# An Application of Branching Process In Multi Level Marketing (MLM)

**Book** · August 2013



**1 author:**

Md. Murad Hossain
Bangabandhu Sheikh Mujibur Rahman Science & Technology University

Multi Level Marketing (MLM) is a Business opportunity that is known to all over the world. MLM Company plays an important role directly or indirectly in our economy. Performance of MLM companies can be explained by stochastic Modeling. So to construct a complete picture of MLM Companies activities in terms of their earning pattern and distributor enrollment in their business we have to explore stochastic modeling. For this perspective this book concerned with the Branching Process to see the distribution of down line distributors and the probability of ultimate extinction. This book also linked up with the Markov Chain Estimation for establishing the profit of MLM Company. In this book try to show a comparison between Markov Chain Estimation using Monte Carlo simulation and Markov Chain Estimation using MCMC simulation. Finally make a conclusion about the taking decision of a distributor or a new investor for the investment of capital into a right company.

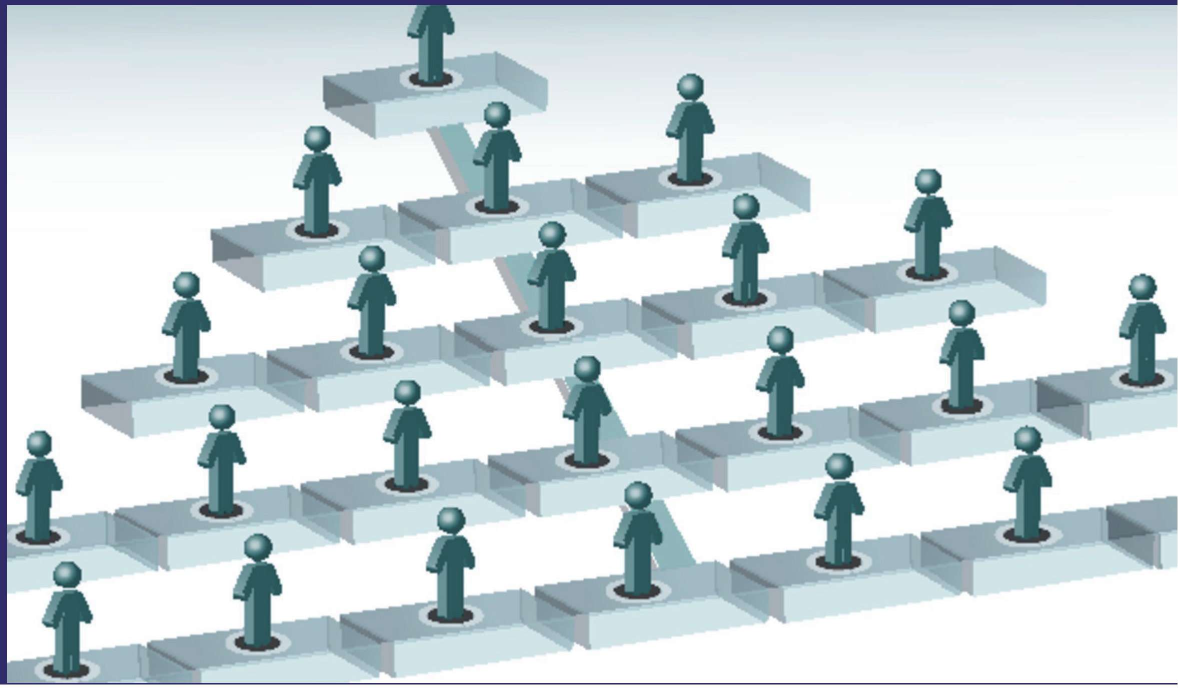
**Md. Murad Hossain**

Md. Murad Hossain, Born in 29 December,1988 in Faridpur, Completed B.Sc and M.S in Statistics from Jahangirnagar University of Bangladesh, Having experiences in research fields. Like to think positively to gather knowledge and interested in research work. Now, working as a GTA in Applied Statistics department of East West University in Bangladesh.

978-3-659-44547-7

**Use of Branching process in MLM**

**Hossain, Rois**



Md. Murad Hossain
Rumana Rois

# An Application Of Branching Process in Multi Level Marketing(MLM)

## Branching process in MLM and comparison with other stochastic Model

LAP LAMBERT Academic Publishing

**Md. Murad Hossain**
**Rumana Rois**

# An Application Of Branching Process in Multi Level Marketing(MLM)

Md. Murad Hossain
Rumana Rois

# An Application Of Branching Process in Multi Level Marketing(MLM)

## Branching process in MLM and comparison with other stochastic Model

LAP LAMBERT Academic Publishing

**Impressum / Imprint**
Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.


Bibliographic information published by the Deutsche Nationalbibliothek: The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.


Coverbild / Cover image: www.ingimage.com

Verlag / Publisher:
LAP LAMBERT Academic Publishing
ist ein Imprint der / is a trademark of
AV Akademikerverlag GmbH & Co. KG
Heinrich-Böcking-Str. 6-8, 66121 Saarbrücken, Deutschland / Germany
Email: info@lap-publishing.com

Herstellung: siehe letzte Seite /
Printed at: see last page
**ISBN: 978-3-659-44547-7**

Zugl. / Approved by: Jahangirnagar university, Savar, Dhaka-1342,Thesis. 2011.

Dedicated
To
my beloved parents

# *Acknowledgement*

From my innermost heart I wish to take the novel opportunity to express all praises and gratefulness to the almighty "GOD" who enables me to complete the work and manages everything properly.

I wish to offer my best regards and profound gratitude and indebtedness to my reverend teacher Rumana Rois, Assistant Professor, Department of Statistics, Jahangirnagar University, Savar, Dhaka for selecting the topic and his indispensable guidance, Keen interest, unparalleled stimulating, continuous encouragement and invaluable suggestions given to me throughout the tenure of my book and for the extension of his helping hand as well as at the time of this dissertation.

I would like to pay my respect and gratefulness to all of my honorable teachers of the Department of Statistics, Jahangirnagar University, for their encouragement and suggestions during my studies in this University.

I offer my cordial thanks to my friends Mintu, Harun, Shohel, Rana, Habibullah Mukta, Sumi, Senjuti, Suly, Eva, Nahid for their warmful and constant helps in my research.

## Contents


List of Tables iv-v

List of Figures vi-vii

Abstract viii

**Chapter 1 Introduction** 1-4

1.1 Background of the study 1-2

1.2 Objectives of the study 2

1.3 Data description 3

1.4 Outline of the Research 3

1.5 Computational issue 4

1.6 Limitation 4

**Chapter 2 Literature review** 5-35

2.1 Introduction 5

2.2 Illustration of MLM Companies 6-7

2.3 Review of Stochastic Process 7

2.3.1 Parameters of Stochastic Process 8

2.3.2 Classification of Stochastic Process 8-10

2.3.2.1 Stochastic Process with Discrete State and Time Space 8-9

2.3.2.2 Stochastic Process with Discrete State and Continuous time Space 9

2.3.2.3 Stochastic Process with Continuous State Space and Discrete Time Space 9-10

2.3.2.4 Stochastic Process with Continuous State and Time Space 10

2.4 Different Stochastic Models 10-21

2.4.1 Markov Process 10-11

2.4.1.1 Discrete space Markov Process 11-12

2.4.1.2 Markov Process with Continuous State Space 12

2.4.2 Markov Chain 12-13

2.4.3 Semi Markov Process 13

2.4.4 Renewal Process 13

| | | |
|---|---|---|
| | 2.4.5 Branching process | 13-21 |
| | 2.4.5.1 Mathematical formulation of Branching process | 14-15 |
| | 2.4.5.2 Branching model | 15-16 |
| | 2.4.5.3 Assumptions of Branching process | 16 |
| | 2.4.5.4 Application of Branching process | 16 |
| | 2.4.5.5 The Generating function of Branching process | 17-18 |
| | 2.4.5.6 Mean and Variance of the branching process | 18 |
| | 2.4.5.7 Moment generating function of the branching process | 18 |
| | 2.4.5.8 Determining the Probability of Extinction of the Population | 19 |
| | 2.4.6 Galton-Watson Branching process | 20-21 |
| | 2.4.6.1 Mathematical definition of Galton-Watson Branching process | 21 |
| | 2.4.6.2 Extinction criterion for Galton–Watson process | 21 |
| 2.5 | Inference of Stochastic Process | 22-29 |
| | 2.5.1 Estimation of Stochastic Models | 22 |
| | 2.5.1.1 Estimation for transition probability | 22-25 |
| | 2.5.2 Hypothesis testing | 25-29 |
| | 2.5.2.1 Testing transition probability matrix | 25-26 |
| | 2.5.2.2 Likelihood ratio test | 26 |
| | 2.5.2.3 Test for the Stationary of the transition probability matrix | 26-27 |
| | 2.5.2.4 Test for the order of Markov chain | 27-28 |
| | 2.5.2.5 Test for the first order Markov dependence | 28-29 |
| 2.6 | Model selection criteria | 29-33 |
| | 2.6.1 The $R^2$ Criterion | 30 |
| | 2.6.2 Adjusted: $R^2$ | 30 |
| | 2.6.3 Akaike Information Criterion (AIC) | 30-31 |
| | 2.6.4 Schwarz Information Criterion (SIC) | 31 |
| | 2.6.5 Mallows's $\boldsymbol{C_p}$ Criterion | 31-32 |

| | | |
|---|---|---|
| | 2.6.6 Forecast Chi-Square $\chi^2$ | 32 |
| | 2.6.7 Minimum Description Length (MDL) principle | 32-33 |
| 2.7 | Metropolis-Hasting Algorithm | 33 |
| 2.8 | Kolmogorov-Smirnov Goodness-of-Fit Test | 34 |
| 2.9 | Conclusion | 35 |
| **Chapter 3** | **Result and analysis** | 36-76 |
| 3.1 | Introduction | 36 |
| 3.2 | Univariate Analysis | 36-46 |
| | 3.2.1 Data consideration | 37-41 |
| | 3.2.2 Graphical Representations | 42-46 |
| 3.3 | Illustration of Branching Process | 46-54 |
| | 3.3.1 Data structure | 48 |
| | 3.3.2 Down line clients (Offspring) Distribution | 49-50 |
| | 3.3.3 Different Generations Structure of down line Clients | 51-55 |
| | 3.3.4 Ultimate Extinction | 56-57 |
| 3.4 | Estimating Markov Chain with Generated Profit | 57-76 |
| | 3.4.1 Markov Chain Estimation using Monte Carlo Simulation | 57-65 |
| | 3.4.2 Markov Chain Estimation using MCMC Simulation (Metropolish Hasting Algorithm) | 65-73 |
| | 3.4.3 Long Run Probabilities | 73-75 |
| | 3.4.4 Kolmogorov Smirnov goodness of fit test | 75-76 |
| 3.5 | Conclusion | 76 |
| **Chapter 4** | **Summary and Conclusion** | 77-79 |
| **References** | | 80-82 |
| **Appendix** | **Program 1** | 83-84 |
| | **Program 2** | 84-85 |
| | **Program 3** | 85-86 |
| | **Program 4** | 87-88 |
| | **Program 5** | 88-89 |
| | **Program 6** | 89 |

# List of the Tables


| Table no | Title of the Table | Page no |
|---|---|---|
| 2.1 | Brief Information of MLM companies in Bangladesh | 6 |
| 3.1 | Original annual net profit of MLM Company D2L | 37 |
| 3.2 | Original annual net profit of MLM Company M.net | 37 |
| 3.3 | Descriptive statistics for original Profit variable of MLM Company D2L | 37 |
| 3.4 | Descriptive statistics for simulated Profit variable using Markov Chain (MC) of MLM Company D2L | 38 |
| 3.5 | Descriptive statistics for simulated Profit variable using MCMC (Metropolish-Hasting algorithm) of MLM Company D2L | 39 |
| 3.6 | Descriptive statistics for original Profit variable of MLM Company M.net | 39 |
| 3.7 | Descriptive statistics for Simulated Profit variable using Markov Chain (MC) of MLM Company M.net | 40 |
| 3.8 | Descriptive statistics for Simulated Profit by MCMC (Metropolish-Hasting algorithm) of MLM Company M.net | 41 |
| 3.9 | Distribution of offspring of the surveyed individual of MLM Company D2L. | 47 |
| 3.10 | Distribution of offspring of the surveyed individual of MLM Company M.net. | 47 |
| 3.11 | Mean and variance of Down Line Clients of MLM company D2L for Different Generations | 55 |
| 3.12 | Mean and variance of Down Line Clients of MLM company M.net for Different Generations | 55 |
| 3.13 | Classification of States in the profit of D2L for Markov Chain model during year2007-2008 to 2010-2011 | 58 |
| 3.14 | Classification of States in the profit of M.net for Markov Chain model during year 2006 to 2010 | 58 |

| | | |
|---|---|---|
| 3.15 | Classification of States in the profit of D2L for Markov Chain Estimation using MCMC Simulation during year2007-2008 to 2010-2011 Markov Chain Estimation using MCMC Simulation | 66 |
| 3.16 | Classification of States in the profit of M.net for Markov Chain Estimation using MCMC Simulation during year 2006 to 2010 Markov Chain Estimation using MCMC Simulation | 66 |
| 3.17 | Long run probabilities by Markov Chain Estimation using Monte Carlo (MC) Simulation for MLM Company D2L | 73 |
| 3.18 | Long run probabilities by Markov Chain Estimation using Monte Carlo Simulation (MC) for MLM Company M.net | 74 |
| 3.19 | Long run probabilities by Markov Chain Estimation using MCMC Simulation (Metropolish Hasting Algorithm) for MLM Company D2L | 74 |
| 3.20 | Long run probabilities by Markov Chain Estimation using MCMC Simulation (Metropolish Hasting Algorithm) for MLM Company D2L for MLM Company M.net | 75 |
| 3.21 | Two-sample Kolmogorov-Smirnov test for generated profit of D2L and M.net using Metropolis Hasting algorithm | 76 |

## List of Figures


| Figure no | Figure titles | Page no |
|---|---|---|
| Figure 2.1 | Stochastic Process with Discrete State & Time Space | 8 |
| Figure 2.2 | Stochastic Process with Discrete State and Time Space | 9 |
| Figure 2.3 | Stochastic Process with Continuous State Space and Discrete Time Space | 9 |
| Figure 2.4 | Stochastic Process with Continuous State and Time Space | 10 |
| Figure 2.5 | Normal random numbers to the K-S test | 34 |
| Figure 3.1 | Histogram of simulated Profit variable using Markov Chain (MC)of M LM company D2L | 42 |
| Figure 3.2 | Histogram for simulated profit variable by MCMC (Metropolish-Hasting algorithm) of MLM company D2L | 42 |
| Figure 3.3 | Histogram for Simulated Profit variable by Markov Chain (MC) of MLM Company M.net | 43 |
| Figure 3.4 | Histogram for Simulated Profit variable by MCMC (Metropolish-Hasting algorithm) of MLM Company M.net | 43 |
| Figure 3.5 | Box-plot for simulated profit variable by Markov Chain (MC) for MLM Company D2L | 44 |
| Figure 3.6 | Box-plot for simulated profit variable by MCMC (Metropolish-Hasting algorithm) for MLM Company D2L | 44 |
| Figure 3.7 | Box-plot for simulated profit variable by Markov Chain (MC) of MLM Company M.net | 45 |
| Figure 3.8 | Box-plot for simulated profit variable by MCMC (Metropolish-Hasting algorithm) of MLM Company M.net | 45 |
| Figure 3.9 | structure of the data of MLM companies | 46 |
| Figure 3.10 | Tree diagram of the surveyed individual of MLM Company D2L with their corresponding probabilities (up to2$^{nd}$ generation). | 49 |
| Figure 3.11 | Tree diagram of the surveyed individual of MLM Company | 50-51 |

| | | |
|---|---|---|
| | D2L with their corresponding probabilities (up to 3rd generation). | |
| Figure 3.12 | Tree diagram of the surveyed individual of MLM Company M.net with their corresponding probabilities (up to2nd generation). | 52 |
| Figure 3.13 | Tree diagram of the surveyed individual of MLM Company M.net with their corresponding probabilities (up to 3rd generation). | 53-54 |

# Abstract

This book considers Branching process as well as Markov Chain models to construct a complete picture of Multi-Level Marketing (MLM) companies' activities in terms of their earning pattern and distributors' enrollment in their business. MLM companies are playing much unexplained role in our economy, so exploring their future performance will directly and indirectly play a significant role in the development of the economy of Bangladesh. As a developing country, the risk of investment to any unknown business is very high for any people of Bangladesh. However, this risk becomes dangerous when the investment associate with the savings of poor people. One of the most important factors in modeling of MLM companies' activities is to forecast the future profit and number of Clients involvement to know the long run performance of the Company. This information can be extract by stochastic modeling. We illustrate our theoretical knowledge of stochastic modeling in two MLM Companies data: the Destiny-2000 and Medsit.net.

