3 = \frac{2}{n} - \min_{i,j=1,\dots,n} \left\{ D_n{}^+(x_j, y_i) \mid x_j > x_i, y_j < y_i \right\}.$$

(3) Compute the maximum distance among the projections of the observed points on the right unit square border, $D_n{}^4 = \frac{1}{n} - \min\limits_{i=1,\dots,n} D_n{}^+(1, y_i)$.

(4) Compute the maximum distance among the projections of the observed points on the top unit square border, $D_n{}^5 = \frac{1}{n} - \min\limits_{i=1,\dots,n} D_n{}^+(x_i, 1)$.

(5) Compute the maximum distance, $D_n = \max\left\{ D_n{}^1, D_n{}^2, D_n{}^3, D_n{}^4, D_n{}^5 \right\}$.

Now for the data given in Table 2.3, we observe the Kolmogorov distance as follows

D_n^1	D_n^2	D_n^3	D_n^4	D_n^5	D_n
0.0975	0.4423	0.0118	0.0359	0.0328	0.4423

From Table 1 (Justel et al. (1997)), the Monte-Carlo approximation to the percentiles of the bivariate Kolmogorov-Smirnov Statistic distribution D_n is accepted at 0.0025 percentiles for $n = 20$. Thus the transformed data fits for the bivariate Pareto II distribution given in (2.6) at 0.0025 percentiles. Also the data is fitted graphically and is given below. With this we conclude this chapter.

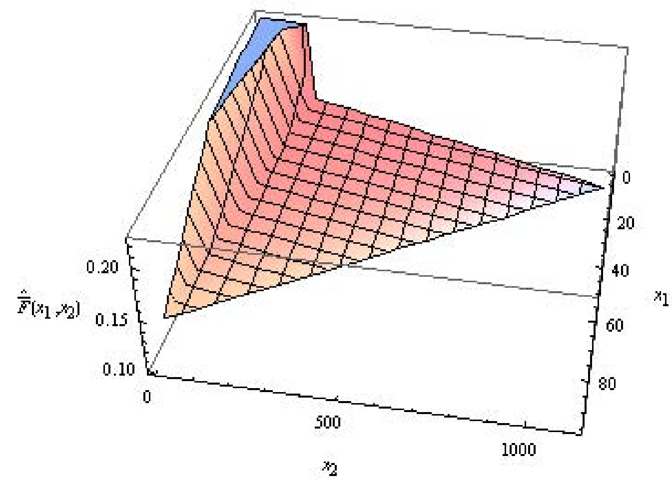


Figure 2.3: Bivariate Empirical Survival Plot for the data given in Table 2.3.

Chapter 3

Characterizations of Bivariate Pareto Distributions

3.1 Introduction

The identification of an appropriate model that can describe the properties of a particular set of data is a basic problem in analyzing statistical data. Characterization problem usually identifies some unique property possessed by a distribution and it helps to obtain an exact model followed by the observations through the consideration of the physical characteristics that governs the pattern of the data. A large volume of work is available in the literature (Galambos and Kotz (1978), Azlarov and Volodin (1986)). As an example, the constancy of the failure rate and lack of memory property are the characteristics of an exponential distribution. The intimate relationship between the Pareto and exponential distributions permits us to obtain analogous characterizations of the former from characterizations of the latter. Many characterizations of the exponential are translations of the lack of memory property. In fact characterizations of Pareto and exponential distributions are in essence variants of the memory less property

Some results in this chapter is published in Asha and Jagathnath (2006, 2008).

of the exponential and are essentially reduced to solving a Cauchy type functional equation or its integral variants. However, when discussing lack of memory property in the higher dimension, it ceases to have a unique extension (See Galambos & Kotz (1978)) and each definition manifests itself in a unique way. For results based on bivariate lack of memory property, see Roy (2004) and for discrete case Sun and Basu (1995), Nair and Asha (1997), Asha et al. (2003) to mention a few cases. The variant form of the lack of memory property has been referred to as the multivariate dullness property by Veenus and Nair (1994), Hanagal (1996), Yeh (2004 a,b).

This chapter contains three sections. In section 3.2, the $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution is characterized under certain conditions by the bivariate dullness property discussed in Hanagal (1996). Apart from these, results based on rescaling and transformations are also discussed. In section 3.3 we consider characterizations $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ and $BP\ II(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ distributions based on reliability characteristics.

3.2 Characterizations using Dullness Property and its Variants

In this section we extend the characterizations results mentioned in Chapter one to the bivariate case. These characterizations are meaningful in the income distribution context. The extended form of the univariate dullness property is given as follows.

Definition 3.1 (Veenus and Nair (1994), Hanagal (1996), Yeh (2004 a,b)) The distribution of $\left(X_1,X_2\right)$ is called dull at the point $\left(t_1,t_2\right)$ whenever

$$P\left[X_1 \geq s_1t_1, X_2 \geq s_2t_2 \mid X_1 \geq t_1, X_2 \geq t_2\right] = P\left[X_1 \geq s_1, X_2 \geq s_2\right] \tag{3.1}$$

for all $s_1,s_2 \geq 1$ and $t_1,t_2 \geq 1$.

In fact, (3.1) can be equivalently written as

$$P\left[Y_1 \geq y_1 + t_1', Y_2 \geq y_2 + t_2'\right] = P\left[Y_1 \geq y_1, Y_2 \geq y_2\right] P\left[Y_1 \geq t_1', Y_2 \geq t_2'\right] \quad (3.2)$$

where $Y_i = \log X_i > 0$ with $t_i' = \log t_i > 0$.

From Galambos and Kotz (1978) it follows that $Y_i's,\ i=1,2$ are independent exponential and so (X_1, X_2) is a bivariate Pareto with independent marginals. Relaxing conditions on (3.1) and defining bivariate dullness property as in Veenus and Nair (1994), Hanagal (1996), the distribution of (X_1, X_2) is called dull at the point (t,t) where $t>1$, whenever

$$P\left[X_1 \geq s_1 t, X_2 \geq s_2 t \mid X_1 \geq t, X_2 \geq t\right] = P\left[X_1 \geq s_1, X_2 \geq s_2\right] \quad (3.3)$$

for all $s_1, s_2 \geq 1$.

The distribution of (X_1, X_2) is called totally dull if (3.3) holds true for all (s_1, s_2) and $t>1$. The equation (3.3) can be equivalently written as

$$P\left[Y_1 \geq y_1 + t', Y_2 \geq y_2 + t'\right] = P\left[Y_1 \geq y_1, Y_2 \geq y_2\right] P\left[Y_1 \geq t', Y_2 \geq t'\right] \quad (3.4)$$

where $Y_i = \log X_i > 0, t' = \log t > 0$.

If $\bar{G}(y_1, y_2)$ denotes the survival function of (Y_1, Y_2) and $\bar{G}_i(y_i)$ denotes the marginal survival functions of Y_i, then it follows from Galambos and Kotz (1978) that

$$\bar{G}(y_1, y_2) = \begin{cases} e^{-(\alpha_1+\alpha_2)y_2}\bar{G}_1(y_1 - y_2);\ 0 \leq y_2 \leq y_1 \\ e^{-(\alpha_1+\alpha_2)y_1}\bar{G}_2(y_2 - y_1);\ 0 \leq y_1 \leq y_2 \end{cases} \quad (3.5)$$

for some $\alpha_1, \alpha_2 > 0$.

The equations (3.4) and (3.5) are nothing but the bivariate lack of memory property (Galambos and Kotz (1978)) and (3.5) is the class of exponential minima. Accordingly, we can have a general class of distribution with Pareto minima as

$$\bar{F}(x_1,x_2)=\begin{cases} x_2^{-(\alpha_1+\alpha_2)}\bar{F}_1\left(\dfrac{x_1}{x_2}\right); 1\le x_2\le x_1 \\ x_1^{-(\alpha_1+\alpha_2)}\bar{F}_2\left(\dfrac{x_2}{x_1}\right); 1\le x_1\le x_2 \end{cases} ;\alpha_1+\alpha_2>0. \tag{3.6}$$

That the class specified in (3.6) satisfies the dullness property (3.3). Hence, we have the following theorem.

Theorem 3.1 A random vector (X_1,X_2) with survival function $\bar{F}(x_1,x_2)$ in the support of $(1,\infty)\times(1,\infty)$ satisfies (3.3) if and only if $\bar{F}(x_1,x_2)$ belongs to the class specified by (3.6).

Corollary 3.1 Let (X_1,X_2) be a bivariate random variable with mixture Pareto marginals specified as

$$\bar{F}_i(x_i)=\frac{\alpha_{3-i}}{\alpha_1+\alpha_2-\alpha_i'}x_i^{-\alpha_i'}+\left(\frac{\alpha_i-\alpha_i'}{\alpha_1+\alpha_2-\alpha_i'}\right)x_i^{-(\alpha_1+\alpha_2)}; 1\le x_i,\ \alpha_i,\alpha_i'>0,\ i=1,2 \tag{3.7}$$

then (X_1,X_2) has a $BP\ I(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ distribution with survival function specified by

$$\bar{F}(x_1,x_2)=\frac{\alpha_{3-i}}{\alpha_1+\alpha_2-\alpha_i'}x_i^{-(\alpha_1+\alpha_2-\alpha_i')}x_{3-i}^{-\alpha_i'}+\left(\frac{\alpha_i-\alpha_i'}{\alpha_1+\alpha_2-\alpha_i'}\right)x_{3-i}^{-(\alpha_1+\alpha_2)};$$
$$1\le x_{3-i}\le x_i,\ \alpha_i,\alpha_i'>0,\ i=1,2, \tag{3.8}$$

if and only if the distribution of (X_1,X_2) is totally dull.

Proof

If (X_1,X_2) is distributed as (3.8), then to prove that

$$\bar{F}(x_1t,x_2t)=\bar{F}(x_1,x_2)\bar{F}(t,t) \text{ for all } x_1,x_2,t\ge 1 \tag{3.9}$$

is straightforward.

Conversely, from (3.6) it follows that

$$\bar{F}(x_1,x_2)=\begin{cases} x_2^{-(\alpha_1+\alpha_2)}\left[\dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}\left(\dfrac{x_1}{x_2}\right)^{-\alpha_1'}+\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)\left(\dfrac{x_1}{x_2}\right)^{-(\alpha_1+\alpha_2)}\right]; \\ \hfill 1\le x_2\le x_1 \\ x_1^{-(\alpha_1+\alpha_2)}\left[\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}\left(\dfrac{x_2}{x_1}\right)^{-\alpha_2'}+\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)\left(\dfrac{x_2}{x_1}\right)^{-(\alpha_1+\alpha_2)}\right]; \\ \hfill 1\le x_1\le x_2 \end{cases}.$$

From which it follows that,

$$\bar{F}(x_1,x_2)=\begin{cases} \left[\dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}x_1^{-\alpha_1'}x_2^{-(\alpha_1+\alpha_2-\alpha_1')}+\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)x_1^{-(\alpha_1+\alpha_2)}\right]; \\ \hfill 1\le x_2\le x_1 \\ \left[\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}x_2^{-\alpha_2'}x_1^{-(\alpha_1+\alpha_2-\alpha_2')}+\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)x_2^{-(\alpha_1+\alpha_2)}\right]; \\ \hfill 1\le x_1\le x_2 \end{cases}.$$

Hence the result.

Corollary 3.2 Let (X_1,X_2) be a bivariate random variable with Pareto marginals specified by

$$\bar{F}_i(x_i)=x_i^{-(\lambda_i+\lambda_{12})};1\le x_i,\ \lambda_i,\lambda_{12}>0,\ i=1,2 \tag{3.10}$$

then (X_1,X_2) has a Marshall-Olkin type bivariate Pareto I (Muliere and Scarcini (1987), Veenus and Nair (1994), Hanagal (1996)) distribution with survival function specified by

$$\bar{F}(x_1,x_2)=x_1^{-\lambda_1}x_1^{-\lambda_2}\left\{\max(x_1,x_2)\right\}^{-\lambda_{12}};1\le x_i,\ \lambda_i,\lambda_{12}>0,\ i=1,2 \tag{3.11}$$

if and only if the distribution of (X_1,X_2) is totally dull.

A few bivariate Pareto distributions belong to the class specified in (3.6) is given below.

Table 3.1 Bivariate Pareto model possessing dullness property

No.	Bivariate Survival function
1.	Pareto [Veenus & Nair (1994), Hanagal (1996)], $\bar{F}(x_1,x_2)=(x_1)^{-\alpha_1}(x_2)^{-\alpha_2}\{\max(x_1,x_2)\}^{-\alpha_0}$; $x_1,x_2\ge 1,\ \alpha_i>1,\ i=1,2.$
2.	Pareto [Yeh (2004 a)], $\bar{F}(x_1,x_2)=(x_1)^{-\alpha}(x_2)^{-\alpha}\max\left((x_1)^{-\alpha},\ (x_2)^{-\alpha}\right);\ x_1,x_2\ge 1,\ \alpha>1$
3.	Mixture Pareto, $\bar{F}(x_1,x_2)=kx_1^{-\alpha_1}x_2^{-(\alpha_2-\alpha_1)}+(1-k)x_1^{-(\alpha_1-\alpha_2)}x_2^{-\alpha_1};\ x_1,x_2\ge 1,$ $\alpha_i>0,\ i=1,2,\ 0<k<1.$
4.	Bivariate Pareto [Asha & Jagathnath Krishna (2008)], $\bar{F}(x_1,x_2)=\begin{cases}\dfrac{\alpha_1\alpha_2'}{(\alpha_1+\alpha_2-\alpha_2')}(x_1)^{-(\alpha_1+\alpha_2-\alpha_2')}(x_2)^{-\alpha_2'} + \dfrac{(\alpha_2-\alpha_2')}{(\alpha_1+\alpha_2-\alpha_2')}(x_2)^{-(\alpha_1+\alpha_2)};\ 1\le x_1\le x_2\\ \dfrac{\alpha_2\alpha_1'}{(\alpha_1+\alpha_2-\alpha_1')}(x_2)^{-(\alpha_1+\alpha_2-\alpha_1')}(x_1)^{-\alpha_1'} + \dfrac{(\alpha_1-\alpha_1')}{(\alpha_1+\alpha_2-\alpha_1')}(x_1)^{-(\alpha_1+\alpha_2)};\ 1\le x_2\le x_1\end{cases}$ $\sigma>0,\ \alpha_i,\alpha_i'>0,\ \alpha_1+\alpha_2\ne\alpha_i',\ i=1,2.$

As typical in the case of lack of memory property, the dullness property (3.3) also manifests itself on characterizations of the bivariate Pareto. One such characterization is based on rescaling is discussed below.

Theorem 3.2 Let (X_1, X_2) be a random vector and suppose there exist $\sigma,\ \alpha_i,\ \alpha_i' > 0$ such that $x_i > \sigma,\ i = 1,2,$ then the following two statements are equivalent.

(i) (X_1, X_2) has a $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution

(ii) Let (X_1, X_2) be a bivariate random vector with marginals specified as in (3.7). Then for any $\sigma < y < x_i,\ i = 1,2,$ the bivariate truncation of (X_1, X_2) is equivalent to the bivariate rescaling in the form

$$P[X_1 > x_1, X_2 > x_2 \mid X_1 > y, X_2 > y] = P\left[\frac{1}{\sigma} yX_1 > x_1, \frac{1}{\sigma} yX_2 > x_2\right]. \tag{3.12}$$

Proof

Suppose (X_1, X_2) has a $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution, then for any $x_1, x_2 > y > \sigma,$ the bivariate truncation of (X_1, X_2) is

$$P[X_1 > x_1, X_2 > x_2 \mid X_1 > y, X_2 > y] = \frac{P[X_1 > x_1, X_2 > x_2]}{P[X_1 > y, X_2 > y]} \tag{3.13}$$

$$= \frac{\left(\dfrac{\alpha_i}{\alpha_1+\alpha_2-\alpha_{3-i}'}\right)\left(\dfrac{x_i}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_{3-i}')}\left(\dfrac{x_{3-i}}{\sigma}\right)^{-\alpha_{3-i}'} + \left(\dfrac{\alpha_{3-i}-\alpha_{3-i}'}{\alpha_1+\alpha_2-\alpha_i'}\right)\left(\dfrac{x_{3-i}}{\sigma}\right)^{-(\alpha_1+\alpha_2)}}{\left(\dfrac{\alpha_i}{\alpha_1+\alpha_2-\alpha_{3-i}'}\right)\left(\dfrac{y}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_{3-i}')}\left(\dfrac{y}{\sigma}\right)^{-\alpha_{3-i}'} + \left(\dfrac{\alpha_{3-i}-\alpha_{3-i}'}{\alpha_1+\alpha_2-\alpha_i'}\right)\left(\dfrac{y}{\sigma}\right)^{-(\alpha_1+\alpha_2)}}$$

$$= \left(\frac{\alpha_i}{\alpha_1+\alpha_2-\alpha_{3-i}'}\right)\left(\frac{x_i}{y}\right)^{-(\alpha_1+\alpha_2-\alpha_{3-i}')}\left(\frac{x_{3-i}}{y}\right)^{-\alpha_{3-i}'} + \left(\frac{\alpha_{3-i}-\alpha_{3-i}'}{\alpha_1+\alpha_2-\alpha_i'}\right)\left(\frac{x_{3-i}}{y}\right)^{-(\alpha_1+\alpha_2)} \tag{3.14}$$

$$\sigma < x_i < x_{3-i} < y < \infty.$$

On the other hand, rescaling of (X_1, X_2) is

$$P\left[\frac{1}{\sigma} yX_1 > x_1, \frac{1}{\sigma} yX_2 > x_2\right] = P\left[X_1 > \frac{\sigma x_1}{y}, X_2 > \frac{\sigma x_2}{y}\right]$$

$$= \left(\frac{\alpha_i}{\alpha_1+\alpha_2-\alpha_{3-i}'}\right)\left(\frac{x_i}{y}\right)^{-(\alpha_1+\alpha_2-\alpha_{3-i}')}\left(\frac{x_{3-i}}{y}\right)^{-\alpha_{3-i}'} + \left(\frac{\alpha_{3-i}-\alpha_{3-i}'}{\alpha_1+\alpha_2-\alpha_i'}\right)\left(\frac{x_{3-i}}{y}\right)^{-(\alpha_1+\alpha_2)}$$

$$\sigma < x_i < x_{3-i} < y < \infty.$$

Thus proving the first part of the theorem.

To prove the converse, it follows from

$$P\left[X_1 > x_1, X_2 > x_2 \mid X_1 > y, X_2 > y\right] = P\left[\frac{1}{\sigma} yX_1 > x_1, \frac{1}{\sigma} yX_2 > x_2\right]$$

that

$$P\left[X_1 > x_1, X_2 > x_2\right] = P\left[X_1 > y, X_2 > y\right] P\left[\frac{1}{\sigma} yX_1 > x_1, \frac{1}{\sigma} yX_2 > x_2\right] \tag{3.15}$$

$$;\sigma < y < x_i < \infty, i = 1,2.$$

Consider the transformation $Z_i = \left(\dfrac{X_i}{\sigma}\right)$ and $x_i = w_i y$, then $\dfrac{w_i y}{\sigma} > 1$.

Equation (3.15) now becomes

$$P\left[Z_1 > \frac{w_1 y}{\sigma}, Z_2 > \frac{w_2 y}{\sigma}\right] = P\left[Z_1 > \frac{y}{\sigma}, Z_2 > \frac{y}{\sigma}\right] P\left[Z_1 > w_1, Z_2 > w_2\right] \tag{3.16}$$

for all $\dfrac{w_i y}{\sigma} > 1, \dfrac{y}{\sigma} > 1.$

Since X_i has marginals specified as in (3.7) it follows that the marginals of Z_i is $\bar{F}_i(\sigma z_i),\ z_i > 1, i = 1,2.$

Further from equation (3.16) and Corollary 3.2 we obtain the survival function of (Z_1, Z_2) as

$$\bar{F}(z_1, z_2) = \begin{cases} (\sigma z_2)^{-(\alpha_1+\alpha_2)}\, \bar{F}_1\left(\dfrac{z_1}{z_2}\right); 1 \le z_2 \le z_1 \\ (\sigma z_1)^{-(\alpha_1+\alpha_2)}\, \bar{F}_2\left(\dfrac{z_2}{z_1}\right); 1 \le z_1 \le z_2 \end{cases}.$$

Simple transformation enables us to obtain $\bar{F}(x_1, x_2)$ as in (i).

Hence the result.

Theorem 3.3 Let (X_1, X_2) be a random vector as defined above and suppose there exist $\sigma, \alpha_i, \alpha_i' > 0$, such that $x_i \ge \sigma$, $i = 1,2$ and $(Y_1, Y_2) = (SX_1, SX_2)$ with $Y_i = SX_i, i = 1,2$ where S is continuous random variable defined over $0 < S < 1$, and is independent of (X_1, X_2) then for any $(y_1, y_2) > (\sigma, \sigma)$

$$P\left[Y_1 > ty_1, Y_2 > tY_2 \mid Y_1 > t\sigma, Y_2 > t\sigma\right] = P\left[X_1 > y_1, X_2 > y_2\right] \tag{3.17}$$

holds for all $t \geq 1$, if and only if (X_1, X_2) has a $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution.

Proof

Since $(Y_1, Y_2) = (SX_1, SX_2)$, S and (X_1, X_2) are independent, the joint survival function of (Y_1, Y_2) is derived by total probability rule as for any $(Y_1, Y_2) > (\sigma, \sigma)$ as

$$\begin{aligned}\bar{F}_Y(y_1, y_2) &= P\left[Y_1 > y_1, Y_2 > y_2\right] \\ &= \int_0^1 P\left[SX_1 > y_1, SX_2 > y_2 \mid S = s\right] dF_S(s) \\ &= \int_0^1 P\left[X_1 > \frac{1}{s} y_1, X_2 > \frac{1}{s} y_2\right] dF_S(s) \\ &= \int_0^1 \bar{F}_X\left[\frac{1}{s} y_1, \frac{1}{s} y_2\right] dF_S(s)\end{aligned}$$

where $F_S(.)$ is the cumulative distribution function of S, $\bar{F}_Y(.)$, $\bar{F}_X(.)$ denotes the survival function of (Y_1, Y_2) and (X_1, X_2) respectively so that (3.17) becomes

$$\begin{aligned}&\bar{F}_Y(ty_1, ty_2) - \bar{F}_Y(t\sigma, t\sigma)\bar{F}_X(y_1, y_2) \\ &\quad = \int_0^1 \left\{\bar{F}_X\left(\frac{t}{s} y_1, \frac{t}{s} y_2\right) - \bar{F}_X\left(\frac{t}{s}\sigma, \frac{t}{s}\sigma\right)\bar{F}_X(y_1, y_2)\right\} dF_S(s)\end{aligned} \tag{3.18}$$

for all $t \geq 1$.

If (X_1, X_2) has a $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution, the integrand on the RHS of (3.18) becomes zero and hence we have

$$\bar{F}_Y(ty_1, ty_2) - \bar{F}_Y(t\sigma, t\sigma)\bar{F}_Y(y_1, y_2) = 0 \tag{3.19}$$

$$P\left[Y_1 > ty_1, Y_2 > ty_2 \mid Y_1 > t\sigma, Y_2 > t\sigma\right] = P\left[X_1 > y_1, X_2 > y_2\right].$$

To prove the converse let (3.19) holds for all $t \geq 1$. Equivalently (3.19) can be written as $\bar{F}_Y(ty_1, ty_2) = \bar{F}_Y(t\sigma, t\sigma)\bar{F}_X(ty_1, ty_2)$ for all $t \geq 1$. From (3.19) we infer that the integrand on the right hand side of (3.18) is zero almost everywhere. Then there exist a subset (0,1), say A, where $A=\{s \mid s\in(0,1), \bar{F}_X\left(t \mid sy_1, sy_2\right) \neq \bar{F}_X\left(t \mid s\sigma, s\sigma\right)\bar{F}_X\left(y_1, y_2\right)\}$ such that $P(A)=0$.

