ON THE ANALYSIS OF BIVARIATE LIFETIME DATA - SOME MODELS AND APPLICATIONS

Thesis Submitted to the

Cochin University of Science and Technology

for the Award of the Degree of

Doctor of Philosophy

under the Faculty of Science

by

Vincent Raja A

Department of Statistics
Cochin University of Science and Technology
Cochin-682022.

Dedicated To
My Family...

CERTIFICATE

Certified that all the relevant corrections and modifications suggested by the audience during pre-synopsis seminar and recommended by the Doctoral committee of the candidate have been incorporated in the thesis.

Cochin- 22

April 2017

Dr. Asha Gopalakrishnan

Professor,

Department of Statistics,

Cochin University of

Science and Technology.

CERTIFICATE

This is to certify that the thesis entitled **"ON THE ANALYSIS OF BIVARIATE LIFETIME DATA-SOME MODELS AND APPLICATIONS"** is a bonafide record of work done by **Mr. Vincent Raja A** under my guidance in the Department of Statistics, Cochin University of Science and Technology and that no part of it has been included anywhere previously for the award of any degree or title.

Cochin-22
April 2017

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

DECLARATION

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

Cochin- 22 **Vincent Raja A**

April 2017

Acknowledgements

I would like to express my greatest appreciation and filled with gratitude to my supervising guide Dr. Asha Gopalakrishnan, Professor in Statistics, Department of Statistics, Cochin University of Science and Technology, Cochin-22, Kerala, India. Madam, you have been a tremendous mentor for me. I want to thank you for encouraging my research and moulding me to grow as a researcher. Your advice on both research and teaching has brought great impact on my career.

I sincerely thank Dr.P.G. Sankaran. Professor and Head, Department of Statistics, CUSAT for his constant encouragement and support.

I thank Dr. Unnikrishnan Nair, Retired Professor, Dr. V. K. Ramachandran Nair, Retired Professor, Dr. N. Balakrishna, Professor, Dr. K. C. James, Professor, and Dr. S. M. Sunoj, Professor, Dr. G. Rajesh, Assistant Professor, Dr. K.G. Geetha, Lecturer (Deputation), Department of Statistics, CUSAT, for their valuable suggestions and help to complete this endeavour. I express my sincere gratefulness to all the former teachers, who provided insight and expertise. I extend my sincere thanks to all non-teaching staff of the Department of Statistics, CUSAT for their kind cooperation.

I wish to extend my profound appreciation and gratitude to Dr. Nalini Ravishanker, Professor in Statistics, Department of Statistics, University of Connecticut, Storrs, USA, Dr. José María Sarabia, Professor in Statistics, Department of Economics, University of Cantabria, Santander, Spain and Dr. I. Elbatal, Professor, Department of Mathematical Statistics, Institute of Statistical Studies and Research (ISSR). Cairo University, Egypt for their colloborative research works.

I would like to thank my fellow research scholars in my department for the constructive comments and support throughout this venture.

I owe my appreciation to Cochin University of Science and Technology, Cochin-22, Kerala, India for providing me financial support to carry out this work under University Junior Research Fellowship and University Senior Research Fellowship.

I would like to mark my greatest appreciation and gratitude to the management and administration of University of Guyana, Georgetown, Guyana, South America for their unending support and encouragement.

Finally, I must express my very profound gratitude to my parents, my wife, my loving children and other family members for providing me with unfailing support and continuous encouragement throughout my study. I owe everything to them.

I remember all my best friends and well wishers who have been constantly encouraging me throughout for the successful completion of this project. I extend my heartfelt gratitute to all of them.

Above all, I render my sincere gratitude and bow before the Almighty for granting me the knowledge, wisdom and understanding to proceed successfully.

Vincent Raja A

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In the past two decades modelling and analysing of bivariate lifetime data have brought a great interest and enthusiasm among researchers in the area of reliability theory and survival analysis. Bivariate lifetime data includes a parallel clustered data in which a system has more than one failure which are observed in parallel and do not satisfy any ordered restriction, like the times to initial contraction of disease of rats in the litter mates study (Mantel et al. (1977)), the time to deterioration level or time to reaction of a treatment for pairs of lungs, kidneys, eyes or ears of humans (Gross et al. (1971)), times to onset of blindness for the treated eye and the untreated eye in a diabetic retionpathy study data (Huster et al. (1989)), the inception times of defect and the additional times to failure of cable insulation specimens study (Stone (1978)) or the lifetimes of two motors working in a parallel system (ReliaSoft (2003)) in a reliability context.

The present thesis focuses on developing some new bivariate models for such data by taking into account particular features exhibited by the data. Broadly speaking, two types of bivariate models have been focussed on - conditionally specified models and load sharing models. First we propose a class of bivariate distributions with specified transmuted conditional distributions. Univariate transmuted distributions are based on a transmuted map (Shaw & Buckley (2009)) and was initially used for introducing skewness to a symmetric baseline distribution. However, there is no specific requirement that the baseline distribution is symmetric and hence transmuted distributions (Shuaib et al. (2016)) are of recent interest as a rich class of distributions for modelling data. Conditionally specified models are apt to model casual failures as in cable insulation specimens data (Stone (1978)) which describe the failure phenomenon called electrical treeing. The observed data here consist of two variables (Y_1, Y_2) where Y_1 is the time to inception of a defect and Y_2 is the subsequent additional time to specimen failure. This motivated us to propose bivariate distributions with transmuted conditionals.

However when dependence between (Y_1, Y_2) is characterized in terms of the change in stochastic behaviour of the surviving component, on failure of a component, it becomes necessary to take into account this feature in modelling. These type of models are popularly referred to as load share models. The bivariate exponential distribution (Freund (1961)) is an apt model for such system when lifetimes are exponential. Generalizations of these models form the topic of study in the later part of the present Thesis.

Before embarking on these models we first recall few basic definitions and results which will be used in the sequel. We also discuss a particular example of transmuted distribution in detail which form the basis of developing bivariate models discussed in later chapters.

1.1 Basic Concepts

Let (Y_1, Y_2) be random variables representing lifetimes, not necessarily independent. Then the bivariate cumulative distribution function and survivor function for $y_1 \geq 0$ and $y_2 \geq 0$ are defined respectively as

$$F(y_1, y_2) = P(Y_1 \leq y_1, Y_2 \leq y_2) \tag{1.1}$$

and

$$S(y_1, y_2) = P(Y_1 \geq y_1, Y_2 \geq y_2). \tag{1.2}$$

or equivalently

$$S(y_1, y_2) = 1 - F(y_1, \infty) - F(\infty, y_2) + F(y_1, y_2), \tag{1.3}$$

for continuous random variables (Y_1, Y_2), where the marginal distributions of Y_1 and Y_2 are defined as $F_1(y_1) = F(y_1, \infty)$ and $F_2(y_2) = F(\infty, y_2)$. Similarly, we can define the marginal survivor functions as $S_1(y_1) = S(y_1, 0)$ and $S_2(y_2) = S(0, y_2)$.

If the second order derivative exists then the corresponding bivariate probability density function is defined by

$$f(y_1, y_2) = \frac{\partial^2 F(y_1, y_2)}{\partial y_1 \partial y_2} = \frac{\partial^2 S(y_1, y_2)}{\partial y_1 \partial y_2}, \tag{1.4}$$

and the marginal densities are respectively,

$$f_1(y_1) = \frac{dF(y_1, \infty)}{dy_1} = \frac{-dS(y_1, 0)}{dy_1}$$

and

$$f_2(y_2) = \frac{dF(\infty, y_2)}{dy_2} = \frac{-dS(0, y_2)}{dy_2}.$$

In rest of the discussions, we confine to absolutely continuous random variables Y_1, Y_2 with distribution $F(y_1, y_2)$ with support $\mathscr{S} \text{x} \mathscr{S} = \{(y_1, y_2) | y_i \in \mathscr{S},\ i = 1, 2\}$ and $\mathscr{S} = (0, \infty)$.

1.2 Bivariate Failure Rate

Basu (1971) defined the bivariate failure (hazard) rate to be

$$r(y_1, y_2) = \frac{f(y_1, y_2)}{S(y_1, y_2)}.$$

It is shown in Basu (1971) that, except for the case of independence, there does not exist any absolutely continuous bivariate exponential distribution with constant bivariate failure rate and marginal exponential distributions.

Johnson & Kotz (1975) defined the hazard gradient as the vector,

$$(r_1(y_1, y_2), r_2(y_1, y_2))^{\mathcal{T}} = -\nabla \log S(y_1, y_2), \tag{1.5}$$

where $\nabla = \left(\frac{\partial}{\partial y_1}, \frac{\partial}{\partial y_2}\right)$, $r_1(y_1, y_2)$ is the hazard rate of the conditional distribution of Y_1 given $\{Y_2 > y_2\}$ and $r_2(y_1, y_2)$ is the hazard rate of the conditional distribution of Y_2 given $\{Y_1 > y_1\}$, that is

$$r_1(y_1, y_2) = r_{Y_1|Y_2>y_2}(y_1) = -\frac{\partial}{\partial y_1} \log S(y_1, y_2) = \frac{1}{S(y_1, y_2)} \int_{y_2}^{\infty} f(y_1, v)dv, \tag{1.6}$$

and

$$r_2(y_1, y_2) = r_{Y_2|Y_1>y_1}(y_2) = -\frac{\partial}{\partial y_2} \log S(y_1, y_2) = \frac{1}{S(y_1, y_2)} \int_{y_1}^{\infty} f(u, y_2)du. \tag{1.7}$$

The most popular version of bivariate failure rate proposed by Cox (1972) in a vector form for $\boldsymbol{y} = (y_1, y_2)$ is defined as

$$\boldsymbol{\lambda(y)} = \left(\lambda(y), \lambda_{12}(y_1|y_2), \lambda_{21}(y_2|y_1)\right), \tag{1.8}$$

where,

$$\lambda(y) = \lambda_{10}(y) + \lambda_{20}(y), \;\; y_1 = y_2 = y,$$

$$\lambda_{i0}(y) = \lim_{\Delta y \to 0+} \frac{P\left(y \le Y_i < y + \Delta y | y \le Y_1, y \le Y_2\right)}{\Delta y}, \;\; i = 1, 2, \tag{1.9}$$

$$\lambda_{21}(y_2|y_1) = \lim_{\Delta y_2 \to 0+} \frac{P\left(y_2 \le Y_2 < y_2 + \Delta y_2 | y_2 \le Y_2, Y_1 = y_1\right)}{\Delta y_2}, \;\; y_1 < y_2, \tag{1.10}$$

and

$$\lambda_{12}(y_1|y_2) = \lim_{\Delta y_1 \to 0+} \frac{P\left(y_1 \le Y_1 < y_1 + \Delta y_1 | y_1 \le Y_1, Y_2 = y_2\right)}{\Delta y_1}, \;\; y_2 < y_1. \tag{1.11}$$

The appropriateness of using Cox failure rate for load sharing dependence which we will discuss later is discussed in Singpurwalla (2006) and Jagathnath (2010) and references cited therein.

Let X denote the lifetime of a component so that its reliability $S(x) = P(X > x)$

and $S'(x) = -f(x)$. The univariate failure rate or hazard rate is defined as

$$r(x) = \lim_{\Delta x \to 0} \frac{P[x \leq X < x + \Delta x | X > x]}{\Delta x}. \tag{1.12}$$

When $f(.)$ is the probability density function of X, (1.12) can be equivalently written as

$$\begin{aligned} r(x) &= \frac{f(x)}{S(x)} \\ &= \frac{d}{dx}\left[-\log S(x)\right]. \end{aligned} \tag{1.13}$$

Mean time to failure (MTTF):

The expected life or the mean time to failure (MTTF) of the component is given by

$$E(X) = \int_0^\infty S(t)dt. \tag{1.14}$$

1.3 Global Dependence Measures

A measure of dependence explains in a manner how closely the variables Y_1 and Y_2 are related. Correlation is the global dependence measure which explains the relationship between Y_1 and Y_2 when $E(Y_2|Y_1)$ or $E(Y_1|Y_2)$ are linearly independent. If it is not linearly independent then we have to search for other nonparametric global dependence measures which are commonly based on concordance and discordance.

Let (y_{1i}, y_{2i}) and (y_{1j}, y_{2j}), $i \neq j = 1, 2$, be called concordant if $(y_{1i} - y_{1j})(y_{2i} - y_{2j}) > 0$ and discordant if $(y_{1i} - y_{1j})(y_{2i} - y_{2j}) < 0$. Geometrically, two distinct points (y_{11}, y_{21}), (y_{12}, y_{22}) in the plane are said to be concordant, if the line segment connecting them has positive slope and discordant if the line segment has negative slope.

Two most popular and important measures of dependence based on the concordance and discordance are Kendall's tau and Spearman's rho. These two dependence measures provide an alternative measure to the linear correlation coefficient for non-normal distributions. The relationship between these two global dependence measures is explained in Nelsen (2006) and Joe (1997). These two measures are explained in Sections 1.3.1 and 1.3.2 respectively.

1.3.1 Kendall's Tau

Kendall's tau (τ_K) is defined as the probability of concordance minus the proability of discordance. That is,

$$\begin{aligned}\tau_K &= P\left[(Y_{1i}-Y_{1j})(Y_{2i}-Y_{2j})>0\right]-P\left[(Y_{1i}-Y_{1j})(Y_{2i}-Y_{2j})<0\right]\\ &= 4\int F(y_1,y_2)dF(y_1,y_2)-1\\ &= 4E[F(Y_1,Y_2)]-1.\end{aligned}$$

The range of possible values for τ_K is $[-1,1]$.

1.3.2 Spearman's Rho

Spearman's rho (ρ_S) is another global dependence measure, which is also defined based on concordance and discordance. Let (Y_{11},Y_{12}), (Y_{21},Y_{22}), (Y_{31},Y_{32}) be three independent pairs of random variables with a common distribution function F. Then ρ_S is defined to be proportional to the probability of concordance minus the probability of discordance for the two pairs (Y_{11},Y_{12}) and (Y_{21},Y_{32}), that is,

$$\begin{aligned}\rho_S &= 3\left\{P\left[(Y_{11}-Y_{21})(Y_{12}-Y_{32})>0\right]-P\left[(Y_{11}-Y_{21})(Y_{12}-Y_{32})<0\right]\right\}\\ &= 12\int\int F_1(y_1)F_2(y_2)dF(y_1,y_2)-3\\ &= 12\int\int S(y_1,y_2)dF_1(y_1)dF_2(y_2)-3.\end{aligned}$$

The range of possible values for ρ_S is $[-1,1]$.

1.3.3 Positive and Negative Quadrant Dependence

If the random variables Y_1 and Y_2 are said to be positively(negatively) quadrant dependent ($PQD(NQD)$), if

$$P(Y_1 > y_1, Y_2 > y_2) \geq (\leq) P(Y_1 > y_1) P(Y_2 > y_2), \text{ for all } y_1 \text{ and } y_2,$$

or equivalently

$$P(Y_1 \leq y_1, Y_2 \leq y_2) \geq (\leq) P(Y_1 \leq y_1) P(Y_2 \leq y_2), \text{ for all } y_1 \text{ and } y_2 \in \mathscr{S}.$$

Among the global dependence measures - total positivity of order 2 ($TP2$) or reverse rule of order 2 ($RR2$) is considered to be a strong measure of dependence and in the next section we recall this global measure.

1.3.4 Total Positivity of Order 2

The global dependence measures total positive of order 2 ($TP2$) functions and reverse rule of order 2 ($RR2$) functions are defined as follows.

Definition 1.3.1. *Let Y_1 and Y_2 have a joint probability density function $f(.,.)$. Then $f(y_1, y_2)$ is said to be totally positive (reverse order) of order 2 $TP2$ ($RR2$) if for all $y_{11} < y_{12}$, $y_{21} < y_{22}$,*

$$f(y_{11}, y_{21}) f(y_{12}, y_{22}) - f(y_{11}, y_{22}) f(y_{21}, y_{12}) \geq 0 \ \ (\leq 0). \tag{1.15}$$

These properties are the strongest of all dependence concepts existing in the literature. For other global dependence measures one can refer to Joe (1997).

1.4 Local Measures of Dependence

The global dependence measures are defined from the moments of the distribution on the whole plane and it can be zero when Y_1 and Y_2 are not independent. Therefore, it is necessary to study the dependence locally. Holland & Wang (1987) defined a local dependence measure, $\eta(y_1, y_2) = \frac{\partial^2}{\partial y_1 \partial y_2} \log f(y_1, y_2)$ and established the relation between local dependence function with $TP2$ and $RR2$ properties which is stated in Theorem 1.4.1.

Theorem 1.4.1. *Let $f(y_1, y_2)$ be the joint probability density function of (Y_1, Y_2) with support on a set $\mathscr{S}^2$. Then $f(y_1, y_2)$ is $TP2$ $(RR2)$ if and only if $\eta(y_1, y_2) > 0$ (< 0).*

1.4.1 Cross Ratio Function

In order to study the dependence between the random variables Y_1 and Y_2 we consider the local dependence function. The cross ratio function (CRF) is a local dependence function related to the hazards of events and was originally introduced by Clayton (1978) and studied by Oakes (1989). It is the ratio of the hazard of Y_1 given that Y_2 the second component failed at time y_2 to the hazard of Y_1 given that the second component had not failed at time y_2. It helps to quantify the association in bivariate survival data. The two event times Y_1 and Y_2 are independent if $\mathscr{C}(y_1, y_2) = 1$, positively correlated if $\mathscr{C}(y_1, y_2) > 1$, and negatively correlated if $\mathscr{C}(y_1, y_2) < 1$ (Kalbfleisch & Prentice (2002)). The CRF at time point (y_1, y_2) is given by

$$\mathscr{C}(y_1, y_2) = \frac{S(y_1, y_2) S_{12}(y_1, y_2)}{S_1(y_1, y_2) S_2(y_1, y_2)}, \tag{1.16}$$

where $S_j(y_1, y_2) = \frac{\partial S(y_1, y_2)}{\partial y_j}, j = 1, 2$ and $S_{12}(y_1, y_2) = \frac{\partial^2 S(y_1, y_2)}{\partial y_1 \partial y_2}$. It is a measure of choice for assessing the time varying dependence.

1.5 Some Methods of Constructing Bivariate Distributions

The construction, study and applications of bivariate distributions is one of the essential areas of research in statistics, and it remains to be an active field of research. Recently, several books have been published containing the theory about bivariate non-normal distributions. Hutchinson & Lai (1990), Joe (1997), Arnold et al. (1999a), Kotz & Nadarajah (2000), Kotz & Nadarajah (2004), Nelsen (2006) to mention a few. In this section we discuss some methods for constructing bivariate distributions. For comprehensive and more detailed reviews on methods for constructions of discrete and continuous bivariate distributions one can refer to (Lai (2004), Lai (2006)) and Sarabia & Gómez (2008).

Some of the popular methods which gained considerable attention and interest in the recent literature and the methods used in the present thesis to construct few new general class of bivariate distributions are chosen for our discussion. Note that these are general methods not confined to non-negative random variables.

1.5.1 Marginal Transformation Method

The general description about the marginal transformation method is defined as follows. Let Y_1 and Y_2 be two continuous random variables with probability density functions f_1 and f_2 respectively. Let F_1 and F_2 be their corresponding distribution functions. If we start with a bivariate distribution $F(y_1, y_2)$ (with density $f(y_1, y_2)$) and apply monotone transformations $Y_1 \to Y_1^*$ and $Y_2 \to Y_2^*$. Now, the new distribution $F^*(y_1^*, y_2^*)$ retains the same bivariate structure as the original F, with changes in the marginals, that is, F_1 becoming F_1^* and F_2 becoming F_2^*. Some of the familiar examples in the univariate situations are transforming the normal distribution so that it becomes lognormal, transforming the exponential so that it becomes Weibull. A well known set of distributions constructed through the marginal transformation method is mainly contributed by Johnson (1949), who started with bivariate normal and transformed Y_1 and/or Y_2 to lognormal, logit-normal, and $sinh^{-1}$-normal. This method popularly known as *translation method*. Kimeldorf & Sampson (1978) derived a new bivariate distribution with desired marginals for a given bivariate cumulative distribution function F with marginal cumulative distribution functions F_1 and F_2 by considering the cumulative distribution function $F(F_1^{-1}(G_1(y_1)), F_2^{-1}(G_2(y_2)))$, where $F_i^{-1}(t) = \inf\{x : F_i(x) \geq t\}, i = 1, 2$, the baseline cumulative distribution $G_1(.)$ is fixed and does not depend on y_2, similarly, $G_2(.)$ is fixed and does not depend on y_1. Frêchet (1951) argued that every univariate distribution can be generalized in many ways. He came up with the idea of boundary distributions for all bivariate distributions with specified marginals.

1.5.2 Copula Method

For the past three decades there has been a great interest in the study of uniform representation (also known as copulas) for constructing bivariate distributions. This is the form the distribution takes when Y_1 and Y_2 are transformed so that each marginal will have a uniform distribution over the range 0 to 1. For example, let us consider

$$F(y_1, y_2) = y_1 y_2 \left[1 + \theta(1 - y_1)(1 - y_2)\right], \tag{1.17}$$

for y_1 and y_2 between 0 and 1, with $-1 \leq \theta \leq 1$. By setting $y_2 = 1$, we see that the distribution of Y_1 is uniform, $F_1 = y_1$; similarly, setting $y_1 = 1$, the distribution of Y_2 becomes uniform, $F_2 = y_2$. The copula given in (1.17) is known as Farlie-Gumbel-Morgenstern copula.

If we require to convert the marginals to be exponential, then let us consider $F_1(y_1) = 1 - e^{-y_1}$ and $F_2(y_2) = 1 - e^{-y_2}$. Now, replacing y_1 as $F_1(y_1) = 1 - e^{-y_1}$ and y_2 by $F_2(y_2) = 1 - e^{-y_2}$ in (1.17), we get

$$F(y_1, y_2) = \left(1 - e^{-y_1}\right)\left(1 - e^{-y_2}\right)\left[1 + \theta e^{-(y_1 + y_2)}\right], \tag{1.18}$$

which is Gumbel bivariate exponential distribution (Gumbel (1960)). Subsequently Gumbel (1961) also studied bivariate logistic distributions.

One can construct a bivariate distribution after determining the coupula C by using, Sklar's theorem given in (1.19).

$$F(y_1, y_2) = C\left(F_1(y_1), F_2(y_2)\right). \tag{1.19}$$

Marshall & Olkin (1988) considered a general method for generating bivariate distributions using the mixture

$$F(y_1, y_2) = \int\int K(F_1^{\theta_1}, F_2^{\theta_2})dG(\theta_1, \theta_2), \tag{1.20}$$

where K is a copula, G is a mixing distribution. Therefore, various choices of G and K will lead to a variety of distributions with marginals as parameters. Note that F_1 and F_2 here are not the marginals of F. Joe (1993) studied the properties of a group of copulas given by Marshall & Olkin (1988). For an extensive study on copulas one can refer to Nelsen (2006).

1.5.3 Method of Mixing and Compounding

A simple and easy way of constructing a bivariate distribution is by the method of mixing along with two distributions. Particularly, if F_1 and F_2 are two bivariate distribution functions, then the new bivariate distribution is given by

$$F(y_1, y_2) = \theta F_1(y_1, y_2) + (1 - \theta)F_2(y_1, y_2),\ 0 \leq \theta \leq 1. \tag{1.21}$$

For exmaples one can refer to Fréchet (1960) and Mardia (1970).

1.5.4 Trivariate Reduction Method

The trivariate reduction or variables in common technique is another method for constructing bivariate distributions. The basic idea is to create a pair of dependent random variables from three or more random variables. The initial random variables are assumed to be independent. The functions that connect initial variables are generally elementary functions. The general definition is given by

$$\begin{aligned} Y_1 &= \mathscr{D}_1(\mathscr{E}_{Y_1}, \mathcal{C}_{Y_1Y_2}), \\ Y_2 &= \mathscr{D}_2(\mathscr{E}_{Y_2}, \tilde{\mathcal{C}}_{Y_1Y_2}), \end{aligned}$$

where $\mathscr{E}_{Y_1}$, $\mathscr{E}_{Y_2}$ represent two sets containing the specific variables of Y_1, Y_2 respectively, ($\mathcal{C}_{Y_1Y_2}$, $\tilde{\mathcal{C}}_{Y_1Y_2}$) sets containing the common or latent variables and $\mathscr{D}_1$, $\mathscr{D}_2$ are the functions that connect the initial variables. Vernic (1997), Vernic (2000) studied bivariate generalized Poisson distribution. Olkin & Liu (2003) proposed the following method for constructing bivariate beta distribution. Let $Y_{1i} \sim \mathcal{G}(a_i, 1)$, $i = 1, 2, 3$, three independent gamma variables with unit scale parameters, and define

$$Y_1 = \frac{Y_{11}}{Y_{11} + Y_{13}},$$
$$Y_2 = \frac{Y_{12}}{Y_{11} + Y_{13}}.$$

The joint probability density function with correlated beta distributions $\mathcal{B}(a_1, a_3)$ and $\mathcal{B}(a_2, a_3)$, $0 \leq y_1, y_2 \leq 1$ is given by

$$f(y_1, y_2; a_1, a_2, a_3) = \frac{y_1^{a_1-1} y_2^{a_2-1} (1-y_1)^{a_1+a_3-1} (1-y_2)^{a_1+a_2-1}}{\mathcal{B}(a_1, a_2, a_3)(1-y_1y_2)^{a_1+a_2+a_3}},$$

where $\mathcal{B}(a_1, a_2, a_3) = \prod_{i=1}^{3} \frac{\Gamma a_i}{\Gamma(\sum_{i=1}^{3} a_i)}$.

Fang et al. (1990) proposed bivariate t-distribution by defining the random variables

$$Y_1 = \frac{Y_{11}}{\sqrt{Y_{13}/n_1}},$$
$$Y_2 = \frac{Y_{12}}{\sqrt{Y_{13}/n_1}},$$

where Y_{11}, Y_{12}, Y_{13} are mutually independent random variables with distributions $Y_{11}, Y_{12} \sim N(0, 1)$ (standard normal) and $Y_{13} \sim \chi^2_{n_1}$ (χ^2-distribution with n_1 degrees of freedom). Note that the marginal distributions are both student t-distribution with n_1 degrees of freedom.

Bivariate F distribution is proposed by Kotz et al. (2000) by considering the random variables Y_{11}, Y_{12} and Y_{13} be mutually independent chi-squared random variables with degrees of freedom $n_i > 0$, $i = 1, 2, 3$, respectively. The corresponding random variables are defined by

$$Y_1 = \frac{Y_{11}/n_1}{Y_{13}/n_3}, \ \ Y_2 = \frac{Y_{12}/n_2}{Y_{13}/n_3}.$$

The marginals $Y_1 \sim F_{n_1,n_3}$ and $Y_2 \sim F_{n_2,n_3}$, which share the degrees of freedom on the denominator.

1.5.5 Frailty Approach

The concept of frailty provides a convenient way to incorporate unobserved covariates (random effect) and associations into models for survival data, where the random effect has a multiplicative effect on the hazard function. For a univariate survival data the frailty model may be formulated in the following manner.

Let X be a survival time with an absolute continuous distribution. A non-negative random variable Z is called frailty or random effect (Vaupel et al. (1979)) if the conditional hazard function given Z has the form:

$$r(x|Z) = Zr_0(x), \tag{1.22}$$

where $r_0(x)$ is the baseline hazard function. The conditional survival for this case is given by

$$S(x|Z) = e^{-ZH(x)}, \tag{1.23}$$

where $H(x) = \int_0^x r_0(t)dt$ is the cumulative baseline hazard. By taking expectation the marginal survival function $S(x)$ is obtained as

$$S(x) = E[S(x|Z)] = E[e^{-ZH(x)}], \tag{1.24}$$

(Hougaard (1984)). Alternatively, (1.24) can be written as

$$S(x) = L[H(x)], \tag{1.25}$$

where L is the Laplace transform of the frailty distribution. Similarly, we can write the bivariate survival function conditional on Z as

$$S(y_1, y_2|Z) = \exp\left[-ZH_1(y_1) + H_2(y_2)\right], \tag{1.26}$$

where $H_i(y_i) = \int_0^{y_i} r_i(t)dt,\ i = 1, 2$ are the integrated hazards of Y_i. Here Y_1 and Y_2 are conditionally independent. The unconditional bivariate survival function is given as,

$$\begin{aligned} S(y_1, y_2) &= \int_z e^{-Z\{H_1(y_1)+H_2(y_2)\}} f(z)dz \\ &= L\left[H_1(y_1) + H_2(y_2)\right]. \end{aligned} \tag{1.27}$$

When Y_1 and Y_2 are dependent, conditionally on Z, the bivariate survival function is

$$S(y_1, y_2|Z) = \exp\left[-ZH(y_1, y_2)\right], \tag{1.28}$$

where $H(y_1, y_2)$ is the bivariate integrated hazard of (Y_1, Y_2). The corresponding unconditional bivariate survival function is

$$\begin{aligned} S(y_1, y_2) &= \int_z e^{-ZH(y_1,y_2)} f(z)dz \\ &= L[H(y_1, y_2)], \end{aligned} \tag{1.29}$$

where L is the Laplace transform of the frailty distribution.

These models assume that the lifetimes are independent conditional on an unobserved covariate or random effect called the "frailty". There is considerable literature on frailty models for lifetime of parallel systems, where it is usually assumed that the failure of some components does not affect the failure *rate* of other operating units (see for example Stefanescu & Turnbull (2012)). Clayton (1978) constructed bivariate distributions with frailty to accommodate certain hereditary disease transmission between parents and children. Lindley & Singpurwalla (1986) study the reliability of two-component systems, where the lifetime of each component follows an exponential baseline distribution shared by a gamma frailty. Hanagal (2011) constructed several bivariate distributions incorporating shared frailty. Inference procedures for frailty models is abundant in literature. We refer to Hougaard (1984) for details.

In the present thesis, Chapter 3 contains a new class of bivariate distribution for load sharing systems when there is dependence between the components induced by some random effect. This model is constructed by applying the frailty approach.

1.5.6 Conditional Specification Method

Let (Y_1, Y_2) be a two dimensional random vector which is absolutely continuous with respect to some product measure $\mu_1 \times \mu_2$ with support $\mathscr{S}(Y_1)$ and $\mathscr{S}(Y_2)$ respectively, which can be finite, countable or uncountable. The distribution function corresponding to Y_1 is F_1 and for Y_2 is F_2. Now, a conditionally specified bivariate distribution is associated with two parametric families of distributions $\mathscr{F}_1 = \{F_1(y_1; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ and $\mathscr{F}_2 = \{F_2(y_2; \boldsymbol{\tau}) : \boldsymbol{\tau} \in \mathscr{T}\}$. The joint distribution (Y_1, Y_2) is required to satisfy the property that for each possible value of y_2 of Y_2, the conditional distribution of $(Y_1|Y_2 = y_2) \in \mathscr{F}_1$ with parameter $\boldsymbol{\theta}$ is possibly dependent on y_2. Also, the conditional distribution of $(Y_2|Y_1 = y_1) \in \mathscr{F}_2$ with parameter $\boldsymbol{\tau}$ is possibly dependent on y_1. Then from the following compatibility theorem we have (Arnold et al. (1999a)),

Theorem 1.5.1. *A joint density $f(y_1, y_2)$, with $f(y_1; \boldsymbol{\theta}(y_2))$ and $f(y_2; \boldsymbol{\tau}(y_1))$ as its conditional densities, will exist if only if*

(i). $N_a = N_b$, and

(ii). there exist functions u and v such that for all y_1, $y_2 \in N_a$,

$$\frac{f(y_1; \boldsymbol{\theta}(y_2))}{f(y_2; \boldsymbol{\tau}(y_1))} = \frac{u(y_1)}{v(y_2)}, \tag{1.30}$$

where

$$\int_{\mathscr{S}(Y_1)} u(y_1) d\mu_1(y_1) < \infty,$$

$$N_a = \{(y_1, y_2) : f(y_1; \boldsymbol{\theta}(y_2)) > 0\},$$

and

$$N_b = \{(y_1, y_2) : f(y_2; \boldsymbol{\tau}(y_1)) > 0\}.$$

Now we are interested in all possible bivariate distributions which have conditionals such that

$$f_{Y_1|Y_2}(y_1|y_2) = f(y_1; \boldsymbol{\theta}(y_2)), \text{ for all } y_1 \in \mathscr{S}(Y_1),\ y_2 \in \mathscr{S}(Y_2), \quad (1.31)$$

and

$$f_{Y_2|Y_1}(y_2|y_1) = f(y_2; \boldsymbol{\tau}(y_1)), \text{ for all } y_1 \in \mathscr{S}(Y_1),\ y_2 \in \mathscr{S}(Y_2). \quad (1.32)$$

For (1.31) and (1.32) to hold there must exist marginal distributions for Y_1 and Y_2 denoted by $f_{Y_1}(y_1)$ and $f_{Y_2}(y_2)$ such that

$$f_{Y_2}(y_2)f(y_1;\, \boldsymbol{\theta}(y_2)) = f_{Y_1}(y_1)f(y_2;\, \boldsymbol{\tau}(y_1)), \text{ for all } y_1 \in \mathscr{S}(Y_1),\ y_2 \in \mathscr{S}(Y_2). \quad (1.33)$$

Then the functional equation (1.33) is solved for $\boldsymbol{\theta}(y_2)$ and $\boldsymbol{\tau}(y_1)$ to obtain a unique bivariate distribution with the specified conditionals. The solution of (1.33) depends on $f(y_1;\, \boldsymbol{\theta}(y_2))$ and $f(y_2;\, \boldsymbol{\tau}(y_1))$. The different challenges in this is discussed in Arnold et al. (1999a).

The pioneering work on construction of bivariate distributions using conditional specifications was given in Patil (1965), where a bivariate power series distribution has been constructed using power series conditional distributions. Besag (1974) discussed conditional specifications in the spatial processes perspective. Castillo & Galambos (1987) brought the importance of functional equations in the study of bivariate distributions with conditionals. Their focus was on normal conditionals. Arnold (1987) studied the bivariate distribution with Pareto conditionals by following the work of Castillo & Galambos (1987). (Arnold & Strauss (1988), Arnold & Strauss (1991)) studied exponential conditionals and also discussed conditionals in prescribed exponential families. Arnold et al. (1999a) developed an algorithm for constructing a bivariate distribution through conditional specification method. Recently, Pulcini (2006) studied bivariate distribution with gamma conditionals.

In Chapter 2, we have investigated bivariate distributions when $f(y_1; \boldsymbol{\theta}(y_2))$ and $f(y_2; \boldsymbol{\tau}(y_1))$ are transmuted distributions. Recently, studying about transmuted distributions, properties and its applications brought a great interest among the researchers. However, most of the works are on univariate transmuted models. Our main goal is to build a new class of bivariate transmuted distributions. Motivated by this task, we initiated a study on a special univariate distribution called "Transmuted

Exponentiated Frêchet distribution". We also studied its properties and some useful applications. The next section of this chapter is fully dedicated to the formulation, properties and applications of this distribution. Some of the results are required for the formulation of bivariate distribution with transmuted conditionals in Chapter 2.

1.6 Transmuted Distributions

The discussions in this section does not confine to lifetime distributions alone. Let $F(x)$ and $G(x)$ be distribution functions with support $\mathscr{S}' \subseteq \mathscr{R}$, the real line. For notations ease we also denote it as F and G sometimes by suppressing the arguments.

1.6.1 Introduction and Literature

A random variable X is said to have transmuted distribution (Shaw & Buckley (2009)) if its cumulative distribution is given by

$$F(x) = (1+\lambda)G(x) - \lambda G(x)^2, \;\; |\lambda| \leq 1, \tag{1.34}$$

where $G(x)$ is the cumulative distribution function of the baseline random variable. Differentiation of (1.34) yields,

$$f(x) = g(x)\left[(1+\lambda) - 2\lambda G(x)\right], \tag{1.35}$$

where $f(x)$ and $g(x)$ are the corresponding probability density functions with cumulative distribution functions $F(x)$ and $G(x)$ respectively. A transmuted random variable X with cumulative distribution function given in (1.34) and probability density function given in (1.35) will be denoted by $X \sim TD(\lambda; G)$. Observe that at $\lambda = 0$ we have the distribution of the base line random variable.

*Some of the results of this section are published in Journal of Statistical Application and Probability (2014) 3(**3**), 379-394. (Elbatal et al. (2014))

1.6.2 Transmutation Map

In this subsection we discuss the transmuted probability distribution. Let F and G be the cumulative distribution functions, with a common support $\mathscr{S}'$. The general rank transmutation as given by Shaw & Buckley (2009) is defined as

$$G_{\mathscr{S}'}(u) = G(F^{-1}(u)).$$

Note that the inverse cumulative distribution function, also known as quantile function, is defined as

$$F^{-1}(u) = \inf_{x\in\mathscr{S}'}\{F(x) \geq u\} \text{ for } u \in [0, 1].$$

The function $G_{\mathscr{S}'}(.)$ maps the unit interval $I = [0, 1]$ into itself, and under suitable assumptions are mutual inverses and they satisfy $G_{\mathscr{S}'}(0) = 0$ and $G_{\mathscr{S}'}(1) = 1$. A quadratic rank transmutation map is defined as

$$G_{\mathscr{S}'}(u) = u + \lambda u(1 - u), \ |\lambda| \leq 1, \ 0 \leq u \leq 1, \tag{1.36}$$

from which it follows that the cumulative distribution functions satisfy the relationship

$$F(x) = (1 + \lambda)G(x) - \lambda G(x)^2. \tag{1.37}$$

When $G(x)$ is absolutely continuous, differentiation of (1.37) yields,

$$f(x) = g(x)[(1 + \lambda) - 2\lambda G(x)], \tag{1.38}$$

where $f(x)$ and $g(x)$ are the corresponding probability density functions associated with cumulative distribution functions $F(x)$ and $G(x)$ respectively. An extensive information about the quadratic rank transmutation map is given in Shaw & Buckley (2009). Observe that at $\lambda = 0$ we have the distribution of the base random variable. The following Lemma prove that the function $f(x)$ given in (1.38) satisfies the property of probability density function.

Lemma 1.6.1. *The funtion $f(.)$ given in* (1.38) *is a well defined probability density function.*

Proof. Rewriting $f(x)$ in (1.38) as $f(x) = g(x)[(1-\lambda(2G(x)-1))]$, we observe that $f(x)$ is non-negative. Also,

$$\begin{aligned}\int_{\mathscr{S}'} f(x)dx &= \int_{\mathscr{S}'} (1+\lambda)g(x)dx - 2\lambda \int_{\mathscr{S}'} g(x)dx \\ &= 1+\lambda-\lambda \\ &= 1\end{aligned}$$

■

Hence $f(x)$ is a well defined probability density function. We call $f(x)$ the transmuted probability density function with base line density $g(x)$. Further properties of this transmuted distribution is studied in Section 1.6.3.

Also many authors have worked with the generalization of some well-known distributions. Aryal & Tsokos (2009) defined the transmuted generalized extreme value distribution and they studied some basic mathematical characteristics of transmuted Gumbel probability distribution and it has been observed that the transmuted Gumbel can be used to model climate data. Also Aryal & Tsokos (2011) presented a new generalization of Weibull distribution, called the transmuted Weibull distribution.

Recently, Aryal (2013) proposed and studied the various structural properties of the transmuted Log-Logistic distribution. Khan & King (2013) introduced the transmuted modified Weibull distribution which extends recent development on transmuted Weibull distribution by Aryal & Tsokos (2011) and they also studied the mathematical properties and maximum likelihood estimation of the unknown parameters. Subsequently, Elbatal & Aryal (2013) presented the transmuted additive Weibull distribution. Also, Elbatal (2013) studied the transmuted modified inverse Weibull distribution. Merovci (2013c) introduced the transmuted Rayleigh distribution, transmuted generalized Rayleigh distribution (Merovci (2013a)), and transmuted Lindley distribution (Merovci (2013b)). Elbatal & Elgarhy (2013) studied the transmuted Quasi Lindley distribution.

Here we make an attempt to establish a generalization for exponentiated Frêchet distribution. Frêchet distribution was introduced by a French mathematician named

Maurice Frêchet (1878-1973) who had identified before one possible limit distribution for the largest order statistic. The Frêchet distribution has been shown to be useful for modelling and analysis of several extreme events ranging from accelerated life testing to modelling earthquakes, floods, rain fall, sea currents and wind speeds. Therefore Frêchet distribution is well suited to characterize random variables of large features. Applications of the Frêchet distribution in various fields given in Harlow (2002) showed that it is an important distribution for modelling the statistical behaviour of materials properties for a variety of engineering applications. Nadarajah & Kotz (2008) discussed the sociological models based on Frêchet random variables. Further, Zaharim et al. (2009) applied Frêchet for analyzing the wind speed data. Mubarak (2012) studied the Frêchet progressive type-II censored data with binomial removals. The Frêchet distribution is a special case of the generalized extreme value distribution. This type-II extreme value distribution (Frêchet) case is equivalent to taking the reciprocal of values from a standard Weibull distribution. The cumulative distribution function and probability density function for Frêchet distribution are given respectively by

$$F(x,\theta,\beta) = e^{-\left(\frac{\theta}{x}\right)^{\beta}}, \; x > 0, \; \theta > 0, \; \beta > 0,$$

where the parameter $\beta > 0$ determines the shape of the distribution and $\theta > 0$ is the scale parameter and

$$f(x,\theta,\beta) = \frac{\beta}{\theta}\left(\frac{\theta}{x}\right)^{\beta+1} e^{-\left(\frac{\theta}{x}\right)^{\beta}}, \; x > 0, \; \theta > 0, \; \beta > 0.$$

Recently, a new three-parameter distribution, named as Exponentiated Frêchet (EF) distribution has been introduced by Nadarajah & Kotz (2003) as a generalization of the standard Frêchet distribution. The exponentiated Frêchet distribution is considered to be one of the newest lifetime models. There are over fifty applications ranging from accelerated life testing, earthquakes, floods, horse racing, rainfall, queues in supermarkets, sea currents, wind speeds and track race records (Kotz & Nadarajah (2000)) etc. The cumulative distribution function of the exponentiated Frêchet distribution is given by

$$G(x,\theta,\beta,\alpha) = 1 - \left[1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right]^{\alpha}, \; x > 0, \; \theta > 0, \; \alpha > 0, \; \beta > 0, \tag{1.39}$$

where α is shape parameter. The corresponding probability density function is given by

$$g(x,\theta,\beta,\alpha) = \alpha\beta\theta^{\beta}x^{-(1+\beta)}\, e^{-\left(\frac{\theta}{x}\right)^{\beta}}\left[1-e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right]^{\alpha-1}, x>0,\ \theta>0,\ \alpha>0,\ \beta>0. \tag{1.40}$$

In this section we present a new generalization of exponentiated Frêchet distribution called the transmuted exponentiated Frêchet (TEF) distribution. We will derive the subject distribution using the quadratic rank transmutation map given in (1.36) (Shaw & Buckley (2009)).

1.6.3 General Properties

In this section we study the properties of the transmuted distribution. Many characteristics of the transmuted distribution function is assured by the behaviour of the baseline distribution function. The next theorem shows the relationship between moments for the transmuted distribution once the baseline moments exist.

Theorem 1.6.1. *Let Φ be a non-degenerate measurable function, and let X be a random variable with transmuted distribution as in* (1.34). *If $E_F(\Phi(X))$ denotes the expectation of $\Phi(X)$, then*

$$E_F(\Phi(X)) = (1+\lambda)E_G(\Phi(X)) - 2\lambda E_G[\Phi(X)G(X)]. \tag{1.41}$$

Proof. From (1.38)

$$\begin{aligned} E_F(\Phi(X)) &= \int \Phi(x)[(1+\lambda)g(x) - 2\lambda g(x)G(x)]dx \\ &= (1+\lambda)E_G(\Phi(X)) - 2\lambda \int \Phi(x)g(x)G(x)dx \\ &= (1+\lambda)E_G(\Phi(X)) - 2\lambda E_G[\Phi(X)G(X)] \end{aligned}$$

∎

Corollary 1.6.1. *If $L_G(t)$ denotes the Laplace transform of the base distribution G,* *then the Laplace transform of the transmuted distribution F is given by*

$$L_F(t) = (1+\lambda)L_G(t) - 2\lambda E_G[e^{-Xt}G(X)]; \ |t| < 1.$$

Corollary 1.6.2. *If $\mu_r(F) = \int x^r f(x)dx$ then $\mu_r(F) = (1+\lambda)\,\mu_r(G) - 2\lambda E_G[X^r G(X)]$.*

Theorem 1.6.2. *For $\lambda > 0$,*
(i). F is a convex distribution function implies that G is also a convex distribution function.
(ii). Conversely, if G is a convex distribution then F is convex if and only if,

$$f(x) \geq \frac{2\lambda g^3(x)}{g'(x)}, \text{ for all } x \in \mathscr{S},$$

where

$$g'(x) = \frac{dg(x)}{dx}.$$

Proof. Let F be a convex distribution function. Then by definition, for $\lambda > 0$, $f'(x) > 0$ which in turn implies

$$\begin{aligned}
f(x) &= g(x)[(1+\lambda) - 2\lambda G(x)] \\
f^{'}(x) &= (1+\lambda)g'(x) - 2\lambda g'(x)G(x) - 2\lambda g^2(x) > 0 \\
&\Leftrightarrow \frac{g'(x)}{g(x)} f(x) > 2\lambda g^2(x) \\
&\Leftrightarrow g'(x) > 2\lambda \frac{g^3(x)}{f(x)} \\
&\Rightarrow g'(x) > 0 \text{ for all } x \in \mathscr{S}.
\end{aligned}$$

Hence proving (i).

To prove (ii) observe that G is a convex distribution function implies $g'(x) > 0$ for all $x \in \mathscr{S}$. Hence from (i) it follows that F is a convex distribution , then for $\lambda > 0$,

$$f^{'}(x) = (1+\lambda)g'(x) - 2\lambda g'(x)G(x) - 2\lambda g^2(x) > 0$$

$$\Leftrightarrow \frac{g'(x)}{g(x)} f(x) > 2\lambda g^2(x)$$
$$\Leftrightarrow f(x) \geq \frac{2\lambda g^3(x)}{g'(x)}, \text{ for all } x \in \mathscr{S}.$$

Hence the result. ■

Remark 1.6.1. *This result holds only for distribution with finite range.*

As a first observation note that (1.34) is a linear mixture of the cumulative distribution functions $G(x)$ and $G^2(x)$. Note that $G^2(x)$ is the cumulative distribution function of maximum of a sample of size two of a random variable with cumulative distribution function G. In consequence, the moments of a transmuted distribution can be expressed as a linear combination of the moments of the random variable with cumulative distribution functions G and G^2. Hence we have,

Theorem 1.6.3. *Let $X' \sim G$ be a random variable with cumulative distribution function $G(\cdot)$ and probability density function $g(\cdot)$. Then, if $X \sim TD(\lambda; G)$,*

$$E(X^r) = (1+\lambda)E(X'^r) - \lambda E(X'^r_{2(2)}), \tag{1.42}$$

where $X'_{2(2)}$ represent the maximum of two independent and identically distributed independent copies of X'.

Proof.

$$F_{X_{2(2)}}(x) = G^2(x) \Rightarrow f_{X_{2(2)}}(x) = 2G(x)g(x)$$
$$\Rightarrow E(X^r_{2(2)}) = \int x^r f_{X_{2(2)}}(x)dx = \int x^r 2g(x)G(x)dx$$

From (1.35) we have

$$\begin{aligned} E(X^r) &= \int x^r f(x)dx \\ &= (1+\lambda)\int x^r g(x)dx - \lambda \int x^r 2g(x)G(x)dx \\ &= (1+\lambda)E(X'^r) - \lambda(EX'^r_{2(2)}) \end{aligned}$$

■

Table 1.1: Cumulative distribution function, r^{th} ordinary moments and r^{th} moments of the maximum of a sample of size two

Distribution	Cumulative distribution function	$E(X'^r)$	$E(X'^r_{2(2)})$
Uniform	$F(x) = x, 0 \leq x \leq 1$	$\frac{1}{r+1}$	$\frac{2}{r+2}$
Exponential	$F(x) = 1 - e^{-\beta x}, x \geq 0$	$\frac{\Gamma(r+1)}{\beta^r}$	$\frac{\Gamma(r+1)}{\beta^r}(2 - 2^{-r})$
Frêchet	$F(x) = e^{\frac{-\beta^\alpha}{x^\alpha}}, x \geq 0$	$\beta^r\Gamma(1 - \frac{r}{\alpha})$	$2^{\frac{r}{\alpha}}\beta^r\Gamma(1 - \frac{r}{\alpha})$

Table 1.1 presents the cumulative distribution function, the r^{th} ordinary moments and the r^{th} moments of the maximum of a sample of size two for a selection of random variables. In this way, we can obtain the moments of the corresponding transmuted distribution by linear combination of these moments using Formula (1.42).

The next few results study the ageing properties of the transmuted distribution $F(x)$ in relation to $G(x)$. One of the characteristic in reliability analysis is the hazard rate function. For an absolutely continuous general transmuted distribution it is defined by

$$h_{TD} = \frac{f}{1-F} = \frac{(1+\lambda)g(x) - 2\lambda G(x)g(x)}{1-(1+\lambda)G(x) + \lambda G(x)^2}. \tag{1.43}$$

The following results are now immediate.

Theorem 1.6.4. *For $\lambda < 0$ ($\lambda > 0$) the transmuted distribution $F(x)$ has an increasing failure rate distribution (decreasing failure rate distribution) if and only if $G(x)$ is an increasing failure rate distribution (decreasing failure rate distribution).*

Proof. From (1.43)

$$h_{TD} = h_G(x)\left[1 + \frac{\lambda\bar{G}(x)}{1-\lambda G(x)}\right], \tag{1.44}$$

where, $h_G(x) = \frac{g(x)}{\bar{G}(x)}$ and $\bar{G}(x) = 1 - G(x)$.

The result follows by observing that $\left[1 + \frac{\lambda\bar{G}(x)}{1-\lambda G(x)}\right]$ is increasing whenever for $\lambda \leq 0$. For $\lambda = 1$, $h_{TD} = 2h_G(x)$. Hence when $\lambda < 0$, $h_{TD} > h_G(x)$ which implies that if the baseline distribution G has an increasing failure rate distribution then F has a increasing failure rate distribution. Similarly, for $\lambda > 0$, $h_{TD} < h_G(x)$, which implies that if the baseline distribution G has a decreasing failure rate distribution then F has a decreasing failure rate distribution.

Hence the result. ■

Hence it is evident that in general the transmuted distribution functions do not behave in a similar manner as the base distribution. Hence it is of interest to study the transmuted distributions on specifying different baseline distributions. Motivated by this we study a particular transmuted distribution by taking the baseline distribution $G(x)$ to be exponentiated Frêchet distribution in (1.37). This distribution is discussed elaborately in rest of this chapter.

1.7 Transmuted Exponentiated Frêchet Distribution

In this section we introduce the transmuted exponentiated Frêchet (TEF) distribution. Substituting (1.39) into (1.37) we have the cumulative distribution function of transmuted exponentiated Frêchet (TEF) distribution

$$F_{TEF}(x,\theta,\beta,\alpha,\lambda) = \left[1-\left(1-e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right]\left[1+\lambda\left(1-e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right], \qquad (1.45)$$

$x > 0,\ \theta > 0,\ \alpha > 0,\ \beta > 0,\ |\lambda| \leq 1$, where λ is the transmuted parameter. The corresponding probability density function is given by

$$\begin{aligned} f_{TEF}(x,\theta,\beta,\alpha,\lambda) = {} & \alpha\beta\theta^{\beta}x^{-(1+\beta)}\, e^{-\left(\frac{\theta}{x}\right)^{\beta}} \times (q(x))^{\alpha-1} \\ & \times \left[(1-\lambda) + 2\lambda\,(q(x))^{\alpha}\right],\ x > 0,\ \theta > 0,\ \alpha > 0,\ \beta > 0,\ |\lambda| \leq 1, \end{aligned}$$

where, $q(x) = \left(1-e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)$.

Now f_{TEF} can be written as

$$f_{TEF}(x,\theta,\beta,\alpha,\lambda) = (1-\lambda)\alpha\beta\theta^{\beta}x^{-(1+\beta)}\, e^{-\left(\frac{\theta}{x}\right)^{\beta}}\, (q(x))^{\alpha-1}$$

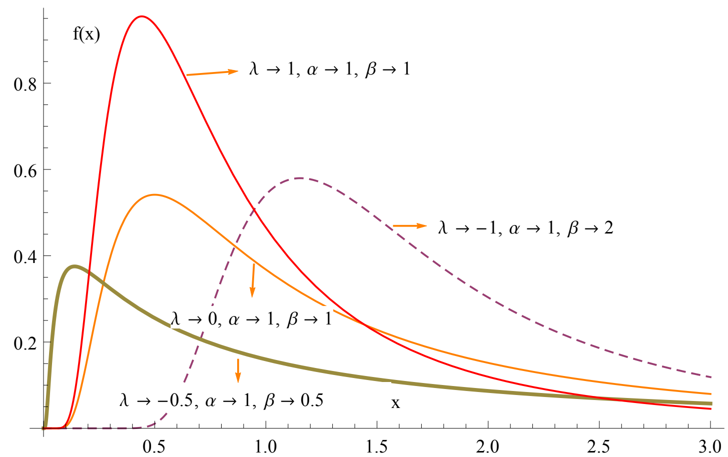


Figure 1.1: Probability density function of Transmuted Exponentiated Frêchet distribution for $\theta = 1$ and different values of λ, α and β

.

$$+2\lambda\alpha\beta\theta^{\beta}x^{-(1+\beta)}\,e^{-\left(\frac{\theta}{x}\right)^{\beta}}(q(x))^{2\alpha-1}. \tag{1.46}$$

It is observed that the transmuted exponentiated Frêchet distribution is an extended model to analyse data and it generalizes some of the widely used distributions. For instance, when $\beta = 1$ it reduces to transmuted exponentiated inverted exponential distribution as discussed in Elbatal (2014). The exponentiated Frêchet distribution is clearly a special case for $\lambda = 0$.

When $\beta = \lambda = 1$ and $\alpha = 0.5$ then the resulting distribution is an inverted exponential distribution with parameter θ (see Abouammoh & Alshingiti (2009)). Figure 1.1 illustrates some of the possible shapes of the probability density function of a transmuted exponentiated Frêchet distribution for selected values of the parameters α, β, λ and for $\theta = 1$. Figure 1.2 explains the effect of varying λ for fixed values of other parameters.

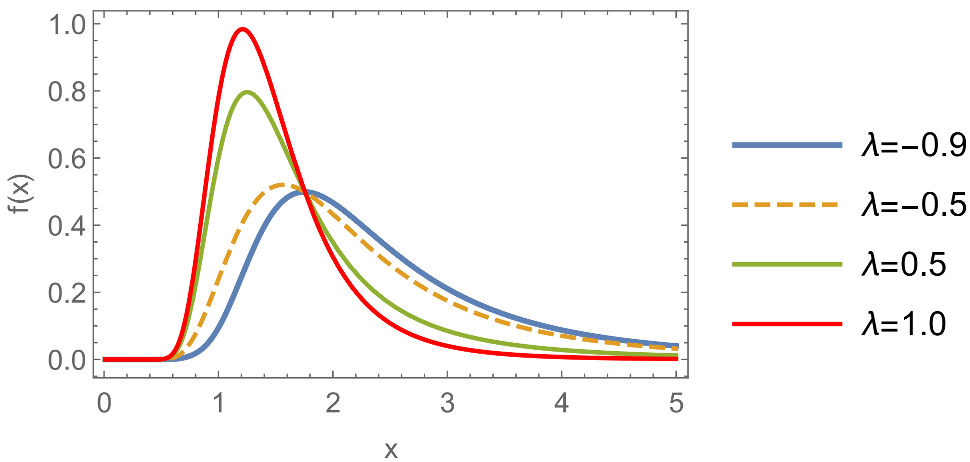


Figure 1.2: Probability density function of Transmuted Exponentiated Frêchet distribution for $\theta = 1, \alpha = 1$ and $\beta = 2$ and different values of λ

.

1.7.1 Statistical Properties of the Transmuted Exponentiated Frêchet Distribution

This section is devoted to studying statistical properties of the TEF distribution, more specifically quantile function, moments and moment generating function.

1.7.1.1 Quantile Function and Random Number Generation

Let X be a random variable with cumulative distribution function F, we can define the quantile function $Q(u) = \inf\{x : F(x) \geq u\}$ for $u \in (0, 1)$. In particular, if F is continuous and strictly increasing then we have $Q(u) = F^{-1}(u)$. We present a method for simulating from the TEF distribution (1.45).

Theorem 1.7.1. *The quantile of the TEF distribution is given by*

$$\begin{aligned} Q(u) &= F^{-1}(u) \\ &= \theta \left\{ -\ln\left[1 - \left(\frac{(\lambda-1) + \sqrt{(\lambda+1)^2 - 4\lambda u}}{2\lambda}\right)^{\frac{1}{\alpha}}\right]\right\}^{\frac{-1}{\beta}} . \end{aligned} \tag{1.47}$$

Let U be a uniform variate on the unit interval (0,1). Thus, by means of the inverse *transformation method, the random variable X is given by*

$$X_u = \theta \left\{ -\ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda+1)^2 - 4\lambda u}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}}, 0 < u < 1. \tag{1.48}$$

Proof. The quantile X_u of the TEF distribution is defined as

$$u = P(X_u \leq x_u) = F(x_u),\ x_u \geq 0.$$

Using the cumulative distribution function of the TEF distribution we have

$$u = F(x_u) = (1+\lambda)\left[1 - \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right] - \lambda \left[1 - \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right]^2, \tag{1.49}$$

that is

$$\lambda \left[1 - \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right]^2 - (1+\lambda)\left[1 - \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right] + u = 0. \tag{1.50}$$

Consider (1.50) as a quadratic equation in $1 - \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}$ as

$$\Delta = 1 + (2 - 4u)\lambda + \lambda^2, \tag{1.51}$$

where Δ is the discriminant of the quadratic equation. The quadratic equation given in (1.50) has roots $\frac{(1+\lambda)\pm\sqrt{\Delta}}{2\lambda}$. These roots exist if Δ is positive.
Now, consider the following cases

- If $\lambda = -1$ then $\Delta = 4u > 0,\ \ u > 0.$
- If $\lambda = 1$ then $\Delta = 4(1-u) > 0,\quad u > 0.$
- Otherwise for $-1 < \lambda < 1$, consider the roots of Δ, as a quadratic form in λ, are

$$\lambda = (2u - 1) \pm 2\sqrt{u^2 - u}.$$

Therefore, $u^2 - u < 0$ for $0 < u < 1$. So the only real roots could occur for $u = 0$ or 1.

- If $u = 0$ then $\lambda = -1$, which contradicts the fact that $-1 < \lambda < 1$.
- If $u = 1$ then $\lambda = 1$, which again contradicts the fact that $-1 < \lambda < 1$.