In this book, branching process is introduced and evaluated to get idea of the distribution of down line distributors. This model is also used in characterizing the distribution of the size of the population for different generations and the probability of extinction of the population.

The Markov Chain model is also established for the annual profit of two companies. We found only 4 years annual profit for destiny-2000 and 5 years for Medsit.net. This information is not adequate for estimating Markov chain. So simulation is carried out in two phase: one is Markov Chain Estimation using Monte Carlo Simulation and other is Markov Chain Estimation using MCMC Simulation (Metropolish Hasting Algorithm).

This book is organized into four chapters. Chapter-1 Introduction, Chapter -2 Literature Review, Chapter-3 Result and Analysis, Chapter-4 Summary and Conclusion and finally References.

# <u>Chapter : 1</u>

# Introduction

### 1.1 Background of the study

The history of the study of Branching process dates back to 1874, when a mathematical model was formulated by Galton and Watson for the problem "extinction of families". The model did not attract much attention for a long time; the situation gradually changed and during the last 60 years much attention has been devoted to it. This is because of the development of interest in the applications of probability theory, in general and also because of the possibility of using the models in a variety of biological, physical and other problems which reproduce similar objects by biological methods or may be physical particles. (See, Medhi,(2009)).

In probability theory, a Branching process is a Markov process that models a population in which each individual in generation n produces some random number of individuals in generation n+1, according to a fixed probability distribution that does not vary from individual to individual. Branching process are used to model reproduction; for example, the individuals might correspond to bacteria, each of which generates 0, 1 or 2 offspring with some probability in a single time unit. Branching process can also be used to model other systems with similar dynamics, e.g. the spread of surnames in genealogy or the propagation of neutrons in a nuclear reactor. Branching process have a wide variety of application in Biological science, Sociological science and Engineering science.

Multi Level Marketing (MLM) is embracing more and more arenas today. MLM route provides employment opportunities to lakhs of people and enhances their social status. The MLM members also get tremendous opportunity to develop themselves personally. This multiple role of MLM companies can be looked at as a social contribution and these companies or cooperatives are emerging as a development oriented social movement. Multi-Level Marketing (MLM) is a business opportunity that goes by many names like - Network marketing, direct selling, person-to person marketing, matrix marketing, or one to one marketing. It is a system whereby the company producing or arranging the product or service and rewards the people who make or ads word of mouth referrals. The people who operate as components in the chain are referred to by any of the following names - Members, Independent Salespeople, Network Marketers, Advisors, Agents, and Distributors. Nobody does not know when started the World famous MLM business. Some are saying it started before the 2nd world war. Nowadays more than 135 countries and more than 12000 company operation their business for selling products and services in MLM system. Bangladesh started business through MLM system in 1998. At present there are 20 to 25 companies and they are operating their business through MLM system. However, this system entered in Bangladesh at 1998 through GGN (Global Guardian Network) but now the Destiny-2000 Limited is one of the largest companies among them. MLM is a recent activity and trend added to the business by time. MLM not only operates as direct marketing channel but also it is being practiced by Multi-Level Marketing, Network Marketing, and Direct Selling etc.

The business of MLM increases as their members increase. So the performance of the MLM companies can be observed by their member's involvement. Nevertheless, profit of the company is one of the prime indicators of company's performance. The number of members of the company is varying with time. As a result, to explore the future performance of MLM companies in Bangladesh we need to use stochastic models in our analysis. Stochastic models deal with the discrete or continuous state variable(s) which are indexed by time. As the member of MLM companies are reproduced another member, so the Branching Process, a stochastic process, will be the most appropriate model at this point of view.

This thesis concerned with "On the use of Branching process in Multi Level Marketing (MLM) in the context of Bangladesh". So, we considered two renowned MLM companies such as MLM Company Destiny-2000 Ltd and Medsit.net

**Objectives of the study**

To estimate a Branching Process model to explore the future performance of Multi-Level Marketing system in Bangladesh.

To find the probability of extinction.

To check the properties of Branching Process in the context of Multi-level Marketing.

To compare the fitted Branching Process with the discrete-time Markov chain.

**Data description**

We have collected data from two MLM companies. The two MLM companies are Destiny 2000 Ltd and Medsit.net. In the MLM sector a new distributor (member) pays a registration fee and joins the business either as a direct associate of the company, or under an existing distributor called a sponsor. The sponsor does not sponsor the new distributor with money, but rather with knowledge and advice about the business. Both the sponsor and the new distributor then go on to each sponsor more new distributors and so on. In this way a network of distributors of the products of the company is built. We have taken two individuals from two MLM companies whose has so many offspring to their down line. For the purpose of calculating long run probabilities, ultimate extinction we take these individual as a unit.

We have also collected last 4 to 5 years profit related data from both MLM companies. To explore the Markov Process we use this profit related data. This types of data helps us to find future structure of profit of two MLM companies.

For data collection, we have used non-probability sampling method as well as convenience sampling method.

**Outline of the research**

This research is organized into four chapters. Chapter-1 includes background of the study, objectives of the research and concept on data consideration. Chapter -2 reviews the literature of illustration of MLM companies, reviews of stochastic process, classification of stochastic process, different stochastic models such Markov process, Markov chain, Semi Markov process, Renewal process, Branching Process, Galton-Watson branching Process , Mean and variance of branching process, probability extinction criteria for Branching process. Inference of stochastic process and concept of different model selection criteria.Chapter-3 explores all the results and analysis. This chapter includes offspring distribution, tree diagram of offspring distribution, Estimating Markov Chain with Generated Profit, Kolmogorov smirnov goodness of fit test. A summary with some Concluding remarks and some idea for further research is contained in the final chapter.

## Computational issue

For arranging the data we have used Microsoft Excel and then all the models and limiting probabilities to generate results calculated by R-program. (R-version 2.14.2)

## Limitation

When we start our research work, there was no inconvenient situation for MLM companies in Bangladesh. But after some day's corruption are marked to MLM companies by Bangladesh Bank. Bangladesh bank regulate strict rule to some MLM companies such as Destiny 2000 Ltd. It banned all kinds of activities of this company. For this restriction all MLM companies did not agree to give data officially. Instead of this limitation, to complete my research I have to collect data un-officially from MLM companies. If there is no restriction on getting sufficient data more analysis and prediction would be continued.

# Chapter : 2

# Literature review

### 2.1 Introduction

Multi-Level Marketing (MLM) is a business opportunity that goes by many names like - Network marketing, direct selling, person-to person marketing, matrix marketing, or one to one marketing. Bangladesh started business through MLM system in 1998 (mlmbd.blogspot.com).

The business of MLM increases as their members increase. So the performance of the MLM companies can be observed by their member's involvement. Nevertheless, profit of the company is one of the prime indicators of company's performance. The number of members of the company is varying with time. As a result, to explore the future performance of MLM companies in Bangladesh we need to use stochastic models in our analysis. Stochastic models deal with the discrete or continuous state variable(s) which are indexed by time. As the member of MLM companies are reproduced another member, so the Branching Process, a stochastic process, will be the most appropriate model at this point of view.

Branching processes are used to model reproduction. A central question in the theory of branching processes is the probability of ultimate extinction, where no individuals exist after some finite number of generations. By this probability of ultimate extinction in Branching Process we can estimate the future fortune of MLM companies in Bangladesh. In this report, we are interested to illustrate the Branching Process to evaluate the long run performance of MLM companies in Bangladesh.

The organization of this chapter is as follows: in section 2.1 introduction of this chapter has described, we illustrate some MLM companies in section 2.2, in section 2.3 we review the stochastic process. Section 2.4 represents different stochastic models, section 2.5 represents inference of stochastic process, section 2.6 explains model selection criteria, section 2.7 includes Metropolis-Hasting Algorithm, section 2.8 represents Kolmogorov-Smirnov Goodness-of-Fit test finally section 2.9 makes some concluding remarks.

## 2.2 Illustration of MLM Companies

Network marketing often called Multi Level Marketing is perhaps one of the least understood methods of moving products or services into the market place. It is a system whereby the company producing or arranging the product or service and rewards the people who make or ads word of mouth referrals. The people who operate as components in the chain are referred to by any of the following names - Members, Independent Salespeople, Network Marketers, Advisors, Agents, and Distributors. Nobody does not know when started the World famous MLM business. Some are saying it started before the $2^{nd}$ world war. Nowadays more than 135 countries and more than 12000 company operation their business for selling products and services in MLM system (mlmbd.blogspot.com).

**Table 2.1:** Brief Information of MLM companies in Bangladesh

| Company Name | Starting time | Types of business | Product Information |
|---|---|---|---|
| Destiny 2000 Ltd | 2000 | Product Basis | Treeplantation package, share package, kalozira, Tv, Refrigerator etc. |
| Unipay2u Ltd | 2004 | Investment site | Selling investment packages |
| Neway-bd Ltd | 2000 | Product Basis | STL, Toshiba Rangs, LaPana, Guardian, Das etc |
| Bravo Ltd | 2006 | Earning & learning | IT. related training courses |
| Bangladesh.dxn2u Ltd | 1993 | Product Basis | Different products |
| Herbal life Pvt Ltd | 1980 | Product Basis | Protein shakes, protein snacks, nutrition, energy and fitness supplements and personal care products |
| Medsit.net Ltd | 2008 | Earning & learning | IT. related training courses |
| E-Links International Pvt. Ltd. | 2003 | Product Basis | Different electronics products |

Bangladesh started business through MLM system in 1998. At present there are 20 to 25 companies and they are operating their business through MLM system. However, this system entered in Bangladesh at 1998 through GGN (Global Guardian Network) but now the Destiny-2000 Limited is one of the largest companies among them. MLM is a recent activity and trend added to the business by time. MLM not only operates as direct marketing channel but also it is being practiced by Multi-Level Marketing, Network Marketing and Direct Selling etc. The business of MLM increases as their members increase. So the performance of the MLM companies can be observed by their member's involvement. Nevertheless, profit of the company is one of the prime indicators of company's performance. The number of members of the company is varying with time.

### 2.3 Review of Stochastic Process

Before share the concept of stochastic process we need to know the concept of random variable. Any variable of a sample space of a random experiment is called random variable. For instance, the sample space of the experiment for tossing two fair coins is $\Omega = \{HH, HT, TH, TT\} = \{w_{1,}w_2, w_3, w_4\}$. Then each outcomes of the sample space $\Omega$ of the experiment indicates as a random variable separately. On the basis of sample space, random variable is generally of two types, discrete random variable and continuous random variable. (See, Roy. (2000))

A stochastic process is a family of random variables that describe the evolution through time of a process. It is denoted by$\{X(t), t \in T\}$ $\{Xn; n \geq 0\}$ $or$ $\{X_n\}_{n=0}^{\infty}$. Where $t$ is often termed as time and $X(t)$, as state of the process. For instance $X(t)$, be a random variable representing the number of trucks passing Paturia ferry terminal during a particular time $t$ then $\{X(t), t \in T\}$ is a family of random variables and hence is a stochastic process. Again consider a simple experiment like throwing a true die. Suppose that $X_n$ is the outcome of the $n^{th}$ throw $n \geq 1$. Then$\{X_n; n \geq 0\}$ is a family of random variables such that for a distinct value of $n = (1,2, \dots)$. One gets a distinct random variable $X_n$; $\{X_n; n \geq 1\}$ $X_n$; constitutes a stochastic process. Even if we consider a random event occurring in time, such as, number of telephone calls received at a switch board. Suppose that $X(t)$ is the random variable which represents the number of incoming calls in an interval $(0, t)$ of duration $t$ units. The number of calls within a fixed interval of specified duration, say, one unit of time, is a random variable $X(1)$

and the family $\{X(t); t \in T\}$ constitutes a stochastic process $T = [(0, \infty)]$. (See, Roy. (2000))

### 2.3.1 Parameters of Stochastic Process

In order to Specify Stochastic Process we need to illustrate parameter space and state space. Where parameters of stochastic process includes parameter space and state space. Parameter space contains the information regarding time and state space contains state. In a stochastic process $\{X(t); t \in T\}, T$ is defined as index set or parameter space and state space is defined as the set of all possible values that the random variable $X(t)$ can take. State space may be discrete or continuous. In a stochastic process $\{X(t); t \in T\}$ , if the index set $T$ is countable then the process is known as discrete state space. For example, if $X$ is the total number of sixes appearing in the first $n$ throws of a die, the set of possible values of $X_n$ is the finite set of non-negative integers $0,1,2, \dots, n$.. Here, the state space of $X_n$ is discrete. Or number of incoming call in a telephone exchange in$(0, t)$ is a discrete state space. The state space of a stochastic process $\{X(t); t \in T\}$ is called continuous state space if the set of possible values of $X(t)$ is in a interval on a real line$(0, \infty)$. Suppose that $X(t)$represents the maximum temperature at a particular place in $(0, t)$, then set of possible values of $X(t)$ is continuous (see, J.Medhi.(2009)).

### 2.3.2 Classification of Stochastic Process

Types of stochastic process depending on state space and time or parameter are as follows-

Stochastic Process with Discrete State and Time Space.

Stochastic Process with Discrete State Space and Continuous Time Space.

Stochastic Process with Continuous State Space and Discrete Time Space.

Stochastic Process with Continuous State and Time Space.

#### 2.3.2.1 Stochastic Process with Discrete State and Time Space

In stochastic process $\{X(t); t \in T\}$ if the state space $X(t)$ is discrete and index set $T$ is discrete then the process is known as discrete state space and discrete time space. (See, J.Medhi. (2009)).

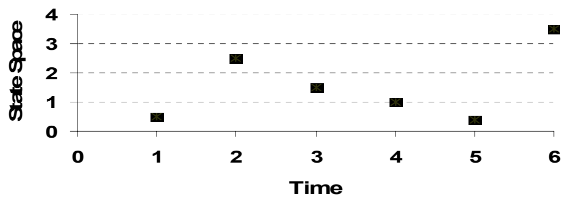


*__Figure: 2.1__ Stochastic Processes with Discrete State & Time Space*

Example: Number of customers arrived in the supermarket per hour

**2.3.2.2 Stochastic Process with Discrete State and Continuous time Space**

In stochastic process $\{X(t); t \in T\}$ if the index set $T$ be an interval on a real line $T = (0, t)$, and the state space $X(t)$ is discrete then the process is known as stochastic process with discrete state space & continuous time space. (See, J.Medhi.(2009)).

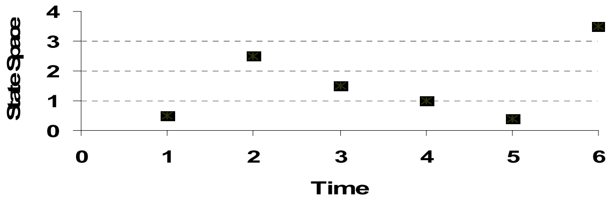



*__Figure: 2.2__ Stochastic Process with Discrete State and Time Space*

Example: the waiting time of the customers before receiving the service per hour

**2.3.2.3 Stochastic Process with Continuous State Space and Discrete Time Space**

In a stochastic process $\{X(t); t \in T\}$ , if the set of possible values of $X(t)$ is an interval on a real line $(0, \infty)$ and the index set, $T$ is discrete then the process is known as continuous state space and discrete time space. (See, J.Medhi. (2009) )

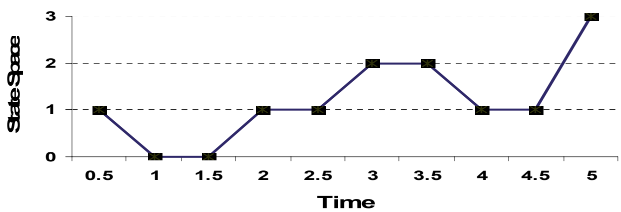


*Figure: 2.3 Stochastic Process with Continuous State Space and Discrete Time Space*

Example: number of incoming calls at a switchboard at time t.