Thus for any given $s\in(0,1)\setminus A$, the following identity

$$P\left[X_1 > \frac{t}{s}y_1, X_2 > \frac{t}{s}y_2\right] = P\left[X_1 > \frac{t}{s}\sigma_1, X_2 > \frac{t}{s}\sigma_2\right]P\left[X_1 > y_1, X_2 > y_2\right] \quad (3.20)$$

holds for all $t \geq 1$.

In particular, let $t=1$, then (3.20) reduces to

$$P\left[X_1 > \frac{1}{s}y_1, X_2 > \frac{1}{s}y_2\right] = P\left[X_1 > \frac{1}{s}\sigma, X_2 > \frac{1}{s}\sigma\right]P\left[X_1 > y_1, X_2 > y_2\right]$$

or

$$P\left[X_1 > \frac{1}{s}y_1, X_2 > \frac{1}{s}y_2 \mid X_1 > \frac{1}{s}\sigma, X_2 > \frac{1}{s}\sigma\right] = P\left[\frac{1}{s}X_1 > \frac{1}{s}y_1, \frac{1}{s}X_2 > \frac{1}{s}y_2\right]. \quad (3.21)$$

Equation (3.21) says that random vector (X_1, X_2) has the truncation invariant property at level vector $\left(\frac{1}{s}\sigma, \frac{1}{s}\sigma\right)(>(\sigma,\sigma))$ for any given $s\in(0,1)\setminus A$, then according to Theorem 3.2, this truncation invariant property is sufficient condition for (X_1, X_2) to have $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution. Hence the result.

3.3 Characterizations Based on Reliability Concepts

In the univariate case under the assumption of existence of moments, a characterization based on truncation is equivalent to linearity of mean residual lifetime (Kotz and Shanbhag (1980)). It is thus interesting to investigate if the characterizations based on the truncation reflect on the the reliability concepts like failure rate, bivariate mean residual life function. The following theorem looks into these aspects. The bivariate failure rate and related concepts like mean residual life function (MRL), vitality function plays a crucial role in reliability and survival analysis. This concept is extended to higher dimensions in more than one way. When considering the reliability characteristics in the bivariate case there is a need to choose an appropriate extension which reflects the ageing characteristics of the bivariate distribution. Here we advocate the Cox's failure rate $\underline{\lambda}(\underline{x})$ given in equation (1.22). In the following theorems, the $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution is characterized using the concepts of Cox's failure rate and bivariate mean residual life. The result for $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution follows as special case. The next theorem discusses the manifestation of the dullness property on the Cox's failure rate, $\underline{\lambda}(\underline{x})$.

Theorem 3.4 Let $\left(X_1,X_2\right)$ be a bivariate random variable with survival function $\bar{F}\left(x_1,x_2\right)$. Then $\bar{F}\left(x_1,x_2\right)$ belongs to the class of distribution specified by (3.6) if and only if the Cox's failure rate satisfies

$$\left(\lambda(tx),\lambda_{12}\left(tx_1 \mid tx_2\right),\lambda_{21}\left(tx_2 \mid tx_1\right)\right)=\left(\frac{\lambda(x)}{t},\frac{\lambda_{12}\left(x_1 \mid x_2\right)}{t},\frac{\lambda_{21}\left(x_2 \mid x_1\right)}{t}\right) \qquad (3.22)$$

for all $x, x_1, x_2, t>0$.

Proof

If $\bar{F}\left(x_1,x_2\right)$ belongs to the class of distribution specified by (3.6), then it follows that $\left(X_1,X_2\right)$ satisfy (3.3). Then the Cox's failure rate $\underline{\lambda}(\underline{x})$ is evaluated for $\bar{F}\left(tx_1,tx_2\right)$ as follows. By definition (1.22) for $x_1=x_2$,

$$\begin{aligned}\lambda(x)&=\frac{\partial}{\partial x}\log\left(\bar{F}\left(tx,tx\right)\right)\\&=\left[\frac{\partial}{\partial tx}\log\bar{F}\left(tx,tx\right)\right]\frac{\partial(xt)}{\partial x}\\&=\lambda(tx)\,t\end{aligned}$$

or

$$\lambda(tx) = \frac{\lambda(x)}{t}. \tag{3.23}$$

Now for $x_1 > x_2$, we have from (3.3)

$$\frac{\partial}{\partial x_2}\bar{F}(x_1 t, x_2 t) = \left(\frac{\partial}{\partial x_2}\bar{F}(x_1, x_2)\right)\bar{F}(t,t)$$

which is equivalent to writing,

$$\left(\frac{\partial}{\partial x_2 t}\bar{F}(x_1 t, x_2 t)\right)t = \left(\frac{\partial}{\partial x_2}\bar{F}(x_1, x_2)\right)\bar{F}(t,t).$$

Taking logarithm on both sides, we get

$$\log\left(\frac{\partial}{\partial x_2 t}\bar{F}(x_1 t, x_2 t)\right) + \log t = \log\left(\frac{\partial}{\partial x_2}\bar{F}(x_1, x_2)\right) + \log\bar{F}(t,t)$$

$$\frac{\partial}{\partial x_1}\log\left(\frac{\partial}{\partial x_2 t}\bar{F}(x_1 t, x_2 t)\right) = \frac{\partial}{\partial x_1}\log\left(\frac{\partial}{\partial x_2}\bar{F}(x_1, x_2)\right)$$

or

$$\frac{\partial}{\partial x_1 t}\log\left(\frac{\partial}{\partial x_2 t}\bar{F}(x_1 t, x_2 t)\right)t = \frac{\partial}{\partial x_1}\log\left(\frac{\partial}{\partial x_2}\bar{F}(x_1, x_2)\right)$$

giving

$$\lambda_{12}(x_1 t \mid x_2 t) = \frac{\lambda_{12}(x_1 \mid x_2)}{t}. \tag{3.24}$$

Similarly,

$$\lambda_{21}(x_2 t \mid x_1 t) = \frac{\lambda_{21}(x_2 \mid x_1)}{t}, \text{ for } x_1 < x_2. \tag{3.25}$$

Conversely, the general solution of $\lambda(xt) = \frac{1}{t}\lambda(x)$ is

$$\lambda(x) = \frac{c}{x} \text{ for some } c > 0.$$

Since $\lambda(x)$ denotes the failure rate of the $\min(X_1, X_2)$, it now follows that the class of distribution satisfying (3.6) is a Pareto minima class. Hence $\bar{F}(x,x) = x^{-(\alpha_1+\alpha_2)}$, for some $\alpha_1 + \alpha_2 > 0$, $x > 1$.

Also observing that (3.24) implies

$$\frac{\left(\frac{\partial^2}{\partial x_1 t\,\partial x_2 t}\bar{F}\left(x_1 t, x_2 t\right)\right)t^2}{\left(\frac{\partial}{\partial x_2 t}\bar{F}\left(x_1 t, x_2 t\right)\right)t} = \frac{\left(\frac{\partial^2}{\partial x_1\,\partial x_2}\bar{F}\left(x_1, x_2\right)\right)\bar{F}(t,t)}{\left(\frac{\partial}{\partial x_2}\bar{F}\left(x_1, x_2\right)\right)\bar{F}(t,t)}$$

or

$$\frac{\partial}{\partial x_1}\left(\log\frac{\partial}{\partial x_2}\bar{F}\left(x_1 t, x_2 t\right)\right) = \frac{\partial}{\partial x_1}\left(\log\frac{\partial}{\partial x_2}\bar{F}\left(x_1, x_2\right)\right)\bar{F}(t,t)$$

which in turn imply,

$$\bar{F}\left(x_1 t, x_2 t\right) = \bar{F}\left(x_1, x_2\right)\bar{F}(t,t) + c(t),\ x_2 \le x_1,\ t \ge 1. \tag{3.26}$$

For $x_1 = x_2 = 1$, it follows $c(t) = 0,$ hence satisfying equation (3.3).

Similarly proceeding for equation (3.25) we can conclude that

$$\bar{F}\left(x_1 t, x_2 t\right) = \bar{F}\left(x_1, x_2\right)\bar{F}(t,t),\ x_1 \le x_2,\ t \ge 1. \tag{3.27}$$

Hence the result.

Thus we have proved that a bivariate random variable satisfies the dullness property if and only if for all $x, x_1, x_2, t > 0,$ the equation (3.22) holds. In particular if we assume Pareto marginals for X_i, we have the following corollary.

Corollary 3.3 Let $\left(X_1, X_2\right)$ be a bivariate random variable with mixture Pareto marginals specified as in (3.7) then $\left(X_1, X_2\right)$ has a $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution with survival function specified by (3.8), if and only if the distribution of $\left(X_1, X_2\right)$ satisfies (3.22).

Corollary 3.5 Let $\left(X_1, X_2\right)$ be a bivariate random variable with Pareto marginals specified as in (3.10) then $\left(X_1, X_2\right)$ has a Marshall – Olkin type (Muliere and Scarsini (1987), Veenus and Nair (1994), Hanagal (1996)) bivariate Pareto distribution with survival function as given in (3.11) if and only if the distribution of $\left(X_1, X_2\right)$ satisfies (3.22).

The constancy of the product of the mean residual life function and failure rate characterizes the univariate Pareto distribution. For the $BP\ II(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ distribution the Cox's failure rate $\underline{\lambda}(\underline{x})$ is given by

$$\underline{\lambda}(\underline{x})=\left(\frac{(\alpha_1+\alpha_2)}{(\sigma+x-\mu)},\frac{\alpha_1'}{(\sigma+x_1-\mu)},\frac{\alpha_2'}{(\sigma+x_2-\mu)}\right) \tag{3.28}$$

and the corresponding bivariate mean residual life function is given by

$$\underline{m}(\underline{x})=\left(\frac{(\sigma+x-\mu)}{(\alpha_1+\alpha_2-1)},\frac{(\sigma+x_1-\mu)}{(\alpha_1'-1)},\frac{(\sigma+x_2-\mu)}{(\alpha_2'-1)}\right); \quad x,x_i>\mu,\ \mu,\sigma,\alpha_i,\alpha_i'>0,\ i=1,2. \tag{3.29}$$

with $m_{10}(x)=\dfrac{\alpha_1(\sigma+x-\mu)}{(\alpha_1+\alpha_2-1)(\alpha_1+\alpha_2)}$, $m_{20}(x)=\dfrac{\alpha_2(\sigma+x-\mu)}{(\alpha_1+\alpha_2-1)(\alpha_1+\alpha_2)}$.

In the next theorem we investigate the component wise product of (3.28) and (3.29) and conditions for a possible extension of the univariate characterization result.

Theorem 3.5 Let (X_1,X_2) be bivariate random variable in the support of $(\mu,\infty)\times(\mu,\infty)$ for $\mu>0$. Let $p_i=P[X_i<X_{3-i}]$, $i=1,2$ be known, then

$$\underline{\lambda}(\underline{x})\underline{m}(\underline{x})=\left(\frac{k}{k-1},\frac{k_1}{k_1-1},\frac{k_2}{k_2-1}\right). \tag{3.30}$$

For constants $k>1, k_i>1, i=1,2$ where $\underline{\lambda}(\underline{x})$ is of the form,

$$\underline{\lambda}(\underline{x})=\left(g(x),g_1(x_1),g_2(x_2)\right) \tag{3.31}$$

and satisfies

$$\frac{\lambda(x)}{\lambda_{ij}(x_i\mid x_j)}=\frac{k-x\lambda(x)}{k_i-x_i\lambda_{ij}(x_i\mid x_j)}=\frac{\lambda_{ij}(x_i\mid x_j)}{\lambda_{ji}(x_j\mid x_i)}=\frac{k_i-x_i\lambda_{ij}(x_i\mid x_j)}{k_j-x_j\lambda_{ji}(x_j\mid x_i)} \tag{3.32}$$

for $i \neq j = 1, 2,$ if and only if (X_1, X_2) has a $BP\ II(\mu, \sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2')$ distribution.

Proof

Let (X_1, X_2) be $BP\ II(\mu, \sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2')$ distribution, then from (3.28) and (3.29) it follows that

$$\underline{\lambda}(\underline{x})\underline{m}(\underline{x}) = \left(\frac{\alpha_1 + \alpha_2}{(\alpha_1 + \alpha_2 - 1)}, \frac{\alpha_1'}{\alpha_1' - 1}, \frac{\alpha_2'}{\alpha_2' - 1} \right).$$

Conversely assume that $\underline{\lambda}(\underline{x})\underline{m}(\underline{x}) = \left(\frac{k}{k-1}, \frac{k_1}{k_1 - 1}, \frac{k_2}{k_2 - 1} \right)$.

Then by using the inter-relationship between Cox's failure rate and BMRL given in (1.35) we have,

$$\lambda_{10}(x)m(x) = p_1 + \frac{d}{dx} m_{10}(x);\ x > \mu \tag{3.33}$$

$$\lambda_{20}(x)m(x) = p_2 + \frac{d}{dx} m_{20}(x);\ x > \mu \tag{3.34}$$

$$\lambda(x)m(x) = 1 + \frac{d}{dx} m(x);\ x > \mu \tag{3.35}$$

$$\lambda_{12}(x_1 \mid x_2) m_{12}(x_1 \mid x_2) = 1 + \frac{\partial}{\partial x_1} m_{12}(x_1 \mid x_2);\ x_1 > x_2 > \mu \tag{3.36}$$

and

$$\lambda_{21}(x_2 \mid x_1) m_{21}(x_2 \mid x_1) = 1 + \frac{\partial}{\partial x_2} m_{21}(x_2 \mid x_1);\ x_2 > x_1 > \mu\,. \tag{3.37}$$

From (3.35) it follows that

$$1 + \frac{d}{dx} m(x) = \frac{k}{k-1}.$$

Integrating with respect to x,

$$m(x) = \frac{x}{k-1} + c;\ x > \mu\,. \tag{3.38}$$

From (3.35) and (3.37), it follows that

$$m_{12}(x_1 \mid x_2) = \frac{x_1}{k_1 - 1} + c_1(x_2)$$

and

$$m_{21}(x_2 \mid x_1) = \frac{x_2}{k_2 - 1} + c_2(x_1)\,.$$

Now from equation (1.35) and (3.31), it follows that

$$\lambda(x) = \frac{k}{x + (k-1)c}\,, \tag{3.39}$$

$$\lambda_{12}(x_1 \mid x_2) = \frac{k_1}{x_1 + (k_1 - 1)c_1(x_2)} \tag{3.40}$$

and

$$\lambda_{21}(x_2 \mid x_1) = \frac{k_2}{x_2 + (k_2 - 1)c_2(x_1)}. \tag{3.41}$$

$\lambda(x)$ satisfies the condition (3.32) implies that

$$(k-1)c = (k_1 - 1)c_1(x_2) = (k_2 - 1)c_2(x_1)\,.$$

Consider the equation (3.38), by putting $x = \mu,$ we obtain

$$(k-1)c = (k-1)m(\mu) - \mu.$$

Hence (3.39) becomes,

$$\lambda(x) = \frac{k}{x - \mu + (k-1)m(\mu)}$$

as $(k-1)c = (k_1 - 1)c_1(x_2) = (k_2 - 1)c_2(x_1),$ we have

$$\underline{\lambda}(\underline{x}) = \left(\frac{k}{x - \mu + (k-1)m(\mu)}, \frac{k_1}{x_1 - \mu + (k-1)m(\mu)}, \frac{k_2}{x_2 - \mu + (k-1)m(\mu)} \right). \tag{3.42}$$

The equation (3.42) is the reciprocal linear in x, x_1 and x_2 respectively and from equation (3.28), it follows that (X_1, X_2) has $BP\ II(\mu, \sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2')$ distribution with

$$\sigma = (k-1)m(\mu),\ \ \alpha_1 = p_1 k,\ \ \alpha_2 = p_2 k,\ \ \alpha_1 + \alpha_2 = k,\ \ \alpha_1' = k_1,\ \ \alpha_2' = k_2\,.$$

Hence the result.

When considering the reliability characteristics in the bivariate case there is a need to choose an appropriate extension which reflects the ageing characteristics of the bivariate distribution. Thus in this chapter we advocate the use of Cox's failure rate and related reliability concepts for studying models exhibiting load sharing dependence. This forms a reasonable motivation to formulate a definition of information measure for the residual life distribution corresponding to a bivariate distribution with load sharing dependence. Accordingly the next chapter deals with this.

Chapter 4

Residual Measure of Uncertainty for Bivariate Distributions with Load Sharing Dependence

4.1 Introduction

In reliability, survival analysis, actuary and many other fields the study of duration is a subject of interest. Capturing effects of the age t of an individual or an item under study on the information about the remaining life time is important in many applications. The entropy function that takes age into consideration was introduced by Ebrahimi (1996) by modifying the information measure defined by Shannon (1948). This chapter deals with the uncertainty associated with a two component parallel system having load sharing dependence.

The chapter is organized into three sections. In section 4.2, the concept of bivariate residual entropy function is introduced along with the properties, monotonicity and uniqueness. Section 4.3 deals with characterizations of life time models using the concept of bivariate residual entropy function. The equivalence of bivariate dullness property and BLMP in terms of bivariate residual entropy

---

Some results in this chapter is published in Asha, Jagathnath and Nair (2009).

function has also been established. The relationship of the bivariate mean residual life function with the proposed measure is also discussed.

4.2 Bivariate Residual Entropy Function

As previously assumed let (X_1, X_2) is a vector of non-negative random variables. Typically we think of $X_i,\ i=1,2$ to be the lifetime of a two component parallel system. However if we are considering a parallel system where failure of the system consists of failure of one of the component first and the eventual failure of the other, it is useful to take this knowledge when considering the still surviving components residual life. In this case Cox's (1972) failure rate is an apt bivariate failure rate to model the failure times. The underlying probability density function can be expressed in terms of Cox's (1972) failure rate through the expression given in (1.23). A suitable measure to model the mean residual life of the system is the bivariate mean residual life function defined in (1.33). These definitions form a reasonable motivation to formulate a definition of information measure for the residual life distributions corresponding to the two stage of failure in the system. The first stage corresponds to a state when both the components are functioning while the second stage corresponds to a state when one of the components has failed.

Accordingly we propose a new measure of bivariate residual entropy vector

$$H(f,t,t_1,t_2) = \left(H(f_Z,t), H_{12}(f,t_1,t_2), H_{21}(f,t_1,t_2)\right)$$

applicable to a two component parallel system. Here $H\left(f_Z,t\right)$ corresponds to the first stage while $H_{12}(f,t_1,t_2)$ and $H_{21}(f,t_1,t_2)$ corresponds to the second stage of failure of the system. When the system is in a state of perfect functioning (that is both the components are working),

$$H(f_Z,t)=1-\frac{1}{\bar{F}_Z(t)}\int_t^{\infty}\lambda(x)\bar{F}_Z(x)\log\lambda(x)dx;\ t>0 \tag{4.1}$$

measures the uncertainty contained in the conditional density of $(X-t)$ given that both the component has survived time t. The second stage of the system would be when one of the components has failed. The

$$H_{12}(f,t_1,t_2)=1-\frac{1}{\left[\frac{\partial}{\partial u}\bar{F}(t_1,u)\right]_{u=t_2}}\int_{t_1}^{\infty}\lambda_{12}(x_1\mid t_2)\left[\frac{\partial}{\partial u}\bar{F}(x_1,u)\right]_{u=t_2}\log\lambda_{12}(x_1\mid t_2)dx_1$$

$$;t_1>t_2 \tag{4.2}$$

measures the expected uncertainty associated with the random variable (X_1-t_1) when the first component has survived beyond t_1 and the second component has failed at time t_2. Similarly,

$$H_{21}(f,t_1,t_2)=1-\frac{1}{\left[\frac{\partial}{\partial u}\bar{F}(u,t_2)\right]_{u=t_1}}\int_{t_2}^{\infty}\lambda_{21}(x_2\mid t_1)\left[\frac{\partial}{\partial u}\bar{F}(u,x_2)\right]_{u=t_1}\log\lambda_{21}(x_2\mid t_1)dx_2$$

$$;t_1<t_2 \tag{4.3}$$

measures the expected uncertainty associated with the random variable (X_2-t_2) when the first component has survived beyond t_2 and the second component has failed at time t_1.

Thus bivariate residual entropy function associated with the random vector $\left((X-t\mid X>t),(X_1-t_1\mid X_1>t_1,X_2=t_2),(X_2-t_2\mid X_2>t_2,X_1=t_1)\right)$ is defined as

$$H(f,t,t_1,t_2)=\left(H(f_Z,t),H_{12}(f,t_1,t_2),H_{21}(f,t_1,t_2)\right). \tag{4.4}$$

When X_1 and X_2 are independent $H(f,t,t_1,t_2)$ reduces to

$$\left(H(f_Z,t),H(f_1,t_1),H(f_2,t_2)\right), \tag{4.5}$$

where $H(f_Z,t)$ is as defined in (4.1) and $H(f_i,t_i)$ are the marginal entropy of X_i's, $i=1,2$ defined in (1.46). Also observe that

$$\frac{d}{dt}H\left(f_Z,t\right)=\lambda(t)\log\lambda(t)-\lambda(t)\left[1-H\left(f_Z,t\right)\right]$$

$$=\lambda(t)\log\lambda(t)-\lambda(t)\log\lambda(t)-\int_t^\infty\frac{\frac{d}{dx}\lambda(x)}{\lambda(x)}\bar{F}_Z(x)$$

$$\leq(\geq)0\text{, according as }\frac{d}{dx}\lambda(x)\geq(\leq)0.$$

Similarly,

$$\frac{\partial}{\partial t_1}H_{12}\left(f,t_1,t_2\right)=\lambda_{12}(t_1\mid t_2)\log\lambda_{12}(t_1\mid t_2)-\lambda_{12}(t_1\mid t_2)\left[1-H_{12}\left(f,t_1,t_2\right)\right]$$

$$=\lambda_{12}(t_1\mid t_2)\log\lambda_{12}(t_1\mid t_2)-\lambda_{12}(t_1\mid t_2)\log\lambda_{12}(t_1\mid t_2)$$

$$-\int_{t_1}^\infty\frac{\frac{\partial}{\partial x_1}\lambda_{12}(x_1\mid t_2)}{\lambda_{12}(x_1\mid t_2)}\left[\frac{\partial}{\partial u}\bar{F}\left(x_1,u\right)\right]_{u=t_2}dx_1$$

which implies $\frac{\partial}{\partial t}H_{12}\left(f,t_1,t_2\right)\leq(\geq)0$ according as $\frac{\partial}{\partial x_1}\lambda_{12}\left(x_1\mid t_2\right)\leq(\geq)0$. Similar argument holds for $H_{21}\left(f,t_1,t_2\right)$. Thus we can say that monotonicity of Cox's failure rate gives sufficient condition for monotonicity of $H(f,t,t_1,t_2)$.

It can also be observed that bivariate residual entropy function $H(f,t,t_1,t_2)$ is not invariant under non-singular transformations. If $Y=\phi(X)$ and $Y_i=\phi(X_i),\ i=1,2$ are one to one transformations, then

$$H(h_Y,t)=H(f_Z,t)-E\left[\log J(Y)\mid X>t\right] \tag{4.6}$$

and

$$H_{ij}(h,t_1,t_2)=H_{ij}(f,t_1,t_2)-E\left[\log J(Y_i)\mid X_i>t_i,X_j=t_j\right],\quad i,j=1,2,\ i\neq j \tag{4.7}$$

where $h_Y(y)$ is the density function corresponding to the random variable Y and $h(y_i)$ that of the random variable $Y_i,\ i=1,2.$ $J(Y)=\left|\frac{\partial\phi^{-1}(y)}{\partial y}\right|$, $J(Y_i)=\left|\frac{\partial\phi_i^{-1}(y_i)}{\partial y}\right|$, $i=1,2$ are the Jacobian of transformation and the expectation of (4.6) is taken with respect to the residual distribution of X and that corresponding to (4.7) is with respect to the residual distribution of X_i given X_j has failed at t_j.