Thus, there are no real roots of Δ as a quadratic in λ. Therefore, Δ has the same sign in the range $-1 \leq \lambda \leq 1$, hence $\Delta > 0$.
Since $\Delta \geq 0$, then

$$\begin{aligned} 1 - \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha} &= \frac{(1+\lambda) - \sqrt{\Delta}}{2\lambda} \\ \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right) &= \left(1 - \frac{(1+\lambda) - \sqrt{\Delta}}{2\lambda}\right)^{\frac{1}{\alpha}} \\ e^{-\left(\frac{\theta}{x}\right)^{\beta}} &= \left\{1 - \left(1 - \frac{(1+\lambda) - \sqrt{\Delta}}{2\lambda}\right)^{\frac{1}{\alpha}}\right\}. \end{aligned}$$

Finally, we obtain the quantile X_u of the TEF distribution as

$$X_u = \theta \left\{ - \ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda+1)^2 - 4\lambda u}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}}, 0 < u < 1.$$

Hence the proof. ■

1.7.1.2 Skewness and Kurtosis

The Skewness, of Bowley (Kenney (2013)) is defined by

$$S_K = \frac{Q_{0.75} - 2Q_{0.5} + Q_{0.25}}{Q_{0.75} - Q_{0.25}},$$

and Kurtosis (see Moors (1988)) based on octiles is defined by

$$K_u = \frac{Q_{0.875} - Q_{0.625} - Q_{0.375} + Q_{0.125}}{Q_{0.75} - Q_{0.25}}.$$

where,

$$Q_{0.125} = \theta \left\{ -\ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda + 1)^2 - \frac{\lambda}{2}}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}},$$

$$Q_{0.25} = \theta \left\{ -\ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda + 1)^2 - \lambda}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}},$$

$$Q_{0.375} = \theta \left\{ -\ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda + 1)^2 - \frac{3\lambda}{2}}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}},$$

$$Q_{0.5} = \theta \left\{ -\ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda + 1)^2 - 2\lambda}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}},$$

$$Q_{0.625} = \theta \left\{ -\ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda + 1)^2 - \frac{5\lambda}{2}}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}},$$

$$Q_{0.75} = \theta \left\{ -\ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda + 1)^2 - 3\lambda}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}},$$

$$Q_{0.875} = \theta \left\{ -\ln \left[1 - \left(\frac{(\lambda - 1) + \sqrt{(\lambda + 1)^2 - \frac{7\lambda}{2}}}{2\lambda} \right)^{\frac{1}{\alpha}} \right] \right\}^{\frac{-1}{\beta}}.$$

1.7.1.3 Moments

Theorem 1.7.2. *If X follows Transmuted Exponentiated Frechet ($TEF(\theta, \beta, \alpha, \lambda)$), where $\theta > 0,\ \alpha > 0,\ \beta > 0,\ |\lambda| \leq 1$, and λ is the transmuted parameter, then the r^{th} moment of X (μ_r') is given by the following*

$$\mu_r' = \sum_{j=0}^{\infty}(-1)^j \theta^r (1+j)^{-\left(1-\frac{r}{\beta}\right)} \Gamma(1-\frac{r}{\beta}) \tag{1.52}$$
$$\left[(1-\lambda)\binom{\alpha-1}{j} + 2\lambda\binom{2\alpha-1}{j}\right].$$

Proof. Let X be a random variable with density function (1.46). The r^{th} ordinary moment of the TEF distribution is given by

$$\mu_r' = E(X^r) = \int_0^\infty x^r f(x)dx,$$

$$\begin{aligned}\mu_r' =&(1-\lambda)\alpha\beta\theta^\beta \int_0^\infty x^{r-\beta-1}\, e^{-\left(\frac{\theta}{x}\right)^\beta} \left[1-e^{-\left(\frac{\theta}{x}\right)^\beta}\right]^{\alpha-1} dx \\ &+ 2\lambda\alpha\beta\theta^\beta \int_0^\infty x^{r-\beta-1}\, e^{-\left(\frac{\theta}{x}\right)^\beta} \left[1-e^{-\left(\frac{\theta}{x}\right)^\beta}\right]^{2\alpha-1} dx.\end{aligned} \tag{1.53}$$

Setting

$$\left[1-e^{-\left(\frac{\theta}{x}\right)^\beta}\right]^{\alpha-1} = \sum_{j=0}^{\infty}(-1)^j \binom{\alpha-1}{j} e^{-j\left(\frac{\theta}{x}\right)^\beta}, \tag{1.54}$$

and substituting from (1.54) into (1.53) we get

$$\begin{aligned}\mu_r' =&(1-\lambda)\sum_{j=0}^{\infty}(-1)^j \binom{\alpha-1}{j} \alpha\beta\theta^\beta \int_0^\infty x^{r-\beta-1}\, e^{-(j+1)\left(\frac{\theta}{x}\right)^\beta} dx \\ &+ 2\lambda\sum_{j=0}^{\infty}(-1)^j \binom{2\alpha-1}{j} \alpha\beta\theta^\beta \int_0^\infty x^{r-\beta-1}\, e^{-(j+1)\left(\frac{\theta}{x}\right)^\beta} dx.\end{aligned} \tag{1.55}$$

let $(j+1)\left(\frac{\theta}{x}\right)^\beta = t$, we get

$$\mu'_r = \sum_{j=0}^{\infty} (-1)^j \theta^r (1+j)^{-\left(1-\frac{r}{\beta}\right)} \Gamma(1-\frac{r}{\beta}) \times K(j), \tag{1.56}$$

where,

$$K(j) = \left[(1-\lambda) \binom{\alpha-1}{j} + 2\lambda \binom{2\alpha-1}{j} \right].$$

This completes the proof. ■

Property 1.7.1. *The moment equation given in* (1.56) *is a convergent series by Alternating Series Test. Therefore all the moments exist for transmuted exponentiated Frêchet distribution.*

Proof. The moment equation in (1.56) can be written as

$$\begin{aligned} \mu'_r = & \sum_{j=0}^{\infty} (-1)^j \theta^r (1+j)^{-\left(1-\frac{r}{\beta}\right)} \Gamma(1-\frac{r}{\beta})(1-\lambda) \binom{\alpha-1}{j} \\ & + \sum_{j=0}^{\infty} (-1)^j \theta^r (1+j)^{-\left(1-\frac{r}{\beta}\right)} \Gamma(1-\frac{r}{\beta})(2\lambda) \binom{2\alpha-1}{j}. \end{aligned} \tag{1.57}$$

Let

$$a_j = (1+j)^{-\left(1-\frac{r}{\beta}\right)} \binom{\alpha-1}{j}$$

and

$$\lim_{j->\infty} a_j = \lim_{j->\infty} \frac{1}{(1+j)^{\left(1-\frac{r}{\beta}\right)}} \frac{(\alpha-1)!}{j!(\alpha-j-1)!} = 0.$$

Similarly, let

$$b_j = (1+j)^{-\left(1-\frac{r}{\beta}\right)} \binom{2\alpha-1}{j}$$

and

$$\lim_{j->\infty} b_j = \lim_{j->\infty} \frac{1}{(1+j)^{\left(1-\frac{r}{\beta}\right)}} \frac{(2\alpha-1)!}{j!(2\alpha-j-1)!} = 0.$$

For $j = 0$, $a_j = 1$ and $b_j = 1$. Now (1.57) can be rewritten as

$$\mu'_r = (1-\lambda)\theta^r\Gamma(1-\frac{r}{\beta})\sum_{j=1}^{\infty}(-1)^j a_j + 2\lambda\theta^r\Gamma(1-\frac{r}{\beta})\sum_{j=1}^{\infty}(-1)^j b_j. \tag{1.58}$$

Also, it is observed that $a_j > a_{j+1}$ and $b_j > b_{j+1}$. Thus by alternating series test (1.57) is a convergent series. Hence the moments exist. ■

1.7.1.4 Moment Generating Function

In this subsection we derived the moment generating function of TEF distribution.

Theorem 1.7.3. *If X has TEF distribution, then the moment generating function $M_X(t)$ has the following form.*

$$M_X(t) = \sum_{r=0}^{\infty}\sum_{j=0}^{\infty}\frac{t^r}{r!}(-1)^j\theta^r(1+j)^{-(1-\frac{r}{\beta})}\Gamma(1-\frac{r}{\beta}) \times K(j), \tag{1.59}$$

where,

$$K(j) = \left[(1-\lambda)\binom{\alpha-1}{j} + 2\lambda\binom{2\alpha-1}{j}\right].$$

Proof. We start with the well known definition of the moment generating function given by

$$\begin{aligned}
M_X(t) &= \int_0^{\infty} e^{tx} f_{TEF}(x)dx \\
&= \sum_{r=0}^{\infty}\frac{t^r}{r!}x^r f_{TEF}(x)dx = \sum_{r=0}^{\infty}\frac{t^r}{r!}\mu'_r \\
&= \sum_{r=0}^{\infty}\sum_{j=0}^{\infty}\frac{t^r}{r!}(-1)^j\theta^r(1+j)^{-(1-\frac{r}{\beta})}\Gamma(1-\frac{r}{\beta}) \\
&\times \left[(1-\lambda)\binom{\alpha-1}{j} + 2\lambda\binom{2\alpha-1}{j}\right],
\end{aligned} \tag{1.60}$$

which completes the proof. ■

Property 1.7.2. *The moment generating function in* (1.59) *is convergent.*

Proof. The moment generating function in (1.59) can be written as $\sum_{r=1}^{\infty}\sum_{j=1}^{\infty}(-1)^j b(j,r)$, where

$$\begin{aligned} b(j,r) &= \frac{t^r}{r!}(-1)^j \theta^r (1+j)^{-(1-\frac{r}{\beta})}\Gamma(1-\frac{r}{\beta}) \\ &\times \left[(1-\lambda)\binom{\alpha-1}{j} + 2\lambda\binom{2\alpha-1}{j}\right], \end{aligned}$$

and

$$\lim_{j,r->\infty} b(j,r) = 0.$$

Hence by double series test (1.59) is a convergent series. ■

1.7.1.5 Distribution of the Order Statistics

In fact, the order statistics have many applications in reliability and life testing. The order statistics arise in the study of reliability of a system. Let X_1, $X_2, ..., X_n$ be a simple random sample from $TEF(\theta, \beta, \alpha, \lambda, x)$ with cumulative distribution function and probability density function as in (1.45) and (1.46), respectively. Let $X_{1(n)} \leq X_{2(n)} \leq ... \leq X_{n(n)}$ denotes; the order statistics obtained from this sample. Note that, in reliability literature, $X_{i(n)}$ denotes; the lifetime of an $(n-i+1)-$ out$-$ of$-$ n system irrespective of whether they are iid or not. Then the probability density function of $X_{i(n)}$, $1 \leq i \leq n$ is given by

$$f_{i(n)}(x) = \frac{1}{\beta(i, n-i+1)}\left[F(x,\boldsymbol{\tau})\right]^{i-1}\left[1-F(x,\boldsymbol{\tau})\right]^{n-i} f(x,\boldsymbol{\tau}), \tag{1.61}$$

where $\boldsymbol{\tau} = (\alpha, \beta, \theta, \lambda)$. Also, the joint probability density function of $X_{i(n)}$ and $X_{j(n)}$, $1 \leq i \leq j \leq n$ is given by

$$\begin{aligned} f_{i(j)(n)}(x_i, x_j) =& C\ \left[F(x_i)\right]^{i-1}\left[F(x_j)-F(x_i)\right]^{j-i-1} \\ & \times \left[1-F(x_j)\right]^{n-j} f(x_i) f(x_j), \end{aligned} \tag{1.62}$$

where

$$C = \frac{n!}{(i-1)!(j-i-1)!(n-j)!}$$

We defined the first order statistic $X_{1(n)} = \text{Min}(X_1, X_2, ..., X_n)$, the last order statistic as $X_{n(n)} = \text{Max}(X_1, X_2, ..., X_n)$ and the median is $X_{(n+1)/2}$.

1.7.1.6 Distribution of Minimum, Maximum and Median

Let $X_1,\ X_2, ...,\ X_n$ be independently identically distributed order random variables from the transmuted Exponentiated Frêchet distribution having first, last and median of the probability density function are given by the following

$$\begin{aligned} f_{1(n)}(x) =& n\left[1 - F(x, \boldsymbol{\tau})\right]^{n-1} f(x, \boldsymbol{\tau}) \\ =& n\left\{1 - \left[1 - h_{(1)}^{\alpha}\right]\left[1 + \lambda h_{(1)}^{\alpha}\right]\right\}^{n-1} \\ &\times\ \alpha\beta\theta^{\beta} x_{(1)}^{-(1+\beta)}(1 - h_{(1)}) h_{(1)}^{\alpha-1} \\ &\times\left[(1-\lambda) + 2\lambda h_{(1)}^{\alpha}\right], \end{aligned} \tag{1.63}$$

$$\begin{aligned} f_{n(n)}(x) =& n\left[F(x_{(n)}, \boldsymbol{\tau})\right]^{n-1} f(x_{(n)}, \boldsymbol{\tau}) \\ =& n\left\{\left[1 - h_{(n)}^{\alpha}\right]\left[1 + \lambda h_{(n)}^{\alpha}\right]\right\}^{n-1} \\ &\times\ \alpha\beta\theta^{\beta} x_{(n)}^{-(1+\beta)}(1 - h_{(n)}) h_{(n)}^{\alpha-1}\left[(1-\lambda) + 2\lambda h_{(n)}^{\alpha}\right], \end{aligned} \tag{1.64}$$

and

$$\begin{aligned} f_{\frac{n+1}{2}(n)}(\widetilde{x}) =& \frac{(2m+1)!}{m!m!}(F(\widetilde{x}))^m (1 - F(\widetilde{x}))^m f(\widetilde{x}) \\ =& \frac{(2m+1)!}{m!m!}\left\{\left[1 - h_{(m+1)}^{\alpha}\right]\left[1 + \lambda h_{(m+1)}^{\alpha}\right]\right\}^m \\ &\times\left\{1 - \left[1 - h_{(m+1)}^{\alpha}\right]\left[1 + \lambda h_{(m+1)}^{\alpha}\right]\right\}^m \\ &\times\ \alpha\beta\theta^{\beta} x_{(m+1)}^{-(1+\beta)}(1 - h_{(m+1)}) h_{(m+1)}^{\alpha-1} \\ &\times\left[(1-\lambda) + 2\lambda h_{(m+1)}^{\alpha}\right], \end{aligned} \tag{1.65}$$

where $h_{(s)} = \left(1 - e^{-\left(\frac{\theta}{x_{(s)}}\right)^{\beta}}\right)$ and $m = \frac{n+1}{2}$.

Joint Distribution of the i^{th} and j^{th} Order Statistics

The joint distribution of the i^{th} and j^{th} order statistics from TEF distribution is

$$\begin{aligned}
f_{i(j)(n)}(x_i, x_j) =& C \ [F(x_i)]^{i-1} [F(x_j) - F(x_i)]^{j-i-1} \\
& \times [1 - F(x_j)]^{n-j} f(x_i) f(x_j) \\
=& C \left\{ \left[1 - h^{\alpha}_{(i)}\right] \left[1 + \lambda h^{\alpha}_{(i)}\right] \right\}^{i-1} \\
& \times \left\{ \left[1 - h^{\alpha}_{(j)}\right] \left[1 + \lambda h^{\alpha}_{(j)}\right] - \left[1 - h^{\alpha}_{(i)}\right] \left[1 + \lambda h^{\alpha}_{(i)}\right] \right\}^{j-i-1} \\
& \times \left\{ 1 - \left[1 - h^{\alpha}_{(j)}\right] \left[1 + \lambda h^{\alpha}_{(j)}\right] \right\}^{n-j} \\
& \times \alpha\beta\theta^{\beta} x_{(i)}^{-(1+\beta)} (1 - h_{(i)}) h^{\alpha-1}_{(i)} \left[(1-\lambda) + 2\lambda h^{\alpha}_{(i)}\right] \\
& \times \alpha\beta\theta^{\beta} x_{(j)}^{-(1+\beta)} (1 - h_{(j)}) h^{\alpha-1}_{(j)} \left[(1-\lambda) + 2\lambda h^{\alpha}_{(j)}\right],
\end{aligned} \tag{1.66}$$

where

$$h_{(s)} = \left(1 - e^{-\left(\frac{\theta}{x_{(s)}}\right)^{\beta}}\right).$$

1.7.1.7 Reliability Characteristics

The reliability function or the survival function of the TEF distribution is defined as

$$\begin{aligned}
S(x) =& 1 - F(x) \\
=& 1 - \left[1 - \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right] \left[1 + \lambda \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right], \ |\lambda| \leq 1, \beta > 0, \alpha > 0.
\end{aligned} \tag{1.67}$$

For the TEF distribution the hazard function defined in (1.43) is given by

$$h(x) = \frac{f(x)}{S(x)}$$

$$h(x) = \frac{A(x;\boldsymbol{\tau}) \times B(x;\boldsymbol{\tau})}{C(x;\boldsymbol{\tau})}, \tag{1.68}$$

where,

$$\boldsymbol{\tau} = (\alpha, \beta, \theta, \lambda),$$

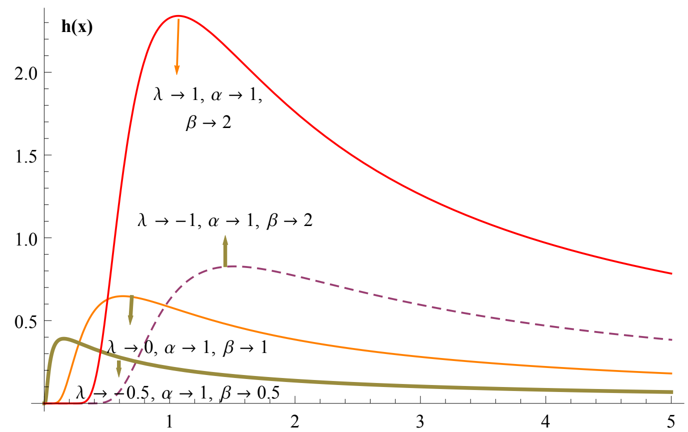


Figure 1.3: Hazard rate function of Transmuted Exponentiated Frêchet distribution for $\theta = 1$ and different values of λ, α and β

.

$$A(x;\boldsymbol{\tau}) = \alpha\beta\theta^{\beta}x^{-(1+\beta)}e^{-\left(\frac{\theta}{x}\right)^{\beta}}\left[1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right]^{\alpha-1},$$

$$B(x;\boldsymbol{\tau}) = \left[(1-\lambda) + 2\lambda\left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right],$$

and

$$C(x;\boldsymbol{\tau}) = 1 - \left[1 - \left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right] \left[1 + \lambda\left(1 - e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right].$$

The transmuted exponentiated Frêchet distribution belongs to the class of distributions that admits upside down bathtub curves. increasing and decreasing hazard rate. In Figure 1.3, some choices of λ, β, α are given for which the TEF exhibits a upside down bathtub (UBT) hazard rate. Conditions under which they are IFR and DFR is discussed in the following theorems.

Theorem 1.7.4. *If $\alpha = \theta = \lambda = 1$, then the failure rate is increasing if $\beta < 0$ and is decreasing if $\beta > 0$.*

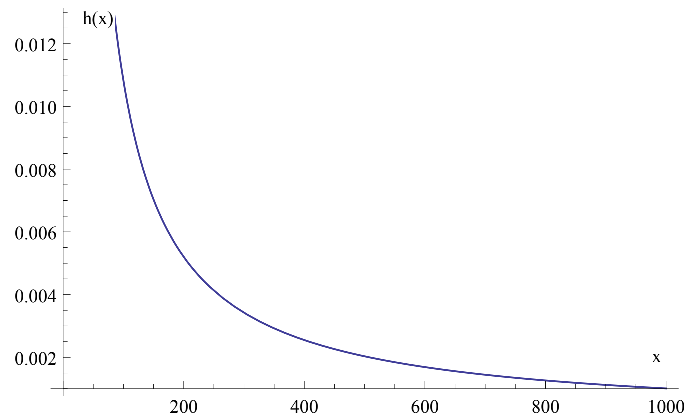


Figure 1.4: Hazard rate function of Transmuted Exponentiated Frêchet distribution for $\alpha = \beta = 1$

.

Proof. If $\alpha = \theta = \lambda = 1$ then

$$h(x) = \frac{2\beta \left(\frac{1}{x}\right)^{1+\beta}}{\left(e^{\left(\frac{1}{x}\right)^{\beta}} - 1\right)},$$

which is increasing for $\beta < 0$ and is decreasing for $\beta > 0$.

∎

Theorem 1.7.5. *If $\beta = \alpha = 1$ then the failure rate is monotonically decreasing for both $\lambda < 0$ and $\lambda > 0$.*

Proof. If $\beta = \alpha = 1$ then we have (Figure 1.4)

$$h(x) = \frac{\theta}{x^2} \left[\frac{1}{\left(e^{\left(\frac{\theta}{x}\right)} - 1\right)} + \frac{\lambda}{\left(e^{\left(\frac{\theta}{x}\right)} - \lambda\right)} \right].$$

It can be easily verified that $h(x)$ is decreasing for both $\lambda < 0$ and $\lambda > 0$. Note that

$$h(0) = \infty \quad and \quad h(\infty) = 0$$

(see Figure 1.4). ∎

The cumulative hazard function of the transmuted Exponentiated Frêchet distribution is denoted by $H(x)$ and is defined as

$$H(x) = -\ln\left|\left[1-\left(1-e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right]\left[1+\lambda\left(1-e^{-\left(\frac{\theta}{x}\right)^{\beta}}\right)^{\alpha}\right]\right|. \tag{1.69}$$

Similar to the hazard rate function, we can also illustrate the behaviour of the cumulative hazard rate function for different choices of parameters.

1.7.1.8 Estimation and Inference

In this section, we determine the maximum likelihood estimates ($MLEs$) of the parameters of the TEF distribution from complete samples only. Let $X_1, X_2, ..., X_n$ be a random sample of size n from $TEF(\theta,\beta,\alpha,\lambda)$ and the log-likelihood is given by

$$\begin{aligned}\ell &= n\log\alpha + n\log\beta + n\beta\log\theta - (1+\beta)\sum_{i=1}^{n}\log x_i - \theta^{\beta}\sum_{i=1}^{n}x_i^{-\beta}\\ &+ (\alpha-1)\sum_{i=1}^{n}\log\left[1-e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right] + \sum_{i=1}^{n}\log\left[(1-\lambda)+2\lambda\left(1-e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha}\right]. \end{aligned}\tag{1.70}$$

The log-likelihood can be maximized either directly or by solving the non-linear likelihood equations obtained by differentiating (1.70). The components of the score vector are given by

$$\begin{aligned}\frac{\partial\ell}{\partial\theta} &= \frac{n\beta}{\theta} - \beta\theta^{\beta-1}\sum_{i=1}^{n}x_i^{-\beta} + \beta(\alpha-1)\sum_{i=1}^{n}\frac{e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\left(\frac{\theta}{x_i}\right)^{\beta-1}}{x_i\left[1-e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right]}\\ &+ \sum_{i=1}^{n}\frac{2\lambda\beta\alpha\left(1-e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha-1}e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\left(\frac{\theta}{x_i}\right)^{\beta-1}}{x_i\left[(1-\lambda)+2\lambda\left(1-e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha}\right]}, \end{aligned}\tag{1.71}$$

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \log\left[1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right]$$
$$+ \sum_{i=1}^{n} \frac{2\lambda\left(1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha} \log\left(1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)}{\left[(1-\lambda) + 2\lambda\left(1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha}\right]}, \tag{1.72}$$

$$\frac{\partial \ell}{\partial \beta} = \frac{n}{\beta} + n\log\theta - \sum_{i=1}^{n} \log x_i - \sum_{i=1}^{n} \log\left(\frac{\theta}{x_i}\right)\left(\frac{\theta}{x_i}\right)^{\beta}$$
$$+(\alpha-1)\sum_{i=1}^{n} \frac{\log\left(\frac{\theta}{x_i}\right) e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\left(\frac{\theta}{x_i}\right)^{\beta}}{\left[1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right]}$$
$$+2\lambda\alpha\sum_{i=1}^{n} \frac{Q(x_i)}{\left[(1-\lambda) + 2\lambda\left(1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha}\right]}, \tag{1.73}$$

where,

$$Q(x_i) = \left(1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha-1} \log\left(\frac{\theta}{x_i}\right) e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\left(\frac{\theta}{x_i}\right)^{\beta},$$

and

$$\frac{\partial \ell}{\partial \lambda} = \sum_{i=1}^{n} \frac{2\left(1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha} - 1}{\left[(1-\lambda) + 2\lambda\left(1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right)^{\alpha}\right]} = 0. \tag{1.74}$$

We describe an effective profile likelihood approach for the model in (1.46) by maximizing the likelihood. The log-likelihood equations are presented in (1.71) to (1.74). The likelihood equations are very difficult to solve and it may be tedious to obtain maximum likelihood estimators (MLE) by Newton-Raphson procedure. We propose the following estimation method. Let $\tilde{\boldsymbol{\lambda}} = (\tilde{\boldsymbol{\tau}}_1, \tilde{\boldsymbol{\tau}}_2)$ where $\tilde{\boldsymbol{\tau}}_1 = (\alpha, \theta)$, $\tilde{\boldsymbol{\tau}}_2 = (\lambda, \beta)$. In the first stage, we estimate $\tilde{\boldsymbol{\tau}}_1$ by maximizing the profile likelihood of $\tilde{\boldsymbol{\tau}}_1$ and once, an estimate of $\tilde{\boldsymbol{\tau}}_1$ is obtained, the estimates of $\tilde{\boldsymbol{\tau}}_2$ can be obtained by substituting the estimates of $\tilde{\boldsymbol{\tau}}_1$. This process is continued iteratively till all the estimates converge to yield the MLE $\hat{\tilde{\boldsymbol{\lambda}}}$ of $\tilde{\boldsymbol{\lambda}}$. The computation is carried out using "FindMaximum" function of Mathematica 10. Section 1.7.1.9 presents a detailed simulation study to illustrate the estimation approach. Also, all the second order derivatives exist. Thus we have the inverse dispersion matrix given by

$$\begin{pmatrix} \widehat{\theta} \\ \widehat{\beta} \\ \widehat{\alpha} \\ \widehat{\lambda} \end{pmatrix} \sim N \left[\begin{pmatrix} \theta \\ \beta \\ \alpha \\ \lambda \end{pmatrix}, \begin{pmatrix} \widehat{V_{\theta\theta}} & \widehat{V_{\theta\beta}} & \widehat{V_{\theta\alpha}} & \widehat{V_{\theta\lambda}} \\ \widehat{V_{\beta\theta}} & \widehat{V_{\beta\beta}} & \widehat{V_{\beta\alpha}} & \widehat{V_{\beta\lambda}} \\ \widehat{V_{\alpha\theta}} & \widehat{V_{\alpha\beta}} & \widehat{V_{\alpha\alpha}} & \widehat{V_{\alpha\lambda}} \\ \widehat{V_{\lambda\theta}} & \widehat{V_{\lambda\beta}} & \widehat{V_{\lambda\alpha}} & \widehat{V_{\lambda\lambda}} \end{pmatrix} \right]. \tag{1.75}$$

Under the conditions that are fulfilled for parameters in the interior of the parameter space, but not on the boundary, the asymptotic distribution of the element of the 4 x 4 observed information matrix for the TEF distribution is $\sqrt{n}(\hat{\tilde{\boldsymbol{\lambda}}} - \tilde{\boldsymbol{\lambda}}) \sim N_4(0, V^{-1})$, where V is the expected information matrix. Thus, the expected information matrix is

$$V^{-1} = -E \begin{bmatrix} V_{\theta\theta} & V_{\theta\beta} & V_{\theta\alpha} & V_{\theta\lambda} \\ & V_{\beta\beta} & V_{\beta\alpha} & V_{\beta\lambda} \\ & & V_{\alpha\alpha} & V_{\alpha\lambda} \\ & & & V_{\lambda\lambda} \end{bmatrix},$$

where

$$V_{\lambda\lambda} = \frac{\partial^2 \ell}{\partial \lambda^2} = \sum_{i=1}^{n} \frac{\left(2\left(q(x_i)\right)^{\alpha} - 1\right)^2}{\left(2\lambda\left(q(x_i)\right)^{\alpha} - \lambda + 1\right)^2}, \tag{1.76}$$

$$V_{\alpha\alpha} = \frac{\partial^2 \ell}{\partial \alpha^2} = \frac{n}{\alpha^2} - \sum_{i=1}^{n} \left(\frac{2\lambda\left(q(x_i)\right)^{\alpha} \log\left(q(x_i)\right)^2}{2\lambda\left(q(x_i)\right)^{\alpha} - \lambda + 1} - \frac{4\lambda^2\left(q(x_i)\right)^{2\alpha} \log\left(q(x_i)\right)^2}{\left(2\lambda\left(q(x_i)\right)^{\alpha} - \lambda + 1\right)^2} \right), \tag{1.77}$$

$$\begin{aligned} V_{\theta\lambda} &= \frac{\partial^2 \ell}{\partial \theta \partial \lambda} \\ &= -\sum_{i=1}^{n} \left(\frac{2\alpha\beta\left(q(x_i)\right)^{\alpha-1} e^{-\left(\frac{\theta}{x_i}\right)^{\beta}} \left(\frac{\theta}{x_i}\right)^{\beta-1}}{2x_i\lambda\left(q(x_i)\right)^{\alpha} - \lambda + 1} \right) \\ &+ \sum_{i=1}^{n} \left(\frac{2\alpha\beta\lambda\left(q(x_i)\right)^{\alpha-1}\left(2\left(q(x_i)\right)^{\alpha} - 1\right) e^{-\left(\frac{\theta}{x_i}\right)^{\beta}} \left(\frac{\theta}{x_i}\right)^{\beta-1}}{x_i\left(2\lambda\left(q(x_i)\right)^{\alpha} - \lambda + 1\right)^2} \right), \end{aligned} \tag{1.78}$$

where,

$$q(x_i) = \left(1 - e^{-\left(\frac{\theta}{x_i}\right)^{\beta}}\right).$$

Similarly, we can find the other components as

$$V_{\theta\theta} = \frac{\partial^2 \ell}{\partial \theta^2},\ V_{\beta\beta} = \frac{\partial^2 \ell}{\partial \beta^2},\ V_{\alpha\theta} = \frac{\partial^2 \ell}{\partial \alpha \partial \theta}, V_{\beta\lambda} = \frac{\partial^2 \ell}{\partial \beta \partial \lambda}$$
$$V_{\alpha\beta} = \frac{\partial^2 \ell}{\partial \alpha \partial \beta},\ V_{\theta\beta} = \frac{\partial^2 \ell}{\partial \theta \partial \beta},\ V_{\alpha\lambda} = \frac{\partial^2 \ell}{\partial \alpha \partial \lambda},$$

By solving this inverse dispersion matrix these solutions will yield asymptotic variance and covariances of these maximum likelihood (ML) estimators for $\widehat{\lambda}, \widehat{\theta}, \widehat{\alpha}$ and $\widehat{\beta}$. Using (1.75), we approximate $100(1-\gamma)\%$ confidence intervals for λ, β, θ and α, and are determined respectively as

$$\widehat{\theta} \pm z_{\frac{\gamma}{2}} \sqrt{\widehat{V_{\theta\theta}}},\ \widehat{\beta} \pm z_{\frac{\gamma}{2}} \sqrt{\widehat{V_{\beta\beta}}},\ \widehat{\alpha} \pm z_{\frac{\gamma}{2}} \sqrt{\widehat{V_{\alpha\alpha}}}\ and\ \widehat{\lambda} \pm z_{\frac{\gamma}{2}} \sqrt{\widehat{V_{\lambda\lambda}}},$$

where z_{γ} is the upper 100_{γ} the percentile of the standard normal distribution.

Simulation Study

We carried out a simulation study in order to evaluate the performance of the profile likelihood estimation. Sample generation of (x_i), $i = 1, 2, \ldots, n$ was carried out by using the algorithm given in Section 1.7.1.1. We generated 1000 samples of sizes $n = 25$, $n = 75$, and $n = 150$ with true values $\theta = 1.25$, $\alpha = 0.75$, $\beta = 1.5$, and $\lambda = 0.6$. The MLEs were obtained using the procedure described in Section 1.7.1 and the average bias across the 1000 samples was computed. The average root mean square error (RMSE) from the 1000 samples was calculated as $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{\tilde{\boldsymbol{\lambda}}}_i - \tilde{\boldsymbol{\lambda}}_i)}$. The approximate variance-covariance matrix of the MLEs was obtained as the inverse of the observed information matrix as given in (1.75). The absolute biases, RMSEs and coverage probabilities(C.P) for the confidence intervals are provided in Table 1.2. We observed from the simulation study that the biases, RMSEs decrease and the coverage probability improves as the sample size increases. We also observed that the rate of convergence improved with increasing sample size. The graph 1.5 shows the stability graph for parameter estimates after 15 iterations. The plots for the profile likelihood estimates for different parameters support that there is a global maximum for each parameter (See).

Table 1.2: Absolute Bias, RMSE and Coverage probability (C.P) of α, θ, β and λ based on 1000 replications for transmuted exponentiated Frêchet distribution obtained using maximum likelihood method

Parameters	α	θ	β	λ
True Values	1.0	1.25	1.5	0.6
		n=25		
Absolute Bias	0.0536	0.0412	0.0578	0.0183
RMSE	0.0040	0.0025	0.0049	0.0050
C.P	0.8962	0.8064	0.8596	0.8335
		n=75		
Absolute Bias	0.0452	0.0362	0.0358	0.0141
RMSE	0.0028	0.0020	0.0018	0.0032
C.P	0.9041	0.8785	0.8683	0.9078
		n=150		
Absolute Bias	0.0365	0.0133	0.0326	0.0130
RMSE	0.0019	0.0003	0.0015	0.0026
C.P	0.9421	0.9132	0.9013	0.9318

The plots for the profile likelihood estimates for different parameters support that there is a global maximum for each parameter (See Figure 1.6).

1.8 Organisation of Thesis

In this chapter we discussed some of the basic concepts which are necessary for the present thesis. Several methods of constructing bivariate distributions are presented. Conditional specification method and frailty approach are of our special interest. In order to satisfy our intension of developing a bivariate distributions with transmuted conditionals an example of transmuted distributions, the transmuted exponentiated Frêchet (TEF) distribution was studied in detail. These results formed the basis for formulation of a general bivariate class in Chapter 2.

The profile likelihood method of estimation is used to estimate the parameters involved. The reliability behaviour of the subject distribution is studied. A simula-

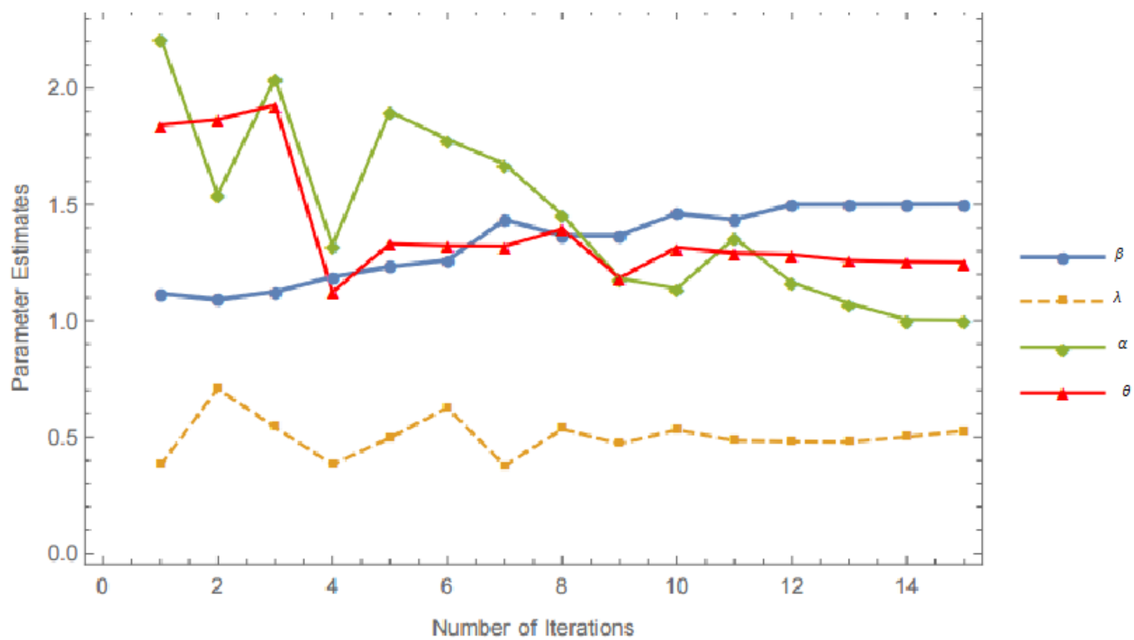


Figure 1.5: The stability graph for the simulated parameters estimates of method of maximum likelihood for transmuted exponentiated Frêchet distribution for 15 iterations

.

tion study is conducted to show the effectiveness of our estimation procedure. The study also reveals that as the sample size increases the biases and RMSEs decreases considerably.

We are hoping to show that the transmuted distributions are a rich class of distribution and is useful to give a more flexible model. It is also helpful in analysing bivariate data. Accordingly, in Chapter 2 an attempt is made to introduce bivariate transmuted distribution. We proposed a general class of bivariate transmuted distribution with transmuted conditionals. We studied the general properties of this model and provided examples with different baseline distributions. Profile likelihood method has been used to estimate the parameters involved. A simulation study has been conducted to show the appropriateness and the effectiveness of our estimation procedure. Two data sets have been analysed which are already published in the literature and showed the superiority of our model. It may be noted that results in Chapter 2 are not restricted to lifetime data though the examples are restricted to lifetime data sets.

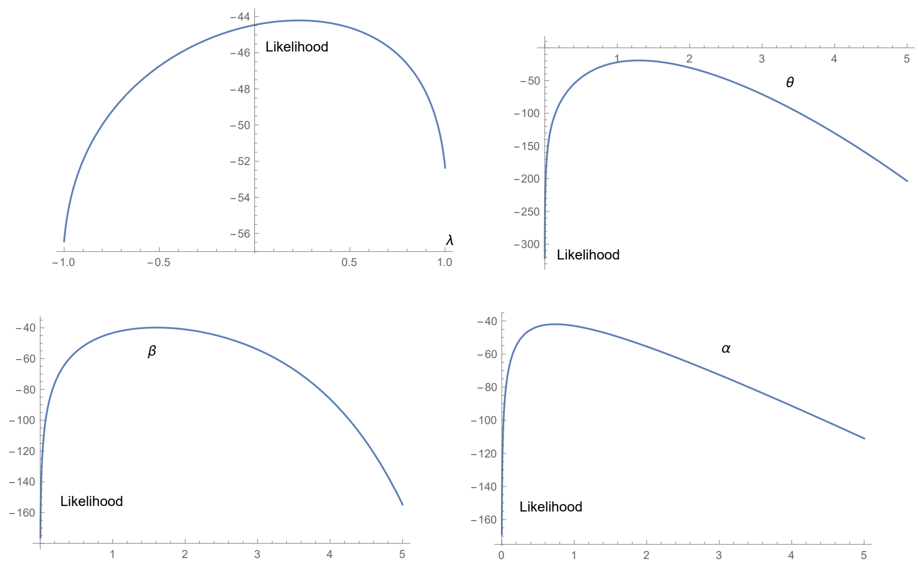


Figure 1.6: Plot for profile likelihood estimate (a) λ when ($\alpha = 1.05,\ \theta = 1.29,\ \beta = 1.56$) (b) θ when ($\alpha = 1.05,\ \lambda = 0.62,\ \beta = 1.56$) (c) β when ($\alpha = 1.05,\ \theta = 1.29,\ \lambda = 0.62$)(d) α when ($\lambda = 0.62,\ \theta = 1.29,\ \beta = 1.56$).

In the reliability and survival context, it is always important to study the dependence structure between components. In particular, a load sharing effect. In a load sharing system, the probability of failure of any component will depend on the working status of the other components (Kvam & Lu (2007)). There are many situations in practice where the failure of a unit could redistribute the workload of the other operating units in the system, thus potentially increasing the failure rate of the operating units. The basic assumption in a two-component load sharing system is that while the system can function even after one of the components has failed, the failure of the component may put additional load on the surviving component and this affects the functioning of the system due to stochastic changes in its residual lifetime. In most situations, an increased load results in a higher component failure rate (Liu (1998)). Examples of such systems include (a) a twin engine aircraft like Boeing's 777 (Singpurwalla (1995)), or (b) mechanical systems ("Reliability in Engineering Design" (n.d.)), or (c) paired organs like eyes, kidneys or lungs (Daniels (1945)), to mention a few.

Freund (1961) constructed a bivariate exponential distribution suitable for load sharing framework. Recently, Asha et al. (2016) provided a general class of bivariate distribution to model load sharing systems.

Another flexible tool for modelling dependent times to failure is the “frailty model”. Frailty models have been widely used to study dependent lifetimes in reliability and survival analysis framework (Clayton (1978)) and Hougaard (1984)). These models assume that the lifetimes are independent conditional on an unobserved covariate or random effect, called the “frailty”. There is considerable literature on frailty models for lifetime of parallel systems, where it is usually assumed that the failure of some components does not affect the failure *rate* of other operating units. Inference procedures for frailty models is abundant in literature. We refer to Hougaard (2000) for details.

However, there are many situations where it is physically meaningful to incorporate the dependence induced both by the frailty and *the dependence due to load sharing* in studying lifetimes of a multi-component system.

Accordingly in Chapter 3, a general class of bivariate distributions for load sharing models with frailty and covariates have been introduced. We study some general properties and provide examples for the model. A general estimation procedure is discussed and general method of generating bivariate samples has been explained. The use of α- stable as frailty distribution was introduced by Hougaard (1986). More references and applications for α- stable as frailty distribution are found in Hougaard (2000), Singpurwalla (1995), Wassell et al. (1999), Ravishanker & Dey (2000).

In Chapter 4, a particular example assuming positive stable frailty and Weibull baseline in the general model presented in Chapter 3 is studied. Profile likelihood method is applied to estimate the unknown parameters and a simulation study is conducted to show the effectiveness of our estimation procedure. Two data sets are analysed, first one is the motor data (ReliaSoft (2003)) which has no covariates and no censoring. Second one is the well analysed diabetic rrtinopathy study (DRS) data (Huster et al. (1989)) with one covariate and censoring.

Sometimes in load share models there is a critical time or threshold time for change in parameter. During this critical time the failure of a part in a system can trigger the failure of successive parts. Such a failure may happen in many types of systems, including power transmission, computer networking, finance, human bodily systems, bridges and so on. Modelling these type of failures involving in a system are known as cascading models. These type of failures are referred to as cascading failures (Lindley & Singpurwalla (2002)).

In chapter 5, we propose a general class of bivariate distribution for cascading failures by extending the works of Lindley & Singpurwalla (2002), Swift (2008) and Asha et al. (2016) using Cox total failure rate (Cox (1972)). We studied the model extensively with a special example by considering exponential baseline. Method of moments and L-moments have been used to estimate the unknown parameters. Simulation study was conducted to show the effectiveness of our estimation procedures and provided evidences for the better performances of L-moments. A real life data set been analysed to show our model applicability.

Finally, the present thesis concludes in Chapter 6 with a detailed discussion on results and conclusions. Some of the possible future works are listed in this chapter. We have proposed a new class of bivariate distribution for discrete load share models. Examples for the proposed model are constructed by considering different baseline distributions such as geometric, discrete Weibull, S distribution and Waring distribution. General properties like the joint survival function, the marginal survival function of the proposed model is presented. General estimation procedures are explained. We consider the simulation study and estimation of the unknown parameters as immediate future work. The present thesis ends with some open problems which will be taken up as future works.

Chapter 2

A Class of Bivariate Distributions with Transmuted Conditionals

2.1 Introduction

As seen in Chapter 1, modelling with transmuted distributions has gained lot of interest among the researchers and academicians in the past two decades. A lot of focus has been given on the univariate transmuted distributions where it is treated as a flexible model, which not only accommodates additional skewness in the data but also is a rich class of distributions because of the additional parameter that it incorporates. On the bivariate note, bivariate and multivariate generalizations of the transmuted distributions have been proposed by Bourguignon et al. (2016). Other than this, not much work is reported in this area. Motivated by this, we have made an attempt to establish a new class of bivariate distributions with transmuted conditionals. The construction of the proposed distribution is by method of conditional specification, discussed in detail in Section 1.5.6.

The contents of this Chapter are the following. In Section 2.2 we obtain the most general bivariate distribution with transmuted conditionals. In Section 2.3 we study the basic properties of the model, including marginal and conditional distributions.

*Some of the results of this Chapter are communicated for publication.

The generation of the random samples, cross-moments, dependence measures of the general model are discussed in Section 2.4. Five specific models namely transmuted uniform conditionals, transmuted normal conditionals, transmuted exponential conditionals, transmuted Weibull conditionals and transmuted exponentiated Frêchet conditionals are discussed in Section 2.5. Estimation methodologies such as method of moments and method of maximum likelihood are presented in Section 2.6. A simulation study is conducted in Section 2.7 for illustrating both method of moments estimation and method of maximum likelihood estimation. In Section 2.8 we have considered two well analysed data sets for our model application. The first one is the cable insulation failure time data. It was originally published and studied by Stone (1978) and further analysed by Lawless (2011). Recently, Pulcini (2006) analysed this data set in the context of forewarning or primer event using bivariate distribution with gamma conditionals. Second, we considered a data set consisting of two components parallel systems. This was originally published and analysed by Murthy et al. (2004). Pulcini (2006) further analysed this data set in the context of forewarning or primer event using bivariate distribution with gamma conditionals. Both the above data sets were analysed using our model and in comparison with the model proposed by Pulcini (2006), it is established that our model is a better fit. Finally, discussion and summary are presented in Section 2.9.

2.2 Bivariate Distributions with Transmuted Conditionals

In this section we present the bivariate distribution with transmuted conditionals.

Let (Y_1, Y_2) be an absolutely continuous bivariate random vector with distribution function $F(y_1, y_2)$ and support $\mathscr{R}^2$, where $\mathscr{R}$, the real line. We want to consider all possible joint distributions for (Y_1, Y_2) with the following properties:

- For each y_2 the conditional distribution of Y_1 given $Y_2 = y_2$ is distributed according to (1.34), with parameter $\lambda_1(y_2)$, which may depend on y_2, where the baseline cumulative distribution function $G_1(.)$ is fixed and does not depend on y_2.
- For each y_1 the conditional distribution of Y_2 given $Y_1 = y_1$ is distributed

according to (1.34), with parameter $\lambda_2(y_1)$, which may depend on y_1, and the baseline cumulative distribution function $G_2(.)$ is fixed and does not depend on y_1.

In consequence, we seek the most general random variable (Y_1, Y_2) such that the conditional distributions admit the stochastic representation:

$$Y_1|Y_2 = y_2 \sim TD(\lambda_1(y_2); G_1), \tag{2.1}$$

$$Y_2|Y_1 = y_1 \sim TD(\lambda_2(y_1); G_2), \tag{2.2}$$

for $y_1 \in \mathscr{R}$ and $y_2 \in \mathscr{R}$, where $\lambda_1(y_2)$, $\lambda_2(y_1)$, are unknown functions and $G_i(.), i = 1, 2$, are the baseline cumulative distribution functions. In the next theorem we obtain the most general model satisfying (2.1) and (2.2).

Theorem 2.2.1. *The most general bivariate joint probability density function with conditional distributions* (2.1) *and* (2.2) *is given by,*

$$f(y_1, y_2; \boldsymbol{\lambda}) = k(\boldsymbol{\lambda}) \left[1 + 2\lambda_{10}G_1(y_1) + 2\lambda_{01}G_2(y_2) + 4\lambda_{11}G_1(y_1)G_2(y_2)\right] g_1(y_1)g_2(y_2), \tag{2.3}$$

where, for $\boldsymbol{\lambda} = (\lambda_{10}, \lambda_{01}, \lambda_{11})$, *for constants* $\lambda_{10}, \lambda_{01}, \lambda_{11} \in \mathscr{S}'$, *where* $\mathscr{S}' \subseteq \mathscr{R}$ *such that*

$$\lambda_1(y_2) = -\frac{\lambda_{10} + 2\lambda_{11}G_2(y_2)}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11})G_2(y_2)},$$

$$\lambda_2(y_1) = -\frac{\lambda_{01} + 2\lambda_{11}G_1(y_1)}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11})G_1(y_1)},$$

and

$$k(\boldsymbol{\lambda}) = \frac{1}{1 + \lambda_{10} + \lambda_{01} + \lambda_{11}}, \tag{2.4}$$

$$|\lambda_{10} + \lambda_{11}| \leq |1 + \lambda_{10} + \lambda_{01} + \lambda_{11}|. \tag{2.5}$$

Proof. From conditions (2.1) and (2.2) the conditional densities are given by,

$$f_{Y_1|Y_2}(y_1|y_2) = (1 + \lambda_1(y_2))g_1(y_1) - 2\lambda_1(y_2)g_1(y_1)G_1(y_1), \tag{2.6}$$

$$f_{Y_2|Y_1}(y_2|y_1) = (1 + \lambda_2(y_1))g_2(y_2) - 2\lambda_2(y_1)g_2(y_2)G_2(y_2), \tag{2.7}$$

where $\lambda_1(y_2)$ and $\lambda_2(y_1)$ are unknown functions. Now, if we write the joint probability density function as product of marginals $f_{Y_i}(y_i)$, $i = 1, 2$, by conditional densities in both senses, we obtain the functional equation from the fact that

$$f_{Y_1|Y_2}(y_1|y_2)f_{Y_2}(y_2) = f_{Y_2|Y_1}(y_2|y_1)f_{Y_1}(y_1),$$

as,

$$\begin{aligned}&(1+\lambda_1(y_2))g_1(y_1)f_{Y_2}(y_2) - 2\lambda_1(y_2)g_1(y_1)G_1(y_1)f_{Y_2}(y_2) = \\ &(1+\lambda_2(y_1))g_2(y_2)f_{Y_1}(y_1) - 2\lambda_2(y_1)g_2(y_2)G_2(y_2)f_{Y_1}(y_1),\end{aligned} \tag{2.8}$$

or

$$\begin{aligned}&\tilde{\lambda}_1(y_2)g_1(y_1)f_{Y_2}(y_2) - 2\lambda_1(y_2)g_1(y_1)G_1(y_1)f_{Y_2}(y_2) \\ &-\left[\tilde{\lambda}_2(y_1)g_2(y_2)f_{Y_1}(y_1) - 2\lambda_2(y_1)g_2(y_2)G_2(y_2)f_{Y_1}(y_1)\right] = 0,\end{aligned} \tag{2.9}$$

where $\tilde{\lambda}_1(y_2) = 1 + \lambda_1(y_2)$, $\tilde{\lambda}_2(y_1) = 1 + \lambda_2(y_1)$.

Observe that the functional equation (2.9) is of the form $\sum_{k=1}^{4} p_k(y_1)q_k(y_2) = 0$, where,

$$p_1(y_1) = g_1(y_1),\ p_2(y_1) = g_1(y_1)G_1(y_1),\ p_3(y_1) = \tilde{\lambda}_2(y_1)f_{Y_1}(y_1),\ p_4(y_1) = \lambda_2(y_1)f_{Y_1}(y_1)$$

and

$$q_1(y_2) = \tilde{\lambda}_1(y_2)f_{Y_2}(y_2),\ q_2(y_2) = -2\lambda_1(y_2)f_{Y_2}(y_2),$$

$$q_3(y_2) = -g_2(y_2) \text{ and } q_4(y_2) = 2g_2(y_2)G_2(y_2).$$

Hence from Theorem 1.3 in page 13 of Arnold et al. (1999b), solution for (2.9) can be written in the form

$$\begin{pmatrix} g_1(y_1) \\ g_1(y_1)G_1(y_1) \\ \tilde{\lambda}_2(y_1)f_{Y_1}(y_1) \\ \lambda_2(y_1)f_{Y_1}(y_1) \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} g_1(y_1) \\ g_1(y_1)G_1(y_1) \end{bmatrix}, \tag{2.10}$$

$$\begin{pmatrix} \tilde{\lambda}_1(y_2)f_{Y_2}(y_2) \\ -2\lambda_1(y_2)f_{Y_2}(y_2) \\ -g_2(y_2) \\ 2g_2(y_2)G_2(y_2) \end{pmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} g_2(y_2) \\ 2g_2(y_2)G_2(y_2) \end{bmatrix}, \tag{2.11}$$

since the system of equations $\{g_1(y_1), g_1(y_1)G_1(y_1)\}$ are mutually linearly independent and again $\{g_2(y_2),\, g_2(y_2)G_2(y_2)\}$ are mutually linearly independent. Then the constants a_{ij} and b_{ij}, $i = j = 1, 2$ satisfy

$$\begin{bmatrix} 1 & 0 & a_{11} & a_{21} \\ 0 & 1 & a_{12} & a_{22} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ -1 & 0 \\ 0 & 1 \end{bmatrix} = 0. \tag{2.12}$$

Now, by solving the equation (2.12) we obtain the solution as $b_{11} = a_{11}$, $b_{12} = -a_{21}$, $b_{21} = a_{12}$ and $b_{22} = -a_{22}$. Therefore we can rewrite the system of equations in (2.10) and (2.11) as

$$\begin{aligned} \tilde{\lambda}_2(y_1)f_{Y_1}(y_1) &= a_{11}g_1(y_1) + a_{12}g_1(y_1)G_1(y_1) \\ \lambda_2(y_1)f_{Y_1}(y_1) &= a_{21}g_1(y_1) + a_{22}g_1(y_1)G_1(y_1) \\ \tilde{\lambda}_1(y_2)f_{Y_2}(y_2) &= a_{11}g_2(y_2) - 2a_{21}g_2(y_2)G_2(y_2) \\ -2\lambda_1(y_2)f_{Y_2}(y_2) &= a_{12}g_2(y_2) - 2a_{22}g_2(y_2)G_2(y_2). \end{aligned} \tag{2.13}$$

Now by taking the ratio of first two equations in (2.13) we get

$$\frac{\tilde{\lambda}_2(y_1)}{\lambda_2(y_1)} = \frac{1 + \lambda_2(y_1)}{\lambda_2(y_1)} = \frac{a_{11} + a_{12}G_1(y_1)}{a_{21} + a_{22}G_1(y_1)}, \tag{2.14}$$

which implies that

$$\lambda_2(y_1) = \frac{a_{21} + a_{22}G_1(y_1)}{a_{11} - a_{21} + (a_{12} - a_{22})G_1(y_1)}. \tag{2.15}$$

Substituting (2.15) into the second equation of (2.13) we get

$$f_{Y_1}(y_1) = \frac{(a_{21} + a_{22}G_1(y_1))g_1(y_1)}{a_{21} + a_{22}G_1(y_1)} \times (a_{11} - a_{21} + (a_{12} - a_{22})G_1(y_1))$$

$$= (a_{11} - a_{21} + (a_{12} - a_{22})G_1(y_1))\, g_1(y_1). \tag{2.16}$$

Similarly, by taking the ratio of the third and the fourth equation of (2.13) we get

$$\frac{\tilde{\lambda}_1(y_2)}{\lambda_1(y_2)} = \frac{1 + \lambda_1(y_2)}{-2\lambda_1(y_2)} = \frac{a_{11} - 2a_{21}G_2(y_2)}{a_{12} - 2a_{22}G_2(y_2)}, \tag{2.17}$$

which implies that

$$\lambda_1(y_2) = \frac{-a_{12} + 2a_{22}G_2(y_2)}{2a_{11} + a_{12} + (4a_{21} - a_{22})G_2(y_2)}. \tag{2.18}$$

Substituting (2.18) into the fourth equation of (2.13) we get

$$\begin{aligned} f_{Y_2}(y_2) &= \frac{a_{12} - 2a_{22}G_2(y_2)g_2(y_2)}{-2\lambda_1(y_2)} \\ &= \frac{(a_{12} - 2a_{22}G_2(y_2))g_2(y_2)}{a_{12} - a_{22}G_2(y_2)} \times (2a_{11} + a_{12} + (4a_{21} - 2a_{22})G_2(y_2)) \\ &= \left(\frac{2a_{11} + a_{12}}{2} + (2a_{21} - a_{22})G_2(y_2)\right) g_2(y_2). \end{aligned} \tag{2.19}$$

Now,

$$1 + \lambda_2(y_1) = \frac{a_{11} + a_{12}G_1(y_1)}{a_{11} - a_{21} + (a_{12} - a_{22})G_1(y_1)}. \tag{2.20}$$

Substituting (2.20) into (2.8) and using the fact that $f(y_1, y_2) = f_{Y_1|Y_2}(y_1|y_2)f_{Y_2}(y_2)$, we get

$$\begin{aligned} f(y_1, y_2) &= \left\{\frac{a_{11} + a_{12}G_1(y_1)}{a_{11} - a_{21} + (a_{12} - a_{22})G_1(y_1)} - \frac{2a_{21} - 2a_{21}G_1(y_1)}{a_{11} - a_{21} + (a_{12} - a_{22})G_1(y_1)}G_2(y_2)\right\} \\ &\times g_2(y_2)\,(a_{11} - a_{21} + (a_{12} - a_{22})G_1(y_1))\, g_1(y_1) \\ &= [a_{11} + a_{12}G_1(y_1) - 2a_{21}G_2(y_2) + 2a_{22}G_1(y_1)G_2(y_2)]\, g_1(y_1)g_2(y_2);\ a_{11} \neq 0. \end{aligned} \tag{2.21}$$

For $a_{11} \neq 0$, dividing (2.21) by a_{11} throughout, we get

$$f(y_1, y_2) = a_{11}\left[1 + \frac{a_{12}}{a_{11}}G_1(y_1) - \frac{2a_{21}}{a_{11}}G_2(y_2) + \frac{2a_{22}}{a_{11}}G_1(y_1)G_2(y_2)\right] g_1(y_1)g_2(y_2). \tag{2.22}$$

By re-parametrising $\frac{a_{12}}{a_{11}} = 2\lambda_{10}$, $\frac{-2a_{21}}{a_{11}} = 2\lambda_{01}$, and $\frac{2a_{22}}{a_{11}} = 4\lambda_{11}$ we get

$$f(y_1, y_2; \boldsymbol{\lambda}) = k(\boldsymbol{\lambda}) \left[1 + 2\lambda_{10}G_1(y_1) + 2\lambda_{01}G_2(y_2) + 4\lambda_{11}G_1(y_1)G_2(y_2)\right] g_1(y_1)g_2(y_2), \tag{2.23}$$

From $\int_{y_2}\int_{y_1} f(y_1, y_2; \boldsymbol{\lambda})dy_1dy_2 = 1$, we obtain the value of the normalizing constant $k(\boldsymbol{\lambda})$, as follows,

$$k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} \left[1 + 2\lambda_{10}G_1(y_1) + 2\lambda_{01}G_2(y_2) + 4\lambda_{11}G_1(y_1)G_2(y_2)\right] g_1(y_1)g_2(y_2)dy_1dy_2 = 1, \tag{2.24}$$

or equivalently,

$$\begin{aligned} &= k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} g_1(y_1)g_2(y_2)dy_1dy_2 \\ &+ 2\lambda_{10}k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} g_1(y_1)g_2(y_2)G_1(y_1)dy_1dy_2 \\ &+ 2\lambda_{01}k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} g_1(y_1)g_2(y_2)G_2(y_2)dy_1dy_2 \\ &+ 4\lambda_{11}k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} g_1(y_1)g_2(y_2)G_1(y_1)G_2(y_2)dy_1dy_2 = 1. \end{aligned} \tag{2.25}$$

Now by applying the fact that, $2\int_x g(x)G(x)dx = 1$, in (2.25) we get

$$k(\boldsymbol{\lambda}) \left[1 + \lambda_{10} + \lambda_{01} + \lambda_{11}\right] = 1,$$

to obtain

$$k(\boldsymbol{\lambda}) = \frac{1}{1 + \lambda_{10} + \lambda_{01} + \lambda_{11}}. \tag{2.26}$$

Now from (2.15) and (2.18) it follows that

$$\lambda_1(y_2) = -\frac{\lambda_{10} + 2\lambda_{11}G_2(y_2)}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11})G_2(y_2)},$$

$$\lambda_2(y_1) = -\frac{\lambda_{01} + 2\lambda_{11}G_1(y_1)}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11})G_1(y_1)}.$$

To retrieve the condition on the parameters the marginal density function $f_{Y_1}(y_1)$ is obtained as

$$
\begin{aligned}
f_{Y_1}(y_1) &= k(\boldsymbol{\lambda}) \int_{y_2} \left[1 + 2\lambda_{10}G_1(y_1) + 2\lambda_{01}G_2(y_2) + 4\lambda_{11}G_1(y_1)G_2(y_2)\right] g_1(y_1)g_2(y_2)dy_2 \\
&= k(\boldsymbol{\lambda}) \int_{y_2} g_1(y_1)g_2(y_2)dy_2 + 2\lambda_{10}k(\boldsymbol{\lambda}) \int_{y_2} g_1(y_1)G_1(y_1)g_2(y_2)dy_2 \\
&+ 2\lambda_{01}k(\boldsymbol{\lambda}) \int_{y_2} g_1(y_1)g_2(y_2)G_2(y_2)dy_2 + 4\lambda_{11}k(\boldsymbol{\lambda}) \int_{y_2} g_1(y_1)g_2(y_2)G_1(y_1)G_2(y_2)dy_2.
\end{aligned}
$$

Again, by applying the fact that $2\int_x g(x)G(x)dx = 1$, we get

$$
\begin{aligned}
f_{Y_1}(y_1) &= k(\boldsymbol{\lambda}) \left[g_1(y_1) + 2\lambda_{10}g_1(y_1)G_1(y_1) + 2\lambda_{01}g_1(y_1) + 2\lambda_{11}g_1(y_1)G_1y_1\right] \\
&= k(\boldsymbol{\lambda})g_1(y_1) \left[1 + \lambda_{01} + 2\lambda_{10}G_1(y_1) + 2\lambda_{11}G_1(y_1)\right] \\
&= k(\boldsymbol{\lambda}) \left[1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11})G_1(y_1)\right] g_1(y_1). \qquad (2.27)
\end{aligned}
$$

Hence the marginal density in (2.27) can be written as

$$
f_{Y_1}(y_1) = \left(1 - \frac{(\lambda_{10} + \lambda_{11})}{(1 + \lambda_{01} + \lambda_{10} + \lambda_{11})}\right) g_1(y_1) - \left(\frac{-2(\lambda_{10} + \lambda_{11})}{1 + \lambda_{01} + \lambda_{10} + \lambda_{11}}\right) g_1(y_1)G_1(y_1). \qquad (2.28)
$$

Similarly, the marginal density $f_{Y_2}(y_2)$ is obtained as

$$
f_{Y_2}(y_2) = \left(1 - \frac{(\lambda_{01} + \lambda_{11})}{(1 + \lambda_{01} + \lambda_{10} + \lambda_{11})}\right) g_2(y_2) - \left(\frac{-2(\lambda_{01} + \lambda_{11})}{1 + \lambda_{01} + \lambda_{10} + \lambda_{11}}\right) g_2(y_2)G_2(y_2). \qquad (2.29)
$$

Once again note that marginal distributions $f_{Y_i}(y_i),\ i = 1, 2$ are transmuted with the stochastic representation as,

$$
Y_1 \sim TD\left(\lambda_1' = \frac{-(\lambda_{10} + \lambda_{11})}{1 + \lambda_{10} + \lambda_{01} + \lambda_{11}}; G_1\right), \qquad (2.30)
$$

and

$$
Y_2 \sim TD\left(\lambda_2' = \frac{-(\lambda_{01} + \lambda_{11})}{1 + \lambda_{01} + \lambda_{10} + \lambda_{11}}; G_2\right). \qquad (2.31)
$$

It now follows from (2.30) and (2.31) that

$$
|\lambda_{10} + \lambda_{11}| \leq |1 + \lambda_{10} + \lambda_{01} + \lambda_{11}| \qquad (2.32)
$$

and

$$|\lambda_{01} + \lambda_{11}| \leq |1 + \lambda_{10} + \lambda_{01} + \lambda_{11}|, \tag{2.33}$$

there by establishing condition (2.5). ■

Remark 2.2.1. *A bivariate distribution with transmuted conditionals and joint probability density function* (2.23) *will be denoted by* $(Y_1, Y_2) \sim BTC(\boldsymbol{\lambda}; G_1, G_2)$.