**Stochastic Process with Continuous State and Time Space**

In a stochastic process $\{X(t); t \in T\}$ , if the set of possible values of $X(t)$ is continuous and index set $T$ be an interval on a real line $T(0, t)$ then the process is known as continuous state & time space. (See, J.Medhi. (2009)

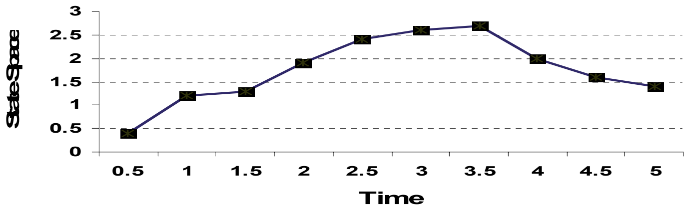



*Figure: 2.4 Stochastic Process with Continuous State and Time Space*

Example: Maximum temperature at a particular place in an interval (0, t)

## 2.4 Different Stochastic Models:

For completing research interest we need to know about some stochastic models, which are related with the objectives of analysis. At the following pages some stochastic models are defined.

### 2.4.1 Markov Process:

In probability theory and statistics, a Markov process, named after the Russian mathematician Andrey Markov, is a time-varying random phenomenon for which a specific property (the Markov property) holds. In a common description, a stochastic process with the Markov property, or memorylessness, is one for which conditional on the present state of the system, its future and past are independent.

Markov processes arise in probability and statistics in one of two ways. A stochastic process, defined via a separate argument, may be shown (mathematically) to have Markov property and as a consequence to have the properties that can be deduced from this for all Markov processes. Of more practical importance is the use of the assumption that the Markov property holds for a certain random process in order to construct.

In modeling terms, assuming that the Markov property holds is one of a limited number of simple ways of introducing statistical dependence into a model for a stochastic process in such a way that allows the strength of dependence at different lags to decline as the lag increases. Often, the term Markov chain is used to mean a Markov process which has a discrete ( finite or countable) state space. Usually a Markov chain would be defined for a discrete set of times (i.e. a discrete-time Markov Chain) ( See, Everitt, 2002 ) although some authors use the same terminology where "time" can take continuous values. Also see continuous time Markov process. ( See, Dodge ).

If $\{X(t), t \in T\}$ is a stochastic process such that, given the value$X(s)$, the values of $(t), t > s$, do not depend on the values of $X(u), u < s$, then the process is said to be a Markov process. It can be written as, if for $t_1 < t_2 < \cdots < t_n < t$

$$\Pr\left\{a \leq X(t) \leq b \,\middle|\, X(t_1) = x_1, \ldots, X(t_n) = x_n\right\}$$

$$= \Pr\{\, a \leq X(t) \leq b \,|\, X(t_n) = x_n \qquad (2.4.1)$$

Then the process $\{X(t), t \in T\}$ is a Markov process. A discrete parameter Markov process is known as a Markov chain. (See, Ross. (2000)).

There are two types of Markov process based on state space such as

Discrete state space Markov process

Continuous state space Markov process

**2.4.1.1 Discrete state space Markov process**

Let X(t) be a continuous parameter Markov process with state space N = {0,1,2,….i,…j,….}. Further let X(t) be time homogeneous then the probability of a transition from state i to state j during the time interval from epoch T to epoch T + t does not depend on the initial time T but depends only on the elapsed time t and on the initial and terminal states i to j. W can thus write

$$\Pr\{X(T+t)=j|X(T)=i\} = P_{ij}(t), \quad i,j = 0,1,2,\dots\dots, t \ge 0. \tag{2.4.2}$$

Where $0 \le P_{ij}(t) \le 1$ for each i,j,t and $\sum_j P_{ij}(t) = 1$.

Some examples of discrete state space Markov processes are

Branching Process
Galton-Watson Process
Birth and Death Process ( Continuous parameter or time space Markov process)
Quasi Birth and Death Process

**2.4.1.2 Markov Process with Continuous State Space**

Poisson process is a process in continuous time with a discrete state space. Here in a small interval of time $\Delta t$, there is either no change of state or there is only one change, the probability of more than one change being of the order of $\Delta t$. In such a process, changes of state occur continually all the time and the state space is continuous.

Some examples of Markov Process with Continuous State Space are

Brownian Motion Process
Wiener Process
Kolmogorov Equation Process
Ornstein-Uhenbeck Process

**2.4.2 Markov Chain:**

The stochastic process$\{x_n, n \ge 0\}$ is called a Markov chain, if for $j, k, j_1, \dots\ j_{n-1} \in N\ |I$

$$\Pr\left[X_n = k \mid X_{n-1} = j, X_{n-2} = j_1 \dots\ X_0 = j_{n-1}\right]$$
$$= \Pr\left[X_n = k \mid X_{n-1} = j\right] = P_{jk} \tag{2.4.3}$$

Whenever the first member is defined .The outcomes are called the states of the Markov Chain; if $X_n$ has the outcome $j(i.e,\ X_n = \text{j})$ the process is said to be at state $j$ at $n^{th}$ trial. $P_{jk}$ denote the transition probability. (See, J.Medhi. (2009)).

A continuous time stochastic process $\{X(t); t \ge 0\}$ is said to be continuous time Markov chain if the conditional probability of the failure given that the present $X(s)$ and the past $X(u), 0 \le u < s$ , depends only on the present and independent of the past. That is, the stochastic process $\{X(t); t \geq 0\}$ is a continuous-time Markov chain if for all $s, t \geq 0$ and non-negative integers $i, j, X(u); 0 \leq u < s$

$$\Pr\left\{X(t+s) = j \middle| X(s) = i, X(u) = x(u); 0 \leq u < s\right\} = \Pr\left\{X(t+s) = j \middle| X(s) = i\right\} = P_{ij}(t) \quad (2.4.4)$$

In other words, a continuous time Markov chain is a stochastic process than moves from state to state in accordance with a Markov chain but in such a way that the amount of time it spends in each state before moving to the next state is an exponential distribution. (See, Ross (2000)).

**2.4.3 Semi Markov process:**

The amount of time that process spends in each state before making a transition is random with the proportion of time that the process is in state i, is given by an equation

$$P_j = \frac{\pi_i \mu_i}{\sum_{j=1}^{N} \pi_j \mu_j}, \quad i = 1,2, \dots, N \quad (2.4.5)$$

Such a process is called semi Markov process. If the amount of time that the process spends in each state before making a transition is identically 1 , then the semi Markov process is called Markov chain.(see, Ross(2000)).

**2.4.4 Renewal Process:**

A counting process for which the inter arrival times are independently distributed and identically distributed random variable with an arbitrary distribution is known as renewal process. Suppose$\{x_1 x_2 \dots \dots x_n\}$ be a sequence of interarrival times of a counting process $\{N(t), t \geq 0\}$. Let time until the 1st arrival $x_1$ has a distribution $F$ again let $x_2$ is the time between 1st and 2nd arrivals, independent of $x_1$ has the same distribution $F$ and so on, then $\{N(t), t \geq 0\}$.is known as renewal process.(See, Ross,(2000)).

**2.4.5 Branching process**

A Branching process is a mathematical description of the growth of a population for which the individual produces off-springs according to stochastic lows. A typical process is the following form.

The Branching process was proposed by Galton and the probability extinction was first obtained by Watson by considering the probability generating function for the number of children in the $n^{th}$ generation. This mathematical model was known as the Galton-Watson Branching Process and had been studied thoroughly in literature. ( See, Jagers, 1975)

In probability theory, a branching process is a Markov process that models a population in which each individual in generation n produces some random number of individuals in generation $n + 1$, according to a fixed probability distribution that does not vary from individual to individual. Branching processes are used to model reproduction; for example, the individuals might correspond to bacteria, each of which generates 0, 1, or 2 offspring with some probability in a single time unit. Branching processes can also be used to model other systems with similar dynamics, e.g., the spread of surnames in genealogy or the propagation of neutrons in a nuclear reactor.

A central question in the theory of Branching processes is the probability of ultimate extinction, where no individuals exist after some finite number of generations. It is not hard to show that, starting with one individual in generation zero, the expected size of generation n equals $\mu^n$ where $\mu$ is the expected number of children of each individual. If $\mu$ is less than 1, then the expected number of individuals goes rapidly to zero, which implies ultimate extinction with probability 1 by Markov's inequality. Alternatively, if $\mu$ is greater than 1, then the probability of ultimate extinction is less than 1 (but not necessarily zero; consider a process where each individual either dies without issue or has 100 children with equal probability). If $\mu$ is equal to 1, then ultimate extinction occurs with probability 1 unless each individual always has exactly one child. In theoretical ecology, the parameter $\mu$ of a branching process is called the basic reproductive rate.

### 2.4.5.1 Mathematical formulation of Branching process

The most common formulation of a branching process is that of the Galton-Watson process. Let $Z_n$ denote the state in period n (often interpreted as the size of generation n), and let $X_{n,i}$ be a random variable denoting the number of direct successors of member i in period n, where $X_{n,i}$ are iid over all $n \in \{0,1,2,...\}$ and $i \in \{1,...,Z_n\}$. Then the recurrence equation is

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i} \qquad (2.4.6)$$

with $Z_0 = 1$. Alternatively, one can formulate a branching process as a random walk. Let $S_i$ denote the state in period i, and let $X_i$ be a random variable that is iid over all i. Then the recurrence equation is

$$S_{i+1} = S_i + X_{i+1} - 1 = \sum_{j=1}^{i+1} X_j - i \qquad (2.4.7)$$

with $S_0 = 1$. To gain some intuition for this formulation, one can imagine a walk where the goal is to visit every node, but every time a previously unvisited node is visited, additional nodes are revealed that must also be visited. Let $S_i$ represent the number of revealed but unvisited nodes in period i, and let $X_i$ represent the number of new nodes that are revealed when node i is visited. Then in each period, the number of revealed but unvisited nodes equals the number of such nodes in the previous period, plus the new nodes that are revealed when visiting a node, minus the node that is visited. The process ends once all revealed nodes have been visited.

#### 2.4.5.2 Branching model

Consider an organism, cell, particle or an individual, whose life time has length one unit. At the end of its life time, it produces a random number ξ of off-spring where

$$P[\xi = k] = a_k, (k = 0,1,\dots), (\sum a_k = 1). \qquad (2.4.8)$$

Each one of the off-spring at the end of its life time produces off-springs the number of which has the same distribution.

Let $X_n$ be the number of particles at the end of the $n^{th}$ generation. We shall take $X_0 = 1$, since there is only one individual initially. The number of particles in the $(n+1)^{th}$ generation are the off-spring of those of the $n^{th}$ generation. If $X_n = i$ and $\xi_i$ denote the off-spring of the i individuals.

$$P[X_{n+1} = j \mid X_{n1} = i_1, X_{n2} = i_2, \dots X_n = i], \qquad n_1 < n_2 < \cdots < n,$$

$$= P[X_{n+1} = j \mid X_n = i],$$

$$= P[\xi_1 + \xi_2 + \cdots \xi_i = j] \qquad (2.4.9)$$

This equation depends on i and the off-spring distribution$\{a_k\}$. $\{X_n\}$ is Markov chain with state space (0,1,2,…).

The model as given above is called Galton-Watson branching process. ( See, Bhat, 2000 )

**2.4.5.2 Assumptions of Branching process**

Even though some of the assumptions such as independent identical off-spring distributions, non overlapping of generations, are stringent, because of the simplicity of the model. Branching process has found applications in many areas. ( See, Bhat, 2000 )

The number of individuals initially present, denoted by $X_0 = 1$ is called the size of the $0^{th}$ generation.

All offspring of the $0^{th}$ generation constitute the first generation and their number is denoted by $X_1$.

In general $\{ X_n\}$ denote the size of the $n^{th}$ generation and follows that $\{ X_n, n = 0, 1, \dots\}$ is a Markov chain having as its state space that the set of non negative integers.

To find the characteristics of the distribution of the population for different generation.

To find the probability of extinction of the population.

**2.4.5.3 Application of Branching process**

It is used as a model of survival and extinction of family names when records are available on the family tree. Genetics use the model for studying the survival extinction of a mutant gene. Biologists have assumed this model in discussing the growth of bacterial or microbial populations or of cells. An epidemic starts with one infected person, who transmits the disease to others. (See, Bhat, 2000)

Branching process have a wide variety of application in –

Biological science,

Sociological science and

Engineering science.

### 2.4.5.4 The Generating function of Branching process

Since the convolution of discrete distributions are best studied using their generating functions, we shall proceed to derive the same for $\{X_n\}$. As before we shall take $X_0$= 0, unless otherwise stated. Let $\emptyset_n(s)$ be the probability generating function of $\{X_n\}$, given $X_0$= 1.

Then

$$\emptyset_{n+1}(s) = \sum_{k=0}^{\infty} P[X_{n+1} = k|\, X_0 = 1]\, s^k,$$

$$= \sum_k \sum_j P[X_{n+1} = k|\, X_n = j]\, P[X_n = j|\, X_0 = 1] s^k,$$

$$= \sum_k s^k \sum_j P[X_n = j|\, X_0 = 1]\, P[\xi_1 + \xi_2 + \cdots \xi_i = k],$$

$$= \sum_j P[X_n = j|\, X_0 = 1] \sum_k s^k P[\textstyle\sum_1^j \xi_i = k],$$

$$= \sum_j P[X_n = j|\, X_0 = 1]\, [A(s)]^j,$$

$$= \emptyset_n(A(s)), (n = 1,2,3, \dots), \tag{2.4.10}$$

Where $\emptyset_1(s) = A(s)$, is the function of $X_1$. Therefore

$$\emptyset_2(s) = \emptyset_1(A(s)) = A\,(\,A\,(s)\,),$$

$$\emptyset_3(s) = A\,(\,A\,(A\,(s)\,)\,),$$

And in general

$$\emptyset_n(s) = \underbrace{A\,(A\,(..\left(A\left(A(s)\right)\right)..)}_{\text{n times}}, \tag{2.4.11}$$

The expression also implies that, when $X_0 = 1$.

If $\{X_n\}$ be a Branching process with $a_k = (1-p)p^k, k = 0,1, \dots$ then ,

$$A(s) = q/(1-ps)$$

$$\emptyset_2(s) = \frac{q}{1-\frac{pq}{q-ps}}$$

$$\vdots$$

$$\emptyset_n(s) = q/(1 - pq + pq/(1 - pq + pq/(1 - pq \ldots pq/(1 - ps))))$$
(2.4.12)

Which is a continued fraction?. ( See, Bhat, 2000 )

### 2.4.5.5 Mean and Variance of the branching process

For $X_0 = 1$, $m = E(X_1) = E(k) = A'$ is the off-spring mean. Let $\sigma^2 = \mathrm{Var}(X_1) = \mathrm{Var}(k)$ be its variance. Then

$$E(X_{n+1}) = \emptyset'_{n+1}(1) = \emptyset'_n(A(s))A'|_{s=1},$$

$$= m^{n+1};\ n = 1,2,\ldots \qquad (2.4.13)$$

If $X_0 = i$, then $E(X_{n+1}) = i\, m^{n+1}$.

Variance of $X_n$ is

$$\mathrm{Var}(X_n) = \emptyset''_n(1) + \emptyset'_n(1) - [\emptyset'_n(1)]^2, \text{ thus}$$

$$\mathrm{Var}(X_n) = \sigma^2 m^{n-1}\left(\frac{m^n - 1}{m-1}\right), \text{if } m \neq 1$$

$$= n\sigma^2, \text{if } m = 1 \qquad (2.4.14)$$

If $X_0 = i$, then $E(X_n) = i\, m^n$.

$$\mathrm{Var}(X_n) = i\sigma^2 m^{n-1}\left(\frac{m^n - 1}{m-1}\right), \text{if } m \neq 1$$

$$= in\sigma^2, \text{ if } m \qquad (2.4.15)$$

### 2.4.5.6 Moment generating function of the branching process

Moments generating function of the Branching process with the Branching Model

$$a_k = (1 - p)p^k, k = 0,1, \ldots \text{ is}$$

$$M_X(e^t) = \frac{p\, e^t}{(1 - q\, e^t)}, \qquad (2.4.16)$$

### 2.4.5.7 Determining the Probability of Extinction of the Population

Let $\pi_0$ denote the probability of die out (under the assumption that $X_0 = 1$). More formally,

$$\pi_0 = \lim_{n\to\infty} P\{X_n = 0 | X_0 = 1\}$$

The problem of determining the value of $\pi_0$ was first raised in connection with the extinction of family surnames by Galton in 1889.