One of the attractive features of residual entropy was that it uniquely determines the underlying distribution function (Belzunce et al. (2004)). It is now natural to ask if the bivariate residual entropy enjoys this property. As such it does not uniquely determine the underlying distribution; however under certain conditions it is possible to get a unique representation. The following theorem discusses the condition under which such a representation can be obtained.

Theorem 4.1 If (X_1, X_2) has an absolutely continuous distribution $F(x_1, x_2)$ with $p_i = P\left[X_i < X_j\right],\ i, j = 1,2,\ i \neq j,$ known and $H(f_Z, t)$ non decreasing in t, $H_{12}(f, t_1, t_2)$ non decreasing in t_1 and $H_{21}(f, t_1, t_2)$ non decreasing in t_2, then $H(f, t, t_1, t_2)$ uniquely determine the underlying distribution.

Proof

From definition of bivariate residual entropy function, we have,

$$\int_t^\infty \lambda(x)\bar{F}_Z(x)\log\lambda(x)dx = \bar{F}_Z(t)\left[1 - H(f_Z, t)\right],$$

$$\int_{t_1}^\infty \lambda_{12}(x_1 \mid t_2)\left[\frac{\partial \bar{F}(x_1, u)}{\partial u}\right]_{u=t_2} \log\lambda_{12}(x_1 \mid t_2)dx_1 = \left[\frac{\partial \bar{F}(x_1, u)}{\partial u}\right]_{u=t_2}\left[1 - H_{12}(f, t_1, t_2)\right],$$

$$\int_{t_2}^\infty \lambda_{21}(x_2 \mid t_1)\left[\frac{\partial \bar{F}(u, x_2)}{\partial u}\right]_{u=t_1} \log\lambda_{21}(x_2 \mid t_1)dx_2 = \left[\frac{\partial \bar{F}(u, x_2)}{\partial u}\right]_{u=t_1}\left[1 - H_{21}(f, t_1, t_2)\right].$$

Differentiating the above equations with respect to t, t_1 and t_2, we obtain

$$\frac{d}{dt}H(f_Z, t) = \lambda(x)\left[H(f_Z, t) + \log\lambda(t) - 1\right], \tag{4.8}$$

$$\frac{\partial}{\partial t_1}H_{12}(f, t_1, t_2) = \lambda_{12}(t_1 \mid t_2)\left[H_{12}(f, t_1, t_2) + \log\lambda_{12}(t_1 \mid t_2) - 1\right], \tag{4.9}$$

$$\frac{\partial}{\partial t_2}H_{21}(f, t_1, t_2) = \lambda_{21}(t_2 \mid t_1)\left[H_{21}(f, t_1, t_2) + \log\lambda_{21}(t_2 \mid t_1) - 1\right]. \tag{4.10}$$

Considering equation (4.8), and proceeding in similar lines as in Belzunce et al. (2004), for a fixed $t>0$, $\lambda(t)$ is a positive solution of the equation,

$$g(y)=y\left[H\left(f_Z,t\right)+\log y-1\right]-\frac{d}{dt}H\left(f_Z,t\right)=0 \tag{4.11}$$

since $g(0)=-\frac{d}{dt}H\left(f_Z,t\right)\leq 0$, $g(+\infty)=+\infty$ and $\frac{d}{dy}g\left(y\right)=H\left(f_Z,t\right)+\log y.$

Now considering (4.9) and (4.10), for $t_1,t_2>0$, $\lambda_{12}(t_1\,|\,t_2)$ and $\lambda_{21}(t_2\,|\,t_1)$ are respectively the positive solution of the following equations

$$g(y_1)=y_1\left[H_{12}(f,t_1,t_2)-\log y_1-1\right]-\frac{\partial}{\partial t_1}H_{12}(f,t_1,t_2)=0 \tag{4.12}$$

and

$$g(y_2)=y_2\left[H_{21}(f,t_1,t_2)-\log y_2-1\right]-\frac{\partial}{\partial t_2}H_{21}(f,t_1,t_2)=0\,. \tag{4.13}$$

Proceeding in similar arguments as in equation (4.11), equations (4.12) and (4.13) has positive solution $\lambda_{12}(t_1\,|\,t_2)$ and $\lambda_{21}(t_2\,|\,t_1)$ respectively for all t_1 and t_2. Thus when $p_i,\ i=1,2$ is known the distribution recovered using unique expression of Cox's failure rate given in equation (1.23).

In particular if $H(f,t,t_1,t_2)$ is of the form $\left(k,k_1(t_2),k_2(t_1)\right)$ where k is some constant and $k_i(t_j),\ i=1,2,\ i\neq j$ are functions of $t_j,\ j=1,2$ then solving (4.11), (4.12) and (4.13) we have

$$\underline{\lambda}(\underline{t})=\left(e^{1-k},e^{1-k_1(t_2)},e^{1-k_2(t_1)}\right).$$

Few distributions belonging to this class are given in Table 4.2.

4.3 Characterizations Using Bivariate Residual Entropy Function

In univariate case the Pareto distribution is characterized by the dullness property. The residual information of this distribution is independent of the age of the component. Extending this to the bivariate case, if X_1 and X_2 are independent it can be inferred from (3.1) that X_i's are independent Pareto and is characterized by

$$\bar{F}(x_1t_1, x_2t_2) = \bar{F}(x_1, x_2)\bar{F}(t_1, t_2)$$

for all $x_1, x_2, t_1, t_2 \geq 0$. Since this model is of no practical use we consider the extension of dullness discussed in Hanagal (1996).

A random vector (X_1, X_2) with survival function $\bar{F}(x_1, x_2)$ in the support of $(1,\infty)\times(1,\infty)$ is said to have bivariate dullness property if

$$\bar{F}(x_1t, x_2t) = \bar{F}(x_1, x_2)\bar{F}(t, t) \tag{4.14}$$

for all x_1, x_2 and $t \geq 0$.

A general form of $\bar{F}(x_1, x_2)$ having bivariate dullness property (Asha and Jagathnath (2006)) is given in (3.6), which is the class of distribution having Pareto minima. The next theorem explores the implication of bivariate dullness on $H(f, t, t_1, t_2)$.

Theorem 4.2 A bivariate random vector (X_1, X_2) with survival function $\bar{F}(x_1, x_2)$ satisfies bivariate dullness property (4.14), if and only if for all $t' > 1$.

$$\begin{aligned}&\left(H(f_Z, tt'), H_{12}(f, t_1t', t_2t'), H_{21}(f, t_1t', t_2t')\right)\\ &\quad = \left(H(f_Z, t) + \log t', H_{12}(f, t_1, t_2) + \log t', H_{21}(f, t_1, t_2) + \log t'\right).\end{aligned} \tag{4.15}$$

Proof

The Theorem 3.4 established the equivalence of bivariate dullness property in terms of Cox failure rate as

$$\left(\lambda(tt'), \lambda_{12}(t_1t' \mid t_2t'), \lambda_{21}(t_2t' \mid t_1t')\right) = \left(\frac{\lambda(t)}{t'}, \frac{\lambda_{12}(t_1 \mid t_2)}{t'}, \frac{\lambda_{21}(t_2 \mid t_1)}{t'}\right). \tag{4.16}$$

Since $\lambda(t)$ is unique positive solution of (4.11), we have

$$\lambda(t) = e^{1-H(f_Z, t)}, \tag{4.17}$$

then

$$\lambda(tt') = e^{1-H(f_Z, tt')}. \tag{4.18}$$

Now from (4.16) it follows that

$$\frac{\lambda(t)}{t'} = e^{1-H(f_Z,\, tt')}.$$

Using (4.17), the above equation can be written as

$$e^{1-H(f_Z,\, t)} = t'\, e^{1-H(f_Z,\, tt')}.$$

Now by taking logarithm on both sides, we get

$$1 - H\left(f_Z, t\right) = \log t' + 1 - H\left(f_Z, tt'\right)$$

i.e.

$$H\left(f_Z, tt'\right) = H\left(f_Z, t\right) + \log t'. \tag{4.19}$$

Also $\lambda_{12}(t_1 \mid t_2)$ is unique positive solution of (4.12), we have

$$\lambda_{12}(t_1 \mid t_2) = e^{1-H_{12}(f,\, t_1, t_2)}, \tag{4.20}$$

then

$$\lambda_{12}(t_1 t' \mid t_2 t') = e^{1-H_{12}(f,\, t_1 t', t_2 t')}. \tag{4.21}$$

Now from (4.16) it follows that

$$\frac{\lambda_{12}(t_1 \mid t_2)}{t'} = e^{1-H_{12}(f,\, t_1 t', t_2 t')}.$$

Using (4.20), the above equation can be written as

$$e^{1-H_{12}(f,\, t_1, t_2)} = t'\, e^{1-H_{12}(f,\, t_1 t', t_2 t')}.$$

Now by taking logarithm on both sides, we get

$$1 - H_{12}\left(f, t_1, t_2\right) = \log t' + 1 - H_{12}\left(f, t_1 t', t_2 t'\right).$$

i.e.

$$H_{12}\left(f, t_1 t', t_2 t'\right) = H_{12}\left(f, t_1, t_2\right) + \log t'. \tag{4.22}$$

Similarly we can obtain

$$H_{21}\left(f, t_1 t', t_2 t'\right) = H_{21}\left(f, t_1, t_2\right) + \log t'. \tag{4.23}$$

Thus from (4.19), (4.22) and (4.23), the equation (4.15) follows. Now if (4.15) is satisfied then (4.16) holds true and hence the bivariate dullness property (4.14) is satisfied. Conversely if (4.14) holds then (4.15) follows in a very straight forward manner from (4.1), (4.2) and (4.3).

Table 4.1 summaries a few members of the class characterized by Theorem 4.2.

The intimate relationship between the Pareto and exponential distribution permits us to obtain analogous characterization of the former from the latter, thus we have the following corollary.

Corollary 4.1 A bivariate random vector with survival function $\overline{F}\left(x_1, x_2\right)$ satisfies BLMP

$$\overline{F}(x_1+t, x_2+t) = \overline{F}(x_1, x_2)\overline{F}(t,t) \tag{4.24}$$

if and only if for all $t' > 0$.

$$\begin{aligned}&\left(H(f_Z, t+t'), H_{12}(f, t_1+t', t_2+t'), H_{21}(f, t_1+t', t_2+t')\right)\\ &\qquad = \left(H(f_Z, t), H_{12}(f_1, t_1, t_2), H_{21}(f, t_1, t_2)\right).\end{aligned} \tag{4.25}$$

Proof

Basu and Sun (1997) has established the equivalence of BLMP in terms of Cox failure rate as

$$\left(\lambda(t+t'), \lambda_{12}(t_1+t' \mid t_2+t'), \lambda_{21}(t_2+t' \mid t_1+t')\right) = \left(\lambda(t), \lambda_{12}(t_1 \mid t_2), \lambda_{21}(t_2 \mid t_1)\right). \tag{4.26}$$

Since $\underline{\lambda}(\underline{t})$ is unique positive solution of (4.11), (4.12) and (4.13) and proceeding as in Theorem 4.2, it follows that if (4.25) is satisfied then (4.26) holds true and hence the BLMP (4.24) is satisfied. Conversely if (4.24) holds, then (4.25) follows in a very straight forward manner from (4.1), (4.2) and (4.3).

Table 4.2 summaries a few members of the class characterized by Corollary 4.1.

Asadi and Ebrahimi (2000) characterized the generalized Pareto distribution by the relationship between residual entropy and mean residual life function. In the following theorem the extension of this result to the bivariate case is investigated.

Theorem 4.3 Let $\left(X_1, X_2\right)$ be non-negative random variable admitting an absolutely continuous distribution function in the support of $(\mu, R)\times(\mu, R)$, $0 \le \mu < R < \infty$ with $E\left[X_i\right] < \infty,\ i = 1,2,$ then

$$H\left(f_Z, t\right) = 1 - k + \log m(t) \text{ for all } t > 0$$

and

$$H_{ij}\left(f, t_1, t_2\right) = 1 - k_i + \log m_{ij}(t_i \mid t_j) \text{ for all } t_i > t_j,\ i \ne j = 1,2. \tag{4.27}$$

only if

$$\underline{\lambda}(\underline{t})\underline{m}(\underline{t}) = \left(e^{k}, e^{k_1}, e^{k_2}\right) \tag{4.28}$$

for a constant vector $\underline{k} = \left(k, k_1, k_2\right)$.

Proof

Differentiating equations in (4.27) with respect to t, t_1 and t_2, we get

$$\frac{d}{dt} H(f_Z, t) = \frac{\left(\frac{d}{dt} m(t)\right)}{m(t)},$$

$$\frac{\partial}{\partial t_1} H_{12}(f, t_1, t_2) = \frac{\left(\frac{\partial}{\partial t_1} m_{12}(t_1 \mid t_2)\right)}{m_{12}(t_1 \mid t_2)} \tag{4.29}$$

and

$$\frac{\partial}{\partial t_2} H_{21}(f, t_1, t_2) = \frac{\left(\frac{\partial}{\partial t_2} m_{21}(t_2 \mid t_1)\right)}{m_{21}(t_2 \mid t_1)}.$$

Substituting (4.29) and (4.27) in (4.8), (4.9) and (4.10), we get

$$\frac{d}{dt} m(t) = m(t)\lambda(t)\left[\log \lambda(t) + \log m(t) - k\right],$$

$$\frac{\partial}{\partial t_1} m_{12}(t_1 \mid t_2) = m_{12}(t_1 \mid t_2)\lambda_{12}(t_1 \mid t_2)\left[\log \lambda_{12}(t_1 \mid t_2) + \log m_{12}(t_1 \mid t_2) - k_1\right]$$

and

$$\frac{\partial}{\partial t_2} m_{21}(t_2 \mid t_1) = m_{21}(t_2 \mid t_1)\lambda_{21}(t_2 \mid t_1)\left[\log \lambda_{21}(t_2 \mid t_1) + \log m_{21}(t_2 \mid t_1) - k_2\right]. \tag{4.30}$$

Using the inter-relationship between Cox failure rate and BMRL given in equation (1.35), (4.30) becomes

$$\begin{aligned}
&c(t) = c(t)\left[\log c(t) - k\right] \\
&c_{12}(t_1 \mid t_2) = c_{12}(t_1 \mid t_2)\left[\log c_{12}(t_1 \mid t_2) - k_1\right] \qquad (4.31) \\
&c_{21}(t_2 \mid t_1) = c_{21}(t_2 \mid t_1)\left[\log c_{21}(t_2 \mid t_1) - k_2\right].
\end{aligned}$$

where

$$\begin{aligned}
&c(t) = \lambda(t) m(t) \\
&c_{12}(t_1 \mid t_2) = \lambda_{12}(t_1 \mid t_2) m_{12}(t_1 \mid t_2) \\
&c_{21}(t_2 \mid t_1) = \lambda_{21}(t_2 \mid t_1) m_{21}(t_2 \mid t_1). \qquad (4.32)
\end{aligned}$$

Once again differentiating the expressions in (4.31) with respect to t, t_1 and t_2, we get

$$\begin{aligned}
&\frac{d}{dt} c(t) = \left[\log c(t) - k\right] \frac{d}{dt} c(t) + \frac{d}{dt} c(t) \\
&\frac{\partial}{\partial t_1} c_{12}(t_1 \mid t_2) = \left[\log c_{12}(t_1 \mid t_2) - k_1\right] \frac{\partial}{\partial t_1} c_{12}(t_1 \mid t_2) + \frac{\partial}{\partial t_1} c_{12}(t_1 \mid t_2) \\
&\frac{\partial}{\partial t_2} c_{21}(t_2 \mid t_1) = \left[\log c_{21}(t_2 \mid t_1) - k_2\right] \frac{\partial}{\partial t_2} c_{21}(t_2 \mid t_1) + \frac{\partial}{\partial t_2} c_{21}(t_2 \mid t_1).
\end{aligned}$$

which implies

$$\log c(t) \frac{d}{dt} c(t) = k \frac{d}{dt} c(t)$$

Thus we get $c(t) = e^k$.

Similarly,

$$\begin{aligned}
&c_{12}(t_1 \mid t_2) = e^{k_1}, \\
&c_{21}(t_2 \mid t_1) = e^{k_2}.
\end{aligned}$$

Therefore, $\left(c(t), c_{12}(t_1 \mid t_2), c_{21}(t_2 \mid t_1)\right) = \left(e^k, e^{k_1}, e^{k_2}\right)$.

Hence the theorem.

In particular when $\underline{k} = (0,0,0)$ it follows that $\underline{\lambda}(\underline{t})\underline{m}(\underline{t}) = (1,1,1)$. From (1.35), we have

$$\frac{d}{dt}m(t)=\frac{\partial}{\partial t_i}m_{ij}\left(t_i \mid t_j\right)=0,\ j=1,2,\ i\neq j\,.$$

Under the assumption

$$m_{ij}\left(t_i \mid 0\right)=m_{ij}\left(t_i \mid t_j\right)\ \text{for all}\ t_i>t_j,\ i,\,j=1,2,\ i\neq j \tag{4.33}$$

we have

$$m\left(t\right)=m\left(0\right)\ \text{for all}\ t>0$$

and

$$m_{ij}\left(t_i \mid t_j\right)=m_{ij}\left(0 \mid 0\right)\ \text{for all}\ t_i>t_j.$$

Once again from (1.35) it follows that

$$\underline{\lambda}(\underline{t})=\left(\frac{1}{m},\frac{1}{m_{12}},\frac{1}{m_{21}}\right),$$

where m denotes $m(0)$ and m_{ij} denotes $m_{ij}(0 \mid 0)$. Given

$$p_i=P\left[X_i>X_j\right] \tag{4.34}$$

it follows from unique expression of the Cox's failure rate and the condition $\int_0^{\infty}\int_0^{\infty} f(x_1,x_2)dx_1dx_2=1$ that

$$f(t_1,t_2)=\begin{cases}\alpha_1\alpha_2'\ \mathrm{e}^{[-\alpha_2' t_2-(\alpha_1+\alpha_2-\alpha_2')t_1]};\ 0<t_1<t_2\\ \alpha_2\alpha_1'\ \mathrm{e}^{[-\alpha_1' t_1-(\alpha_1+\alpha_2-\alpha_1')t_2]};\ 0<t_2<t_1\end{cases} \tag{4.35}$$

where $\alpha_i=\dfrac{p_i}{m}$ and $\alpha_i'=\dfrac{1}{m_{ij}},\ i,\,j=1,2,\ i\neq j$ which is the Freund (1961) bivariate exponential distribution.

Thus along with Table 4.3 we have proved the following corollary.

Corollary 4.2 Let $\left(X_1,X_2\right)$ be a non-negative random vector admitting an absolutely continuous distribution function such that $E\left[X_i\right]<\infty,\ i=1,2.$ Then under conditions (4.33) and (4.34),

$$H\left(f_Z,t\right)=1+\log m(0)\ \text{for all}\ t>0$$

$$H_{ij}\left(f,t_i,t_j\right)=1+\log m_{ij}(0 \mid 0)\ \text{for all}\ t_i>t_j,\ i,\,j=1,2,\ i\neq j.$$

If and only if (X_1, X_2) has Freund (1961) bivariate exponential given by (4.35). When $k, k_1, k_2 > 0,$ it then follows that $\underline{\lambda}(\underline{t})\underline{m}(\underline{t}) > (1,1,1)$. Now let $\underline{\lambda}(\underline{t})\underline{m}(\underline{t}) > (a, a_1, a_2)$, where $a, a_1, a_2 > 1$. Apart from conditions (4.33) and (4.34) assume that

$$\frac{m}{a-1} = \frac{m_{12}}{a_1 - 1} = \frac{m_{21}}{a_2 - 1}. \tag{4.36}$$

Then from (1.35) it follows that

$$\underline{\lambda}(\underline{t}) = \left(\frac{a \mid (a-1)}{t + m \mid (a-1)}, \frac{a_1 \mid (a_1 - 1)}{t_1 + m \mid (a-1)}, \frac{a_2 \mid (a_2 - 1)}{t_2 + m \mid (a-1)} \right).$$

From the unique expression of Cox failure rate (1.23) and since $\int\limits_{\mu}^{\infty}\int\limits_{\mu}^{\infty} f(x_1, x_2) dx_1 dx_2 = 1$, we have

$$f(t_1, t_2) = \begin{cases} \dfrac{\alpha_1 \alpha_2'}{\sigma^2} \left(1 + \dfrac{t_1 - \mu}{\sigma}\right)^{-(\alpha_1 + \alpha_2 - \alpha_2' + 1)} \left(1 + \dfrac{t_2 - \mu}{\sigma}\right)^{-(\alpha_2' + 1)} & ;\ \mu < t_1 < t_2 \\ \dfrac{\alpha_2 \alpha_1'}{\sigma^2} \left(1 + \dfrac{t_2 - \mu}{\sigma}\right)^{-(\alpha_1 + \alpha_2 - \alpha_1' + 1)} \left(1 + \dfrac{t_1 - \mu}{\sigma}\right)^{-(\alpha_1' + 1)} & ;\ \mu < t_2 < t_1 \end{cases} \tag{4.37}$$

where $\alpha_i = \dfrac{p_i a}{a-1}$, $\alpha_i' = \dfrac{a_i}{a_i - 1}$ and $\sigma = \mu + \dfrac{m}{a-1}$, $i = 1, 2$ which is the bivariate Pareto II distribution (Asha and Jagathnath, 2008).

A similar argument holds for $k, k_1, k_2 < 0$ or equivalently $a, a_1, a_2 < 1$ and the joint density is obtained as

$$f(t_1, t_2) = \begin{cases} \dfrac{\alpha_1 \alpha_2'}{R^2} \left(1 - \dfrac{t_1}{R}\right)^{(\alpha_1 + \alpha_2 - \alpha_2' - 1)} \left(1 - \dfrac{t_2}{R}\right)^{(\alpha_2' - 1)} & ; 0 < t_1 < t_2 < R \\ \dfrac{\alpha_2 \alpha_1'}{R^2} \left(1 - \dfrac{t_2}{R}\right)^{(\alpha_1 + \alpha_2 - \alpha_1' - 1)} \left(1 - \dfrac{t_1}{R}\right)^{(\alpha_1' - 1)} & ; 0 < t_2 < t_1 < R \end{cases} \tag{4.38}$$

where $\alpha_i = \dfrac{p_i a}{1-a}$, $\alpha_i' = \dfrac{a_i}{1 - a_i}$, $R = \dfrac{m}{1-a}$, $\mu = 0$, $i = 1, 2$, which is a bivariate finite range distribution.

Corollary 4.3 Let $\left(X_1, X_2\right)$ be non-negative random variable admitting an absolutely continuous distribution function in the support of $(\mu, R)\times(\mu, R)$, $0 \le \mu < R < \infty$ with $E\left[X_i\right] < \infty$, $i = 1, 2$, then under the condition (4.33), (4.34) and (4.35)

$$H\left(f_Z, t\right) = 1 - k + \log m(t) \text{ for all } t > o$$

$$H_{ij}\left(f, t_1, t_2\right) = 1 - k_i + \log m_{ij}(t_i \mid t_j) \text{ for all } t_i > t_j,\ i \ne j = 1, 2.$$

If and only if

(i) For $k, k_1, k_2 > 0$, $\left(X_1, X_2\right)$ follows a bivariate Pareto distribution specified in (4.37).

(ii) For $k, k_1, k_2 < 0$, $\left(X_1, X_2\right)$ follows a bivariate finite range distribution specified in (4.38).

The next theorem will establish the characteristic property of the bivariate Pareto I distribution using the relationship between the bivariate residual entropy function and the bivariate geometric vitality function.