2.3 General Properties

In this section we study the basic properties of the bivariate random variable with joint probability density function (2.23).

Property 2.3.1. *The marginal density of* (Y_1, Y_2) *with* $BTC(\boldsymbol{\lambda}; G_1, G_2)$ *is given by*

$$f_{Y_1}(y_1) = k(\boldsymbol{\lambda})\left[1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11})G_1(y_1)\right] g_1(y_1),$$

and

$$f_{Y_2}(y_2) = k(\boldsymbol{\lambda})\left[1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11})G_2(y_2)\right] g_2(y_2).$$

Property 2.3.2. *The conditional density of* (Y_1, Y_2) *with* $BTC(\boldsymbol{\lambda}; G_1, G_2)$ *is given by*

$$f_{Y_1|Y_2}(y_1|y_2) = \left[\frac{(1 + 2\lambda_{01}G_2(y_2))\, g_1(y_1)}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11})G_2(y_2)} + \frac{(2\lambda_{10} + 4\lambda_{11}G_2(y_2))\, G_1(y_1)g_1(y_1)}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11})G_2(y_2)}\right],$$

and

$$f_{Y_2|Y_1}(y_2|y_1) = \left[\frac{(1 + 2\lambda_{10}G_1(y_1))\, g_2(y_2)}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11})G_1(y_1)} + \frac{(2\lambda_{01} + 4\lambda_{11}G_1(y_1))\, G_2(y_2)g_2(y_2)}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11})G_1(y_1)}\right].$$

We verify that the conditional distribution of (2.23) with $BTC(\boldsymbol{\lambda}; G_1, G_2)$ are transmuted. The conditional density of (2.23) for $Y_1|Y_2$ is obtained as

$$f_{Y_1|Y_2}(y_1|y_2) = \frac{f(y_1, y_2; \boldsymbol{\lambda})}{f_{Y_2}(y_2)}$$

$$= \left[\frac{(1 + 2\lambda_{01} G_2(y_2))\, g_1(y_1)}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11}) G_2(y_2)} + \frac{(2\lambda_{10} + 4\lambda_{11} G_2(y_2))\, G_1(y_1) g_1(y_1)}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11}) G_2(y_2)} \right]. \tag{2.34}$$

The conditional density in (2.34) can be written in the form

$$\begin{aligned} f_{Y_1|Y_2}(y_1|y_2) = &\left(1 - \frac{\lambda_{10} + 2\lambda_{11} G_2(y_2)}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11}) G_2(y_2)}\right) g_1(y_1) \\ &- 2\left(\frac{-(\lambda_{10} + 2\lambda_{11} G_2(y_2))}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11} G_2(y_2))} g_1(y_1) G_1(y_1)\right), \end{aligned} \tag{2.35}$$

and hence the conditional distribution of $f_{Y_1|Y_2}(y_1|y_2)$ is transmuted with stochastic representation as

$$Y_1|Y_2 = y_2 \sim TD\left(\lambda_1(y_2); G_1\right),$$

where,

$$\lambda_1(y_2) = -\frac{\lambda_{10} + 2\lambda_{11} G_2(y_2)}{1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11}) G_2(y_2)}.$$

and similarly the conditional density for $Y_2|Y_1$ can be written in the form

$$\begin{aligned} f_{Y_2|Y_1}(y_2|y_1) = &\left(1 - \frac{\lambda_{01} + 2\lambda_{11} G_1(y_1)}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11}) G_1(y_1)}\right) g_2(y_2) \\ &- 2\left(\frac{-(\lambda_{01} + 2\lambda_{11} G_1(y_1))}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11} G_2(y_2))} g_2(y_2) G_2(y_2)\right), \end{aligned} \tag{2.36}$$

with stochastic representation as

$$Y_2|Y_1 = y_1 \sim TD\left(\lambda_2(y_1); G_2\right),$$

where,

$$\lambda_2(y_1) = -\frac{\lambda_{01} + 2\lambda_{11} G_1(y_1)}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11}) G_1(y_1)}. \tag{2.37}$$

Property 2.3.3. *The $BTC(\boldsymbol{\lambda}, G_1, G_2)$ reduces to the product of the marginals with $Y_1 \sim TD(-\frac{\lambda_{10}}{1+\lambda_{10}}, G_1)$ and $Y_2 \sim TD(-\frac{\lambda_{01}}{1+\lambda_{01}}, G_2)$ if and only if $\lambda_{11} = \lambda_{01}\lambda_{10}$.*

Proof. For the choice of $\lambda_{11} = \lambda_{01}\lambda_{10}$, (2.23) reduces to

$$f(y_1, y_2; \boldsymbol{\lambda}) = \frac{1}{[1 + \lambda_{10} + \lambda_{01} + \lambda_{01}\lambda_{10}]} \left[1 + 2\lambda_{10} G_1(y_1) + 2\lambda_{01} G_2(y_2) + 4\lambda_{01}\lambda_{10} G_1(y_1) G_2(y_2)\right]$$

$$
\begin{aligned}
&\times g_1(y_1)g_2(y_2)\\
&= \frac{1}{(1+\lambda_{10})(1+\lambda_{01})}\left[(1+2\lambda_{10}G_1(y_1))g_1(y_1)(1+2\lambda_{01}G_2(y_2))g_2(y_2)\right]\\
&= \left(\frac{(1+2\lambda_{10}G_1(y_1))}{(1+\lambda_{10})}\right) g_1(y_1) \left(\frac{(1+2\lambda_{01}G_2(y_2))}{(1+\lambda_{01})}\right) g_2(y_2).
\end{aligned}
$$

This can be further factorised into the following form

$$
f(y_1, y_2; \boldsymbol{\lambda}) = \left(\frac{1}{1+\lambda_{10}} + \frac{2\lambda_{10}}{1+\lambda_{10}} G_1(y_1)\right) g_1(y_1) \left(\frac{1}{1+\lambda_{01}} + \frac{2\lambda_{01}}{1+\lambda_{01}} G_2(y_2)\right) g_2(y_2),
$$

which is the case of independence with marginals $Y_1 \sim TD(-\frac{\lambda_{10}}{1+\lambda_{10}}, G_1)$ and $Y_2 \sim TD(-\frac{\lambda_{01}}{1+\lambda_{01}}, G_2)$.

Conversely, if $(Y_1, Y_2) \sim BTC(\boldsymbol{\lambda}; G_1, G_2)$, and Y_1 and Y_2 are independent then

$$
\lambda_i(y_j) = k_i;\ i \neq j = 1, 2.
$$

where k_i's are some constant in $\mathscr{S}'$. Observe that $\lambda_1(y_2) = k_1$ implies

$$
G_1(y_2)\left[(\lambda_{10} + \lambda_{11})k_1 - \lambda_{11}\right] = \lambda_{01} - k_1(1 + \lambda_{01}), \tag{2.38}
$$

from which it follows that

$$
k_1 = \frac{\lambda_{01}}{1+\lambda_{01}}.
$$

Similarly, it is shown that

$$
k_2 = \frac{\lambda_{10}}{1+\lambda_{10}}.
$$

Substituting for k_1 in (2.38) we have $\lambda_{11} = \lambda_{10}\lambda_{01}$, and hence the result. ■

2.4 Simulation

The generation of bivariate samples (y_{1i}, y_{2i}), $i = 1, 2, ..., n$ from (2.3) is direct by taking into account that,

$$Y_1 \sim TD\left(\lambda_1'; G_1\right),$$

$$Y_2|Y_1 = y_1 \sim TD\left(\lambda_2(y_1); G_2\right),$$

where λ_1' and $\lambda_2(y_1)$ are defined in (2.30) and (2.37), respectively.
Let

$$u_1 = F_{Y_1}(y_1) = (1+\lambda_1')G_1(y_1) - \lambda_1'(G_1(y_1))^2,$$

that is

$$\lambda_1'(G_1(y_1))^2 - (1+\lambda_1')G_1(y_1) + u_1 = 0. \tag{2.39}$$

Consider (2.39) as a quadratic equation in $G_1(y_1)$, then

$$\Delta_1 = (1+\lambda_1')^2 - 4\lambda_1' u_1, \tag{2.40}$$

where Δ_1 is the discriminant of the quadratic equation. The quadratic equation in (2.39) has roots $\frac{(1+\lambda_1')\pm\sqrt{\Delta_1}}{2\lambda_1'}$.
If we denote by $Q_{G_i}(u)$ the quantile function of $G_i,\ i = 1, 2$, then the real root is (see Theorem 1.7.1 on page 27),

$$\begin{aligned} G_1(y_1) &= \left(\frac{1}{2\lambda_1'}\left(1+\lambda_1' - \sqrt{(1+\lambda_1')^2 - 4\lambda_1' u_1}\right)\right) \\ y_1 &= Q_{G_1}\left(\frac{1}{2\lambda_1'}\left(1+\lambda_1' - \sqrt{(1+\lambda_1')^2 - 4\lambda_1' u_1}\right)\right). \end{aligned} \tag{2.41}$$

Similarly,

$$u_2 = F_{Y_2|Y_1}(y_2|y_1) = (1+\lambda_2(y_1))G_2(y_2) - \lambda_2(y_1)(G_2(y_2))^2,$$

that is

$$\lambda_2(y_1)(G_2(y_2))^2 - (1+\lambda_2(y_1))G_2(y_2) + u_2 = 0. \tag{2.42}$$

Consider (2.42) as a quadratic equation in $G_2(y_2)$, then

$$\Delta_2 = (1+\lambda_2(y_1))^2 - 4\lambda_2(y_1)u_2, \tag{2.43}$$

where Δ_2 is the discriminant of the quadratic equation. The quadratic equation in (2.42) has roots $\frac{(1+\lambda_2(y_1))\pm\sqrt{\Delta_2}}{2\lambda_2(y_1)}$. Then the real root is (from Theorem 1.7.1 on page 27),

$$\begin{aligned} G_2(y_2) &= \left(\frac{1}{2\lambda_2(y_1)}\left(1+\lambda_2(y_1)-\sqrt{(1+\lambda_2(y_1))^2-4\lambda_2(y_1)u_2}\right)\right) \\ y_2 &= Q_{G_2}\left(\frac{1}{2\lambda_2(y_1)}\left(1+\lambda_2(y_1)-\sqrt{(1+\lambda_2(y_1))^2-4\lambda_2(y_1)u_2}\right)\right). \end{aligned} \tag{2.44}$$

Thus, from (2.41) and (2.44), a sample (y_1, y_2) from a bivariate transmuted conditional distribution is obtained by

$$y_1 = Q_{G_1}\left(\frac{1}{2\lambda_1'}\left(1+\lambda_1'-\sqrt{(1+\lambda_1')^2-4\lambda_1'u_1}\right)\right),$$

$$y_2 = Q_{G_2}\left(\frac{1}{2\lambda_2(y_1)}\left(1+\lambda_2(y_1)-\sqrt{(1+\lambda_2(y_1))^2-4\lambda_2(y_1)u_2}\right)\right),$$

where $u_i,\ i=1,2$ are independent and identically distributed random samples from a uniform distribution in $[0,1]$.

2.4.1 Cross-Moments

Let $X_{1i} \sim G_1$ and $X_{2i} \sim G_2, i = 1,2$ be two sets of random variables with cumulative distribution functions G_1 and G_2 and probability density functions g_1 and g_2 respectively. We denote the ordinary moments by $\mu_r'(X) = E(X^r)$. Then, if $(Y_1, Y_2) \sim BTC(\boldsymbol{\lambda}; G_1, G_2)$, the cross-moments $E(Y_1^{r_1}Y_2^{r_2})$ are given by,

$$\begin{aligned} E(Y_1^{r_1}Y_2^{r_2}) &= \int_{y_2}\int_{y_1} y_1^{r_1}y_2^{r_2} f(y_1,y_2)dy_1dy_2 \\ &= \int_{y_2}\int_{y_1} y_1^{r_1}y_2^{r_2} k(\boldsymbol{\lambda})\left[1+2\lambda_{10}G_1(y_1)+2\lambda_{01}G_2(y_2)+4\lambda_{11}G_1(y_1)G_2(y_2)\right] \\ &\times g_1(y_1)g_2(y_2)dy_1dy_2 \end{aligned}$$

$$
\begin{aligned}
&= k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} y_1^{r_1} y_2^{r_2} g_1(y_1) g_2(y_2) dy_1 dy_2 \\
&+ 2\lambda_{10} k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} y_1^{r_1} y_2^{r_2} g_1(y_1) g_2(y_2) G_1(y_1) dy_1 dy_2 \\
&+ 2\lambda_{01} k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} y_1^{r_1} y_2^{r_2} g_1(y_1) g_2(y_2) G_2(y_2) dy_1 dy_2 \\
&+ 4\lambda_{11} k(\boldsymbol{\lambda}) \int_{y_2}\int_{y_1} g_1(y_1) g_2(y_2) G_1(y_1) G_2(y_2) dy_1 dy_2. \qquad (2.45)
\end{aligned}
$$

Now by applying Theorem 1.6.3, page 23, we get $E(Y_1^{r_1} Y_2^{r_2})$ as,

$$
\frac{\mu'_{r_1}(X_1)\mu'_{r_2}(X_2) + \lambda_{10}\mu'_{r_1}(X_{1(2)})\mu'_{r_2}(X_2) + \lambda_{01}\mu'_{r_1}(X_1)\mu'_{r_2}(X_{2(2)}) + \lambda_{11}\mu'_{r_1}(X_{1(2)})\mu'_{r_2}(X_{2(2)})}{1 + \lambda_{01} + \lambda_{10} + \lambda_{11}}, \qquad (2.46)
$$

where $X_{1(2)} = \text{Max}\{X_{11}, X_{12}\}$, $X_{2(2)} = \text{Max}\{X_{21}, X_{22}\}$. Also, let ($X_{11}$,$X_{12}$) be two independent and identically distributed independent copies of X_1 and (X_{21},X_{22}) be two independent and identically distributed independent copies of X_2.

2.4.2 Dependence Measures

In this section, we consider conditions under which $BTC(\boldsymbol{\lambda}, G_1, G_2)$ is $TP2$ and $RR2$.

Theorem 2.4.1. *Let* $(Y_1, Y_2) \sim BTC(\boldsymbol{\lambda}; G_1, G_2)$ *a bivariate distribution with transmuted conditionals and joint probability density function* (2.3). *Then,*

- *If* $\lambda_{11} > \lambda_{01}\lambda_{10}$, $f_{Y_1,Y_2}(y_1, y_2)$ *is* $TP2$,
- *If* $\lambda_{11} < \lambda_{01}\lambda_{10}$, $f_{Y_1,Y_2}(y_1, y_2)$ *is* $RR2$.

Proof. From (2.3), we have

$$
\begin{aligned}
\log f(y_1, y_2) &= \log k(\boldsymbol{\lambda}) + \log g_1(y_1) + \log g_2(y_2) \\
&\quad + \log\left[1 + 2\lambda_{10}G_1(y_1) + 2\lambda_{01}G_2(y_2) + 4\lambda_{11}.G_1(y_1)G_2(y_2)\right]
\end{aligned}
$$

Now,

$$\frac{\partial \log f(y_1,y_2)}{\partial y_1} = \frac{g_1'(y_1)}{g_1(y_1)} + \frac{2\lambda_{10}g_1(y_1) + 4\lambda_{11}g_1(y_1)G_2(y_2)}{[1+2\lambda_{10}G_1(y_1)+2\lambda_{01}G_2(y_2)+4\lambda_{11}G_1(y_1)G_2(y_2)]}.$$

Next,

$$\frac{\partial^2 \log f(y_1,y_2)}{\partial y_1 \partial y_2} = \frac{4\lambda_{11}g_1(y_1)g_2(y_2) - 4\lambda_{01}\lambda_{10}g_1(y_1)g_2(y_2)}{[1+2\lambda_{10}G_1(y_1)+2\lambda_{01}G_2(y_2)+4\lambda_{11}G_1(y_1)G_2(y_2)]^2}.$$

Hence, the local dependence function is given by,

$$\eta(y_1,y_2) = \frac{4(\lambda_{11}-\lambda_{01}\lambda_{10})g_1(y_1)g_2(y_2)}{[1+2\lambda_{10}G_1(y_1)+2\lambda_{01}G_2(y_2)+4\lambda_{11}G_1(y_1)G_2(y_2)]^2}. \tag{2.47}$$

It is quite evident from (2.47) that the local dependence function $\eta(y_1,y_2) \geq 0$ when $\lambda_{11} \geq \lambda_{01}\lambda_{10}$ and $\eta(y_1,y_2) \leq 0$ when $\lambda_{11} \leq \lambda_{01}\lambda_{10}$. Hence, from Theorem 1.4.1, page 8, we obtain the result. ■

2.4.3 Bivariate Cumulative Distribution, Survival and Hazard Rate Functions

The bivariate cumulative distribution function has a simple closed form and is given by

$$F(y_1,y_2;\boldsymbol{\lambda}) = \frac{G_1(y_1)G_2(y_2) + \lambda_{01}G_1(y_1)G_2^2(y_2) + \lambda_{10}G_2(y_2)G_1^2(y_1) + \lambda_{11}G_1^2(y_1)G_2^2(y_2)}{1+\lambda_{01}+\lambda_{10}+\lambda_{11}}. \tag{2.48}$$

Taking into account that $S(y_1,y_2) = 1 - F_{Y_1}(y_1) - F_{Y_2}(y_2) + F(y_1,y_2)$, the bivariate survival function is given by,

$$\begin{aligned} S(y_1,y_2;\boldsymbol{\lambda}) = 1 &- k(\boldsymbol{\lambda})\left[(1+\lambda_{01})G_1(y_1) + (\lambda_{10}+\lambda_{11})G_1^2(y_1)\right] \\ &- k(\boldsymbol{\lambda})\left[(1+\lambda_{10})G_2(y_2) + (\lambda_{01}+\lambda_{11})G_2^2(y_2)\right] \\ &+ k(\boldsymbol{\lambda})\left[G_1(y_1)G_2(y_2)\left(1+\lambda_{01}G_2(y_2)\right) + \left(\lambda_{10}+\lambda_{11}G_2(y_2)\right)G_2(y_2)G_1^2(y_1)\right]. \end{aligned} \tag{2.49}$$

The bivariate hazard rate function defined by Basu (1971) is obtained combining (2.3) with (2.49) to get,

$$r(y_1, y_2) = \frac{f(y_1, y_2)}{S(y_1, y_2)}$$
$$= \frac{k(\boldsymbol{\lambda})\mathscr{G}_1(y_1, y_2; \boldsymbol{\lambda})}{1 - k(\boldsymbol{\lambda})\left[\lambda_{01}(\mathscr{G}_2(y_1, y_2)) + \lambda_{10}\mathscr{G}_3(y_1, y_2) + \lambda_{11}\mathscr{G}_4(y_1, y_2) + \mathscr{G}_5(y_1, y_2)\right]}, \tag{2.50}$$

where,

$$\mathscr{G}_1(y_1, y_2; \boldsymbol{\lambda}) = (1 + 2\lambda_{10}G_1(y_1) + 2\lambda_{01}G_2(y_2) + 4\lambda_{11}G_1(y_1)G_2(y_2))\, g_1(y_1)g_2(y_2),$$
$$\mathscr{G}_2(y_1, y_2) = G_1(y_1) + G_2^2(y_2) - G_1(y_1)G_2^2(y_2),$$
$$\mathscr{G}_3(y_1, y_2) = G_2(y_2) + G_1^2(y_1) - G_1^2(y_1)G_2(y_2),$$
$$\mathscr{G}_4(y_1, y_2) = G_1^2(y_1) + G_2^2(y_2) - G_1^2(y_1)G_2^2(y_2),$$
$$\mathscr{G}_5(y_1, y_2) = G_1(y_1) + G_2(y_2) - G_1(y_1)G_2(y_2).$$

If $(Y_1, Y_2) \sim BTC(\boldsymbol{\lambda}; G_1, G_2)$, we have the Johnson & Kotz (1975) hazard gradient in (1.6) and (1.7) as,

$$r_1(y_1, y_2) = -\frac{\left[1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11})G_1 + G_2 + 2\lambda_{10}G_1G_2 + \lambda_{01}G_2^2 + 2\lambda_{11}G_1G_2^2\right] g_1 k(\boldsymbol{\lambda})}{S(y_1, y_2)}, \tag{2.51}$$

and

$$r_2(y_1, y_2) = -\frac{\left[1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11})G_2 + G_1 + 2\lambda_{01}G_1G_2 + \lambda_{10}G_1^2 + 2\lambda_{11}G_1^2G_2\right] g_2 k(\boldsymbol{\lambda})}{S(y_1, y_2)}, \tag{2.52}$$

where we have omitted the arguments of the functions $G_i(y_i)$ and $g_i(y_i)$, $i = 1, 2$.

The monotonicity of the conditional hazard rate of Y_1 given $Y_2 > y_2$ can be obtained by using the result, the conditional hazard rate is decreasing in y_2 for every y_1 (Shaked (1977)). Hence we have the following lemma.

Lemma 2.4.1. *If $f(y_1, y_2)$ is $TP2$ ($RR2$), the conditional hazard rate $r_1(y_1, y_2)$ of Y_1 given $Y_2 > y_2$ is decreasing (increasing) in y_2.*

Proof. By using the above result and Theorem 2.4.1, if $\lambda_{11} > \lambda_{01}\lambda_{10}(<)$, we conclude that $r_1(y_1, y_2)$ is decreasing (increasing) in y_2 and $r_2(y_1, y_2)$ is decreasing (increasing) in y_1. ■

It is of interest to quantify the association between the failure times in bivariate data through local dependence measure. If $(Y_1, Y_2) \sim BTC(\boldsymbol{\lambda}; G_1, G_2)$, we have the Clayton's cross-ratio function Clayton (1978) in (1.16) as,

$$\mathscr{C}(y_1, y_2) = \frac{k(\boldsymbol{\lambda})S(y_1, y_2; \boldsymbol{\lambda})\mathscr{G}_1(y_1, y_2; \boldsymbol{\lambda})}{S_1(y_1, y_2; \boldsymbol{\lambda})S_2(y_1, y_2; \boldsymbol{\lambda})} \tag{2.53}$$

where,

$$\begin{aligned} S_1(y_1, y_2; \boldsymbol{\lambda}) &= k(\boldsymbol{\lambda}) \left\{g_1 G_2[(\lambda_{01} + 2\lambda_{11}G_2)G_2]\right\} \\ &\quad - k(\boldsymbol{\lambda}) \left\{(1 + \lambda_{01})g_1 + 2G_1 g_1[(\lambda_{10} + \lambda_{11}) - (\lambda_{10} + \lambda_{11}G_2)G_2]\right\}, \\ S_2(y_1, y_2; \boldsymbol{\lambda}) &= k(\boldsymbol{\lambda}) \left\{g_2 G_1[(\lambda_{10} + 2\lambda_{11}G_1)G_1]\right\} \\ &\quad - k(\boldsymbol{\lambda}) \left\{(1 + \lambda_{10})g_2 + 2G_2 g_2[(\lambda_{01} + \lambda_{11}) - (\lambda_{01} + \lambda_{11}G_1)G_1]\right\}, \end{aligned}$$

where we have omitted the arguments of the functions $G_i(y_i)$ and $g_i(y_i)$, $i = 1, 2$.

2.4.4 Concomitants of Order Statistics

Let (Y_{1i}, Y_{2i}), $i = 1, 2, ..., n$ be n independent random variables from a bivariate distribution. If we arrange the Y_1 variates in ascending order as $Y_{1_{[1(n)]}} \leq Y_{1_{[2(n)]}} \leq ... \leq Y_{1_{[n(n)]}}$, then the Y_2 variates corresponding to these order statistics are denoted by $Y_{2_{[1(n)]}} \leq Y_{2_{[2(n)]}} \leq ... \leq Y_{2_{[n(n)]}}$, and termed the concomitants of the order statistic. In particular, for $r = 1, 2, ..., n$, we denote $Y_{2_{[r(n)]}}$ the concomitant of the r^{th} order statistic. The density function of the concomitant of the r^{th} order statistics of the first component, is given by (David & Nagaraja (2003)).

$$f_{[r(n)]}(y_2) = \int_{y_1} f_{Y_2|Y_1}(y_2|y_1) f_{[r(n)]}(y_1) dy_1, \tag{2.54}$$

where $f_{[r(n)]}(y_1)$ is the probability density function of the r^{th} order statistic of Y_1 from a sample of size n.

Theorem 2.4.2. *Let* $(Y_1, Y_2) \sim BTC(\boldsymbol{\lambda}; G_1, G_2)$ *be a bivariate distribution with transmuted conditionals and joint probability density function* (2.3)*. Then, the distribution of the concomitant of the* r^{th} *order statistic is again transmuted with distribution,*

$$Y_{2_{[r(n)]}} \sim TD\left(\int_{y_1} \lambda_2(y_1) f_{[r(n)]}(y_1) dy_1; G_2\right), \tag{2.55}$$

where $f_{[r(n)]}(y_1)$ *is the density function of the* r^{th} *order statistic of* Y_1 *from a sample of size* n*.*

Proof. The density function of the concomitant of the r^{th} order statistic of the first component is,

$$f_{[r(n)]}(y_2) = \int_{y_1} f_{Y_2|Y_1}(y_2|y_1) f_{[r(n)]}(y_1) dy_1,$$

and then substituting $f_{Y_2|Y_1}(y_2|y_1)$ from its expression in (2.36) we get,

$$\begin{aligned} f_{[r(n)]}(y_2) &= \int_{y_1} \left(1 - \frac{\lambda_{01} + 2\lambda_{11} G_1(y_1)}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11}) G_1(y_1)}\right) g_2(y_2) f_{[r(n)]}(y_1) dy_1 \\ &\quad - 2 \int_{y_1} \left(\frac{-(\lambda_{01} + 2\lambda_{11} G_1(y_1))}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11} G_2(y_2))} g_2(y_2) G_2(y_2)\right) f_{[r(n)]}(y_1) dy_1 \\ &= \left\{\left(1 + \int_{y_1} \lambda_2(y_1) f_{[r(n)]}(y_1) dy_1\right) - 2\left(\int_{y_1} \lambda_2(y_1) f_{[r(n)]}(y_1) dy_1\right) G_2(y_2)\right\} g_2(y_2), \end{aligned}$$

where,

$$\lambda_2(y_1) = \left(1 - \frac{\lambda_{01} + 2\lambda_{11} G_1(y_1)}{1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11}) G_1(y_1)}\right).$$

Hence, the r^{th} order statistic $Y_{2_{[r(n)]}}$ is transmuted with stochastic representation as

$$Y_{2_{[r(n)]}} \sim TD\left(\int \lambda_2(y_1) f_{[r(n)]}(y_1) dy_1; G_2\right). \tag{2.56}$$

Thus we obtain (2.55). ∎

2.5 Some Examples

In this section we consider some examples of bivariate distributions with transmuted conditional distributions by considering specific baseline distributions for $G(x)$.

2.5.1 Bivariate Distributions with Transmuted Uniform Conditionals

Let X be a transmuted uniform distribution with cumulative distribution function,

$$F(x;\lambda) = (1+\lambda)x - \lambda x^2, 0 \leq x \leq 1,\ |\lambda| \leq 1. \tag{2.57}$$

Now, we are to consider the most general bivariate distribution with uniform conditionals of the form (2.57).

Using Theorem 2.2.1 with $G_i(z) = z, i = 1, 2$, the joint probability density function in (2.3)) becomes,

$$f(y_1, y_2; \boldsymbol{\lambda}) = \frac{1 + 2\lambda_{10}y_1 + 2\lambda_{01}y_2 + 4\lambda_{11}y_1y_2}{1 + \lambda_{01} + \lambda_{10} + \lambda_{11}},\ 0 \leq y_1,\ y_2 \leq 1. \tag{2.58}$$

The marginal distributions are again transmuted uniforms,

$$Y_1 \sim TD(\lambda_1';\, y_1), \quad Y_2 \sim TD(\lambda_2';\, y_2),$$

where $\lambda_i', i = 1, 2$ are defined in (2.30) and (2.31).

Using formula (2.46) the cross moments $E(Y_1^{r_1}Y_2^{r_2})$ are given by,

$$\left\{\frac{1}{(1+r_1)(1+r_2)} + \frac{2\lambda_{10}}{(2+r_1)(1+r_2)} + \frac{2\lambda_{01}}{(1+r_1)(2+r_2)} + \frac{4\lambda_{11}}{(2+r_1)(2+r_2)}\right\} k(\boldsymbol{\lambda}) \tag{2.59}$$

The covariance takes a simple expression and is given by,

$$Cov\,(Y_1, Y_2) = \frac{\lambda_{11} - \lambda_{10}\lambda_{01}}{36(1 + \lambda_{10} + \lambda_{01} + \lambda_{11})^2}. \tag{2.60}$$

The local dependence function $\eta(y_1, y_2)$ is obtained as

$$\eta(y_1, y_2) = \frac{4(\lambda_{11} - \lambda_{10}\lambda_{01})}{[1 + 2\lambda_{10}y_1 + 2\lambda_{01}y_2 + 4\lambda_{11}y_1y_2]^2}. \tag{2.61}$$

2.5.2 Bivariate Distributions with Transmuted Normal Conditionals

Let X be a transmuted normal random variable with probability density function ,

$$f(x; \lambda) = (1 + \lambda)\phi(x) - 2\lambda\phi(x)\Phi(x), \ |\lambda| \leq 1. \tag{2.62}$$

where $\phi(x)$ and $\Phi(x)$ denote respectively, the probability density function and the cumulative distribution function of the standard normal distribution. In this case, $G_i(z) = \Phi(z), \ i = 1, 2$, and using (2.3) we obtain the joint probability density function,

$$f(y_1, y_2; \boldsymbol{\lambda}) = \frac{4\phi(y_1)\phi(y_2)(1 + 2\lambda_{10}\Phi(y_1) + 2\lambda_{01}\Phi(y_2) + 4\lambda_{11}\Phi(y_1)\Phi(y_2)}{1 + \lambda_{01} + \lambda_{10} + \lambda_{11}}, \tag{2.63}$$

where $y_1, y_2 \in \mathscr{R}$.

The regression functions are non-linear. The conditional mathematical expectations are given by,

$$E(Y_1|Y_2 = y_2) = \frac{\lambda_{10} + 2\lambda_{11}\Phi(y_2)}{\sqrt{\pi}(1 + \lambda_{10} + 2(\lambda_{01} + \lambda_{11})\Phi(y_2))}, \tag{2.64}$$

$$E(Y_2|Y_1 = y_1) = \frac{\lambda_{01} + 2\lambda_{11}\Phi(y_1)}{\sqrt{\pi}(1 + \lambda_{01} + 2(\lambda_{10} + \lambda_{11})\Phi(y_1))}. \tag{2.65}$$

The local dependence function $\eta(y_1, y_2)$ is obtained as

$$\eta(y_1, y_2) = \frac{4(\lambda_{11} - \lambda_{10}\lambda_{01})\phi(y_1)\phi(y_2)}{[1 + 2\lambda_{10}\Phi(y_1) + 2\lambda_{01}\Phi(y_2) + 4\lambda_{11}\Phi(y_1)\Phi(y_2)]^2}. \tag{2.66}$$

2.5.3 Bivariate Distributions with Transmuted Exponential Conditionals

Now, let us consider the transmuted exponential distribution with probability density function,

$$f(x;\lambda,\beta) = (1-\lambda)\beta e^{-\beta x} + 2\lambda\beta e^{-2\beta x},\ x \geq 0,\ |\lambda| \leq 1. \tag{2.67}$$

In this case, $G_i(z) = 1 - e^{-\beta_i z},\ i = 1, 2$. Then using (2.3) we obtain the joint probability density function,

$$\begin{aligned} f(y_1, y_2; \boldsymbol{\lambda}) = &\big[1 + 2\lambda_{10}(1 - e^{-\beta_1 y_1}) + 2\lambda_{01}(1 - e^{-\beta_2 y_2}) + 4\lambda_{11}(1 - e^{-\beta_1 y_1})(1 - e^{-\beta_2 y_2})\big] \\ &\times k(\boldsymbol{\lambda})\beta_1\beta_2 e^{-\beta_1 y_1 - \beta_2 y_2}. \end{aligned} \tag{2.68}$$

The marginal distributions are again transmuted exponentials,

$$Y_1 \sim TD(\lambda_1';\ 1 - e^{-\beta_1 y_1}),\ \ Y_2 \sim TD(\lambda_2';\ 1 - e^{-\beta_2 y_2}).$$

This cross-moments are given by,

$$E(Y_1^{r_1} Y_2^{r_2}) = k(\boldsymbol{\lambda})\frac{\Gamma(r_1+1)\Gamma(r_2+1)}{\beta_1^{r_1}\beta_2^{r_2}}\left\{1 + 2\lambda_{10}c_{r_1} + 2\lambda_{01}c_{r_2} + 4\lambda_{11}c_{r_1}c_{r_2}\right\}, \tag{2.69}$$

where $c_r = 2 - 2^{-r}$. The covariance takes the expression,

$$\begin{aligned} Cov(Y_1, Y_2) = &k(\boldsymbol{\lambda})\left(\frac{1 + 3\lambda_{01} + 3\lambda_{10} + 9\lambda_{11}}{\beta_1\beta_2}\right) \\ &- \frac{[k(\boldsymbol{\lambda})]^2}{\beta_1\beta_2}\left(1 + 3\lambda_{01} + 2\lambda_{10} + 6\lambda_{11}\right)\left(1 + 2\lambda_{01} + 3\lambda_{10} + 6\lambda_{11}\right). \end{aligned} \tag{2.70}$$

A contour plot and plot for probability density function of bivariate transmuted distribution with exponential conditionals are shown in Figure 2.1 and Figure 2.2 respectively.

The local dependence function is obtained as

$$\eta(y_1, y_2) = \frac{4(\lambda_{11} - \lambda_{10}\lambda_{01})\beta_1\beta_2 e^{-(\beta_1 y_1 + \beta_2 y_2)}}{[1 + 2\lambda_{10}(1 - e^{-\beta_1 y_1}) + 2\lambda_{01}(1 - e^{-\beta_2 y_2}) + 4\lambda_{11}(1 - e^{-\beta_1 y_1})(1 - e^{-\beta_2 y_2})]^2} \tag{2.71}$$

2.5.4 Bivariate Distributions with Transmuted Weibull Conditionals

Now, let us consider the transmuted Weibull distribution with probability density function,

$$f(x;\lambda,\gamma) = (1-\lambda)\gamma x^{\gamma-1}e^{-x^{\gamma}} + 2\lambda\gamma x^{\gamma-1}e^{-2x^{\gamma}},\ x \geq 0,\ |\lambda| \leq 1. \tag{2.72}$$

In this case, $G_i(z) = 1 - e^{-z^{\gamma_i}}$, $i = 1, 2$. Then using (2.3) we obtain the joint probability density function,

$$\begin{aligned} f(y_1, y_2; \boldsymbol{\lambda}) = &\left[1 + 2\lambda_{10}\left(1 - e^{-y_1^{\gamma_1}}\right) + 2\lambda_{01}\left(1 - e^{-y_2^{\gamma_2}}\right) + 4\lambda_{11}\left(1 - e^{-y_1^{\gamma_1}}\right)\left(1 - e^{-y_2^{\gamma_2}}\right)\right] \\ &\times k(\boldsymbol{\lambda})\gamma_1 y_1^{\gamma_1-1}e^{-y_1^{\gamma_1}}\gamma_2 y_2^{\gamma_2-1}e^{-y_2^{\gamma_2}}. \end{aligned} \tag{2.73}$$

The marginal distributions are again transmuted Weibull conditionals,

$$Y_1 \sim TD\left(\lambda_1';\, 1 - e^{-y_1^{\gamma_1}}\right),\ Y_2 \sim TD\left(\lambda_2';\, 1 - e^{-y_2^{\gamma_2}}\right).$$

The cross-moments $E[Y_1^{r_1}Y_2^{r_2}]$ are given by,

$$\begin{aligned} &\left[2^{\left(\frac{r_1}{\gamma_1}+\frac{r_2}{\gamma_2}\right)} + 2^{\frac{r_1}{\gamma_1}}\lambda_{01}\left(2^{\frac{r_2+\gamma_2}{\gamma_2}}\right) + 2^{\frac{r_2}{\gamma_2}}\lambda_{10}\left(2^{\frac{r_1+\gamma_1}{\gamma_1}}\right) + \lambda_{11}\left(1 - 2^{\frac{r_1+\gamma_1}{\gamma_1}} + 2^{2+\frac{r_1}{\gamma_1}+\frac{r_2}{\gamma_2}} - 2^{\frac{r_2+\gamma_2}{\gamma_2}}\right)\right] \\ &\qquad \times k(\boldsymbol{\lambda})\Gamma\left(\frac{r_1+\gamma_1}{\gamma_1}\right)\Gamma\left(\frac{r_2+\gamma_2}{\gamma_2}\right)2^{-\left(\frac{r_1}{\gamma_1}+\frac{r_2}{\gamma_2}\right)}. \end{aligned} \tag{2.74}$$

The covariance between Y_1 and Y_2 is given by

$$\begin{aligned} Cov(Y_1, Y_2) = &\, 2^{-\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}}k(\boldsymbol{\lambda})\Gamma\left(1+\frac{1}{\gamma_1}\right)\Gamma\left(1+\frac{1}{\gamma_2}\right)\left[\lambda_{11}(1-2^{1+\frac{1}{\gamma_1}}) - 2^{\frac{1}{\gamma_1}}\lambda_{01}\right] \\ &- 2^{-\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}}[k(\boldsymbol{\lambda})]^2\left(2^{\frac{1}{\gamma_2}}(1+2\lambda_{01}+\lambda_{10}+2\lambda_{11}) - \lambda_{01} - \lambda_{11}\right) \\ &\times \left(2^{\frac{1}{\gamma_1}}(1+\lambda_{01}+2\lambda_{10}+2\lambda_{11}) - \lambda_{10} - \lambda_{11}\right)\Gamma\left(1+\frac{1}{\gamma_1}\right)\Gamma\left(1+\frac{1}{\gamma_2}\right) \\ &+ 2^{-\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}}k(\boldsymbol{\lambda})2^{\frac{1}{\gamma_2}}\left(2^{\frac{1}{\gamma_1}}(1+2\lambda_{01}+2\lambda_{10}+4\lambda_{11}) - \lambda_{10} - 2\lambda_{11}\right) \end{aligned}$$

$$\times \Gamma\left(1+\frac{1}{\gamma_1}\right)\Gamma\left(1+\frac{1}{\gamma_2}\right). \tag{2.75}$$

A contour plot and plot for probability density function of bivariate transmuted distribution with Weibull conditionals are shown in Figure 2.3 and Figure 2.4 respectively.

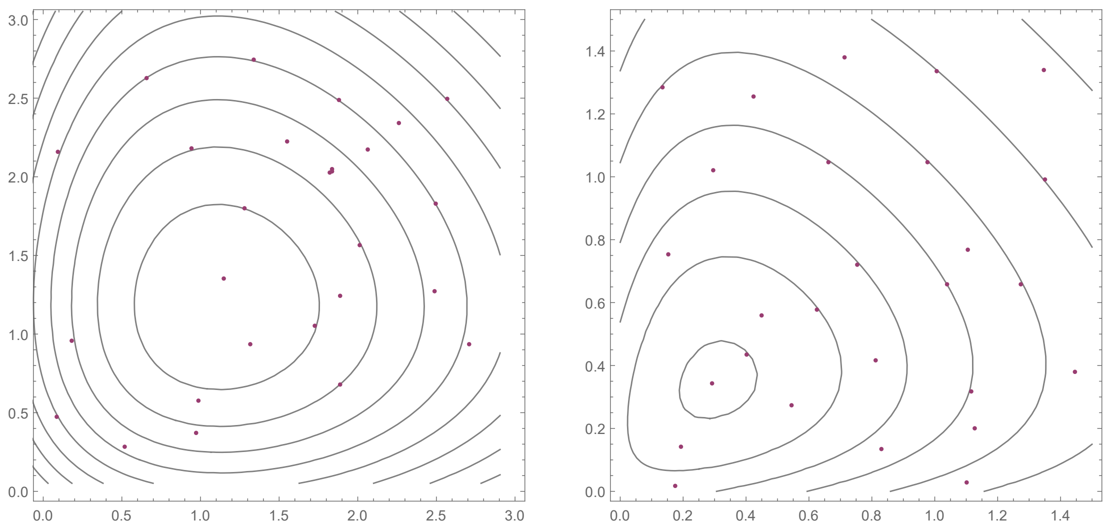


Figure 2.1: Contour plots for joint probability density function of the bivariate distribution with transmuted exponential conditionals for different parameter values with some sample points: (a) ($\lambda_{10} = 1.3$, $\lambda_{01} = 1.2$, $\lambda_{11} = 1.5$, $\beta_1 = 0.3$, $\beta_2 = 0.3$), (b) ($\lambda_{10} = 0.6$, $\lambda_{01} = 0.85$, $\lambda_{11} = 2.1$, $\beta_1 = 1.2$, $\beta_2 = 1.2$).

The local dependence function is obtained as

$$\eta(y_1, y_2) = \frac{4(\lambda_{11} - \lambda_{10}\lambda_{01})\gamma_1\gamma_2 y_1^{\gamma_1-1} y_2^{\gamma_2-1} e^{-(y_1^{\gamma_1}+y_2^{\gamma_2})}}{\left[1 + 2\lambda_{10}\left(1-e^{-y_1^{\gamma_1}}\right) + 2\lambda_{01}\left(1-e^{-y_2^{\gamma_2}}\right) + 4\lambda_{11}\left(1-e^{-y_1^{\gamma_1}}\right)\left(1-e^{-y_2^{\gamma_2}}\right)\right]^2} \tag{2.76}$$

2.5.5 Bivariate Distributions with Transmuted Exponentiated Frêchet Conditionals

The univariate cumulative distribution function for the transmuted exponentiated Frêchet distribution is given in (1.45) and its corresponding probability density func-

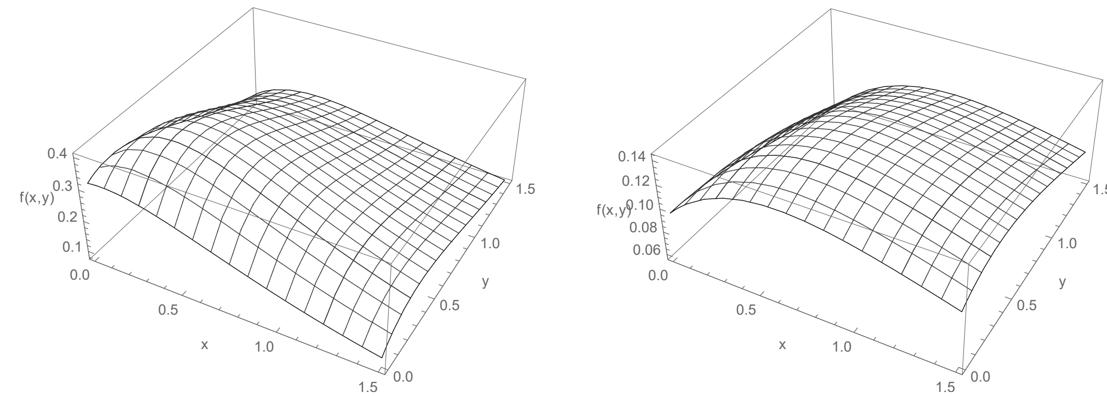


Figure 2.2: Plots for joint probability density function of the bivariate distribution with transmuted exponential conditionals for different parameter values: (a) ($\lambda_{10} = 0.6$, $\lambda_{01} = 0.85$, $\lambda_{11} = 2.1$, $\beta_1 = 1.2$, $\beta_2 = 1.2$), (b) ($\lambda_{10} = 1.6$, $\lambda_{01} = 0.55$, $\lambda_{11} = 2.0$, $\beta_1 = 0.7$, $\beta_2 = 0.7$).

tion is given in (1.46).

In this case,

$$G_i(z) = 1 - \left[1 - e^{-\left(\frac{\theta_i}{z}\right)^{\beta_i}}\right]^{\alpha_i}, \; i = 1, 2.$$

Then using (2.3) we obtain the joint probability density function,

$$\begin{aligned} f(y_1, y_2; \boldsymbol{\lambda}) &= k(\boldsymbol{\lambda}) q_1(y_1) q_2(y_2) \left[1 + 2\lambda_{10}\left(1 - \left[1 - e^{-\left(\frac{\theta_1}{y_1}\right)^{\beta_1}}\right]^{\alpha_1}\right)\right] \\ &+ k(\boldsymbol{\lambda}) q_1(y_1) q_2(y_2) \left[1 + 2\lambda_{01}\left(1 - \left[1 - e^{-\left(\frac{\theta_2}{y_2}\right)^{\beta_2}}\right]^{\alpha_2}\right)\right] \\ &+ 4\lambda_{11} k(\boldsymbol{\lambda}) q_1(y_1) q_2(y_2) \left(1 - \left[1 - e^{-\left(\frac{\theta_1}{y_1}\right)^{\beta_1}}\right]^{\alpha_1}\right)\left(1 - \left[1 - e^{-\left(\frac{\theta_2}{y_2}\right)^{\beta_2}}\right]^{\alpha_2}\right), \end{aligned} \tag{2.77}$$

where

$$q_i(y_i) = \alpha_i \beta_i \theta_i^{\beta_i} y_i^{-(1+\beta_i)} \, e^{-\left(\frac{\theta_i}{y_i}\right)^{\beta_i}} \left[1 - e^{-\left(\frac{\theta_i}{y_i}\right)^{\beta_i}}\right]^{\alpha_i - 1}, \; i = 1, 2.$$

The marginal distributions are again transmuted exponentiated Frêchet conditionals,

$$Y_1 \sim TD\left(\lambda_1'; 1 - \left[1 - e^{-\left(\frac{\theta_1}{y_1}\right)^{\beta_1}}\right]^{\alpha_1}\right), \; Y_2 \sim TD\left(\lambda_2'; 1 - \left[1 - e^{-\left(\frac{\theta_2}{y_2}\right)^{\beta_2}}\right]^{\alpha_2}\right).$$

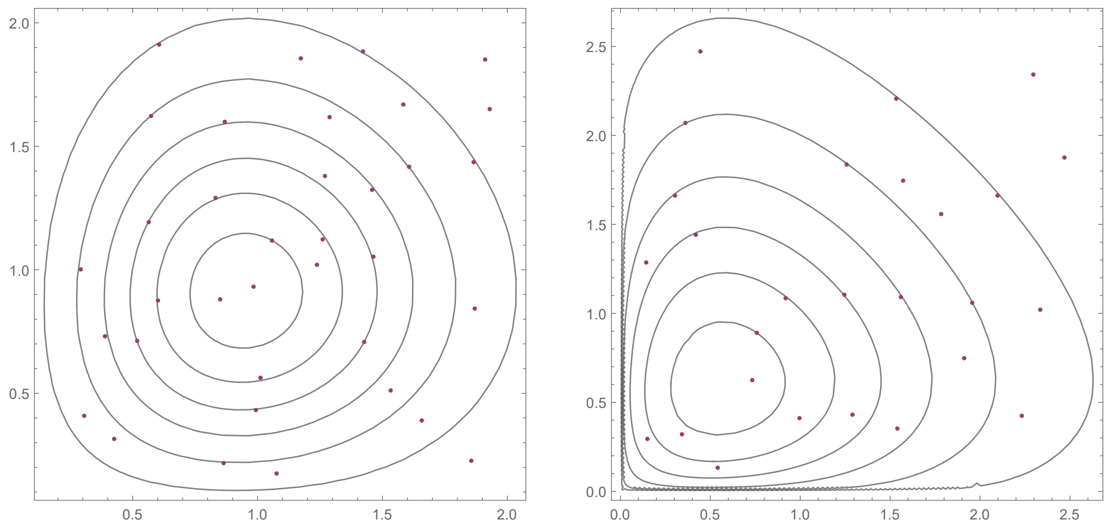


Figure 2.3: Contour plots for joint probability density function of the bivariate distribution with transmuted Weibull conditionals for different parameter values: (a) ($\lambda_{10} = 1.2$, $\lambda_{01} = 0.6$, $\lambda_{11} = 1.8$, $\gamma_1 = 1.9$, $\gamma_2 = 1.9$), (b) ($\lambda_{10} = 0.7$, $\lambda_{01} = 0.9$, $\lambda_{11} = 1.2$, $\gamma_1 = 1.3$, $\gamma_2 = 1.3$).

2.6 Parameter Estimation

In this section we discuss the method of moments and method of maximum likelihood. Let (Y_{1i}, Y_{2i}), $(i = 1, 2, ..., n)$ be a random sample from a bivariate distribution with transmuted conditionals as in (2.3). The unknown parameters involved here are transmuted and baseline parameters. The method of moment is probably the oldest method for constructing an estimator. The advantage of method of moment is that it is quite easy to use. From the cross moments function we establish the basic population moments such as means, variances and covariances then equating it to the corresponding sample moments, we can obtain the estimates and the computation is carried out by Mathematica 10. The parameters are estimated under the method of maximum likelihood by differentiating the likelihood function with respect to each parameter and equating it to zero and solving them simultaneously or by maximizing the likelihood function with respect to the unknown parameters. These techniques are discussed in Section 2.7 and Section 2.8.

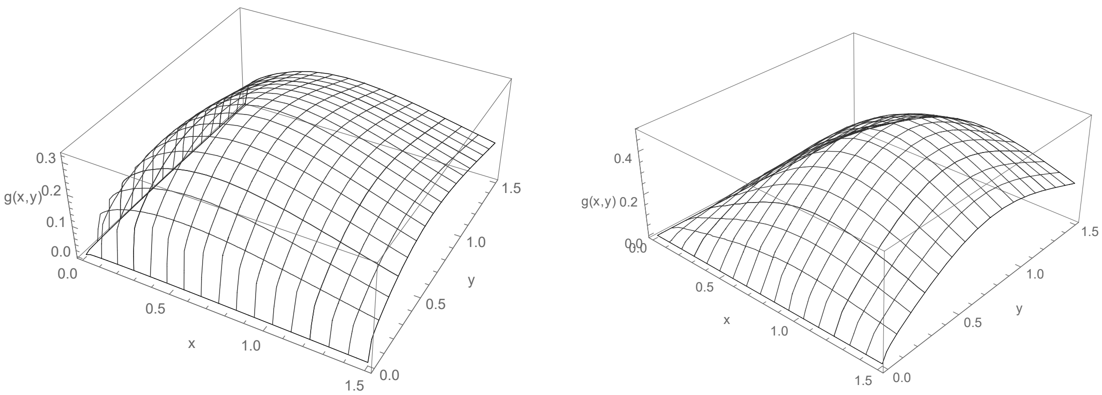


Figure 2.4: Plots for joint probbility density function of the bivariate transmuted distribution with Weibull conditionals for different parameter values: (a) ($\lambda_{10} = 0.6$, $\lambda_{01} = 0.85$, $\lambda_{11} = 2.1$, $\gamma_1 = 1.2$, $\gamma_2 = 1.2$), (b) ($\lambda_{10} = 1.6$, $\lambda_{01} = 0.55$, $\lambda_{11} = 2.0$, $\gamma_1 = 1.6$, $\gamma_2 = 1.6$).

2.6.1 Moment Estimates

In this section we propose moment estimates. The parameters are estimated by equating theoretical moment to sample moments, and by solving the moment equations in a exact way or numerically, once the sample means $\bar{y}_1 = \frac{\sum_{i=1}^{n} y_{1i}}{n}$ and $\bar{y}_2 = \frac{\sum_{i=1}^{n} y_{2i}}{n}$, variances $S^2_{y_1} = \frac{\sum_{i=1}^{n}(y_{1i}-\bar{y}_1)^2}{n}$ and $S^2_{y_2} = \frac{\sum_{i=1}^{n}(y_{2i}-\bar{y}_2)^2}{n}$, and covariance $S^2_{y_1,y_2} = \frac{\sum_{i=1}^{n}(y_{1i}-\bar{y}_1)^2(y_{2i}-\bar{y}_2)^2}{n}$ are obtained.

2.6.1.1 Bivariate distributions with transmuted uniform conditionals

The cross moments for the bivariate distribution with transmuted uniform conditionals is given in (2.59). By considering $r_1 = 1$ and $r_2 = 0$ in (2.59) we get,

$$E(Y_1) = \left\{ \frac{1}{2} + \frac{2\lambda_{10}}{3} + \frac{\lambda_{01}}{2} + \frac{2\lambda_{11}}{3} \right\} k(\boldsymbol{\lambda}).$$

If $r_1 = 0$ and $r_2 = 1$ in (2.59), we get

$$E(Y_2) = \left\{ \frac{1}{2} + \frac{\lambda_{10}}{2} + \frac{2\lambda_{01}}{3} + \frac{2\lambda_{11}}{3} \right\} k(\boldsymbol{\lambda}).$$

If $r_1 = 1$ and $r_2 = 1$ in (2.59), we get

$$E(Y_1Y_2) = \left\{\frac{1}{4} + \frac{\lambda_{10}}{3} + \frac{\lambda_{01}}{3} + \frac{4\lambda_{11}}{9}\right\} k(\boldsymbol{\lambda}).$$

The corresponding sample moments are $\bar{y}_1 = \frac{\sum_{i=1}^{n} y_{1i}}{n}$, $\bar{y}_2 = \frac{\sum_{i=1}^{n} y_{2i}}{n}$ and $\overline{y_1y_2} = \frac{\sum_{i=1}^{n} y_{1i}y_{2i}}{n}$. Now equating the population moments with the samples moments and solving it analytically, we otain the moment estimates as,

$$\lambda_{01}^{*} = \frac{3(2 - 3\bar{y}_1 + 6\overline{y_1y_2} - 4\bar{y}_2)}{2(-4 + 6\bar{y}_1 - 9\overline{y_1y_2} + 6\bar{y}_2)}, \tag{2.78}$$

$$\lambda_{10}^{*} = \frac{3(2 - 4\bar{y}_1 + 6\overline{y_1y_2} - 3\bar{y}_2)}{2(-4 + 6\bar{y}_1 - 9\overline{y_1y_2} + 6\bar{y}_2)}, \tag{2.79}$$

$$\lambda_{11}^{*} = \frac{3(-1 + 2\bar{y}_1 - 4\overline{y_1y_2} + 2\bar{y}_2)}{2(-4 + 6\bar{y}_1 - 9\overline{y_1y_2} + 6\bar{y}_2)}. \tag{2.80}$$

2.6.1.2 Bivariate distributions with transmuted normal conditionals

The general cross moment $E[Y_1^{r_1}Y_2^{r_2}]$ is given by

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} k[\boldsymbol{\lambda}] y_1^{r_1} y_2^{r_2} \phi(y_1)\phi(y_2) \left(1 + 2\lambda_{10}\Phi(y_1) + 2\lambda_{01}\Phi(y_2) + 4\lambda_{11}\Phi(y_1)\Phi(y_2)\right) dy_1 dy_2,$$

where $\phi(.)$ and $\Phi(.)$ are respectively, the probability density function and the cumulative distribution function of the standard normal distribution. In this case the cross moments are functions of the Gauss hypergeometric function ${}_2F_1$, and the moment equations must be solved numerically.

2.6.1.3 Bivariate distributions with transmuted exponential conditionals

The cross moments for the bivariate distribution with transmuted exponential conditionals is given in (2.69). By considering $r_1 = 1$ and $r_2 = 0$ in (2.69) we get,

$$E[Y_1] = \frac{k(\boldsymbol{\lambda})}{\beta_1} \left[1 + 3\lambda_{10} + 2\lambda_{01} + 6\lambda_{11}\right]. \tag{2.81}$$

If $r_1 = 0$ and $r_2 = 1$ in (2.69) we get,

$$E[Y_2] = \frac{k(\boldsymbol{\lambda})}{\beta_2}\left[1 + 2\lambda_{10} + 3\lambda_{01} + 6\lambda_{11}\right]. \tag{2.82}$$

If $r_1 = 2$ and $r_2 = 0$ in (2.69) we get,

$$E[Y_1^2] = \frac{k(\boldsymbol{\lambda})}{\beta_1^2}\left[1 + \frac{7}{2}\lambda_{10} + 2\lambda_{01} + 7\lambda_{11}\right]. \tag{2.83}$$

If $r_1 = 0$ and $r_2 = 2$ in (2.69) we get,

$$E[Y_2^2] = \frac{k(\boldsymbol{\lambda})}{\beta_2^2}\left[1 + 2\lambda_{10} + \frac{7}{2}\lambda_{01} + 7\lambda_{11}\right]. \tag{2.84}$$

If $r_1 = 1$ and $r_2 = 1$ in (2.69) we get,

$$E[Y_1Y_2] = \frac{k(\boldsymbol{\lambda})}{\beta_1\beta_2}\left(1 + 3\lambda_{10} + 3\lambda_{01} + 9\lambda_{11}\right). \tag{2.85}$$

$$\begin{aligned} V[Y_1] &= E[Y_1^2] - [E(Y_1)]^2 \\ &= \frac{[k(\boldsymbol{\lambda})]^2}{\beta_1^2}\left[1 + 5\lambda_{10}(2 + \lambda_{10}) + 18\lambda_{11} + 6\lambda_{10}\lambda_{11} - 8\lambda_{11}^2 + \lambda_{01}(2 + 6\lambda_{10} + 8\lambda_{11})\right], \end{aligned} \tag{2.86}$$

$$\begin{aligned} V[Y_2] &= E[Y_2^2] - [E(Y_2)]^2 \\ &= \frac{[k(\boldsymbol{\lambda})]^2}{\beta_2^2}\left[1 + 5\lambda_{01}^2 + 2\lambda_{11}(9 - 4\lambda_{11}) + 2\lambda_{01}(5 + 3\lambda_{10} + 3\lambda_{11}) + \lambda_{10}(2 + 8\lambda_{11})\right], \end{aligned} \tag{2.87}$$

and

$$\begin{aligned} Cov[Y_1, Y_2] &= E[Y_1Y_2] - E[Y_1]E[Y_2] \\ &= k(\boldsymbol{\lambda})\left(\frac{1 + 3\lambda_{01} + 3\lambda_{10} + 9\lambda_{11}}{\beta_1\beta_2}\right) \\ &\quad - \frac{[k(\boldsymbol{\lambda})]^2}{\beta_1\beta_2}\left(1 + 3\lambda_{01} + 2\lambda_{10} + 6\lambda_{11}\right)\left(1 + 2\lambda_{01} + 3\lambda_{10} + 6\lambda_{11}\right). \end{aligned} \tag{2.88}$$

Solving the equations from (2.81)-(2.88) for λ_{10}, λ_{01}, λ_{11}, β_1 and β_2 by equating with corresponding samples moments we obtain the moment estimates numerically.