We first note that $\pi_0 = 1$ if $\mu < 1$. This follows since

$$\mu^n = E(X_n) = \emptyset'_{n+1}(1) = \emptyset'_n(A(s))A'|_{s=1}, \quad n = 1,2,\ldots$$

Since $\mu^n \to 0$ when $\mu < 1$,

Then the Probability of Extinction of the Population is given by

$$P\{X_n = 0\} \to 1$$

In fact it can be shown that $\pi_0 = 1$, when $\mu = 1$. When $\mu > 1$, it turns out that $\pi_0 < 1$, and an equation determining $\pi_0$ may be derived by conditioning on the number of offspring of the initial individual as follows

$$\pi_0 = \sum_{k=0}^{\infty} P\{\text{Population dis out} \mid X_1 = k\}P_k$$

No, given that$X_1 = k$, the population will eventually die out if and only if each of the k families started by the members of the first generation eventually dies out.

Since each family is assumed to act independently and since the probability that any particular family dies out is just $\pi_0$, this yields

$$P\{\text{Population dis out} \mid X_1 = k\} = \pi_0{}^k$$

and thus $\pi_0$ satisfies

$$\pi_0 = \sum_{k=0}^{\infty} \pi_0{}^k P_k \qquad (2.4.17)$$

in fact when $\mu > 1$, it can be shown that $\pi_0$ is smallest positive number satisfying equation (2.4.17) .

### 2.4.6 Galton-Watson Branching process

The Galton–Watson Branching process is a branching stochastic process arising from Francis Galton’s statistical investigation of the extinction of family names.

Galton–Watson survival probabilities for different exponential rates of population growth, if the number of children of each parent node can be assumed to follow a Poisson distribution. For $\lambda \leq 1$, eventual extinction will occur with probability 1. But the probability of survival of a new type may be quite low even if $\lambda > 1$ and the population as a whole is experiencing quite strong exponential increase. There was concern amongst the Victorians that aristocratic surnames were becoming extinct.

Galton originally posed the question regarding the probability of such an event in the Educational Times of 1873, and the Reverend Henry William Watson replied with a solution. Together, they then wrote an 1874 paper entitled On the probability of extinction of families. Galton and Watson appear to have derived their process independently of the earlier work by I. J. Bienayme; see Heyde and Seneta 1977. For a detailed history see Kendall (1966 and 1975). Assume, as was taken for granted in Galton's time, that surnames are passed on to all male children by their father. Suppose the number of a man's sons to be a random variable distributed on the set { 0, 1, 2, 3, ... }.

Further suppose the numbers of different men's sons to be independent random variables, all having the same distribution. Then the simplest substantial mathematical conclusion is that if the average number of a man's sons is 1 or less, then their surname will almost surely die out, and if it is more than 1, then there is more than zero probability that it will survive for any given number of generations.

Modern applications include the survival probabilities for a new mutant gene, or the initiation of a nuclear chain reaction, or the dynamics of disease outbreaks in their first generations of spread, or the chances of extinction of small population of organisms; as well as explaining (perhaps closest to Galton's original interest) why only a handful of males in the deep past of humanity now have any surviving male-line descendants, reflected in a rather small number of distinctive human Y- chromosome DNA haplo groups. A corollary of high extinction probabilities is that if a lineage has survived, it is likely to have experienced, purely by chance, an unusually high growth rate in its early generations at least when compared to the rest of the population. ( See, Bruss, 1984 )

### 2.4.6.1 Mathematical definition of Galton-Watson Branching process

A Galton–Watson process is a stochastic process $\{X_n\}$ which evolves according to the recurrence formula $X_0 = 1$ and

$$X_{n+1} = \sum_{j=1}^{X_n} \xi_j^{(n+1)}$$

where for each n, $\xi_j^{(n)}$ is a sequence of IID natural number-valued random variables. The extinction probability (i.e. the probability of final extinction) is given by

$$\lim_{n\to\infty} \Pr(X_n = 0)$$

This is clearly equal to zero if each member of the population has exactly one descendent. Excluding this case (usually called the trivial case) there exists a simple necessary and sufficient condition. ( See, Bruss, 1984 )

### 2.4.6.2 Extinction criterion for Galton–Watson process

In the non-trivial case the probability of final extinction is equal to one if $E\{\xi_1\} \leq 1$ and strictly less than one if $E\{\xi_1\} > 1$.

The process can be treated analytically using the method of probability generating functions. If the number of children $\xi_j$ at each node follows a Poisson distribution, a particularly simple recurrence can be found for the total extinction probability $x_n$ for a process starting with a single individual at time n = 0:

$$x_{n+1} = e^{\lambda(x_n - 1)} \tag{2.4.18}$$

giving the curves plotted above. ( See, Bruss, 1984 )

### 2.5 Inference of Stochastic Process

One of the main objectives of Statistics is to draw inferences about a population from the analysis of a sample drawn from that population. Two important problems in statistical inference are:

Estimation.

Testing of hypothesis.

#### 2.5.1 Estimation of Stochastic Models:

In stochastic process estimation is an important part of statistical inference. For getting information about the parameter we need to estimate parameter of Markov chain. Estimated value of the parameter helps to take decision about the transition probability. Estimation technique for inference of transition probability are described as follows-

##### 2.5.1.1 Estimation for transition probability:

For purpose of inference a Markov chain can be observed in two ways:

One observation of a chain of great length or Observations on a large number of realizations of same chain. Observe a finite Markov chain with $m$ states $(1,2,\dots,m)$until $n$ transitions take place. Let $n_{ij}$ be the number of transitions from $i\ to\ j (i,j = 1,2,\dots,m)$ Let

$$\sum_{j=1}^{m} n_{ij} = n_i \qquad (2.5.1)$$

These numbers can be represented as:

| | 1 | 2 | 3 | … | $m$ | Total |
|---|---|---|---|---|---|---|
| 1 | $n_{11}$ | $n_{12}$ | $n_{13}$ | $\cdots$ | $n_{1m}$ | $n_1$ |
| 2 | $n_{21}$ | $n_{22}$ | $n_{23}$ | $\cdots$ | $n_{2m}$ | $n_2$ |
| 3 | $n_{31}$ | $n_{32}$ | $n_{33}$ | $\cdots$ | $n_{3m}$ | $n_3$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | $\vdots$ |
| $m$ | $n_{m1}$ | $n_{m2}$ | $n_{m3}$ | $\cdots$ | $n_{mm}$ | $n_m$ |
| Total | | | | | | $n$ |

Let the stationary transition matrix of the Markov chain is $P$ , given by

$$P=\begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \dots & P_{2m} \\ \vdots & \vdots & \dots & \vdots \\ P_{m1} & P_{m2} & \dots & P_{mm} \end{bmatrix} \tag{2.5.2}$$

We are interested in the estimation of the element $P_{ij}$ For a given initial state $i$ and a number of trials $n_i$, the sample of transition counts $(n_{i1}, n_{i2}, \dots, n_{im})$ can be considered as a sample of size $n_i$ from a multinomial distribution with probabilities $(p_{i1}, p_{i2}, \dots, p_{im})$, such that $\sum_{j=1}^{m} p_{ij} = 1$. The probability of the outcome can be given as:

$$\frac{n_{i1}!}{n_{i1}!n_{i2}!\dots n_{im}!} p_{i1}^{n} p_{i2}^{n} \dots p_{i3}^{n} \tag{2.5.3}$$

$where\ \sum_{j=1}^{m} n_{ij} = n_i,\ \sum_{j=1}^{n} P_{ij} = 1$

Extending this argument for the m initial states$(1,2,\dots,m)$, when the breakdown of the total number of trials n into $(n_1, n_2, \dots, n_3)$ is given, the probability of the realization of transition is given by,

$$\prod_{i=1}^{m} \frac{n_{i1}}{n_{i1}!n_{i2}!\dots n_{im}} p_{i1}^{n} p_{i2}^{n} \dots p_{im}^{n} \tag{2.5.4}$$

It can be shown that this distribution is independent of the probability elements $P_{ij}$ and therefore, without going into its explicit form, we shall denote it by $A$. The likelihood function and natural logarithm $L(p)$ can be expressed as:

$$f(p) = A \prod_{i=1}^{m} \frac{n_{i1}}{n_{i1}!n_{i2}!\dots n_{im}} p_{i1}^{n} p_{i2}^{n} \dots p_{im}^{n} \tag{2.5.5}$$

$$L(p) = lnB + L(p) = lnB + \sum_{i=1}^{m} \sum_{j=1}^{m} n_{ij} lnp_{ij} \tag{2.5.6}$$

Where $lnB$ contains all terms independent of $P_{ij}$'s to derive maximum likelihood elements, we maximize (2.4.6) under the condition $\sum_{j=1}^{n} P_{ij} = 1, (i = 1,2,\dots,m)$. Incorporating this condition in (2.5.6) we can write:

$$L(p) = lnB + \sum_{i=1}^{m} \sum_{j=1}^{m} n_{ij} \ln p_{ij} + \sum_{i=1}^{m} n_{im}\ \ln\left(1 - p_{i1} - p_{i2} - \cdots\ p_{im-1}\right) \tag{2.5.7}$$

For a specific value i, we have:

$L_i(p) = \ln B + \sum_{j=1}^{m} n_{ij}\ \ln p_{ij} + n_{ij}\ \ln(1 - p_{i1} - p_{i2} - \cdots - p_{i3})$ (2.5.8)

Differentiating with respect to $p_{ij}$ and setting it equal to zero, we get

$$\frac{n_{i1}}{p_{i1}} = \frac{n_{im}}{1 - p_{i1} - p_{i2} - \cdots\ p_{im-1}} = 0$$

$$\frac{n_{i2}}{p_{i2}} = \frac{n_{im}}{1 - p_{i1} - p_{i2} - \cdots\ p_{im-1}} = 0$$

$$\vdots$$

$$\frac{n_{im-1}}{p_{im-1}} = \frac{n_{im}}{1 - p_{i1} - p_{i2} - \cdots\ p_{im-1}} = 0$$

Combining these equations we may write:

$$\frac{n_{i1}}{p_{i1}} = \frac{n_{i2}}{p_{i2}} = \cdots \frac{n_{im-1}}{p_{im-1}} = \frac{n_{im}}{1 - p_{i1} - p_{i2} - \cdots\ p_{im-1}}$$

This leads us to write

$$\frac{n_{i1}}{n_{i1}} p_{i1} = p_{i1}$$

$$\frac{n_{i2}}{n_{i1}} p_{i1} = p_{i2}$$

$$\vdots$$

$$\frac{n_{im}}{n_{i1}} p_{i1} = 1 - p_{i1} - p_{i2} - \ldots - p_{im-1} \quad (2.5.9)$$

Adding both sides we get,

$$\frac{n_{i1} + n_{i2} + \ldots + n_{im}}{n_{i1}} p_{i1} = 1 \quad (2.5.10)$$

Which yields the estimate:

$$\hat{p}_{i1_1} = \frac{n_{i1}}{n_{i2}} \quad (2.5.11)$$

In a similar manner, we can derive estimates of other elements. Thus we get:

$$\hat{p}_{ij} = \frac{n_{ij}}{n_j}\ ; \quad i, j = 1,2, \ldots, m \quad (2.5.12)$$

(See, Bhat, (2000))

### 2.5.2 Hypothesis testing:

The objectives of the statistical analysis are to infer about the population parameter. The inference is draw on the basis of sample statistics and using some statistical procedure. The first step of the procedure is to assume a value about population parameter. The statistical procedure to varify the assumption on the basis of sample statistic is known as test of hypothesis. There are several hypothesis testing problems related to Markov chain. The major ones one of the following:

Tests for a given transition probability matrix.

Tests for the stationary of the transition probability matrix.

Tests for the order of the Markov chain.

Tests for the first order Markov dependence.

.

#### 2.5.2.1 Testing transition probability matrix

We consider the null hypothesis

$$H_0 : P = P_0$$

For large $n$, it can be shown that $n_{ij}$ are asymptotically normally distributed and that the static $n_i^{1/2}(\hat{P}_{ij} - P_{ij})$ has the normal distribution with mean 0 and variance $P_{ij}(1 - P_{ij})$. Thus we can write the test static as:

$$\sum_{j=1}^{m} \frac{n_i(\hat{P}_{ij} - P_{ij}^0)^2}{P_{ij}^0} \sim \chi^2_{(m-1)} \quad ; \ i = 1,2,\ldots,m \qquad (2.5.13)$$

Distributed as $\chi^2$ $with$ $(m-1)$ degrees of freedom. In obtaining, (2.5.1) we have assumed that all $P_{ij}^0$ are non-zero. If there are some zero elements in the $i^{th}$ row, only the non-zero elements should be considered in (2.5.1) and the degrees of freedom should be decreased by the number of zeros in it.

Thus we can write:

$$\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{n_i(\hat{P}_{ij}-P_{ij}^0)^2}{P_{ij}^0} \qquad (2.5.14)$$

Has a $\chi^2\ with\ (m-1)$ degrees of freedom, where $d$ is the number of zeros in $P^0$ and the summation in (2.5.14) is taken only over $(i,j)$ for which $P_{ij}^0 > 0$. (See, Bhat, (2000))

**2.5.2.2 Likelihood ratio test**

The likelihood ratio criteria for $H_0: P = P^0$can be obtained as

$$\Lambda = \frac{f(P^0)}{f(\hat{P})}$$

Where $f(\hat{P})$ is the maximized value of the likelihood function. From statistical theory, it is known that when $H_0$ is true.$-2\ ln\Lambda\sim\chi^2$ with $m(m-1)$ degrees of freedom. In this case we have

$$-2\ ln\Lambda = 2[L(\hat{P}) - L(P^0)] = 2\sum_{i=1}^{m}\sum_{j=1}^{m} n_{ij} ln\frac{n_{ij}}{n_i P_{ij}^0} \qquad (2.5.15)$$

(See, Bhat, (2000))

**2.5.2.3 Test for the Stationarity of the transition probability matrix**

$P_{ij}^t$ be the one step transition probability of a time dependent process $X(t)$, such that,

$$P_{ij}^t = P\{X(t+1) = j|\ X(t) = i\} \qquad (2.5.16)$$

Let $n_{ij}^t$ be the number of transition $i \to j$ during the $t^{th}$ transition of a process. For a given initial state $i$, the transition counts $n_{ij}^t; t = 1,2,\dots,T$ can be represented as

| | 1 | 2 | 3 | $\cdots$ | $m$ |
|---|---|---|---|---|---|
| 1 | $n_{i1}^{1}$ | $n_{i2}^{1}$ | $n_{i3}^{1}$ | $\cdots$ | $n_{im}^{1}$ |
| 2 | $n_{i1}^{2}$ | $n_{i2}^{2}$ | $n_{i3}^{2}$ | $\cdots$ | $n_{im}^{2}$ |
| 3 | $n_{i1}^{3}$ | $n_{i2}^{3}$ | $n_{i3}^{3}$ | $\cdots$ | $n_{im}^{3}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ |
| $T$ | $n_{i1}^{T}$ | $n_{i2}^{T}$ | $n_{i3}^{T}$ | $\cdots$ | $n_{im}^{T}$ |

The maximum likelihood estimates of $P_{ij}^{t}$ can be obtained:

$$P_{ij}^{t} = \frac{n_{ij}^{t}}{n_{i}^{t-1}} \; where, n_{i}^{t-1} = \sum_{j=1}^{m} n_{ij}^{t} \qquad (2.5.17)$$

Now we can consider the null hypothesis,

$$H_0 : P_{ij}^{t} = P_{ij} (t = 1,2, \dots, T)$$

Then the maximum likelihood function is given by $f(\hat{P}^{t})$ and the likelihood ration criterion $\Lambda$ is given by

$$\Lambda = \frac{f(P^{t})}{f(P)}$$

Here, $-2 \; ln\Lambda = 2\left[L(\hat{P}) - L(P^{0})\right] = 2\sum_{i=1}^{T}\sum_{i=1}^{m}\sum_{j=1}^{m} n^{t} \; ln \frac{n_{ij}^{t}}{n_{ij}^{t-1}P_{ij}}$ (2.5.18)

Under $H_0$, $-2 \; ln\Lambda \sim \chi^2$ with $(T-1)[m(m-1)]$ degrees of freedom. (See, Bhat, (2000)).