Theorem 4.4 Let $\left(X_1, X_2\right)$ be a random vector admitting an absolutely continuous distribution function in the support of $(\sigma, \infty)\times(\sigma, \infty)$ with bivariate geometric vitality function $GV\left(t, t_1, t_2\right)$ (1.42), bivariate residual entropy $H(f, t, t_1, t_2)$ and $p_i = P\left[X_i > X_j\right]$, $i, j = 1, 2$, $i \ne j$ known, then the relation

$$H\left(f_Z, t\right) - \log GV(t) = c,$$

$$H_{12}\left(f, t_1, t_2\right) - \log GV_{12}(t_1 \mid t_2) = c_1$$

and

$$H_{21}\left(f, t_1, t_2\right) - \log GV_{21}(t_2 \mid t_1) = c_2 \tag{4.39}$$

where c, c_1, c_2 are constants holds for all $t, t_1, t_2 > 0$ if and only if (X_1, X_2) follows bivariate Pareto I ($BPI(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2')$) distribution specified (2.7)

Proof

Let $BPI(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ as given in (2.7) then from definition of bivariate residual entropy function and bivariate geometric vitality function we have,

$$H\left(f,t,t_1,t_2\right)=\left(\frac{\alpha_1+\alpha_2+1}{\alpha_1+\alpha_2}-\log\left(\frac{\alpha_1+\alpha_2}{t}\right),\frac{\alpha_1'+1}{\alpha_1'}-\log\left(\frac{\alpha_1'}{t_1}\right),\frac{\alpha_2'+1}{\alpha_2'}-\log\left(\frac{\alpha_2'}{t_2}\right)\right)$$

and

$$\log GV\left(t,t_1,t_2\right)=\left(\log t+\frac{1}{\alpha_1+\alpha_2},\log t_1+\frac{1}{\alpha_1'},\log t_2+\frac{1}{\alpha_2'}\right).$$

Then

$$H\left(f_Z,t\right)-\log GV(t)=1-\log(\alpha_1+\alpha_2)$$
$$H_{12}\left(f,t_1,t_2\right)-\log GV_{12}(t_1\mid t_2)=1-\log(\alpha_1')$$
$$H_{21}\left(f,t_1,t_2\right)-\log GV_{21}(t_2\mid t_1)=1-\log(\alpha_2').$$

Now conversely assume that equation (4.39) holds and using (1.42) and (4.4), we will get the following expressions

$$(c-1)\bar{F}_T(t)=-\int_t^\infty \lambda(x)\bar{F}_X(x)\log\lambda(x)dx-\int_t^\infty \lambda(x)\bar{F}_X(x)\log x dx. \tag{4.40}$$

$$\begin{aligned}(c_1-1)\left[\frac{\partial\bar{F}(t_1,u)}{\partial u}\right]_{u=t_2}&=-\int_{t_1}^\infty \lambda_{12}(x_1\mid t_2)\left[\frac{\partial\bar{F}(x_1,u)}{\partial u}\right]_{u=t_2}\log\lambda_{12}(x_1\mid t_2)dx_1\\&\quad-\int_{t_1}^\infty \lambda_{12}(x_1\mid t_2)\left[\frac{\partial\bar{F}(x_1,u)}{\partial u}\right]_{u=t_2}\log x_1dx_1.\end{aligned} \tag{4.41}$$

$$\begin{aligned}(c_2-1)\left[\frac{\partial\bar{F}(u,t_2)}{\partial u}\right]_{u=t_1}&=-\int_{t_2}^\infty \lambda_{21}(x_2\mid t_1)\left[\frac{\partial\bar{F}(u,x_2)}{\partial u}\right]_{u=t_1}\log\lambda_{12}(x_2\mid t_1)dx_2\\&\quad-\int_{t_2}^\infty \lambda_{21}(x_2\mid t_1)\left[\frac{\partial\bar{F}(u,x_2)}{\partial u}\right]_{u=t_1}\log x_2dx_2.\end{aligned} \tag{4.42}$$

Differentiating (4.40), (4.41) and (4.42) with respect to t,t_1,t_2, we get

$$\log\left[t\lambda(t)\right]=1-c$$
$$\log\left[t_1\lambda_{12}(t_1\mid t_2)\right]=1-c_1$$

$$\log\left[t_2\lambda_{21}(t_2 \mid t_1)\right] = 1 - c_2.$$

i.e.

$$\lambda(t) = \frac{e^{1-c}}{t},\ \lambda_{12}(t_1 \mid t_2) = \frac{e^{1-c_1}}{t_1},\ \lambda_{21}(t_2 \mid t_1) = \frac{e^{1-c_2}}{t_2}.$$

That is $\underline{\lambda}(\underline{t}) = \left(\frac{e^{1-c}}{t}, \frac{e^{1-c_1}}{t_1}, \frac{e^{1-c_2}}{t_2}\right)$.

The Cox's (1972) total failure rate uniquely determines the distribution and $\underline{\lambda}(\underline{t})$ is reciprocal linear in t, t_1, t_2 respectively characterizes $BPI\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distribution with $\alpha_1 + \alpha_2 = e^{1-c}$, $\alpha_1' = e^{1-c_1}$, $\alpha_2' = e^{1-c_2}$, $\alpha_1 = p_1 e^{1-c}$ and $\alpha_2 = p_2 e^{1-c}$.

The very same arguments that motivated a definition of the bivariate residual entropy function makes us seek an appropriate inequality measure for a bivariate data form a model with load sharing dependence. Hence in the next chapter we look into inequality measures for bivariate distributions with load sharing dependence.

Table 4.1 Class of distribution possessing bivariate dullness property

Bivariate survival function	Bivariate residual entropy function
Bivariate Pareto I (Asha & Jagathnath (2006)) distribution $\bar{F}(x_1,x_2)=\begin{cases}\frac{1}{\alpha_1+\alpha_2-\alpha_2'}\left\{\alpha_1\, x_1^{-(\alpha_1+\alpha_2-\alpha_2')}x_2^{-\alpha_2'}+(\alpha_2-\alpha_2')\,x_2^{-(\alpha_1+\alpha_2)}\right\};\ 1\le x_1\le x_2\\ \frac{1}{\alpha_1+\alpha_2-\alpha_1'}\left\{\alpha_2\, x_2^{-(\alpha_1+\alpha_2-\alpha_1')}x_1^{-\alpha_1'}+(\alpha_1-\alpha_1')\,x_1^{-(\alpha_1+\alpha_2)}\right\};\ 1\le x_2\le x_1\end{cases}$ $\alpha_i>0,\alpha_i'>0,\alpha_1+\alpha_2\ne\alpha_i',i=1,2.$	$\left(\frac{\alpha_1+\alpha_2+1}{\alpha_1+\alpha_2}-\log\left(\frac{\alpha_1+\alpha_2}{t}\right),\ \frac{\alpha_1'+1}{\alpha_1'}-\log\left(\frac{\alpha_1'}{t_1}\right),\ \frac{\alpha_2'+1}{\alpha_2'}-\log\left(\frac{\alpha_2'}{t_2}\right)\right)$
Mulerie and Scarcini (1987) bivariate Pareto I distribution $\bar{F}(x_1,x_2)=x_1^{-\alpha_1}x_2^{-\alpha_2}\left\{\max\left(x_1,x_2\right)\right\}^{-\alpha_0};x_1,x_2\ge 1$ $\alpha_0,\alpha_1,\alpha_2>0$	$\left(\frac{\alpha_1+\alpha_2+\alpha_0+1}{\alpha_1+\alpha_2+\alpha_0}-\log\left(\frac{\alpha_1+\alpha_2+\alpha_0}{t}\right),\ \frac{\alpha_1+\alpha_0+1}{\alpha_1+\alpha_0}-\log\left(\frac{\alpha_1+\alpha_0}{t_1}\right),\ \frac{\alpha_2+\alpha_0+1}{\alpha_2+\alpha_0}-\log\left(\frac{\alpha_2+\alpha_0}{t_2}\right)\right)$
Independent Pareto I distribution $\bar{F}(x_1,x_2)=x_1^{-\alpha_1}x_2^{-\alpha_2};x_1,x_2\ge 1,\ \alpha_1,\alpha_2>0$	$\left(\frac{\alpha_1+\alpha_2+1}{\alpha_1+\alpha_2}-\log\left(\frac{\alpha_1+\alpha_2}{t}\right),\ \frac{\alpha_1+1}{\alpha_1}-\log\left(\frac{\alpha_1}{t_1}\right),\frac{\alpha_2+1}{\alpha_2}-\log\left(\frac{\alpha_2}{t_2}\right)\right)$

Table 4.2 Class of distribution possessing BLMP

Bivariate survival function	Bivariate residual entropy function
Freund (1961) bivariate exponential $$\bar{F}(x_1,x_2)=\begin{cases}\dfrac{1}{\alpha_1+\alpha_2-\alpha_2'}\left\{\alpha_1\ e^{-\alpha_2' x_2-(\alpha_1+\alpha_2-\alpha_2')x_1}+(\alpha_2-\alpha_2')e^{-(\alpha_1+\alpha_2)x_2}\right\};\ 0\le x_1\le x_2\\ \dfrac{1}{\alpha_1+\alpha_2-\alpha_1'}\left\{\alpha_2\ e^{-\alpha_1' x_1-(\alpha_1+\alpha_2-\alpha_1')x_2}+(\alpha_1-\alpha_1')e^{-(\alpha_1+\alpha_2)x_1}\right\};\ 0\le x_2\le x_1\end{cases}$$ $\alpha_i>0,\alpha_i'>0,\alpha_1+\alpha_2\ne\alpha_i',i=1,2.$	$\left(1-\log(\alpha_1+\alpha_2),1-\log\alpha_1',1-\log\alpha_2'\right)$
Marshall and Olkin (1967) bivariate exponential $$\bar{F}(x_1,x_2)=\begin{cases}e^{(-\lambda_1 x_1-(\lambda_2+\lambda_{12})x_2)};\ 0\le x_1\le x_2\\ e^{(-\lambda_2 x_2-(\lambda_1+\lambda_{12})x_1)};\ 0\le x_2\le x_1\end{cases}$$ $\lambda_1,\lambda_2,\lambda_{12}>0$	$(1-\log\lambda,$ $1-\dfrac{\lambda_1\lambda}{(\lambda_1+\lambda_2)}\log\left(\dfrac{\lambda_2\lambda(\lambda_1+\lambda_{12})}{(\lambda_1+\lambda_2)}\right.$ $1-\dfrac{\lambda_2\lambda}{(\lambda_1+\lambda_2)}\log\left(\dfrac{\lambda_1\lambda(\lambda_2+\lambda_{12})}{(\lambda_1+\lambda_2)}\right.$ $\lambda=\lambda_2+\lambda_1+\lambda_{12}$
Independent exponential distribution $\bar{F}(x_1,x_2)=e^{(-\lambda_1 x_1-\lambda_2 x_2)};\ x_1,x_2\ge 0,\ \lambda_1,\lambda_2>0$	$\left(1-\log(\lambda_1+\lambda_2),1-\log\lambda_1,1-\log\lambda_2\right)$

Table 4.3 Expression for total failure rate, BMRL and bivariate residual entropy function.

Bivariate density function	$\underline{\lambda}(\underline{t})$	$\underline{m}(\underline{t})$	$H(f,t,t_1,t_2)$
Bivariate Freund exponential distribution (1961) $f(x_1,x_2)=\begin{cases}\alpha_1\alpha_2' \, e^{-\alpha_2' x_2-(\alpha_1+\alpha_2-\alpha_2')x_1}; & 0<x_1<x_2\\ \alpha_2\alpha_1' \, e^{-\alpha_1' x_1-(\alpha_1+\alpha_2-\alpha_1')x_2}; & 0<x_2<x_1\end{cases}$ $\alpha_i>0, \alpha_i'>0, \alpha_1+\alpha_2\neq\alpha_i', i=1,2.$	$(\alpha_1+\alpha_2, \alpha_1', \alpha_2')$	$\left(\frac{1}{\alpha_1+\alpha_2}, \frac{1}{\alpha_1'}, \frac{1}{\alpha_2'}\right)$	$(1-\log(\alpha_1+\alpha_2), 1-\log\alpha_1', 1-\log\alpha_2')$
Bivariate Pareto II distribution (Asha & Jagathnath (2008)) $f(x_1,x_2)=\begin{cases}\frac{\alpha_1\alpha_2'}{\sigma^2}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_2'+1)}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-(\alpha_2'+1)}; & \mu<x_1<x_2\\ \frac{\alpha_2\alpha_1'}{\sigma^2}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-(\alpha_1+\alpha_2-\alpha_1'+1)}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-(\alpha_1'+1)}; & \mu<x_2<x_1\end{cases}$ $\alpha_i>1, \alpha_i'>1, \alpha_1+\alpha_2\neq\alpha_i', i=1,2.$	$\left(\frac{\alpha_1+\alpha_2}{\sigma+t-\mu}, \frac{\alpha_1'}{\sigma+t_1-\mu}, \frac{\alpha_2'}{\sigma+t_2-\mu}\right)$	$\left(\frac{\sigma+t-\mu}{\alpha_1+\alpha_2-1}, \frac{\sigma+t_1-\mu}{\alpha_1'-1}, \frac{\sigma+t_2-\mu}{\alpha_2'-1}\right)$	$\left(\frac{\alpha_1+\alpha_2+1}{\alpha_1+\alpha_2}-\log\left(\frac{\alpha_1+\alpha_2}{\sigma+t-\mu}\right), \frac{\alpha_1'+1}{\alpha_1'}-\log\left(\frac{\alpha_1'}{\sigma+t_1-\mu}\right), \frac{\alpha_2'+1}{\alpha_2'}-\log\left(\frac{\alpha_2'}{\sigma+t_2-\mu}\right)\right)$
Bivariate Finite Range distribution $f(x_1,x_2)=\begin{cases}\frac{\alpha_1\alpha_2'}{R^2}\left(1-\frac{x_1}{R}\right)^{(\alpha_1+\alpha_2-\alpha_2'-1)}\left(1-\frac{x_2}{R}\right)^{(\alpha_2'-1)}; & 0<x_1<x_2<R\\ \frac{\alpha_2\alpha_1'}{R^2}\left(1-\frac{x_2}{R}\right)^{(\alpha_1+\alpha_2-\alpha_1'-1)}\left(1-\frac{x_1}{R}\right)^{(\alpha_1'-1)}; & 0<x_2<x_1<R\end{cases}$ $\alpha_i>0, \alpha_i'>0, \alpha_1+\alpha_2\neq\alpha_i', R>0, i=1,2.$	$\left(\frac{\alpha_1+\alpha_2}{R-t}, \frac{\alpha_1'}{R-t_1}, \frac{\alpha_2'}{R-t_2}\right)$	$\left(\frac{R-t}{\alpha_1+\alpha_2+1}, \frac{R-t_1}{\alpha_1'+1}, \frac{R-t_2}{\alpha_2'+1}\right)$	$\left(\frac{\alpha_1+\alpha_2-1}{\alpha_1+\alpha_2}-\log\left(\frac{\alpha_1+\alpha_2}{R-t}\right), \frac{\alpha_1'-1}{\alpha_1'}-\log\left(\frac{\alpha_1'}{R-t_1}\right), \frac{\alpha_2'-1}{\alpha_2'}-\log\left(\frac{\alpha_2'}{R-t_2}\right)\right)$

Chapter 5

Inequality Measures for Bivariate Distributions with Load Sharing Dependence

5.1 Introduction

In the univariate case inequality measures related to or identical to the Lorenz order have gained general acceptance. Another measure of inequality for univariate populations which stands out from the rest in terms of acceptance and applicability is the Gini index.

It is only natural to seek appropriate extensions of the Lorenz curves and Gini index to higher dimensions. Some early works suggesting these extensions can be found in Taguchi (1972 a), Lunetta (1972), De Simoni (1979). Taguchi (1972 a) defined the concentration surface of a two dimensional random vector and extended the notions of concentration surface to complete surface called the Lorenz manifold. In Arnold (1983) a parametric representation of the bivariate Lorenz curve is given as

$$L(u,v)=\frac{\int_0^x\int_0^y \xi\eta f_{12}(\xi,\eta)d\xi d\eta}{E[XY]} \tag{5.1}$$

Some results in this chapter have been communicated in Asha and Jagathnath (2010) where f_{12} denotes the joint income density and $f_i, i=1,2$ denotes the marginals corresponding to the non-negative random variables X and Y respectively. Here $u=\int_0^x f_1(\xi)d\xi$ and $v=\int_0^y f_2(\eta)d\eta$. Koshevoy (1995) provides a definition in higher dimensions in terms of the Lorenz zonoids and the inequality measures for multivariate distributions are given in Arnold (2005).

Let $\underline{X}$ denote a k-dimensional non-negative vector with positive finite expectations. Let $\psi^k(\underline{x})$ be a measurable mapping from $R^k \to [0,1]$. Then the Lorenz zonoid of $\underline{X}$ denoted by $L(\underline{X})$ is defined by

$$L(\underline{X})=\left\{\int \psi(\underline{x})dF(\underline{x}), \int \frac{x_1\psi(\underline{x})}{E[X_1]}dF(\underline{x}),\ldots,\int \frac{x_k\psi(\underline{x})}{E[X_k]}dF(\underline{x})\right\}, \psi(k)\in \psi^k. \quad (5.2)$$

This definition is consistent with the univariate definition and higher order extensions are straight forward.

The Gini index is another popular inequality measure defined in terms of geometric features of the Lorenz curve. It represents twice the area between Lorenz curve of X and the line of equality. The Gini index has also been extended to higher dimensions. Mosler (2002) extended it as the volume of the Lorenz zonoid and calls it the Gini zonoid index, Weymark (2004) describes parameterized family of multivariate generalized Gini indices. A Gini index for truncated multivariate distributions was proposed by Sathar et al. (2007) which was consistent with that of Ord et al. (1983) for the univariate case.

However when studying distributions in higher dimensions there is a need to choose an appropriate measure that reflect appropriate aggregation aspects of the inequality when comparisons are to be made. This choice is very much related to the dependency enjoyed by the multivariate distributions. In this chapter we consider the load sharing dependence. This dependence finds exclusive use in socio-economic problems, where income from multiple sources is considered. To take a very simple situation one could envisage a unit with two sources of independent income. For example, it could be the income from farming two different crops or income of a couple to the household. Both the source work independently till one of the source is unable to generate an income. Then the other source has either generates more income or perhaps be affected adversely. Thus the income distribution of the surviving source undergoes a parameter change and hence a load-sharing model is now apt to model the dependence. The measures (5.1) and (5.2) fail to reflect the inequality aspect of the data. A simple and effective way to formulate a measure to reflect the aggregate aspects of the inequalities in the data is to consider the two different states of the income generating source. Firstly, both the sources are generating an income and secondly, only one source is functional. Hence it is now more adequate to study the related one dimensional distribution

(i) $P\left[Z \geq x\right] = \bar{F}_Z(x)$, where $Z = min\left(X_1, X_2\right)$

(ii) $P\left[X_2 \geq x \mid X_1 = x_1\right] = \left[-\frac{\partial}{\partial u}\bar{F}(u,x)\right]_{u=x_1}$

(iii) $P\left[X_1 \geq x \mid X_2 = x_2\right] = \left[-\frac{\partial}{\partial u}\bar{F}(x,u)\right]_{u=x_2}$

together than the bivariate distribution $F\left(x_1, x_2\right)$. In this chapter we study the inequality measures specifically, Lorenz function and Gini index of these distributions and characterize the original bivariate distribution using these measures.

The proposed study is organized into three sections. In the second section, we propose the definitions of the income measures, Lorenz curve and Gini index which reflect the inequalities in the data taking into account the information regarding the two different states of the income generating source. We also study the theoretical properties of the inequality measures. In the third section, characterizations using these measures are discussed.

5.2 Definitions and Relationships

Let (X_1, X_2) be a vector of non-negative random variables admitting absolutely continuous distribution function $F(x_1, x_2)$ and density function $f(x_1, x_2)$. Let $\bar{F}(x_1, x_2) = P[X_1 \geq x_1, X_2 \geq x_2]$. Also assume that $E[X_i] < \infty, i = 1, 2$, then (X_1, X_2) could represent income from two different crops or income in a household unit from two independent sources. When the income from a unit goes below a threshold level say $x, x \in R^+$, then the load of generating more income falls on the other source. Assume that the income is reported only if at least one of the sources has an income that exceeds x. Then the possibilities are

(i) Income from both the sources exceed x.

(ii) At least one of the sources have an income larger than x, while the other has an income less than x.

In the former situation if $Z = min(X_1, X_2)$ and $F_Z, f_Z, \bar{F}_Z$ denotes the distribution function, density function and survival function respectively, associated with Z, then the proportion of units when both the sources are generating income greater than x, and up to t is given by

$$P_x(t) = \frac{\int_x^t dF_Z(y)}{\bar{F}_Z(x)}. \tag{5.3}$$

The cumulative income share of this population is given by

$$L(P_x(t)) = \frac{\int_x^t y dF_Z(y)}{\int_x^\infty \text{y } dF_Z(y)}, x < t. \tag{5.4}$$

Thus $(P_x(t), L(P_x(t)))$ represents the inequality in income of the population when both the sources generate an income of at least x. However when the income from one source falls below the threshold value x, then the above measure ceases to be an apt measure. Assume that the income from the second source is $x_2, x_2 < x$ and income of the first source is greater than x, then

$$P_{x|x_2}(t) = \frac{\int_x^t f(y, x_2)dy}{\left[-\frac{\partial}{\partial u}\bar{F}(x,u)\right]_{u=x_2}}, \; x_2 < x < t \tag{5.5}$$

denotes the proportion of population whose first resource has an income greater than x and up to t while the second source generates an income $x_2 < x$. The cumulative income share of the population (5.5) is

$$L\left(P_{x|x_2}(t)\right) = \frac{\int_x^t \mathrm{y}\, f(y, x_2)dy}{\int_x^\infty \mathrm{y}\, f(y, x_2)dy}, \; x_2 < x. \tag{5.6}$$

Similar interpretation follows for

$$L\left(P_{x|x_1}(t)\right) = \frac{\int_x^t \mathrm{y}\, f(x_1, y)dy}{\int_x^\infty \mathrm{y}\, f(x_1, y)dy}, \; x_1 < x \tag{5.7}$$

where

$$P_{x|x_1}(t) = \frac{\int_x^t f(x_1, y)dy}{\left[-\frac{\partial}{\partial u}\bar{F}(u,x)\right]_{u=x_1}}, \; x_1 < x < t. \tag{5.8}$$

The three measures namely

$$\left(P_x(t), L\left(P_x(t)\right)\right),\; \left(P_{x|x_2}(t), L\left(P_{x|x_2}(t)\right)\right) \text{ and } \left(P_{x|x_1}(t), L\left(P_{x|x_1}(t)\right)\right) \tag{5.9}$$

are consistent with the univariate definition of the Lorenz curve and reflect inequality in a bivariate income data under specific situations mentioned in (i) and (ii).

Note: Since the above definitions can be seen as univariate measures in their domains of definition it is not difficult to observe that

(i) $P_x(t), P_{x|x_i}(t), i=1,2$ are continuous in $[0,1]$.

(ii) $L\left(P_x(t)\right), L\left(P_{x|x_i}(t)\right), i=1,2 \to 0(1)$ according as $P_x(t), P_{x|x_i}(t) \to 0(1)$.

(iii) $L\left(P_x(t)\right), L\left(P_{x|x_i}(t)\right), i=1,2$ are increasing in x.

(iv) $L\left(P_x(t)\right), L\left(P_{x|x_i}(t)\right), i=1,2$ is convex in $P_x(t), P_{x|x_i}(t)$ respectively.