2.6.1.4 Bivariate distributions with transmuted Weibull conditionals

The cross moments for the bivariate distribution with transmuted Weibull conditionals is given in (2.74). By considering $r_1 = 1$ and $r_2 = 0$ in (2.74) we get,

$$E[Y_1] = k(\boldsymbol{\lambda})\Gamma\left(\frac{1+\gamma_1}{\gamma_1}\right)2^{\frac{-1}{\gamma_1}} \times \left[2^{\frac{1}{\gamma_1}} + 2^{\frac{1}{\gamma_1}}\lambda_{01} + \lambda_{10}\left(2^{\frac{1+\gamma_1}{\gamma_1}} - 1\right) - \lambda_{11}\left(1 - 2^{2+\frac{1}{\gamma_1}} + 2^{\frac{1+\gamma_1}{\gamma_1}}\right)\right]. \tag{2.89}$$

If $r_1 = 0$ and $r_2 = 1$ in (2.74) we get,

$$E[Y_2] = k(\boldsymbol{\lambda})\Gamma\left(\frac{1+\gamma_2}{\gamma_2}\right)2^{\frac{-1}{\gamma_2}} \times \left[2^{\frac{1}{\gamma_2}} + 2^{\frac{1}{\gamma_2}}\lambda_{10} + \lambda_{01}\left(2^{\frac{1+\gamma_2}{\gamma_2}} - 1\right) - \lambda_{11}\left(1 - 2^{2+\frac{1}{\gamma_2}} + 2^{\frac{1+\gamma_2}{\gamma_2}}\right)\right]. \tag{2.90}$$

If $r_1 = 2$ and $r_2 = 0$ in (2.74) we get,

$$E[Y_1^2] = k(\boldsymbol{\lambda})\Gamma\left(\frac{2+\gamma_1}{\gamma_1}\right)2^{\frac{-2}{\gamma_1}} \times \left[2^{\frac{2}{\gamma_1}} + 2^{\frac{2}{\gamma_1}}\lambda_{01} + \lambda_{10}\left(2^{\frac{2+\gamma_1}{\gamma_1}} - 1\right) - \lambda_{11}\left(1 - 2^{2+\frac{2}{\gamma_1}} + 2^{\frac{2+\gamma_1}{\gamma_1}}\right)\right]. \tag{2.91}$$

If $r_1 = 0$ and $r_2 = 2$ in (2.74) we get,

$$E[Y_2^2] = k(\boldsymbol{\lambda})\Gamma\left(\frac{2+\gamma_2}{\gamma_2}\right)2^{\frac{-2}{\gamma_2}} \times \left[2^{\frac{2}{\gamma_2}} + 2^{\frac{2}{\gamma_2}}\lambda_{10} + \lambda_{01}\left(2^{\frac{2+\gamma_2}{\gamma_2}} - 1\right) - \lambda_{11}\left(1 - 2^{2+\frac{2}{\gamma_2}} + 2^{\frac{2+\gamma_2}{\gamma_2}}\right)\right]. \tag{2.92}$$

If $r_1 = 1$ and $r_2 = 1$ in (2.74) we get,

$$E[Y_1Y_2] = 2^{-\left(\frac{1}{\gamma_1}+\frac{1}{\gamma_2}\right)}k(\boldsymbol{\lambda})\Gamma\left(\frac{1+\gamma_1}{\gamma_1}\right)\Gamma\left(\frac{1+\gamma_2}{\gamma_2}\right) \times \left\{2^{\frac{1}{\gamma_1}+\frac{1}{\gamma_2}} + 2^{\frac{1}{\gamma_1}}\left(2^{\frac{1+\gamma_2}{\gamma_2}} - 1\right)\lambda_{01} + 2^{\frac{1}{\gamma_2}}\left(2^{\frac{1+\gamma_1}{\gamma_1}} - 1\right)\lambda_{10}\right\}$$

$$
\begin{aligned}
&+\left\{\lambda_{11}\left(1-2^{\frac{1+\gamma_1}{\gamma_1}}+2^{2+\frac{1}{\gamma_1}+\frac{1}{\gamma_2}}-2^{\frac{1+\gamma_2}{\gamma_2}}\right)\right\} \\
&\times 2^{-\left(\frac{1}{\gamma_1}+\frac{1}{\gamma_2}\right)} k(\boldsymbol{\lambda})\Gamma\left(\frac{1+\gamma_1}{\gamma_1}\right)\Gamma\left(\frac{1+\gamma_2}{\gamma_2}\right). \qquad (2.93)
\end{aligned}
$$

$$
\begin{aligned}
V[Y_1] &= E[Y_1^2]-(E[Y_1])^2 \\
&= -4^{\frac{-1}{\gamma_1}}[k(\boldsymbol{\lambda})]^2\Gamma\left(1+\frac{1}{\gamma_1}\right)^2\left[\lambda_{10}+\lambda_{11}-2^{\frac{1}{\gamma_1}}\left(1+\lambda_{01}+2\lambda_{10}+2\lambda_{11}\right)\right]^2 \\
&+4^{\frac{-1}{\gamma_1}}k(\boldsymbol{\lambda})\Gamma\left(\frac{2+\gamma_1}{\gamma_1}\right)\left[4^{\frac{1}{\gamma_1}}\left(1+\lambda_{01}+2\lambda_{10}+2\lambda_{11}\right)-\lambda_{10}+\lambda_{11}\right]. \qquad (2.94)
\end{aligned}
$$

$$
\begin{aligned}
V[Y_2] &= E[Y_2^2]-(E[Y_2])^2 \\
&= -4^{\frac{-1}{\gamma_2}}[k(\boldsymbol{\lambda})]^2\Gamma\left(1+\frac{1}{\gamma_2}\right)^2\left[\lambda_{01}+\lambda_{11}-2^{\frac{1}{\gamma_2}}\left(1+2\lambda_{01}+\lambda_{10}+2\lambda_{11}\right)\right]^2 \\
&+4^{\frac{-1}{\gamma_2}}k(\boldsymbol{\lambda})\Gamma\left(\frac{2+\gamma_2}{\gamma_2}\right)\left[4^{\frac{1}{\gamma_2}}\left(1+2\lambda_{01}+\lambda_{10}+2\lambda_{11}\right)-\lambda_{01}+\lambda_{11}\right]. \qquad (2.95)
\end{aligned}
$$

$$
\begin{aligned}
Cov(Y_1,Y_2) &= E[Y_1Y_2]-E[Y_1]E[Y_2] \\
&= 2^{-\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}}k(\boldsymbol{\lambda})\Gamma\left(1+\frac{1}{\gamma_1}\right)\Gamma\left(1+\frac{1}{\gamma_2}\right)\left[\lambda_{11}(1-2^{1+\frac{1}{\gamma_1}})-2^{\frac{1}{\gamma_1}}\lambda_{01}\right] \\
&-2^{-\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}}[k(\boldsymbol{\lambda})]^2\left(2^{\frac{1}{\gamma_2}}(1+2\lambda_{01}+\lambda_{10}+2\lambda_{11})-\lambda_{01}-\lambda_{11}\right) \\
&\times\left(2^{\frac{1}{\gamma_1}}(1+\lambda_{01}+2\lambda_{10}+2\lambda_{11})-\lambda_{10}-\lambda_{11}\right)\Gamma\left(1+\frac{1}{\gamma_1}\right)\Gamma\left(1+\frac{1}{\gamma_2}\right) \\
&+2^{-\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}}k(\boldsymbol{\lambda})2^{\frac{1}{\gamma_2}}\left(2^{\frac{1}{\gamma_1}}(1+2\lambda_{01}+2\lambda_{10}+4\lambda_{11})-\lambda_{10}-2\lambda_{11}\right) \\
&\times\Gamma\left(1+\frac{1}{\gamma_1}\right)\Gamma\left(1+\frac{1}{\gamma_2}\right). \qquad (2.96)
\end{aligned}
$$

Solving the equations from (2.89)-(2.96) for λ_{10}, λ_{01}, λ_{11}, γ_1 and γ_2 and applying the procedure explained in 2.6, we can obtain the moment estimates numerically. This has been explained in Section 2.7 and 2.8 by a simulation study and some real time applications.

2.6.2 Method of Maximum Likelihood

In this section we derive the maximum likelihood estimator (MLE) of the parameters of the $BTC(\boldsymbol{\lambda}; G_1, G_2)$ family defined in (2.3). Let $(y_{1i}, y_{2i}), i = 1, 2, ..., n$ be a random sample of size n from (2.3), where we assume that both baseline functions are $G_1(y_1; \boldsymbol{\lambda})$ and $G_2(y_2; \boldsymbol{\lambda})$. The log-likelihood function for $\boldsymbol{\lambda} = (\lambda_{10}, \lambda_{01}, \lambda_{11})$ may be written,

$$\begin{aligned} \ell(\boldsymbol{\lambda}) = n \log k(\boldsymbol{\lambda}) + \sum_{i=1}^{n} \log g_1(y_{1i}) + \sum_{i=1}^{n} \log g_2(y_{2i}) \\ + \sum_{i=1}^{n} \log [1 + 2\lambda_{10} G_1(y_{1i}) + 2\lambda_{01} G_2(y_{2i}) + 4\lambda_{11} G_1(y_{1i}) G_2(y_{2i})]. \end{aligned} \tag{2.97}$$

The general log-likelihood equations are given as follows:

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{10}} = \sum_{i=1}^{n} \left[\frac{2G_1(y_{1i})}{1 + 2\lambda_{10} G_1(y_{1i}) + 2\lambda_{01} G_2(y_{2i}) + 4\lambda_{11} G_1(y_{1i}) G_2(y_{2i})} \right] - nk(\boldsymbol{\lambda}) = 0, \tag{2.98}$$

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{01}} = \sum_{i=1}^{n} \left[\frac{2G_2(y_{2i})}{1 + 2\lambda_{10} G_1(y_{1i}) + 2\lambda_{01} G_2(y_{2i}) + 4\lambda_{11} G_1(y_{1i}) G_2(y_{2i})} \right] - nk(\boldsymbol{\lambda}) = 0, \tag{2.99}$$

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{11}} = \sum_{i=1}^{n} \left[\frac{4G_1(y_{1i}) G_2(y_{2i})}{1 + 2\lambda_{10} G_1(y_{1i}) + 2\lambda_{01} G_2(y_{2i}) + 4\lambda_{11} G_1(y_{1i}) G_2(y_{2i})} \right] - nk(\boldsymbol{\lambda}) = 0. \tag{2.100}$$

2.6.2.1 Bivariate distributions with transmuted uniform conditionals

The log-likelihood equation for the model given in (2.58) is defined as follows:

$$\ell(\boldsymbol{\lambda}) = \sum_{i=1}^{n} \log [1 + 2\lambda_{10} y_{1i} + 2\lambda_{01} y_{2i} + 4\lambda_{11} y_{1i} y_{2i}] + n \log k(\boldsymbol{\lambda}) \tag{2.101}$$

The log-likelihood can be maximized either directly or by solving the non-linear likelihood equations obtained by differentiating (2.101). The log-likelihood equations are given by

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{10}} = \sum_{i=1}^{n} \left[\frac{2y_{1i}}{1 + 2\lambda_{10}y_{1i} + 2\lambda_{01}y_{2i} + 4\lambda_{11}y_{1i}y_{2i}} \right] - nk(\boldsymbol{\lambda}), \tag{2.102}$$

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{01}} = \sum_{i=1}^{n} \left[\frac{2y_{2i}}{1 + 2\lambda_{10}y_{1i} + 2\lambda_{01}y_{2i} + 4\lambda_{11}y_{1i}y_{2i}} \right] - nk(\boldsymbol{\lambda}), \tag{2.103}$$

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{11}} = \sum_{i=1}^{n} \left[\frac{2y_{1i}y_{2i}}{1 + 2\lambda_{10}y_{1i} + 2\lambda_{01}y_{2i} + 4\lambda_{11}y_{1i}y_{2i}} \right] - nk(\boldsymbol{\lambda}) = 0. \tag{2.104}$$

We can find the estimates of the unknown parameters by maximum likelihood method by setting the above non-linear equations (2.101) - (2.104) to zero and solve them simultaneously. Therefore, we have to use mathematical software to get the MLE of the unknown parameters. Also, all the second order derivatives exist. Thus we have the inverse dispersion matrix is given by

$$\begin{pmatrix} \widehat{\lambda_{10}} \\ \widehat{\lambda_{01}} \\ \widehat{\lambda_{11}} \end{pmatrix} \sim N \left[\begin{pmatrix} \lambda_{10} \\ \lambda_{01} \\ \lambda_{11} \end{pmatrix}, \begin{pmatrix} \widehat{V_{\lambda_{10}\lambda_{10}}} & \widehat{V_{\lambda_{10}\lambda_{01}}} & \widehat{V_{\lambda_{10}\lambda_{11}}} \\ \widehat{V_{\lambda_{01}\lambda_{10}}} & \widehat{V_{\lambda_{01}\lambda_{01}}} & \widehat{V_{\lambda_{01}\lambda_{11}}} \\ \widehat{V_{\lambda_{11}\lambda_{10}}} & \widehat{V_{\lambda_{11}\lambda_{01}}} & \widehat{V_{\lambda_{11}\lambda_{11}}} \end{pmatrix} \right]. \tag{2.105}$$

Under the conditions that are fulfilled for parameters in the interior of the parameter space, but not on the boundary, the asymptotic distribution of the element of the 3 x 3 observed information matrix for the bivariate distributions with transmuted uniform conditionals is $\sqrt{n}(\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) \sim N_3(0, V^{-1})$, where V is the expected information matrix. Thus, the inverse of the expected information matrix is

$$V^{-1} = -E \begin{bmatrix} V_{\lambda_{10}\lambda_{10}} & V_{\lambda_{10}\lambda_{01}} & V_{\lambda_{10}\lambda_{11}} \\ & V_{\lambda_{01}\lambda_{01}} & V_{\lambda_{01}\lambda_{11}} \\ & & V_{\lambda_{11}\lambda_{11}} \end{bmatrix},$$

where

$$V_{\lambda_{10}\lambda_{10}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{10}^2} = n\,[k(\boldsymbol{\lambda})]^2 + \sum_{i=1}^{n} \left[\frac{-4y_{1i}^2}{(1 + 2\lambda_{10}y_{1i} + 2\lambda_{01}y_{2i} + 4\lambda_{11}y_{1i}y_{2i})^2} \right], \tag{2.106}$$

$$V_{\lambda_{01}\lambda_{01}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{01}^2} = n\,[k(\boldsymbol{\lambda})]^2 + \sum_{i=1}^{n} \left[\frac{-4y_{2i}^2}{(1 + 2\lambda_{10}y_{1i} + 2\lambda_{01}y_{2i} + 4\lambda_{11}y_{1i}y_{2i})^2} \right], \tag{2.107}$$

$$V_{\lambda_{11}\lambda_{11}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{11}^2} = n\left[k(\boldsymbol{\lambda})\right]^2 + \sum_{i=1}^{n}\left[\frac{-16y_{1i}^2 y_{2i}^2}{\left(1+2\lambda_{10}y_{1i}+2\lambda_{01}y_{2i}+4\lambda_{11}y_{1i}y_{2i}\right)^2}\right], \quad (2.108)$$

$$V_{\lambda_{10}\lambda_{01}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{10}\partial \lambda_{01}} = n\left[k(\boldsymbol{\lambda})\right]^2 + \sum_{i=1}^{n}\left[\frac{-4y_{1i}y_{2i}}{\left(1+2\lambda_{10}y_{1i}+2\lambda_{01}y_{2i}+4\lambda_{11}y_{1i}y_{2i}\right)^2}\right], \quad (2.109)$$

$$V_{\lambda_{10}\lambda_{11}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{10}\partial \lambda_{11}} = n\left[k(\boldsymbol{\lambda})\right]^2 + \sum_{i=1}^{n}\left[\frac{-8y_{1i}^2 y_{2i}}{\left(1+2\lambda_{10}y_{1i}+2\lambda_{01}y_{2i}+4\lambda_{11}y_{1i}y_{2i}\right)^2}\right], \quad (2.110)$$

$$V_{\lambda_{01}\lambda_{11}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{01}\partial \lambda_{11}} = n\left[k(\boldsymbol{\lambda})\right]^2 + \sum_{i=1}^{n}\left[\frac{-8y_{1i}y_{2i}^2}{\left(1+2\lambda_{10}y_{1i}+2\lambda_{01}y_{2i}+4\lambda_{11}y_{1i}y_{2i}\right)^2}\right]. \quad (2.111)$$

By solving this inverse dispersion matrix these solutions will yield asymptotic variance and covariances of these ML estimators for $\widehat{\lambda_{10}}, \widehat{\lambda_{01}}$ and $\widehat{\lambda_{11}}$. Using (2.105), we approximate $100(1-\gamma)\%$ confidence intervals for $\lambda_{10}, \lambda_{01}$ and λ_{11} are determined respectively as

$$\widehat{\lambda_{10}} \pm z_{\frac{\gamma}{2}}\sqrt{\widehat{V_{\lambda_{10}\lambda_{10}}}},\ \widehat{\lambda_{01}} \pm z_{\frac{\gamma}{2}}\sqrt{\widehat{V_{\lambda_{01}\lambda_{01}}}},\ \text{and}\ \widehat{\lambda_{11}} \pm z_{\frac{\gamma}{2}}\sqrt{\widehat{V_{\lambda_{11}\lambda_{11}}}},$$

where z_γ is the upper 100_γ the percentile of the standard normal distribution.

2.6.2.2 Bivariate distributions with transmuted normal conditionals

The log-likelihood equation for the model given in (2.63) is defined as follows:

$$\begin{aligned}\ell(\boldsymbol{\lambda}) = n\log 4 + \sum_{i=1}^{n}\log\phi(x_i) + \sum_{i=1}^{n}\log\phi(y_i) + n\log k(\boldsymbol{\lambda}) \\ + \sum_{i=1}^{n}\log\left[1+2\lambda_{10}\Phi(x_i)+2\lambda_{01}\Phi(y_i)+4\lambda_{11}\Phi(x_i)\Phi(y_i)\right]. \quad (2.112)\end{aligned}$$

The log-likelihood can be maximized either directly or by solving the non-linear likelihood equations obtained by differentiating (2.112). The log-likelihood equations are given by

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{10}} = \sum_{i=1}^{n} \left[\frac{2\Phi(y_{1i})}{1 + 2\lambda_{10}\Phi(y_{1i}) + 2\lambda_{01}\Phi(y_{2i}) + 4\lambda_{11}\Phi(y_{1i})\Phi(y_{2i})} \right] - nk(\boldsymbol{\lambda}), \quad (2.113)$$

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{01}} = \sum_{i=1}^{n} \left[\frac{2\Phi(y_{2i})}{1 + 2\lambda_{10}\Phi(y_{1i}) + 2\lambda_{01}\Phi(y_{2i}) + 4\lambda_{11}\Phi(y_{1i})\Phi(y_{2i})} \right] - nk(\boldsymbol{\lambda}), \quad (2.114)$$

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{11}} = \sum_{i=1}^{n} \left[\frac{4\Phi(y_{1i})\Phi(y_{2i})}{1 + 2\lambda_{10}\Phi(y_{1i}) + 2\lambda_{01}\Phi(y_{2i}) + 4\lambda_{11}\Phi(y_{1i})\Phi(y_{2i})} \right] - nk(\boldsymbol{\lambda}), \quad (2.115)$$

Given the observed data, (y_{1i}, y_{2i}) , $i = 1, 2, \ldots, n$, we wish to find the value of $\boldsymbol{\lambda}$ that maximizes $\ell(\boldsymbol{\lambda})$.

2.6.2.3 Bivariate distributions with transmuted Weibull conditionals

The log-likelihood equation for the model given in (2.3) with one-parameter Weibull marginals is defined as follows:

$$\ell(\boldsymbol{\lambda}) = \sum\nolimits_{i=1}^{n} \log \left[1 + 2\lambda_{10}\left(1 - e^{-y_{1i}^{\gamma_1}}\right) + 2\lambda_{01}\left(1 - e^{-y_{2i}^{\gamma_2}}\right) + 4\lambda_{11}\left(1 - e^{-y_{1i}^{\gamma_1}}\right)\left(1 - e^{-y_{2i}^{\gamma_2}}\right) \right]$$
$$+ n\log\gamma_1 + n\log\gamma_2 - \sum_{i=1}^{n} y_{1i}^{\gamma_1} - \sum_{i=1}^{n} y_{2i}^{\gamma_2} + n\log k(\boldsymbol{\lambda}), \quad (2.116)$$

where $\boldsymbol{\lambda} = (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$, being $\boldsymbol{\tau}_1 = (\lambda_{10}, \lambda_{01}, \lambda_{11})$ and $\boldsymbol{\tau}_2 = (\gamma_1, \gamma_2)$.

The log-likelihood can be maximized either directly or by solving the non-linear likelihood equations obtained by differentiating (2.116). The components of the score vector are given by

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{10}} = \sum_{i=1}^{n} \left[\frac{2\left(1 - e^{-y_{1i}^{\gamma_1}}\right)}{1 + 2\lambda_{10}\left(1 - e^{-y_{1i}^{\gamma_1}}\right) + 2\lambda_{01}\left(1 - e^{-y_{2i}^{\gamma_2}}\right) + 4\lambda_{11}\left(1 - e^{-y_{1i}^{\gamma_1}}\right)\left(1 - e^{-y_{2i}^{\gamma_2}}\right)} \right]$$

$$-\, nk(\boldsymbol{\lambda}), \tag{2.117}$$

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{01}} = \sum_{i=1}^{n}\left[\frac{2\left(1-e^{-y_{2i}^{\gamma_2}}\right)}{1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)}\right] - nk(\boldsymbol{\lambda}), \tag{2.118}$$

$$\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \lambda_{11}} = \sum_{i=1}^{n}\left[\frac{4\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)}{1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)}\right] - nk(\boldsymbol{\lambda}), \tag{2.119}$$

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \gamma_1} &= \frac{n}{\gamma_1}+\sum_{i=1}^{n}\log y_{1i}-\sum_{i=1}^{n}\log[y_{1i}][y_{1i}]^{\gamma_1}\\ &+\sum_{i=1}^{n}\left[\frac{2\lambda_{10}e^{-y_{1i}^{\gamma_1}}\log[y_{1i}][y_{1i}]^{\gamma_1}+4\lambda_{11}e^{-y_{1i}^{\gamma_1}}\log[y_{1i}][y_{1i}]^{\gamma_1}\left(1-e^{-y_{2i}^{\gamma_2}}\right)}{1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)}\right],\end{aligned} \tag{2.120}$$

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\lambda})}{\partial \gamma_2} &= \frac{n}{\gamma_2}+\sum_{i=1}^{n}\log y_{2i}-\sum_{i=1}^{n}\log[y_{2i}][y_{2i}]^{\gamma_2}\\ &+\sum_{i=1}^{n}\left[\frac{2\lambda_{01}e^{-y_{2i}^{\gamma_2}}\log[y_{2i}][y_{2i}]^{\gamma_2}+4\lambda_{11}e^{-y_{2i}^{\gamma_2}}\log[y_{2i}][y_{2i}]^{\gamma_2}\left(1-e^{-y_{1i}^{\gamma_1}}\right)}{1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)}\right].\end{aligned} \tag{2.121}$$

We can find the estimates of the unknown parameters by maximum likelihood method by setting the above non-linear equations (2.117) - (2.121) to zero and solve them simultaneously. Therefore, we have to use mathematical software to get the MLE of the unknown parameters. Also, all the second order derivatives exist. Thus we have the inverse dispersion matrix is given by

$$\begin{pmatrix} \widehat{\lambda_{10}} \\ \widehat{\lambda_{01}} \\ \widehat{\lambda_{11}} \\ \widehat{\gamma_1} \\ \widehat{\gamma_2} \end{pmatrix} \sim N \left[\begin{pmatrix} \lambda_{10} \\ \lambda_{01} \\ \lambda_{11} \\ \gamma_1 \\ \gamma_2 \end{pmatrix}, \begin{pmatrix} \widehat{V_{\lambda_{10}\lambda_{10}}} & \widehat{V_{\lambda_{10}\lambda_{01}}} & \widehat{V_{\lambda_{10}\lambda_{11}}} & \widehat{V_{\lambda_{10}\gamma_1}} & \widehat{V_{\lambda_{10}\gamma_2}} \\ \widehat{V_{\lambda_{01}\lambda_{10}}} & \widehat{V_{\lambda_{01}\lambda_{01}}} & \widehat{V_{\lambda_{01}\lambda_{11}}} & \widehat{V_{\lambda_{01}\gamma_1}} & \widehat{V_{\lambda_{01}\gamma_2}} \\ \widehat{V_{\lambda_{11}\lambda_{10}}} & \widehat{V_{\lambda_{11}\lambda_{01}}} & \widehat{V_{\lambda_{11}\lambda_{11}}} & \widehat{V_{\lambda_{11}\gamma_1}} & \widehat{V_{\lambda_{11}\gamma_2}} \\ \widehat{V_{\gamma_1\lambda_{10}}} & \widehat{V_{\gamma_1\lambda_{01}}} & \widehat{V_{\gamma_1\lambda_{11}}} & \widehat{V_{\gamma_1\gamma_1}} & \widehat{V_{\gamma_1\gamma_2}} \\ \widehat{V_{\gamma_2\lambda_{10}}} & \widehat{V_{\gamma_2\lambda_{01}}} & \widehat{V_{\gamma_2\lambda_{11}}} & \widehat{V_{\gamma_2\gamma_1}} & \widehat{V_{\gamma_2\gamma_2}} \end{pmatrix} \right], \tag{2.122}$$

Under the conditions that are fulfilled for parameters in the interior of the parameter space, but not on the boundary, the asymptotic distribution of the element of the 5 x 5 observed information matrix for the bivariate distributions with transmuted Weibull conditionals is $\sqrt{n}(\widehat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) \sim N_5(0, V^{-1})$, where V is the expected information matrix. Thus, the inverse of the expected information matrix is

$$V^{-1} = -E \begin{bmatrix} V_{\lambda_{10}\lambda_{10}} & V_{\lambda_{10}\lambda_{01}} & V_{\lambda_{10}\lambda_{11}} & V_{\lambda_{10}\gamma_1} & V_{\lambda_{10}\gamma_2} \\ & V_{\lambda_{01}\lambda_{01}} & V_{\lambda_{01}\lambda_{11}} & V_{\lambda_{01}\gamma_1} & V_{\lambda_{01}\gamma_2} \\ & & V_{\lambda_{11}\lambda_{11}} & V_{\lambda_{11}\gamma_1} & V_{\lambda_{11}\gamma_2} \\ & & & V_{\gamma_1\gamma_1} & V_{\gamma_1\gamma_2} \\ & & & & V_{\gamma_2\gamma_2} \end{bmatrix},$$

where

$$\begin{aligned} V_{\lambda_{10}\lambda_{10}} &= \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{10}^2} \\ &= \sum_{i=1}^{n} \left[\frac{4\left(1-e^{-y_{1i}^{\gamma_1}}\right)^2}{\left(1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)\right)^2} \right] \\ &\quad + n\left[k(\boldsymbol{\lambda})\right]^2, \end{aligned} \tag{2.123}$$

$$\begin{aligned} V_{\lambda_{01}\lambda_{01}} &= \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{01}^2} \\ &= \sum_{i=1}^{n} \left[\frac{4\left(1-e^{-y_{2i}^{\gamma_2}}\right)^2}{\left(1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)\right)^2} \right] \\ &\quad + n\left[k(\boldsymbol{\lambda})\right]^2, \end{aligned} \tag{2.124}$$

$$
\begin{aligned}
V_{\lambda_{11}\lambda_{11}} &= \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \lambda_{11}^2} \\
&= \sum_{i=1}^{n}\left[\frac{16\left(1-e^{-y_{1i}^{\gamma_1}}\right)^2\left(1-e^{-y_{2i}^{\gamma_2}}\right)^2}{\left(1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)\right)^2}\right] \\
&\quad + n\left[k(\boldsymbol{\lambda})\right]^2,
\end{aligned}
\tag{2.125}
$$

$$
\begin{aligned}
V_{\gamma_1\gamma_1} &= \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \gamma_1^2} \\
&= \frac{n}{\gamma_1^2} + \sum_{i=1}^{n} \log[y_{1i}]^2 y_{1i}^{\gamma_1} \\
&+ \sum_{i=1}^{n}\left[\frac{\left(2e^{-y_{1i}^{\gamma_1}}\log(y_{1i})y_{1i}^{\gamma_1}[\lambda_{10}+2\lambda_{11}(1-e^{-y_{2i}^{\gamma_2}})]\right)^2}{\left(1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)\right)^2}\right] \\
&+ \sum_{i=1}^{n}\left[\frac{\left(2e^{-y_{1i}^{\gamma_1}}\log(y_{1i})^2 y_{1i}^{\gamma_1}\right)\left(\lambda_{10}(1-y_{1i}^{\gamma_1})+2\lambda_{11}(1-e^{-y_{2i}^{\gamma_2}})(1-y_{1i}^{\gamma_1})\right)}{\left(1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)\right)}\right]
\end{aligned}
\tag{2.126}
$$

$$
\begin{aligned}
V_{\gamma_2\gamma_2} &= \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial \gamma_2^2} \\
&= \frac{n}{\gamma_2^2} + \sum_{i=1}^{n} \log[y_{2i}]^2 y_{2i}^{\gamma_2} \\
&+ \sum_{i=1}^{n}\left[\frac{\left(2e^{-y_{2i}^{\gamma_2}}\log(y_{2i})y_{2i}^{\gamma_2}[\lambda_{01}+2\lambda_{11}(1-e^{-y_{1i}^{\gamma_1}})]\right)^2}{\left(1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)\right)^2}\right] \\
&+ \sum_{i=1}^{n}\left[\frac{\left(2e^{-y_{2i}^{\gamma_2}}\log(y_{2i})^2 y_{2i}^{\gamma_2}\right)\left(\lambda_{01}(1-y_{2i}^{\gamma_2})+2\lambda_{11}(1-e^{-y_{1i}^{\gamma_1}})(1-y_{2i}^{\gamma_2})\right)}{\left(1+2\lambda_{10}\left(1-e^{-y_{1i}^{\gamma_1}}\right)+2\lambda_{01}\left(1-e^{-y_{2i}^{\gamma_2}}\right)+4\lambda_{11}\left(1-e^{-y_{1i}^{\gamma_1}}\right)\left(1-e^{-y_{2i}^{\gamma_2}}\right)\right)}\right]
\end{aligned}
\tag{2.127}
$$

Similarly, we can find the other elements by using the following equations:

$$V_{\lambda_{10}\lambda_{01}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{10}\partial\lambda_{01}},\ V_{\lambda_{10}\lambda_{11}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{10}\partial\lambda_{11}},\ V_{\lambda_{01}\lambda_{11}} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{01}\partial\lambda_{11}},\ V_{\lambda_{10}\gamma_1} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{10}\partial\gamma_1},$$
$$V_{\lambda_{10}\gamma_2} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{10}\partial\gamma_2},\ V_{\lambda_{01}\gamma_1} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{01}\partial\gamma_1},\ V_{\lambda_{01}\gamma_2} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{01}\partial\gamma_2},\ V_{\lambda_{11}\gamma_1} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{11}\partial\gamma_1},$$
$$V_{\lambda_{11}\gamma_2} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\lambda_{11}\partial\gamma_2},\ V_{\gamma_1\gamma_2} = \frac{\partial^2 \ell(\boldsymbol{\lambda})}{\partial\gamma_1\partial\gamma_2}.$$

By solving this inverse dispersion matrix these solutions will yield asymptotic variance and covariances of these ML estimators for $\widehat{\lambda_{10}}, \widehat{\lambda_{01}}, \widehat{\lambda_{01}}, \widehat{\gamma_1}$ and $\widehat{\gamma_2}$. Using (2.122), we approximate $100(1-\gamma)\%$ confidence intervals for $\lambda_{10}, \lambda_{01}$, γ_1 and γ_2 are determined respectively as

$$\widehat{\lambda_{10}} \pm z_{\frac{\gamma}{2}}\sqrt{\widehat{V_{\lambda_{10}\lambda_{10}}}},\ \widehat{\lambda_{01}} \pm z_{\frac{\gamma}{2}}\sqrt{\widehat{V_{\lambda_{01}\lambda_{01}}}},\ \widehat{\lambda_{11}} \pm z_{\frac{\gamma}{2}}\sqrt{\widehat{V_{\lambda_{11}\lambda_{11}}}},\ \widehat{\beta_1} \pm z_{\frac{\gamma}{2}}\sqrt{\widehat{V_{\gamma_1\gamma_1}}}, \text{ and } \widehat{\gamma_2} \pm z_{\frac{\gamma}{2}}\sqrt{\widehat{V_{\gamma_2\gamma_2}}},$$

where z_γ is the upper 100_γ the percentile of the standard normal distribution.

We describe an effective profile likelihood approach. The model in (2.68) and (2.73) are with parameters $\boldsymbol{\tau}_1$ and $\boldsymbol{\tau}_2$ where $\boldsymbol{\tau}_1$ is the vector transmuted parameters and $\boldsymbol{\tau}_2$ is the vector baseline parameters. Denote $\ell(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$ as the log-likelihood function. For each value of $\boldsymbol{\tau}_1$, $\ell_1(\boldsymbol{\tau}_1)$ is the maximum of the log-likelihood function over the remaining parameters. The profile likelihood function for $\boldsymbol{\tau}_1$ is

$$\ell_1(\boldsymbol{\tau}_1) = \max_{\boldsymbol{\tau}_2} \ell(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2).$$

This maximization is done numerically and the procedure is explained as follows. Let $\tilde{\boldsymbol{\lambda}} = (\tilde{\boldsymbol{\tau}}_1, \tilde{\boldsymbol{\tau}}_2)$ where $\tilde{\boldsymbol{\tau}}_1 = (\lambda_{10}, \lambda_{01}, \lambda_{11})$, $\tilde{\boldsymbol{\tau}}_2 = (\gamma_1, \gamma_2)$ (for transmuted Weibull conditionals). In the first stage, we estimate $\tilde{\boldsymbol{\tau}}_1$ by maximizing the profile likelihood of $\tilde{\boldsymbol{\tau}}_1$ and once, an estimate of $\tilde{\boldsymbol{\tau}}_1$ is obtained, the estimates of $\tilde{\boldsymbol{\tau}}_2$ can be obtained by substituting the estimates of $\tilde{\boldsymbol{\tau}}_1$. We set the moment estimates as the initial values. This process is continued iteratively till all the estimates converge to yield the MLE $\widehat{\tilde{\boldsymbol{\lambda}}}$ of $\tilde{\boldsymbol{\lambda}}$. The computation is carried out using "FindMaximum" function of Mathematica 10. Section 2.7 presents a detailed simulation study to illustrate the estimation approach and Section 2.8 presents three illustrations using a real time data sets.

2.7 Simulation Study

We carried out a simulation study in order to evaluate the performance of the profile likelihood estimation. For simplicity we considered the bivariate model with transmuted Weibull conditionals given in (2.73). Sample generation of (y_{1i}, y_{2i}), $i = 1, 2, \ldots, n$ was carried out by using the algorithm given in Section 2.4. We generated 1000 samples of sizes $n = 25$, $n = 75$, and $n = 150$ with $\lambda_{10} = 0.8$, $\lambda_{01} = 0.9$, $\lambda_{11} = 1.1$, $\gamma_1 = 1.9$, and $\gamma_2 = 1.9$. The MLEs were obtained using the procedure described in Section 2.6 and the average bias across the 1000 samples was computed. The average root mean square error (RMSE) from the 1000 samples was calculated as $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{\tilde{\boldsymbol{\lambda}}}_i - \tilde{\boldsymbol{\lambda}}_i)}$. The approximate variance-covariance matrix of the MLEs was obtained as the inverse of the observed information matrix. The biases, RMSEs are provided in Tables 2.1. We observed from the simulation study that the biases and RMSE's decrease as the sample sizes increase. We also observed that the rate of convergence improved with increasing sample size. The simulation study shows that, for the given set of true parameter values, the Maximum Likelihood estimators perform well comparing to the moment estimators. Figure 2.5 shows the stability graph of simulated parameters estimates for 25 iterations for a single simulated sample.

Table 2.1: Absolute Bias, RMSE, for $\lambda_{10}, \lambda_{01}, \lambda_{11}, \gamma_1$ and γ_2 of Bivariate distributions with transmuted Weibull conditionals based on 1000 replications by method of moments estimation (MME) and method of maximum likelihood estimation (MLE) for a simulated data

	MME					MLE				
	λ_{10}	λ_{01}	λ_{11}	γ_1	γ_2	λ_{10}	λ_{01}	λ_{11}	γ_1	γ_2
True Values	0.8	0.9	1.1	1.9	1.9	0.8	0.9	1.1	1.9	1.9
n=25										
Absolute Bias	0.4475	0.4745	0.3664	0.0665	0.0773	0.3396	0.2600	0.1093	0.0548	0.0598
RMSE	0.2790	0.3389	0.3605	0.2998	0.3190	0.0714	0.1036	0.0988	0.0407	0.0568
n=75										
Absolute Bias	0.4027	0.3972	0.3514	0.0605	0.0667	0.1956	0.2241	0.0966	0.0526	0.0536
RMSE	0.2437	0.2324	0.3424	0.1596	0.1633	0.0501	0.0680	0.0896	0.0326	0.0352
n=150										
Absolute Bias	0.3018	0.3153	0.3236	0.0269	0.0297	0.1283	0.1502	0.0837	0.0240	0.0267
RMSE	0.1355	0.2295	0.1534	0.1168	0.1135	0.0209	0.0295	0.0439	0.0047	0.0054

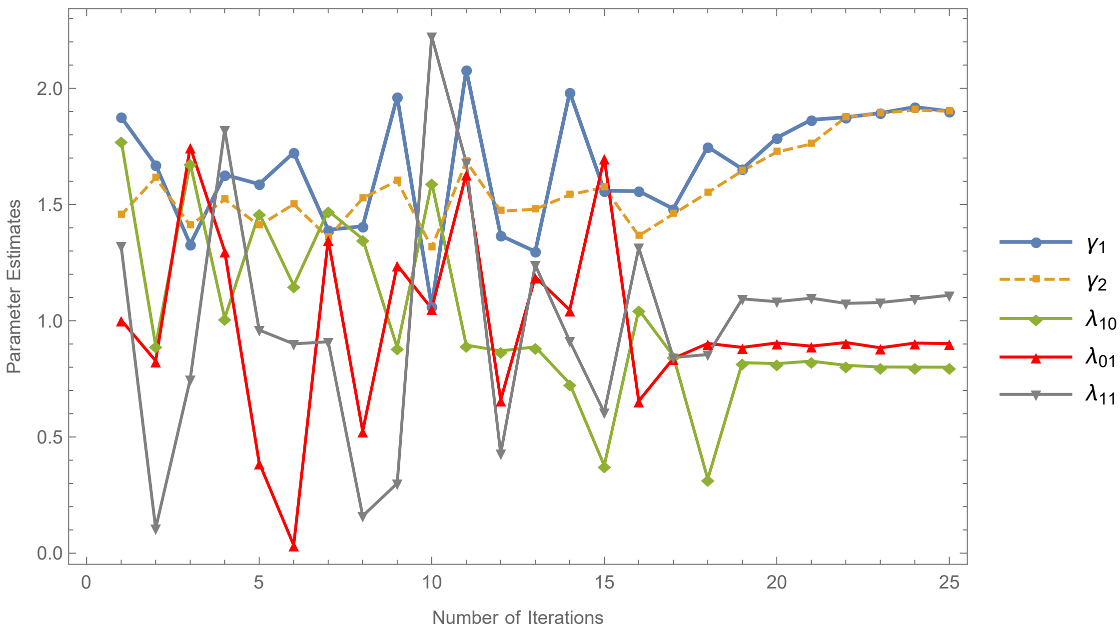


Figure 2.5: The stability graph for the simulated parameters for maximum likelihood estimates of bivariate distributions with transmuted Weibull conditionals for 25 iterations.

2.8 Data Analysis

2.8.1 Data Set 1: Reliability Analysis of Cable Insulation Specimens

We demonstrate the application of Bivariate distributions with transmuted Weibull conditionals for a reliability data. The reliability data set consists of relative failure of 20 epoxy electrical cable insulation specimens that worked under 55 kilovolts voltage conditions constantly. The data was originally reported by Stone (1978). Lawless (2011) described the failure phenomenon called electrical treeing. In this process there is considered to be an inception or initiation period in which it looks as if nothing is happening under the microscopic scanner, but after some point in time there appears a tiny defect in the material and then onwards the defect grows and eventually causing the failure of the insulation. The observed data consist of two variables - one is time to inception of the defect Y_1 (in minutes) and second is the subsequent additional time to specimen failure Y_2 (in minutes). Pulcini (2006) further analyzed the data using a bivariate distribution with gamma conditionals in the context of forewarning or primer event and applied the maximum likelihood method. The data set has been presented in Table 2.2

Table 2.2: Failure time data (in minutes) of cable insulation specimens

Specimen	Inception Time(Y_1)	Additional Time(Y_2)
1	228	30
2	106	8
3	246	66
4	700	72
5	473	25
6	155	7
7	414	30
8	1374	90
9	128	4
10	1227	39
11	254	46
12	435	85
13	1155	85
14	195	27
15	117	27
16	724	21
17	300	96

We estimated the parameters using the procedure explained in Section 2.6. The computation is carried out using "FindMaximum" and "FindRoot" function of Mathematica 10. We compared this fitted model with the fitted bivariate distribution with gamma conditionals proposed by Pulcini (2006). The estimates with S.E., 95% confidence interval and AIC values are provided in Table 2.3. From the AIC values we infer that bivariate distribution with transmuted Weibull conditionals is a better model for fitting cable insulation specimen data. We fitted the marginals of the bivariate distribution with transmuted Weibull conditionals and the Kolmogorov-Smirnov (K-S) test revealed that both the marginals give good fits for cable insulation specimen data. For the marginal, Y_1, the K-S test statistic is 0.3713 and for the marginal, Y_2, the value is 0.3768 and we accept the null-hypothesis that the model given in (2.73) fits well for the cable insulation specimen data at 0.01 level of significance since $D_{0.01,17} = 0.381$.

Table 2.3: Estimates for bivariate distributions with transmuted conditionals for cable insulation specimens in data set 1

Model	Estimates					
$BTC(\boldsymbol{\lambda}, 1-e^{-z_i^{\gamma_i}})$	$\widehat{\lambda_{10}}$	$\widehat{\lambda_{01}}$	$\widehat{\lambda_{11}}$	$\widehat{\gamma_1}$	$\widehat{\gamma_2}$	AIC
Moment Estimate	-0.4265	-0.4267	0.2043	0.0006	0.0005	
ML Estimate	0.6547	0.9315	1.4369	1.0923	1.0783	90.28
S.E	0.3275	0.5828	0.4140	0.0430	0.0412	
95% Lower CL	0.0128	-0.2108	0.6255	1.0080	0.9975	
95% Upper CL	1.2966	2.0738	2.2483	1.1766	1.1591	
Pulcini (2006)	$\widehat{a}$	$\widehat{b}$	$\widehat{p}$	$\widehat{c}$	$\widehat{d}$	AIC
Moment Estimate	1.470	0.0030	0.9993	3.419	0.1106	
ML Estimate	1.680	0.0035	0.9992	1.982	0.0710	632.46
S.E	0.1281	0.0003	0.0001	0.1533	0.0074	
95% Lower CL	1.0020	0.0019	0.9984	1.1730	0.0351	
95% Upper CL	2.816	0.0063	0.9999	3.348	0.1437	

2.8.2 Data Set 2: Reliability Analysis of Two-component Parallel Systems

The second example that we consider to show our model application is the data set given in Murthy et al. (2004). The data set consists of 9 two-component systems connected in parallel. Let Y_1 be the failure time of component A and Y_2 be the failure time of component B. The data set is presented in Table 2.4. Pulcini (2006) further analyzed the data set by considering bivariate distributions with Gamma conditionals. We analyzed the data set using our model and estimated the parameters by the procedure explained in Section 2.6. The computation is carried out using

Table 2.4: Failure time data (in minutes) of two-component parallel systems

Specimen	Component A (Y_1)	Component B (Y_2)
1	77.2	156.6
2	74.3	108.0
3	9.6	12.4
4	251.6	108.0
5	134.9	84.1
6	115.7	51.2
7	195.7	289.8
8	42.2	59.1
9	27.8	35.5

"FindMaximum" and "FindRoot" function of Mathematica 10. We compared this fitted model with the fitted bivariate distribution with gamma conditionals proposed by Pulcini (2006). The estimates with S.E., 95% confidence interval and AIC values are provided in Table 2.5. From the AIC values we infer that bivariate distribution with transmuted Weibull conditionals is a better model for fitting two-component parallel systems. We fitted the marginals of the bivariate distribution with transmuted Weibull conditionals and the Kolmogorov-Smirnov (K-S) test revealed that both the marginals give good fits for the two-component parallel systems. For the marginal, Y_1, the K-S test statistic is 0.5019 and for the marginal, Y_2, the value is 0.5031 and we accept the null-hypothesis that the model given in (2.73) fits well for the cable insulation specimen data at 0.01 level of significance since $D_{0.01,9} = 0.513$.

2.9 Discussion and Summary

The model proposed in (2.3) is a general and a rich class of bivariate distributions with transmuted conditionals. It is interesting to note that the marginals are also transmuted. Profile likelihood method is applied for estimating the parameters involved in the model. Simulation study reveals that maximum likelihood estimation method should be preferred to method of moments in estimating the parameters.

From the data analysis it is observed that the bivariate distribution with trans-

Table 2.5: Estimates for bivariate distributions with transmuted conditionals for two-component parallel systems in data set 2

Model	Estimates					
$BTC(\boldsymbol{\lambda}, 1-e^{-z_i^{\gamma_i}})$	$\widehat{\lambda_{10}}$	$\widehat{\lambda_{01}}$	$\widehat{\lambda_{11}}$	$\widehat{\gamma_1}$	$\widehat{\gamma_2}$	AIC
Moment Estimate	-0.3487	-0.3482	0.0223	0.0332	0.0301	
ML Estimate	0.8922	0.8962	1.8530	0.1921	0.1928	263.12
S.E	0.6359	0.5753	0.4573	0.0125	0.0125	
95% Lower CL	-0.3542	-0.2314	0.9567	0.1676	0.1683	
95% Upper CL	2.1386	2.0238	2.7493	0.2166	0.2173	
Pulcini (2006)	$\widehat{a}$	$\widehat{b}$	$\widehat{p}$	$\widehat{c}$	$\widehat{d}$	AIC
Moment Estimate	2.088	0.0280	0.9922	6.599	0.239	
ML Estimate	1.964	0.0264	0.9825	1.836	0.156	366.92
S.E	0.2863	0.0044	0.0025	0.266	0.039	
95% Lower CL	0.956	0.0116	0.9703	0.898	0.045	
95% Upper CL	4.032	0.0598	0.9948	3.753	0.538	

muted Weibull conditionals is a better model comparing to bivariate distribution with gamma conditionals (Pulcini (2006)). The claim is well supported by AIC. The Kolmogorov-Smirnov test revealed that both the marginals for bivariate distribution with transmuted Weibull conditionals fit well for the two data sets.

Remark 2.9.1. *The data set in Example 2 consists of two-component systems connected in parallel. The K-S test revealed that the data set fits well for our model given in* (2.73). *When modelling with transmuted conditionals we have not taken into consideration the different characteristics of the system, for example the dependence structure between components. In particular, for Example 2 we could suspect a load* *sharing effect, where failure of one component affects the stochastic behaviour of the surviving component. To check whether there is any load sharing effect on the system, a nonparametric test procedure proposed by Deshpande et al. (2010) is performed. Consider an appeal sample* $(Y_{1i}, Y_{2i}),\ i = 1, 2, ..., n$.

Here the test statistic

$$U = \frac{1}{\binom{n}{2}} \sum_{i,j=1,(i<j)}^{n} \frac{1}{2}[h(Y_{1i}, Y_{2j}) + h(Y_{1j}, Y_{2i})] = \left(\frac{1}{36}\right)\frac{33}{2},$$

and the value of the test statistic is

$$Z = \frac{U - E(U)}{\sigma_U} = -5.3364,$$

where,

$$\begin{aligned} E(U) &= \frac{3m-1}{2(2m-1)} = \frac{5}{6}, \\ \sigma_U^2 &= \frac{4m(3m-1)\left(-3m^2+6m-1\right)(m-1)}{3m-2} + \frac{4(3m-1)\left(4m^2-3m+1\right)}{2m-1} \\ &- \frac{8m^2(m+1)(3m-1)\left(m^2-6m+1\right)}{2m-1} + \frac{4}{3}(m-1)^2 \\ &+ \frac{4}{3}(m+1)(m-1)^2 - \frac{4(3m-1)^2}{4(2m-1)^2} + \frac{4m}{3m-2} = \frac{2}{45}, \end{aligned}$$

where m *is the number of components in the system. In the present context, we take* $m = 2$. *Thus the null hypothesis that there is no load sharing effect is rejected at 5% significance level in favour of the alternative hypothesis indicating a load sharing effect.*

Hence it is imperative to take into account this feature of the system while modelling its lifetime. Accordingly in the next chapter we model bivariate lifetimes accommodating the dependence behaviours enjoyed by them. In particular we consider the load sharing behaviour of the system. In the rest of the thesis we fully concentrate on modelling of load sharing systems and its related areas.

Chapter 3

Modelling Load Share Data with Shared Frailty

3.1 Introduction

In a load sharing system, the probability of failure of any component will depend on the working status of the other components (Kvam & Lu (2007)). There are many situations in practice where the failure of a unit could redistribute the workload of the other operating units in the system, thus potentially increasing the failure rate of the operating units. Load sharing was first discussed in stress-strength models for fibre bundles (Daniels (1945))). Statistical models for studying times to failure in load sharing systems are valuable in many application areas including biomedical studies, manufacturing, material testing, software reliability, etc. These models characterize and estimate the mechanism of load change after a component within the system fails. In particular, the basic assumption in a two-component load sharing system is that while the system can function even after one of the components has failed, the failure of the component may put additional load on the surviving component and this affects the functioning of the system due to stochastic changes in its residual lifetime. In most situations, an increased load results in a higher component failure rate (Liu

*Some of the results of this chapter are published in Applied Stochastic Models in Business and Industry. (Asha et al. (2017))

(1998)). Examples of such systems include (a) a twin engine aircraft like Boeing's 777 (Singpurwalla (1995)), or (b) mechanical systems ("Reliability in Engineering Design" (n.d.)), or (c) paired organs like eyes, kidneys or lungs (Daniels (1945)), to mention a few.

Following the pioneering research on load sharing models by Daniels (1945), there has been considerable research activity in this area; see the excellent review article by Dewan & Naik-Nimbalkar (2010). Freund's bivariate exponential distribution Freund (1961) is an effective model for load sharing systems. In a two-component load sharing system, suppose T_1 and T_2 are non-negative random variables representing the lifetimes of two-components A and B respectively, when they are first put to test. If component B fails before A does, i.e., if $T_2 < T_1$, the lifetime distribution of A changes, and suppose we denote the failure time by T_1^*. Eventually, the system fails when component A also fails, and we observe the random bivariate failure times (T_1^*, T_2) where $T_1^* > T_2$. On the other hand, if A fails before B so that $T_1 < T_2$, the lifetime distribution of B changes and its failure time is denoted by T_2^*, say. The system fails when component B fails eventually, and one finally observes the bivariate random variables (T_1, T_2^*). To set notation, if we denote the lifetimes of the components A and B as the non-negative random variables (Y_1, Y_2), then one observes $Y_1 = T_1^*$, $Y_2 = T_2$, if $Y_1 > Y_2$, and $Y_1 = T_1$, $Y_2 = T_2^*$, if $Y_1 < Y_2$.

Assume that T_1 and T_2 are independent and have exponential distributions with respective failure rates θ_1 and θ_2, $\theta_i > 0,\ i = 1, 2$. It is further assumed that T_1^* and T_2^* also have exponential distributions with respective failure rates θ_1' and θ_2', $\theta_i' > 0,\ i = 1, 2$. The joint probability density function of (Y_1, Y_2) is (Freund (1961))

$$f(y_1, y_2) = \begin{cases} \theta_1'\theta_2 e^{-\theta_1' y_1} e^{-(\theta_1+\theta_2-\theta_1')y_2}; & y_1 > y_2 \\ \theta_1\theta_2' e^{-(\theta_1+\theta_2-\theta_2')y_1} e^{-\theta_2' y_2}; & y_2 > y_1,\ y_i > 0,\ i = 1, 2. \end{cases} \tag{3.1}$$

An extension to a model with Weibull component lifetime distributions is discussed in Lu (1989), Spurrier & Weier (1981) and Shaked (1984). In Asha et al. (2016) generalization of (3.1) is proposed as follows. Let $S(.)$ and $r(.)$ respectively denote the baseline survival and hazard functions. They assumed that T_1 and T_2 are independently distributed with respective survival functions $[S(.)]^{\theta_1}$, $\theta_1 > 0$ and $[S(.)]^{\theta_2}$, $\theta_2 > 0$. Again, T_1^* and T_2^* are assumed to have survival functions

$[S(.)]^{\theta_1'}$, $\theta_1' > 0$ and $[S(.)]^{\theta_2'}$, $\theta_2' > 0$, respectively. The joint probability density function of the failure times (Y_1, Y_2) under the generalized model is

$$f(y_1, y_2) = \begin{cases} \theta_1'\theta_2 f(y_1)f(y_2)[S(y_2)]^{(\theta_1+\theta_2-\theta_1'-1)}[S(y_1)]^{\theta_1'-1}; & y_1 > y_2 \\ \theta_1\theta_2' f(y_1)f(y_2)[S(y_1)]^{(\theta_1+\theta_2-\theta_2'-1)}[S(y_2)]^{\theta_2'-1}; & y_2 > y_1. \end{cases} \tag{3.2}$$

Note that these load sharing models take into consideration the *dependence* between the failure times of the system components.

Model fitting and sophisticated techniques for inference of parameters of load sharing models have been addressed in the recent literature. Hanagal (2011) discussed inference for a modified Freund's exponential distribution model. Kvam & Pena (2005) discussed estimating load sharing models in a dynamic reliability systems framework under an equal load share rule. Kim & Kvam (2004) derived methods for statistical inference on load-share parameters based on the maximum likelihood principle when the load share rule is unknown. Deshpande et al. (2010) described a general semiparametric family of distributions for load share systems and proposed a nonparametric test for the dependence between failure times.

As seen in 1.5.5 an alternate flexible tool for modelling dependent times to failure is the "frailty model". Frailty models have been widely used to study dependent lifetimes in reliability and survival analysis framework (Wassell et al. (1995), Ma & Krings (2008) and Hougaard (2000)).

There are many situations where it is physically meaningful to incorporate the dependence induced both by the frailty and *the dependence due to load sharing* in studying lifetimes of a multi-component system. For instance, often in reliability data, the covariates are not measured or ignored. In such situations it is advantageous to analyze the data by accommodating the frailty aspect to the model. Wassell et al. (1995) advocated the use of frailty models to study the manufacturing effects on respirator cartridges. Another work in this direction was given in the random hazards model (Lu & Bhattacharyya (1990)) and more recently by Hanagal (2010). In this chapter, we describe a generalized model framework for the bivariate load sharing model with frailty and covariates. A re-parameterized model in Hanagal (2011) is a special case of our model.

The format of this chapter follows. In Section 3.2, we present the model formulation for a generalized conditional model for load sharing with a frailty (random effect) and observed covariates is presented. General properties of the model studied extensively. We also derived the generalized unconditional model by integrating out the frailty component. Section 3.3 derives a general cross-ratio function for the model presented in Section 3.2 and we also studied the properties in detail. Section 3.4 discuss few examples for the general model by considering different frailty distributions namely gamma distribution, power variance family of distributions and inverse Gaussian distribution with Weibull baseline. Section 3.5 describes the bivariate hazard gradient for general model proposed by Johnson & Kotz (1975) and Cox (1972). A general description about the parameter estimation for the general model presented in 3.2 is explained by maximizing the profile likelihood of the unknown parameters with and without censoring are given in Section 3.6. A general simulation study has been explained in 3.7. The chapter ends with a discussion summary in Section 3.8.

3.2 Generalized Bivariate Load Sharing Model with Frailty and Covariates

This section describes the generalized bivariate load sharing model with frailty and covariates. Section 3.2.1 derives the conditional model given the frailty and describes some useful properties. Section 3.2.2 shows the unconditional model and discusses several special cases.

3.2.1 Conditional Model Given Frailty

Let Z denote the random frailty effect associated with the two-component parallel system in (3.2) and let $\boldsymbol{X}$ denote the vector of observed covariates. As mentioned in Sections 1.1 and 1.2, $S(.)$ and $r(.)$ are respectively the baseline survival and hazard functions. Let (Y_1, Y_2) denote the lifetime of the system. The random frailty is assumed to have a multiplicative effect on the conditional failure rates as follows. Following the notation in Cox (1972), for $i = 1, 2$, let $\lambda_{i0}(y_i)$ denote the failure rate of the i^{th} component when both components are functioning at time y_i, and for $i \neq j = 1, 2$, let $\lambda_{ij}(y_i|y_j)$ denote the failure rate of the ith component given that the jth component has failed at time y_j:

$$\begin{aligned}\lambda_{i0}(y) &= \lim_{\Delta y \to 0+} \frac{P\left(y \leq Y_i < y + \Delta y | y \leq Y_1, y \leq Y_2\right)}{\Delta y}, y_1 = y_2 = y \\ \lambda_{ij}(y_i|y_j) &= \lim_{\Delta y_i \to 0+} \frac{P\left(y_i \leq Y_i < y_i + \Delta y_i | y_i \leq Y_i, Y_j = y_j\right)}{\Delta y_i}, \ y_j < y_i.\end{aligned} \tag{3.3}$$

Given $\boldsymbol{X}$, and $Z = z$,

$$\begin{aligned}\lambda_{10}(y|z, \boldsymbol{X}) &= z\theta_1 r(y) e^{\boldsymbol{X}\boldsymbol{\beta}}, y_1 = y_2 = y > 0 \\ \lambda_{20}(y|z, \boldsymbol{X}) &= z\theta_2 r(y) e^{\boldsymbol{X}\boldsymbol{\beta}}, y_1 = y_2 = y > 0 \\ \lambda_{12}(y_1|y_2, z, \boldsymbol{X}) &= z\theta_1' r(y_1) e^{\boldsymbol{X}\boldsymbol{\beta}}, y_1 > y_2 \\ \lambda_{21}(y_2|y_1, z, \boldsymbol{X}) &= z\theta_2' r(y_2) e^{\boldsymbol{X}\boldsymbol{\beta}}, y_1 < y_2.\end{aligned} \tag{3.4}$$

For more details on (3.3) we refer to Shaked & Shanthikumar (2015). Recalling that $S(y|z, \boldsymbol{X}) = \exp\left[-ze^{\boldsymbol{X}\boldsymbol{\beta}} \int_0^y r(u)du\right]$, the bivariate density function of (Y_1, Y_2) conditioned on the frailty and observed covariates $f((y_1, y_2)|z, \boldsymbol{X})$ follows from (3.3) and (3.4), as (Cox (1972))

$$\begin{aligned} f(y_1, y_2|z, \boldsymbol{X}) &= \exp\left[-\int_0^{y_1} \{\lambda_{10}(u|z, \boldsymbol{X}) + \lambda_{20}(u|z, \boldsymbol{X})\}du - \int_{y_1+0}^{y_2} \lambda_{21}(u|y_1, z, \boldsymbol{X})du\right] \\ &\times \lambda_{10}(y_1\,|z, \boldsymbol{X})\, \lambda_{21}(y_2\,|\,y_1, z, \boldsymbol{X}) \\ &= \exp\left[-z(\theta_1 + \theta_2)e^{\boldsymbol{X}\boldsymbol{\beta}} \int_0^{y_1} r(u)du\right] \exp\left[-\int_0^{y_2} z\theta_2' e^{\boldsymbol{X}\boldsymbol{\beta}} r(u)du\right] \\ &\times \exp\left[\int_0^{y_1} z\theta_2' e^{\boldsymbol{X}\boldsymbol{\beta}} r(u)du\right] z\theta_1 \frac{f(y_1)}{S(y_1)} e^{\boldsymbol{X}\boldsymbol{\beta}} z\theta_2' \frac{f(y_2)}{S(y_2)} e^{\boldsymbol{X}\boldsymbol{\beta}} \\ &= z^2 e^{2\boldsymbol{X}\boldsymbol{\beta}} \theta_1 \theta_2' f(y_1) f(y_2) [S(y_1)]^{ze^{\boldsymbol{X}\boldsymbol{\beta}}(\theta_1+\theta_2-\theta_2')-1} [S(y_2)]^{ze^{\boldsymbol{X}\boldsymbol{\beta}}\theta_2'-1} \end{aligned} \tag{3.5}$$

for $y_2 > y_1$.