#### 2.5.2.4 Test for the order of Markov chain

Let

$$P_{ijk} = P(X_n = k| \; X_{n-2} = i) \qquad (2.5.19)$$

And $n_{ijk}$ be the corresponding transition count. Also let$n_{ij}^{t} = \sum_k n_{ijk}$. The estimates of $P_{ijk}$ is obtained as

$$P_{ijk} = \frac{n_{ij}}{n_{ij}^{t}}$$

Now consider the null hypothesis that the Markov chain is a first order Markov chain against the alternative that it is of order to can be given as $H_0: P_{ijk} = P_{ijk}(i, j, k = 1,2, \dots, m)$. To test the Hypothesis we consider the following test static

$$\sum_{i,j,k} \frac{n_{ij}^{t}(\hat{P}_{ijk} - \hat{P}_{jk})}{\hat{P}_{jk}} \tag{2.5.20}$$

which has a $\chi^2$ distribution with $m(m-1)^2$ degrees of freedom. The likelihood ratio test static can be written as,

$$-2\, ln\Lambda = 2n_{ijk}\left(\ln \hat{P}_{ijk} - \ln \hat{P}_{jk}\right) = 2\sum_{i,j,k} n_{ijk} \ln \frac{n_{ijk}}{n_{ij}^{t}} \frac{n_i}{n_{ij}} \tag{2.5.21}$$

Which has a $\chi^2$distribution with $m(m-1)^2$ degrees of freedom. Now if we test the null hypothesis that the Markov chain is of order $(r-1)$ against the alternative hypothesis that it is of order $r$ , the test static is

$$-2\, ln\Lambda = 2\sum_{i,j,k} n_{ijk} \ln \frac{n_{ijk}}{n_{ij}^{t}} \frac{n_i}{n_{ij}} \sim \chi^2 \tag{2.5.22}$$

With $m^{r-1}(m-1)^2$ degrees of freedom. (See, Bhat, (2000))

**2.5.2.5 Test for the first –order Markov dependence:**

We consider the null hypothesis that the observations are independent against the alternative that the process is a first order Markov chain. The test statistics are obtained by using an appropriate $P^0$. The appropriate $P^0$ should have identical rows under the hypothesis of independence.

Let $P^0$ consist of m identical rows $\pi = (\pi_1, \pi_2, \dots, \pi_m)$. When these probabilities are not known, their maximum likelihood estimates can be determined as follows:

Let $n_{ij} = \sum_{j=1}^{m} n_{ij}$ The log likelihood function can be written as

$$L(\pi) = \ln B + \sum_{j=1}^{m} n_{ij} \ln \pi_j \tag{2.5.23}$$

Hence the log likelihood function leads to the maximum likelihood estimate

$\hat{\pi}_j = \frac{n_j}{n}$

The $\chi^2$ test statistic can be written as

$$\sum_{i=1}^{m} \sum_{j}^{m} (n_{ij} - \frac{\frac{n_i n_j}{n}}{\frac{n_i n_j}{n}})^2 \quad (2.5.24)$$

The likelihood ration statistic can be written as:

$$-2 \, ln\Lambda = 2 \sum_{i=1}^{m} \sum_{j=1}^{m} n_{ij} \, ln \left( \frac{n_{ij}}{\frac{n_i n_j}{n}} \right) \quad (2.5.25)$$

has a $\chi^2$ distribution with $(m-1)^2$ degrees of freedom. (See, Bhat, (2000))

## 2.6 Model selection criteria

One of the most important assumption of classical linear regression model is that model used in the analysis is "correctly" specified. We encounter the problem of model specification error or model specification bias. To search the correct model  we need to identify some model selection criteria. Several criteria are used for this purpose. Such as

$a) R^2$

b) Adjusted $R^2 = (\bar{R}^2)$

c) Akaike information criterion (AIC)

d) Schwarz Information Criterion (SIC)

e) Mallows's $C_p$ Criterion

f) MDL

### 2.6.1 The $R^2$ Criterion:

We know that one of the measures of goodness of fit of a regression model is $R^2$, which, as we know, is defined as:

$$R^2 = \frac{ESS}{Tss} = 1 - \frac{Rss}{Tss} \qquad (2.6.1)$$

$R^2$, thus defined, of necessity lies between 0 and 1. The closer it is to 1, the better is the fit. But there are problems with $R^2$.It measures in sample goodness of fit in the sense of how close an estimated $Y$ value is to its actual value in the given sample. There is no guarantee that it will forecast well out of sample observations .In comparing two or more $R^2$'s, the dependent variable must be the same. The more importantly, a $R^2$cannot fall when more variables are added to the model. Therefore, there is every temptation to play the game of "maximizing the $R^2$ by simply adding more variables to the model. Of course, adding more variables to the model may increase $R^2$ but it may also increase the variance of forecast error. (See, Dmodar N. Guajarati)).

### 2.6.2 Adjusted: $R^2$

As a penalty for adding regressors to increase the $R^2$ value, Henry Theil developed the adjusted $R^2$, denoted by $\bar{R}^2$, can be written as

$$R^2 = \frac{ESS/(n-k)}{TSS/(n-1)} = 1 - (1 - R^2)\frac{n-1}{n-k} \qquad (2.6.2)$$

The adjusted $R^2$ will increase only if the absolute $t$ value of the added variable is greater that 1. For comparative purposes, therefore, $\overline{R}^2$ is a better measure that $R^2$. But again in this measure, the regressand must be the same for the comparison to be valid. (see, Dmodar N. Gujrati))

### 2.6.3 Akaike Information Criterion (AIC)

The idea of imposing a penalty for adding regressors to the model has been carried further in the AIC criterion, which is defined as:

$$AIC = e^{2k/n}\frac{\sum \hat{u}_i^2}{n} = e^{2k/n}\frac{RSS}{n} \qquad (2.6.3)$$

Where $k$ is the number of regressors and $n$ is the number of observations. For mathematical convenience, (1) is written as

$$\ln AIC = \frac{2\mathrm{k}}{\mathrm{n}} + \ln \frac{RSS}{n} \tag{2.6.4}$$

Where, ln$AIC$ is the natural log of $AIC$ and $\frac{2\mathrm{k}}{\mathrm{n}}$ is the penalty factor.

In comparing two or more models, the model with the lowest value of $\boldsymbol{AIC}$ is preferred. One advantage of $\boldsymbol{AIC}$ is that it is useful for not only in-sample but also out of sample forecasting performance of a regression model. Also, it is useful for both nested and non-nested models. It has been used to determine the lag length in an $AR(p)$ model. .(see, (Dmodar N. Guajarati))

### 2.6.4 Schwarz Information Criterion (SIC)

Similar in spirit to the AIC, the SIC criterion is defined as:

$$SIC = n^{k/n} \frac{\sum \hat{u}_i^2}{n} = n^{k/n} \frac{RSS}{n} \tag{2.6.5}$$

Or in log-form:

$$\ln SIC = \frac{k}{n} \ln n + ln \frac{RSS}{n} \tag{2.6.6}$$

where $\frac{k}{n} \ln n$ is the penalty factor. SIC imposes a harsher penalty than AIC. Like AIC, the lower the value of SIC, the better the model. Again, like AIC, SIC can be used to compare in sample or out of sample forecasting performance of a model. .(see, Dmodar N. Gujrati,(4th ed), P.-538)

### 2.6.5 Mallows's $C_p$ Criterion

Suppose we have a model consisting of $k$ regressors, including the intercept. Let $\hat{\sigma}^2$ as usual be the estimator of the true$\sigma^2$. But suppose that we only choose $p$ regressors $(p \leq k)$ and obtain the RSS form the regression using these $p$ regressors. Let $RSS_p$ denote the residual sum of squares using the $p$ regressors. Now, C. P. Mallows has developed the following criterion for model selection, known as the $C_p$ criterion:

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - (n - 2p) \tag{2.6.7}$$

Where $n$ is the number of observations.We know that $E(\hat{\sigma}^2)$ is an unbiased estimator of the true $\sigma^2$. Now, if the model with $p$ regressors is adequate in that it does not suffer from lack of fit, it can be shown that $E(RSS_p) = (n - 2p)\sigma^2$. In consequence, it is true approximately that

$$E(C_p) \approx \frac{(n-p)\sigma^2}{\sigma^2} - (n - 2p) \approx p \qquad (2.6.8)$$

in choosing a model according to the $C_p$ criterion, we would look for a model that has a low $C_p$ value, about equal to $p$. In other words, following the principle of parsimony, we will choose a model with $p$ regressors $(p < k)$ that gives a fairly good fit to the data. In practice, one usually plots $C_p$ computed from (2.8.8) against $p$. An "adequate" model will show up as a point close to the $C_p = p$ line, as can be seen from figure 1. As this figure shows, Model A may be preferable to Model B, as it is closer to the $C_p = p$ line than model B. (see, Dmodar N. Gujrati).

### 2.6.6 Forecast Chi-Square $\chi^2$

Suppose we have a regression model based on $n$ observations and suppose we want to use it to forecast the (mean) values of the regressand for an additional $t$ observations. It is a good idea to save part of the sample data to see how the estimated model forecasts the observations not included in the sample, the post sample period.

Now the forecast $\chi^2$ test is defined as follows:

$$Forcast,\ \chi^2 = \frac{\sum_{i=n+1}^{n+t} \hat{u}_i^2}{\hat{\sigma}^2} \qquad (2.6.9)$$

where $\hat{u}_i$ is the forecast error made for period $i = (n+1, n+2, \ldots, n+t)$, using the parameters obtained from the fitted regression and the values of the regressors in the post sample period. $\hat{\sigma}^2$ is the usual OLS estimator of $\sigma^2$ based on the fitted regression.

If we hypothesize that the parameter values have not changed between the sample and post sample periods, it can be shown that the statistic given in (2.8.9) following the chi-square distribution with $t$ degrees of freedom, where $t$ is the number of periods for which the forecast is made. .(see, Dmodar N. Gujrati,).

### 2.6.7 Minimum Description Length (MDL) principle

The model selection approach implied by the MDL principle requires that the value of

$$L_{MDL}(x,\theta) = -\lg(p(x \mid \theta)) + \sum_{1 \le i \le k} d_i.$$

(2.6.10)

be minimized through the choice of the model θ (Rissanen 1978)

## 2.7 Metropolis-Hasting Algorithm

In statistics and in statistical physics, the Metropolis–Hastings algorithm is a Markov chain Monte Carlo (MCMC) method for obtaining a sequence of random samples from a probability distribution for which direct sampling is difficult. This sequence can be used to approximate the distribution (i.e., to generate a histogram), or to compute an integral (such as an expected value). Metropolis–Hastings and other MCMC algorithms are generally used for sampling from multi-dimensional distributions, especially when the number of dimensions is high. For single dimensional distributions, other methods are usually available (e.g. adaptive rejection sampling) that can directly return independent samples from the distribution, and are free from the problem of auto-correlated samples that is inherent in MCMC methods. (en.wikipedia.org/wiki/Metropolish-Hasting_algorithm)

Algorithm for Metropolish–Hasting

Given $x^{(t)}$,

1. Generate $Y_t \sim q(y|x^{(t)})$.

2. Take $X^{(t+1)} = \begin{cases} Y_t \ \ with\, probability\, p\left(x^{(t)}, Y_t\right), \\ x^{(t)} \ \ with\, probability\ 1 - p\left(x^{(t)}, Y_t\right), \end{cases}$

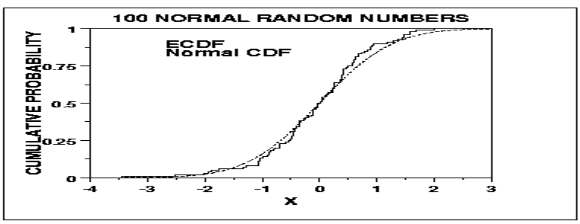
Where $p(x,y) = min\left\{\frac{f(y)}{f(x)}\frac{q(x|y)}{q(y|x)}, 1\right\}$.

## 2.8 Kolmogorov-Smirnov Goodness-of-Fit Test

The Kolmogorov-Smirnov test (Chakravart, Laha, and Roy, 1967) is used to decide if a sample comes from a population with a specific distribution.The Kolmogorov-Smirnov (K-S) test is based on the empirical distribution function (ECDF). Given *N ordered* data points $Y_1, Y_2, \ldots, Y_N$, the ECDF is defined as

$$E_N = \frac{n(i)}{N} \qquad (2.8.1)$$

Where $\boldsymbol{n(i)}$the number of points is less than $\boldsymbol{Y_i}$ and the $\boldsymbol{Y_i}$ are ordered from smallest to largest value. This is a step function that increases by $1/N$ at the value of each ordered data point.
The graph below is a plot of the empirical distribution function with a normal cumulative distribution function for 100 normal random numbers. The K-S test is based on the maximum distance between these two curves.




*Source: Chakravart, Laha, and Roy, 1967*
***Figure: 2.5*** *Normal random numbers to the K-S test*

The Kolmogorov-Smirnov test is defined by:

$H_0$: The data follow a specified distribution

$H_a$: The data do not follow the specified distribution

The Kolmogorov-Smirnov test statistic is defined as

$$D = \max_{1 \le i \le N} \left( F(Y_i) - \frac{i-1}{N}, \frac{i}{N} - F(Y_i) \right)$$

Where $\boldsymbol{F}$ is the theoretical cumulative distribution of the distribution being tested which must be a continuous distribution (i.e., no discrete distributions such as the binomial or Poisson), and it must be fully specified (i.e., the location, scale, and shape parameters cannot be estimated from the data).

## 2.9 Conclusions

In this chapter we try to introduce some MLM companies in Bangladesh. Review of different stochastic models is presented in section 2.4. In this chapter we also try to introduce inference of stochastic process that are used in estimation and test of hypothesis of a stochastic process. As our research focused on “On The use of Branching Process in Multi-Level Marketing in the Context of Bangladesh” so the analysis includes three phases, in the first phase we simulate the profit of MLM companies by using independent observation as well as related observation of the process. The later will do by Metropolish Hasting algorithm, which have discussed in section (2.7). In the second phase, we will check the underlying assumption of our generated variables using Kolmogorov-Smirnov Goodness-of-Fit Test that have presented in section (2.8). Finally, we will fit the Branching Process and Markov chain model for our data and make some concluding remarks for our study.

# Chapter: 3

# Result and analysis

### 3.1 Introduction

In this chapter we represent all the analysis and outcomes of our Research. This thesis concern with “On the use of Branching process in Multi level marketing (MLM) in the context of Bangladesh”. So, we collected data from two MLM companies. Companies are Destiny 2000 Ltd (D2L) and Medsit.net (M.net). We consider two individuals for two companies. For two individuals we observed that, under each individual there are few individuals of both sides. Under immediate of individuals there are two sides left and right sides. In both sides new offspring are produced. So we will observe the performance of Branching process for each individual. We will consider the data of total profit for both MLM companies to calculate the Markov chain and metro-polish-hasting algorithm as well as kolmogorov smirnov goodness of fit test.

The organization of this chapter is as follows: in section 3.1 introduction, in section 3.2 we describe the Univariate analysis, in section 3.3 we represent the illustration of Branching process ( in section 3.3.1 data structure, in section 3.3.2 offspring distribution, in section 3.3.3 ultimate extinction) , in section 3.4 we explain the Estimating Markov Chain with Generated Profit (In section 3.4.1 we explain Markov Chain Estimation using Monte Carlo Simulation, In section 3.4.2 we describe the Markov Chain Estimation using MCMC Simulation 3.4.3 long run probabilities, 3.4.4 Kolmogorov smirnov goodness of fit test ).