The Gini index is a popular inequality measure closely related to the Lorenz curve, though there are various definitions for the Gini index unrelated to Lorenz curve. But it should also be noted the various definitions agree with each other. One definition closely associated to the Lorenz function is given by

$$G_x = 2\int_x^\infty P_x(t)\frac{d}{dt}L\left(P_x(t)\right)-1. \tag{5.10}$$

Analogously using the inequality measures in (5.3) to (5.8), we have the definition for Gini index as $\underline{G}(\underline{x})=\left(G_x, G_{x|x_2}, G_{x|x_1}\right)$, where G_x is defined in (5.10) and

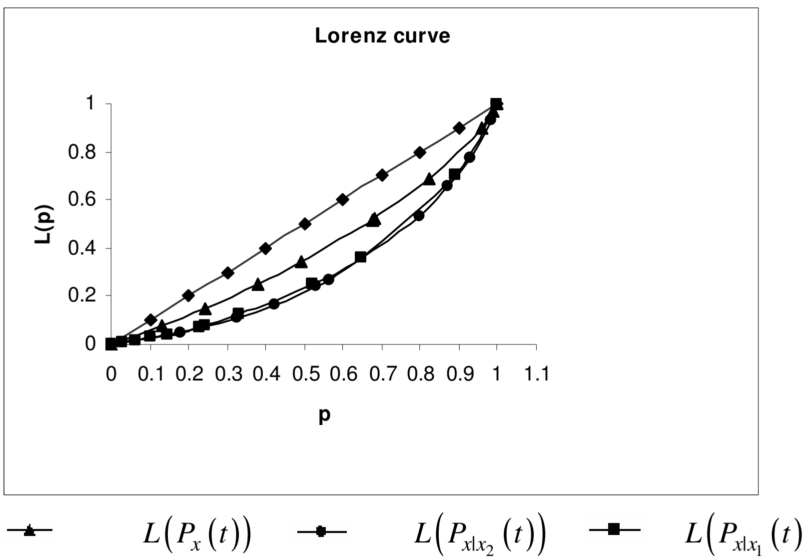
$$G_{x|x_i} = 2\int_x^\infty P_{x|x_i}(t)\frac{d}{dt}L\left(P_{x|x_i}(t)\right)-1, x_i < x, i=1,2. \tag{5.11}$$

We can use the Lorenz curve to obtain a different interpretation of lifetime data. Let us illustrate the Lorenz function for a Freund bivariate exponential (refer Table 5.2) data (Kim and Kvam (2004)). The data is shown to exhibit a load sharing dependence (Deshpande et al. (2007)). The estimates of the parameters are obtained as $\alpha_1 = 0.18, \alpha_2 = 0.35, \alpha_1' = 0.22$ and $\alpha_2' = 0.29$. The values of the Lorenz functions are given in Table 5.1.

Table 5.1 Values of the Lorenz functions

$P_x(t)$	$L(P_x(t))$	$P_{x\mid x_2}(t)$	$L(P_{x\mid x_2}(t))$	$P_{x\mid x_1}(t)$	$L(P_{x\mid x_1}(t))$
0.9275	0.7716	0.1814	0.0471	0.0314	0.0074
0.2354	0.0672	0.2354	0.0672	0.0618	0.0154
0.3240	0.1070	0.3240	0.1070	0.1043	0.0278
0.5646	0.2679	0.4231	0.1629	0.1450	0.0412
0.9833	0.9271	0.5329	0.2415	0.2275	0.0729
0.1814	0.0471	0.5646	0.2679	0.2430	0.0796
0.4231	0.1629	0.7962	0.5305	0.3298	0.1219
0.5329	0.2415	0.8705	0.6535	0.5213	0.2479
0.7962	0.5305	0.9275	0.7716	0.6480	0.3630
0.8705	0.6535	0.9833	0.9271	0.8900	0.7017

Figure 5.1 Lorenz curve corresponding to the failure time data



From the Figure 5.1, it can be inferred that there is less disparity among the samples when both the units have exceeded the truncation point (assumed as unity in this case) than the Lorenz curve when one of the components lie below unity.

As suggested in Chandra and Singpurwalla (1981) it is interesting to use certain ideas used in reliability theory to derive the theoretical properties of the

Lorenz curve and Gini index. This motivated us to study the relationship between the bivariate Lorenz curve and the bivariate mean residual life function, defined as

$$\underline{m}(\underline{x}) = \left(m(x), m_{12}\left(x \mid x_2\right), m_{21}\left(x \mid x_1\right)\right) \tag{5.12}$$

where

$$m(x) = \frac{\int_x^\infty \bar{F}_Z(y)dy}{\bar{F}_Z(x)}$$

$$m_{ij}\left(x \mid x_j\right) = \frac{\int_x^\infty \left(y - x_j\right) f\left(y, x_j\right) dy}{\left[-\dfrac{\partial \bar{F}(x,u)}{\partial u}\right]_{u=x_j}}, i \neq j = 1, 2.$$

The mean residual life function defined above is related to the Cox's (1972) failure rate $\underline{\lambda}(\underline{x}) = \left(\lambda(x), \lambda_{12}\left(x \mid x_2\right), \lambda_{21}\left(x \mid x_1\right)\right)$ by

$$\frac{1 + \dfrac{d}{dx} m(x)}{m(x)} = \lambda(x) = \frac{f_Z(x)}{\bar{F}_Z(x)}$$

$$\frac{1 + \dfrac{\partial}{\partial x} m_{ij}(x \mid x_j)}{m_{ij}(x \mid x_j)} = \lambda_{ij}(x \mid x_j) = \frac{\dfrac{\partial^2}{\partial x\, \partial x_j} \bar{F}(x, x_j)}{\left[-\dfrac{\partial}{\partial u} \bar{F}(x,u)\right]_{u=x_j}}, x_j < x, i = 3 - j,\ j = 1, 2.$$

(5.13)

Theorem 5.1 Let $\left(X_1, X_2\right)$ be a bivariate random variable admitting an absolutely continuous distribution function and having finite expectations. Further if $L\left(P_x(t)\right), L\left(P_{x|x_2}(t)\right), L\left(P_{x|x_1}(t)\right)$ are differentiable, then

$$-\frac{d}{dt} \log\left[1 - L\left(P_x(t)\right)\right] = \frac{t\left(1 + \dfrac{d}{dt} m(t)\right)}{m(t)\left[t + m(t)\right]} \tag{5.14}$$

$$-\frac{\partial}{\partial t} \log\left[1 - L\left(P_{x|x_2}(t)\right)\right] = \frac{t\left(1 + \dfrac{\partial}{\partial t} m_{12}(t \mid x_2)\right)}{m_{12}(t \mid x_2)\left[t + m_{12}(t \mid x_2)\right]} \tag{5.15}$$

$$-\frac{\partial}{\partial t}\log\left[1-L\left(P_{x|x_1}(t)\right)\right]=\frac{t\left(1+\frac{\partial}{\partial t}m_{21}(t\mid x_1)\right)}{m_{21}(t\mid x_1)\left[t+m_{21}(t\mid x_1)\right]}. \tag{5.16}$$

Proof

From (5.4), we have

$$L\left(P_x(t)\right)=\frac{\int_x^t ydF_Z(y)}{\int_x^\infty ydF_Z(y)}.$$

$$1-L\left(P_x(t)\right)=\frac{\int_t^\infty ydF_Z(y)}{\int_x^\infty ydF_Z(y)}$$

$$=\frac{t\overline{F}_Z(t)+\int_t^\infty \overline{F}_Z(y)dy}{\mu_Z(x)}$$

where $\mu_Z(x)=\int_x^\infty ydF_Z(y)$.

$$\mu_Z(x)\left[1-L\left(P_x(t)\right)\right]=t\overline{F}_Z(t)+\int_t^\infty \overline{F}_Z(y)dy$$

From (5.12) the above relation becomes

$$\overline{F}_Z(x)\left[x+m(x)\right]\left[1-L\left(P_x(t)\right)\right]=\overline{F}_Z(t)\left[t+m(t)\right]. \tag{5.17}$$

Note that, since (5.17) represent the cumulative income larger than t, it is only natural that right hand side of (5.17) is independent of x since $t > x$.

Since $\frac{d}{dt}L\left(P_x(t)\right)=\frac{t\,f(t)}{\overline{F}_Z(x)\left[x+m(x)\right]}$, we have using (5.17),

$$\left(\frac{1}{1-L\left(P_x(t)\right)}\right)\frac{d}{dt}L\left(P_x(t)\right)=\frac{t\,\lambda(t)}{\left[t+m(t)\right]}$$

where $\lambda(t)$ is as in (5.13). Substituting for $\lambda(t)$, we have

$$-\frac{d}{dt}\log\left[1-L\left(P_x(t)\right)\right]=\frac{t\left(1+\frac{d}{dt}m(t)\right)}{m(t)\left[t+m(t)\right]}$$

Now consider $L\left(P_{x|x_2}(t)\right)$ as in (5.6)

$$L\left(P_{x|x_2}(t)\right)=1-\frac{\int_t^\infty y\, f(y,x_2)dy}{\int_x^\infty y\, f(y,x_2)dy},\ x_2<x$$

or

$$1-L\left(P_{x|x_2}(t)\right)=\frac{\int_t^\infty y\, f(y,x_2)dy}{\int_x^\infty y\, f(y,x_2)dy},\ x_2<x$$

$$=\frac{t\left[\frac{\partial}{\partial u}\bar{F}(t,u)\right]_{u=x_2}+\int_t^\infty\left[\frac{\partial}{\partial u}\bar{F}(y,u)\right]_{u=x_2}dy}{\mu(x,x_2)}$$

where $\mu(x,x_2)=\int_x^\infty y\, f(y,x_2)dy,\ x_2<x$.

So that

$$\mu(x,x_2)\left[1-L\left(P_{x|x_2}(t)\right)\right]=\left[\frac{\partial}{\partial u}\bar{F}(t,u)\right]_{u=x_2}\left[t+m_{12}\left(t\mid x_2\right)\right],\ x_2<t.$$

Since $\frac{\partial}{\partial u}L\left(P_{x|x_2}(t)\right)=\frac{t\, f\left(t,x_2\right)}{\mu(x,x_2)}$, we have

$$-\frac{\partial}{\partial t}\log\left[1-L\left(P_{x|x_2}(t)\right)\right]=\frac{tf\left(t,x_2\right)}{\left[\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\left[t+m_{12}\left(t\mid x_2\right)\right]}$$

$$=\frac{t\left[1+\frac{\partial}{\partial t}m_{12}\left(t\mid x_2\right)\right]}{m_{12}\left(t\mid x_2\right)\left[t+m_{12}\left(t\mid x_2\right)\right]}.$$

Proceeding in a similar manner with $L\left(P_{x|x_1}(t)\right)$ we can prove that

$$-\frac{\partial}{\partial t}\log\left[1-L\left(P_{x|x_1}(t)\right)\right]=\frac{t\left[1+\frac{\partial}{\partial t}m_{21}(t\mid x_1)\right]}{m_{21}(t\mid x_1)\left[t+m_{21}(t\mid x_1)\right]}.$$

Remark 5.1 In equation (5.17), when $x=0,$ it reduces to the result of Chandra and Singpurwalla (1981).

It is natural to investigate if there exist a similar relationship between the bivariate Gini index and bivariate mean residual life function. The next theorem deals with this.

Theorem 5.2 Under the usual assumptions and conditions specified in Theorem 5.1, the following relationships hold,

$$\left[1-G_x\right]\left[x+m(x)\right]=x+\frac{1}{\left[\bar{F}_Z(x)\right]^2}\int_x^{\infty}\left[\bar{F}_Z(x)\right]^2 dt \tag{5.18}$$

$$\left[1-G_{x|x_2}\right]\left[x+m_{12}(x\mid x_2)\right]=x+\frac{1}{\left\{\left[\frac{\partial}{\partial u}\bar{F}(x,u)\right]_{u=x_2}\right\}^2}\int_x^{\infty}\left\{\left[\frac{\partial}{\partial u}\bar{F}(t,u)\right]_{u=x_2}\right\}^2 dt \tag{5.19}$$

$$\left[1-G_{x|x_1}\right]\left[x+m_{21}(x\mid x_1)\right]=x+\frac{1}{\left\{\left[\frac{\partial}{\partial u}\bar{F}(u,x)\right]_{u=x_1}\right\}^2}\int_x^{\infty}\left\{\left[\frac{\partial}{\partial u}\bar{F}(u,t)\right]_{u=x_1}\right\}^2 dt. \tag{5.20}$$

Proof

From (5.10) it follows that

$$G_x=2\int_x^{\infty}P_x(t)\frac{dL\left(P_x(t)\right)}{dt}-1.$$

Which from (5.3) and (5.4) can be rewritten as

$$G_x = \frac{2\int_x^\infty \left[\frac{F_Z(t)-F_Z(x)}{\bar{F}_Z(x)}\right] t\ f_Z(t)dt}{\int_x^\infty y\ f_Z(y)dy} - 1.$$

Which is same as

$$G_x \int_x^\infty y\ f_Z(y)dy = 2\int_x^\infty \left[\frac{F_Z(t)-F_Z(x)}{\bar{F}_Z(x)}\right] t\ f_Z(t)dt - \int_x^\infty y\ f_Z(y)dy$$

or

$$G_x \int_x^\infty y\ f_Z(y)dy = 2\int_x^\infty \frac{F_Z(t)\ t\ f_Z(t)}{\bar{F}_Z(x)}dt - \frac{2\left[1-\bar{F}_Z(x)\right]}{\bar{F}_Z(x)}\int_x^\infty t\ f_Z(t)dt - \int_x^\infty y\ f_Z(y)dy$$

$$= 2\int_x^\infty \frac{F_Z(t)\ t\ f_Z(t)}{\bar{F}_Z(x)}dt - \frac{2}{\bar{F}_Z(x)}\int_x^\infty t\ f_Z(t)dt + 2\int_x^\infty t\ f_Z(t)dt - \int_x^\infty y\ f_Z(y)dy$$

So that

$$G_x \int_x^\infty y\ f_Z(y)dy = 2\int_x^\infty \frac{F_Z(t)\ t\ f_Z(t)}{\bar{F}_Z(x)}dt - \frac{2}{\bar{F}_Z(x)}\int_x^\infty t\ f_Z(t)dt + \int_x^\infty t\ f_Z(t)dt.$$

Observing that $\dfrac{\int_x^\infty y\ f_Z(y)dy}{\bar{F}_Z(x)} = x + m(x)$, we have

$$G_x\left[x+m(x)\right] = 2\int_x^\infty \frac{F_Z(t)\ t\ f_Z(t)}{\left[\bar{F}_Z(x)\right]^2}dt - \frac{2}{\left[\bar{F}_Z(x)\right]^2}\int_x^\infty t\ f_Z(t)dt + \left[x+m(x)\right]$$

or

$$\left[1-G_x\right]\left[x+m(x)\right] = \frac{2}{\left[\bar{F}_Z(x)\right]^2}\int_x^\infty t\ f_Z(t)\bar{F}_Z(t)dt. \tag{5.21}$$

Consider $I = \int_x^\infty t\ f_Z(t)\bar{F}_Z(t)dt$

$$= -\int_x^\infty \left(t\ \bar{F}_Z(t)\right)d\bar{F}_Z(t)$$

$$= -\left(t\ \bar{F}_Z(t)\right)\int_x^\infty d\bar{F}_Z(t) + \int_x^\infty \left[-\left(t\ f(t)\right)\int d\bar{F}_Z(t)\right]dt - \int_x^\infty \left[\bar{F}_Z(t)\int d\bar{F}_Z(t)\right]dt$$

$$= -\left(t\,\overline{F}_Z(t)\right)\int_x^\infty d\overline{F}_Z(t) - \int_x^\infty t\ f(t)\overline{F}_Z(t)dt + \int_x^\infty \left[\overline{F}_Z(t)\right]\left[\overline{F}_Z(t)\right]dt$$

$$= x\left[\overline{F}_Z(x)\right]^2 - \int_x^\infty t\ f(t)\overline{F}_Z(t)dt + \int_x^\infty \left[\overline{F}_Z(t)\right]^2 dt\,.$$

Thus

$$2I = x\left[\overline{F}_Z(x)\right]^2 + \int_x^\infty \left[\overline{F}_Z(t)\right]^2 dt$$

Substituting for $2I$ in (5.21) we have the result.

To prove the second equality, from (5.11), (5.5) and (5.6), it follows that

$$G_{x|x_2} = 2\int_x^\infty P_{x|x_2}(t)\frac{dL\left(P_{x|x_2}(t)\right)}{dt} - 1.$$

$$= \frac{2\int_x^\infty \left\{\left[\frac{\partial F(t,u)}{\partial u}\right]_{u=x_2} - \left[\frac{\partial F(x,u)}{\partial u}\right]_{u=x_2}\right\} t\ f(t,x_2)dt}{\left[-\frac{\partial \overline{F}(x,u)}{\partial u}\right]_{u=x_2} \int_x^\infty y\ f(y,x_2)dy} - 1\,.$$

That is

$$G_{x|x_2}\int_x^\infty y\ f(y,x_2)dy = \frac{2\int_x^\infty \left[-\frac{\partial \overline{F}(t,u)}{\partial u}\right]_{u=x_2} t\ f(t,x_2)dt}{\left[-\frac{\partial \overline{F}(x,u)}{\partial u}\right]_{u=x_2}} + 2\int_x^\infty t\ f(t,x_2)dt - \int_x^\infty y\ f(y,x_2)dy.$$

Dividing through out by $\left[-\frac{\partial \overline{F}(x,u)}{\partial u}\right]_{u=x_2}$, we get

$$G_{x|x_2}\frac{\int_x^\infty y\ f(y,x_2)dy}{\left[-\frac{\partial \overline{F}(x,u)}{\partial u}\right]_{u=x_2}} = \frac{2}{\left\{\left[-\frac{\partial \overline{F}(x,u)}{\partial u}\right]_{u=x_2}\right\}^2}\int_x^\infty \left[-\frac{\partial \overline{F}(t,u)}{\partial u}\right]_{u=x_2} t\ f(t,x_2)dt +$$

$$\frac{2}{\left[-\frac{\partial \overline{F}(x,u)}{\partial u}\right]_{u=x_2}}\int_x^\infty t\ f(t,x_2)dt - \frac{1}{\left[-\frac{\partial \overline{F}(x,u)}{\partial u}\right]_{u=x_2}}\int_x^\infty y\ f(y,x_2)dy.$$

From (5.12) we have

$$G_{x|x_2}\left[x+m_{12}\left(x \mid x_2\right)\right]=\frac{2}{\left\{\left[-\frac{\partial\bar{F}(x,u)}{\partial u}\right]_{u=x_2}\right\}^2}\int_x^\infty\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2} t\, f(t,x_2)dt$$

$$+\left[x+m_{12}\left(x \mid x_2\right)\right].$$

Thus

$$\left[1-G_{x|x_2}\right]\left[x+m_{12}\left(x \mid x_2\right)\right]=\frac{-2}{\left\{\left[-\frac{\partial\bar{F}(x,u)}{\partial u}\right]_{u=x_2}\right\}^2}\int_x^\infty\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2} t\, f(t,x_2)dt. \tag{5.22}$$

Now consider the integral given by

$$\int_x^\infty t\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2} f(t,x_2)dt=\int_x^\infty t\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\frac{\partial}{\partial t}\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}$$

$$=t\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\int_x^\infty\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}-\int_x^\infty -t\left[\frac{\partial^2\bar{F}(t,u)}{\partial t\partial u}\right]_{u=x_2}\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}dt$$

$$-\int_x^\infty\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}dt$$

$$=-x\left\{\left[-\frac{\partial\bar{F}(x,u)}{\partial u}\right]_{u=x_2}\right\}^2+\int_x^\infty t\left[\frac{\partial^2\bar{F}(t,u)}{\partial t\partial u}\right]_{u=x_2}\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}dt$$

$$-\int_x^\infty\left\{\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\right\}^2dt$$

$$=-x\left\{\left[-\frac{\partial\bar{F}(x,u)}{\partial u}\right]_{u=x_2}\right\}^2+\int_x^\infty t\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\frac{\partial}{\partial t}\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}dt$$

$$-\int_x^\infty\left\{\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\right\}^2dt$$

i.e.

$$2\int_x^\infty t\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\frac{\partial}{\partial t}\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}=$$

$$-x\left\{\left[-\frac{\partial\bar{F}(x,u)}{\partial u}\right]_{u=x_2}\right\}^2-\int_x^\infty\left\{\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}\right\}^2 dt$$

where

$$f\left(t,x_2\right)=\frac{\partial}{\partial t}\left[-\frac{\partial\bar{F}(t,u)}{\partial u}\right]_{u=x_2}.$$

Substituting the above expression in (5.22), we get (5.19).

Proceeding in similar arguments with (5.11), (5.7) and (5.8), the third equality can be claimed.

5.3 Characterizations

In the backdrop of Theorem 5.1 and Theorem 5.2 it is of interest to look into characterizations based on the Lorenz function and Gini index. One general result in this direction is given as follows.

Theorem 5.3 Under the usual notations the relation

$$G_x=\frac{\left(1-k\right)m(x)}{x+m(x)},\ x>0 \tag{5.23}$$

$$G_{x|x_j}=\frac{\left(1-k_i\right)m_{ij}(x\mid x_j)}{x+m_{ij}(x\mid x_j)},\ x>x_j>0,\ i\neq j=1,2 \tag{5.24}$$

holds for a bivariate vector $\left(X_1,X_2\right)$ with

(i) $k,k_i=\frac{1}{2},\ i=1,2$ if and only if $\left(X_1,X_2\right)$ is distributed as Freund bivariate exponential distribution.

(ii) $k,k_i<\frac{1}{2},\ i=1,2$ if and only if $\left(X_1,X_2\right)$ is distributed as bivariate Pareto I distribution.

(iii) $k,k_i>\frac{1}{2},\ i=1,2$ if and only if $\left(X_1,X_2\right)$ is distributed as bivariate finite range distribution.