Equivalently,

$$f((y_1, y_2)|\boldsymbol{X}, z) = \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \frac{d}{dy_1}[S(y_1)]^{z\theta_1' e^{\boldsymbol{X\beta}}} \frac{d}{dy_2}[S(y_2)]^{z(\theta_1+\theta_2-\theta_1')e^{\boldsymbol{X\beta}}}. \tag{3.6}$$

Similarly, for $y_2 < y_1$

$$\begin{aligned} f(y_1, y_2|z, \boldsymbol{X}) &= \exp\left[-\int_0^{y_2} \{\lambda_{10}(u|z, \boldsymbol{X}) + \lambda_{20}(u|z, \boldsymbol{X})\}du - \int_{y_2+0}^{y_1} \lambda_{12}(u|y_2, z, \boldsymbol{X})du\right] \\ &\times \lambda_{20}(y_2\,|z, \boldsymbol{X})\,\lambda_{12}(y_1\,|\,y_2, z, \boldsymbol{X}) \\ &= \exp\left[-z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}\int_0^{y_2} r(u)du\right]\exp\left[-\int_0^{y_1} z\theta_1' e^{\boldsymbol{X\beta}} r(u)du\right] \\ &\times \exp\left[\int_0^{y_2} z\theta_1' e^{\boldsymbol{X\beta}} r(u)du\right] z\theta_2 \frac{f(y_2)}{S(y_2)} e^{\boldsymbol{X\beta}} z\theta_1' \frac{f(y_1)}{S(y_1)} e^{\boldsymbol{X\beta}} \\ &= z^2 e^{2\boldsymbol{X\beta}}\theta_2\theta_1' f(y_1) f(y_2)[S(y_2)]^{ze^{\boldsymbol{X\beta}}(\theta_1+\theta_2-\theta_1')-1}[S(y_1)]^{ze^{\boldsymbol{X\beta}}\theta_1'-1}. \end{aligned} \tag{3.7}$$

Equivalently,

$$f(y_1, y_2|\boldsymbol{X}, z) = \frac{\theta_1}{\theta_1 + \theta_2 - \theta_2'} \frac{d}{dy_2}[S(y_2)]^{z\theta_2' e^{\boldsymbol{X\beta}}} \frac{d}{dy_1}[S(y_1)]^{z(\theta_1+\theta_2-\theta_2')e^{\boldsymbol{X\beta}}},\ y_2 > y_1. \tag{3.8}$$

The general bivariate density for load share with frailty and observed covariates is given by

$$f((y_1, y_2)|z, \boldsymbol{X}) = z^2 e^{2\boldsymbol{X\beta}}\theta_i'\theta_j f(y_1) f(y_2)[S(y_j)]^{ze^{\boldsymbol{X\beta}}(\theta_1+\theta_2-\theta_i')-1}[S(y_i)]^{ze^{\boldsymbol{X\beta}}\theta_i'-1},\ y_i > y_j. \tag{3.9}$$

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

The corresponding joint survival function conditional on $\boldsymbol{X}$ and Z is derived as

$$S(y_1, y_2|\boldsymbol{X}, z) = \int_{y_1}^{\infty}\int_{y_2}^{\infty} f(t_1, t_2|\boldsymbol{X}, z)dt_2dt_1.$$

For $y_1 \geq y_2$

$$
\begin{aligned}
S(y_1, y_2|\boldsymbol{X}, z) &= \int_{y_1}^{\infty}\int_{y_2}^{t_1} f(t_1 > t_2)dt_2dt_1 + \int_{y_1}^{\infty}\int_{t_1}^{\infty} f(t_1 < t_2)dt_2dt_1 \\
&= \int_{y_1}^{\infty}\left[\int_{y_2}^{t_1} \frac{\theta_2}{\theta_1+\theta_2-\theta_1'}\frac{d}{dt_1}[S(t_1)]^{z\theta_1' e^{\boldsymbol{X\beta}}}\frac{d}{dt_2}[S(t_2)]^{z(\theta_1+\theta_2-\theta_1')e^{\boldsymbol{X\beta}}}dt_2\right]dt_1 \\
&+ \int_{y_1}^{\infty}\left[\int_{t_1}^{\infty} \frac{\theta_1}{\theta_1+\theta_2-\theta_2'}\frac{d}{dt_2}[S(t_2)]^{z\theta_2' e^{\boldsymbol{X\beta}}}\frac{d}{dt_1}[S(t_1)]^{z(\theta_1+\theta_2-\theta_2')e^{\boldsymbol{X\beta}}}dt_2\right]dt_1 \\
&= \frac{\theta_2[S(y_1)]^{z\theta_1' e^{\boldsymbol{X\beta}}}}{\theta_1+\theta_2-\theta_1'}\left[[S(y_2)]^{z(\theta_1+\theta_2-\theta_1')e^{\boldsymbol{X\beta}}} - [S(y_1)]^{z(\theta_1+\theta_2-\theta_1')e^{\boldsymbol{X\beta}}}\right] \\
&+ [S(y_1)]^{z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}}. \qquad (3.10)
\end{aligned}
$$

Similarly, for $y_2 \geq y_1$

$$
\begin{aligned}
S((y_1, y_2)|z, \boldsymbol{X}) &= \frac{\theta_1[S(y_2)]^{z\theta_2' e^{\boldsymbol{X\beta}}}}{\theta_1+\theta_2-\theta_2'}\left[[S(y_1)]^{z(\theta_1+\theta_2-\theta_2')e^{\boldsymbol{X\beta}}} - [S(y_2)]^{z(\theta_1+\theta_2-\theta_2')e^{\boldsymbol{X\beta}}}\right] \\
&+ [S(y_2)]^{z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}}. \qquad (3.11)
\end{aligned}
$$

In general, the bivariate survival function for load share with frailty and covariates is given by

$$
S(y_1, y_2|z, \boldsymbol{X}) = [1-k_{ij}][S(y_i)]^{z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}} + k_{ij}\left[\frac{S(y_i)}{S(y_j)}\right]^{z\theta_i' e^{\boldsymbol{X\beta}}}[S(y_j)]^{z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}};\ y_i \geq y_j \qquad (3.12)
$$

where, $k_{ij} = \frac{\theta_j}{\theta_1+\theta_2-\theta_i'}$, when $\theta_1+\theta_2 \neq \theta_i'$, $i \neq j = 1, 2$.

Remark 3.2.1. *When $\theta_1 + \theta_2 = \theta_i'$, for $i \neq j = 1, 2$, the joint survival function is given by*

$$
S(y_1, y_2|z, \boldsymbol{X}) = [S(y_i)]^{z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}}\left[1 + ze^{\boldsymbol{X\beta}}\theta_j\left(\log\, S(y_j) - \log\, S(y_i)\right)\right];\ y_i \geq y_j. \qquad (3.13)
$$

Property 3.2.1. *The probability density function specified in* (3.9) *reduces to a model with independent marginals whenever $\theta_i = \theta_i'$, for $i = 1, 2$.*

This is straight forward by substituting $\theta_i = \theta_i'$ for $i = 1, 2$, in which case,

$$f(y_1, y_2|z, \boldsymbol{X}) = \frac{d[S(y_1)]^{z\theta_1 e^{\boldsymbol{X\beta}}}}{dy_1} \times \frac{d[S(y_2)]^{z\theta_2 e^{\boldsymbol{X\beta}}}}{dy_2}.$$

Property 3.2.2. *The marginal survival function of Y_i, denoted as $S_{Y_i}(y_i|z)$ is specified by $S(y_1, 0) = S(y_1)$ and $S(0, y_2) = S(y_2)$. In general, the marginal survival function for load share with frailty and covariates is given by*

$$S_{Y_i}(y_i|z) = [1 - k_{ij}][S(y_i)]^{z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}} + k_{ij}\left[S(y_i)\right]^{z\theta_i' e^{\boldsymbol{X\beta}}}; \; y_i \geq 0,$$

for $\theta_1 + \theta_2 \neq \theta_i', \; i \neq j = 1, 2,$ and by

$$S_{Y_i}(y_i|z) = [S(y_i)]^{z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}}\left[1 - ze^{\boldsymbol{X\beta}}\theta_j \; (\log \; S(y_i))\right];$$

for $\theta_1 + \theta_2 = \theta_i', \; i \neq j = 1, 2.$

Property 3.2.3. *The conditional survival function for load share with frailty and covariates is given by*

$$S(y_i|y_j) = \frac{\frac{d}{dy_j} S(y_1, y_2|\boldsymbol{X}, z)}{\frac{d}{dy_j} S(Y_i|\boldsymbol{X}, z)}; \text{ for } y_i \geq y_j, \; i \neq j, \; = 1, 2.$$

The general conditional survival function for load share with frailty and covariates is given by

$$S(y_i|y_j) = \frac{\theta_j\left[S(y_i)\right]^{z\theta_i' e^{\boldsymbol{X\beta}}}\left[S(y_j)\right]^{ze^{\boldsymbol{X\beta}}(\theta_1+\theta_2-\theta_i'-1)}}{(1-k_{ij})(\theta_1+\theta_2)\left[S(y_j)\right]^{ze^{\boldsymbol{X\beta}}(\theta_1+\theta_2)-1} + k_{ij}\theta_j'\left[S(y_j)\right]^{z\theta_j' e^{\boldsymbol{X\beta}-1}}}; y_i \geq y_j, i \neq j = 1, 2. \tag{3.14}$$

Property 3.2.4. *The survival distribution of $V = \min{(Y_1, Y_2)}$ is the proportional hazards model specified by*

$$S_V(v|z) = [S(v)]^{z(\theta_1+\theta_2)e^{\boldsymbol{X\beta}}}; \;\; v \geq 0.$$

Property 3.2.5. *The survival distribution of $W = \text{Max}(Y_1, Y_2)$ is the proportional* *hazards model specified by*

$$
\begin{aligned}
P[\text{Max}(Y_1, Y_2) \le w] &= P[Y_1 \le w, Y_2 \le w] \\
&= 1 - P[Y_1 > w] - P[Y_2 > w] + P[Y_1 > w, Y_2 > w]
\end{aligned}
\tag{3.15}
$$

Then if $W = \text{Max}(Y_1, Y_2)$

$$
S_W(w|\boldsymbol{X}) = 1 - \sum_{i=1}^{2} S_{Y_i}(w|\boldsymbol{X}) - L_z(\Psi_1(w)),\ w \ge 0,
\tag{3.16}
$$

where $S_{Y_i}(w|\boldsymbol{X})$ *is*

$$
S_{Y_i}(w|\boldsymbol{X}) = [1 - k_{ij}]L_z(\Psi_1(w)) + k_{ij}L_z(\theta_i' H(w));\ w \ge 0,\ i, j = 1, 2,\ i \ne j.
\tag{3.17}
$$

for $\theta_1 + \theta_2 \ne \theta_i'$, $i = 1, 2$ *and*

$$
S_{Y_i}(w|\boldsymbol{X}) = L_z(\Psi_1(w)) - \left[e^{\boldsymbol{X\beta}} \theta_2 \left(\log\, S(w)\right) \frac{\partial}{\partial s} L_z(s) \right]_{s=(\Psi_1(w))}.
\tag{3.18}
$$

for $\theta_1 + \theta_2 = \theta_i'$.

Property 3.2.6. *Without the observed covariates,* (3.9) *reduces to the model specified by*

$$
f((y_1, y_2)|z) = z^2 \theta_i' \theta_j f(y_1) f(y_2) [S(y_j)]^{z(\theta_1 + \theta_2 - \theta_i') - 1} [S(y_i)]^{z\theta_i' - 1}; y_i > y_j,\ i \ne j = 1, 2.
\tag{3.19}
$$

Remark 3.2.2. *When* $S(y) = e^{-y}$, *model* (3.19) *reduces to the Freund's bivariate exponential model (Freund (1961)) for a given frailty* Z *with probability density function*

$$
f(y_1, y_2|z) = z^2 \theta_i' \theta_j e^{-z\theta_i' y_i} e^{-z(\theta_1 + \theta_2 - \theta_i') y_j};\ y_i > y_j,\ i \ne j = 1, 2.
\tag{3.20}
$$

Remark 3.2.3. *When* $S(y) = e^{-y^\gamma}, \gamma > 0$, *the model reduces to the Weibull extension of the Freund model (Lu (1989)) for a given frailty* Z *with probability density function*

$$
f(y_1, y_2|z) = z^2 \theta_i' \theta_j \gamma^2 y_i^{\gamma - 1} y_j^{\gamma - 1} e^{-z\theta_i' y_i^\gamma} e^{-z(\theta_1 + \theta_2 - \theta_i') y_j^\gamma};\ y_i > y_j,\ i \ne j = 1, 2.
\tag{3.21}
$$

3.2.2 Unconditional Model

Let $g(z)$ denote the density function of the unobserved frailty random variable Z, and let $L_z(s)$ denote its Laplace transform. Integrating out z in (3.12), and recalling that $S(y_i|z,\boldsymbol{X}) = e^{-zH(y_i)e^{\boldsymbol{X\beta}}}$, and also observing that $\int_z e^{-zH(y_i)e^{\boldsymbol{X\beta}}}g(z)dz = L_z[H(y_i)e^{\boldsymbol{X\beta}}]$, where $H(y_i)$ is the cumulative hazard function of Y_i, $i=1,2$, we get the unconditional survival function for $y_2 \geq y_1$, when $\theta_1+\theta_2 \neq \theta_i'$ as

$$\begin{aligned} S(y_1,y_2|\boldsymbol{X}) &= \int_{z=0}^{\infty} S(y_1,y_2|z,\boldsymbol{X})g(z)dz \\ &= [1-k_{21}]\int_z e^{-z\Psi_1(y_2)}g(z)dz + k_{21}\int_z e^{-z\Psi_{21}(y_1,y_2)}g(z)dz \\ &= [1-k_{21}]L_z[\Psi_1(y_2)] + k_{21}\left\{L_z[\Psi_{21}(y_1,y_2)]\right\}, \end{aligned}$$

where $\Psi_1(y_2) = (\theta_1+\theta_2)H(y_2)e^{\boldsymbol{X\beta}}$ and $\Psi_{21}(y_1,y_2) = [\theta_2'H(y_2) + (\theta_1+\theta_2-\theta_2')H(y_1)]\, e^{\boldsymbol{X\beta}}$.

When $y_1 \geq y_2$, the unconditional joint survival function $S(y_1,y_2|\boldsymbol{X})$ is given by

$$S(y_1,y_2|\boldsymbol{X}) = [1-k_{12}]L_z[\Psi_1(y_1)] + k_{12}\left\{L_z[\Psi_{12}(y_1,y_2)]\right\}$$

where $\Psi_1(y_1) = (\theta_1+\theta_2)H(y_1)e^{\boldsymbol{X\beta}}$ and $\Psi_{12}(y_1,y_2) = [\theta_1'H(y_1) + (\theta_1+\theta_2-\theta_1')H(y_1)]\, e^{\boldsymbol{X\beta}}$.

In general we can write,

$$S(y_1,y_2|\boldsymbol{X}) = [1-k_{ij}]L_z(\Psi_1(y_i)) + k_{ij}Lz(\Psi_{ij}(y_1,y_2));\ y_i \geq y_j, \tag{3.22}$$

where $\Psi_1(y_i) = (\theta_1+\theta_2)H(y_i)e^{\boldsymbol{X\beta}}$, $\Psi_{ij}(y_1,y_2) = [\theta_i'H(y_i) + (\theta_1+\theta_2-\theta_i')H(y_j)]e^{\boldsymbol{X\beta}}$ and $\theta_1+\theta_2 \neq \theta_i',\ i \neq j = 1,2$.

Proceeding in the same manner for $\theta_1+\theta_2 = \theta_i',\ i=1,2$, we obtain the unconditional survival function for $y_i \geq y_j$ as

$$S(y_1,y_2|\boldsymbol{X}) = L_z(\Psi_1(y_i)) + \left[e^{\boldsymbol{X\beta}}\theta_2\left(\log\, S(y_j) - \log\, S(y_i)\right)\left(\frac{\partial}{\partial s}L_z(s)\bigg|_{s=\Psi_1(y_i)}\right)\right]. \tag{3.23}$$

3.2.3 The Joint Density Function for the Unconditional Model

The joint density function corresponding to (3.22) is derived as

$$f(y_1, y_2|\boldsymbol{X}) = \frac{\partial^2 S(y_1, y_2|\boldsymbol{X})}{\partial y_1 \partial y_2}.$$

For $y_1 > y_2$

$$\begin{aligned} f(y_1, y_2|\boldsymbol{X}) &= \frac{\partial^2 \left\{[1 - k_{12}]L_z(\Psi_1(y_1)) + k_{12}Lz(\Psi_{12}(y_1, y_2))\right\}}{\partial y_1 \partial y_2} \\ &= \theta_1' \theta_2 r(y_1) r(y_2) e^{2\boldsymbol{X\beta}} \left[\frac{\partial^2 L_z(s)}{\partial s^2}\right]_{s=\Psi_{12}(y_1,y_2)}. \end{aligned} \tag{3.24}$$

Similarly, for $y_2 > y_1$

$$\begin{aligned} f(y_1, y_2|\boldsymbol{X}) &= \frac{\partial^2 \left\{[1 - k_{21}]L_z(\Psi_1(y_2)) + k_{21}Lz(\Psi_{21}(y_1, y_2))\right\}}{\partial y_1 \partial y_2} \\ &= \theta_2' \theta_1 r(y_1) r(y_2) e^{2\boldsymbol{X\beta}} \left[\frac{\partial^2 L_z(s)}{\partial s^2}\right]_{s=\Psi_{21}(y_1,y_2)}. \end{aligned} \tag{3.25}$$

Proceeding in the same manner, the joint density function for $\theta_1 + \theta_2 = \theta_i'$ is directly derived from (3.23) as

$$f(y_1, y_2|\boldsymbol{X}) = \frac{\partial^2 S(y_1, y_2|\boldsymbol{X})}{\partial y_1 \partial y_2}$$

For $y_1 > y_2$

$$\begin{aligned} f(y_1, y_2|\boldsymbol{X}) &= \frac{\partial^2 \left\{ L_z(\Psi_1(y_1)) + \left[e^{\boldsymbol{X\beta}} \theta_2 \left(\log\, S(y_2) - \log\, S(y_1)\right) \left(\frac{\partial}{\partial s} L_z(s) \Big|_{s=\Psi_1(y_1)} \right) \right] \right\}}{\partial y_1 \partial y_2} \\ &= \theta_1' \theta_2 r(y_1) r(y_2) e^{2\boldsymbol{X\beta}} \left[\frac{\partial^2 L_z(s)}{\partial s^2}\right]_{s=\Psi_{12}(y_1,y_2)}. \end{aligned} \tag{3.26}$$

Similarly, for $y_2 > y_1$

$$f(y_1, y_2|\boldsymbol{X}) = \theta_2'\theta_1 r(y_1)r(y_2)e^{2\boldsymbol{X\beta}}\left[\frac{\partial^2 L_z(s)}{\partial s^2}\right]_{s=\Psi_{21}(y_1,y_2)}. \tag{3.27}$$

In general, the density corresponding to both (3.22) and (3.23) is

$$f(y_1, y_2|\boldsymbol{X}) = \theta_i'\theta_j r(y_1)r(y_2)e^{2\boldsymbol{X\beta}}\left[\frac{\partial^2 L_z(s)}{\partial s^2}\right]_{s=\Psi_{ij}(y_1,y_2)}; \; y_i > y_j, \;\; i \neq j = 1, 2. \tag{3.28}$$

Observing that $\Psi_{ii}(y_i, y_i) = \Psi_1(y_i)$, it can be directly seen that the survival function of $V = \min(Y_1, Y_2)$ is

$$S_V(v|\boldsymbol{X}) = L_z(\Psi_1(v)), \;\; v \geq 0.$$

Also, the unconditional marginal survival function of Y_i, for $\theta_1 + \theta_2 \neq \theta_i'$ has the form

$$S_{Y_i}(y_i|\boldsymbol{X}) = [1 - k_{ij}]L_z(\Psi_1(y_i)) + k_{ij}L_z(\theta_i' H(y_i)); \; y_i \geq 0, \; i \neq j = 1, 2.$$

For $\theta_1 + \theta_2 = \theta_i'$,

$$S_{Y_i}(y_i|\boldsymbol{X}) = L_z(\Psi_1(y_i)) - \left[e^{\boldsymbol{X\beta}}\theta_2 \left(\log\; S(y_i)\right) \frac{\partial}{\partial s} L_z(s)\right]_{s=(\Psi_1(y_i))}.$$

Remark 3.2.4. *Observe when $\theta_i = \theta_i'$, $i = 1, 2$ the components are independent and the unconditional model given in* (3.22) *becomes*

$$S(y_1, y_2|\boldsymbol{X}) = L_z\left[\left(H_1(y_1) + H_2(y_2)\right)e^{\boldsymbol{X}\beta}\right]. \tag{3.29}$$

3.3 The Cross Ratio Function

It is of interest to quantify the association between the failure times in bivariate survival data. Clayton's local cross-ratio function (CRF) (Clayton (1978)) discussed in 1.4.1 describes the time-varying dependence and is defined at (y_1, y_2) by

$$\mathscr{C}(y_1, y_2) = \frac{S(y_1, y_2|\boldsymbol{X})S_{12}(y_1, y_2|\boldsymbol{X})}{S_1(y_1, y_2|\boldsymbol{X})S_2(y_1, y_2|\boldsymbol{X})},$$

where $S_j(y_1, y_2|\boldsymbol{X}) = \frac{\partial S(y_1,y_2|\boldsymbol{X})}{\partial y_j}$, $j = 1, 2$ and $S_{12}(y_1, y_2|\boldsymbol{X}) = \frac{\partial^2 S(y_1,y_2|\boldsymbol{X})}{\partial y_1 \partial y_2}$. For the unconditional bivariate survival functions in (3.22) and (3.23) the $CRF's$, are respectively

$$\mathscr{C}_1(y_1, y_2) = \frac{\frac{\partial}{\partial y_i}\left[\log \frac{\partial}{\partial s} L_z(\Psi_{ij}(y_1, y_2)\right]}{\frac{\partial}{\partial y_i}\left[\log\left\{(1 - k_{ij})L_z(\Psi_1(y_i)) + k_{ij} L_z(\Psi_{ij}(y_1, y_2))\right\}\right]};\ y_i > y_j \tag{3.30}$$

and

$$\mathscr{C}_2(y_1, y_2) = \frac{\frac{\partial}{\partial s}\left[\log \frac{\partial L_z(s)}{\partial s}\right]_{s=\Psi_1(y_i)} \frac{\partial}{\partial y_i}\Psi_1(y_i)}{\frac{\partial}{\partial y_i}\log\left[L_z(\Psi_1(y_i)) + \left(e^{\boldsymbol{X\beta}}\theta_2(\log S(y_j) - \log S(y_i))\frac{\partial}{\partial s}L_z(s)\right)_{s=\Psi_1(y_i)}\right]}. \tag{3.31}$$

The CRF in the absence of frailty obtained from (3.12) and (3.13) with $z = 1$ are respectively

$$\mathscr{C}_1^L(y_1, y_2) = \frac{\theta_i'\left[1 + \frac{k_{ij}[S(y_i)]^{e^{\boldsymbol{X\beta}}(\theta_i'-1)}[S(y_j)]^{e^{\boldsymbol{X\beta}}(\theta_1+\theta_2-\theta_i')}}{(1-k_{ij})[S(y_i)]^{e^{\boldsymbol{X\beta}}(\theta_1+\theta_2-1)}}\right]}{(\theta_1 + \theta_2)\left[1 + \frac{k_{ij}[S(y_i)]^{e^{\boldsymbol{X\beta}}(\theta_i'-1)}[S(y_j)]^{e^{\boldsymbol{X\beta}}(\theta_1+\theta_2-\theta_i')}}{(1-k_{ij})[S(y_i)]^{e^{\boldsymbol{X\beta}}(\theta_1+\theta_2-1)}}\right]}; y_i > y_j \tag{3.32}$$

and

$$\mathscr{C}_2^L(y_1, y_2) = \frac{1 + e^{\boldsymbol{X\beta}}\theta_j(\log S(y_j) - \log S(y_i))}{\frac{\theta_i}{\theta_1+\theta_2} + e^{\boldsymbol{X\beta}}\theta_j(\log S(y_j) - \log S(y_i))}; y_i > y_j, i \neq j = 1, 2. \tag{3.33}$$

Remark 3.3.1. *Observe that the CRF depends on (Y_1, Y_2) only through the distribution of the frailty variable Z and baseline distribution $S(y_i)$.*

Remark 3.3.2. *The $CRFs$ under the absence of observed covariates are obtained by replacing $e^{\boldsymbol{X\beta}}\theta_i$ and $e^{\boldsymbol{X\beta}}\theta_i'$ by θ_i and θ_i' in (3.32) and (3.33) respectively. This is because for $\mathscr{C}_1^L(y_1, y_2)$ and $\mathscr{C}_2^L(y_1, y_2)$, the effect of the observed covariates is multiplicative on the load sharing parameters thus retaining the dependence structure under the presence of observed covariates.*

It is of interest to investigate conditions under which the cross ratio function of the models (3.22) and (3.23), indicates independence or positive dependence. This is discussed below.

Property 3.3.1. *For* $\theta_i = \theta_i'$, $i = 1, 2$, $\mathscr{C}_1(y_1, y_2) = 1$ *for all* y_1, y_2 *if and only if* Z *is degenerate at 1.*

Proof. That $\mathscr{C}_1(y_1, y_2) = 1$ for all y_1, y_2 when Z is degenerate at 1 follows directly from (3.30). Conversely, let $\mathscr{C}_1(y_1, y_2) = 1$, then from (3.30) it follows that

$$\frac{\partial L_z(\Psi_1(y_i))}{\partial y_i} = L_z(\Psi_1(y_i))$$

which on solution provides $L_z(s) = e^{-s}$. ∎

Property 3.3.2. *For* Z *degenerate at 1 and* $\theta_1 + \theta_2 = \theta_i'$, $\mathscr{C}_2(y_1, y_2) = \mathscr{C}_2^L(y_1, y_2) > 1$ *for all* $y_i > y_j$; $i \neq j = 1, 2$.

Proof. The proof follows in a straight forward manner from (3.31) and (3.33). ∎

Property 3.3.3. $\mathscr{C}_1(y_1, y_2) > (<)\ 1$, *according as*

$$\frac{\partial L_z(\Psi_{ij}(y_1, y_2))}{\partial y_i} > (<)(1 - k_{ij}) L_z(\Psi_1(y_i)) + k_{ij} L_z(\Psi_{ij}(y_1, y_2)).$$

3.4 Examples

A family of generalized bivariate load sharing models with different frailty distributions is listed in Examples 3.4.1 - 3.4.3.

Example 3.4.1. *Let the frailty random variable* Z *follow a one-parameter gamma distribution with density function* $f(z) = \frac{\alpha^\alpha z^{\alpha-1} e^{-\alpha z}}{\Gamma(\alpha)}, z > 0, \alpha > 0$ *and Laplace transform* $L_z(s) = \left[1 + \frac{s}{\alpha}\right]^{-\alpha}, \alpha > 0$. *Then, the load share Gamma frailty model is*

$$f(y_1, y_2 | \boldsymbol{X}) = \theta_i' \theta_j r(y_1) r(y_2) e^{\boldsymbol{X\beta}} \left(\frac{1+\alpha}{\alpha}\right) \times \left[1 + \frac{\Psi_{ij}(y_1, y_2)}{\alpha}\right]^{-(\alpha+2)}; \ y_i > y_j, \ i \neq j = 1, 2. \tag{3.34}$$

Example 3.4.2. *Let the frailty random variable Z belong to the power variance family (PVF) of distributions with density function*

$$\begin{aligned} f(z) = & \; e^{\sigma z + \frac{\sigma}{\alpha}} \frac{1}{z} \sum_{k=1}^{\infty} \left[\frac{\left(\frac{z^{\alpha}}{(\alpha\sigma^{\alpha-1})} \right)^{k}}{k!\Gamma(-k\alpha)} \right], z > 0, \\ = & \quad 0 \text{ therwise and } k\alpha > 0; \end{aligned}$$

and Laplace transform $L_z(s) = e^{-\frac{\sigma\{(1+\frac{s}{\sigma})^{\alpha}-1\}}{\alpha}}$. Then the load share PVF frailty model is

$$\begin{aligned} f(y_1, y_2|\boldsymbol{X}) = \theta_i'\theta_j r(y_1) r(y_2) e^{2\boldsymbol{X\beta}} \left(1 + \Psi_{ij}(y_1, y_2)\right)^{(\alpha-2)} e^{-\frac{\sigma}{\alpha}\left[\left(1+\frac{\Psi_{ij}(y_1,y_2)}{\alpha}\right)^{\alpha}-1\right]} \\ \times \left[(1 + \Psi_{ij}(y_1, y_2))^{\alpha} + (\alpha - 1)\right]; \; \alpha > 0, \; y_i > y_j, \quad i \neq j = 1, 2. \end{aligned} \tag{3.35}$$

The positive stable distribution is a sub-class; of PVF distributions. In Chapter 4, we discuss this special case of Example 3.4.2 and its properties in detail.

Example 3.4.3. *Let the frailty random variable Z follow the inverse Gaussian distribution with density function $f(z) = \left[\frac{1}{2\pi\sigma^2}\right]^{\frac{1}{2}} z^{\frac{-3}{2}} e^{\frac{-(z-1)^2}{2z\sigma^2}}; z > 0, \sigma^2 > 0$ and Laplace transform $L_z(s) = \exp\left[\frac{1-(1+2\sigma^2 s)^{\frac{1}{2}}}{\sigma^2}\right]$. Then the load share inverse Gaussian frailty model is*

$$\begin{aligned} f(y_1, y_2|\boldsymbol{X}) = \theta_i'\theta_j r(y_1) r(y_2) e^{2\boldsymbol{X\beta}} \\ \times \left. \frac{\partial^2 \left[\exp\left(\frac{1-(1+2\sigma^2 s)^{\frac{1}{2}}}{\sigma^2} \right) \right]}{\partial s} \right|_{s=\Psi_{ij}(y_1,y_2)} ; \; y_i > y_j, \; i \neq j = 1, 2. \end{aligned} \tag{3.36}$$

3.5 Bivariate Hazard Gradient

Johnson & Kotz (1975) hazard gradient in (1.6) and (1.7) for the unconditional distribution in (3.22) is obtained as,

$$r_i(y_1, y_2) = -\frac{\partial}{\partial y_i} \log\left\{[1 - k_{ij}] L_z(\Psi_1(y_i)) + k_{ij} Lz(\Psi_{ij}(y_1, y_2))\right\}, \; i \neq j = 1, 2. \tag{3.37}$$

The Cox hazard gradient (3.3) for the unconditional distribution in (3.22) is derived as

$$\begin{aligned} \lambda_{i0}(y) &= \frac{\partial}{\partial y}\left[\log\ L_z(\Psi_1(y))\right],\ i=1,2,\ \text{and} \\ \lambda_{ij}(y_i|y_j) &= \frac{\partial}{\partial y_i}\left[\log\ \frac{\partial L_z(\Psi_{ij}(y_1,y_2))}{\partial y_j}\right];\ y_i > y_j,\ i \neq j = 1,2. \end{aligned} \qquad (3.38)$$

3.6 Parameter Estimation

In this section, we consider parameter estimation of the bivariate load sharing model with frailty and covariates, where the lifetimes may be subject to censoring. Different approaches have been proposed in the literature under different frailty assumptions. Klein et al. (1992) considered simultaneous estimation of parameters using EM-algorithm in case of gamma frailties. Wang et al. (1995) applied EM-algorithm for estimation in the positive stable frailty model with the frailties regarded as missing data. Xue & Brookmeyer (1996) used EM algorithm for fitting bivariate log-normal frailty model. Lam & Kuk (1997) advocated a unified marginal likelihood for parameter estimation in frailty models. Fine et al. (2003) established a simple estimation procedure for the positive stable frailty model by considering both the conditional and marginal hazards of the Cox form. Martinussen & Pipper (2005) proposed a likelihood-based estimation procedure for the positive stable frailty model. Mallick et al. (2008) developed Markov Chain Monte Carlo algorithms to facilitate Bayesian inference to estimate the parameters involved in a bivariate positive stable frailty model. Hanagal (2011) gave a general discussion on estimation of frailty models.

Let us suppose that there are n independent pairs of components or organs under study, and the r^{th} pair of the components have lifetimes (y_{1r}, y_{2r}). We use the following notations for defining, n_1 = number of observations for which $y_{1r} > y_{2r}$ and n_2 = number of observations for which $y_{1r} < y_{2r}$. Now, we want to estimate the parameters of the generalized model given in (3.28). The log-likelihood based on the sample of size n is given by

$$\ell(\tilde{\tau}) = \sum_{r=1}^{n_1} \log\left[\frac{\partial H(y_{1r})}{\partial y_{1r}}\right] + \sum_{r=1}^{n_1} \log\left[\frac{\partial H(y_{2r})}{\partial y_{2r}}\right] + \sum_{i=1}^{n_2} \log\left[\frac{\partial H(y_{1r})}{\partial y_{1r}}\right]$$

$$+\sum_{r=1}^{n_2}\log\left[\frac{\partial H(y_{2r})}{\partial y_{2r}}\right]+n_1\ \log(\theta_1'\theta_2)+n_2\ \log(\theta_1\theta_2')+\sum_{r=1}^{n}\tilde{x}_r\tilde{\beta}_r$$
$$+\sum_{r=1}^{n_1}\frac{\partial^2\left[L_z[\Psi_{12}(y_{1r},y_{2r})]\right]}{\partial y_{1r}\partial y_{2r}}+\sum_{r=1}^{n_2}\frac{\partial^2\left[L_z[\Psi_{21}(y_{1r},y_{2r})]\right]}{\partial y_{1r}\partial y_{2r}} \tag{3.39}$$

for $y_{ir}\geq y_{jr},\ i\neq j=1,2.$

The proposed model given in (3.28) can also be extended when the observed data has some censoring cases. Suppose we have n pairs of components or organs under study, and the r^{th} pair of the components have lifetimes (y_{1r},y_{2r}) and a censoring time (w_r), then the lifetime associated with the r^{th} pair of components is given by

$$\begin{aligned}(Y_{1r},Y_{2r}) &= (y_{1r},y_{2r})\,;\ \ \max\,(y_{1r},y_{2r})<w_r\\ &= (y_{1r},w_r);\ \ \ y_{1r}<w_r<y_{2r}\\ &= (w_r,y_{2r});\ \ \ y_{2r}<w_r<y_{1r}\\ &= (w_r,w_r);\ \ \ w_r<\min\,(y_{1r},y_{2r}).\end{aligned}$$

We use the following notations to define n_1=number of observations for which $y_{2r}<y_{1r}<w_r$, n_2 = number of observations for which $y_{1r}<y_{2r}<w_r$, n_3 = number of observations for which $y_{2r}<w_r<y_{1r}$, n_4 = number of observations for which $y_{1r}<w_r<y_{2r}$ and n_5 = number of observations for which $w_r<\min(y_{1r},y_{2r})$. We are interested in estimating $\tilde{\boldsymbol{\tau}}=(\theta_1,\theta_2,\theta_1',\theta_2',\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma})$, with $\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}$ denoting the frailty, regression, and baseline parameters respectively. The likelihood function given the data is

$$\prod_{r=1}^{n_1}f_{1r}\prod_{r=1}^{n_2}f_{2r}\prod_{r=1}^{n_3}f_{3r}\prod_{r=1}^{n_4}f_{4r}\prod_{r=1}^{n_5}S(w_r,w_r|\boldsymbol{X}), \tag{3.40}$$

where

$$f_{1r}=k_{12}\ \frac{\partial^2 L_z\left(\Psi_{12}(y_{1r},y_{2r})\right)}{\partial y_{1r}\partial y_{2r}};\ \ y_{2r}<y_{1r}<w_r,$$
$$f_{2r}=k_{21}\ \frac{\partial^2 L_z\left(\Psi_{21}(y_{1r},y_{2r})\right)}{\partial y_{1r}\partial y_{2r}};\ \ y_{1r}<y_{2r}<w_r,$$
$$f_{3r}=\int_{y_2}\theta_i'\theta_j r(w_r)r(y_2)e^{2\boldsymbol{X}\boldsymbol{\beta}}\left[\frac{\partial^2 L_z(s)}{\partial s^2}\right]_{s=\Psi_{ij}(w_r,y_2)};\ \ y_{2r}<w_r<y_{1r},$$

$$f_{4r} = \int_{y_1} \theta_i' \theta_j r(y_1) r(w_r) e^{2\boldsymbol{X\beta}} \left[\frac{\partial^2 L_z(s)}{\partial s^2} \right]_{s=\Psi_{ij}(y_1, w_r)} ; \;\; y_{1r} < w_r < y_{2r},$$

and $S(w_r, w_r|\boldsymbol{X})$ is as defined in (3.22).

Given the observed data, y_1, y_2 and $\boldsymbol{x}$, we wish to find the value of $\tilde{\boldsymbol{\tau}}$, where $\boldsymbol{\tau}$ is a vector that contains parameters from load sharing, frailty, regression, and baseline distributions that maximizes the log-likelihood $\ell(\tilde{\boldsymbol{\tau}})$. The likelihood equations for (3.39) and (3.40) can be obtained by solving first order partial derivatives of the likelihood and equating to zero. The likelihood equations are most likely non-linear in nature and they are difficult to solve. It may be tedious to obtain maximum likelihood estimators (MLEs) by Newton-Raphson procedure. From the reliability literature on frailty we observe that the solutions do not converge for the specified sample sizes in the Newton-Raphson procedure and the method of maximum likelihood (ML) fails to estimate all the parameters simultaneously.

Hanagal (2011) advocates the best linear unbiased predictor (BLUP) method as one of the estimation procedures to solve the likelihood equations derived from (3.39) and (3.40). In this method the likelihood function can be written in the following form:

$$\begin{aligned} h = \log L(y_1, y_2, \boldsymbol{x}, \boldsymbol{\beta}, z) &= \log L_1(y_1, y_2, \boldsymbol{x}|\boldsymbol{\beta}, z) + \log L_2(z) \\ &= \ell_1(y_1, y_2, \boldsymbol{x}|\boldsymbol{\beta}, z) + \ell_2(z). \end{aligned} \tag{3.41}$$

The BLUP method is based on the maximization of the sum of the above two components. The first component in (3.41) is the likelihood of failure times and observed covariates and the second component is the likelihood of frailty model. One can refer to McGilchrist & Aisbett (1991), McGilchrist (1993), McGilchrist (1994), Yau & McGilchrist (1998), Noh et al. (2006) for the estimation of parameters using BLUP method incorporating lognormal as frailty model. Hanagal (2011) also adopted the two-stage ML method. This method is quite similar to the profile likelihood method. This is possible when the Laplace transform of the choice of the frailty model supports the profiling method. In the first stage, estimate the parameters θ_1', θ_2', β, frailty parameter and baseline parameter by ML method by considering θ_1 and θ_2

known. In the second stage, estimate the parameters by ML method after substituting MLEs of the parameters obtained from the first stage. Then re-substitute the estimates obtained from the first stage and estimate the parameters θ_1 and θ_2. Continue this iterative procedure until the convergence meet in both stages. This estimation method is an alternative one when closed form in MLEs is not possible and iterative procedures fail to converge. For asymptotic properties one can refer to Hanagal (2005).

3.7 Simulation Study

In this section, we established an algorithm to generate the samples (y_{1r}, y_{2r}), $r = 1, 2, ..., n$ for the model in (3.28) and conducting a simulation study in order to evaluate the performance of our estimation procedure. For covariates $\boldsymbol{X}$ we assume to follow Normal distribution with mean zero and standard deviation σ.

Therefore the conditional survival function for an individual or component for given frailty $Z = z$ and covariates $\boldsymbol{X}$ at time $y_1 > 0$ and $y_2 > 0$ is given by

$$S(y_1, y_2|z, \boldsymbol{X}) = e^{-zH_0(y_1,y_2)\epsilon},$$

where $\epsilon = e^{\boldsymbol{X\beta}}$.
Sample generation of (y_{1r}, y_{2r}) is explained through the following procedure.

(a). Generate a random sample of size n from Uniform distribution $[0, 1]$ and name it as u.

(b). Generate a random sample of size n from Uniform distribution $[0, 1]$ and name it as v.

(c). Generate a random sample of size n from Uniform distribution $[0, 1]$ and name it as w.

(d). Generation of frailty random sample (z) of size n depends on the choice of frailty model.

(e). Generate $\boldsymbol{X}$ from $N(0, \sigma)$.

(f). Compute $\epsilon = e^{\boldsymbol{X}\boldsymbol{\beta}}$ with regression coefficients $\boldsymbol{\beta}$.

(g). If $u \leq \frac{\theta_1}{\theta_1+\theta_2}$, $S_0(y_1|Z, \boldsymbol{X}) = (1-v)^{\frac{1}{\theta_1}}$.
Therefore, $(y_1|Z, \boldsymbol{X}) = S_0^{-1}((1-v)^{\frac{1}{\theta_1}})$, and
$(y_2|Z, \boldsymbol{X}) = S_0^{-1}(S_0(y_1|Z, \boldsymbol{X})(1-w)^{\frac{1}{\theta'_2}})$.

(h). If $u > \frac{\theta_1}{\theta_1+\theta_2}$, $S_0(y_2|Z, \boldsymbol{X}) = ((1-v)^{\frac{1}{\theta_2}})$, Therefore, $(y_2|\boldsymbol{X}, z) = S_0^{-1}(1-v)^{\frac{1}{\theta_2}}$ and $(y_1|Z, \boldsymbol{X}) = S_0^{-1}(S_0(y_2|Z, \boldsymbol{X})(1-w)^{\frac{1}{\theta'_1}})$ (see Asha et al. (2016)).

The estimation procedure is explained in detail in Chapter 4.

3.8 Discussion and Summary

We have proposed a general load share model with frailty and covariates by using the frailty approach discussed in 1.5.5. We studied the general properties of the general model in (3.22). The local dependence measure, cross-ratio function is presented for the general model and studied its properties. Some examples have been discussed by considering different frailty distributions namely, gamma frailty, power variance family and inverse Gaussian frailty distributions. Two popular bivariate hazard gradients such as Johnson & Kotz (1975) and Cox (1972) have been discussed for the general model. General estimation procedures for the model has been discussed. A general algorithm for conducting simulation study for the proposed model is presented.

It is of interest to further investigate the class of distribution in (3.22). In the next chapter we consider a particular example where in the baseline is bivariate Weibull distribution and frailty is distributed as positive stable distribution.

Chapter 4

Load Share Positive Stable Frailty Model with Covariates

4.1 Introduction

In this section we consider a particular example of the model in (3.28), with the widely used Weibull baseline cumulative hazard, $H(y) = y^{\gamma}$ and the positive α-stable frailty (Oakes (1989)) with probability density function and Laplace transform given respectively by

$$f(z) = -\frac{1}{\pi z}\sum_{l=1}^{\infty}\frac{\Gamma(l\alpha+1)}{l!}\left(-z^{-\alpha}\right)^{l}\sin(l\alpha\pi);\ z>0,\ 0<\alpha<1, \tag{4.1}$$

and

$$L_z(s) = E\{e^{-sZ}\} = e^{-s^{\alpha}}. \tag{4.2}$$

For basic properties of the distribution given in (4.1), one can refer to Duchateau & Janssen (2007). The reason behind choosing positive stable frailty over gamma frailty or other frailty models is that from the survival analysis literature Shih (1998), Glidden (1999), Fan et al. (2000) have observed that the gamma frailty specification

*Some of the results of this chapter are published in Applied Stochastic Models in Business and Industry. Asha et al. (2017)

may not fit well. One important observation made by Hougaard (2000) is that the positive stable model induces high early dependence whereas the most popular gamma frailty model exhibits high late dependence. This property is well observed in familial relationships of the ages of onset of diseases with etiologic heterogeneity, where genetic cases occur early and long-term survivors are meekly correlated. The gamma frailty model has predictive hazard ratios which are time invariant and may not be suitable for modelling failures include genetic factors (Fine et al. (2003)). The rest of this chapter is organised in the following manner. Model formulation is presented in Section 4.2. Model properties are discussed in Section 4.3. Bivariate hazard gradient in Section 4.4, Cross ratio function in Section 4.5 is discussed in detail. Section 4.6 deals with the parameter estimation and we employ the profile likelihood technique to estimate the model parameters. A simulation study is conducted in Section 4.7 to show the efficiency of our estimation procedure. Two data sets are analysed in data analysis Section in 4.8 and 4.9. Finally we discuss the results and draw some conclusions in Section 4.10.

4.2 Model Formulation

Now, for $y_2 \geq y_1$ and $\theta_1 + \theta_2 \neq \theta_i'$ the model (3.22) reduces to

$$\begin{aligned} S((y_1, y_2)|\boldsymbol{X}) &= \left[1 - \frac{\theta_1}{\theta_1 + \theta_2 - \theta_2'}\right] e^{-[(\theta_1+\theta_2)y_2^\gamma e^{\boldsymbol{X}\boldsymbol{\beta}}]^\alpha} \\ &+ \left[\frac{\theta_1}{\theta_1 + \theta_2 - \theta_2'}\right] e^{-[(\theta_2' y_2^\gamma + (\theta_1+\theta_2-\theta_2')y_1^\gamma)e^{\boldsymbol{X}\boldsymbol{\beta}}]^\alpha} \\ &= (1 - k_{21})e^{-[\varphi_1(y_2)]^\alpha} + k_{21}e^{-[\varphi_{21}(y_1,y_2)]^\alpha}. \end{aligned} \tag{4.3}$$

Similarly, for $y_1 \geq y_2$

$$\begin{aligned} S((y_1, y_2)|\boldsymbol{X}) &= \left[1 - \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'}\right] e^{-[(\theta_1+\theta_2)y_1^\gamma e^{\boldsymbol{X}\boldsymbol{\beta}}]^\alpha} \\ &+ \left[\frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'}\right] e^{-[(\theta_1' y_1^\gamma + (\theta_1+\theta_2-\theta_1')y_2^\gamma)e^{\boldsymbol{X}\boldsymbol{\beta}}]^\alpha} \\ &= (1 - k_{12})e^{-[\varphi_1(y_1)]^\alpha} + k_{12}e^{-[\varphi_{12}(y_1,y_2)]^\alpha}. \end{aligned} \tag{4.4}$$

For $y_i \geq y_j$ and $\theta_1 + \theta_2 \neq \theta'_i$, the joint survival function for load share positive stable frailty model is given by

$$S(y_1, y_2|\boldsymbol{X}) = (1 - k_{ij})e^{-[\varphi_1(y_i)]^\alpha} + k_{ij}e^{-[\varphi_{ij}(y_1,y_2)]^\alpha}. \tag{4.5}$$

Proceeding in the same manner for $y_i > y_j$ and $\theta_1+\theta_2 = \theta_i$, the joint survival function given in (3.23) reduces to

$$S(y_1, y_2|\boldsymbol{X}) = e^{-[\varphi_1(y_i)]^\alpha} + e^{-[\boldsymbol{X\beta}]}\theta_2(y_i^\gamma - y_j^\gamma)\frac{\partial}{\partial\varphi_1(y_i)}e^{[\varphi_1(y_i)]^\alpha}. \tag{4.6}$$

The corresponding bivariate density function for $y_i > y_j$ given in (4.5) and (4.6) simplifies to

$$\begin{aligned} f(y_1, y_2|\boldsymbol{X}) &= \theta_i\theta'_j\alpha\gamma^2(y_1y_2)^{\gamma-1}e^{2\boldsymbol{X\beta}-[\varphi_{ij}(y_1,y_2)]^\alpha} \\ &\times [\varphi_{ij}(y_1, y_2)]^{\alpha-2}(1 + \alpha([\varphi_{ij}(y_1, y_2)]^\alpha - 1)), \end{aligned} \tag{4.7}$$

where, $\varphi_{ij}(y_1, y_2) = e^{\boldsymbol{X\beta}}\left(\theta'_i y_i^\gamma + (\theta_1 + \theta_2 - \theta'_i)y_j^\gamma\right)$ and $\varphi_1(y_i) = e^{\boldsymbol{X\beta}}(\theta_1 + \theta_2)y_i^\gamma,\ i \neq j = 1, 2$. The plot of survival function in (4.5) is given in Figure 4.1.

4.3 Properties

Remark 4.3.1. *Re-parameterize $\theta'_1 = a\theta_1,\ \theta'_2 = b\theta_2$; then the joint survival function given in* (4.5) *is a generalization of the Weibull extension of the bivariate exponential of Freund with positive stable frailty model (Hanagal (2011)) with covariates. Re-parameterizing* (4.5) *without covariates reduces to the model in Hanagal (2011) and is given by*

$$S(y_1, y_2) = \begin{cases} \frac{\theta_2(1-b)e^{-[(\theta_1+\theta_2)y_2^\gamma]^\alpha} + \theta_1 e^{-[(\theta_1+\theta_2(1-b))y_1^\gamma + b\theta_2 y_2^\gamma]^\alpha}}{\theta_1+\theta_2(1-b)}; & y_1 \leq y_2 \\ \frac{\theta_1(1-a)e^{-[(\theta_1+\theta_2)y_1^\gamma]^\alpha} + \theta_2 e^{-[(\theta_2+\theta_1(1-a))y_2^\gamma + a\theta_1 y_1^\gamma]^\alpha}}{\theta_2+\theta_1(1-a)}; & y_2 \leq y_1 \end{cases} \tag{4.8}$$

Property 4.3.1. *The survival distribution of $V = \min(Y_1, Y_2)$ is a Weibull distribu-*

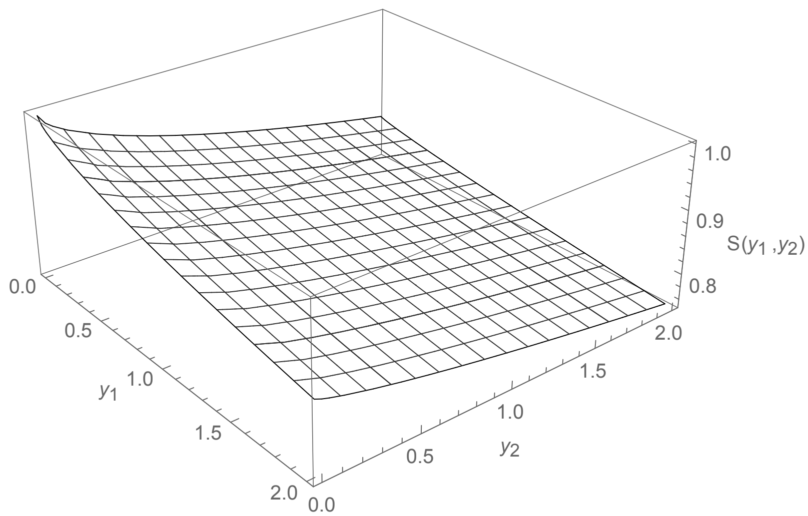


Figure 4.1: Plot for the survival function in (4.5) with parameters $\alpha = 0.8,\ \theta_1 = 0.05,\ \theta_2 = 0.07,\ \theta_1' = 0.09,\ \theta_2' = 0.11,\ \gamma = 0.8$ and $\beta = 0$.

tion:

$$S_V(v|\boldsymbol{X}) = e^{-[(\theta_1+\theta_2)^{\alpha} e^{\alpha \boldsymbol{X\beta}}] v^{\gamma\alpha}},\ v > 0. \tag{4.9}$$

Property 4.3.2. *The marginal distributions $S_{Y_i}(.)$, $i = 1, 2$ of (4.6) are Weibull mixtures for $\theta_1 + \theta_2 \neq \theta_i'$ given as*

$$\begin{aligned} S_{Y_i}(y_i|\boldsymbol{X}) &= [1 - k_{ij}] e^{-[(\theta_1+\theta_2)^{\alpha} y_i^{\gamma\alpha} e^{\alpha \boldsymbol{X\beta}}]} \\ &+ [k_{ij}] e^{-[(\theta_i')^{\alpha} y_i^{\gamma\alpha} e^{\alpha \boldsymbol{X\beta}}]},\ 0 < \alpha < 1,\ y_i > 0. \end{aligned} \tag{4.10}$$

and

$$S_{Y_i}(y_i|\boldsymbol{X}) = e^{-[\varphi_1(y_i)]^{\alpha}} - \left[e^{\boldsymbol{X\beta}} \theta_2 \left(\log S(y_i) \frac{\partial e^{-s^{\alpha}}}{\partial s} \right) \Bigg|_{s = e^{\boldsymbol{X\beta}} (\theta_1+\theta_2) y_i^{\gamma}} \right]$$

for $\theta_1 + \theta_2 = \theta_i'$, $0 < \alpha < 1$, $y_i > 0$.

4.4 Bivariate Hazard Gradient for Load Share Positive Stable Frailty Model

Johnson & Kotz (1975) hazard gradient in (1.6) and (1.7) for the distribution in (4.5) is obtained as,

$$r_i(y_1, y_2) = -\frac{\partial}{\partial y_i}\left[\log\left((1-k_{ij})e^{-[\varphi_1(y_i)]^\alpha} + k_{ij}e^{-[\varphi_{ij}(y_1,y_2)]^\alpha}\right)\right];\, y_i > y_j,\ \ i \neq j = 1, 2. \tag{4.11}$$

Similarly, for (4.6) Johnson and Kotz hazard gradient is obtained as,

$$r_i(y_1, y_2) = -\frac{\partial}{\partial y_i}\left[\log\left(e^{-[\varphi_1(y_i)]^\alpha} + e^{-[\boldsymbol{X\beta}]}\theta_2(y_i^\gamma - y_j^\gamma)\frac{\partial}{\partial\varphi_1(y_i)}e^{[\varphi_1(y_i)]^\alpha}\right)\right];\, y_i > y_j, \tag{4.12}$$

$i \neq j = 1, 2.$

The Cox hazard gradient (3.3) for the distribution in (4.5), reduces to

$$\begin{aligned}\lambda_{i0}(y) &= \frac{\alpha\gamma\theta_i\left(\varphi_1(y)\right)^\alpha}{y(\theta_1+\theta_2)},\ \ i = 1, 2,\ \ y > 0\\ \lambda_{ij}(y_i|y_j) &= \frac{y_i^{\gamma-1}\gamma\theta_i'[(1-\alpha)+\alpha(e^{\boldsymbol{X\beta}}\varphi_{ji}(y_1,y_2))^\alpha}{\varphi_{ji}(y_1,y_2)};\ y_i > y_j,\ \theta_1+\theta_2 > \theta_i',\ i \neq j = 1, 2.\end{aligned} \tag{4.13}$$

and

$$\lambda_{ij}(y_i|y_j) = \frac{\theta_1\theta_2'}{\theta_2}r(y_i)e^{\boldsymbol{X\beta}}\left[\frac{\partial}{\partial s}\log\frac{\partial}{\partial s}L_z(s)\bigg|_{s=\Psi_1(y_i)}\right];\, y_i > y_j,\ \theta_1+\theta_2 = \theta_i',\ i \neq j = 1, 2. \tag{4.14}$$

Figure 4.2 shows a plot of the bivariate hazard gradient given in (4.13) for some choices of parameter values. The following properties show that the monotone behavior of the failure rates of components depends on (a) the baseline hazard parameters and the frailty parameter when both components are functioning, and (b) only the baseline hazard parameters when one component fails, if $\theta_1 + \theta_2 > \theta_i'$, $i = 1, 2$.

Property 4.4.1. *The component failure rate $\lambda_{i0}(y)$, $i = 1, 2$ is monotonically increasing, constant, or decreasing at $\alpha\gamma > 1$, $\alpha\gamma = 1$ or $\alpha\gamma < 1$.*

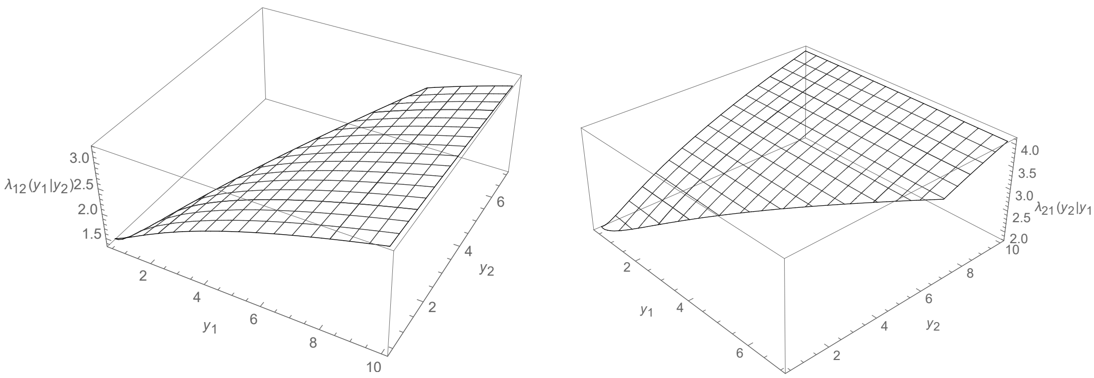


Figure 4.2: Plot for hazard gradients (a) $\lambda_{12}(y_1|y_2)$, (b) $\lambda_{21}(y_2|y_1)$ when $\theta_1 = 0.5$, $\theta_2 = 0.8$, $\theta_1' = 0.9$, $\theta_2' = 1.2$, $\gamma = 1.75$, $\alpha = 0.8$ and $\beta = 0$.

This is shown in Figure 4.3 for some choices of the parameter values.