### 3.2 Univariate Analysis

Univariate analysis consists of descriptive statistic including measures of locations and measures of dispersions. Measures of locations indicate the data is how far from the central point and on the other hand measures of dispersion indicate how spread the data is. Different tabular and graphical presentation can be evaluated by Univariate analysis.

For Univariate analysis we find the descriptive statistics, histogram, and Box-plot of the profit of MLM companies Destiny 2000 Ltd (D2L) and Medsit.net (M.net) are as follows

### 3.2.1 Data consideration

In this section we have collected annual profit of Destiny 2000 Ltd (D2L) during the year 2007- 2011 and annual profit of Medsit.net (M.net) during 2006-2010 from the annual report of that companies separately. Then we have simulated profit weekly considering the original mean and standard deviation using Markov process (MP) and Metropolish-Hasting Algorithm related R- program respectively which are shown in Appendix.

**Table 3.1** Original annual net profit of MLM Company D2L

| **Studied year** | **Amount of profit in Taka** |
|---|---|
| 2007-2008 | 190155923 |
| 2008-2009 | 304946238 |
| 2009-2010 | 126660692 |
| 2010-2011 | 208329899 |

**Table 3.2** Original annual net profit of MLM Company M.net

| **Studied year** | **Amount of profit in dollar($)** |
|---|---|
| 2006 | 143139 |
| 2007 | 191451 |
| 2008 | 221190 |
| 2009 | 203346 |
| 2010 | 290533 |

**Table 3.3** Descriptive statistics for original Profit variable of MLM Company D2L

| Measure | values | Measure | values |
|---|---|---|---|
| Mean | 207523188 | Skewness | 0.648414705 |
| Standard Error | 36892034.6 | Range | 178285546 |
| Median | 199242911 | Minimum | 126660692 |
| Standard Deviation | 63898878.3 | Maximum | 304946238 |
| Sample Variance | 5.44409E+15 | Sum | 830092752 |
| Kurtosis | 1.450346874 | count | 4 |

Table 3.3 illustrates different descriptive measures of profit of D2L such as mean, median, standard deviation, standard error, sample variance, range, minimum, maximum, sum, kurtosis skewness and finally count. However, these summary information is not quite relevant as these are based on only for observation. On the other hand, in practical situation it is not possible to collect more information on this variable. So we use simulation to generate the profit of D2L and hence collect summary measures to approximate the real situation. As the profit of one year has a significant impact on the next year profit, so we need to generate related sequence of profit. Hence, Monte Carlo simulations are conducted to simulate the profit of D2L in two ways. At first, we generate the profit by using Markov process and in the second phase, we generate the profit by (MCMC) method using M-H algorithm.

**Table 3.4** Descriptive statistics for simulated Profit variable using Markov Chain (MC) of MLM Company D2L

| **Measure** | **values** | **Measure** | **values** |
|---|---|---|---|
| Mean | 207523188 | Skewness | -0.097142968 |
| Standard Error | 4533866.412 | Range | 346996141 |
| Median | 201687133.5 | Minimum | 26642054 |
| Standard Deviation | 63898878.3 | Maximum | 373638195 |
| Sample Variance | 4.08307E+15 | Sum | 42813229360 |
| Kurtosis | -0.46535679 | count | 208 |

Table 3.4 presents some summary descriptive information of our generated profit data of D2L using Markov Process. The average profit of D2L is 207523188 taka and median profit is 201687133.5 taka with a standard deviation 63898878.3 taka. So, there is a clear evidence of symmetric distribution (as Mean>Median>Mode), which is also explained by the coefficient of Skewness -0.097142968 (close to zero).

**Table 3.5** Descriptive statistics for simulated Profit variable using MCMC (Metropolish-Hasting algorithm) of MLM Company D2L

| Measure | values |
|---|---|
| Mean | 203399624.5 |
| Standard Error | 4642912.772 |
| Median | 202724780 |
| Mode | 201759111 |
| Standard Deviation | 66961040.27 |
| Sample Variance | 4.48378E+15 |
| Kurtosis | -0.479059609 |
| Skewness | -0.076909785 |
| Range | 294539204 |
| Minimum | 44568310 |
| Maximum | 339107514 |
| Sum | 42307121891 |
| count | 208 |

Table 3.5 describes some summary descriptive information of our generated profit data of D2L using Metropolish Hasting algorithm. The average profit of D2L is 203399624.5 taka, median profit is 202724780 taka and mode profit is 201759111 with a standard deviation 66961040.27 taka. So, there is a clear evidence of symmetric distribution (as Mean>Median>Mode), which is also explained by the coefficient of Skewness -0.076909785 (close to zero).

**Table 3.6** Descriptive statistics for original Profit variable of MLM Company M.net

| Measure | values | Measure | Values |
|---|---|---|---|
| Mean | 209931.8 | Skewness | 0.598312381 |
| Standard Error | 23947.32905 | Range | 147394 |
| Median | 203346 | Minimum | 143139 |
| Standard Deviation | 53547.85563 | Maximum | 290533 |
| Sample Variance | 2867372843 | Sum | 1049659 |
| Kurtosis | 1.463043404 | count | 5 |

Table 3.6 illustrates different descriptive measures of profit of M.net such as mean, median, standard deviation, standard error, sample variance, range, minimum, maximum, sum, kurtosis skewness and finally count. However, this summary information is not quite relevant as these are based on only 4 observations. On the other hand, in practical situation it is not possible to collect more information on this variable. So we use simulation to generate the profit of M.net and hence collect summary measures to approximate the real situation. As the profit of one year has a significant impact on the next year profit, so we need to generate related sequence of profit. Hence, Monte Carlo simulations are conducted to simulate the profit of M.net in two ways. At first, we generate the profit by using Markov process and in the second phase, we generate the profit by (MCMC) method using M-H algorithm.

**Table 3.7** Descriptive statistics for Simulated Profit variable using Markov Chain (MC) of MLM Company M.net

| **Measure** | **values** |
|---|---|
| Mean | 209931.8 |
| Standard Error | 3111.561011 |
| Median | 209700.5 |
| Mode | 184307 |
| Standard Deviation | 53547.86 |
| Sample Variance | 2867373311 |
| Kurtosis | -0.268908511 |
| Skewness | -0.140937508 |
| Range | 262693 |
| Minimum | 68356 |
| Maximum | 331049 |
| Sum | 54419780 |
| count | 260 |

Table 3.7 represents some summary descriptive information of our generated profit data of M.net using Markov Process. The average profit of M.net is 209931.8 dollar, median profit is 209700.5 dollar and mode profit is 184307 dollar with a standard deviation 53547.86 dollar. So, there is a clear evidence of symmetric distribution (as Mean>Median>Mode), which is also explained by the coefficient of Skewness -0.140937508.

**Table 3.8** Descriptive statistics for Simulated Profit by MCMC (Metropolish-Hasting algorithm) of MLM Company M.net

| **Measure** | **values** |
|---|---|
| Mean | 199646.692 |
| Standard Error | 3158.204207 |
| Median | 207112.1 |
| Mode | 207259.8 |
| Standard Deviation | 50924.51268 |
| Sample Variance | 2593305991 |
| Kurtosis | 1.529883525 |
| Skewness | -0.730437879 |
| Range | 298042.68 |
| Minimum | 298042.68 |
| Maximum | 310778.6 |
| Sum | 51908139.93 |
| count | 260 |

Table 3.8 describes some summary descriptive information of our generated profit data of M.net using Metropolish Hasting algorithm. The average profit of M.net is 199646.692 dollar, median profit is 207112.1 dollar and mode profit is 207259.8 dollar with a standard deviation 50924.51268 dollar. So, there is a clear evidence of negative skewed of shape.(as Mean<Median<Mode), which is also explained by the coefficient of Skewness -0.730437879.

### 3.2.2 Graphical Representations

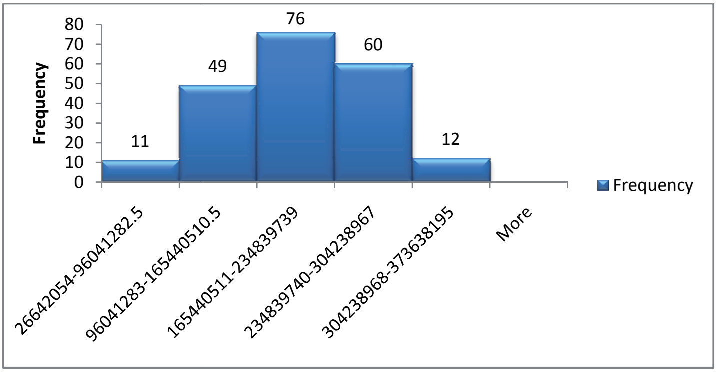



***Figure 3.1*** *Histogram of simulated Profit variable using Markov Chain (MC) of MLM Company D2L*

Figure 3.1 Indicates that overall shape of the distribution is symmetric. Lowest frequency of weekly profit lie between class interval 26642054-96041282.5 and highest frequency of weekly profit lie between class intervals 165440511-234839739.

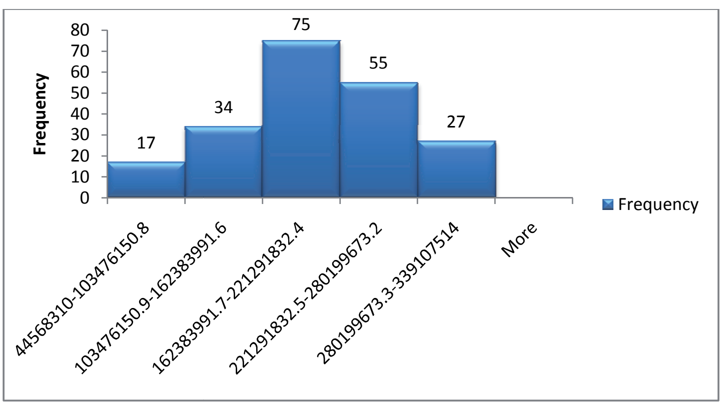



***Figure 3.2*** *Histogram for simulated profit variable by MCMC(Metropolish-Hasting algorithm) of MLM company D2L*

Figure 3.2 Indicates that overall shape of the distribution is symmetric. Lowest frequency of weekly profit lie between class interval 44568310-103476510.8and highest frequency of weekly profit lie between class intervals 162383991.7-221291832.4.

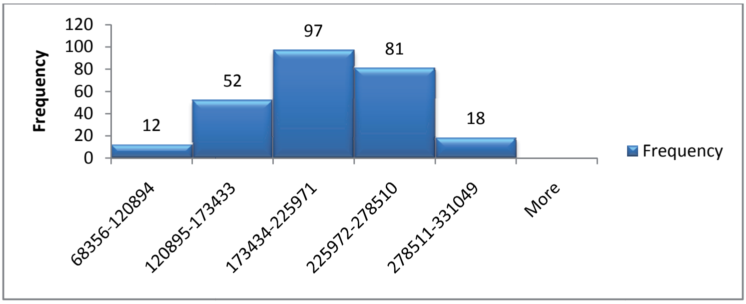



***Figure 3.3*** *Histogram for Simulated Profit variable by Markov Chain (MC) of MLM Company M.net*

Figure 3.3 Indicates that overall shape of the distribution is symmetric. Lowest frequency of weekly profit lie between class interval 68356-120894.8and highest frequency of weekly profit lie between class intervals 173434-225971.

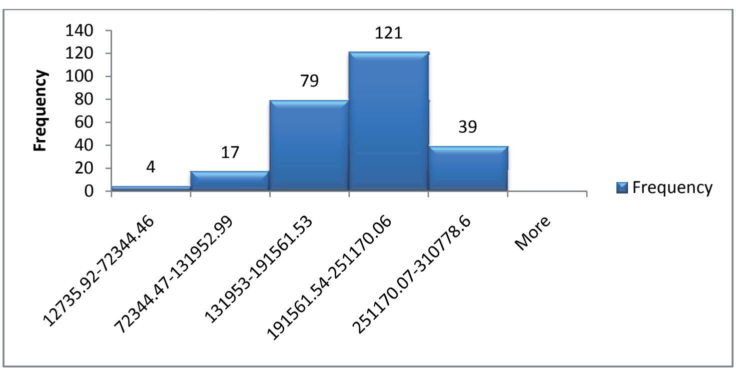



***Figure 3.4*** *Histogram for Simulated Profit variable by MCMC (Metropolish-Hasting algorithm) of MLM Company M.net*

Figure 3.4 describes that overall shape of the distribution is assymetric. Lowest frequency of weekly profit lie between class interval 12735.92-72344.46.8and highest frequency of weekly profit lie between class intervals 191561.54-251170.

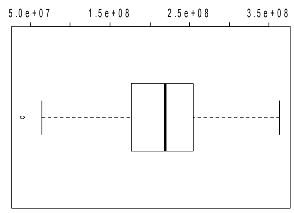



***Figure 3.5*** *Box-plot for simulated profit variable by Markov Chain (MC) for MLM Company D2L*

Figure 3.5 shows that overall shape of the distribution is symmetric, whereas centered 50% observations also express a symmetry shape. There is an evidence of a outlier in the generated observation.

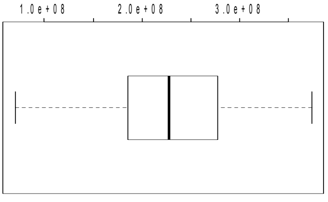



***Figure 3.6*** *Box-plot for simulated profit variable by MCMC (Metropolish-Hasting algorithm) for MLM Company D2L*

Figure 3.6 explains that overall shape of the distribution is symmetric, whereas centered 50% observations also express a symmetry shape. There is an evidence of outlier in the generated observation.

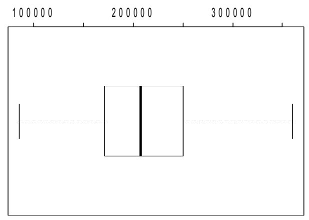



***Figure 3.7*** *Box-plot for simulated profit variable by Markov Chain (MC) of MLM Company M.net*

Figure 3.7 represents that overall shape of the distribution is symmetric, whereas centered 50% observations also express a symmetry shape. There is an evidence of a outlier in the generated observation.

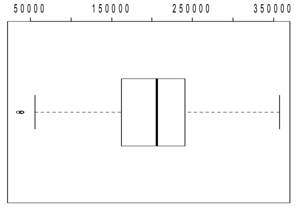



***Figure 3.8*** *Box-plot for simulated profit variable by MCMC (Metropolish-Hasting algorithm) of MLM Company M.net*

Figure 3.8 represents that overall shape of the distribution is negatively skewed, whereas 50% observations does not express a symmetry shape. There is an evidence of outlier in the generated observation.

### 3.3 Illustration of Branching Process

#### 3.3.1 Data structure

There are two types data are used in my thesis from two MLM companies in respect of Bangladesh. Firstly we have used personal data from individuals for illustration of Branching process. We consider two individuals from two MLM companies. We count each individual as a producer of offspring or client. We observed that immediate under each individual there may exist at most two individuals. We consider each individual as an unit of offspring producer for further generation

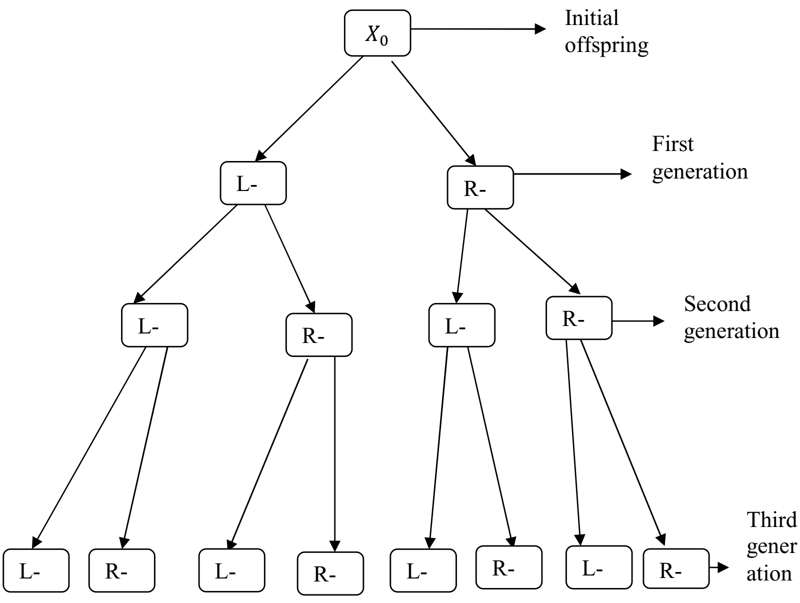



***Figure: 3.9*** *structure of the data of MLM companies individuals.*

L-H indicates left hand of a client whether R-H indicates a right hand.Then we use the data about the total profit of two MLM companies for some years to explore the Markov Chain. This type of profit data will help us to predict the future profit of MLM Company.