Proof

The if part of the theorem can be verified from Table 5.2. To prove the converse, consider (5.23), the equation becomes

$$G_x = \frac{m(x)}{x+m(x)} - \frac{k\ m(x)}{x+m(x)}$$

$$= 1 - \frac{x}{x+m(x)} - \frac{k\ m(x)}{x+m(x)}$$

$$1 - G_x = \frac{x+k\ m(x)}{x+m(x)}. \tag{5.25}$$

From equation (5.25), the equation (5.18) becomes

$$x + \frac{1}{\left[\bar{F}_Z(x)\right]^2} \int_x^\infty \left[\bar{F}_Z(t)\right]^2 dt = x + km(x).$$

or

$$\int_x^\infty \left[\bar{F}_Z(t)\right]^2 dt = k\ m(x)\left[\bar{F}_Z(x)\right]^2. \tag{5.26}$$

Differentiating (5.26) with respect to t, we get

$$-\left[\bar{F}_Z(x)\right]^2 = -2k\ m(x)\bar{F}_Z(x) f_Z(x) + k\left[\bar{F}_Z(x)\right]^2 \frac{d}{dx} m(x)$$

or

$$-1 = \frac{-2k\ m(x) f_Z(x)}{\bar{F}_Z(x)} + k\frac{d}{dx} m(x).$$

Now using (5.13), this equation becomes

$$-1 = -2k\ m(x)\lambda(x) + k\left[\lambda(x)m(x) - 1\right]$$

or

$$1 = k\ m(x)\lambda(x) + k$$

so that

$$m(x)\lambda(x) = \frac{1-k}{k}. \tag{5.27}$$

Once again using (5.13) we have

$$\frac{d}{dx} m(x) = \frac{1-2k}{k}$$

which on integration with respect to x gives

$$m(x) = \left(\frac{1-2k}{k}\right)x + c$$

or

$$\lambda(x) = \frac{1-k}{(1-2k)x + c}. \tag{5.28}$$

Now consider the second equation (5.24), which is equivalent to

$$1 - G_{x|x_j} = \frac{x + k_i\ m_{ij}(x \mid x_j)}{x + m_{ij}(x \mid x_j)}. \tag{5.29}$$

Now equation (5.29) implies

$$x + \frac{1}{\left\{\left[-\frac{\partial \bar{F}(x,u)}{\partial u}\right]_{u=x_2}\right\}^2} \int_x^\infty \left\{\left[-\frac{\partial \bar{F}(t,u)}{\partial u}\right]_{u=x_2}\right\}^2 dt = x + k_1 m_{12}\left(x \mid x_2\right).$$

or

$$\int_x^\infty \left\{\left[-\frac{\partial \bar{F}(t,u)}{\partial u}\right]_{u=x_2}\right\}^2 dt = k_1\ m_{12}\left(x \mid x_2\right)\left\{\left[-\frac{\partial \bar{F}(x,u)}{\partial u}\right]_{u=x_2}\right\}^2.$$

Proceeding in the same manner as for (5.26) we will arrive at

$$m_{12}(x \mid x_2)\lambda_{12}(x \mid x_2) = \frac{1-k_1}{k_1}.$$

Using (5.13) we have

$$\lambda_{12}(x \mid x_2) = \frac{1-k_1}{\left(1-2k_1\right)x + c_1},\ x_2 < x. \tag{5.30}$$

Similar arguments hold for

$$G_{x|x_1} = \frac{\left(1-k_2\right)m_{21}(x \mid x_1)}{x + m_{21}(x \mid x_1)},\ x_1 < x$$

to get

$$\lambda_{21}(x \mid x_1) = \frac{1-k_2}{\left(1-2k_2\right)x + c_2},\ x_2 < x. \tag{5.31}$$

Thus from Cox (1972) uniqueness property given in (1.23), we have

$$f(x_1,x_2)=\begin{cases}\exp\left[-\int_0^{x_1}\frac{(1-k)/(1-2k)}{u+c/(1-2k)}du-\int_{x_1}^{x_2}\frac{(1-k_2)/(1-2k_2)}{u+c_2/(1-2k_2)}du\right]\frac{p_1\,(1-k)/(1-2k)}{x_1+c/(1-2k)}\frac{(1-k_2)/(1-2k_2)}{x_2+c_2/(1-2k_2)} \\ \qquad ;x_1<x_2 \\ \exp\left[-\int_0^{x_2}\frac{(1-k)/(1-2k)}{u+c/(1-2k)}du-\int_{x_2}^{x_1}\frac{(1-k_1)/(1-2k_1)}{u+c_1/(1-2k_1)}du\right]\frac{p_2\,(1-k)/(1-2k)}{x_2+c/(1-2k)}\frac{(1-k_1)/(1-2k_1)}{x_1+c_1/(1-2k_1)} \\ \qquad ;x_2<x_1\end{cases}. \tag{5.32}$$

Which when $k, k_i = \frac{1}{2}, i = 1, 2$, becomes

$$f(x_1,x_2)=\begin{cases}\alpha_1\alpha_2'\exp[-\alpha_2'x_2-(\alpha_1+\alpha_2-\alpha_2')x_1];\, 0<x_1<x_2<\infty \\ \alpha_2\alpha_1'\exp[-\alpha_1'x_1-(\alpha_1+\alpha_2-\alpha_1')x_2];\, 0<x_2<x_1<\infty\end{cases}$$

which is the Freund's (1961) bivariate exponential distribution with $\alpha_1+\alpha_2=\frac{1}{2c}$, $\alpha_i=\frac{p_i}{2c}$ and $\alpha_i'=\frac{1}{2c_i}, i=1,2.$

When $k, k_i < \frac{1}{2}, i=1,2$ and $c, c_i = 0,$ it becomes

$$f(x_1,x_2)=\begin{cases}\frac{\alpha_1\alpha_2'}{\sigma^{-(\alpha_1+\alpha_2)}}\ x_1^{-(\alpha_1+\alpha_2-\alpha_1'+1)}\ x_2^{-(\alpha_2'+1)};\ \sigma<x_1<x_2<\infty \\ \frac{\alpha_2\alpha_1'}{\sigma^{-(\alpha_1+\alpha_2)}}\ x_2^{-(\alpha_1+\alpha_2-\alpha_2'+1)}\ x_1^{-(\alpha_1'+1)};\ \sigma<x_2<x_1<\infty\end{cases}$$

which is the bivariate Pareto I $\left(BPI(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')\right)$ distribution (Asha and Jagathnath (2008)) with parameters

$\alpha_1+\alpha_2=\frac{1-k}{1-2k},\ \alpha_i=\frac{p_i(1-k)}{1-2k}$ and $\alpha_i'=\frac{1-k_i}{1-2k_i}, i=1,2.$

When $k, k_i > \frac{1}{2},$ and $\frac{c}{1-2k}=\frac{c_i}{1-2k_i}, i=1,2,$ it becomes

$$f(x_1,x_2)=\begin{cases}\frac{\alpha_1\alpha_2'}{R^2}\left(1-\frac{x_1}{R}\right)^{(\alpha_1+\alpha_2-\alpha_1'-1)}\left(1-\frac{x_2}{R}\right)^{(\alpha_2'-1)};0<x_1<x_2<R \\ \frac{\alpha_2\alpha_1'}{R^2}\left(1-\frac{x_2}{R}\right)^{(\alpha_1+\alpha_2-\alpha_2'-1)}\left(1-\frac{x_1}{R}\right)^{(\alpha_1'-1)};0<x_2<x_1<R\end{cases}$$

which is the bivariate finite range distribution with parameters, $\alpha_1+\alpha_2=\dfrac{k-1}{2k-1}$, $\alpha_i=\dfrac{p_i(k-1)}{2k-1}$, $\alpha_i'=\dfrac{k_i-1}{2k_i-1}, i=1,2$ and $R=\dfrac{c}{1-2k}$.

Remark 5.2 Under conditions of the Theorem 5.3, equation (5.23) and (5.24) are equivalent to stating

$$1-G_x=\frac{x}{x+m(x)}+\frac{k}{\lambda(x)}$$

and

$$1-G_{x|x_j}=\frac{x}{x+m_{ij}(x\mid x_j)}+\frac{k_i}{\lambda_{ij}(x\mid x_j)}.$$

This result helps us to deduce many results which are analogous to popular results in the univariate case. The piecewise constancy of the bivariate Gini index can be deduced from Theorem 5.3, which in a way extends the truncation invariance property (Ord et. al (1983)) of the univariate Pareto I distribution.

Remark 5.3 Under the conditions in Theorem 3.1, the bivariate Gini index is of the form

$$\underline{G}(\underline{x})=\left(g,g_1,g_2\right) \tag{5.33}$$

where $0<g,g_i<1, i=1,2$ if and only if $\left(X_1,X_2\right)$ has a bivariate Pareto I distribution (refer Table 5.2).

Remark 5.4 The quantity $\dfrac{m(x)}{x+m(x)}$ is referred to as income gap ratio (Belzunce et al. (1998)). It measures the proportion of the people, whose income from both the sources are greater than the threshold value x. So $\dfrac{m_{ij}(x\mid x_j)}{x+m_{ij}(x\mid x_j)}$ can be used to measure the proportion of the people whose income from one source falls below the threshold value x, when $x > x_j$. Thus

$$\underline{\beta}(\underline{x}) = \left(\frac{m(x)}{x+m(x)}, \frac{m_{12}(x \mid x_2)}{x+m_{12}(x \mid x_2)}, \frac{m_{21}(x \mid x_1)}{x+m_{21}(x \mid x_1)} \right)$$

can be viewed as a bivariate income gap ratio for the rich.

Corollary 5.1 Under the conditions of Theorem 5.3, the following statements are equivalent.

(i) The bivariate Gini index is of the form (5.33) if and only if the bivariate income gap ratio is of the form

$$\left(\frac{m(x)}{x+m(x)}, \frac{m_{12}(x \mid x_2)}{x+m_{12}(x \mid x_2)}, \frac{m_{21}(x \mid x_1)}{x+m_{21}(x \mid x_1)} \right) = (\beta, \beta_1, \beta_2),$$

where β, β_1 and β_2 are some constant such that $0 < \beta, \beta_i < 1, i = 1, 2.$

(ii) (X_1, X_2) has a bivariate Pareto I distribution.

Table 5.2: Expression for Gini index and income gap ratio

Bivariate density function	$\underline{\lambda}(\underline{x})$	$\underline{m}(\underline{x})$	$\underline{G}(\underline{x})$	$\underline{\beta}(\underline{x})$
Bivariate exponential by Freund (1961) $f(x_1,x_2)=\begin{cases}\alpha_1\alpha_2' \exp[-\alpha_2' x_2-(\alpha_1+\alpha_2-\alpha_2')x_1];\ 0<x_1<x_2\\ \alpha_2\alpha_1' \exp[-\alpha_1' x_1-(\alpha_1+\alpha_2-\alpha_1')x_2];\ 0<x_2<x_1\end{cases}$ $\alpha_i>0,\ \alpha_i'>0,\ \alpha_1+\alpha_2\neq\alpha_i',\ i=1,2.$	$(\alpha_1+\alpha_2,\ \alpha_1',\alpha_2')$	$\left(\frac{1}{\alpha_1+\alpha_2},\ \frac{1}{\alpha_1'},\frac{1}{\alpha_2'}\right)$	$\left(\frac{1}{2((\alpha_1+\alpha_2)x+1)},\ \frac{1}{2(\alpha_1'x+1)},\frac{1}{2(\alpha_2'x+1)}\right)$	$\left(\frac{1}{((\alpha_1+\alpha_2)x+1)},\ \frac{1}{(\alpha_1'x+1)},\frac{1}{(\alpha_2'x+1)}\right)$
$BPI(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2')$ (Asha & Jagathnath Krishna, 2008) $f(x_1,x_2)=\begin{cases}\frac{\alpha_1\alpha_2'}{\sigma^{-(\alpha_1+\alpha_2)}}\ x_1^{-(\alpha_1+\alpha_2-\alpha_1'+1)}\ x_2^{-(\alpha_2'+1)};\ \sigma<x_1<x_2\\ \frac{\alpha_2\alpha_1'}{\sigma^{-(\alpha_1+\alpha_2)}}\ x_2^{-(\alpha_1+\alpha_2-\alpha_2'+1)}\ x_1^{-(\alpha_1'+1)};\ \sigma<x_2<x_1\end{cases}$ $\alpha_i>1,\ \alpha_i'>1,\ \alpha_1+\alpha_2\neq\alpha_i',\ \sigma>0,\ i=1,2.$	$\left(\frac{\alpha_1+\alpha_2}{x},\ \frac{\alpha_1'}{x},\frac{\alpha_2'}{x}\right)$	$\left(\frac{x}{\alpha_1+\alpha_2-1},\ \frac{x}{\alpha_1'-1},\ \frac{x}{\alpha_2'-1}\right)$	$\left(\frac{1}{(2(\alpha_1+\alpha_2)-1)},\ \frac{1}{(2\alpha_1'-1)},\frac{1}{(2\alpha_2'-1)}\right)$	$\left(\frac{1}{\alpha_1+\alpha_2},\ \frac{1}{\alpha_1'},\frac{1}{\alpha_2'}\right)$
Bivariate finite range distribution $f(x_1,x_2)=\begin{cases}\frac{\alpha_1\alpha_2'}{R^2}\left(1-\frac{x_1}{R}\right)^{(\alpha_1+\alpha_2-\alpha_1'-1)}\left(1-\frac{x_2}{R}\right)^{(\alpha_2'-1)};\ 0<x_1<x_2<R\\ \frac{\alpha_2\alpha_1'}{R^2}\left(1-\frac{x_2}{R}\right)^{-(\alpha_1+\alpha_2-\alpha_2'-1)}\left(1-\frac{x_1}{R}\right)^{(\alpha_1'-1)};\ 0<x_2<x_1<R\end{cases}$ $\alpha_i>0,\ \alpha_i'>0,\ \alpha_1+\alpha_2\neq\alpha_i',\ R>0,\ i=1,2.$	$\left(\frac{\alpha_1+\alpha_2}{R-x},\ \frac{\alpha_1'}{R-x},\frac{\alpha_2'}{R-x}\right)$	$\left(\frac{R-x}{\alpha_1+\alpha_2+1},\ \frac{R-x}{\alpha_1'+1},\frac{R-x}{\alpha_2'+1}\right)$	$\left(\frac{(\alpha_1+\alpha_2)(R-x)}{(2\alpha+1)(R+\alpha x)},\ \frac{\alpha_1'(R-x)}{(2\alpha_1'+1)(R+\alpha_1'x)},\ \frac{\alpha_2'(R-x)}{(2\alpha_2'+1)(R+\alpha_2'x)}\right)$	$\left(\frac{(R-x)}{(R+(\alpha_1+\alpha_2)x)},\ \frac{(R-x)}{(R+\alpha_1'x)},\ \frac{(R-x)}{(R+\alpha_2'x)}\right)$

Chapter 6

A General Representation

6.1 Introduction

From the study so far we saw that properties enjoyed by the Freund bivariate exponential distribution can be translated to other distributions transformed from the Freund bivariate exponential distribution. In fact all its properties can be translated into a property of an arbitrary bivariate continuous distribution. With the aim of doing so in this chapter, we explore the representation of the bivariate Pareto distributions in terms of uniform random variables. This idea is heavily borrowed from the idea of copulas. A copula is a function $C(u,v)$ from I^2 to I where $I^2 = \{(x_1, x_2) \mid 0 < x_i < 1, i = 1, 2\}$ with the following properties.

1. $C(u,v)$ is a grounded function. That means for all u, v in I,

   $C(u,0) = 0 = C(0,v)$.

2. $C(u,1) = u$ and $C(1,v) = v$ (6.1)

3. For every u_1, u_2, v_1, v_2 in I such that $u_1 \le u_2$ and $v_1 \le v_2$,

   $$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \ge 0.$$

Sklar (1959) proposed a theorem which is central to the theory of copulas and is the foundation of many, if not most of the applications of that theory to statistics. This theorem elucidates the role that copulas play in the relationship between multivariate distribution functions and their univariate marginals. The statement is as follows.

Let $H(x_1,x_2)$ be a joint distribution function with marginals $F_1(x_1)$ and $F_2(x_2)$, then there exists a copula C such that for any $x_1,x_2 \in \overline{R}$, $\overline{R}$ is the extended real line $[-\infty,\infty]$,

$$H(x_1,x_2)=C\left(F_1(x_1),F_2(x_2)\right). \tag{6.2}$$

If $F_1(x_1)$ and $F_2(x_2)$ are continuous then C is unique otherwise C is uniquely determined on $Range\,F \times Range\,G$. Conversely if C is the copula and F_1 and F_2 are marginals then the function H defined by (6.2) is the joint distribution function. If the function $C(u,v)$ satisfies only the properties 1 and 3 in (6.1) then it is called a pseudo copula. This was introduced by Fermanian and Wegkamp (2004) to study the dynamic dependence structure. It was also shown by the same authors that Sklar's theorem can be extended to pseudo copulas too.

A natural question that arises is that is there a relationship between univariate and joint survival functions analogous to the one between univariate and joint distribution functions as embodied in Sklar's theorem. This paved way to the concept of survival copulas.

If $\overline{H}(x_1,x_2)=P\left[X_1 \geq x_1, X_2 \geq x_2\right]$, the marginals of $\overline{H}(x_1,x_2)$ are $\overline{H}(x_1,-\infty)$ and $\overline{H}(-\infty,x_2)$ which are univariate survival functions $\overline{F}_1(x_1)$ and $\overline{F}_2(x_2)$ respectively, then

$$\begin{aligned}\overline{H}(x_1,x_2) &= 1-F_1(x_1)-F_2(x_2)+H(x_1,x_2)\\ &= \overline{F}_1(x_1)+\overline{F}_2(x_2)-1+C\left(F_1(x_1),F_2(x_2)\right)\\ &= \overline{F}_1(x_1)+\overline{F}_2(x_2)-1+C\left(1-\overline{F}_1(x_1),1-\overline{F}_2(x_2)\right).\end{aligned}$$

Denoting $\hat{C}$ as a function from I^2 to I by

$$\hat{C}(u,v)=u+v-1+C(1-u,1-v),$$

we have

$$\bar{H}\left(x_1, x_2\right) = \hat{C}\left(\bar{F}_1\left(x_1\right), \bar{F}_2\left(x_2\right)\right)$$

where $\hat{C}(u,v)$ is referred to as the survival copula of X_1 and X_2. Secondly $\hat{C}(u,v)$ couples the joint survival functions to its univariate marginals in a manner completely analogous to the way in which a copula connects the joint distribution to its marginals. If $\bar{C}(u,v)$ is the joint survival function for two uniform $(0,1)$ random variables whose joint distribution function is copula $C(u,v)$, then

$$\hat{C}(1-u,1-v) = \bar{C}\left(u,v\right) = 1-u-v+C\left(u,v\right).$$

It is easy to prove that a survival copula $\hat{C}(u,v)$ is a copula since

(i) $\hat{C}(u,v)$ is a grounded function, that is

$$\begin{aligned} \hat{C}(u,0) = \hat{C}(0,v) &= u-1+C(1,1-v) \\ &= u-1+1-u \\ &= 0 \end{aligned}$$

(ii) For every u,v in I, (6.3)

$$\begin{aligned} \hat{C}(u,1) &= u+C(1-u,0) \\ &= u \end{aligned}$$

and

$$\hat{C}(1,v) = v.$$

(iii) For every u_1, u_2, v_1, v_2 in I such that $u_1 \le u_2$ and $v_1 \le v_2$,

$$\hat{C}(u_2,v_2) - \hat{C}(u_2,v_1) - \hat{C}(u_1,v_2) + \hat{C}(u_1,v_1) \ge 0.$$

Barnett (1980) gave the survival copula called Gumbel-Barnett copula which generalizes the dependence in the Gumbel's bivariate exponential distribution specified by $\bar{F}(x_1,x_2) = e^{-\left(x_1+x_2+\theta x_1 x_2\right)}$; $x_1, x_2 > 0$. The corresponding survival copula is given by

$$\hat{C}(u,v) = uve^{-\theta \ln(u)\ln(v)}.$$

The second Gumbel's exponential distribution specified by

$$\bar{F}(x_1, x_2) = \left(1-e^{-x_1}\right)\left(1-e^{-x_2}\right)\left(1+\theta e^{-(x_1+x_2)}\right);\ x_1, x_2 > 0$$

corresponds to the Farlie-Gumbel-Morgenstern copula is given by

$$C(u,v) = uv + \theta uv(1-u)(1-v);\ 0 < u < 1,\ 0 < v < 1.$$

The Marshall-Olkin (1967) bivariate exponential distribution specified by the survival function

$$\bar{F}(x_1, x_2) = e^{-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_{12} max(x_1, x_2)};\ x_1, x_2 > 0,\ \lambda_1, \lambda_2, \lambda_{12} > 0$$

has the survival copula given by

$$\hat{C}(u,v) = \min\left(u^{1-\alpha}v, uv^{1-\beta}\right). \tag{6.4}$$

This family is known both as the Marshall-Olkin family and the generalized Caudras-Auge family. Other copulas related to exponential distribution are detailed in Genest and Mackay (1986), Joe and Hu (1996), Joe (1997), Nelsen (1999) and Joe and Ma (2000). This now enables a study of the Marshall-Olkin type Pareto distributions (Veenus and Nair (1994), Hanagal (1996), Yeh (2004 a,b)) and other transformed distribution by studying these copulas.

But in the case of the Freund bivariate Pareto distribution discussed in this thesis, there does not exist an analytical expression of the copula though they could be evaluated numerically. This turns out to be a great drawback as an analytical expression for the copula helps in providing a very convenient model for studying the properties with tools that are scale free.

The major advantage of studying a general representation in terms of uniform variants is that certain properties are same for all distributions in a particular equivalence class. If $H(u,v)$ denotes the uniform representation of a distribution $F(x_1, x_2)$ and $G(x_1, x_2)$ belongs to the same equivalence class if $H_F(u,v) = H_G(u,v)$.

This motivates us to give a representation in terms of uniform variates which would enable us to study properties of the Freund bivariate exponential distribution and its transformation like the $BP\ I\left(\sigma, \alpha_1, \alpha_2, \alpha_1', \alpha_2'\right)$ distributions and $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution discussed in this thesis under a unified frame work. Accordingly in the next section we give a representation and call it the uniform representation. We also give examples of the distributions that are derived using this representation. These distributions are the bivariate distributions obtained by transforming the Freund bivariate exponential distribution by suitable transformations. They include bivariate Weibull distribution (Lu (1989)), $BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ and $BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distributions (Asha and Jagathnath (2008)). In the third section, the reliability property particularly, the total failure rate of the general representation is given. This expression enables us to directly compute the failure rate of the distributions having this representation once we know the uniform translate. This is illustrated in Table 6.1. A general property analogous to the dullness property is defined for the uniform representation and a characterization of this class is discussed. With this chapter we conclude the thesis by briefly stating the direction of future course of study.

6.2 The General Representation

As in the previous section we adopt the representation U and V for uniform variates. Obviously (U,V) ranges over the unit square. Consider the uniform representation

$$h(u,v)=\begin{cases}\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}u^{\alpha_1+\alpha_2-\alpha_2'}v^{\alpha_2'}+\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}v^{\alpha_1+\alpha_2}; 0\le u\le v\le 1\\[2ex] \dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}v^{\alpha_1+\alpha_2-\alpha_1'}u^{\alpha_1'}+\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}u^{\alpha_1+\alpha_2}; 0\le v\le u\le 1\end{cases}. \quad (6.5)$$

$$\alpha_i,\alpha_i'>0,\ \alpha_1+\alpha_2\ne\alpha_i',\ i=1,2.$$

It should be observed that (6.5) is a finite mixture of two pseudo-survival copulas.

$$h_1(u,v)=\begin{cases} u^{\alpha_1+\alpha_2-\alpha_2'}v^{\alpha_2'}; u\le v \\ v^{\alpha_1+\alpha_2-\alpha_1'}u^{\alpha_1'}; v\le u \end{cases} \tag{6.6}$$

and

$$h_2(u,v)=\max\left(u,v\right)^{\alpha_1+\alpha_2}. \tag{6.7}$$

Hence it follows that $h(u,v)$ is also a pseudo-survival copula as (i) and (iii) of (6.3) are satisfied. The survival function associated with the pseudo-survival copula (6.5) is given by

$$\overline{h}(u,v)=h(1-u,1-v)$$

i.e.,

$$\overline{h}(u,v)=\begin{cases} \dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}(1-u)^{\alpha_1+\alpha_2-\alpha_2'}(1-v)^{\alpha_2'}+\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)(1-v)^{\alpha_1+\alpha_2}; \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 0\le u\le v\le 1 \\ \dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}(1-v)^{\alpha_1+\alpha_2-\alpha_1'}(1-u)^{\alpha_1'}+\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)(1-u)^{\alpha_1+\alpha_2}; \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 0\le v\le u\le 1 \end{cases}. \tag{6.8}$$

Since $h(u,v)$ is a pseudo-survival copula, evidently not a bivariate uniform distribution. Some members belonging to this class are worked out below.