Property 4.4.2. *Under the condition $\theta_1 + \theta_2 - \theta_i' > 0$, the component failure rate $\lambda_{ij}(y_i|y_j),\ i \neq j = 1, 2$ is increasing whenever the baseline distribution has increasing failure rate, or equivalently when $\gamma > 1$.*

Proof. From (4.13), observe that

$$\frac{\partial \lambda_{ij}(y_i|y_j)}{\partial y_i} = \left[(1-\alpha) + \alpha e^{\alpha \boldsymbol{X\beta}} (\varphi_{ji})^{\alpha}\right] \frac{d}{dy_i}\left[\frac{y_i^{\gamma-1}\gamma\theta_i'}{\varphi_{ji}}\right]. \tag{4.15}$$

It is straight forward to show that $\frac{d}{dy_i}\left[\frac{y_i^{\gamma-1}\gamma\theta_i'}{\varphi_{ji}}\right] \geq 0$ for $\gamma > 1$. Under the stated conditions, every term in (4.15) is positive, which proves the assertion. ■

4.5 The Cross Ratio Function for the Load Share Positive Stable Frailty Model

The CRF corresponding to (4.5) is given by

$$\mathscr{C}_1(y_1, y_2) = \frac{(1-\alpha[(1-(\varphi_{ij}(y_1,y_2))^{\alpha}])\theta_i' r(y_i) e^{\boldsymbol{X\beta}}\left[(1-k_{ij})e^{-[\varphi_1(y_i)]^{\alpha}} + k_{ij}e^{-[\varphi_{ij}(y_1,y_2)]^{\alpha}}\right]}{\varphi_{ij}(y_1,y_2)\left[\alpha r(y_i) e^{\boldsymbol{X\beta}}[(1-k_{ij})[\varphi_1(y_i)]^{\alpha}]e^{-[\varphi_1(y_i)]^{\alpha}}(\theta_1+\theta_2)+\theta_j[(\varphi_{ij}(y_1,y_2))^{\alpha}e^{-[\varphi_{ij}(y_1,y_2)]^{\alpha}}\right]}, \tag{4.16}$$

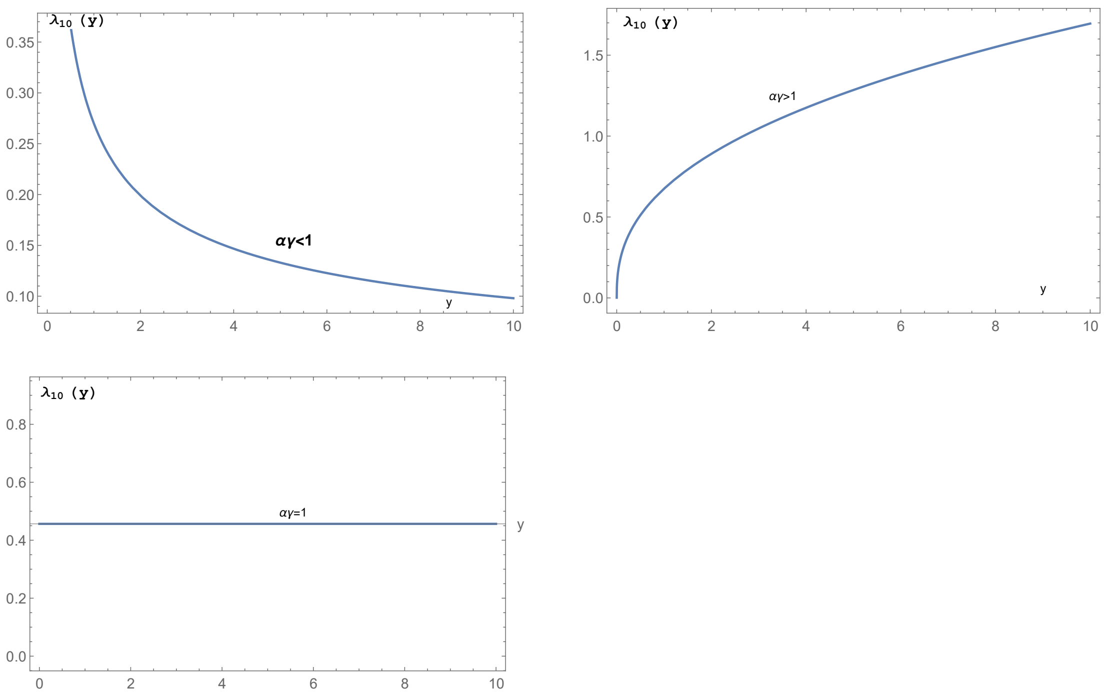


Figure 4.3: Plot for (a) $\lambda_{10}(y)$ when $(\alpha = 0.8,\ \gamma = 0.7)$, (b) $\lambda_{10}(y)$ when $(\alpha = 0.8,\ \gamma = 1.75)$, (c) $\lambda_{10}(y)$ when $(\alpha = 0.5,\ \gamma = 2.0)$ for $\theta_1 = 0.5$, $\theta_2 = 0.7$, $\theta_1' = 0.9$, $\theta_2' = 1.5$ and $\beta = 0$.

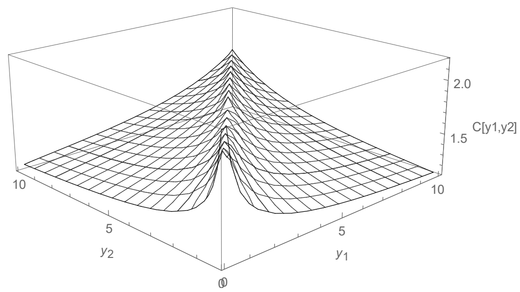
and the CRF corresponding to (4.6) is given by

$$\mathscr{C}_2(y_1, y_2) = \left.\frac{\frac{\partial}{\partial s}\log\left[\frac{\partial L_z(s)}{\partial s}\right]\left[L_z(s) - \left(e^{\boldsymbol{X\beta}}\theta_2\left[\log S(y_j) - \log S(y_i)\frac{\partial L_z(s)}{\partial s}\right]\right)\right]}{(\theta_1+\theta_2)r(y_i)e^{\boldsymbol{X\beta}}\left[\frac{\partial L_z(s)}{\partial s} - \theta_2 e^{\boldsymbol{X\beta}}\frac{\partial^2 L_z(s)}{\partial s^2}\left(\log S(y_j) - \log S(y_i)\right)\right]}\right|_{s=\Psi_1(y_i)}. \tag{4.17}$$

The CRF given in (4.16) is plotted for situations with and without positive stable frailty in Figure 4.4 for some choices of parameter values.

4.6 Parameter Estimation

In this section, we consider parameter estimation of the bivariate load share positive stable frailty model and covariates, where the lifetimes may be subject to censoring.



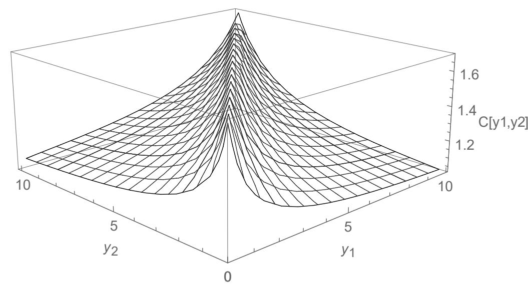


Figure 4.4: (a) CRF with positive stable frailty , (b) CRF without positive stable frailty when $\theta_1 = 0.7$, $\theta_2 = 0.9$, $\theta_1' = 1.2$, $\theta_2' = 1.5$, $\gamma = 0.7$, $\alpha = 0.8$ and $\beta = 0$.

Various approaches have been proposed in the literature for estimating the parameters involving positive stable frailty model. Wang et al. (1995) applied EM-algorithm for estimation in the positive stable frailty model with the frailties regarded as missing data. Fine et al. (2003) proposed a simple estimation procedure for a proportional hazards regression model for clustered survival data in which the dependence is generated by positive stable distribution. Martinussen & Pipper (2005) discussed a likelihood based estimation procedure for the positive stable frailty model. Mallick et al. (2008) developed Markov Chain Monte Carlo algorithms to facilitate Bayesian inference to estimate the parameters involved in a bivariate positive stable frailty model. Hanagal (2011) proposed two-stage ML method and BLUP method for bivariate Weibull extensions with positive stable frailty models.

Let $\tau = (\theta_1, \theta_2, \theta_1', \theta_2', \alpha, \beta, \gamma)$. Suppose that there are n independent pairs of components or organs under study, and the r^{th} pair of the components have lifetimes (y_{1r}, y_{2r}). Now, we want to estimate the parameters of the load share positive stable frailty model given in (4.7). The log-likelihood based on the sample of size n is given by

$$\begin{aligned}\ell(\tilde{\boldsymbol{\tau}}) &= n_1 \ \log \ \theta_1 + n_2 \ \log \ \theta_2' + n_2 \ \log \ \theta_1' + n_2 \ \log \ \theta_2 + n \ \log \ \alpha + 2n \ \log \ \gamma \\ &+ (\gamma - 1) \sum_{r=1}^{n} (\log \ y_{1r} + \log \ y_{2r}) + (\alpha - 2) \sum_{r=1}^{n_1} \log \ [\varphi_{12}(y_{1r}, y_{2r})] \\ &+ \sum_{r=1}^{n_1} \log \ [1 + \alpha[\varphi_{12}(y_{1r}, y_{2r})]^{\alpha} - 1] - \sum_{r=1}^{n_1} [\varphi_{12}(y_{1r}, y_{2r})]^{\alpha} - \sum_{r=1}^{n_2} [\varphi_{21}(y_{1r}, y_{2r})]^{\alpha} \\ &+ \sum_{r=1}^{n_2} \log [1 + \alpha([\varphi_{21}(y_{1r}, y_{2r})]^{\alpha} - 1)] + (\alpha - 2) \sum_{r=1}^{n_2} \log [\varphi_{21}(y_{1r}, y_{2r})] \, , \end{aligned} \tag{4.18}$$

for $y_{ir} \geq y_{jr}$, $i \neq j = 1, 2$. Given the observed data, y_{1r}, y_{2r} and $\boldsymbol{x}_r$, $r = 1, 2, ..., n$, we find the value of $\boldsymbol{\tau}$ that maximizes $\ell(\boldsymbol{\tau})$. Analytical solution of the likelihood equations and numerical maximization of the likelihood function are both extremely cumbersome. We propose the following estimation method. Let $\boldsymbol{\tau} = (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$ where $\boldsymbol{\tau}_1 = (\theta_1, \theta_2)$, $\boldsymbol{\tau}_2 = (\theta_1', \theta_2', \alpha, \beta, \gamma)$. We set $\theta_i = n[\sum_{r=1}^{n} y_{ir}]^{-1}$, $i = 1, 2$, to provide initial values of $\boldsymbol{\tau}_1$. Using "FindMaximum" function of Mathematica 10 we maximize the (profile) likelihood of $\boldsymbol{\tau}_2$ given the initial values of $\boldsymbol{\tau}_1$. Next, these estimates of $\boldsymbol{\tau}_2$ enable us to write the profile likelihood for the $\boldsymbol{\tau}_1$ which we maximize. This process is continued iteratively until all the estimates converge to yield the profile MLE $\widehat{\boldsymbol{\tau}}$. Hanagal (2011) showed that the two-stage MLE $\widehat{\boldsymbol{\tau}}$ has asymptotic properties similar to that of the MLE.

4.7 Simulation Study

We carried out a simulation study in order to evaluate the performance of the profile likelihood estimation. For simplicity we considered only a single covariate X_1 which follows a normal distribution with mean zero and variance 0.5. We assumed a Weibull baseline distribution.

Sample generation of (y_{1r}, y_{2r}), $r = 1, 2, ..., n$ was carried out by generating three sets of random samples of size n from the uniform $(0, 1)$ distribution; denote these by u_1, u_2, and u_3 respectively. We then generated a random sample of size n from a positive stable distribution with density given in (4.1), by using the model

$$z_r = E_r^{\left(\frac{-(1-\alpha)}{\alpha}\right)} (\sin(\xi_r))^{\frac{-1}{\alpha}} \times \sin(\alpha \xi_r) \times \sin[(1-\alpha)\xi_r]^{\frac{1-\alpha}{\alpha}} \tag{4.19}$$

(McKenzie (1982)). The covariate X_1 was generated from $N(0, \sigma^2)$, with $\sigma^2 = 0.5$. The bivariate sample (y_{1r}, y_{2r}), $r = 1, 2, ..., n$ for the distribution in (4.5) was generated by using the algorithm given in Asha et al. (2016). Thus, the samples are generated as follows:

- if $u_{1r} \leq \frac{\theta_1}{\theta_1+\theta_2}$, then $(y_{1r}|Z, X_1) = S_0^{-1}((1-u_{2r})^{\frac{1}{\theta_1}})$, and $(y_{2r}|Z, X_1) = S_0^{-1}(S_0(y_{1r}|Z, X_1)(1-u_{3r})^{\frac{1}{\theta_2'}})$. For Weibull baseline

$y_{1r} = [\frac{-1}{z_r\theta_1 e^{x_1\beta}}\log(1-u_{2r})]^{\frac{1}{\gamma}}$ and $y_{2r} = [y_{1r} + (\frac{-1}{z_r\theta_2' e^{x_1\beta}}\log(1-u_{3r}))^{\frac{1}{\gamma}}]$.

- if $u_{1r} > \frac{\theta_1}{\theta_1+\theta_2}$, then $(y_{2r}|Z, X_1) = S_0^{-1}((1-u_{2r})^{\frac{1}{\theta_2}})$, and $(y_{1r}|Z, X_1) = S_0^{-1}(S_0(y_2|Z, X_1)(1-u_{3r})^{\frac{1}{\theta_1'}})$. For Weibull baseline $y_{2r} = [\frac{-1}{z_r\theta_2 e^{x_1\beta}}\log(1-u_{2r})]^{\frac{1}{\gamma}}$ and $y_{1r} = [y_{2r} + (\frac{-1}{z_r\theta_1' e^{x_1\beta}}\log(1-u_{3r}))^{\frac{1}{\gamma}}]$.

We generated 1000 samples of sizes $n = 25$ and $n = 150$ from $S(y_1, y_2|z, X_1)$ in (4.5) with $S(y) = e^{-y^\gamma}$, $\alpha = 0.5$, $\beta = 0.5$, $\gamma = 0.7$, $\theta_1 = 0.3$, $\theta_2 = 0.5$, $\theta_1' = 0.9$ and $\theta_2' = 2.1$. The profile MLE of $\boldsymbol{\tau} = (\theta_1, \theta_2, \theta_1', \theta_2', \alpha, \beta, \gamma)$ is obtained. The average absolute bias across the 1000 samples was computed as $\frac{1}{n}\sum_{i=1}^{n}|(\widehat{\boldsymbol{\tau}}_i - \boldsymbol{\tau}_i)|$. The average root mean square error (RMSE) from the 1000 samples was calculated as $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{\boldsymbol{\tau}}_i - \boldsymbol{\tau}_i)^2}$.

The absolute biases, RMSEs are provided in Table 4.1. We observed from the simulation study that the absolute biases and RMSE's decrease as the sample sizes increase. we also observed that the rate of convergence improved with increasing sample size.

4.8 Data Set 1: Reliability of a System with Two Motors

We illustrate the load share model with frailty for a reliability data set which consists of a parallel system with two motors. When both motors function, the load is shared between them. If one of the motors fails, the entire load is then shifted to the surviving motor. The system fails when both motors fail. The data was originally published and analysed in Relia Soft, Reliability Edge Home ReliaSoft (2003). Recently Sutar & Naik-Nimbalkar (2014) analysed this data in a load sharing perspective using accelerated failure time (AFT) models and showed that the data satisfies load sharing properties. Table 4.2 shows the time to failure data for 18 such systems. This data set has no observed covariates and no censoring. We analysed this data using our model and showed that there is a dependence induced by frailty in the data apart from load sharing dependence. We estimated the parameters using the procedure explained in Section 4.6. The computation was carried out using "FindMaximum" function of

Table 4.1: Absolute bias, RMSE for $\tilde{\boldsymbol{\tau}} = (\theta_1, \theta_2, \theta_1', \theta_2', \alpha, \beta, \gamma)$ based on 1000 replications

Parameters	True values	Absolute bias	RMSE
		n=25	
θ_1	0.3	0.0639	0.0593
θ_2	0.5	0.0593	0.0262
θ_1'	0.9	0.0414	0.1104
θ_2'	2.1	0.0817	0.1661
α	0.5	0.0708	0.0840
γ	0.7	0.0839	0.1162
β	0.5	0.0366	0.0643
		n=150	
θ_1	0.3	0.0468	0.0402
θ_2	0.5	0.0314	0.0105
θ_1'	0.9	0.0265	0.0821
θ_2'	2.1	0.0670	0.1507
α	0.5	0.0654	0.0661
γ	0.7	0.0632	0.0856
β	0.5	0.0311	0.0492

Mathematica 10. We fit and compare the following models. Model 1: Load share positive stable frailty model, Model 2: Load share model (no frailty) and Model 3: Positive stable frailty model (no load share). The estimates with S.E., the lower and the upper limits (LCL and UCL) of the 95% confidence intervals and the Akaike information criterion (AIC) values for each model are provided in Table 4.3. Model 1 had the smallest AIC and gave the best in-sample fit for the motor data providing support for the existence of a frailty effect which contributes to the dependence apart from the dependence between the components induced by load sharing. We fitted the marginals of Model 1 given in (4.6) and the Kolmogorov-Smirnov (K-S) test revealed that both the marginals gave good fit for motor data. For the marginal, Y_1, the K-S test statistic was 0.3405 and for the marginal, Y_2, the value was 0.3612 and we accept the null-hypothesis that the model given in (4.6) fits well for the motor data at 0.01 level of significance since $D_{0.01,18} = 0.371$.

Table 4.2: Time to failure (in days) data set for two motors in a load sharing configuration

System	Time to failure for motor A (Y_1)	Time to failure for motor B (Y_2)	Event description
1	102	65	B Failed First
2	84	148	A Failed First
3	88	202	A Failed First
4	156	121	B Failed First
5	148	123	B Failed First
6	139	150	A Failed First
7	245	156	B Failed First
8	235	172	B Failed First
9	220	192	B Failed First
10	207	214	A Failed First
11	250	212	B Failed First
12	212	220	A Failed First
13	213	265	A Failed First
14	220	275	A Failed First
15	243	300	A Failed First
16	300	248	B Failed First
17	257	330	A Failed First
18	263	350	A Failed First

4.8.1 Cross Ratio Illustration with Motor Data

The cross ratio function for the best fitting model (Model 1) was given in (4.16). The parameter estimates are $\widehat{\theta_1} = 0.021$, $\widehat{\theta_2} = 0.051$, $\widehat{\theta_1'} = 0.275$, $\widehat{\theta_2'} = 0.299$, $\widehat{\gamma} = 0.732$, $\widehat{\alpha} = 0.642$ with AIC value 480.5. The best fitting model with just load share ignoring frailty (Model 2) is given by

$$f(y_1, y_2) = \theta_i'\theta_j\gamma^2 y_i^{\gamma-1} y_j^{\gamma-1} e^{-\theta_i' y_i^\gamma} e^{-(\theta_1+\theta_2-\theta_i')y_j^\gamma};\ y_i > y_j,\ i \neq j = 1, 2. \qquad (4.20)$$

The parameter estimates are $\widehat{\theta_1} = 0.021$, $\widehat{\theta_2} = 0.016$, $\widehat{\theta_1'} = 0.218$, $\widehat{\theta_2'} = 0.198$, $\widehat{\gamma} = 0.822$, with AIC value 494.9. The corresponding cross ratio function is given by

$$\mathscr{C}(y_1, y_2) = \frac{\theta_i \left\{ \theta_j e^{y_i^\gamma(2(\theta_1+\theta_2)-\theta_i')} - (\theta_i' - \theta_i) e^{y_i^\gamma(\theta_1+\theta_2)+y_j^\gamma(\theta_1+\theta_2-\theta_i')} \right\}}{(\theta_1+\theta_2)(\theta_i - \theta_i') e^{y_i^\gamma(\theta_1+\theta_2)+y_j^\gamma(\theta_1+\theta_2-\theta_i')} + \theta_i'\theta_j e^{y_i^\gamma(2(\theta_1+\theta_2)-\theta_i')}}. \qquad (4.21)$$

The cross ratio comparison for Model 1 and Model 2 is presented in Table 4.4. From the results we observe that the Model 1 exhibits the dependence between Y_1 and Y_2

more than two times for almost all the systems comparing to Model 2.

4.8.2 Model Selection Criteria

We are using AIC comparison for model selection. We also compared load share Gamma frailty model given in (3.34) with AIC 532.9. To reconfirm our chosen model, we propose a novel cross validation technique. Cross validation is a popular method for model evaluation. Here, we evaluate the conditional survival probabilities for three different models namely, load share positive stable frailty (Model 1), load share without positive stable frailty (Model 2) and positive stable frailty model without load sharing (Model 3). The model indicating least conditional survival probability at the observed data point is deemed to have a better fit for the data.

The conditional survival function for load share frailty model in (5.12) is given

$$S(y_i|y_j) = \frac{\theta_j \left[S(y_i)\right]^{z\theta_i' e^{\boldsymbol{X\beta}}} \left[S(y_j)\right]^{ze^{\boldsymbol{X\beta}}(\theta_1+\theta_2-\theta_i'-1)}}{(1-k_{ij})(\theta_1+\theta_2)\left[S(y_j)\right]^{ze^{\boldsymbol{X\beta}}(\theta_1+\theta_2)-1} + k_{ij}\theta_j' \left[S(y_j)\right]^{z\theta_j' e^{\boldsymbol{X\beta}-1}}}; y_i \geq y_j, \tag{4.22}$$

$i \neq j = 1, 2$. When $S(y) = e^{-y^\gamma}$ and the frailty random variable Z follows a positive stable distribution (4.22) becomes

$$S(y_i|y_j) = \frac{e^{\boldsymbol{X\beta} - [\varphi_{ij}(y_1,y_2)]^\alpha + [\varphi_1 y_i]^\alpha + \left(\theta_j' y_j^\gamma e^{\boldsymbol{X\beta}}\right)^\alpha} y_j^\gamma \theta_j [\varphi_{ij}(y_1, y_2)]^{\alpha-1}}{k_{ij}\left(e^{[\varphi_1 y_i]^\alpha}\theta_i \left[\theta_j' y_j^\gamma e^{\boldsymbol{X\beta}}\right]^{\alpha} + e^{[\varphi_1 y_i]^\alpha}[\varphi_1 y_i]^\alpha \theta_j' - \theta_j\right)}; y_i \geq y_j, i \neq j = 1, 2. \tag{4.23}$$

The conditional survival function for load share without frailty model is given as

$$S(y_i|y_j) = \frac{\theta_j \left[S(y_i)\right]^{\theta_i' e^{\boldsymbol{X\beta}}} \left[S(y_j)\right]^{e^{\boldsymbol{X\beta}}(\theta_1+\theta_2-\theta_i'-1)}}{(1-k_{ij})(\theta_1+\theta_2)\left[S(y_j)\right]^{e^{\boldsymbol{X\beta}}(\theta_1+\theta_2)-1} + k_{ij}\theta_j' \left[S(y_j)\right]^{\theta_j' e^{\boldsymbol{X\beta}-1}}}; y_i \geq y_j, \tag{4.24}$$

$i \neq j = 1, 2$. When $S(y) = e^{-y^\gamma}$, (4.24) becomes

$$S(y_i|y_j) = \frac{\theta_j(\theta_1+\theta_2-\theta_j') \exp\left[e^{\boldsymbol{X\beta}}\left(y_j^\gamma(\theta_i'+\theta_j') - \theta_i' y_i^\gamma\right)\right]}{\theta_i\theta_j' \exp\left[(\theta_1+\theta_2) e^{\boldsymbol{X\beta}} y_j^\gamma\right] + (\theta_1+\theta_2)(\theta_j-\theta_j') \exp\left[e^{\boldsymbol{X\beta}}\theta_j' y_j^\gamma\right]}. \tag{4.25}$$

The conditional survival function for frailty without load share model in (3.29) with

$S(y) = e^{-y^\gamma}$ and the frailty random variable Z follows a positive stable distribution is given as

$$S(y_i|y_j) = e^{-\left[(y_i^\gamma + y_j^\gamma)^\alpha - (y_i^\gamma)^\alpha\right]} y_i^{-\gamma(\alpha-1)} (y_i^\gamma + y_j^\gamma)^{\alpha-1};\ y_i \geq y_j, i \neq j = 1, 2. \qquad (4.26)$$

In the leave-one-out cross-validation method, each training set is created by taking all the units except one, the test set being the sample unit that is held out. Thus, for n samples, we have n different training sets and n different test sets. This cross-validation procedure is especially useful with small data sets. For each training set, the estimated parameters are used to compute conditional survival probability of failure of the second component of the hold-out unit given that the first component had failed. The best model will correspond to the smallest conditional survival probability.

Table 4.5 presents a comparison between these three modes using the leave-one-out cross-validation: Load share positive stable frailty model (Model 1), load share without positive stable frailty model (Model 2), and positive stable frailty model without load sharing (Model 3). Column 1 shows the failure times for the held-out system. Columns 2-4 show the conditional survival probabilities from (4.23), (4.25) and (4.26) respectively. From the results we observe that in fifteen out of eighteen systems (83.33%) the load sharing with frailty model performs well. Figure enables a good visual comparison of these conditional survival probabilities from the three models and overwhelmingly supports the load share model with frailty.

4.9 Data Set 2: Diabetic Retinopathy Study Data

The second illustration is a well analysed Diabetic Retinopathy Study (DRS) data (Huster et al. (1989)). This data concerns the time for onset of blindness of human eyes. Though well studied this has not been analyzed incorporating both the load sharing and frailty perspectives together so far. Diabetic retinopathy is a complication associated with diabetes mellitus consisting of abnormalities in the micro vasculature within the retina of the eye. The Diabetic Retinopathy Study (DRS) was began in 1971 to study the effectiveness of laser photocougulation in delaying the onset of blindness in patients with diabetic retinopathy. The study was mainly conducted for

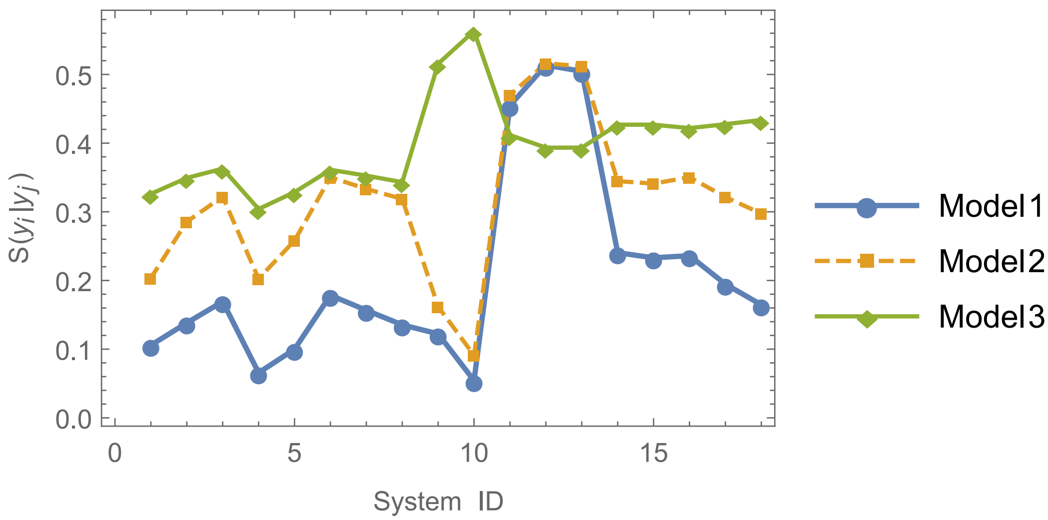


Figure 4.5: Comparison of conditional survival probabilities for the three models with motor data.

patients affected with diabetic retinopathy in both eyes and visual acuity of $\frac{20}{100}$ or better in both eyes.

Each patient had one eye randomized to laser treatment and the other eye kept without treatment for observation. The time to event for each eye was the time from initiation of the treatment to the time when visual acuity dropped below $\frac{5}{200}$ in two consecutive visits referred as "blindness". There is a built-in lag time of 6.5 months. The visits were every 4 months. For data analysis purpose survival times are therefore considered with the actual time to blindness in months, subtracted by the smallest possible time-to-event, which was estimated as 6.5 months.

Huster et al. (1989) analysed the data set and indicated possible treatment effect in delaying the onset of blindness since by ignoring censoring the sample mean survival time for the treated eye is 38.87 months while the same for the untreated eye is 32.29 months. Huster et al. (1989) considered exponential and Weibull as possible marginals distributions and showed Weibull marginals fit well for the data comparing to exponential marginals. They also proved that there exists positive correlation between failure times of the two eyes. Again, Sahu & Dey (2000) analyzed this data and showed that the failure rate for the adult patients is higher than the failure rate for the juveniles, on average.

We further analyse the data using our model. We observe that when the untreated eye fails first the failure rate of the treated eye is increasing. Similarly when the treated eye fails first the failure rate of the untreated eye is increasing. There is a positive association induced by the frailty 'genetic factor' in time for onset of blindness for treated eye and the untreated eye. The age covariate has an effect on time to blindness. These measures further confirm the findings of Huster et al. (1989). Though, this data has been analyzed by many authors but never been looked upon in a load sharing perspective.

For our analysis we considered the complete data set presented in Huster et al. (1989) except simultaneous failures ($N = 181$). In our model the basic assumption is that initially for given frailty, the eyes have independent failure rates with dependency arriving only at the time one of these two eyes fails. If the untreated eye fails first, we investigate if the treated eye is disturbed. Similarly, when treated eye fails first, we observe how the failure rate of the untreated eye alters. Also, we study the dependence induced by the frailty. Our main interest here is to estimate the parameters $\boldsymbol{\tau} = (\theta_1, \theta_2, \theta_1', \theta_2', \alpha, \gamma, \beta)$.

4.9.1 Likelihood Contributions

Let Y_1 be the onset of blindness for the treated eye and Y_2 be the onset of blindness for the untreated eye. For each data point (y_{1r}, y_{2r}) (excluding simultaneous failues)one of the following five censoring situations can happen:

(i). $y_{1r} < y_{2r}$; treated eye fails before untreated eye and eventually both of them fail.

(ii). $y_{1r} > y_{2r}$; untreated eye fails before treated eye and eventually both of them fail.

(iii). $y_{1r} \leq w_r < y_{2r}$; the untreated eye is censored at w_r and the treated eye fails before w_r.

(iv). $y_{2r} \leq w_r < y_{1r}$; the treated eye is censored at w_r and the untreated eye fails before w_r.

(v). $w_r < \text{Min}(y_{1r}, y_{2r})$; both eyes are censored at the same time.

The likelihood contributions for the model given in (4.6)are as follows:

$$\begin{aligned} f_{1r} &= k_{12}\ \frac{\partial^2 L_z\left(\Psi_{12}(y_{1r}, y_{2r})\right)}{\partial y_{1r}\partial y_{2r}};\ \ y_{2r} < y_{1r} < w_r, \\ &= \theta_1\theta_2'\alpha\gamma^2(y_{1r}y_{2r})^{\gamma-1}e^{\alpha\boldsymbol{X\beta}-[\varphi_{21}(y_{1r},y_{2r})]^\alpha}[\varphi_{21}(y_{1r}, y_{2r})]^{\alpha-2} \\ &\times\left(1+\left(\alpha[\varphi_{21}((y_{1r}, y_{2r})]^\alpha - 1]\right)\right);\ y_{1r} < y_{2r} \leq w_r, \end{aligned} \tag{4.27}$$

$$\begin{aligned} f_{2r} &= k_{21}\ \frac{\partial^2 L_z\left(\Psi_{21}(y_{1r}, y_{2r})\right)}{\partial y_{1r}\partial y_{2r}};\ \ y_{1r} < y_{2r} < w_r, \\ &= \theta_1'\theta_2\alpha\gamma^2(y_{1r}y_{2r})^{\gamma-1}e^{\alpha\boldsymbol{X\beta}-[\varphi_{12}(y_{1r},y_{2r})]^\alpha}[\varphi_{12}(y_{1r}, y_{2r})]^{\alpha-2} \\ &\times\left(1+\left(\alpha[\varphi_{12}((y_{1r}, y_{2r})]^\alpha - 1]\right)\right);\ y_{2r} < y_{1r} \leq w_r, \end{aligned} \tag{4.28}$$

$$\begin{aligned} f_{3r} &= \int_{y_2}\theta_i'\theta_j r(w_r)r(y_2)e^{2\boldsymbol{X\beta}}\left[\frac{\partial^2 L_z(s)}{\partial s^2}\right]_{s=\Psi_{ij}(w_r,y_2)} \\ &= \theta_1\alpha\gamma(y_{1r})^{\gamma-1}e^{\boldsymbol{X\beta}-[\varphi_{21}(y_{1r},w_r)]^\alpha}\times[\varphi_{21}(y_{1r}, w_r)]^{\alpha-1};\ y_{1r} \leq w_r < y_{2r}, \end{aligned} \tag{4.29}$$

$$\begin{aligned} f_{4r} &= \int_{y_1}\theta_i'\theta_j r(y_1)r(w_r)e^{2\boldsymbol{X\beta}}\left[\frac{\partial^2 L_z(s)}{\partial s^2}\right]_{s=\Psi_{ij}(y_1,w_r)};\ \ y_{1r} < w_r < y_{2r}, \\ &= \theta_2\alpha\gamma(y_{2r})^{\gamma-1}e^{\boldsymbol{X\beta}-[\varphi_{12}(w_r,y_{2r})]^\alpha}[\varphi_{12}(w_r, y_{2r})]^{\alpha-1};\ y_{2r} \leq w_r < y_{1r}, \end{aligned} \tag{4.30}$$

and

$$S(w_r, w_r|\boldsymbol{X}) = e^{-[(\theta_1+\theta_2)w_r^\gamma e^{\boldsymbol{X\beta}}]^\alpha}. \tag{4.31}$$

4.9.2 Data Analysis

We have considered one covariate for the analysis, type of diabetes (1-Juvenile, 2-Adult) and it is denoted by X_1 with censoring. We analysed this data using our model and showed that there is a dependence induced by frailty in the data apart from load sharing dependence. We estimated the parameters using the procedure explained in Section 4.6. The computation was carried out using "FindMaximum" function of Mathematica 10. We fit and compare the following models. Model 1: Load share positive stable frailty model, Model 2: Load share model (no frailty) and Model 3: Positive stable frailty model (no load share). The estimates with S.E., the lower and the upper limits (LCL and UCL) of the 95% confidence intervals and the Akaike information criterion (AIC) values for each model are provided in Table 4.6. From Table 4.6 we obtained the estimates as $\hat{\theta}_1 \simeq 0.1927$, $\hat{\theta}_2 \simeq 0.3100$, $\hat{\theta}'_1 \simeq 0.9366$, $\hat{\theta}'_2 \simeq 0.9937$, $\hat{\gamma} \simeq 0.3340$, $\hat{\alpha} \simeq 0.1464$ and $\hat{\beta} \simeq 0.7286$. From the analysis, we conclude that when the untreated eye fails first, $\theta_1 = 0.1927$ increases to $\theta'_1 = 0.9366$. Similarly, when the treated eye fails first, $\theta_2 = 0.3100$ increases to $\theta'_2 = 0.9937$. This from (3.4) signifies that the failure rate of the surviving eye is increasing.

Model 1 had the smallest AIC and gave the best in-sample fit for the diabetic retinopathy data providing support for the existence of a frailty effect which contributes to the dependence apart from the dependence between the organs induced by load sharing. The covariate '$\beta = 0.7286 \simeq e^{\beta}$=2.07' times, indicates that the type of diabetic shows a positive impact on the onset of blindness. Since the age covariate has value 1 for juvenile and 2 for adult patients and β is positive, we can conclude that the onset of blindness for adults is 2 times faster than that of the juvenile patients. These findings append the findings of Huster et al. (1989), Hanagal & Sharma (2011) and Sahu & Dey (2000).

The K-S test reveals that both the marginals fit well for the diabetic retinopathy data. For the marginal Y_1 the K-S test statistic value is 0.15 which is less than $D_{0.01,42} = 0.2403$, therefore we accept the null hypothesis at 0.01 significance level that the data follows the distribution in (4.7). Similarly, for the marginal Y_2 the K-S test statistic value is 0.1774 which is less than $D_{0.01,82} = 0.1800$. Therefore we accept the null hypothesis at 0.01 significance level that the data follows the distribution in (4.7).

4.10 Discussion and Summary

We have proposed a load share model with frailty to explain the dependence among bivariate failure times. On failure of one component, the surviving component may have extra load which also increases the stress level until the failure of the second component. For various choices of the failure time distributions, we proposed families of bivariate distributions which include load sharing, frailty, and covariates in (3.28). For various choices of the parameters θ_i, θ_i', $i = 1, 2$ and without frailty or covariates, our model reduces to the models given by Freund (1961), Lu (1989), Asha et al. (2016) and Hanagal (2011).

From the analysis of the data set 1 we observed that the frailty plays a significant role in the dependence between failure times, apart from the load sharing dependence. The cross ratio function showed that there is a positive association in the failure times for motor A and motor B. The AIC also supports the load share positive stable frailty model (Model 1). The K-S test revealed that both the marginals fit well under this model.

From the cross validation we observed that in most of the cases (83.33%) the load share frailty model gives least conditional survival probability thereby predicting the failures of the second component more accurately than the other two models. Notice that for motors representing the system IDs 12, 13 and 14, both the models incorporating load share perform poorly compare to the positive stable frailty model. We suspect that it is more likely that some external force had affected the three consecutive motors in the systems mentioned. It is desirable to further investigate the circumstances under which these motors were performed.

From the data analysis for data set 2 we observed that the frailty plays a significant role in the dependence between failure times, apart from the load sharing dependence. Also, we observed that the failure of the treated eye increases the failure rate of the untreated eye. Similarly, the failure of the untreated eye increases the failure rate of the treated eye. These findings append the findings of Huster et al. (1989), Sahu et al. (1997) and Hanagal & Sharma (2011). The full model, that is, load share with positive stable frailty model is the best fitted model and the AIC also supports this claim. The K-S test revealed that both the marginals fit well under this model.

As extensively discussed, in a load share system, failure of a component affects the stochastic behaviour of the surviving components thereby increasing/decreasing the failure rate. In many practical situations there is a critical time after which the surviving components regain their original failure rate. This phenomenon is widely seen in many types of systems, including power transmission, computer networking, finance, human bodily systems, bridges and so on. We refer these type of failures involving in a system as cascading failures and is discussed in Lindley & Singpurwalla (2002). In the next chapter we attempt to model data sets from such a system.

Table 4.3: Estimates for reliability models for the motor data

Model	Parameter Estimates						
	$\widehat{\theta}_1$	$\widehat{\theta}_2$	$\widehat{\theta}'_1$	$\widehat{\theta}'_2$	$\widehat{\gamma}$	$\widehat{\alpha}$	AIC
			Model 1				
Estimates	0.021	0.051	0.275	0.299	0.732	0.642	480.5
S.E	0.002	0.006	0.03	0.028	0.016	0.025	
LCL	0.017	0.038	0.216	0.245	0.683	0.610	
UCL	0.024	0.063	0.333	0.354	0.780	0.674	
			Model 2				
Estimates	0.021	0.016	0.218	0.198	0.822		494.9
S.E	0.002	0.001	0.003	0.003	0.011		
LCL	0.017	0.014	0.212	0.193	0.801		
UCL	0.024	0.019	0.224	0.204	0.843		
			Model 3				
Estimates					0.762	0.227	558
S.E					0.056	0.018	
LCL					0.653	0.113	
UCL					0.872	0.331	

Table 4.4: Cross ratio comparison for motor data

Y_1	Y_2	CR for Model 1	CR for Model 2	Ratio
102	65	4.96	1.73	2.85
156	121	5.77	1.80	3.21
148	123	6.71	1.83	3.66
245	156	4.21	1.70	2.48
235	172	4.88	1.76	2.78
220	192	7.03	1.85	3.79
250	212	6.30	1.83	3.43
300	248	5.70	1.81	3.14
84	148	4.02	1.67	2.41
88	202	3.55	1.58	2.24
139	150	6.30	1.89	3.34
207	214	6.64	1.91	3.47
212	220	6.56	1.91	3.43
213	265	4.79	1.79	2.67
220	275	4.75	1.79	2.65
243	300	4.76	1.79	2.65
257	330	4.53	1.77	2.56
263	350	4.36	1.75	2.49

Table 4.5: Cross validation comparison for motor data

Deleted sample point	Model 1	Model 2	Model 3
(101, 65)	0.1058	0.2034	0.3261
(156, 121)	0.1383	0.2857	0.3493
(148, 123)	0.1693	0.3220	0.3628
(245, 156)	0.0656	0.2026	0.3041
(235, 172)	0.0998	0.2581	0.3286
(220, 192)	0.1785	0.3504	0.3611
(250, 212)	0.1567	0.3332	0.3526
(300, 248)	0.1354	0.3188	0.3432
(84, 148)	0.1223	0.1620	0.5156
(88, 202)	0.0543	0.0910	0.5632
(139, 150)	0.4549	0.4702	0.4117
(207, 214)	0.5135	0.5162	0.3857
(212, 220)	0.5046	0.5119	0.3934
(213, 265)	0.2404	0.3446	0.3750
(220, 275)	0.2330	0.3411	0.3739
(243, 300)	0.2362	0.3504	0.3722
(257, 330)	0.1947	0.3215	0.3687
(263, 350)	0.1646	0.2974	0.3660

Table 4.6: Estimates for survival models for the diabetic retinopathy study data

Model	$\widehat{\theta_1}$	$\widehat{\theta_2}$	$\widehat{\theta'_1}$	$\widehat{\theta'_2}$	$\widehat{\gamma}$	$\widehat{\beta}$	$\widehat{\alpha}$	AIC
			Model 1					
Estimates	0.1927	0.3100	0.9366	0.9937	0.3340	0.7286	0.1464	1899.96
S.E	0.0073	0.0115	0.0493	0.0340	0.0045	0.0561	0.0015	
LCL	0.1784	0.2957	0.8399	0.9270	0.3253	0.6187	0.1435	
UCL	0.2070	0.3243	1.0333	1.0604	0.3427	0.8385	0.1493	
			Model 2					
Estimates	0.7548	0.8297	0.9684	1.0081	0.4525	-1.7174		1952.388
S.E	0.0177	0.0096	0.0171	0.0222	0.0023	0.0093		
LCL	0.7201	0.7950	0.9350	0.9646	0.4475	-1.6993		
UCL	0.7895	0.8644	1.0018	1.0516	0.4575	-1.7355		
			Model 3					
Estimates					0.3930	0.7517	0.4344	3053.8
S.E					0.0022	0.0021	0.0012	
LCL					0.3889	0.7475	0.4323	
UCL					0.3971	0.7559	0.4364	

Chapter 5

Modelling Cascading Failure Data

5.1 Introduction

Cascading failures are failures where an initial failure alters the structure function of the system which triggers a series of subsequent failures. These type of failures are very common in electrical power grids (Chang & Wu (2011)), paired organs in biological sciences (Gross et al. (1971)), computer networking (Epema et al. (1996)), financial institutions and banks (Wheelock & Wilson (2000)), bridges (Komatsu & Sakimoto (1977)) and sports (Kim & Kvam (2004)) to name a few.

Lindley & Singpurwalla (2002) discussed the concept of cascading failures within the framework of reliability theory. They extended the Freund's bivariate exponential model to model cascading failures.

The Freund's bivariate exponential model provides a suitable framework for developing models of cascading failures. We recall that in the Freund's bivariate exponential model the lifetimes of the two-components behave as if they are independent, until one of the components fail, after which the remaining component suffers an increased/decreased stress. These types of models are also referred popularly as load sharing models. For more on load sharing models and its generalizations we refer

[*]Some of the results of this Chapter are published in Communications in Statistics-Simulation and Computation. Asha & Raja (2017)

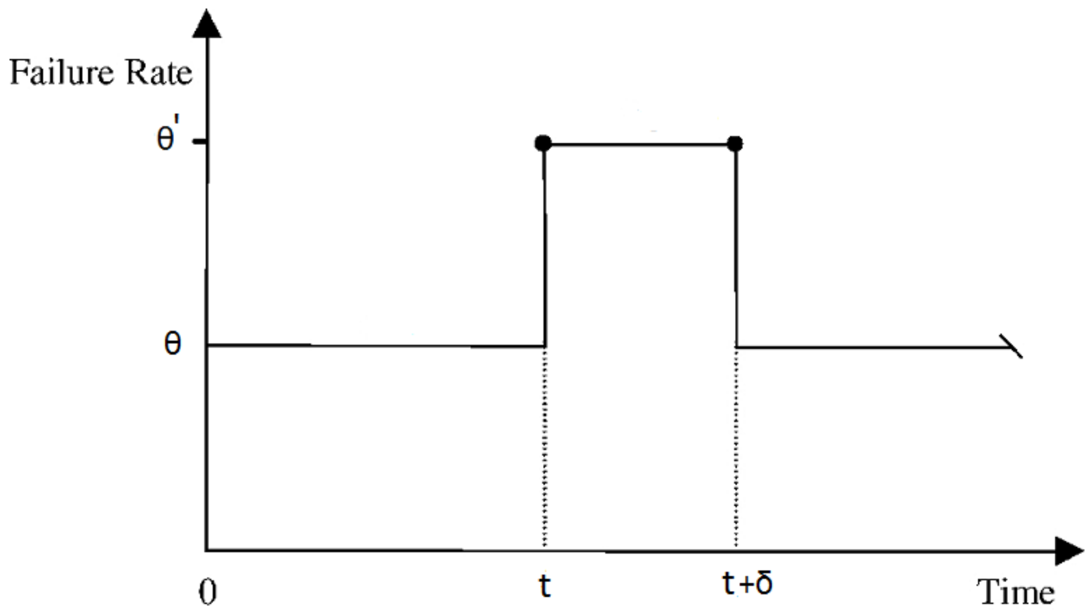


Figure 5.1: Failure rate of the second component to fail in an exponential cascade model.

to Daniels (1945), Rosen (1964), Coleman (1958), Gross et al. (1971), Lynch (1999), Singpurwalla (1995) Hollander & Peña (1995), Peña (2006), Kim & Kvam (2004), McCool (2006), Deshpande et al. (2010) and see the references cited therein.

In Lindley & Singpurwalla (2002) the Freund's model is modified such that the change in the parameter of the surviving component reverts back to the original value after a threshold time. Assuming that components are identically distributed as exponential θ and the renewed parameter is θ', then for a threshold time δ on failure of the first component at time point t, the failure rate is as illustrated in Figure 5.1. Generalization of the above model to cover monotone failure rate functions, multiple components and random values of the threshold time δ is discussed in Swift (2008).

Instead of arbitrarily distributed lifetime we modify the model in Swift (2008) by considering the components to have a proportional hazards model. The model formulation is done by applying the Cox proportional hazard model discussed in Section 1.2 and is presented in Section 5.2. In Section 5.3 we discuss the exponential cascading failure model as a special case of the model in Section 5.2. This is shown to generalize the models of Lindley & Singpurwalla (2002) and Swift (2008). In Section 5.4 we employ the method of L-moments and method of moments estimation to estimate the parameters of the model. A simulation study has been discussed in Section 5.5. The usefulness of our model is illustrated by considering a cricket data. Cascading failures are common in the game of cricket. "Cascading of wickets" and "Cascading failure of a team" are some of the phrases used by the experts to define sequence of batting failures and sequence of losses by a particular team in a tournament respectively. For our model we considered the time spent by two opening batters at the crease and if one opener fails, how the failure rate of the other opening batsman changes is explained through a data analysis. We analyse a data set from www.espncricinfo.com ESPNCricinfo (2013) and is analyzed using this model in Section 5.6.

5.2 Model Formulation and Properties

We consider the system discussed in Section 3.1 with some modifications. Let T_1 and T_2 be random variables representing the lifetimes of components A and B respectively in a two-component parallel system when they are first put on a test. If component B fails before A, or equivalently if $T_1 > T_2$, the lifetime distribution of A changes up to a critical period δ. Let it be denoted by T_1^*. After the time interval δ the lifetime of A is governed by the original random variable T_1. Finally the system fails when component A fails. The same applies for B when A fails first with an analogous explanation for T_2^*.

In order to derive a general class of bivariate distributions, it is further assumed here that T_1 and T_2 are independently distributed having survival functions of the form $[S(.)]^{\theta_1}$ and $[S(.)]^{\theta_2}$ respectively. It is further assumed that T_1^* and T_2^* have survival functions $[S(.)]^{\theta_1'}$ and $[S(.)]^{\theta_2'}$ respectively. If we denote the lifetimes of A and B as Y_1 and Y_2, then the dependence between Y_1 and Y_2 is essentially such that the failure of the component B changes the parameter of the life distribution of A from θ_1 to θ_1' up to a critical time $\delta > 0$, while the failure of the component A, changes the parameter of the life distribution of the component B from θ_2 to θ_2' up to a critical time $\delta > 0$. This dependency, which is explained in (1.9),(1.10),(1.11) and

(3.4) reflects in the failure rate behaviour of the component as follows. If $r(.)$ denotes the baseline failure rate corresponding to $S(.)$, then

$$\begin{aligned}
\lambda_{10}(y) &= \theta_1 r(y),\ y \geq 0 \\
\lambda_{20}(y) &= \theta_2 r(y),\ y \geq 0 \\
\lambda_{12}(y_1|y_2) &= \theta_1' r(y_1),\ y_2 < y_1 < y_2 + \delta \\
\lambda_{12}(y_1|y_2) &= \theta_1 r(y_1),\ y_1 \geq y_2 + \delta \\
\lambda_{21}(y_2|y_1) &= \theta_2' r(y_2),\ y_1 < y_2 < y_1 + \delta \\
\lambda_{21}(y_2|y_1) &= \theta_2 r(y_2),\ y_2 \geq y_1 + \delta
\end{aligned} \tag{5.1}$$

where $\lambda_{10}(y),\ \lambda_{20}(y),\ \lambda_{21}(y_2|y_1),\ \lambda_{12}(y_1|y_2)$ are failure rate functions (Cox (1972)) defined as

$$\begin{aligned}
\lambda_{i0}(y) &= \lim_{\Delta y \to 0+} \frac{P\left(y \leq Y_i < y + \Delta y | y \leq Y_1, y \leq Y_2\right)}{\Delta y},\ i = 1, 2 \\
\lambda_{21}(y|u) &= \lim_{\Delta y \to 0+} \frac{P\left(y \leq Y_2 < y + \Delta y | y \leq Y_2, Y_1 = u\right)}{\Delta y},\ u < y
\end{aligned} \tag{5.2}$$

with a similar definition for $\lambda_{12}(y|u)$. In terms of (5.2) the bivariate probability density function $f(y_1, y_2)$ is given by (Cox (1972))

$$\begin{aligned}
f(y_1, y_2) = \exp&\left[-\int_0^{y_1 - 0} \{\lambda_{10}(u) + \lambda_{20}(u)\} du - \int_{y_1 + 0}^{y_2 - 0} \lambda_{21}(u|y_1) du\right] \\
&\times \lambda_{10}(y_1)\, \lambda_{21}(y_2\,|\,y_1),\ y_2 \geq y_1\,.
\end{aligned} \tag{5.3}$$

For $y_1 < y_2 < y_1 + \delta$, the underlying model is

$$\begin{aligned}
f(y_1, y_2) &= \exp\left[-\int_0^{y_1} [\theta_1 r_0(y) + \theta_2 r_0(y)] du + \int_{y_1}^{y_2} [\theta_2' r_0(y_2)] dy_2\right] \theta_1 r_0(y_1) \theta_2' r_0(y_2) \\
&= \exp\left[-(\theta_1 + \theta_2) \int_0^{y_1} r_0(y) du\right] \exp\left[-\int_0^{y_2} \theta_2' r_0(y_2) dy_2\right] \\
&\times \exp\left[\int_0^{y_1} \theta_2' r_0(y - 2) du\right] \theta_1 \frac{f(y_1)}{S(y-1)} \theta_2' \frac{f(y_2)}{S(y_2)} \\
&= [S(y_1)]^{\theta_1 + \theta_2} [S(y_2)]^{\theta_2'} [S(y_1)]^{-\theta_2' - 1} \theta_1 \theta_2' f(y_1) f(y_2) [S(y_2)]^{-1}.
\end{aligned} \tag{5.4}$$

Therefore for $y_1 < y_2 < y_1 + \delta$

$$f(y_1, y_2) = \theta_1\theta_2' f(y_1)f(y_2)[S(y_1)]^{\theta_1+\theta_2-\theta_2'-1}[S(y_2)]^{\theta_2'-1},\ y_1 < y_2 < y_1 + \delta. \quad (5.5)$$

Now for $y_2 > y_1 + \delta$, the bivariate probability density function $f(y_1, y_2)$ becomes

$$f(y_1, y_2) = \exp\left[-\int_0^{y_1}\{\lambda_{10}(u) + \lambda_{20}(u)\}du - \int_{y_1}^{y_2}\lambda_{21}(u|y_1)du\right]\lambda_{10}(y_1)\,\lambda_{21}(y_2\,|\,y_1),\ y_2 \geq y_1 \quad (5.6)$$

$$f(y_1, y_2) = \exp\left[-(\theta_1 + \theta_2)\int_0^{y_1} r_o(y)dy\right]\exp\left[-\int_{y_1}^{y_1+\delta}\theta_2' r_0(u)du - \int_{y_1+\delta}^{y_2}\theta_2 r_0(u)du\right]$$
$$\times\,\theta_1 f(y_1)[S(y_1)]^{-1}\theta_2 f(y_2)[S(y_2)]^{-1}. \quad (5.7)$$

Recalling that

$$S(y) = e^{-\int_0^y r_0(u)du},$$

$$f(y_1, y_2) = \theta_1\theta_2 f(y_1)f(y_2)[S(y_1)]^{\theta_1+\theta_2-\theta_2'-1}[S(y_2)]^{\theta_2-1}[S(y_1+\delta)]^{\theta_2'-\theta_2},\ y_2 > y_1 + \delta. \quad (5.8)$$

Similarly, we can write the bivariate density function $f(y_1, y_2)$ for the region $y_2 < y_1 < y_2 + \delta$ as

$$f(y_1, y_2) = \theta_2\theta_1' f(y_1)f(y_2)[S(y_2)]^{\theta_1+\theta_2-\theta_1'-1}[S(y_1)]^{\theta_1'-1},\ y_2 < y_1 < y_2 + \delta. \quad (5.9)$$

And for the region $y_1 > y_2 + \delta$ the bivariate density function $f(y_1, y_2)$ is given by

$$f(y_1, y_2) = \theta_1\theta_2 f(y_1)f(y_2)[S(y_2)]^{\theta_1+\theta_2-\theta_1'-1}[S(y_1)]^{\theta_1-1}[S(y_2+\delta)]^{\theta_1'-\theta_1},\ y_1 > y_2 + \delta. \quad (5.10)$$

Hence the model (5.1) is retrieved from equations (5.5), (5.8), (5.9) and (5.10). Hence the joint probability function of (Y_1, Y_2) for the underlying model is derived as

$$f(y_1, y_2) = \begin{cases} \theta_2\,\theta_1'\,f(y_2)f(y_1)[S(y_2)]^{\theta_1+\theta_2-\theta_1'-1}[S(y_1)]^{\theta_1'-1}, y_2 < y_1 < y_2+\delta \\ \theta_1\,\theta_2\,f(y_2)f(y_1)[S(y_2)]^{\theta_1+\theta_2-\theta_1'-1}[S(y_1)]^{\theta_1-1}[S(y_2+\delta)]^{\theta_1'-\theta_1}, y_1 > y_2+\delta \\ \theta_1\,\theta_2'\,f(y_2)f(y_1)[S(y_1)]^{\theta_1+\theta_2-\theta_2'-1}[S(y_2)]^{\theta_2'-1}, y_1 < y_2 < y_1+\delta \\ \theta_1\,\theta_2\,f(y_2)f(y_1)[S(y_1)]^{\theta_1+\theta_2-\theta_2'-1}[S(y_2)]^{\theta_2-1}[S(y_1+\delta)]^{\theta_2'-\theta_2}, y_2 > y_1+\delta \end{cases} \quad (5.11)$$

5.2.1 Alternative Formulation

Alternatively, the model given in (5.11) can also be derived in conditional probability approach. Now let us look at the joint probability density function of Y_1 and Y_2 for $y_2 < y_1 < y_2 + \delta$,

$$f(y_1, y_2) = \lim_{dy_1, dy_2 -> 0} \frac{P[y_1 \leq Y_1 \leq y_1 + dy_1, y_2 \leq Y_2 + dy_2]}{dy_1 dy_2}$$

$$\begin{aligned}
&P[y_1 \leq T_1^* \leq y_1 + dy_1, y_2 \leq T_2 \leq y_2 + dy_2] \\
&= P[y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2, T_1 > T_2, T_2 < T_1^* < T_2 + \delta, y_1 \leq T_1^* \leq y_1 + dy_1] \\
&= P[y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2] \times P[T_1 > T_2 | y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2] \\
&\times P[T_2 < T_1^* < T_2 + \delta | y_2 \leq \text{Min}(T_1, T_2 \leq y_2 + dy_2, T_1 > T_2)] \\
&\times P[y_1 < T_1^* \leq y_1 + dy_1 | y_2 \leq \text{Min}(T_1, T_2 \leq y_2 + dy_2, T_1 > T_2, T_2 < T_1^* < T_2 + \delta)].
\end{aligned}$$

Now

$$\lim_{dy_2 -> 0} \frac{P[y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2]}{dy_2} = -\frac{d}{dy_2}[S(y_2)]^{\theta_1 + \theta_2},$$

$$P[T_1 > T_2 | y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2] = \frac{\theta_2}{\theta_1 + \theta_2},$$

$$\begin{aligned}
&P[T_2 < T_1^* < T_2 + \delta | y_2 \leq \text{Min}(T_1, T_2 \leq y_2 + dy_2, T_1 > T_2)] \\
&\qquad = \frac{P[y_2 < T_1^* < y_2 + \delta]}{P[T_1^* > y_2]} = \frac{S(y_2)]^{\theta_1'} - [S(y_2 + \delta)]^{\theta_1'}}{[S(y_2)]^{\theta_1'}}.
\end{aligned}$$

Since the event $[y_2 \leq \text{Min}(T_1, T_2 \leq y_2 + dy_2, T_1 > T_2)]$ is equivalent to $[T_1^* > y_2]$ and finally

$$P[y_1 < T_1^* < y_1 + dy_1 \,|\, y_2 \leq \text{Min}(T_1, T_2 \leq y_2 + dy_2, T_1 > T_2, T_2 < T_1^* < T_2 + \delta)]$$

$$= \frac{P[y_1 < T_1^* < y_1 + dy_1]}{P[y_2 < T_1^* < y_2 + \delta]} = \frac{-\frac{d[S(y_1)]^{\theta_1'}}{dy_1}}{[S(y_2)]^{\theta_1'} - [S(y_2 + \delta)]^{\theta_1'}},$$

so that $y_2 < y_1 < y_2 + \delta$

$$f(y_1, y_2) = \frac{\theta_2}{\theta_1 + \theta_2} \left[\frac{d}{dy_2}[S(y_2)]^{\theta_1 + \theta_2} \frac{[S(y_2)]^{\theta_1'} - [S(y_2 + \delta)]^{\theta_1'}}{[S(y_2)]^{\theta_1'}} \frac{\frac{d[S(y_1)]^{\theta_1'}}{dy_1}}{[S(y_2)]^{\theta_1'} - [S(y_2 + \delta)]^{\theta_1'}} \right]$$

$$f(y_1, y_2) = \theta_2 \, \theta_1' \, f(y_2) f(y_1) [S(y_2)]^{\theta_1 + \theta_2 - \theta_1' - 1} [S(y_1)]^{\theta_1' - 1}.$$

This is the same expression we obtained in the earlier set up with Freund dependency for $y_2 < y_1$,

$$f(y_1, y_2) = \lim_{dy_1, dy_2 -> 0} \frac{P[y_1 \leq Y_1 \leq y_1 + dy_1, y_2 \leq Y_2 + dy_2]}{dy_1 dy_2}$$

Now for $y_1 > y_2 + \delta$ it follows that

$$\begin{aligned}
&P[y_1 \leq Y_1 \leq y_1 + dy_1, y_2 \leq Y_2 + dy_2] \\
&= P[y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2, T_1 > T_2, T_1^* > T_2 + \delta, y_1 \leq T_1 \leq y_1 + \delta] \\
&= P[y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2] \times P[T_1 > T_2 | y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2] \\
&\times P[T_1^* > T_2 + \delta | y_2 \leq \text{Min}(T_1, T_2 \leq y_2 + dy_2, T_1 > T_2)] \\
&\times P[y_1 < T_1 \leq y_1 + dy_1 | y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2, T_1 > T_2, T_1^* > T_2 + \delta)].
\end{aligned}$$

Now

$$\lim_{dy_2 -> 0} \frac{P[y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2]}{dy_2} = -\frac{d}{dy_2} [S(y_2)\,]^{\theta_1 + \theta_2}$$

$$P[T_1 > T_2 \,|\, y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2] = \frac{\theta_2}{\theta_1 + \theta_2}$$

$$P[T_2 < T_1^* < T_2 + \delta | \, y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2, T_1 > T_2]$$

Since the event $[y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2, T_1 > T_2]$ is equivalent to $[T_1^* > y_2]$

$$= \frac{P[T_1^* > y_2 + \delta]}{P[T_1^* > y_2]} = \frac{S[(y_2 + \delta)]^{\theta_1'}}{S[(y_2)]^{\theta_1'}}.$$

Finally, consider

$$P[y_1 \leq T_1 \leq y_1 + dy_1 | \, y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2, T_1 > T_2, T_1^* > T_2 + \delta].$$

The event $(y_2 \leq \text{Min}(T_1, T_2) \leq y_2 + dy_2, T_1 > T_2, T_1^* > T_2 + \delta)$ is equivalent to $T_1 > y_2 + \delta$,

$$\lim_{dy_1 -> 0} \frac{P[y_1 \leq T_1 \leq y_1 + dy_1 | \, T_1 > y_2 + \delta]}{dy_1} = \frac{-\frac{d[S(y_1)]^{\theta_1}}{dy_1}}{[S(y_2 + \delta)]^{\theta_1}}$$

so that we have

$$f(y_1, y_2) = \frac{d}{dy_2}[S(y_2)\,]^{\theta_1 + \theta_2} \times \frac{\theta_2}{\theta_1 + \theta_2} \times \frac{[S(y_2 + \delta)]^{\theta_1'}}{[S(y_2)]^{\theta_1'}} \times \frac{\frac{d[S(y_2)]^{\theta_1}}{dy_1}}{[S(y_2 + \delta)]^{\theta_1}}, \; y_1 > y_2 + \delta$$

or

$$f(y_1, y_2) = \theta_2\, \theta_1\, f(y_2) f(y_1) [S(y_2)]^{\theta_1 + \theta_2 - \theta_1' - 1} [S(y_1)]^{\theta_1 - 1} [S(y_2 + \delta)]^{\theta_1' - \theta_1}, \; y_1 > y_2 + \delta \quad \blacksquare$$

Thus

$$f(y_1, y_2) = \begin{cases} \theta_2\, \theta_1'\, f(y_2) f(y_1) [S(y_2)]^{\theta_1 + \theta_2 - \theta_1' - 1} [S(y_1)]^{\theta_1' - 1}, y_2 < y_1 < y_2 + \delta \\ \theta_2\, \theta_1\, f(y_2) f(y_1) [S(y_2)]^{\theta_1 + \theta_2 - \theta_1' - 1} [S(y_1)]^{\theta_1 - 1} [S(y_2 + \delta)]^{\theta_1' - \theta_1}, y_1 > y_2 + \delta \end{cases}$$

Similarly for $y_1 < y_2$,

$$f(y_1, y_2) = \begin{cases} \theta_1\, \theta_2'\, f(y_2) f(y_1) [S(y_1)]^{\theta_1 + \theta_2 - \theta_2' - 1} [S(y_2)]^{\theta_2' - 1}, \; y_1 < y_2 < y_1 + \delta \\ \theta_1\, \theta_2\, f(y_2) f(y_1) [S(y_1)]^{\theta_1 + \theta_2 - \theta_2' - 1} [S(y_2)]^{\theta_2 - 1} [S(y_1 + \delta)]^{\theta_2' - \theta_2}, \; y_2 > y_1 + \delta \end{cases}$$

Thus we have

$$f(y_1, y_2) = \begin{cases} \theta_2\, \theta_1'\, f(y_2) f(y_1) [S(y_2)]^{\theta_1 + \theta_2 - \theta_1' - 1} [S(y_1)]^{\theta_1' - 1}, \; y_2 < y_1 < y_2 + \delta \\ \theta_2\, \theta_1\, f(y_2) f(y_1) [S(y_2)]^{\theta_1 + \theta_2 - \theta_1' - 1} [S(y_1)]^{\theta_1 - 1} [S(y_2 + \delta)]^{\theta_1' - \theta_1}, \; y_1 > y_2 + \delta \\ \theta_1\, \theta_2'\, f(y_2) f(y_1) [S(y_1)]^{\theta_1 + \theta_2 - \theta_2' - 1} [S(y_2)]^{\theta_2' - 1}, \; y_1 < y_2 < y_1 + \delta \\ \theta_1\, \theta_2\, f(y_2) f(y_1) [S(y_1)]^{\theta_1 + \theta_2 - \theta_2' - 1} [S(y_2)]^{\theta_2 - 1} [S(y_1 + \delta)]^{\theta_2' - \theta_2}, \; y_2 > y_1 + \delta \end{cases}$$

Equivalently

$$\begin{aligned} f(y_1,y_2) &= \frac{\theta_2}{\theta_1+\theta_2-\theta_1'}\frac{d[S(y_1)]^{\theta_1'}}{dy_1}\frac{d[S(y_2)]^{\theta_1+\theta_2-\theta_1'}}{dy_2}, \;\; y_2<y_1<y_2+\delta \\ &= \frac{\theta_2}{\theta_1+\theta_2-\theta_1'}\frac{d[S(y_1)]^{\theta_1}}{dy_1}\frac{d[S(y_2)]^{\theta_1+\theta_2-\theta_1'}}{dy_2}[S(y_2+\delta)]^{\theta_1'-\theta_1}, \;\; y_1>y_2+\delta \\ &= \frac{\theta_1}{\theta_1+\theta_2-\theta_2'}\frac{d[S(y_1)]^{\theta_1+\theta_2-\theta_2'}}{dy_1}\frac{d[S(y_2)]^{\theta_2'}}{dy_2}, \;\; y_1<y_2<y_1+\delta \\ &= \frac{\theta_1}{\theta_1+\theta_2-\theta_2'}\frac{d[S(y_1)]^{\theta_1+\theta_2-\theta_2'}}{dy_1}\frac{d[S(y_2)]^{\theta_2}}{dy_2}[S(y_1+\delta)]^{\theta_2'-\theta_2}, \;\; y_2>y_1+\delta. \end{aligned}$$

5.2.2 Properties

Property 5.2.1. *The probability density function specified in* (5.11) *reduces to a model with independent marginals whenever either of the following holds.*

1. $\theta_i=\theta_i', i=1,2$
2. $\delta=0$

Proof. This can be seen directly from (5.11) on substituting $\theta_i=\theta_i'$, $i=1,2$, or $\delta=0$. In both the cases

$$f(y_1,y_2)=\frac{d[S(y_1)]^{\theta_1}}{dy_1}\times\frac{d[S(y_2)]^{\theta_2}}{dy_2}$$

■

Property 5.2.2. *For* $\delta=\infty$ *the probability density function given in* (5.11) *reduces to the model specified by (Asha et al. (2016))*

$$f(y_1,y_2)=\begin{cases}\theta_2\,\theta_1'\,f(y_1)f(y_2)[S(y_2)]^{\theta_1+\theta_2-\theta_1'-1}[S(y_1)]^{\theta_1'-1}, y_1>y_2>0\\ \theta_1\,\theta_2'\,f(y_1)f(y_2)[S(y_1)]^{\theta_1+\theta_2-\theta_2'-1}[S(y_2)]^{\theta_2'-1}, y_2>y_1>0\end{cases} \tag{5.12}$$

Remark 5.2.1. *When* $S(y)=e^{-y}$, *the model* (5.12) *reduces to the Freund bivariate exponential model (Freund (1961)). In this case the joint probability density function* *of* (5.12) *becomes*

$$f(y_1, y_2) = \begin{cases} \theta'_1\theta_2 e^{-\theta'_1 y_1} e^{-(\theta_1+\theta_2-\theta'_1)y_2}, \ y_1 > y_2 > 0 \\ \theta_1\theta'_2 e^{-(\theta_1+\theta_2-\theta'_2)y_1} e^{-\theta'_2 y_2}, \ y_2 > y_1 > 0 \end{cases} \tag{5.13}$$

Remark 5.2.2. *When* $S(y) = e^{-y^\alpha}$, $\alpha > 0$, *the model* (5.12) *reduces to Weibull extensions of the Freund model (Lu (1989)). In this case the joint probability density function of* (5.12) *becomes*

$$f(y_1, y_2) = \begin{cases} \theta'_1\theta_2\alpha^2 y_1^{\alpha-1} y_2^{\alpha-1} e^{-\theta'_1 y_1^\alpha} e^{-(\theta_1+\theta_2-\theta'_1)y_2^\alpha}, y_1 > y_2 > 0 \\ \theta_1\theta'_2\alpha^2 y_1^{\alpha-1} y_2^{\alpha-1} e^{-(\theta_1+\theta_2-\theta'_2)y_1^\alpha} e^{-\theta'_2 y_2^\alpha}, y_2 > y_1 > 0 \end{cases} \tag{5.14}$$

The other examples are provided in Tables 5.1 - 5.2 .