### 3.3.2 Down line clients (Offspring) Distribution

To find the distribution of down line clients firstly we calculate frequency distributions of down line clients and showing the tree diagram from generation to generation based on distributors nature , then calculate the mean and variance of the offspring:

**Table 3.9** Distribution of offspring of the surveyed individual of MLM Company D2L.

| **No. of offspring (X)** | **Frequency** | **Probability ($P_i$)** |
|---|---|---|
| 0 | 5 | .14 |
| 1 | 15 | .43 |
| 2 | 15 | .43 |
| **Total** | **35** | **1** |

**Table 3.10** Distribution of offspring of the surveyed individual of MLM Company M.net.

| **No. of offspring (X)** | **Frequency** | **Probability ($P_i$)** |
|---|---|---|
| 0 | 9 | .14 |
| 1 | 27 | .42 |
| 2 | 28 | .44 |
| **Total** | **64** | **1** |

Mean of the down line clients (offspring) of MLM Company D2L is calculated below:

$$\mu = E(X_1)$$

$$= \sum_{k=0}^{2} kp_k$$

$$= 0 \times p_0 + 1 \times p_1 + 2 \times p_2$$

$$= 0 \times .14 + 1 \times .43 + 2 \times .43$$

$$= 0 + .43 + .86 = 1.29$$

Hence, the variance is;

$$\sigma^2 = var(X_1) \quad = E\left({X_1}^2\right) - \{E(X_1)\}^2$$

$$= 0^2 \times p_0 + 1^2 \times p_1 + 2^2 \times p_2 - (\mu)^2$$

$$= 0^2 \times .14 + 1^2 \times .43 + 2^2 \times .43 - (1.29)^2$$

$$= .4859$$

Since our calculated mean is μ=1.29 which is greater than 1, hence we can say that our obtaining Branching process is super-critical.

Again,

Mean of the down line clients (offspring) of MLM Company M.net is calculated below

$$\mu = E(X_1)$$

$$= \sum_{k=0}^{2} kp_k$$

$$= 0 \times p_0 + 1 \times p_1 + 2 \times p_2$$

$$= 0 \times .14 + 1 \times .42 + 2 \times .44$$

$$= 0 + .42 + .88 = 1.30$$

Hence, the variance is;

$$\sigma^2 = var(X_1) \quad = E\left({X_1}^2\right) - \{E(X_1)\}^2$$

$$= 0^2 \times p_0 + 1^2 \times p_1 + 2^2 \times p_2 - (\mu)^2$$

$$= 0^2 \times .14 + 1^2 \times .42 + 2^2 \times .44 - (1.30)^2$$

$$= .49$$

Since our calculated mean is μ=1.30 which is greater than 1, hence we can say that our obtaining Branching process is super-critical.

**3.3.3 Different Generations Structure of down line Clients:**

Tree diagram of the surveyed individual of MLM Company D2L 1st generation, 2nd generation and at next page 3rd generation is shown.

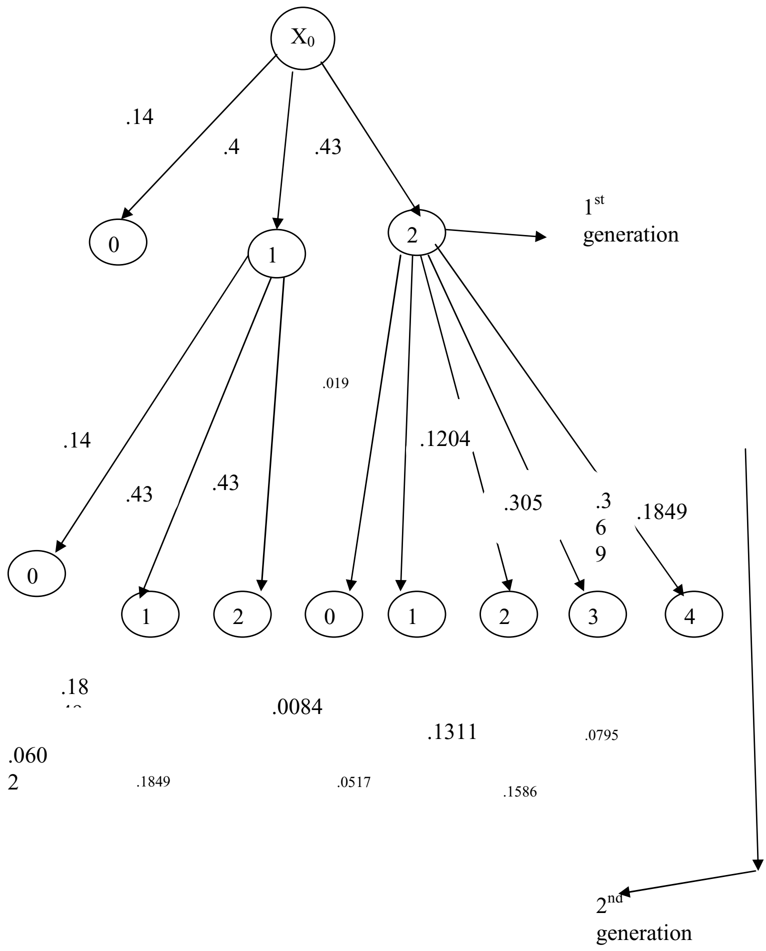
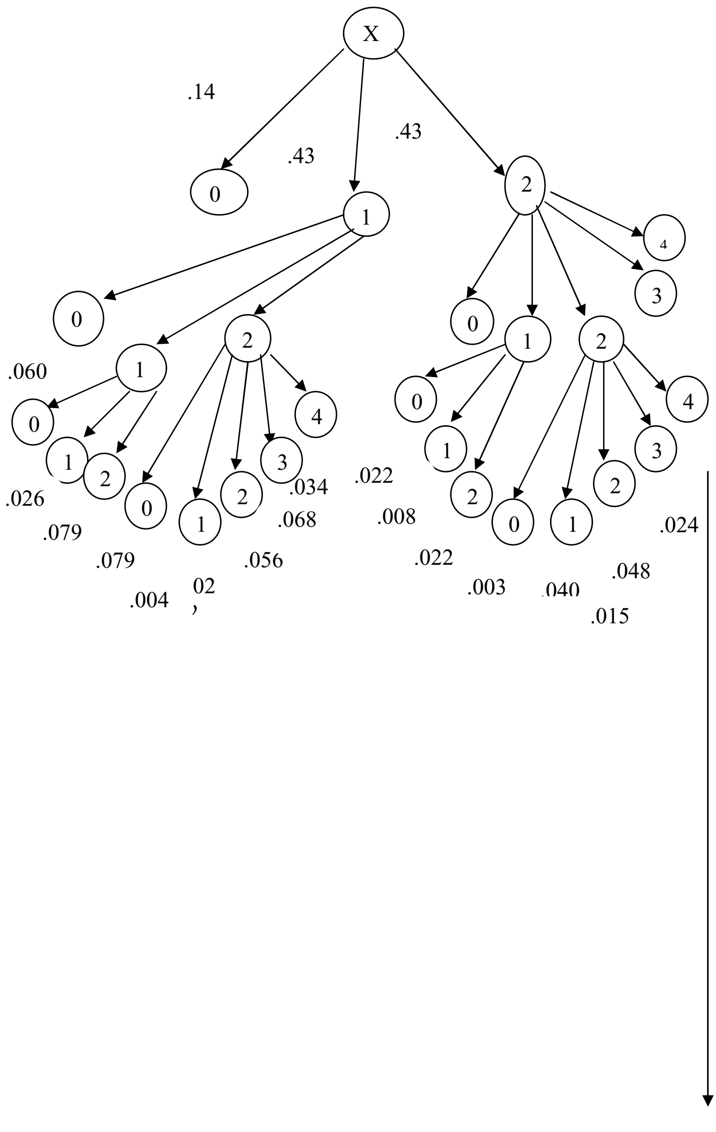



***Figure 3.10*** *Tree diagram of the surveyed individual of MLM Company D2L with their corresponding probabilities (up to2$^{nd}$ generation).*


X
.14
.43
.43
0
1
2
4
3
0
0
1
2
0
1
2
.060
0
1
2
0
1
2
3
4
0
1
2
3
4
.026
.079
.079
.004
02
,
.056
.068
.034
.022
.008
.022
.003
.040
.015
.048
.024

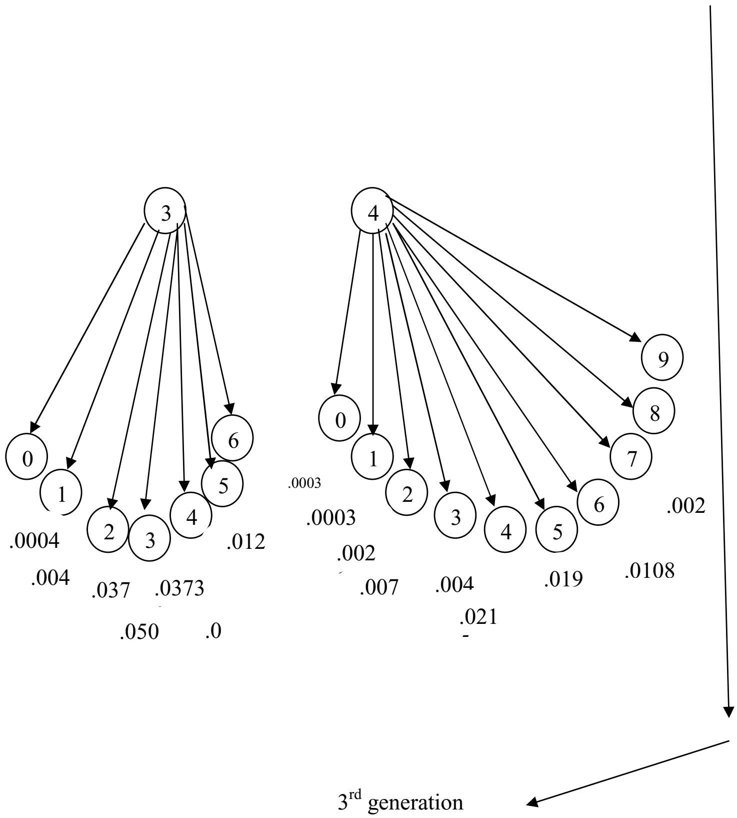



***Figure 3.11*** *Tree diagram of the surveyed individual of MLM Company D2L with their corresponding probabilities (up to 3rd generation).*

Tree diagram for surveyed individual of MLM Company M.net 1st generation, 2nd generation and at next page 3rd generation is shown:

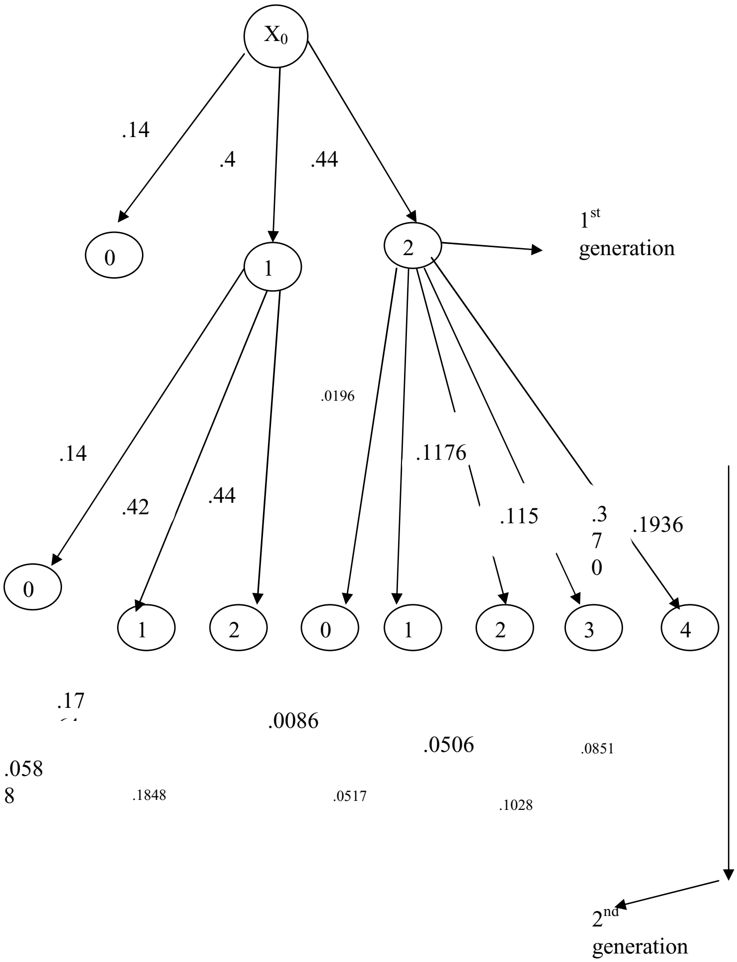
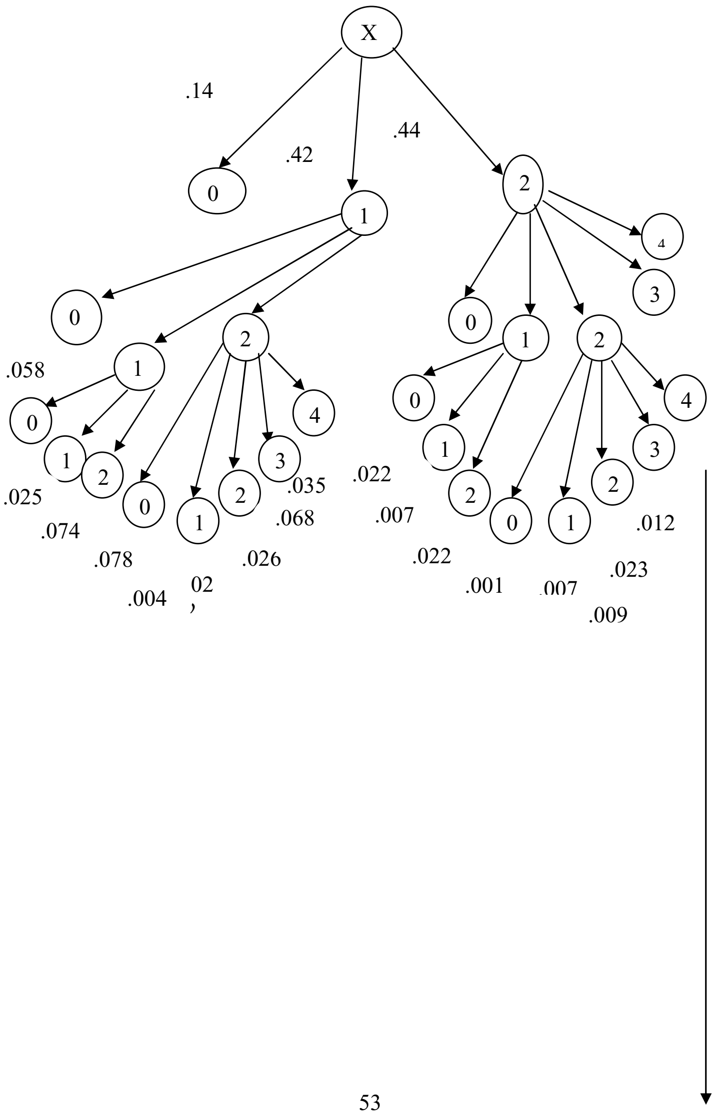



***Figure 3.12*** *Tree diagram of the surveyed individual of MLM Company M.net with their corresponding probabilities (up to2nd generation).*


X
.14
.42
.44
0
1
2
4
3
0
1
2
0
1
2
4
3
.058
.025
.074
.078
.004
02
.026
.068
.035
.022
.007
.022
.001
.007
.009
.023
.012

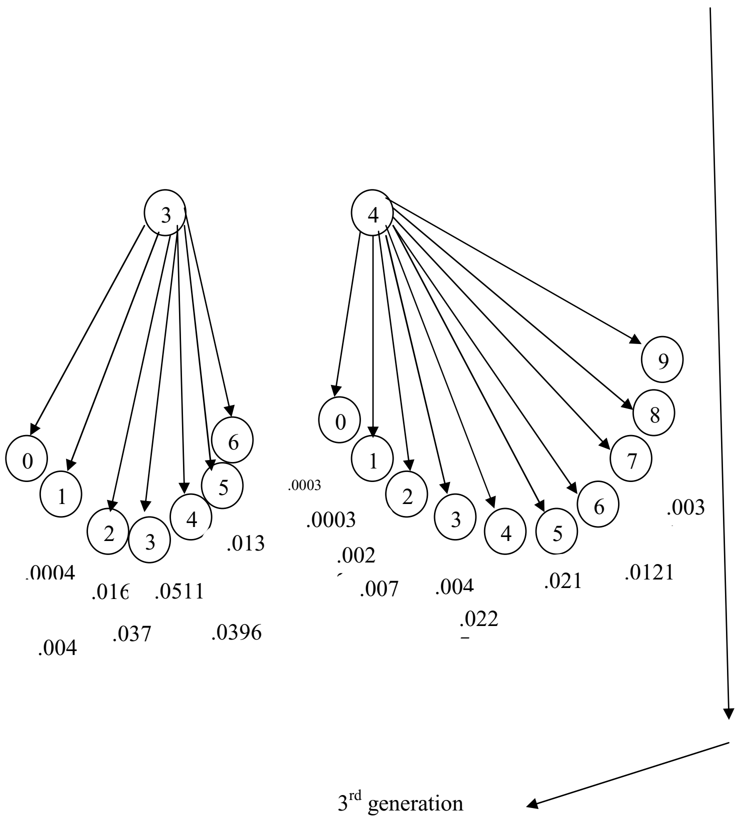



***Figure 3.13*** *Tree diagram of the surveyed individual of MLM Company M.net with their corresponding probabilities (up to 3rd generation).*

We found that the probability of large offspring size is near to 0 for future generations. To establish the similar findings we can calculate the mean and variance of different generations.