1. Freund bivariate exponential distribution (1961).

For, $1-u=e^{-x_1}$ and $1-v=e^{-x_2}$,

$$\overline{F}(x_1,x_2)=\begin{cases} \dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2-\alpha_2')x_1-\alpha_2'x_2}+\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_2}; \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 0\le x_1\le x_2 \\ \dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-(\alpha_1+\alpha_2-\alpha_1')x_2-\alpha_1'x_1}+\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_1}; \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad 0\le x_2\le x_1 \end{cases}.$$

2. Bivariate Weibull distribution (Lu (1989)).

For, $1-u=e^{-(x_1)^{c1}}$ and $1-v=e^{-(x_2)^{c2}}$,

$$\overline{F}(x_1,x_2)=\begin{cases}\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2-\alpha_2')x_1^{c1}-\alpha_2'x_2^{c2}}+\\ \qquad\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_2^{c2}};0\le x_1^{c_1}\le x_2^{c_2}\\ \dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-(\alpha_1+\alpha_2-\alpha_1')x_2^{c2}-\alpha_1'x_1^{c1}}+\\ \qquad\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_1^{c1}};0\le x_2^{c_2}\le x_1^{c_1}\end{cases}.$$

3. Bivariate Pareto I ($BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$) distribution (Table 4.3).

For, $1-u=\left(\dfrac{x_1}{\sigma}\right)^{-c}$ and $1-v=\left(\dfrac{x_2}{\sigma}\right)^{-c}$,

$$\overline{F}(x_1,x_2)=\begin{cases}\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}\left(\dfrac{x_1}{\sigma}\right)^{-c(\alpha_1+\alpha_2-\alpha_2')}\left(\dfrac{x_2}{\sigma}\right)^{-c\alpha_2'}+\\ \qquad\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)\left(\dfrac{x_2}{\sigma}\right)^{-c(\alpha_1+\alpha_2)};\sigma\le x_1\le x_2\\ \dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}\left(\dfrac{x_2}{\sigma}\right)^{-c(\alpha_1+\alpha_2-\alpha_1')}\left(\dfrac{x_1}{\sigma}\right)^{-c\alpha_1'}+\\ \qquad\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)\left(\dfrac{x_1}{\sigma}\right)^{-c(\alpha_1+\alpha_2)};\sigma\le x_2\le x_1\end{cases}.$$

4. Bivariate Pareto II $\left(BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)\right)$ distribution (Table 4.3).

For, $1-u=\left(1+\frac{x_1-\mu}{\sigma}\right)^{-c}$ and $1-v=\left(1+\frac{x_2-\mu}{\sigma}\right)^{-c}$,

$$\overline{F}(x_1,x_2)=\begin{cases} \frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-c(\alpha_1+\alpha_2-\alpha_2')}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-c\alpha_2'}+ \\ \qquad \left(\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)\left(1+\frac{x_2-\mu}{\sigma}\right)^{-c(\alpha_1+\alpha_2)} ;\ \mu\le x_1\le x_2 \\ \frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}\left(1+\frac{x_2-\mu}{\sigma}\right)^{-c(\alpha_1+\alpha_2-\alpha_1')}\left(1+\frac{x_1-\mu}{\sigma}\right)^{-c\alpha_1'}+ \\ \qquad \left(\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)\left(1+\frac{x_1-\mu}{\sigma}\right)^{-c(\alpha_1+\alpha_2)} ;\ \mu\le x_2\le x_1 \end{cases}.$$

5. Bivariate finite range distribution (Table 4.3)

For, $1-u=\left(1-\frac{x_1}{R}\right)^{c}$ and $1-v=\left(1-\frac{x_2}{R}\right)^{c}$,

$$\overline{F}(x_1,x_2)=\begin{cases} \frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}\left(1-\frac{x_1}{R}\right)^{c(\alpha_1+\alpha_2-\alpha_2')}\left(1-\frac{x_2}{R}\right)^{c\alpha_2'}+\left(\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right) \\ \qquad \left(1-\frac{x_2}{R}\right)^{c(\alpha_1+\alpha_2)} ;\ 0\le x_2\le x_1\le R \\ \frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}\left(1-\frac{x_2}{R}\right)^{c(\alpha_1+\alpha_2-\alpha_1')}\left(1-\frac{x_1}{R}\right)^{c\alpha_1'}+\left(\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right) \\ \qquad \left(1-\frac{x_1}{R}\right)^{c(\alpha_1+\alpha_2)} ;\ 0\le x_1\le x_2\le R \end{cases}.$$

6. Bivariate log-logistic distribution.

For, $1-u=\left(1+(\lambda x_1)^{a_1}\right)^{-1}$ and $1-v=\left(1+(\lambda x_2)^{a_2}\right)^{-1}$,

$$\overline{F}(x_1,x_2)=\begin{cases}\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}\left(1+(\lambda x_1)^{a_1}\right)^{-(\alpha_1+\alpha_2-\alpha_2')}\left(1+(\lambda x_2)^{a_2}\right)^{-\alpha_2'}+ \\ \qquad\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)\left(1+(\lambda x_2)^{a_2}\right)^{-(\alpha_1+\alpha_2)} \ ; \ 0\le(\lambda x_1)^{a_1}\le(\lambda x_2)^{a_2} \\ \dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}\left(1+(\lambda x_2)^{a_2}\right)^{-(\alpha_1+\alpha_2-\alpha_1')}\left(1+(\lambda x_1)^{a_1}\right)^{-\alpha_1'}+ \\ \qquad\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)\left(1+(\lambda x_1)^{a_1}\right)^{-(\alpha_1+\alpha_2)} \ ; \ 0\le(\lambda x_2)^{a_2}\le(\lambda x_1)^{a_1}\end{cases}.$$

6.3 Failure Rate of the General Class

In this section we consider how total failure rate can be related to the properties of the general representation given in (6.5) in terms of general representation of the total failure rate (Cox (1972)) is defined as

$$\lambda(x)=-\frac{d}{du_Z}\left[\log h\left(1-u_Z,1-u_Z\right)\right]\frac{du_Z}{dx}\text{, where } Z=min(X_1,X_2)$$

$$\lambda_{12}(x_1 \mid x_2)=\frac{\partial}{\partial u}\left[\log\frac{\partial}{\partial u}h\left(1-u,1-v\right)\right]\frac{\partial u}{\partial x_1};\ x_1>x_2$$

$$\lambda_{21}(x_2 \mid x_1)=\frac{\partial}{\partial u}\left[\log\frac{\partial}{\partial v}h\left(1-u,1-v\right)\right]\frac{\partial v}{\partial x_2};\ x_1<x_2 \qquad (6.9)$$

So that for (6.5), $\underline{\lambda}(\underline{x})$ is now

$$\underline{\lambda}(\underline{x})=\left(\frac{-\left(\alpha_1+\alpha_2\right)\dfrac{d(1-u_Z)}{dx}}{1-u_Z},\frac{-\alpha_1'\dfrac{\partial(1-u)}{\partial x_1}}{1-u},\frac{-\alpha_2'\dfrac{\partial(1-v)}{\partial x_2}}{1-v}\right).$$

Table 6.1 lists the failure rate for distributions defined by representation (6.5).

Table 6.1 Total Failure Rate for the Uniform Representation

Distributions	$1-u$	$1-v$	$\underline{\lambda}(\underline{x})$
Freund's bivariate exponential distribution (1961)	e^{-x_1}	e^{-x_2}	$\left((\alpha_1+\alpha_2),\ \alpha_1',\alpha_2'\right)$
Bivariate Weibull distribution (Lu (1989))	$e^{-\left(\frac{x_1}{\sigma}\right)^c}$	$e^{-\left(\frac{x_2}{\sigma}\right)^c}$	$\left(\frac{c(\alpha_1+\alpha_2)}{\sigma}\left(\frac{x}{\sigma}\right)^{c-1}, \frac{c\alpha_1'}{\sigma}\left(\frac{x_1}{\sigma}\right)^{c-1}, \frac{c\alpha_2'}{\sigma}\left(\frac{x_2}{\sigma}\right)^{c-1}\right)$
$BP\ I\left(\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution	$\left(\frac{x_1}{\sigma}\right)^{-c}$	$\left(\frac{x_2}{\sigma}\right)^{-c}$	$\left(\frac{c(\alpha_1+\alpha_2)}{x}, \frac{c\alpha_1'}{x_1}, \frac{c\alpha_2'}{x_2}\right)$
$BP\ II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution	$\left(1+\frac{x_1-\mu}{\sigma}\right)^{-c}$	$\left(1+\frac{x_2-\mu}{\sigma}\right)^{-c}$	$\left(\frac{c(\alpha_1+\alpha_2)}{\sigma+x-\mu}, \frac{c\alpha_1'}{\sigma+x-\mu}, \frac{c\alpha_2'}{\sigma+x-\mu}\right)$
Bivariate finite range distribution (Table 4.3)	$\left(1-\frac{x_1}{R}\right)^c$	$\left(1-\frac{x_2}{R}\right)^c$	$\left(\frac{c(\alpha_1+\alpha_2)}{(R-x)}, \frac{c\alpha_1'}{(R-x_1)}, \frac{c\alpha_2'}{(R-x_2)}\right)$
Bivariate log-logistic distribution	$\frac{1}{\left(1+(\lambda x_1)^a\right)}$	$\frac{1}{\left(1+(\lambda x_2)^a\right)}$	$\left(\frac{(\alpha_1+\alpha_2)a\lambda^a x^{a-1}}{1+(\lambda x)^a}, \frac{\alpha_1' a\lambda^a x_1^{a-1}}{1+(\lambda x_1)^a}, \frac{\alpha_2' a\lambda^a x_2^{a-1}}{1+(\lambda x_2)^a}\right)$

Another interesting property that has gained vast attention of the researchers is the no-ageing property and its variants like the dullness property. As with other reliability concepts there are various extensions to the bivariate case. We consider the extension in (3.3). A uniform representation is said to have a bivariate dullness property if it verifies

$$h\left((1-t)(1-u),(1-t)(1-v)\right)=h\left((1-t),(1-t)\right)h\left((1-u),(1-v)\right) \tag{6.10}$$

for all $0\le t,u,v\le 1$.

The representations that belong to the class (6.10) are characterized in the following theorem.

Theorem 6.1 The uniform representation $h(u,v)$ where U and V are uniform variates satisfies (6.10), if and only if $h(u,v)$ can be written as

$$h(1-u,1-v)=\begin{cases}(1-u)^c h\left(1,\dfrac{1-v}{1-u}\right); u\le v\\[2ex] (1-v)^c h\left(\dfrac{1-u}{1-v},1\right); v\le u\end{cases}, c>0\,. \tag{6.11}$$

Proof

Let (6.10) be satisfied. Then for $1-v=1-u$,

$$h\left((1-t)(1-u),(1-t)(1-u)\right)=h\left((1-t),(1-t)\right)h\left((1-u),(1-u)\right).$$

From Aczel (1966 p.41) and the fact that $h\left((1-u),(1-v)\right)$ is a survival function it follows thats

$$h\left((1-t),(1-t)\right)=(1-t)^c; c>0$$

so that

$$h\left((1-t)(1-u),(1-t)(1-v)\right)=(1-t)^c h\left((1-u),(1-v)\right). \tag{6.12}$$

Now let $1-v\le 1-u,$ then once again from (6.10) it follows that

$$\begin{aligned}h\left((1-u),(1-v)\right)&=h\left(1-u,1-u\right)h\left(1,\frac{1-v}{1-u}\right)\\&=(1-u)^c h\left(1,\frac{1-v}{1-u}\right), 0\le u\le v\le 1\end{aligned}$$

Similarly,

$$h\left((1-u),(1-v)\right)=(1-v)^c h\left(\frac{1-u}{1-v},1\right), 0\le v\le u\le 1\,.$$

To prove the converse, if $h\left((1-u),(1-v)\right)$ is as given in (6.12) then

$h\big((1-t)(1-u),(1-t)(1-v)\big)=(1-t)^{c}\,h\big((1-u),(1-v)\big)$ for all $0\le t,u,v\le 1$.

That is

$$h\big((1-t)(1-u),(1-t)(1-v)\big)=h\big((1-t),(1-t)\big)\,h\big((1-u),(1-v)\big)$$

for all $0\le t,u,v\le 1$.

Hence the theorem.

Corollary 6.1 The uniform representation given in (6.5) belongs to the class of distributions verifying (6.12).

Proof

Observe that the uniform representation (6.5) can be written as

$$h(1-u,1-v)=\begin{cases}(1-u)^{\alpha_1+\alpha_2}\left(\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}\left(\dfrac{1-v}{1-u}\right)^{\alpha_2'}+\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\left(\dfrac{1-v}{1-u}\right)^{\alpha_1+\alpha_2}\right); & \\ \hfill 0\le u\le v\le 1 & \\ (1-v)^{\alpha_1+\alpha_2}\left(\dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}\left(\dfrac{1-u}{1-v}\right)^{\alpha_1'}+\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\left(\dfrac{1-u}{1-v}\right)^{\alpha_1+\alpha_2}\right); & \\ \hfill 0\le v\le u\le 1. & \end{cases}$$

The proof now follows from Theorem 6.1.

Corollary 6.2 When $1-u=e^{-x_1}$ and $1-v=e^{-x_2}$, it follows from Corollary 6.1 that for $0\le x_1\le x_2$,

$$\begin{aligned}&\frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2-\alpha_2')x_1-\alpha_2'x_2}+\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2)x_2}\\ &\qquad=e^{-(\alpha_1+\alpha_2)x_1}\left(\frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-\alpha_2'(x_2-x_1)}+\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2)(x_2-x_1)}\right)\\ &\qquad=e^{-(\alpha_1+\alpha_2)x_1}\overline{F}_1(x_2-x_1),\end{aligned}$$

where

$$\bar{F}_1\left(x_2-x_1\right)=\left(\frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-\alpha_2'(x_2-x_1)}+\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2)(x_2-x_1)}\right).$$

Similarly for $0\le x_2\le x_1$,

$$\frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-\left(\alpha_1+\alpha_2-\alpha_1'\right)x_2-\alpha_1'x_1}+\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}e^{-(\alpha_1+\alpha_2)x_1}$$

$$=e^{-(\alpha_1+\alpha_2)x_2}\left(\frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-\alpha_1'(x_1-x_2)}+\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}e^{-(\alpha_1+\alpha_2)(x_1-x_2)}\right)$$

$$=e^{-(\alpha_1+\alpha_2)x_2}\bar{F}_2\left(x_1-x_2\right),$$

where

$$\bar{F}_2\left(x_1-x_2\right)=\left(\frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-\alpha_1'(x_1-x_2)}+\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}e^{-(\alpha_1+\alpha_2)(x_1-x_2)}\right).$$

In general by writing

$$\bar{F}(x_1,x_2)=\begin{cases} e^{-(\alpha_1+\alpha_2)x_1}\bar{F}_1\left(x_2-x_1\right);\ 0\le x_1\le x_2 \\ e^{-(\alpha_1+\alpha_2)x_2}\bar{F}_2\left(x_1-x_2\right);0\le x_2\le x_1 \end{cases}.$$

This is the class of exponential minima given in equation (3.5) characterized by the bivariate lack of memory property,

$$\bar{F}\left(x_1+t,x_2+t\right)=\bar{F}\left(x_1,x_2\right)\bar{F}\left(t,t\right)\ \text{for}\ x_1,x_2,t>0.$$

Corollary 6.3 When $1-u=\left(x_1\right)^{-c}$ and $1-v=\left(x_2\right)^{-c}$, it follows from Corollary 6.1,

$$\frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}x_1^{-c(\alpha_1+\alpha_2-\alpha_2')}x_2^{-c\alpha_2'}+\left(\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)x_2^{-c(\alpha_1+\alpha_2)}$$

$$=x_1^{-c(\alpha_1+\alpha_2)}\left(\frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}\left(\frac{x_2^{-c}}{x_1^{-c}}\right)^{\alpha_2'}+\left(\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)\left(\frac{x_2^{-c}}{x_1^{-c}}\right)^{(\alpha_1+\alpha_2)}\right),$$

$$=x_1^{-c(\alpha_1+\alpha_2)}\bar{F}_2\left(\frac{x_2}{x_1}\right);\ 1\le x_1\le x_2,$$

where

$$\bar{F}_2\left(\frac{x_2}{x_1}\right)=\frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}\left(\frac{x_2^{-c}}{x_1^{-c}}\right)^{\alpha_2'}+\left(\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)\left(\frac{x_2^{-c}}{x_1^{-c}}\right)^{(\alpha_1+\alpha_2)}.$$

Similarly for $1\le x_2\le x_1$

$$\frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}x_2^{-c(\alpha_1+\alpha_2-\alpha_1')}x_1^{-c\alpha_1'}+\left(\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)x_1^{-c(\alpha_1+\alpha_2)}$$

$$=x_2^{-c(\alpha_1+\alpha_2)}\left\{\frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}\left(\frac{x_1^{-c}}{x_2^{-c}}\right)^{\alpha_1'}+\left(\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)\left(\frac{x_1^{-c}}{x_2^{-c}}\right)^{(\alpha_1+\alpha_2)}\right\}$$

$$=x_2^{-c(\alpha_1+\alpha_2)}\bar{F}_1\left(\frac{x_1}{x_2}\right);1\le x_2\le x_1,$$

where

$$\bar{F}_1\left(\frac{x_1}{x_2}\right)=\frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}\left(\frac{x_1^{-c}}{x_2^{-c}}\right)^{\alpha_1'}+\left(\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_1'}\right)\left(\frac{x_1^{-c}}{x_2^{-c}}\right)^{(\alpha_1+\alpha_2)}.$$

In general

$$\bar{F}(x_1,x_2)=\begin{cases}x_2^{-c(\alpha_1+\alpha_2)}\bar{F}_1\left(\frac{x_1}{x_2}\right);1\le x_2\le x_1\\ x_1^{-c(\alpha_1+\alpha_2)}\bar{F}_2\left(\frac{x_2}{x_1}\right);1\le x_1\le x_2\end{cases}.$$

This is class of Pareto minima given in (3.6) characterized by the bivariate dullness property,

$$\bar{F}(x_1t,x_2t)=\bar{F}(x_1,x_2)\bar{F}(t,t)\text{ for } x_1,x_2,t>0.$$

Corollary 6.4 When $1-u=e^{-(x_1)^c}$ and $1-v=e^{-(x_2)^c}$, it follows from Corollary 6.1,

$$\frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2-\alpha_2')x_1^c-\alpha_2'x_2^c}+\left(\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_2^c}$$

$$= e^{-(\alpha_1+\alpha_2)x_1^c}\left\{\frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-\alpha_2'(x_2^c-x_1^c)}+\left(\frac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)(x_2^c-x_1^c)}\right\};$$

$$0\le x_1\le x_2 .$$

Similarly for $0\le x_2\le x_1$,

$$\frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-(\alpha_1+\alpha_2-\alpha_1')x_2^c-\alpha_1'x_1^c}+\left(\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_1^c}$$

$$= e^{-(\alpha_1+\alpha_2)x_2^c}\left\{\frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-\alpha_1'(x_1^c-x_2^c)}+\left(\frac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)(x_1^c-x_2^c)}\right\}$$

Thus we have,

$$\overline{F}(x_1,x_2)=\begin{cases}\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2-\alpha_2')x_1^c-\alpha_2'x_2^c}+\\ \qquad\qquad\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_2^c};\ 0\le x_1\le x_2\\ \dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-(\alpha_1+\alpha_2-\alpha_1')x_2^c-\alpha_1'x_1^c}+\\ \qquad\qquad\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_1^c};\ 0\le x_2\le x_1\end{cases}.$$

i.e.,

$$\overline{F}\left(\left(x_1^c+t^c\right)^{1/c},\left(x_2^c+t^c\right)^{1/c}\right)=\begin{cases}e^{-(\alpha_1+\alpha_2)t^c}\left\{\dfrac{\alpha_1}{\alpha_1+\alpha_2-\alpha_2'}e^{-(\alpha_1+\alpha_2-\alpha_2')x_1^c-\alpha_2'x_2^c}\right.\\ \qquad\left.+\left(\dfrac{\alpha_2-\alpha_2'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_2^c}\right\};0\le x_1\le x_2\\ e^{-(\alpha_1+\alpha_2)t^c}\left\{\dfrac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1'}e^{-(\alpha_1+\alpha_2-\alpha_1')x_2^c-\alpha_1'x_1^c}\right.\\ \qquad\left.+\left(\dfrac{\alpha_1-\alpha_1'}{\alpha_1+\alpha_2-\alpha_2'}\right)e^{-(\alpha_1+\alpha_2)x_1^c}\right\};0\le x_2\le x_1\end{cases}$$

Hence, $\overline{F}\left(\left(x_1+t\right)^{1/c},\left(x_2+t\right)^{1/c}\right)=\overline{F}(t,t)\overline{F}\left(x_1,x_2\right)$, which is the extension of the characterization of a univariate Weibull distribution.

Corollary 6.5 The Caudras-Auge copula given in (6.4) verifies the property (6.10).

Proof

From (6.4) we have

$$\min\left(u^{1-\alpha}v, uv^{1-\beta}\right) = \begin{cases} u^{1-\alpha}v;\ v^{\beta} \le u^{\alpha} \\ uv^{1-\beta};\ u^{\alpha} \le v^{\beta} \end{cases}.$$

i.e.,

$$\min\left(u^{1-\alpha}v, uv^{1-\beta}\right) = \begin{cases} u^{2-\alpha}\left(\dfrac{v}{u}\right);\ v^{\beta} \le u^{\alpha} \\ v^{2-\beta}\left(\dfrac{u}{v}\right);\ u^{\alpha} \le v^{\beta} \end{cases}.$$

Hence the result.

6.4 Conclusion

In this chapter we have given a uniform representation of the Freund (1961) bivariate exponential distribution and its transformation. We showed how this representation can be used to infer on the total failure rate of each distribution having this representation, once we know their uniform translates. We also characterize this uniform representation by what we define as general dullness property (6.10). It is further shown that this property implies the bivariate lack of memory property (Marshall and Olkin (1967)) for the Freund's bivariate exponential distribution. It implies the bivariate dullness property (Veenus and Nair (1994), Hanagal (1996), Yeh (2004 a,b)) for the $BP\,I\left(1,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ distribution. It also implies a characterization for the bivariate Weibull distribution. With this we conclude the present thesis, after discussing below the future course of work.

6.5 Future Work

In this thesis the bivariate Pareto distributions, $BP\,I\left(1,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ and $BP\,II\left(\mu,\sigma,\alpha_1,\alpha_2,\alpha_1',\alpha_2'\right)$ were obtained by transforming the Freund bivariate exponential distribution. Many distributional and reliability properties that find application in Reliability and Economic studies were considered. As mentioned earlier these models do not have a straight forward multivariate extension as in the case of Marshall-Olkin (1967) bivariate exponential distribution. An extension to this distribution has been discussed in Weimann (1966) under certain restrictive conditions. It remains to develop a multivariate distribution and study the properties of the same. One generalization is envisaged as follows. Let there be *n* component parallel system, where the components work independently with lifetimes $X_i^1,\ i=1,\dots,n.$ Let $f_i\left(x_i^1\right), i=1,\dots,n$ denote the probability density function of X_i^1 in the first stage. Assume that there occur no simultaneous failure or failure occurs at stages and failed items are not replaced.

Once a failure occurs the distribution of the remaining lifetime of the other component has same distribution but with possibly changed parameter values. Let X_i^j denote the remaining lifetime of the i^{th} component in the j^{th} stage (if it survived). If $Y_1, Y_2, \ldots, Y_n$ denotes the component lifetime, that is

$$
\begin{aligned}
Y_{i_1} &= X_{i_1}^1 \\
Y_{i_2} &= Y_{i_1} + X_{i_2}^2 \\
Y_{i_3} &= Y_{i_2} + X_{i_3}^3 \\
&\ \,. \\
&\ \,. \\
&\ \,. \\
Y_{i_n} &= Y_{i_{n-1}} + X_{i_n}^n.
\end{aligned}
$$

Then the distribution of the Freund extension is derived as

$$f\left(y_1, y_2, \ldots y_n\right) = \prod_{j=1}^{n} f_{i_j}^{j}\left(y_{i_j}\right) \int_{y_{i_j}}^{\infty} \prod_{k=j+1}^{n} f_{i_k}^{j}\left(y_{i_k}\right) dy_{i_k}\,;\ y_{i_1} < y_{i_2} < \ldots < y_{i_n}.$$

It remains to study the properties of this distribution and this will be taken up as a continuation of this research.

The bivariate residual entropy function which is applicable to a two component parallel system is considered in this thesis. Recently the study on the past entropy has received much attention among the researchers. Hence the study on the bivariate past entropy function for a load sharing dependent models, its properties, characterizations and multivariate extensions can be consider for the future research.

The bivariate inequality measures are introduced in this thesis, which are applicable to data that shows a load sharing dependence. Also the income gap ratio for the rich in the bivariate situation is discussed here. The extension of these inequality measures in the multivariate case as well as the development of income gap ratio for the poor is also to be explored. The properties of the bivariate uniform representation and the study on its association measures are yet to be discussed.