Table 5.1: Examples for $f(y_1, y_2)$ in (5.11) for different baseline survival functions

S.No	$S(y)$	$f(y_1, y_2)$
1	e^{-y}	$\begin{cases} \theta_2\,\theta_1'\,e^{-(\theta_1+\theta_2-\theta_1')y_2-\theta_1' y_1}, & y_2 < y_1 < y_2+\delta \\ \theta_2\,\theta_1\,e^{-\theta_2 y_2}e^{-\theta_1 y_1}e^{-(\theta_1'-\theta_1)\delta}, & y_1 > y_2+\delta \\ \theta_1\,\theta_2'\,e^{-(\theta_1+\theta_2-\theta_2')y_1-\theta_2' y_2}, & y_1 < y_2 < y_1+\delta \\ \theta_1\,\theta_2\,e^{-\theta_1 y_1}e^{-\theta_2 y_2}e^{-(\theta_2'-\theta_2)\delta}, & y_2 > y_1+\delta \end{cases}$
2	$e^{-y^{\alpha}}$	$\begin{cases} \theta_1'\theta_2\alpha^2 y_1^{\alpha-1}y_2^{\alpha-1}e^{-\theta_1' y_1^{\alpha}}e^{-(\theta_1+\theta_2-\theta_1')y_2^{\alpha}}; y_2 < y_1 < y_2+\delta \\ \theta_1\theta_2\alpha^2 y_1^{\alpha-1}y_2^{\alpha-1}e^{-\theta_1 y_1^{\alpha}}e^{-(\theta_1+\theta_2-\theta_1')y_2^{\alpha}}e^{-(\theta_1'-\theta_1)(y_2+\delta)^{\alpha}}; y_1 > y_2+\delta \\ \theta_1\theta_2'\alpha^2 y_1^{\alpha-1}y_2^{\alpha-1}e^{-\theta_2' y_2^{\alpha}}e^{-(\theta_1+\theta_2-\theta_2')y_1^{\alpha}}; y_1 < y_2 < y_1+\delta \\ \theta_1\theta_2\alpha^2 y_1^{\alpha-1}y_2^{\alpha-1}e^{-\theta_2 y_2^{\alpha}}e^{-(\theta_1+\theta_2-\theta_2')y_1^{\alpha}}e^{-(\theta_2'-\theta_2)(y_1+\delta)^{\alpha}}; y_2 > y_1+\delta \end{cases}$
3	$(1+\alpha y^{\beta})^{-1}$	$\begin{cases} \theta_1'\theta_2\alpha^2\beta^2 y_1^{\beta-1}y_2^{\beta-1}(1+\alpha y_1^{\beta})^{-(\theta_1'+1)}(1+\alpha y_2^{\beta})^{-(\theta_1+\theta_2-\theta_1'+1)}; y_2 < y_1 < y_2+\delta \\ \theta_1\theta_2\alpha^2\beta^2 y_1^{\beta-1}y_2^{\beta-1}(1+\alpha y_1^{\beta})^{-(\theta_1+1)}(1+\alpha y_2^{\beta})^{-(\theta_1+\theta_2-\theta_1'+1)}(1+\alpha(y_2+\delta)^{\beta})^{(\theta_1-\theta_1')}; y_1 > y_2+\delta \\ \theta_1\theta_2'\alpha^2\beta^2 y_1^{\beta-1}y_2^{\beta-1}(1+\alpha y_2^{\beta})^{-(\theta_2'+1)}(1+\alpha y_1^{\beta})^{-(\theta_1+\theta_2-\theta_2'+1)}; y_1 < y_2 < y_1+\delta \\ \theta_1\theta_2\alpha^2\beta^2 y_1^{\beta-1}y_2^{\beta-1}(1+\alpha y_2^{\beta})^{-(\theta_2+1)}(1+\alpha y_1^{\beta})^{-(\theta_1+\theta_2-\theta_2'+1)}(1+\alpha(y_1+\delta)^{\beta})^{(\theta_2-\theta_2')}; y_2 > y_1+\delta \end{cases}$

Table 5.2: Examples for $f(y_1, y_2)$ in (5.11) for different baseline survival functions

S.No	$S(y)$	$f(y_1, y_2)$
4	$e^{-\frac{a}{b}(e^{by}-1)}$	$\begin{cases} a^2\theta_1'\theta_2 e^{b(y_1+y_2)} e^{\frac{a}{b}(1-e^{by_1})\theta_1'} e^{\frac{a}{b}(1-e^{by_2})(\theta_1+\theta_2-\theta_1')}; y_2 < y_1 < y_2+\delta \\ a^2\theta_1\theta_2 e^{b(y_1+y_2)} e^{\frac{a}{b}(1-e^{by_1})\theta_1} e^{\frac{a}{b}(1-e^{by_2})(\theta_1+\theta_2-\theta_1')} e^{-\frac{a}{b}(e^{b(y_2+\delta)}-1)(\theta_1'-\theta_1)}; y_1 > y_2+\delta \\ a^2\theta_1\theta_2' e^{b(y_1+y_2)} e^{\frac{a}{b}(1-e^{by_2})\theta_2'} e^{\frac{a}{b}(1-e^{by_1})(\theta_1+\theta_2-\theta_2')}; y_1 < y_2 < y_1+\delta \\ a^2\theta_1\theta_2 e^{b(y_1+y_2)} e^{\frac{a}{b}(1-e^{by_2})\theta_2} e^{\frac{a}{b}(1-e^{by_1})(\theta_1+\theta_2-\theta_2')} e^{-\frac{a}{b}(e^{b(y_1+\delta)}-1)(\theta_2'-\theta_2)}; y_2 > y_1+\delta \end{cases}$
5	$\left(1+\frac{y-\mu}{\sigma}\right)^{-\alpha}$	$\begin{cases} \theta_1'\theta_2 \left(\frac{\alpha}{\sigma}\right)^2 \left(\frac{y_1}{\sigma}\right)^{-(\alpha\theta_1'+1)} \left(\frac{y_2}{\sigma}\right)^{-(\alpha(\theta_1+\theta_2-\theta_1')+1)}; y_2 < y_1 < y_2+\delta \\ \theta_1\theta_2 \left(\frac{\alpha}{\sigma}\right)^2 \left(\frac{y_1}{\sigma}\right)^{-(\alpha\theta_1+1)} \left(\frac{y_2}{\sigma}\right)^{-(\alpha(\theta_1+\theta_2-\theta_1')+1)} \left(\frac{y_2+\delta}{\sigma}\right)^{\alpha(\theta_1-\theta_1')}; y_1 > y_2+\delta \\ \theta_1\theta_2' \left(\frac{\alpha}{\sigma}\right)^2 \left(\frac{y_1}{\sigma}\right)^{-(\alpha(\theta_1+\theta_2-\theta_2')+1)} \left(\frac{y_2}{\sigma}\right)^{-(\alpha\theta_2'+1)}; y_1 < y_2 < y_1+\delta \\ \theta_1\theta_2 \left(\frac{\alpha}{\sigma}\right)^2 \left(\frac{y_1}{\sigma}\right)^{-(\alpha(\theta_1+\theta_2-\theta_2')+1)} \left(\frac{y_2}{\sigma}\right)^{-(\alpha\theta_2+1)} \left(\frac{y_1+\delta}{\sigma}\right)^{\alpha(\theta_2-\theta_2')}; y_2 > y_1+\delta \end{cases}$
6	$e^{-(\lambda y+\frac{\eta}{2}y^2)}$	$\begin{cases} \theta_1'\theta_2(\lambda+\eta y_1)(\lambda+\eta y_2) e^{-\theta_1'(\frac{\eta y_1^2}{2}+\lambda y_1)} e^{-(\theta_1+\theta_2-\theta_1')(\frac{\eta y_2^2}{2}+\lambda y_2)}; y_2 < y_1 < y_2+\delta \\ \theta_1\theta_2(\lambda+\eta y_1)(\lambda+\eta y_2) e^{-\theta_1(\frac{\eta y_1^2}{2}+\lambda y_1)} e^{-(\theta_1+\theta_2-\theta_1')(\frac{\eta y_2^2}{2}+\lambda y_2)} e^{-\frac{1}{2}(y_2+\delta)(2\lambda+\eta(y_2+\delta)(\theta_1'-\theta_1)}; y_1 > y_2+\delta \\ \theta_1\theta_2'(\lambda+\eta y_1)(\lambda+\eta y_2) e^{-\theta_2'(\frac{\eta y_2^2}{2}+\lambda y_2)} e^{-(\theta_1+\theta_2-\theta_2')(\frac{\eta y_1^2}{2}+\lambda y_1)}; y_1 < y_2 < y_1+\delta \\ \theta_1\theta_2(\lambda+\eta y_1)(\lambda+\eta y_2) e^{-\theta_2(\frac{\eta y_2^2}{2}+\lambda y_2)} e^{-(\theta_1+\theta_2-\theta_2')(\frac{\eta y_1^2}{2}+\lambda y_1)} e^{-\frac{1}{2}(y_1+\delta)(2\lambda+\eta(y_1+\delta)(\theta_2'-\theta_2)}; y_2 > y_1+\delta \end{cases}$

Property 5.2.3. *For the probability density function defined in* (5.11) $P(X_i > X_{3-i}) = \frac{\theta_i}{\theta_1+\theta_2}, i = 1, 2.$

Proof. For $y_1 > y_2$

$$\int_0^\infty \left[\int_{y_2}^\infty f(y_1, y_2)dy_1\right] dy_2 = \int_0^\infty \int_{y_2}^{y_2+\delta} \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \frac{d[S(y_1)]^{\theta_1'}}{dy_1} \frac{d[S(y_2)]^{\theta_1 + \theta_2 - \theta_1'}}{dy_2} dy_1 dy_2$$
$$+ \int_0^\infty \int_{y_2+\delta}^\infty \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \frac{d[S(y_1)]^{\theta_1}}{dy_1} \frac{d[S(y_2)]^{\theta_1 + \theta_2 - \theta_1'}}{dy_2} [S(y_2 + \delta)]^{\theta_1' - \theta_1} dy_1 dy_2. \tag{5.15}$$

Now

$$\int_0^\infty \int_{y_2}^{y_2+\delta} \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \frac{d[S(y_1)]^{\theta_1'}}{dy_1} \frac{d[S(y_2)]^{\theta_1 + \theta_2 - \theta_1'}}{dy_2} dy_1 dy_2$$
$$= \int_0^\infty \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \frac{d[S(y_2)]^{\theta_1 + \theta_2 - \theta_1'}}{dy_2} [[S(y_2 + \delta)]^{\theta_1'} - [S(y_2)]^{\theta_1'}] dy_2$$
$$= \frac{\theta_2}{\theta_1 + \theta_2} + \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \int_0^\infty [S(y_2 + \delta)]^{\theta_1'} \frac{d[S(y_2)]^{\theta_1 + \theta_2 - \theta_1'}}{dy_2} dy_2. \tag{5.16}$$

Also

$$\int_0^\infty \int_{y_2+\delta}^\infty \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \frac{d[S(y_1)]^{\theta_1}}{dy_1} \frac{d[S(y_2)]^{\theta_1 + \theta_2 - \theta_1'}}{dy_2} [S(y_2 + \delta)]^{\theta_1' - \theta_1} dy_1 dy_2$$
$$= -\int_0^\infty \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \frac{d[S(y_2)]^{\theta_1 + \theta_2 - \theta_1'}}{dy_2} [S(y_2 + \delta)]^{\theta_1' - \theta_1} [S(y_2 + \delta)]^{\theta_1} dy_2$$
$$= -\int_0^\infty \frac{\theta_2}{\theta_1 + \theta_2 - \theta_1'} \frac{d[S(y_2)]^{\theta_1 + \theta_2 - \theta_1'}}{dy_2} [S(y_2 + \delta)]^{\theta_1'} dy_2. \tag{5.17}$$

From (5.16) and (5.17) it follows that

$$\int_0^\infty \left[\int_{y_2}^\infty f(y_1, y_2) dy_1\right] dy_2 = \frac{\theta_2}{\theta_1 + \theta_2}. \tag{5.18}$$

Similarly we can prove that

$$\int_0^\infty \left[\int_{y_1}^\infty f(y_1, y_2) dy_2\right] dy_1 = \frac{\theta_1}{\theta_1 + \theta_2}. \tag{5.19}$$

■

5.3 The Exponential Cascade Model

When $S(y) = e^{-y}$ in the model (5.11), it reduces to the exponential cascade model, which is the main focus in the remaining of this chapter. The model is

$$f(y_1, y_2) = \begin{cases} \theta_2\, \theta_1'\, e^{-(\theta_1 + \theta_2 - \theta_1')y_2 - \theta_1'\, y_1}, & y_2 < y_1 < y_2 + \delta \\ \theta_2\, \theta_1\, e^{-\theta_2\, y_2} e^{-\theta_1\, y_1} e^{-(\theta_1' - \theta_1)\delta}, & y_1 > y_2 + \delta \\ \theta_1\, \theta_2'\, e^{-(\theta_1 + \theta_2 - \theta_2')y_1 - \theta_2'\, y_2}, & y_1 < y_2 < y_1 + \delta \\ \theta_1\, \theta_2\, e^{-\theta_1\, y_1} e^{-\theta_2\, y_2} e^{-(\theta_2' - \theta_2)\delta}, & y_2 > y_1 + \delta \end{cases} \tag{5.20}$$

Remark 5.3.1. *When $S(y) = e^{-y}$, $\theta_i = \theta$ and $\theta_i' = 2\theta$, for $i = 1, 2$, in the model* (5.11)*, then it reduces to the model specified by (Lindley & Singpurwalla (2002)).*

$$f(y_1, y_2) = \begin{cases} 2\theta^2 e^{-2\theta y_1}, y_2 < y_1 < y_2 + \delta \\ \theta^2 e^{-\theta(y_1 + y_2 + \delta)}, y_1 > y_2 + \delta \\ 2\theta^2 e^{-2\theta y_2}, y_1 < y_2 < y_1 + \delta \\ \theta^2 e^{-\theta(y_1 + y_2 + \delta)}, y_2 > y_1 + \delta \end{cases} \tag{5.21}$$

Remark 5.3.2. *When $S(y) = e^{-y}$, $\theta_i = \theta$ and $\theta_i' = c\theta$, for $i = 1, 2$, in the model*

(5.11), *then it reduces to the model specified by (Swift (2008)).*

$$f(y_1, y_2) = \begin{cases} c\theta^2 e^{-c\theta y_1}, y_2 < y_1 < y_2 + \delta \\ \theta^2 e^{-\theta(y_1+y_2+\delta)}, y_1 > y_2 + \delta \\ c\theta^2 e^{-c\theta y_2}, y_1 < y_2 < y_1 + \delta \\ \theta^2 e^{-\theta(y_1+y_2+\delta)}, y_2 > y_1 + \delta \end{cases} \tag{5.22}$$

Theorem 5.3.1. *The probability density function of the time to failure of the system that is* $W = \mathrm{Max}\,[Y_1, Y_2]$ *with lifetime model specified by* (5.20) *is given by*

$$\begin{aligned} f_C(w) &= \frac{\theta_2\,\theta_1'}{\theta_1 + \theta_2 - \theta_1'}\left[e^{-\theta_1' w} - e^{-(\theta_1+\theta_2)w}\right] + \frac{\theta_1\,\theta_2'}{\theta_1 + \theta_2 - \theta_2'}\left[e^{-\theta_2' w} - e^{-(\theta_1+\theta_2)w}\right], w < \delta \\ &= \theta_1\, e^{-\delta(\theta_1' - \theta_1)} e^{-\theta_1 w}\left[1 - e^{-\theta_2(w-\delta)}\right] + \frac{\theta_2\,\theta_1'\, e^{-(\theta_1+\theta_2)w}}{\theta_1 + \theta_2 - \theta_1'}\left[e^{-\delta(\theta_1' - \theta_1 - \theta_2)} - 1\right] \\ &+ \theta_2\, e^{-(\theta_2' - \theta_2)\delta} e^{-\theta_2 w}\left[1 - e^{-\theta_1(w-\delta)}\right] + \frac{\theta_1\,\theta_2'\, e^{-(\theta_1+\theta_2)w}}{\theta_1 + \theta_2 - \theta_2'}\left[e^{-\delta(\theta_2' - \theta_1 - \theta_2)} - 1\right], w \geq \delta. \end{aligned} \tag{5.23}$$

Proof. First let us consider for $w < \delta$

$$f_{C_1}(w) = \int_0^w f(y_1, y_2)dy_2 + \int_0^w f(y_1, y_2)dy_1 = \int_0^w f(w, y_2)dy_2 + \int_0^w f(y_1, w)dy_1$$

$$f_{C_1}(w) = \frac{\theta_2\,\theta_1'}{\theta_1 + \theta_2 - \theta_1'}\left[e^{-\theta_1' w} - e^{-(\theta_1+\theta_2)w}\right] + \frac{\theta_1\,\theta_2'}{\theta_1 + \theta_2 - \theta_2'}\left[e^{-\theta_2' w} - e^{-(\theta_1+\theta_2)w}\right]. \tag{5.24}$$

Now let us consider for $w > \delta$

$$\begin{aligned} f_{C_2}(w) = &\int_0^{w-\delta} [\theta_1\,\theta_2\, e^{-\theta_1 y_1} e^{-\theta_2 y_2} e^{-(\theta_2' - \theta_2)\delta}]dy_1 + \int_{w-\delta}^{w} [\theta_1\,\theta_2'\, e^{-(\theta_1+\theta_2-\theta_2')y_1 - \theta_2' y_2}]dy_1 \\ &+ \int_0^{w-\delta} [\theta_2\,\theta_1\, e^{-\theta_2 y_2} e^{-\theta_1 y_1} e^{-(\theta_1' - \theta_1)\delta}]dy_2 + \int_{w-\delta}^{w} \theta_2\,\theta_1'\, e^{-(\theta_1+\theta_2-\theta_1')y_2 - \theta_1' y_1}]dy_2 \end{aligned} \tag{5.25}$$

$$f_{C_2}(w) = \theta_1\, e^{-(\theta_1' - \theta_1)\delta} e^{-\theta_1 w}\left[1 - e^{-\theta_2(w-\delta)}\right] + \frac{\theta_2\, \theta_1'\, e^{-(\theta_1+\theta_2)w}}{\theta_1 + \theta_2 - \theta_1'}\left[e^{(\theta_1+\theta_2-\theta_1')\delta} - 1\right]$$
$$+ \frac{\theta_1\, \theta_2'\, e^{-(\theta_1+\theta_2)w}}{\theta_1 + \theta_2 - \theta_2'}\left[e^{(\theta_1+\theta_2-\theta_2')\delta} - 1\right] + \theta_2\, e^{-(\theta_2' - \theta_2)\delta} e^{-\theta_2 w}\left[1 - e^{-\theta_1(w-\delta)}\right]. \quad (5.26)$$

Therefore, the time to failure of the system with cascading failures at w is of the form

$$f_C(w) = \frac{\theta_2\, \theta_1'}{\theta_1 + \theta_2 - \theta_1'}\left[e^{-\theta_1' w} - e^{-(\theta_1+\theta_2)w}\right] + \frac{\theta_1\, \theta_2'}{\theta_1 + \theta_2 - \theta_2'}\left[e^{-\theta_2' w} - e^{-(\theta_1+\theta_2)w}\right], w < \delta$$
$$= \theta_1\, e^{-\delta(\theta_1' - \theta_1)} e^{-\theta_1 w}\left[1 - e^{-\theta_2(w-\delta)}\right] + \frac{\theta_2\, \theta_1'\, e^{-(\theta_1+\theta_2)w}}{\theta_1 + \theta_2 - \theta_1'}\left[e^{-\delta(\theta_1' - \theta_1 - \theta_2)} - 1\right]$$
$$+ \theta_2\, e^{-(\theta_2' - \theta_2)\delta} e^{-\theta_2 w}\left[1 - e^{-\theta_1(w-\delta)}\right] + \frac{\theta_1\, \theta_2'\, e^{-(\theta_1+\theta_2)w}}{\theta_1 + \theta_2 - \theta_2'}\left[e^{-\delta(\theta_2' - \theta_1 - \theta_2)} - 1\right], w \geq \delta. \quad (5.27)$$

■

Corollary 5.3.1. *For $\theta_i = \theta$ and $\theta_i' = 2\theta$ for $i = 1, 2$*

$$f_C(w) = 4\theta^2 w e^{-2\theta w}, w < \delta$$
$$= 4\theta^2 \delta e^{-2\theta w} + 2\theta e^{-\theta(w+\delta)} - 2\theta e^{-2\theta w}, w \geq \delta \quad (5.28)$$

(Lindley & Singpurwalla (2002)).

Theorem 5.3.2. *The survival or reliability function is specified by*

$$\begin{aligned} S_C(w) &= \frac{\theta_1(\theta_1+\theta_2-\theta_1')e^{-w\theta_2'} + \theta_2(\theta_1+\theta_2-\theta_2')e^{-w\theta_1'}}{(\theta_1+\theta_2-\theta_1')(\theta_1+\theta_2-\theta_2')} \\ &+ \frac{(\theta_2'(\theta_1'-\theta_1) - \theta_2\theta_1')e^{-w(\theta_1+\theta_2)}}{(\theta_1+\theta_2-\theta_1')(\theta_1+\theta_2-\theta_2')}, \quad w < \delta \\ &= e^{-w\theta_1+\delta(\theta_1-\theta_1')} + e^{-w\theta_2+\delta(\theta_2-\theta_2')} + \frac{(\theta_1'-\theta_1)e^{-w(\theta_1+\theta_2)+\delta(\theta_1+\theta_2-\theta_1')}}{\theta_1+\theta_2-\theta_1'} \\ &- \frac{(\theta_2\theta_1' + \theta_2'(\theta_1'-\theta_1))e^{-w(\theta_1+\theta_2)}}{(\theta_1+\theta_2-\theta_1')(\theta_1+\theta_2-\theta_2')} + \frac{(\theta_2'-\theta_2)e^{-w(\theta_1+\theta_2)+\delta(\theta_1+\theta_2-\theta_2')}}{\theta_1+\theta_2-\theta_2'}, \quad w \geq \delta. \end{aligned} \quad (5.29)$$

Proof. Though, the calculation is cumbersome, it can be directly derived from the result,

$$S_C(w) = \int_w^\delta f_{C_1}(w)dw + \int_\delta^\infty f_{C_2}(w)dw,\ w < \delta$$
$$= \int_w^\infty f_{C_2}(w)dw, \quad w \geq \delta. \tag{5.30}$$

Now

$$\begin{aligned}
S_{C_1}(w) = & \int_w^\delta \left(\frac{(-e^{-w(\theta_1+\theta_2)} + e^{-w\,\theta_1'})\theta_2\theta_1'}{\theta_1+\theta_2-\theta_1'} + \frac{(-e^{-w(\theta_1+\theta_2)} + e^{-w\theta_2'})\theta_1\theta_2'}{\theta_1+\theta_2-\theta_2'} \right) dw \\
& + \int_\delta^\infty \left(e^{-\delta(\theta_1'-\theta_1)} \left[e^{-w\theta_1} - e^{-w\theta_1-(w-\delta)\theta_2} \right] \theta_1 \right) dw \\
& + \int_\delta^\infty \left(e^{-\delta(\theta_2'-\theta_2)} \left[e^{-w\theta_2} - e^{-w\theta_2-(w-\delta)\theta_1} \right] \theta_2 \right) dw \\
& + \int_\delta^\infty \left(\frac{\left[e^{-w(\theta_1+\theta_2)+\delta(\theta_1+\theta_2-\theta_1')} - e^{-w(\theta_1+\theta_2)} \right] \theta_2\theta_1'}{\theta_1+\theta_2-\theta_1'} \right) dw \\
& + \int_\delta^\infty \left(\frac{\left[e^{-w(\theta_1+\theta_2)+\delta(\theta_1+\theta_2-\theta_2')} - e^{-w(\theta_1+\theta_2)} \right] \theta_1\theta_2'}{\theta_1+\theta_2-\theta_2'} \right) dw
\end{aligned} \tag{5.31}$$

and

$$\begin{aligned}
S_{C_2}(w) = & \int_w^\infty \left(e^{-\delta(\theta_1'-\theta_1)-w\theta_1} \left[1 - e^{-(w-\delta)\theta_2} \right] \theta_1 + e^{-\delta(\theta_2'-\theta_2)-w\theta_2} \left[1 - e^{-(w-\delta)\theta_1} \right] \theta_2 \right) dw \\
& + \int_w^\infty \left(\frac{\left[e{-}w(\theta_1+\theta_2) + \delta(\theta_1+\theta_2-\theta_1') - e^{-w(\theta_1+\theta_2)} \right] \theta_2\theta_1'}{\theta_1+\theta_2-\theta_1'} \right) dw \\
& + \int_w^\infty \left(\frac{\left[e^{-w(\theta_1+\theta_2)+\delta(\theta_1+\theta_2-\theta_2')} - e^{-w(\theta_1+\theta_2)} \right] \theta_1\theta_2'}{\theta_1+\theta_2-\theta_2'} \right) dw.
\end{aligned} \tag{5.32}$$

Therefore, the reliability function at w is of the form

$$S_C(w) \quad = \quad \frac{\theta_1(\theta_1+\theta_2-\theta_1')e^{-w\theta_2'} + \theta_2(\theta_1+\theta_2-\theta_2')e^{-w\theta_1'}}{(\theta_1+\theta_2-\theta_1')(\theta_1+\theta_2-\theta_2')}$$

$$
\begin{aligned}
&+ \quad \frac{(\theta_2'(\theta_1' - \theta_1) - \theta_2\theta_1')e^{-w(\theta_1+\theta_2)}}{(\theta_1 + \theta_2 - \theta_1')(\theta_1 + \theta_2 - \theta_2')}, \quad w < \delta \\
&= \quad e^{-w\theta_1+\delta(\theta_1-\theta_1')} + e^{-w\theta_2+\delta(\theta_2-\theta_2')} + \frac{(\theta_1' - \theta_1)e^{-w(\theta_1+\theta_2)+\delta(\theta_1+\theta_2-\theta_1')}}{\theta_1 + \theta_2 - \theta_1'} \\
&- \quad \frac{(\theta_2\theta_1' + \theta_2'(\theta_1' - \theta_1))e^{-w(\theta_1+\theta_2)}}{(\theta_1 + \theta_2 - \theta_1')(\theta_1 + \theta_2 - \theta_2')} + \frac{(\theta_2' - \theta_2)e^{-w(\theta_1+\theta_2)+\delta(\theta_1+\theta_2-\theta_2')}}{\theta_1 + \theta_2 - \theta_2'}, \quad w \geq \delta
\end{aligned}
\tag{5.33}
$$

The equation (5.29) is now retrieved from (5.31) and (5.32). ■

Property 5.3.1. *The failure rate of* (5.27) *is given by*

$$
h_C(w) = \begin{cases} \frac{A(w;\boldsymbol{\tau})+B(w;\boldsymbol{\tau})-C(w;\boldsymbol{\tau})}{D(w;\boldsymbol{\tau})}, w < \delta \\ \frac{E(w;\boldsymbol{\tau})-F(w;\boldsymbol{\tau})-G(w;\boldsymbol{\tau})-H(w;\boldsymbol{\tau})}{I(w;\boldsymbol{\tau})-J(w;\boldsymbol{\tau})}, w \geq \delta \end{cases}
\tag{5.34}
$$

where $\boldsymbol{\tau} = (\theta_1, \theta_2, \theta_1', \theta_2', \delta)$,

$$
\begin{aligned}
A(w;\boldsymbol{\tau}) &= e^{(\theta_1+\theta_2-\theta_2')w}[\theta_2\,\theta_1'\,(\theta_1+\theta_2-\theta_2')], \\
B(w;\boldsymbol{\tau}) &= e^{(\theta_1+\theta_2-\theta_1')w}[\theta_1\,\theta_2'(\theta_1+\theta_2-\theta_1')], \\
C(w;\boldsymbol{\tau}) &= [(\theta_1+\theta_2)(\theta_2\,\theta_1'+\theta_2'(\theta_1-\theta_1'))]e^{(\theta_1'+\theta_2')w}, \\
D(w;\boldsymbol{\tau}) &= \left[\begin{array}{l} (\theta_1+\theta_2-\theta_2')\,\theta_2\,e^{(\theta_1+\theta_2+\theta_2')w} - (\theta_2\,\theta_1'+\theta_2'(\theta_1-\theta_1'))e^{(\theta_1'+\theta_2')w} \\ +\,(\theta_1+\theta_2-\theta_1')\theta_1\,e^{(\theta_1+\theta_2-\theta_1')w} \end{array}\right], \\
E(w;\boldsymbol{\tau}) &= (\theta_1+\theta_2-\theta_2')(\theta_1+\theta_2-\theta_1')\left[\theta_1 e^{-w\,\theta_1-\delta(\theta_1'-\theta_1)} + \theta_2\,e^{-w\,\theta_2-\delta(\theta_2'-\theta_2)}\right], \\
F(w;\boldsymbol{\tau}) &= [(\theta_1+\theta_2-\theta_2')(\theta_1+\theta_2)(\theta_1-\theta_1')]\,e^{-(\theta_1+\theta_2)w-\delta(\theta_1'-\theta_1-\theta_2)}, \\
G(w;\boldsymbol{\tau}) &= [(\theta_1+\theta_2-\theta_1')\,(\theta_1+\theta_2)\,(\theta_2-\theta_2')]\,e^{-(\theta_1+\theta_2)w+(\theta_1+\theta_2-\theta_2')\delta}, \\
H(w;\boldsymbol{\tau}) &= [(\theta_1+\theta_2)(\theta_2\,\theta_1'\,\theta_2')(\theta_1-\theta_1')]\,e^{-(\theta_1+\theta_2)w}, \\
I(w;\boldsymbol{\tau}) &= (\theta_1+\theta_2-\theta_2')(\theta_1+\theta_2-\theta_1')\left[\theta_1 e^{-w\,\theta_1-\delta(\theta_1'-\theta_1)} + \theta_2\,e^{-w\,\theta_2-\delta(\theta_2'-\theta_2)}\right], \\
J(w;\boldsymbol{\tau}) &= (\theta_1-\theta_1')e^{-w(\theta_1+\theta_2)}\left[\begin{array}{l} [(\theta_1+\theta_2-\theta_2')e^{-\delta(\theta_1'-\theta_1-\theta_2)} - \theta_2\,\theta_1'\,\theta_2'] - \\ (\theta_1+\theta_2-\theta_1')(\theta_2'-\theta_2)e^{-\delta(\theta_2'-\theta_1-\theta_2)} \end{array}\right].
\end{aligned}
$$

The probability density function, $f_C(w)$ starts at zero when $w = 0$, attains its maximum as w moves close to 0.5, and then start decreasing towards zero as w approaches to ∞ (see Figure 5.2). The hazard function, $h_C(w)$ increases to a maximum and then remains constant as w approaches to ∞ (see Figure 5.4).

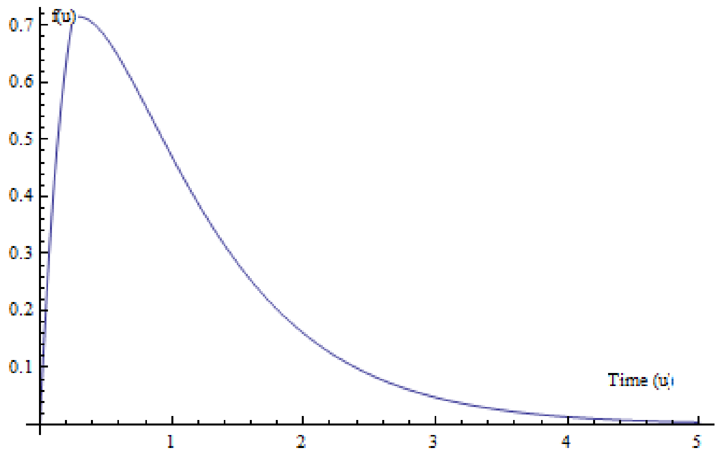


Figure 5.2: Failure time density function for the cascading model when $\theta_1 = 1.25, \theta_2 = 1.25, \theta_1' = 2.0, \theta_2' = 2.0$ *and* $\delta = 0.25$.

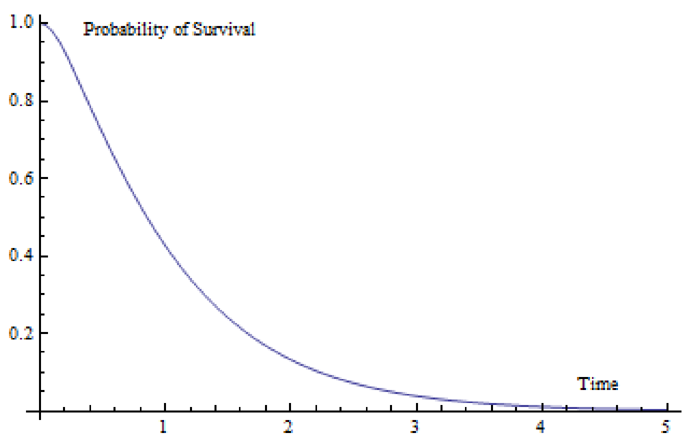


Figure 5.3: Survival function (when $\theta_1 = 1.25, \theta_2 = 1.25, \theta_1' = 2.0, \theta_2' = 2.0, \delta = 0.25$)

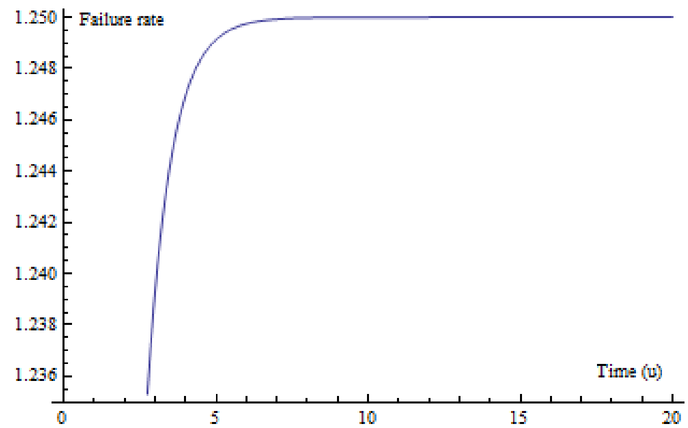


Figure 5.4: Hazard rate function (when $\theta_1 = 1.25, \theta_2 = 1.25, \theta_1' = 2.0, \theta_2' = 2.0, \delta = 0.25$)

The mean time to system failure (MTTF) is defined by $\int_0^\infty S_C(w)dw$. The resulted MTTF is given by

$$MTTF = \frac{\theta_1\,\theta_2(\theta_1\,\theta_1' + (\theta_2 + \theta_1')\,\theta_2') - \theta_1^2\,\theta_1'(\theta_2 - \theta_2')e^{-\delta\,\theta_2'} - \theta_2^2\,\theta_2'(\theta_1 - \theta_1')e^{-\delta\,\theta_1'}}{\theta_1\,\theta_2(\theta_1 + \theta_2)\,\theta_1'\,\theta_2'} \quad (5.35)$$

Remark 5.3.3. *For $\theta_i = \theta$ and $\theta_i' = 2\theta$, $i = 1, 2$ the mean time to system failure is $\frac{1}{\theta} + \frac{e^{-2\delta\theta}}{2\theta}$ as in Lindley & Singpurwalla (2002). When the critical time $\delta \uparrow \infty$, the mean time to system failure is $\frac{\theta_1\,\theta_1' + (\theta_2 + \theta_1')\,\theta_2'}{(\theta_1 + \theta_2)\,\theta_1'\,\theta_2'}$ and when $\delta \downarrow 0$, the mean time to system failure is $\frac{\theta_1^2 + \theta_2^2 + \theta_1\,\theta_2}{\theta_1\,\theta_2(\theta_1 + \theta_2)}$. Now, for $(\theta_i = \theta)$ and $(\theta_i' = 2\theta)$, for $i = 1, 2$ the mean time to system failure would simplify into $\frac{1}{\theta}$ and $\frac{3}{2\theta}$ respectively. (Lindley & Singpurwalla (2002)). Also, for $(\theta_i = \theta)$ and $(\theta_i' = c\theta)$, for $i = 1, 2$ and $\delta \downarrow 0$ the mean time to system failure would simplify into $\frac{3}{2\theta}$ as obtained in Swift (2008). Further, from* (5.35) *it can be inferred that the cascading failure results in a greater mean time to failure than the mean time to failure under Freund's model whenever failure of one component adversely affects the functioning of the surviving component in a two-unit parallel redundant system.*

Remark 5.3.4. *When $\theta_1 + \theta_2 = \theta_i'$ and $\theta_1 + \theta_2 \neq \theta_{3-i}'$, for $i = 1, 2$, then it reduces to the model with the survival or reliability function specified by*

$$S_C(w) = \frac{e^{-w(\theta_1+\theta_2+\theta_{3-i}')}\left(\theta_i e^{w(\theta_1+\theta_2)} + e^{w\theta_{3-i}'}(\theta_{3-i}(1 + w(\theta_1 + \theta_2 - \theta_{3-i}') - \theta_{3-i}')\right)}{(\theta_1 + \theta_2 - \theta_{3-i}')}, \; w < \delta$$

$$= e^{-w\theta_i-\delta\theta_{3-i}} + e^{-w\theta_{3-i}+\delta(\theta_{3-i}-\theta'_{3-i})} + e^{-w(\theta_1+\theta_2)}\left(\delta\theta_{3-i} - \frac{\theta_i}{\theta_1+\theta_2+\theta'_{3-i}}\right)$$
$$+ \frac{(\theta'_{3-i}-\theta_{3-i})e^{-w(\theta_1+\theta_2)+\delta(\theta_1+\theta_2+\theta'_{3-i})}}{\theta_1+\theta_2-\theta'_{3-i}}, \; w \geq \delta. \tag{5.36}$$

Remark 5.3.5. *When $\theta_1+\theta_2=\theta'_i$, for $i=1,2$, then it reduces to the model with the survival or reliability function specified by*

$$S_C(w) = w(\theta_1+\theta_2)e^{-w(\theta_1+\theta_2)} + e^{-w(\theta_1+\theta_2)}; w < \delta$$
$$= \delta(\theta_1+\theta_2)e^{-w(\theta_1+\theta_2)} + e^{-(\theta_1 w+\theta_2\delta)} + e^{-(\theta_1\delta+\theta_2 w)} - e^{-w(\theta_1+\theta_2)}; w \geq \delta. \tag{5.37}$$

5.4 Estimation of the parameters

In quest of explaining the parameter estimation of the failure time distribution function in (5.27), classical procedures such as method of maximum likelihood and method of least squares turn out to be cumbersome as the range of the random variables depend on parameter of interest. Hence, we consider the technique of (i) method of moments and (ii) method of L-moments. It is our primary interest to estimate the change in failure rates and the threshold time δ. The method of moments technique of parameter estimation is applied to the density in (5.27) to obtain the moment equations. The first three moment equations are given as follows:

$$m_1 = \frac{\theta_1(\theta'_2-\theta_2)e^{-\delta\,\theta'_2}}{\theta'_2\,\theta_2(\theta_1+\theta_2)} + \frac{\theta_2(\theta'_1-\theta_1)e^{-\delta\,\theta'_1}}{\theta'_1\,\theta_1(\theta_1+\theta_2)} + \frac{\theta_1\,\theta'_1+\theta'_2(\theta'_1+\theta_2)e^{-\delta\,\theta'_2}}{\theta'_1\,\theta'_2(\theta_1+\theta_2)}, \tag{5.38}$$

$$m_2 = \frac{2\theta_2(\theta'_1-\theta_1)\left[\theta_1(\theta_1+\theta_2)+\theta'_1(\theta_2+\theta_1(2+\delta(\theta_1+\theta_2))\right]e^{-\delta\,\theta'_1}}{\theta_1^2\theta_1'^2(\theta_1+\theta_2)^2}$$
$$+ \frac{2\theta_1(\theta'_2-\theta_2)\left[\theta_1(\theta_2+\theta'_2+\delta\,\theta_2\,\theta'_2)+\theta_2(\theta_2+2\,\theta'_2+\delta\,\theta_2\,\theta'_2)\right]e^{-\delta\,\theta'_2}}{\theta_2^2(\theta_1+\theta_2)^2}$$
$$+ \frac{2\left[\theta_1(\theta_1+\theta_2)\,\theta_1'^2+(\theta_2(\theta_1+\theta_2)+\theta_2\,\theta'_1+\theta_1'^2)\,\theta_2'^2+\theta_1\,\theta_1'^2\,\theta'_2)\right]}{\theta_1'^2\,\theta_2'^2(\theta_1+\theta_2)^2}, \tag{5.39}$$

$$m_3 = \int_0^\infty w^3 f_C(w) dw. \tag{5.40}$$

Alternatively, we could think of substituting the L-moments for the conventional moments. The foundation of this approach is found in Hosking (1990). Basically, L-moments are expectations of certain linear combinations of order statistics. They can be defined for any random variable whose mean exists. The main advantage of L-moments over conventional moments is that L-moments, being linear functions of the data, suffer less from the effects of sampling variability: L-moments are more robust than conventional moments to outliers in the data and enable more secure inferences to be made from small samples about an underlying probability distribution. L-moments sometimes yield more efficient parameter estimates than the maximum likelihood estimates. If $W_{1(n)} \leq W_{2(n)} \leq ... \leq W_{n(n)}$ are the rank statistics of the random sample of size n selected from the distribution W, then the r^{th} L-moment of the random variable W is defined as

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \begin{pmatrix} r-1 \\ k \end{pmatrix} E\, W_{(r-k)(r)}, r = 1, 2, 3, ... \ , \tag{5.41}$$

where $\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 1.$

The expected value of the rank statistic is of the form

$$EW_{j(r)} = \frac{r!}{(j-1)!(r-j)!} \int w[1 - S_C(w)]^{j-1} [S_C(w)]^{r-j} (-dS_C(w)) \tag{5.42}$$

(David (1968)). The L-moments for the distribution in (5.27) are obtained by using (5.42). Observe that the first L-moment is the $E(W)$. The first three L-moments are given as follows:

$$\lambda_1 = \frac{\theta_1(\theta_2' - \theta_2)e^{-\delta\,\theta_2'}}{\theta_2'\,\theta_2(\theta_1 + \theta_2)} + \frac{\theta_2(\theta_1' - \theta_1)e^{-\delta\,\theta_1'}}{\theta_1'\,\theta_1(\theta_1 + \theta_2)} + \frac{\theta_1\,\theta_1' + \theta_2'(\theta_1' + \theta_2)e^{-\delta\,\theta_2'}}{\theta_1'\,\theta_2'(\theta_1 + \theta_2)} \tag{5.43}$$

$$\lambda_2 \;=\; \frac{\theta_1(\theta_2' - \theta_2)e^{-\delta\,\theta_2'}}{\theta_2'\,\theta_2(\theta_1 + \theta_2)} + \frac{\theta_2(\theta_1' - \theta_1)e^{-\delta\,\theta_1'}}{\theta_1'\,\theta_1(\theta_1 + \theta_2)} + \frac{\theta_1^2(\theta_2 - \theta_2')\,(\theta_1 + 2\theta_2 - \theta_2')e^{-2\delta\,\theta_2'}}{2\theta_2'\,\theta_2(\theta_1 + \theta_2)\,(\theta_1 + 2\,\theta_2)\,(\theta_1 + \theta_2 - \theta_2')}$$

$$
\begin{aligned}
&+ \quad \frac{\theta_2^2(\theta_1-\theta_1')\,(2\theta_1+\theta_2-\theta_1')e^{-2\delta\,\theta_1'}}{2\theta_1'\,\theta_1(\theta_1+\theta_2)\,(2\theta_1+\theta_2)\,(\theta_1+\theta_2-\theta_1')} \\
&+ \quad \frac{\theta_1(\theta_2'-\theta_2)\,[\theta_2\,\theta_1'+\theta_2'(\theta_1-\theta_1')]e^{-\delta(\theta_1+\theta_2+\theta_2')}}{(\theta_1+\theta_2)(\theta_1+2\,\theta_2)\,(\theta_1+\theta_2-\theta_1')\,(\theta_1+\theta_2-\theta_2')\,(\theta_1+\theta_2+\theta_2')} \\
&+ \quad \frac{\theta_2(\theta_1'-\theta_1)\,[\theta_2\,\theta_1'+\theta_2'(\theta_1-\theta_1')]e^{-\delta(\theta_1+\theta_2+\theta_1')}}{(\theta_1+\theta_2)(\theta_1+2\,\theta_2)\,(\theta_1+\theta_2-\theta_1')\,(\theta_1+\theta_2-\theta_2')\,(\theta_1+\theta_2+\theta_2')} \\
&+ \quad \frac{\theta_1\,\theta_2\,\left[\theta_1(3\,\theta_1'^2+\theta_1{}'\theta_2'+2\,\theta_2'^2)+\theta_2(2\,\theta_1'^2+\theta_1{}'\theta_2'+3\,\theta_2'^2)\right]}{2(\theta_1+\theta_2)\,\theta_1'\,\theta_2'\,(\theta_1+\theta_2+\theta_1')\,(\theta_1+\theta_2+\theta_2')\,(\theta_1'+\theta_2')} \\
&+ \quad \frac{(\theta_1'+\theta_2')\left\{\theta_1^3\,\theta_1'+\theta_2'(\theta_2+\theta_2')(\theta_2^2+\theta_2\,\theta_1'+\theta_1'^2)+2\,\theta_2(\theta_1'^2+\theta_2'^2)\right\}+\theta_1'\,\theta_1\left[(\theta_1'+\theta_2')^2\right]\left\{\theta_1+\theta_2'\right\}}{2(\theta_1+\theta_2)\,\theta_1'\,\theta_2'\,(\theta_1+\theta_2+\theta_1')\,(\theta_1+\theta_2+\theta_2')\,(\theta_1'+\theta_2')} \\
&+ \quad \frac{\theta_1\,\theta_2\left\{4\theta_1^3+4\theta_2^3-(\theta_1'+\theta_2')\left[8\theta_2^2-3\,\theta_1'\,\theta_2'+\theta_2(5\theta_1'+4\,\theta_2')+8\theta_1^2+17\,\theta_1\,\theta_2-\theta_1(4\theta_1'+5\,\theta_2')\right]+14\theta_1^2\,\theta_2+14\theta_2^2\,\theta_1\right\}e^{-\delta(\theta_1'+\theta_2')}}{(\theta_1+\theta_2)(2\theta_1+\theta_2)(\theta_1+2\theta_2)\,(\theta_1+\theta_2-\theta_1')\,(\theta_1+\theta_2-\theta_2')\,(\theta_1'+\theta_2')},
\end{aligned}
\tag{5.44}
$$

$$
\lambda_3 = \frac{1}{3}E[W_{3(3)} - 2W_{2(3)} + W_{1(3)}]. \tag{5.45}
$$

The expression for the higher order are not presented here due to the complicated expressions. They are obtained with the help of Mathematica Software.

Now we will assume that $w_1, w_2, w_3, ..., w_n$ is a random sample and $w_{1:n} \leq w_{2:n} \leq ... \leq w_{n:n}$ is the ordered sample. The first three sample L-moment are obtained from the equation below by substituting $r = 1, 2, 3$ respectively.

$$
l_r = \begin{pmatrix} n \\ r \end{pmatrix}^{-1} \sum_{1\leq i_1} \sum_{<i_2<...} ... \sum_{<i_r\leq n} r^{-1} \sum_{k=0}^{r-1} (-1)^k \begin{pmatrix} r-1 \\ k \end{pmatrix} w_{i_{r-k:n}}, \; {\scriptstyle r=1,2,...,n.} \tag{5.46}
$$

Sample L-moments can be used similarly as the conventional sample moments because they characterize basic properties of the sample distribution and estimate the corresponding properties of the distribution from which the data sampled. They might be also used to estimate the parameters of this distribution. For asymptotic properties of L-moments one can refer to Hosking (1990).

This has been shown in the following simulation study. A comparison is made for the estimates produced by both the method of moments and method of L-moments and conclude that L-moments perform better.

5.5 Simulation Study

Here a simulation study to determine the biases and Root Mean Squared Errors(RMSEs) of the estimators discussed in previous section is presented. Our main interest here is to estimate the parameters θ_1', θ_2' and δ. For $\theta_1 = \frac{1}{125}$ and $\theta_2 = \frac{1}{81}$ the model in (5.27) reduces to the model given by

$$\begin{aligned} f_C(w) &= \frac{\theta_1'\left(e^{-\theta_1' w} - e^{-0.02w}\right)}{(1.65 - 81\,\theta_1')} + \frac{\theta_2'\left(e^{-\theta_2' w} - e^{-0.02w}\right)}{(2.54 - 125\,\theta_2')}, w < \delta \\ &= \frac{\theta_1'\, e^{-0.02w}\big[e^{-\delta(\theta_1' - 0.02)} - 1\big]}{(1.65 - 81\,\theta_1')} + \frac{\theta_2'\, e^{-0.02w}\big[e^{-\delta(\theta_2' - 0.02)} - 1\big]}{(2.54 - 125\,\theta_2')} \\ &+ 0.01 e^{-\delta\,\theta_2' - 0.01(w-\delta)}\left[1 - e^{-0.008(w-\delta)}\right] \\ &+ 0.008 e^{-\delta\,\theta_1' - 0.008(w-\delta)}\left[1 - e^{-0.01(w-\delta)}\right], w \geq \delta. \end{aligned} \tag{5.47}$$

We consider a simulate sample of sizes $n = 50,\ 75\ \textit{and}\ 100$ for different values of the parameters. For each combination of the parameters and n, we performed 1000 replications of the simulation. The results are presented in Tables 5.3 - 5.4.

Examining from Tables 5.3 - 5.4, we observe that as the sample size increases, biases and RMSE's of $\hat{\theta}_1'$, $\hat{\theta}_2'$ and $\hat{\delta}$ decrease steadily. Furthermore, we wish to make one more important observation is that the method of L-moments gives lesser biases and RMSEs for the estimators $\hat{\theta}_1'$, $\hat{\theta}_2'$ and $\hat{\delta}$ for all the sample sizes comparing to the moment method of estimation.

5.6 Some Applications

5.6.1 Cricket Data

In this section a sports data is analyzed. We consider the opening batters (consist of two players) in a cricket team. The traditional role of openers in cricket, especially in Test matches is to see off the new ball at the start of their team's innings. This task falls to both opening batters, both of whom should have good defensive technique against the hard, swinging, and seaming new ball. According to the experts from

Table 5.3: Absolute Bias, RMSE of $\hat{\theta}'_1$, $\hat{\theta}'_2$ and $\hat{\delta}$ based on 1000 Replications for $f_C(u)$ obtained using method of L-moment and method of moments for $\theta'_1 = 0.4$, $\theta'_2 = 0.6$, and $\delta = 5$.

	L-moments			Method of moments		
	$\widehat{\theta'_1}$	$\widehat{\theta'_2}$	$\widehat{\delta}$	$\widehat{\theta'_1}$	$\widehat{\theta'_2}$	$\widehat{\delta}$
n=50						
Absolute Bias	0.0820	0.0775	0.3887	0.0815	0.1595	0.6200
RMSE	0.1484	0.3217	2.2194	0.1529	0.3378	3.0319
n=75						
Absolute Bias	0.0712	0.0694	0.3236	0.0749	0.1515	0.3763
RMSE	0.1454	0.3216	2.2134	0.1518	0.3358	2.9869
n=100						
Absolute Bias	0.0633	0.0494	0.2609	0.0598	0.1506	0.2900
RMSE	0.1454	0.3205	2.1514	0.1502	0.3316	2.7138

this game, the first one hour is very crucial for the team batting and for the openers, during this time the failure rate of the batters is very high. Once they sustain that period, the ball becomes softer and less shiny as the innings progresses, making it easier for the batting team and scoring runs become much easier.

Here, we consider that the system comprises of two opening batsmen who have been playing together for India since 2001. Opener 1 has played so far 96 Innings. Opener 2 has played so far 180 Innings. They together opened the Innings for India on 87 occasions. The last Innings was on December 2012. We randomly selected 28 Innings score cards.The data is obtained from the official website of ESPN Cricinfo(www.stats.espncricinfo.com) ESPNCricinfo (2013) and it is presented in Table 4. We mainly focused on the average amount of time in which these two openers

Table 5.4: Absolute Bias, RMSE of $\hat{\theta}'_1$, $\hat{\theta}'_2$ and $\hat{\delta}$ based on 1000 Replications for $f_C(u)$ obtained using method of L-moment and method of moments for $\theta'_1 = 0.5$, $\theta'_2 = 0.7$ and $\delta = 6$.

	L-moments			Method of moments		
	$\widehat{\theta'_1}$	$\widehat{\theta'_2}$	$\widehat{\delta}$	$\widehat{\theta'_1}$	$\widehat{\theta'_2}$	$\widehat{\delta}$
n=50						
Absolute Bias	0.1298	0.0793	0.7517	0.2155	0.1590	1.9336
RMSE	0.1334	0.2955	2.0523	0.1444	0.3425	3.7810
n=75						
Absolute Bias	0.1187	0.0687	0.4365	0.1847	0.1500	1.8901
RMSE	0.1269	0.2830	1.9061	0.1439	0.3416	3.6734
n=100						
Absolute Bias	0.1094	0.0572	0.2384	0.1200	0.1123	1.7979
RMSE	0.1243	0.2703	1.7815	0.1412	0.3377	3.4415

spent on the crease while they are batting. It is observed as 125 minutes and 81 minutes respectively. Here failure is considered as loosing one's wicket. In our model the basic assumption is that both these openers have independent failure rates. The dependency comes only at the time one of these two openers fails. If opener 1 fails first, we investigate if opener 2 is disturbed for a critical time period, δ (threshold period) and once opener 2 sustains that critical period his failure rate will revert back to normal. Similarly, when opener 2 fails first, the failure rate of the opener 1 will increase for a critical time period δ and once opener 1 survives that period his failure rate reverts back to normal. Our main interest here is to estimate the parameters θ'_1, θ'_2 and δ, given that $\theta_1 = \frac{1}{125}$ and $\theta_2 = \frac{1}{81}$. The density function in (5.27) now reduces into the form

$$
\begin{aligned}
f_C(w) &= \frac{\theta_1'\left(e^{-\theta_1' w} - e^{-0.02w}\right)}{(1.65 - 81\,\theta_1')} + \frac{\theta_2'\left(e^{-\theta_2' w} - e^{-0.02w}\right)}{(2.54 - 125\,\theta_2')}, w < \delta \\
&= \frac{\theta_1'\, e^{-0.02w}[e^{-\delta(\theta_1' - 0.02)} - 1]}{(1.65 - 81\,\theta_1')} + \frac{\theta_2'\, e^{-0.02w}[e^{-\delta(\theta_2' - 0.02)} - 1]}{(2.54 - 125\,\theta_2')} \\
&+ 0.01e^{-\delta\,\theta_2' - 0.01(w-\delta)}\left[1 - e^{-0.008(w-\delta)}\right] \\
&+ 0.008e^{-\delta\,\theta_1' - 0.008(w-\delta)}\left[1 - e^{-0.01(w-\delta)}\right], w \geq \delta \qquad (5.48)
\end{aligned}
$$

Now, we estimate the parameters $\theta_1', \theta_2', \delta$ by using the method of L-moments and applying Kolmogorov-Smirnov test to fit the cumulative distribution function of the model $f_C(w)$ presented in (5.48).

Table 5.5: Time (In Minutes) spent on the crease for each innings by the two openers

Innings	Opener 1(Y_1)	Opener 2(Y_2)	Innings	Opener 1(Y_1)	Opener 2(Y_2)
1	13	16	2	25	13
3	23	73	4	23	86
5	70	60	6	17	6
7	82	91	8	9	7
9	43	11	10	62	41
11	23	74	12	48	55
13	115	96	14	183	219
15	24	65	16	39	23
17	42	77	18	77	74
19	25	22	20	63	64
21	52	81	22	79	64
23	65	47	24	28	35
25	14	2	26	21	40
27	17	57	28	83	52

* From Table 5.6, we accept the null hypothesis, that is the data presented in Table 4 fits well for Model 1, since the Kolmogorov-Smirnov test statistic, $(D = 0.2301)$ is less than the Kolmogorov-Smirnov table value $(D_{28} = 0.2499)$, at $\alpha = 0.05$ level of significance. Therefore, we have considered Model 1 for our further data analysis. From the L-moments we obtained the estimates as $\hat{\theta}_1' \simeq \frac{1}{5}$, $\hat{\theta}_2' \simeq \frac{1}{11}$, and $\hat{\delta} \simeq 26.1021$. From the analysis, we conclude that when opener 1 fails first, the failure rate of

Table 5.6: Models with Kolmogorov-Smirnov test statistic values

Model	K-S test statistic
Model 1 as in (5.33)	0.2301*
Model 2 as in (5.36)	0.2712
Model 3 as in (5.37)	0.9851

opener 2 increases approximately $\frac{1}{11}$ times for the next 27 minutes; once he survives that critical period (threshold period) his failure rate will revert back to its original value. Similarly, when opener 2 fails first, the failure rate of opener 2 increases approximately $\frac{1}{4}$ times for the next 27 minutes; once he sustains that critical period (threshold period) his failure rate will revert back to its original value.

To test whether the failure of opener 1 does not increase the load on opener 2, against the alternative that the failure of opener 1 increases the load on opener 2, similarly the failure of opener 2 does not increase the load on opener 1 against the alternative that the failure of opener 2 increases the load on opener 1, we refer to the likelihood ratio test in Asha et al. (2016). For our analysis we consider the model given in (5.13). The test procedure is as follows:

$$H_0 : \theta_1 = \theta_1', \ \theta_2 = \theta_2',$$

against the alternative

$$H_1 : \theta_1 \neq \theta_1', \ \theta_2 \neq \theta_2'.$$

The log-likelihood ratio

$$\Lambda = \frac{\ell(\boldsymbol{\lambda})|H_0}{\ell(\boldsymbol{\lambda})|H_1},$$

where $\boldsymbol{\lambda} = (\theta_1, \theta_2, \theta_1', \theta_2')$. Since the log-likelihood ratio, $\Lambda = 0.8316 < 1$ we reject H_0 in favour of H_1: The failure of opener 1 increases the load on opener 2, similarly the failure of opener 2 increases the load on opener 1. Hence, there is sufficient evidence for the presence of load sharing effect in the cricket data.

5.7 Conclusions

The model that we proposed in (5.11) incorporates changes in the performance of a two-component system due to the failure of first or second component. For various choices of distribution function, model (5.11) provides us with families of bivariate distributions which include load sharing and cascading failures.

On failure of one component, the surviving component may have extra stress for a specific period of time, during that time it may fail or it can sustain that period so that its failure rate reverts back to normal. For the various choices of the parameters θ_i, θ_i' and δ, $i = 1, 2$ our model reduces into the existing models such as Freund (1961), Lindley & Singpurwalla (2002), Swift (2008) and Asha et al. (2016). The method of L-moments is an alternative approach since conventional methods of parameter estimation like maximum likelihood is not conducive.

From the data analysis we estimated the 'threshold time' or 'critical time' for Opener 1 when Opener 2 fails first or for opener 2 when opener 1 fails first is estimated as 27 minutes. The approximated crucial time of half an hour after the failure of any one of the openers is universally accepted by most of the pundits in the game of cricket. From the likelihood ratio test we concluded that, the failure of opener 1 increases the load on opener 2, similarly the failure of opener 2 increases the load on opener 1. Therefore, there is a presence of load sharing effect in the cricket data.

Chapter 6

Conclusions and Future Work

6.1 Overall Summary

In Chapter 1, we discussed some of the basic concepts related to the work in the present Thesis. Some of the popular methods for constructing bivariate distributions are reviewed. We introduced and studied extensively a new univariate distribution, namely transmuted exponentiated Frêchet distribution. This study formed the basis of introducing a general class of bivariate distribution with transmuted conditionals.

In Chapter 2, we proposed a new general class of bivariate distributions with transmuted conditionals. We studied the general and the particular properties of the proposed model. Examples for the general model are constructed by considering various baseline distributions such as uniform, normal, exponential, Weibull and exponentiated Frêchet. Method of moments and profile likelihood approach were considered for estimating the parameters. A simulation study is conducted to show the efficiency of our estimation procedures. Two well analysed data sets are considered for model applicability and made comparison with the existing models in the literature. Finally we concluded that bivariate distribution with transmuted conditionals performs better compared to the existing model in the literature.

In Chapter 3, a new general class of distributions for load sharing with frailty and covariates was studied. It is shown that by considering different frailty distributions and re-parametrizing, the general model can reduce to some of the existing models in the literature. Examples for the general model are constructed by considering various frailty distributions such as gamma, inverse-Gaussian and positive stable. A general estimation procedure for the proposed model is discussed and a general algorithm for performing simulation study is also studied in detail.

In Chapter 4, a particular example of the general model proposed in Chapter 3 is studied in detail. We have considered positive stable as frailty model and Weibull as the baseline distribution. The properties which are discussed for the general model in Chapter 3 are also discussed for the new model. The profile likelihood approach is employed to estimate the unknown parameters. A simulation study is conducted to show that the estimation procedure performs well for our model. Two well - analysed data sets are considered for data analysis purposes and conclude that the load share model with frailty and covariates performs well compared to the other competing models which do not take into account load share or frailty.

In chapter 5, we proposed a general class of bivariate distribution for cascading failures by extending the works of Lindley & Singpurwalla (2002), Swift (2008) and Asha et al. (2016) using Cox total failure rate (Cox (1972)). We studied the model extensively with a special example by considering exponential baseline. Method of moments and L-moments have been used to estimate the unknown parameters. Simulation study was conducted to show the effectiveness of our estimation procedures and provided evidences for the better performances of L-moments. A real life data set has been analysed to show the applicability of our model. From the data analysis the 'threshold time' or 'critical time' δ for Opener 1 when Opener 2 fails first or for opener 2 when opener 1 fails first is estimated as 27 minutes. This crucial time is universally accepted by the most of the experts in the game of cricket.

In Chapter 6, we discuss the results and conclusion of the present thesis, we list some of our future research problems. One of the interesting problems is the discrete analogues for bivariate distributions for load sharing models discussed in Chapter 3. We proposed a new class of bivariate distribution for discrete load share models. Examples for the proposed model is constructed by considering different baseline distributions such as geometric, discrete Weibull, S distribution and Waring distribution. General properties of the proposed model is discussed. General estimation procedure for the proposed model is presented.

As a result of this extensive research we put up with some possible open problems which will be considered for future works. In Section 6.2 we discuss in detail about some of our future research problems.

6.2 Work in Progress/Future Work

6.2.1 Discrete Bivariate Distributions for Load Sharing Models

In reliability literature, lifetime data has been analysed by variety of bivariate distributions. Most of the existing models deal with continuous failures rates. Discrete failure rates quite often occur with situation where product life can best be described through non-negative integer valued random variables. For instance lifetime of tyres on a jet fighter is measured by the number of landings it has undergone, the life of a ship is measured by the number of successful voyages it has made, life of a gas lighter is measured by the number of shots, life of a weapon is measured by the number of rounds fired, life of a switch is measured by the number of strokes, life of a motor is measured by the number of complete rotations. Therefore it is important to develop reliability theory for discrete descriptions like its continuous version.

In the existing literature few works have been done in the area of discrete bivariate distributions; particularly, bivariate geometric models and their properties. The bivariate geometric model proposed by Basu & Dhar (1995) is an analogue to the bivariate distribution of Marshall & Olkin (1967). In an alternative approach Dhar (1998) derived a bivariate geometric model which is a discrete analogue to Freund's model. Lee & Cha (2015) proposed two general methods namely, the minimisation and the maximisation methods to generate new class of discrete bivariate distributions and studied some special distributions namely bivariate Poisson, bivariate geometric, bivariate negative binomial and bivariate binomial distributions. Achcar et al. (2016) extended, Basu-Dhar bivariate geometric distribution in the presence of covariates and censored data. They estimated the parameters using Bayesian approach. Recently, Vahid & Kundu (2017) developed discrete bivariate generalized exponential distribution, whose marginals are discrete generalized exponential distribution and applied EM-algorithm to estimate the unknown parameters.

In this section, we propose a general bivariate discrete load share model for two-component parallel system. Both the components are assumed to have independent failure rates and the dependence comes when one of the two-components fails. Model formulation is presented. The general properties of the model are discussed. some of the general estimation procedures are listed. Simulation study and data analysis are aimed for immediate future works.

6.2.1.1 Model Formulation

Let (Y_1, Y_2) be a random vector with support in $\mathscr{N}_+{}^2$, where $\mathscr{N}_+ = \{1, 2, 3, ...,\}$ and $\mathscr{N}_+{}^2 = \{(Y_1, Y_2)|(y_1, y_2) \in \mathscr{N}_+\}$. The cumulative distribution function for the discrete case is defined as $F(x) = P[X \leq x]$ and the corresponding survival function is defined as $S(x) = P[X \geq x]$. Denote the joint probability function for (Y_1, Y_2) by

$$p(y_1, y_2) = P[Y_1 = y_1, Y_2 = y_2],\ (y_1, y_2) \in \mathscr{N}_+{}^2. \tag{6.1}$$

We assume that Y_1 and Y_2 are the discrete lifetimes of two-components, A and B. The system considered here is same as the system in the formulation of model 3.2 except that the lifetimes are discrete and measured after every completed cycle of time. Hence Y_1 and Y_2 are independent discrete random variables in N^2 representing the lifetimes of the components A and B. The discrete bivariate conditional hazard rate function (Shaked et al. (1995)) of (Y_1, Y_2) is now

$$\lambda_1(y) = \theta_1 r(y)(1 - \theta_2 r(y)),\ y \in \mathscr{N}_+ \tag{6.2}$$

$$\lambda_2(y) = \theta_2 r(y)(1 - \theta_1 r(y)),\ y \in \mathscr{N}_+ \tag{6.3}$$

$$\lambda_{12}(y) = \theta_1 \theta_2 r^2(y),\ y \in \mathscr{N}_+ \tag{6.4}$$

$$\lambda_1(y_1|y_2) = P[Y_1 = y_1|Y_1 \geq y_1, Y_2 = y_2]; y_1 > y_2 = \theta_1' r(y),\ (y_1, y_2) \in \mathscr{N}_+{}^2, \tag{6.5}$$

$$\lambda_2(y_2|y_1) = P[Y_2 = y_2|Y_1 = y_1, Y_2 \geq y_2]; y_2 > y_1 = \theta_2' r(y),\ (y_1, y_2) \in \mathscr{N}_+{}^2. \tag{6.6}$$

Clearly, bivariate conditional hazard rate functions given in (6.2)-(6.6) are determined through $p(y_1, y_2)$ (Shaked et al. (1995)).
For $y_1 < y_2$,

$$\begin{aligned} p(y_1, y_2) &= P\{Y_1 = y_1, Y_2 = y_2\} \\ &= P[Y_1 \geq y_1, Y_2 \geq y_2] P[Y_1 = y_1, Y_2 > y_1 | Y_1 \geq y_1, Y_2 \geq y_1] \\ &\times P[Y_2 \geq y_2 | Y_1 = y_1, Y_2 > y_1] P[Y_2 = y_2 | Y_1 = y_1, Y_2 \geq y_2];\ y_1 < y_2 \end{aligned}$$

and

$$p(y_1, y_2) = P[Y_1 \geq y_1] P[Y_2 \geq y_1] \lambda_1(y_1) \prod_{i=y_1+1}^{y_2-1} (1 - \lambda_2(i|y_1)) \lambda_2(y_2|y_1),\ y_1 < y_2, \quad (6.7)$$

$$p(y_1, y_2) = \prod_{i=1}^{y_1} (1 - \theta_1 r(i))(1 - \theta_2 r(i)) \theta_1 r(y_1)(1 - \theta_2 r(y_1)) \prod_{i=y_1+1}^{y_2-1} (1 - \theta_2' r(i)) \theta_2' r(y_2),\ y_1 < y_2. \quad (6.8)$$

Thus for $y_1 < y_2$

$$p(y_1, y_2) = (1 - \theta_2 r(y_2))^{y_1-1} (1 - \theta_1 r(y_1))^{y_1-1} \theta_1 r(y_1)(1 - \theta_2 r(y_1)) \prod_{i=y_1+1}^{y_2-1} (1 - \theta_2' r(i)) \theta_2' r(y_2) \quad (6.9)$$

Similarly, for $y_2 < y_1$

$$p(y_1, y_2) = (1 - \theta_1 r(y_1))^{y_2-1} (1 - \theta_2 r(y_2))^{y_2-1} \theta_2 r(y_2)(1 - \theta_1 r(y_2)) \prod_{i=y_2+1}^{y_1-1} (1 - \theta_1' r(i)) \theta_1' r(y_1) \quad (6.10)$$

and

$$p(y, y) = \prod_{i=1}^{y-1} (1 - \theta_1 r(i))(1 - \theta_2 r(i)) \theta_1 r(i) \theta_2 r(i),\ y_1 = y_2 \quad (6.11)$$

The general joint density function is given by

$$\begin{aligned} p(y_1, y_2) &= (1 - \theta_2 r(y_2))^{y_1-1} (1 - \theta_1 r(y_1))^{y_1-1} \theta_1 r(y_1)(1 - \theta_2 r(y_1)) \\ &\times \prod_{i=y_1+1}^{y_2-1} (1 - \theta_2' r(i)) \theta_2' r(y_2);\ y_1 < y_2 \end{aligned}$$

$$
\begin{aligned}
&= (1-\theta_1 r(y_1))^{y_2-1}(1-\theta_2 r(y_2))^{y_2-1}\theta_2 r(y_2)(1-\theta_1 r(y_2))\\
&\times \prod_{i=y_2+1}^{y_1-1}(1-\theta_1' r(i))\theta_1' r(y_1);\; y_2 < y_1\\
&= \prod_{i=1}^{y-1}(1-\theta_1 r(i))(1-\theta_2 r(i))\theta_1 r(i)\theta_2 r(i),\; y_1 = y_2 = y
\end{aligned}
\tag{6.12}
$$

6.2.1.2 Properties

In this we discuss the various distributional properties of general discrete bivariate load share model given in (6.12).

Property 6.2.1. *The joint survival function of* (Y_1, Y_2) *is given by*

$$
\begin{aligned}
S(y_1,y_2) &= \sum_{i=y_1}^{y_2-1}\theta_1 r(i)(1-\theta_2 r(i)) \times \left[\prod_{j=i+1}^{y_2-1}(1-\theta_2' r(j))\right]\\
&\times \left[\prod_{j=1}^{i-1}(1-\theta_1 r(j)(1-\theta_2 r(j)) - \theta_2 r(j)(1-\theta_1 r(j)) - \theta_1\theta_2 r^2(j)\right]\\
&+ \left[\prod_{j=1}^{y_2-1}(1-\theta_1 r(j)(1-\theta_2 r(j)) - \theta_2 r(j)(1-\theta_1 r(j)) - \theta_1\theta_2 r^2(j)\right]; y_1 \le y_2,\\
&= \sum_{i=y_2}^{y_1-1}\theta_2 r(i)(1-\theta_1 r(i)) \times \left[\prod_{j=i+1}^{y_1-1}(1-\theta_1' r(j))\right]\\
&\left[\prod_{j=1}^{i-1}(1-\theta_1 r(j)(1-\theta_2 r(j)) - \theta_2 r(j)(1-\theta_1 r(j)) - \theta_1\theta_2 r^2(j)\right]\\
&+ \left[\prod_{j=1}^{y_1-1}(1-\theta_1 r(j)(1-\theta_2 r(j)) - \theta_2 r(j)(1-\theta_1 r(j)) - \theta_1\theta_2 r^2(j)\right]; y_2 \le y_1.\\
&= \sum_{i=1}^{y-1}\theta_2 r(i)(1-\theta_1 r(i)) \times \left[\prod_{j=i+1}^{y-1}(1-\theta_1' r(j))\right]\\
&\left[\prod_{j=1}^{i-1}(1-\theta_1 r(j)(1-\theta_2 r(j)) - \theta_2 r(j)(1-\theta_1 r(j)) - \theta_1\theta_2 r^2(j)\right]
\end{aligned}
$$

$$+ \left[\prod_{j=1}^{y-1} (1 - \theta_1 r(j)(1 - \theta_2 r(j)) - \theta_2 r(j)(1 - \theta_1 r(j)) - \theta_1 \theta_2 r^2(j) \right]; y_1 = y_2 = y. \tag{6.13}$$

Proof.

$$S(y_1, y_2) = P[Y_1 \geq y_1, Y_2 \geq y_2]$$

For $y_1 \leq y_2$

$$S(y_1, y_2) = \left\{ \sum_{i=1}^{y_2-1} \lambda_1(i) \left[\prod_{j=1}^{i-1} (1 - \lambda^*(j)) \right] \left[\prod_{j=i+1}^{y_2-1} (1 - \lambda_2(j|i)) \right] \right\} + \prod_{j=1}^{y_2-1} (1 - \lambda^*(j)). \tag{6.14}$$

Now by substituting the conditional hazard rate function given in (6.1)-(6.6) we get

$$\begin{aligned} S(y_1, y_2) = & \sum_{i=y_1}^{y_2-1} \theta_1 r(i)(1 - \theta_2 r(i)) \times \left[\prod_{j=i+1}^{y_2-1} (1 - \theta_2' r(j)) \right] \\ & \times \left[\prod_{j=1}^{i-1} (1 - \theta_1 r(j)(1 - \theta_2 r(j)) - \theta_2 r(j)(1 - \theta_1 r(j)) - \theta_1 \theta_2 r^2(j) \right] \\ & + \left[\prod_{j=1}^{y_2-1} (1 - \theta_1 r(j)(1 - \theta_2 r(j)) - \theta_2 r(j)(1 - \theta_1 r(j)) - \theta_1 \theta_2 r^2(j) \right]; y_1 \leq y_2. \end{aligned} \tag{6.15}$$

Similarly, for $y_2 \leq y_1$

$$S(y_1, y_2) = \left\{ \sum_{i=1}^{y_1-1} \lambda_2(i) \left[\prod_{j=1}^{i-1} (1 - \lambda^*(j)) \right] \left[\prod_{j=i+1}^{y_1-1} (1 - \lambda_1(j|i)) \right] \right\} + \prod_{j=1}^{y_1-1} (1 - \lambda^*(j)). \tag{6.16}$$

Again by substituting the conditional hazard rate function given in (6.1)-(6.6) we get

$$\begin{aligned} S(y_1, y_2) = & \sum_{i=y_2}^{y_1-1} \theta_2 r(i)(1 - \theta_1 r(i)) \times \left[\prod_{j=i+1}^{y_1-1} (1 - \theta_1' r(j)) \right] \\ & \times \left[\prod_{j=1}^{i-1} (1 - \theta_1 r(j)(1 - \theta_2 r(j)) - \theta_2 r(j)(1 - \theta_1 r(j)) - \theta_1 \theta_2 r^2(j) \right] \end{aligned}$$

$$+\left[\prod_{j=1}^{y_1-1}(1-\theta_1 r(j)(1-\theta_2 r(j))-\theta_2 r(j)(1-\theta_1 r(j))-\theta_1\theta_2 r^2(j)\right]; y_2 \le y_1. \tag{6.17}$$

Finally for $y_1 = y_2 = y$

$$S(y_1,y_2)=\left\{\sum_{i=1}^{y-1}\lambda_2(i)\left[\prod_{j=1}^{i-1}(1-\lambda^*(j))\right]\left[\prod_{j=i+1}^{y-1}(1-\lambda_1(j|i))\right]\right\}+\prod_{j=1}^{y-1}(1-\lambda^*(j)). \tag{6.18}$$

From (6.1)-(6.6) we get

$$\begin{aligned}S(y_1,y_2)=&\sum_{i=1}^{y-1}\theta_2 r(i)(1-\theta_1 r(i))\times\left[\prod_{j=i+1}^{y-1}(1-\theta_1' r(j))\right]\\&\times\left[\prod_{j=1}^{i-1}(1-\theta_1 r(j)(1-\theta_2 r(j))-\theta_2 r(j)(1-\theta_1 r(j))-\theta_1\theta_2 r^2(j)\right]\\&+\left[\prod_{j=1}^{y-1}(1-\theta_1 r(j)(1-\theta_2 r(j))-\theta_2 r(j)(1-\theta_1 r(j))-\theta_1\theta_2 r^2(j)\right]; y_1=y_2=y.\end{aligned} \tag{6.19}$$

Thus we retrieve (6.13). ■

Remark 6.2.1. *When $y_1 = y_2 = 1$ it follows that $S(1,1) = 1$.*

Property 6.2.2. *The marginal survival function of Y_j is obtained as $S(y_1, 0) = S(y_1)$ and $S(0, y_2) = S(y_2)$. In general the marginal survival functions for the model in (6.12) is given by*

$$\begin{aligned}S(y_i)=&\sum_{t_i=y_i}^{y_j-1}\sum_{t_j=y_j}^{\infty}\prod_{k=1}^{y_i-1}(1-\theta_i r(k))(1-\theta_j r(k))\prod_{k=y_i+1}^{y_j-1}(1-\theta_j' r(k))\theta_j'\\&+\prod_{k=1}^{y_j-1}(1-\theta_i r(k))(1-\theta_j r(k));\ y_i<y_j,\ i=1,2,\ i\neq j.\end{aligned} \tag{6.20}$$

The marginal densities for the model in (6.12) is given by

$$f(y_i)=-\frac{\partial S(y_i,y_j)}{\partial y_j}$$

$$= -\frac{\partial}{\partial y_j}\left[\sum_{y_i=t_i}^{t_j-1}\sum_{y_j=t_j}^{\infty}\prod_{k=1}^{t_i-1}(1-\theta_i r(k))(1-\theta_j r(k))\prod_{k=t_i+1}^{t_j-1}(1-\theta_j' r(k))\theta_j'\right]$$
$$-\frac{\partial}{\partial y_j}\left[\prod_{k=1}^{y_j-1}(1-\theta_i r(k))(1-\theta_j r(k))\right];\ y_i < y_j,\ i=1,2,\ i\neq j. \tag{6.21}$$

6.2.1.3 Some Examples

Example 6.2.1. *Let the random variables Y_1 and Y_2 have independent geometric distributions. Then the bivariate geometric load share model is given by*

$$\begin{aligned} p(y_1,y_2) &= (1-P_1)^{y_1-1}(1-P_2)^{y_1-1}P_1(1-P_2)(1-P_2')^{y_2-y_1-1}P_2';\ y_1<y_2 \\ &= (1-P_1)^{y_2-1}(1-P_2)^{y_2-1}P_2(1-P_1)(1-P_1')^{y_1-y_2-1}P_1';\ y_2<y_1 \\ &= (1-P_1)^{y-1}(1-P_2)^{y-1}P_1P_2;\ y_1=y_2=y \end{aligned} \tag{6.22}$$

where $\theta_1 P = P_1$, $\theta_2 P = P_2$, $\theta_1' P = P_1'$ and $\theta_2' P = P_2'$.

Example 6.2.2. *Let the random variables Y_1 and Y_2 have independent discrete Weibull distributions with $r(y) = (\frac{y}{m})^{\alpha-1}$; $1 \le y \le m$ then the bivariate discrete Weibull load share model is given by*

$$\begin{aligned} p(y_1,y_2) &= \prod_{i=1}^{t_1-1}\left[1-\theta_1\left(\frac{y(i)}{m}\right)^{\alpha-1}\right]\left[1-\theta_2\left(\frac{y(i)}{m}\right)^{\alpha-1}\right]\theta_1\left(\frac{y_1}{m}\right)^{\alpha-1}\left[1-\theta_2\left(\frac{y_1}{m}\right)^{\alpha-1}\right] \\ &\times\prod_{i=t_1+1}^{t_2-1}\left[1-\theta_2'\left(\frac{y(i)}{m}\right)^{\alpha-1}\right]\theta_2'\left[\left(\frac{y_2}{m}\right)^{\alpha-1}\right];\ y_1<y_2 \\ &= \prod_{i=1}^{t_2-1}\left[1-\theta_1\left(\frac{y(i)}{m}\right)^{\alpha-1}\right]\left[1-\theta_2\left(\frac{y(i)}{m}\right)^{\alpha-1}\right]\theta_2\left(\frac{y_2}{m}\right)^{\alpha-1}\left[1-\theta_1\left(\frac{y_2}{m}\right)^{\alpha-1}\right] \\ &\times\prod_{i=t_2+1}^{t_1-1}\left[1-\theta_1'\left(\frac{y(i)}{m}\right)^{\alpha-1}\right]\theta_1'\left[\left(\frac{y_1}{m}\right)^{\alpha-1}\right];\ y_2<y_1 \\ &= \prod_{i=1}^{t-1}\left[1-\theta_1\left(\frac{y(i)}{m}\right)^{\alpha-1}\right]\left[1-\theta_2\left(\frac{y(i)}{m}\right)^{\alpha-1}\right]\theta_1\left(\frac{y(i)}{m}\right)^{\alpha-1}\theta_2\left(\frac{y(i)}{m}\right)^{\alpha-1};y_1=y_2 \end{aligned} \tag{6.23}$$

Example 6.2.3. *Let the random variables Y_1 and Y_2 have independent S distribution* *with $r(y) = q(1-\pi^y)$, where $y \in \mathscr{N}_+$, $0 < q < 1$, $0 < \pi < 1$, then the bivariate discrete S load share model is given by*

$$\begin{aligned} p(y_1, y_2) &= \prod_{i=1}^{t_1-1} \left[1-\theta_1 q\left(1-\pi^{y(i)}\right)\right]\left[1-\theta_2 q\left(1-\pi^{y(i)}\right)\right]\theta_1 q\left(1-\pi^{y_1}\right)\left[1-\theta_2 q\left(1-\pi^{y_1}\right)\right] \\ &\times \prod_{i=t_1+1}^{t_2-1} \left[1-\theta_2' q\left(1-\pi^{y(i)}\right)\right]\theta_2' q\left(1-\pi^{y_2}\right);\ y_1 < y_2 \\ &= \prod_{i=1}^{t_2-1} \left[1-\theta_2 q\left(1-\pi^{y(i)}\right)\right]\left[1-\theta_1 q\left(1-\pi^{y(i)}\right)\right]\theta_2 q\left(1-\pi^{y_2}\right)\left[1-\theta_1 q\left(1-\pi^{y_2}\right)\right] \\ &\times \prod_{i=t_2+1}^{t_1-1} \left[1-\theta_1' q\left(1-\pi^{y(i)}\right)\right]\theta_1' q\left(1-\pi^{y_1}\right);\ y_2 < y_1 \\ &= \prod_{i=1}^{t-1} \left[1-\theta_1 q\left(1-\pi^{y(i)}\right)\right]\left[1-\theta_2 q\left(1-\pi^{y(i)}\right)\right]\theta_1\theta_2 q^2\left(1-\pi^{y(i)}\right)^2;\ y_1 = y_2 \end{aligned} \tag{6.24}$$

Example 6.2.4. *Let the random variables Y_1 and Y_2 follow independent simple discrete DFR distribution with $r(y) = \frac{c}{y+1}$; $0 \le c \le 1$. Then the bivariate simple discrete DFR load share model is given by*

$$\begin{aligned} p(y_1, y_2) &= \prod_{i=1}^{t_1-1} \left[1-\theta_1\left(\frac{c}{y(i)+1}\right)\right]\left[1-\theta_2\left(\frac{c}{y(i)+1}\right)\right]\theta_1\left(\frac{c}{y_1+1}\right) \\ &\times \left[1-\theta_2\left(\frac{c}{y_1+1}\right)\right]\prod_{i=t_1+1}^{t_2-1}\left[1-\theta_2'\left(\frac{c}{y(i)+1}\right)\right]\theta_2'\left[\left(\frac{c}{y_2+1}\right)\right];\ y_1 < y_2 \\ &= \prod_{i=1}^{t_2-1} \left[1-\theta_1\left(\frac{c}{y(i)+1}\right)\right]\left[1-\theta_2\left(\frac{c}{y(i)+1}\right)\right]\theta_2\left(\frac{c}{y_2+1}\right) \\ &\times \left[1-\theta_1\left(\frac{c}{y_2+1}\right)\right]\prod_{i=t_2+1}^{t_1-1}\left[1-\theta_1'\left(\frac{c}{y(i)+1}\right)\right]\theta_1'\left[\left(\frac{c}{y_1+1}\right)\right];\ y_2 < y_1 \\ &= \prod_{i=1}^{t-1} \left[1-\theta_1\left(\frac{c}{y(i)+1}\right)\right]\left[1-\theta_2\left(\frac{c}{y(i)+1}\right)\right]\theta_1\left(\frac{c}{y(i)+1}\right)\theta_2\left(\frac{c}{y(i)+1}\right);y_1 = y_2 \end{aligned} \tag{6.25}$$

Example 6.2.5. *Let the random variables Y_1 and Y_2 follow independent Waring distribution with $r(y) = \frac{1}{\beta+\phi y}$; $\beta > 1$, $0 < \phi < 1$ then the discrete bivariate Waring* *load share model is given by*

$$\begin{aligned}
p(y_1, y_2) &= \prod_{i=1}^{t_1-1}\left[1-\theta_1\left(\frac{1}{\beta+\phi y(i)}\right)\right]\left[1-\theta_2\left(\frac{1}{\beta+\phi y(i)}\right)\right]\theta_1\left(\frac{1}{\beta+\phi y_1}\right)\\
&\times\left[1-\theta_2\left(\frac{1}{\beta+\phi y_1}\right)\right]\prod_{i=t_1+1}^{t_2-1}\left[1-\theta_2'\left(\frac{1}{\beta+\phi y(i)}\right)\right]\theta_2'\left[\left(\frac{1}{\beta+\phi y_2}\right)\right];\ y_1<y_2\\
&= \prod_{i=1}^{t_2-1}\left[1-\theta_1\left(\frac{1}{\beta+\phi y(i)}\right)\right]\left[1-\theta_2\left(\frac{1}{\beta+\phi y(i)}\right)\right]\theta_2\left(\frac{1}{\beta+\phi y_2}\right)\\
&\times\left[1-\theta_1\left(\frac{1}{\beta+\phi y_2}\right)\right]\prod_{i=t_2+1}^{t_1-1}\left[1-\theta_1'\left(\frac{1}{\beta+\phi y(i)}\right)\right]\theta_1'\left[\left(\frac{1}{\beta+\phi y_1}\right)\right];\ y_2<y_1\\
&= \prod_{i=1}^{t-1}\left[1-\theta_1\left(\frac{1}{\beta+\phi y(i)}\right)\right]\left[1-\theta_2\left(\frac{1}{\beta+\phi y(i)}\right)\right]\\
&\times\theta_1\left(\frac{1}{\beta+\phi y(i)}\right)\theta_2\left(\frac{1}{\beta+\phi y(i)}\right);y_1=y_2=y \qquad (6.26)
\end{aligned}$$

6.2.1.4 Parameter Estimation

In this section we employ the method of maximum likelihood estimation to estimate the unknown parameters of the model given in (6.12). Let us assume that $\mathcal{E} = \{(y_{11}, y_{21}), (y_{12}, y_{22})..., (y_{1n}, y_{2n})\}$ is a bivariate sample of size n. n_1=number of observations for which $y_{1i} < y_{2i}$, n_2 = number of observations for which $y_{1i} > y_{2i}$, n_3 = number of observations for which $y_{1i} = y_{2i} = y_i$.

The general log-likelihood function for the model given in (6.12) can be written as

$$\begin{aligned}
\ell(\boldsymbol{\lambda}) &= (y_{1i}-1)\sum_{i=1}^{n_1}\log[1-\theta_2 r(y_{2i})] + (y_{1i}-1)\sum_{i=1}^{n_1}\log[1-\theta_1 r(y_{1i})] + \theta_1\sum_{i=1}^{n_1}\log[r(y_{1i})]\\
&+\sum_{i=1}^{n_1}\log[1-\theta_2 r(y_{1i})] + \sum_{i=1}^{n_1}\log[1-\theta_2' r(i)\theta_2' r(y_{2i})]\\
&+(y_{2i}-1)\sum_{i=1}^{n_2}\log[1-\theta_1 r(y_{1i})] + (y_{2i}-1)\sum_{i=1}^{n_2}\log[1-\theta_2 r(y_{2i})] + \theta_2\sum_{i=1}^{n_2}\log[r(y_{2i})]
\end{aligned}$$

$$+\sum_{i=1}^{n_2}\log[1-\theta_1 r(y_{2i})]+\sum_{i=1}^{n_2}\log[1-\theta_1' r(i)\theta_1' r(y_{1i})]$$
$$+\sum_{i=1}^{n_3}\log[1-\theta_1 r(y_i)]+\sum_{i=1}^{n_3}\log[1-\theta_2 r(y_i)+\sum_{i=1}^{n_3}\log(\theta_1 r(y_i))+\sum_{i=1}^{n_3}\log(\theta_2 r(y_i)) \quad (6.27)$$

where $\boldsymbol{\lambda}=(\theta_1,\theta_2,\theta_1',\theta_2')$. The MLEs of the unknown parameters can be obtained by maximizing (6.27) with respect to the unknown parameters.

It further remains to investigate various properties enjoyed by (6.12). This will be taken up as immediate future work.

6.2.2 Copula Approach to Bivariate Transmuted Distributions

In the present thesis we have constructed bivariate distributions with two methods namely the conditional specification approach and frailty approach. These models can also be constructed by other methods such as copula models especially the bivariate distributions with transmuted conditionals model. Bayesian approach may be an alternative estimation procedure can be performed.

6.2.3 Multivariate Load Share Models with Frailty

Multivariate extension for load share frailty models and Bayesian parametric approach is another possible future work. Note that extension of this model to higher dimension is not straight forward and every additional dimension needs specific model formulation.

6.2.4 Cascading Models with Random Critical Time

For the cascading model in Chapter 5, the critical time δ in the cascading failure model can be considered as a random variable. By assigning a suitable distribution for the random variable we can construct a new model which can be a further flexible and rich model to analyse cascading failures and the parameter estimation can be performed through Bayesian Technique.

These further extensions and research works are presently in progress.

List of Published / Communicated Papers

(1). Asha, G., **Raja, A. V.,** & Ravishanker, N. (2017). Reliability modelling incorporating load share and frailty. *Applied Stochastic Models in Business & Industry*, 34 (2), 206-223.

(2). Asha, G., & **Raja, A. V.** (2017). An extension of the Freund's bivariate distribution to model cascading failures. *Communications in Statistics-Simulation & Computation*, 46 (7), 5516-5530.

(3). Elbatal, I., Asha, G., & **Raja, A. V.** (2014). Transmuted exponentiated Frêchet distribution: Properties and applications. *Journal of Statistics Applications & Probability*, 3 (3), 379-394.

(4). Sarabia, J.M., **Raja, A. V.,** & Asha, G. (2017): Bivariate distributions with transmuted conditionals: Models and Applications (Communications in Statistics-Theory & Methods, Revision after Submission).

References

Abouammoh, A. M., & Alshingiti, A. M. (2009). Reliability estimation of generalized inverted exponential distribution. *Journal of Statistical Computation & Simulation*, *79*(11), 1301–1315.

Achcar, J. A., Davarzani, N., & Souza, R. M. (2016). Basu–Dhar bivariate geometric distribution in the presence of covariates and censored data: A Bayesian approach. *Journal of Applied Statistics*, *43*(9), 1636–1648.

Arnold, B. C. (1987). Bivariate distributions with Pareto conditionals. *Statistics & Probability Letters*, *5*(4), 263–266.

Arnold, B. C., Castillo, E., & Sarabia, J. M. (1999a). *Conditional specification of statistical models*. Springer Science & Business Media.

Arnold, B. C., Castillo, E., & Sarabia, J. M. (1999b). *Conditional specification of statistical models*. Springer-Verlag.

Arnold, B. C., & Strauss, D. J. (1988). Bivariate distributions with exponential conditionals. *Journal of the American Statistical Association*, *83*(402), 522–527.

Arnold, B. C., & Strauss, D. J. (1991). Bivariate distributions with conditionals in prescribed exponential families. *Journal of the Royal Statistical Society. Series B (Methodological)*, 365–375.

Aryal, G. R. (2013). Transmuted log-logistic distribution. *Journal of Statistics Applications & Probability*, *2*(1), 11 - 20.

Aryal, G. R., & Tsokos, C. P. (2009). On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods & Applications*, *71*, 1401 - 1407.

Aryal, G. R., & Tsokos, C. P. (2011). Transmuted Weibull distribution: A generalization of the Weibull probability distribution. *European Journal of Pure & Applied Mathematics*, *4*(2), 89-102.

Asha, G., Jagathnath, K. K. M., & Kundu, D. (2016). An extension of the Freund's bivariate distribution to model load-sharing systems. *American Journal of Mathematical and Management Sciences*, *35*(3), 1–20.

Asha, G., & Raja, A. V. (2017). An extension of the Freund's bivariate distribution to model cascading failures. *Communications in Statistics-Simulation & Computation*, *46*(7), 5516–5530.

Asha, G., Raja, A. V., & Ravishanker, N. (2017). Reliability modelling incorporating load share and frailty. *Applied Stochastic Models in Business & Industry*, *34*(2), 206–223.

Basu, A. P. (1971). Bivariate failure rate. *Journal of the American Statistical Association*, *66*(333), 103–104.

Basu, A. P., & Dhar, S. K. (1995). Bivariate geometric distribution. *Journal of Applied Statistical Science*, *2*(1), 33–44.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, *36*(2), 192–236.

Bourguignon, M., Ghosh, I., & Cordeiro, G. M. (2016). General results for the transmuted family of distributions and new models. *Journal of Probability & Statistics*, *2016*, 12.

Castillo, E., & Galambos, J. (1987). Bivariate distributions with normal conditionals. *Proceedings of the International Association of Science and Technology for Development*, *2*(1), 59–62.

Chang, L., & Wu, Z. (2011). Performance and reliability of electrical power grids under cascading failures. *International Journal of Electrical Power & Energy Systems*, *33*(8), 1410–1419.

Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, *65*(1), 141–151.

Coleman, B. D. (1958). Statistics and time dependence of mechanical breakdown in fibers. *Journal of Applied Physics*, *29*(6), 968–983.

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, *34*(2), 187–220.

Daniels, H. E. (1945). The statistical theory of the strength of bundles of threads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, *183*(995), 405–435.

David, H. A. (1968). Miscellanea: Gini's mean difference rediscovered. *Biometrika*, *55*(3), 573–575.

David, H. A., & Nagaraja, H. N. (2003). *Order statistics, 3rd Edition.* John Wiley, New York.

Deshpande, J. V., Dewan, I., & Naik-Nimbalkar, U. V. (2010). A family of distributions to model load sharing systems. *Journal of Statistical Planning and Inference*, *140*(6), 1441–1451.

Dewan, I., & Naik-Nimbalkar, U. V. (2010). *Load-Sharing Systems.* Wiley Online Library, Hoboken, NJ, USA.

Dhar, S. K. (1998). Data analysis with discrete analogue of Freund's model. *Journal of Applied Statistical Science*, *7*, 169–183.

Duchateau, L., & Janssen, P. (2007). *The frailty model.* Springer Science & Business Media.

Elbatal, I. (2013). Transmuted modified inverse Weibull distribution: A generalization of the modified inverse Weibull probability distribution. *International Journal of Mathematical Archive*, *4*(8).

Elbatal, I. (2014). Transmuted generalized inverted exponential distribution. *Economic Quality Control*, *28*(2), 128-133.

Elbatal, I., & Aryal, G. (2013). On the transmuted additive Weibull distribution. *Austrian Journal of Statistics*, *42*(2).

Elbatal, I., Asha, G., & Raja, A. V. (2014). Transmuted exponentiated Frêchet distribution: Properties and applications. *Journal of Statistics Applications & Probability*, *3*(3), 379–394.

Elbatal, I., & Elgarhy, M. (2013). Transmuted Quasi Lindley distribution: A generalization of the Quasi Lindley distribution. *International Journal of Pure Applied Science & Technology*, *18*(2), 59-70.

Epema, D. H. J., Livny, M., van Dantzig, R., Evers, X., & Pruyne, J. (1996). A worldwide flock of Condors: Load sharing among workstation clusters. *Future Generation Computer Systems*, *12*(1), 53–65.

ESPNCricinfo. (2013). Match and series archives. *www. cricinfo.com.*

Fan, J., Hsu, L., & Prentice, R. L. (2000). Dependence estimation over a finite bivariate failure time region. *Lifetime Data Analysis*, *6*(4), 343–355.

Fang, K.-T., Kotz, S., & Ng, K. W. (1990). *Symmetric multivariate and related distributions.* Chapman and Hall.

Fine, J. P., Glidden, D. V., & Lee, K. E. (2003). A simple estimator for a shared frailty regression model. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *65*(1), 317–329.

Frêchet, M. (1951). On the correlation tables whose margins are given. *Ann. Univ. Lyon Sect*, *A*(14), 53–77.

Fréchet, M. (1960). On arrays whose margins and bounds are given. *Review of the International Statistical Institute*, 10–32.

Freund, J. E. (1961). A bivariate extension of the exponential distribution. *Journal of the American Statistical Association*, *56*(296), 971–977.

Glidden, D. V. (1999). Checking the adequacy of the gamma frailty model for multivariate failure times. *Biometrika*, 381–393.

Gross, A. J., Clark, V. A., & Liu, V. (1971). Estimation of survival parameters when one of two organs must function for survival. *Biometrics*, 369–377.

Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, *55*(292), 698–707.

Gumbel, E. J. (1961). Bivariate logistic distributions. *Journal of the American Statistical Association*, *56*(294), 335–349.

Hanagal, D. D. (2005). A positive stable frailty regression model in bivariate survival data. *Journal of the Indian Society for Probability and Statistics*, *9*, 35–44.

Hanagal, D. D. (2010). Modeling heterogeneity for bivariate survival data by the Weibull distribution. *Statistical Papers*, *51*(4), 947–958.

Hanagal, D. D. (2011). *Modeling survival data using frailty models.* Chapman & Hall/CRC.

Hanagal, D. D., & Sharma, R. (2011). Analysis of diabetic retinopathy data using shared inverse Gaussian frailty model. *Technical Report: University of Pune, Pune, India.*.

Harlow, D. G. (2002). Applications of the Frêchet distribution function. *International Journal of Materials & Product Technology*, *17*(5), 482-495.

Holland, P., & Wang, Y. (1987). Dependence function for continuous bivariate densities. *Communications in Statistics—Theory & Methods*, *16*(3), 863-876.

Hollander, M., & Peña, E. A. (1995). Dynamic reliability models with conditional proportional hazards. *Lifetime Data Analysis*, *1*(4), 377–401.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, *52*, 105–124.

Hougaard, P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika*, *71*(1), 75–83.

Hougaard, P. (1986). A class of multivanate failure time distributions. *Biometrika*, *73*(3), 671–678.

Hougaard, P. (2000). *Analysis of multivariate survival data.* Springer, New York.

Huster, W. J., Brookmeyer, R., & Self, S. G. (1989). Modelling paired survival data with covariates. *Biometrics*, *45*, 145–156.

Hutchinson, T. P. T. P., & Lai, C. D. (1990). *Continuous bivariate distributions emphasising applications* (Tech. Rep.).

Jagathnath, K. K. M. (2010). *A bivariate Pareto distribution with Freund's dependence structure.* Ph.D Thesis submitted to Cochin University of Science and Technology.

Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis*, *46*(2), 262–282.

Joe, H. (1997). *Multivariate models and multivariate dependence concepts.* CRC Press.

Johnson, N. L. (1949). Bivariate distributions based on simple translation systems. *Biometrika*, *36*(3/4), 297–304.

Johnson, N. L., & Kotz, S. (1975). A vector multivariate hazard rate. *Journal of Multivariate Analysis*, *5*(1), 53–66.

Kalbfleisch, J. D., & Prentice, R. L. (2002). *The statistical analysis of failure time data, second edition.* Hoboken, NJ: John Wiley.

Kenney, J. F. (2013). *Mathematics of Statistics.* D. Van Nostrand Company Inc, Toronto.

Khan, M. S., & King, R. (2013). Transmuted modified Weibull distribution: A generalization of the modified Weibull probability distribution. *European Journal of Pure & Applied Mathematics*, *6*(1).

Kim, H., & Kvam, P. H. (2004). Reliability estimation based on system data with an unknown load share rule. *Lifetime Data Analysis*, *10*(1), 83–94.

Kimeldorf, G., & Sampson, A. R. (1978). Monotone dependence. *The Annals of Statistics*, *6*(4), 895–903.

Klein, J. P., Moeschberger, M., Li, Y., Wang, S., & Flournoy, N. (1992). *Survival Analysis: State of the Art.* Springer.

Komatsu, S., & Sakimoto, T. (1977). Ultimate load carrying capacity of steel arches. *Journal of the Structural Division*, *103*(ASCE 13424 Proceeding).

Kotz, S., Balakrishnan, N., & Johnson, N. L. (2000). Bivariate and trivariate normal distributions. *Continuous Multivariate Distributions: Models and Applications, Volume 1, 2nd Edition*, *1*(2), 251–348.

Kotz, S., & Nadarajah, S. (2000). *Extreme value distributions Vol. (31).* Imperial College Press, London.

Kotz, S., & Nadarajah, S. (2004). *Multivariate t-distributions and their applications.* Cambridge University Press.

Kvam, P. H., & Lu, J.-C. (2007). Load-Sharing Models. *Encyclopedia of Statistics in Quality & Reliability*.

Kvam, P. H., & Pena, E. A. (2005). Estimating load-sharing properties in a dynamic reliability system. *Journal of the American Statistical Association*, *100*(469), 262–272.

Lai, C.-D. (2004). Constructions of continuous bivariate distributions. *Journal of the Indian Society for Probability & Statistics*, *8*, 21–43.

Lai, C.-D. (2006). Constructions of discrete bivariate distributions. *Advances in Distribution Theory, Order Statistics, and Inference*, *1*(1), 29–58.

Lam, K. F., & Kuk, A. Y. C. (1997). A marginal likelihood approach to estimation in frailty models. *Journal of the American Statistical Association*, *92*(439), 985–990.

Lawless, J. F. (2011). *Statistical models and methods for lifetime data* (Vol. 362). John Wiley & Sons.

Lee, H., & Cha, J. H. (2015). On two general classes of discrete bivariate distributions. *The American Statistician*, *69*(3), 221–230.

Lindley, D. V., & Singpurwalla, N. D. (1986). Multivariate distributions for the life lengths of components of a system sharing a common environment. *Journal of Applied Probability*, *23*(2), 418–431.

Lindley, D. V., & Singpurwalla, N. D. (2002). On exchangeable, causal and cascading failures. *Statistical Science*, *17*(2), 209–219.

Liu, H. (1998). Reliability of a load-sharing k-out-of-n: G system: Non-iid components with arbitrary distributions. *IEEE Transactions on Reliability*, *47*(3), 279–284.

Lu, J.-C. (1989). Weibull extensions of the Freund and Marshall-Olkin bivariate exponential models. *IEEE Transactions on Reliability*, *38*(5), 615–619.

Lu, J.-C., & Bhattacharyya, G. K. (1990). Some new constructions of bivariate Weibull models. *Annals of the Institute of Statistical Mathematics*, *42*(3), 543–559.

Lynch, J. D. (1999). On the joint distribution of component failures for monotone load-sharing systems. *Journal of Statistical Planning & Inference*, *78*(1), 13–21.

Ma, Z. S., & Krings, A. W. (2008). Multivariate survival analysis (I): Shared frailty approaches to reliability and dependence modeling. *Aerospace Conference, 2008 IEEE*, *13*(1), 1–21.

Mallick, M., Ravishanker, N., & Kannan, N. (2008). Bivariate positive stable frailty models. *Statistics & Probability Letters*, *78*(15), 2371–2377.

Mantel, N., Bohidar, N. R., & Ciminera, J. L. (1977). Mantel-Haenszel analyses of litter-matched time-to-response data, with modifications for recovery of interlitter information. *Cancer Research*, *37*(11), 3863–3868.

Mardia, K. V. (1970). *Families of bivariate distributions* (Vol. 27). Griffin London.

Marshall, A. W., & Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, *62*(317), 30–44.

Marshall, A. W., & Olkin, I. (1988). Families of multivariate distributions. *Journal of the American Statistical Association*, *83*(403), 834–841.

Martinussen, T., & Pipper, C. B. (2005). Estimation in the positive stable shared frailty Cox proportional hazards model. *Lifetime Data Analysis*, *11*(1), 99–115.

McCool, J. I. (2006). Testing for dependency of failure times in life testing. *Technometrics*, *48*(1).

McGilchrist, C. A. (1993). REML estimation for survival models with frailty. *Biometrics*, 221–225.

McGilchrist, C. A. (1994). Estimation in generalized mixed models. *Journal of the Royal Statistical Society. Series B (Methodological)*, 61–69.

McGilchrist, C. A., & Aisbett, C. W. (1991). Regression with frailty in survival analysis. *Biometrics*, 461–466.

McKenzie, E. D. (1982). Product autoregression: A time series characterization of the gamma distribution. *Journal of Applied Probability*, *19*, 463–468.

Merovci, F. (2013a). Transmuted generalized Rayleigh distribution. *Journal of Statistics Applications & Probability*, *2*(3), 1-12.

Merovci, F. (2013b). Transmuted Lindley distribution. *International Journal of Open Problems in Computer Science and Mathematics*, *6*.

Merovci, F. (2013c). Transmuted Rayleigh distribution. *Austrian Journal of Statistics*, *42*(1), 21 - 31.

Moors, J. J. A. (1988). A quantile alternative for kurtosis. *The statistician*, 25–32.

Mubarak, M. (2012). Parameter estimation based on the Frêchet progressive type II censored data with binomial removals. *International Journal of Quality, Statistics, & Reliability*, *Article ID 245910*.

Murthy, D. N. P., Xie, M., & Jiang, R. (2004). *Weibull models* (Vol. 505). John Wiley & Sons.

Nadarajah, S., & Kotz, S. (2003). The exponentiated Frêchet distribution. *Interstat Electronic Journal*, *9*(12), 1 - 7.

Nadarajah, S., & Kotz, S. (2008). Sociological models based on Frêchet random variables. *Quality & Quantity*, *42*(1), 89-95.

Nelsen, R. B. (2006). An introduction to copulas, 2nd edition. *New York: Springer-Science Business Media*.

Noh, M., Ha, I. D., & Lee, Y. (2006). Dispersion frailty models and HGLMs. *Statistics in medicine*, *25*(8), 1341–1354.

Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, *84*(406), 487–493.

Olkin, I., & Liu, R. (2003). A bivariate beta distribution. *Statistics & Probability Letters*, *62*(4), 407–412.

Patil, G. P. (1965). On a characterization of multivariate distribution by a set of its conditional distributions. *Bulletin of the International Statistical Institute*, *41*(2), 768–769.

Peña, E. A. (2006). Dynamic modeling and statistical analysis of event times. *Statistical Science*, *21*(4), 487–500.

Pulcini, G. (2006). A bivariate distribution for the reliability analysis of failure data in presence of a forewarning or primer event. *Communications in Statistics—Theory & Methods*, *35*(11), 2107–2126.

Ravishanker, N., & Dey, D. K. (2000). Multivariate survival models with a mixture of positive stable frailties. *Methodology & Computing in Applied Probability*, *2*(3), 293–308.

Reliability in Engineering Design. (n.d.).

ReliaSoft. (2003). Using QALT models to analyze system configurations with load sharing. *Reliability Edge Home, www.reliasoft.com*, *3*(3), 1–4.

Rosen, B. W. (1964). Tensile failure of fibrous composites. *The American Institute of Aeronautics and Astronautics Journal*, *2*(11), 1985–1991.

Sahu, S. K., & Dey, D. K. (2000). A comparison of frailty and other models for bivariate survival data. *Lifetime Data Analysis*, *6*(3), 207–228.

Sahu, S. K., Dey, D. K., Aslanidou, H., & Sinha, D. (1997). A Weibull regression model with gamma frailties for multivariate survival data. *Lifetime Data Analysis*, *3*(2), 123–137.

Sarabia, J. M., & Gómez, D. E. (2008). Construction of multivariate distributions: A review of some recent results. *SORT*, *32*(1), 3–36.

Shaked, M. (1977). A family of concepts of dependence for bivariate distributions. *Journal of the American Statistical Association*, *72*(359), 642–650.

Shaked, M. (1984). Extensions of the Freund distribution with applications in reliability theory. *Operations Research*, *32*(4), 917–925.

Shaked, M., & Shanthikumar, J. G. (2015). Multivariate conditional hazard rate functions–An overview. *Applied Stochastic Models in Business & Industry*, *31*(3), 285–296.

Shaked, M., Shanthikumar, J. G., & Valdez-Torres, J. B. (1995). Discrete hazard rate functions. *Computers & Operations Research*, *22*(4), 391–402.

Shaw, W. T., & Buckley, I. R. C. (2009). The alchemy of probability distributions: Beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint arXiv:0901.0434*.

Shih, J. H. (1998). A goodness-of-fit test for association in a bivariate survival model. *Biometrika*, *85*(1), 189–200.

Shuaib, K. M., Robert, K., & Lena, H. I. (2016). Transmuted Kumaraswamy Distribution. *Statistics in Transition New Series*, *17*(2), 183–210.

Singpurwalla, N. D. (1995). Survival in dynamic environments. *Statistical Science*, *10*(1), 86–103.

Singpurwalla, N. D. (2006). *Reliability and Risk: A Bayesian perspective.* John Wiley & Sons.

Spurrier, J. D., & Weier, D. R. (1981). Bivariate survival model derived from a Weibull distribution. *IEEE Transactions on Reliability*, *30*(2), 194–197.

Stefanescu, C., & Turnbull, B. W. (2012). Multivariate frailty models for exchangeable survival data with covariates. *Technometrics*, *48*(3), 411–417.

Stone, G. C. (1978). Statistical analysis of accelerated aging tests on solid electrical insulation. *Unpublished MA Sc. Thesis, University of Waterloo, Waterloo, Ontario, Canada*.

Sutar, S. S., & Naik-Nimbalkar, U. V. (2014). Accelerated failure time models for load sharing systems. *IEEE Transactions on Reliability*, *63*(3), 706–714.

Swift, A. W. (2008). Stochastic models of cascading failures. *Journal of Applied Probability*, *45*(4), 907–921.

Vahid, N., & Kundu, D. (2017). Bivariate discrete generalized exponential distribution. *arXiv:1701.03569v1 [stat.ME]*.

Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, *16*(3), 439–454.

Vernic, R. (1997). On the bivariate generalized Poisson distribution. *Astin Bulletin*, *27*(01), 23–32.

Vernic, R. (2000). A multivariate generalization of the generalized Poisson distribution. *Astin Bulletin*, *30*(01), 57–67.

Wang, S. T., Klein, J. P., & Moeschberger, M. L. (1995). Semi-parametric estimation of covariate effects using the positive stable frailty model. *Applied Stochastic Models & Data Analysis*, *11*(2), 121–133.

Wassell, J. T., Kulczycki, G. W., & Moyer, E. S. (1995). Frailty models of manufacturing effects. *Lifetime Data Analysis*, *1*(2), 161–170.

Wassell, J. T., Wojciechowski, W. C., & Landen, D. D. (1999). Recurrent injury event-time analysis. *Statistics in Medicine*, *18*(23), 3355–3363.

Wheelock, D. C., & Wilson, P. W. (2000). Why do banks disappear? The determinants of US bank failures and acquisitions. *Review of Economics & Statistics*, *82*(1), 127–138.

Xue, X., & Brookmeyer, R. (1996). Bivariate frailty model for the analysis of multivariate survival time. *Lifetime Data Analysis*, *2*(3), 277–289.

Yau, K. K. W., & McGilchrist, C. A. (1998). ML and REML estimation in survival analysis with time dependent correlated frailty. *Statistics in Medicine*, *17*(11), 1201–1213.

Zaharim, A., Najid, S. K., Razali, A. M., & Sopian, K. (2009). Analyzing Malaysian wind speed data using statistical distribution. In *Proceedings of the 4th IASME/WSEAS International Conference on Energy & Environment, Cambridge, UK.*