These problems will be taken up as a future course of work plan.

References

1. Aczel, J. (1966): *Lectures on Functional Equations and their Applications*, Academic Press, New York.
2. Ahsanullah, M. and Kabir, A.B.M.L. (1973): A characterization of the Pareto distribution, *The Canadian Journal of Statistics*, 1(1), 109-112.
3. Arnold, B.C. (1975): Multivariate exponential distributions based on hierarchical successive damage, *Journal of Applied Prob*ability, 12, 142-147.
4. Arnold, B.C. (1983): *Pareto distributions*, International Co-operative Publishing house Fairland, MD.
5. Arnold, B.C. (1987): *Majorization and Lorenz order: A brief introduction: Lecturer notes in Statistics, No. 43*, Springer-Verlang, New York.
6. Arnold, B.C. (1989): *Dependence in conditionally specified distributions, in Topics in Statistical Dependence*, IMS Lecture Notes Monograph No. 16, pp. 13-18, Hayward, California.
7. Arnold, B.C. (1995): *Conditional survival models*, *in recent advances in life-testing and reliability (N. Balakrishnan, ed.)*, 589-601, Boca, Raton, Florida: CRC Press.
8. Arnold, B.C. (2005): Inequality measures for multivariate distributions, *Metron*, LXIII (3), 317-327.

9. Arnold, B.C., Castillo, E. and Sarabia, J.M. (1992): *Conditionally specified distributions, Lecture notes in Statistics- 73*, Springer-Verlag, New York.

10. Arnold, B.C., Castillo, E. and Sarabia, J.M. (1993): Multivariate distributions with generalized Pareto conditionals, *Statistics and Probability Letters*, 17, 361-368.

11. Arnold, B.C. and Zahedi, H. (1988): On multivariate mean residual life functions, *Journal of Multivariate Analysis*, 25, 1-9.

12. Asadi, M. (1999): Multivariate distributions characterized by a relationship between mean residual life and hazard rate, *Metrika*, 49, 121-126.

13. 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.

14. Asadi, M., Ebrahimi, N. and Soofi, E.S. (2005): Dynamic generalized information measures, *Statistics and Probability Letters*, 71, 85-98.

15. Asha, G. and Jagathnath Krishna, K.M. (2006): A note on the characterizations of generalized Marshall-Olkin type multivariate Pareto distributions, *Journal of the Indian Statistical Association*, 44(2), 117-125.

16. Asha, G. and Jagathnath Krishna, K.M. (2008): Modeling and characterizations of a bivariate Pareto distribution, *Journal of Statistical Theory and Applications*, 7(4), 435-452.

17. Asha, G. and Jagathnath Krishna K.M. (2010): Inequality measures for bivariate truncated random variables (communicated).

18. Asha, G., Jagathnath Krishna, K.M. and Nair, K.R.M. (2009): Characterization of bivariate distributions using residual measure of uncertainty, *Journal of Probability and Statistical Science*, 7(2), 139-152.

19. Asha, G., Sankaran, P.G. and Nair, N.U. (2003): Probability models with constant total failure rate, *Communication in Statistics: Theory and Methods,* 32 (6), 1089-1099.
20. Azlarov, T.A. and Volodin, N.A. (1986): *Characterization problems associated with the exponential distribution*, Springer Verlag, New York.
21. Bandhyopandhyay, D. and Basu, A.P. (1990): On a generalization of a model by Lindley and Singpurwall, *Advances in Applied Probability*, 22, 498 – 500.
22. Barnett, V. (1980): Some bivariate uniform distributions, *Communications in Statistics-Theory and Methods*, 9(4), 453-461.
23. Basu, A.P. (1971): Bivariate failure rate, *Journal of American Statistical Association,* 66, 103-104.
24. Basu, A.P. and Sun, K. (1997): Multivariate exponential distributions with constant failure rates, *Journal of Multivariate Analysis*, 61, 159–169.
25. Belzunce, F., Candel, J. and Ruiz, J.M. (1998): Ordering and asymptotic properties of residual income distributions, *Sankhya, Ser. B*, 60, 331-348.
26. Belzunce, F., Navarro, F., Ruiz, J.M. and Del Aguila, Y. (2004): Some results on residual entropy function, *Metrika*, 59, 147-161.
27. Bhattacharya, N. (1963): A property of the Pareto distribution, *Sankhya Ser. B*, 25, 195-196.
28. Buchanan, H.B. and Singpurwalla, N.D. (1977): *Some stochastic characterization of multivariate survival. In the theory and applications of reliability Vol.1, edt. C.P. Tsokos and I.N. Shimi*, Academic Press, New York, 329-348.
29. Chakrabarthy, G. (1982): Estimation of Lorenz curve and Gini coefficient from grouped observations, *Artha Vijnana*, 24, 120-133.
30. Chandra, M. and Singpurwalla, N.D. (1981): Relationships between some notions which are common to reliability theory and economics, *Mathematics of Operations Research*, 6, 113-214.

31. Chiragiev, A. and Landsman, Z. (2009): Multivariate flexible Pareto model: dependency structure, properties and characterizations, *Statistics and Probability letters*, 79(16), 1733-1743.

32. Cover, T.M. and Thomas, J.A. (1991): *Elements of information theory*, John Wiley, New York.

33. Cox, D.R. (1972): Regression models and life tables, *Journal of Royal Statistical Society, Ser.B*, 34, 187 – 220.

34. Cuadras, C.M., Cuadras, D. and Lahlou, Y. (2006): Principal components of the general Pareto distribution with applications, *Journal of Statistical Planning and Inference*, 136(8), 2572-2583.

35. Daniels, H.E. (1945): The Statistical theory of the strength of bundles of threads, *I. Proc. R. Soc. Lond. A*, 183, 404-435.

36. Darbellay, G.A. and Vajda, I. (2000): Entropy expressions for multivariate continuous distributions, *IEEE Transaction on Information Theory. IT*, 46, 709–712.

37. Davison, A.C. and Smith, R.L. (1990): Models for exceedances over high thresholds, *Journal of Royal Statistical Society, Ser. B*, 52(3), 393-442.

38. De Simoni, S. (1979): Nota su un indice di concentrazione per variabili multiple, *Quaderno dell' Istituto di Statistica*, Universita di Bologna, 3, 325- 338.

39. Deshpande, J.V., Isha Dewan and Naik-Nimbalkar, U.V. (2007): Two component load sharing systems with applications to biology, *http://www.isid.ac.in/~statmath/eprints*.

40. Drummond, H., Vazquez, E., Sanchez-Colon, S. Martinez-Gomez, M. and Hudson, R. (2000): Competition for milk in the domestic rabbit: survivors benefit from littermate deaths, *Ethology*, 106, 511-526.

41. Ebrahimi, N. (1996): How to measure uncertainty in residual life time distributions, *Sankhya, Ser. A*, 58, 48-56.

42. Ebrahimi, N. (2000): The maximum entropy method for lifetime distributions, *Sankhya, Ser. A*, 62, 236-243.

43. Ebrahimi, N. and Pellerey, F. (1995): New partial ordering of survival functions based on the notion of uncertainty, *Journal of Applied Probability*, 32, 202-211.
44. Ebrahimi, N., Kirmani, S.N.U.A. and Soofi, E.S. (2007): Multivariate dynamic information, *Journal of Multivariate Analysis*, 98, 328–349.
45. Esary, J.D. and Marshall, A.W. (1974): Multivariate distribution with exponential minimums, *Annals of Statistics,* 2(1), 84-98.
46. Feller, W. (1971): *An introduction to probability theory and its applications, Vol. 2*, Wiley, New York.
47. Fermanian, J.D. and Wegkamp, M. (2004): Time dependent copulas, Working paper, *http://www.crest.fr/pageperso/ifa/fermanianfermanian .htm*.
48. Finkelstein, M.S. (2003): On one class of bivariate distributions, *Statistics and Probability Letters*, 65(1), 1-6.
49. Freund, J. (1961): A bivariate extension of exponential distribution, *Journal of American Statistical Association*, 56, 971 – 977.
50. Frosini, B.V. (1988): *Introduzione alla Statistica*, Roma: La Nuova Italia Scientifica.
51. Galambos, J. and Kotz, S., (1978): *Characterization of probability distributions, Lecture Notes on Mathematics,* Springer-Verlag Berlin Heidelberg, New York .
52. Garren, S.T., Hume, E.R. and Leppert, G,R. (2007): Improved estimation of the Pareto location parameter under squared error loss, *Communications in Statistics- Theory and Methods*, 36, 97-102.
53. Gastwirth, J.L. (1971): A general definition of Lorenz curve, *Econometrica*, 39, 1037-1039.
54. Genest, C. and Mackay, J. (1986): The joy of copulas: Bivariate distributions with uniform marginals, *American Statistician*, 40, 280-285.
55. Gini, C. (1912): *Variabilita e Mutabilita*, Bologna.

56. Gradshteyn, I.S., Ryzhik, I.M. and Jeffrey, A. (1994): *Tables of integrals, series and products*, Accademic Press, New York.

57. Guess, F. and Proschan, F. (1988): *Mean residual life*, *In: Rao Krishnaiah, P.R., Rao, C.R. (Eds.), Handbook of Statistics, vol. 7,* North-Holland, Amsterdam, pp. 215–224.

58. Gumbel, E.J. (1960): Bivariate exponential distributions, *Journal of the American Statistical Association*, 55, 698-707.

59. Gupta, A.K. and Nadarajah, S. (2006): Sums, products, ratios for Freund's bivariate exponential distribution, *Applied Mathematics and Computation*, 173, 1334-1349.

60. Gupta, R.C. and Kirmani, S.N.U.A (2004): Some characterizations of distributions by functions of failure rate and mean residual life, *Communications in Statistics- Theory and Methods*, 33(12) 3115-3131.

61. Hanagal, D.D. (1996): A multivariate Pareto distributions, *Communications in Statistics- Theory and Methods*, 25(7) 1471 – 1488.

62. Hao, Z. and Singh, V.P. (2009): Entropy-based parameter estimation for extended Burr XII distribution, *Stochastic Environmental Research Risk Assessment*, 23(8), 1113-1122.

63. Hollander, M. and Pena, E.A. (1995): Dynamic reliability models with conditional proportional hazards, *Lifetime Data Analysis*, 1, 377-401.

64. Jaynes, E.T. (1961): *Probability Theory in Science and Engineering*, Mc Graw-Hill Company, New York.

65. Joe, H. (1997): *Multivariate models and dependence concepts, Monographs on Statistics and Applied Probability*, 73, Chapman and Hall, London.

66. Joe, H. and Hu, T. (1996): Multivariate distributions from mixtures of max-infinitely divisible distributions, *Journal of Multivariate Analysis*, 57(2), 240-265.

67. Joe, H. and Ma, C. (2000): Multivariate survival functions with a min-stable property, *Journal of Multivariate Analysis*, 75(1), 13-35.

68. Johnson, N.L. and Kotz, S. (1975): A vector multivariate hazard rate, *Journal of Multivariate Analysis*, 5, 53-66.

69. Johnson, N.L., Kotz, S. and Balakrishnan N. (1994): *Continuous univariate distributions: Vol. 1*, John Wiley and Sons, New York.

70. Jupp, P.E. and Mardia, K.V. (1982): A characterization of the multivariate Pareto distribution, *The Annals of* Statistics, 10(3), 1021-1024.

71. Justel, A., Pena, D. and Zamar, R. (1997): A multivariate Kolmogorov-Smirnov test of goodness of fit, *Statistics and Probability Letters*, 35, 251-259.

72. Kapur, J.N. (1989): *Maximum entropy models in science and engineering*, Wiley Eastern Limited, New Delhi.

73. Kim, H. and Kvam, P. (2004): Reliability estimation based on system data with an unknown load share rule, *Lifetime Data Analysis*, 10, 83-94.

74. Klefsjo, B. (1984): Reliability interpretations of some concepts from economics, *Nava Research Logistics Quarterly*, 31, 301-308.

75. Koshevoy, G. (1995): Multivariate Lorenz majorization, *Social Choice and Welfare*, 12, 93-102.

76. Koshevoy, G. and Mosler, K. (1996): The Lorenz zonoid of a multivariate distribution, *Journal of the American Statistical Association*, 91, 873-882.

77. Koshevoy, G. and Mosler, K. (1997): Multivariate Gini indices, *Journal of Multivariate Analysis*, 60, 252-276.

78. Kotz, S and Shanbhag, D.N. (1980): Some new approaches to probability distributions, *Advances in Applied Probability,* 12, 903 – 921.

79. Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000): *Continuous multivariate distributions, Vol. 1: Models and applications*, John Wiley, New York.

80. Krishnaji, N. (1970): Characterization of the Pareto distribution through a model of under-reported incomes, *Econometrica*, 38(2), 251-255.

81. Krishnan, P. (1985): Some demographic applications of the univariate and the bivariate forms of the Pareto distribution, *Proceedings of the 45th Session of the International Statistical Institute*, 1, 267-268.

82. Kupka, J. and Loo, S. (1989): The hazard and vitality measures of ageing, *Journal of Applied Probability*, 26 (3), 532-542.

83. Kvam, P.H. and Pena, E.A. (2005): Estimating load sharing properties in a dynamic reliability system, *Journal of American Statistical Association*, 100, 263-272.

84. Lin, G.D. (2003): Characterizations of the exponential distribution via the residual lifetime, *Sankhya, Ser. A*, 65, 249-258.

85. Lindley, D.V. and Singpurwalla, N.D. (1986): Multivariate distributions for the life lengths of components of a system sharing common environment, *Journal of Applied Probability,* 23, 418-431.

86. Lomax, H.S. (1954): Business failures; Another example of the analysis of failure data, *Journal American Statistical Association*, 49, 847-852.

87. Lorenz, M.O. (1905): Methods of measuring the concentration of wealth, *Publications of the American Statistical Association*, 9, 209-219.

88. Lu, J.C. (1989): Weibull extensions of the Freund and Marshall-Olkin bivariate exponential models, *IEEE Transactions on Reliability*, 38(5), 615-620.

89. Lunetta, G. (1972): Sulla concentrazione delle distribuzioni doppie, In: *Atti della XXVII Riunione Scientifica della Societa Italiana di Statistica, Palermo*, 2, 127-150.

90. Lynch, J.D. (1999): On the joint distribution of component failures for monotone load-sharing systems, *Journal of Statistical Planning and Inference*, 78, 13-21.

91. Malik, H.J. and Trudel, R. (1985): Distribution of the product and the quotient from bivariate *t*, *F* and Pareto distributions, *Communications in Statistics-Theory and Methods*, 14, 2951-2602.

92. Mardia, K.V. (1962): Multivariate Pareto distributions, *Annals of Mathematical Statistics*, 33, 1008-1015

93. Marshall, A.W. (1975): Some comments on the hazard gradient, *Stochastic Processes and their Applications*, 3, 293-300.

94. Marshall, A.W. and Olkin, I. (1967): Multivariable exponential distribution, *Journal of American Statistical Association*, 62, 30–44.

95. Moothathu, T.S.K. (1986): A characterization of power function distribution through a property of the Lorenz curve, *Sankhya: Ser. B*, 48, 262-265.

96. Mosler, K. (2002): *Multivariate Dispersion, Central Regions and Depth: The lift zonoid approach*, Springer Lecture Notes in Statistics, Vol. 165, Springer, New York.

97. Mukherjee, S.P and Roy, D. (1986): Some characterizations of the exponential and related life distributions, *Calcutta Statistical Association Bulletin*, 36, 189- 197.

98. Muliere, P. and Scarcini, M. (1987): Characterization of a Marshall-Olkin type class of distributions. *Annals of Institute of Mathematics,* 39, 429-441.

99. Nadarajah, S. (2005): Exponential Pareto distributions, *Statistics*, 39(3), 255-260.

100. Nadarajah, S. (2008): A bivariate Pareto model for drought, *Stochastic Environmental Research and Risk Assessment*, 23(4), 811-822.

101. Nadarajah, S. and Gupta, A.K. (2008): A product Pareto distribution, *Metrika*, 68,199-208.

102. Nadarajah, S. and Zografos, K. (2005): Expression for Renyi and Shannon entropies for bivariate distributions, *Information Sciences*, 170, 173-189.

103. Nair, K.R.M. and Nair, N.U. (1988): Bivariate mean residual life, *IEEE Transactions on Reliability*, 38, 362-364.

104. Nair, K.R.M. and Rajesh, G. (1998): Characterizations of probability distributions using the residual entropy function, *Journal of Indian Statistical Association*, 36, 157-166.

105. Nair, K. R. M. and Rajesh, G. (2000): Geometric vitality function and its applications to reliability, *IAPQR Transactions*, 25(1), 1-8.

106. Nair, N. U and Asha, G. (1997): Some class of multivariate life distributions in discrete time, *Journal of Multivariate Analysis*, 62, 181-189.

107. Navarro, J., Ruiz, J.M. and Sandoval, C.J. (2008): Properties of systems with two exchangeable Pareto components, *Statistical Papers*, 49, 177-190.

108. Nayak, T.K. (1987): Multivariate Lomax distribution: properties and usefulness in reliability theory. *Journal of Applied Probability,* 24*,* 170-177.

109. Nelsen, R.B. (1999): *An introduction to copulas*, Springer, New York.

110. Ord, J.K., Patil, G.P. and Taillie, C. (1983): Truncated distributions and measures of income inequality, *Sankhya, Ser. B*, 45, 413-430.

111. Pareto, V. (1897): *Cours d' Economic Politique*, Rouge and Cie, Lansanne and Paris.

112. Pickands, J. (1975): Statistical inference using extreme order Statistics, *Annals of Statistics*, 3, 119-131.

113. Quandt, R.E. (1966): Old and new methods of estimation and the Pareto distribution, *Metrika*, 10, 55-82.

114. Revankar, N.S., Hartley, M.J. and Pagano, M. (1974): A characterization of the Pareto distribution, *The Annals of Statistics*, 2(3), 599-601.

115. Rosen, B.W. (1964): Tensile failure of fibrous composites, *AIAA Journal*, 2, 1985-1991.

116. Roy, D. (2004): A Generalization of the lack of memory property and Related characterization Results, *Communications in Statistics - Theory and Methods*, 33, 3145 – 3158.

117. Ruiz, J.M. and Navarro, J. (1994): Characterizations of distributions by relationships between failure rate and mean residual life, *IEEE Transactions on Reliability*, 43(4), 640 - 644.

118. Sankaran, P.G. and Gupta, R.P. (1999): Characterization of lifetime distributions using measure of uncertainty, *Calcutta Statistical Association Bulletin*, 49, 195-196.

119. Sankaran, P.G. and Nair, N.U. (1991): *On bivariate vitality functions*, Proceedings of National Symposium on Distribution Theory.

120. Sankaran, P.G. and Nair, N.U. (1993): A bivariate Pareto model and its applications to reliability, *Naval Research Logistic Quarterly*, 40, 1013-1020.

121. Sathar, A.E.I. (2002): *On truncated versions of certain measures of inequalities and stabilities*, PhD Thesis, Cochin University of Science and Technology.

122. Sathar, A.E.I., Suresh, R.P. and Nair, K.R.M. (2007): A vector valued bivariate Gini index for truncated distributions, *Statistical Papers*, 48, 543-557.

123. Serfling, R.J. (1980): *Approximation theorems of mathematical statistics*, John Wiley, New York.

124. Shaked, M. and Shanthikumar, J.G. (1987): The multivariate hazard construction, *Stochastic Processes and their Applications*, 24, 241-258.

125. Shaked, M. and Shanthikumar, J.G. (1991): Dynamic multivariate mean residual life functions, *Journal of Applied Probability*, 28, 613-629.

126. Shanbag, D.N. and Kotz, S. (1987): Some new approaches to multivariate probability distributions, *Journal of Multivariate Analysis*, 22(2), 189-211.

127. Shannon, C.E. (1948): A mathematical theory of communication, *Bell System Technical Journal*, 27, 379 – 423.

128. Singh, V.P. (1998): *Entropy based parameter estimation in hydrology*, Kluwer Academic Publishers, The Netherlands.

129. Singh, V.P. and Ahmad, M. (2004): A comparative evaluation of the estimators of the three-parameter generalized Pareto distribution, *Journal of Statistical Computation and Simulation*, 74 (2), 91-106.

130. Singh, V.P. and Guo, H. (1995): Parameter estimation for 3-parameter generalized Pareto distribution by the principle of maximum entropy (POME), *Hydrological Sciences- Journal-des Sciences Hydrologiques*, 40 (2), 165-181.

131. Singh, V.P. and Guo, H. (1997): Parameter estimation for 2-parameter generalized Pareto distribution by POME, *Stochastic Hydrology and Hydraulics*, 11, 211-227.

132. Singh, V.P. and Rajagopal A.K. (1986): A new method of parameter estimation for hydrologic frequency analysis, *Hydrological Science and Technology*, 2(3), 33-40.

133. Sklar, A. (1959): *Fonctions de repartition a n dimensions et leurs marges*, Public Institute Statistics, University of Paris, 8, 229-231.

134. Somesh Kumar and Bandhyopadhyay, A.K. (2005): Estimating parameters of a selected Pareto population, *Statistical Methodology*, 2, 121-130.

135. Srivastava, T. and Gupta, H. (2005): Pitman estimator to Pareto scale parameters, *Communications in Statistics-Theory and Methods*, 34, 2235-2255.

136. Sun, K. and Basu, A.P. (1995): A characterization of a bivariate geometric distribution, *Statistics and Probability Letters*, 23 (4), 307-311.

137. Swartz, G.B. (1973): The mean residual life time functions, *IEEE Transactions on Reliability*, 22, 108-109.

138. Taguchi, T. (1972 a): On the two dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case I, *Annals of the Institute of Statistical Mathematics*, 24, 355-382.

139. Taguchi, T. (1972 b): On the two dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case II, *Annals of the Institute of Statistical Mathematics*, 24, 599-619.
140. Talwaker, S. (1980): On the property of dullness of Pareto distribution. *Metrika*, 27, 115-119.
141. Tang, L.C., Lu, Y. and Chew, E.P. (1999): Mean residual life of lifetime distributions, *IEEE Transactions on Reliability*, 48(1), 73 - 78.
142. Tribus, M. (1969): *Rational Descriptions, Decisions and Designs*, Pergamon, New York, USA.
143. Veenus, P and Nair, K. R. M (1994): Characterization of a bivariate Pareto distribution, *Journal of Indian Statistical Association*, 32, 15 – 20.
144. Weimann, D.G. (1966): *A multivariate extension of the exponential distribution*, Ph.D. Thesis, Arizona State University.
145. Wesolowski, J. (1994): *Bivariate distributions via the Pareto conditional distribution and a regression function*, Preprint, Mathematical Institute, Warsaw University of Technology, Warsaw, Poland.
146. Weymark, J.A. (2004): *The normative approach to the measurement of multidimensional inequality*, Working paper #03-W14R, Department of Economics, Vanderbilt University, Nashville.
147. Wiener, N. (1948): *Cybernetics*, John Wiley, New York.
148. Xekalaki, E, and Dimaki, C. (2004): Characterizations of bivariate Pareto and Yule distributions, *Communication in Statistics-Theory and Methods*, 33 (12), 3033-3042.
149. Yeh, H.C. (1994): Some properties of the homogeneous multivariate Pareto IV distributions, *Journal of Multivariate Analysis*, 51, 46-52.
150. Yeh, H.C. 2004(a): Some properties and characterizations for generalized multivariate Pareto distributions, *Journal of Multivariate Analysis,* 88, 47-60.

151. Yeh, H.C. 2004(b): The Generalized Marshall-Olkin type multivariate Pareto distributions, *Communication in Statistics-Theory and Methods*, 33, 1053-1068.

152. Zografos, K. and Nadarajah, S. (2005): Expression for Renyi and Shannon entropies for multivariate distributions, *Statistics and Probability Letters*, 71, 71-84.