We know that,

$$Mean\ E(X_n) = \mu^n \tag{3.3.1}$$

$$\text{Variance}(X_n) = \begin{cases} \frac{\mu^{n-1}(\mu^n-1)}{\mu-1}\sigma^2; & \mu \neq 1 \\ n\sigma^2; & \mu = 1 \end{cases} \tag{3.3.2}$$

For different values of n we find mean and variance of 1 to 20 generations are as follows:

**Table 3.11:** Mean and variance of Down Line Clients of MLM company D2L for Different Generations

| Generation Number | Mean ($X_n$) | Variance ($X_n$) |
|---|---|---|
| 1 | 1.29 | .4859 |
| 2 | 1.67 | 1.435 |
| 3 | 2.15 | 3.197 |
| 4 | 2.77 | 6.363 |
| 5 | 3.57 | 11.935 |
| 6 | 4.60 | 21.597 |
| ⋮ | ⋮ | ⋮ |
| 10 | 12.76 | 194.94 |
| ⋮ | ⋮ | ⋮ |
| 20 | 162.85 | 34235.17 |

**Table 3.12:** Mean and variance of Down Line Clients of MLM company M.net for Different Generations

| Generation Number | Mean ($X_n$) | Variance ($X_n$) |
|---|---|---|
| 1 | 1.30 | .49 |
| 2 | 1.69 | 1.4651 |
| 3 | 2.19 | 3.304 |
| 4 | 2.86 | 6.660 |
| 5 | 3.71 | 12.656 |
| 6 | 4.83 | 23.20 |
| ⋮ | ⋮ | ⋮ |
| 10 | 13.79 | 221.46 |
| ⋮ | ⋮ | ⋮ |
| 20 | 190.05 | 45141.33 |

### 3.3.4 Ultimate Extinction

Probability of ultimate extinction for MLM Company D2L

Since, $\mu > 1$ we have to solve the following equqtions

$$s = p(s)$$

$$=> \; s = \sum_{k=0}^{2} p_k s^k$$

$$=> \; s = p_0 s^0 + \; p_1 s^1 + \; p_2 s^2$$

$$=> \; s = .14 + .43s + .43\, s^2$$

$$=> \; .43\, s^2 - .57s + .14 = 0$$

$$=> \; s = \frac{-(-.57) \pm \sqrt{(-.57)^2 - 4 \times (.43) \times (.14)}}{2 \times .43}$$

$$=> s = \frac{.57 \pm \sqrt{.0841}}{2 \times .43}$$

$$=> s = \frac{.57 \pm .29}{.86}$$

$$=> s = \frac{.57 + .29}{.86}, \frac{.57 - .29}{.86}$$

$$=> s = 1, .325$$

So the ultimate probability of extinction is= .325

Similarly, Probability of ultimate extinction for MLM Company M.net:

Since, $\mu > 1$ we have to solve the following equations

$$s = p(s)$$

$$=> \; s = \sum_{k=0}^{2} p_k s^k$$

$$=> \; s = p_0 s^0 + \; p_1 s^1 + \; p_2 s^2$$

$$=> \quad s = .14 + .42s + .44\, s^2$$

$$=> \ .44\, s^2 - .58s + .14 = 0$$

$$=> \ s = \frac{-(-.58) \pm \sqrt{(-.58)^2 - 4 \times (.44) \times (.14)}}{2 \times .44}$$

$$=> s = \frac{.58 \pm \sqrt{.09}}{2 \times .44}$$

$$=> s = \frac{.58 \pm .30}{.88}$$

$$=> s = \frac{.58 + .30}{.88}, \frac{.58 - .30}{.88}$$

$$=> s = 1, .318181818$$

So the ultimate probability of extinction is= .318181818

### 3.4 Estimating Markov Chain with Generated Profit

Data are generated for the annual profit of Destiny 2000 Ltd (D2L) and Medsit.net (M.net) in two phases for each company. At the first phase, Markov Chain is simulated by generating random numbers using Monte Carlo simulation and Markov Chain Monte Carlo (MCMC) simulation is done by using Metropolis-Hasting algorithm in the second stage.

As the profit follows normal distribution with mean $\mu$ and variance $\sigma^2$ . Hence, we will generate the profit of D2L with mean (207523188) and standard deviation (63898878.3) and the profit of M.net with mean (209931.8) and standard deviation (53547.86) in the first stage. The program codes are given in Appendix A1. In the second stage, the profits of two companies are generated by MCMC using Metropolis-Hasting (MH) Algorithm.

#### 3.4.1 Markov Chain Estimation using Monte Carlo Simulation

Before simulating the profit of two companies, we need to classify different states of this variable. Since the minimum profit of D2L is 26642054 taka and maximum is 373638195 taka, whereas the minimum profit is 68356 dollar and maximum profit is 331049 dollar for M.net. So the states ranges are different for two companies.

**Table 3.13:** Classification of States in the profit of D2L for Markov Chain model during year2007-2008 to 2010-2011

| Generated Profit Range (in taka) | state |
|---|---|
| 26642054-96041282 | 1 |
| 96041283-165440510 | 2 |
| 165440511-234839738 | 3 |
| 234839739-304238966 | 4 |
| 304238967-373638195 | 5 |

**Table 3.14** Classification of States in the profit of M.net for Markov Chain model during year 2006 to 2010

| Generated Profit Range(in dollar) | state |
|---|---|
| 68356-120895 | 1 |
| 120896-173434 | 2 |
| 173435-225973 | 3 |
| 225974-278512 | 4 |
| 278513-331049 | 5 |

Using R-program (Appendix) we get Transition matrix T for the generated profit of D2L in the first phase simulation (Markov chain) is given below:

$$T = \begin{bmatrix} 0 & 1 & 6 & 2 & 1 \\ 2 & 6 & 20 & 9 & 1 \\ 5 & 19 & 43 & 28 & 5 \\ 2 & 9 & 26 & 13 & 1 \\ 1 & 3 & 4 & 0 & 0 \end{bmatrix} \tag{3.4.1}$$

As a result, Transition probability matrix (TPM) for D2L using Markov Chain (MC) simulation during 2007-2008 to 2010-2011 is

$$P = \begin{bmatrix} 0.00000000 & 0.1000000 & 0.6000000 & 0.2000000 & 0.10000000 \\ 0.05263158 & 0.1578947 & 0.5263158 & 0.2368421 & 0.02631579 \\ 0.05000000 & 0.1900000 & 0.4300000 & 0.2800000 & 0.05000000 \\ 0.03921569 & 0.1764706 & 0.5098039 & 0.2549020 & 0.01960784 \\ 0.12500000 & 0.3750000 & 0.5000000 & 0.0000000 & 0.00000000 \end{bmatrix} \qquad (3.4.2)$$

From the TPM matrix we observed that the profit of D2L starts with 26642054-96041282 taka in the first year will be retain in the same range in the next year with probability 0, there is 10% chance in increase the profit to 96041283-165440510 taka, 60% likelihood for increasing the profit to 165440511-234839738 taka, 20% possibility to increase in next state and only 10% chance remaining to earn more 304238967-373638195 in the next year.

When the profit of D2Lstarts with 96041283-165440510 taka in the first year it decreases to 26642054-96041282 taka in the next year with probability .05, there is 16% chance to remain the same profit range, 52% likelihood for increasing the profit to 165440510-234839738 taka, 23% possibility to raise in 234839739-304238966 and only 2% chance remaining to earn 304238967-373638197 in the next year.

When the profit of D2L starts with 165440510-234839738 taka in the first year it decreases to 26642054-96041282 taka in the next year with probability .05, there is 19% chance in decreases the profit to 96041283-165440510 taka, there is 43% chance to remain the same profit range ,28% likelihood for increasing the profit to 234839739 -304238966 taka, 5% possibility to raise in 304238967-373638195 in the next year.

When the profit of D2L starts with 234839739-304238966 taka in the first year it decreases to 26642054-96041282 in the next year with probability .03, there is 18% chance to decrease profit to 96041283-165440510, 50% likelihood for decreasing the profit to 165440510-234839738 taka, 25% possibility to remain on the same range and only 2% chance to increase the profit range 304238967-373638197 in the next year.

When the profit of D2Lstarts with more than 304238967-373638195 taka in the first year it decreases to 26642054-96041282 taka in the next year with probability.013, there is 38% chance to decrease profit range 96041283-165440510, 50% likelihood for decreasing the profit to 165440511-234839738, no chance to decrease in 234839739-304238966 and there is 0% chance to remain the profit in the same range.

$$P^2 = \begin{bmatrix} 0.5560630 & 0.2025836 & 0.4625924 & 0.2426646 & 0.03655315 \\ 0.04720344 & 0.1818580 & 0.4748982 & 0.2556624 & 0.04037803 \\ 0.04873039 & 0.1848618 & 0.4826451 & 0.2467725 & 0.03699020 \\ 0.04473684 & 0.1809837 & 0.4753783 & 0.2573589 & 0.03905381 \\ 0.04473684 & 0.1667105 & 0.4873684 & 0.25381858 & 0.04736842 \end{bmatrix} \quad (3.4.3)$$

From the second order TPM the profit of D2L starts with 26642054-96041282 taka in the first year will be retain in the same range in the next year with probability 0.56, there is 20% chance in increase the profit to 96041283-165440510 taka, 46% likelihood for increasing the profit to 165440511-234839738 taka, 24% possibility to increase in next state and only 3% chance remaining to earn more 304238967-373638195 in the next year.

When the profit of D2Lstarts with 96041283-165440510 millions in the first year it decreases to 26642054-96041282 taka in the next year with probability .047, there is 18% chance to remain the same profit range, 47% likelihood for increasing the profit to 165440510-234839738 taka, 25% possibility to raise in 234839739-304238966 and only 4% chance remaining to earn 304238967-373638197 in the next year.

When the profit of D2L starts with 165440510-234839738 taka in the first year it decreases to 26642054-96041282 taka in the next year with probability .048, there is 18% chance in decreases the profit to 96041283-165440510 taka, there is 48% chance to remain the same profit range, 25% likelihood for increasing the profit to 234839739 -304238966 taka, 4% possibility to raise in 304238967-373638195 in the next year.

When the profit of D2L starts with 234839739-304238966 taka in the first year it decreases to 26642054-96041282 in the next year with probability .044, there is 18% chance to decrease profit to 96041283-165440510, 48% likelihood for decreasing the profit to 165440510-234839738 taka, 26% possibility to remain on the same range and only 3% chance to increase the profit range 304238967-373638197 in the next year.

When the profit of D2Lstarts with more than 304238967-373638195 taka in the first year it decreases to 26642054-96041282 taka in the next year with probability.044, there is 17% chance to decrease profit range 96041283-165440510, 49% likelihood for decreasing the profit to 165440511-234839738, 25% chance to decrease in 234839739-304238966 and there is only 4% chance to remain the profit in the same range.

Similarly we get $P^3, P^4, P^5, P^6, P^7, P^8\ P^9,$ and finally $P^{10}$. Where $P^{10}$ is the convergent point.

$$P^{10} = \begin{bmatrix} 0.04825019 & 0.1834928 & 0.478613 & 0.2511354 & 0.038505865 \\ 0.0482509 & 0.1834928 & 0.478613 & 0.2511354 & 0.038505865 \\ 0.0482509 & 0.1834928 & 0.478613 & 0.2511354 & 0.038505865. \\ 0.0482509 & 0.1834928 & 0.478613 & 0.2511354 & 0.038505865 \\ 0.0482509 & 0.1834928 & 0.47813 & 0.2511354 & 0.038505865 \end{bmatrix} \qquad (3.4.4)$$

From the 10th order TPM the profit of D2L starts with 26642054-96041282 taka in the first year will be retain in the same range in the next year with probability 0.048, there is 18% chance in increase the profit to 96041283-165440510 taka, 48% likelihood for increasing the profit to 165440511-234839738 taka, 25% possibility to increase in next state and only 4% chance remaining to earn more 304238967-373638195 in the next year.

When the profit of D2Lstarts with 96041283-165440510 millions in the first year it decreases to 26642054-96041282 taka in the next year with probability .048, there is 18% chance to remain the same profit range, 48% likelihood for increasing the profit to 165440510-234839738 taka, 25% possibility to raise in 234839739-304238966 and only 4% chance remaining to earn 304238967-373638197 in the next year.

When the profit of D2L starts with 165440510-234839738 taka in the first year it decreases to 26642054-96041282 taka in the next year with probability .048, there is 18% chance in decreases the profit to 96041283-165440510 taka, there is 48% chance to remain the same profit range, 25% likelihood for increasing the profit to 234839739 -304238966 taka, 4% possibility to raise in 304238967-373638195 in the next year.

When the profit of D2L starts with 234839739-304238966 taka in the first year it decreases to 26642054-96041282 in the next year with probability .048, there is 18% chance to decrease profit to 96041283-165440510, 48% likelihood for decreasing the profit to 165440510-234839738 taka, 25% possibility to remain on the same range and only 4% chance to increase the profit range 304238967-373638197 in the next year.

When the profit of D2Lstarts with more than 304238967-373638195 taka in the first year it decreases to 26642054-96041282 taka in the next year with probability.048, there is 18% chance to decrease profit range 96041283-165440510, 48% likelihood for decreasing the profit to 165440511-234839738, 25% chance to decrease in 234839739-304238966 and there is only 4% chance to remain the profit in the same range.

Again Using R-program (Appendix) we get Transition matrix T for the generated profit of M.net in the first phase simulation (Markov chain) is given below:

$$T = \begin{bmatrix} 0 & 3 & 4 & 6 & 3 \\ 5 & 15 & 16 & 13 & 5 \\ 7 & 15 & 31 & 28 & 8 \\ 4 & 17 & 26 & 19 & 8 \\ 0 & 4 & 12 & 8 & 2 \end{bmatrix} \qquad (3.4.5)$$

As a result, Transition probability matrix (TPM) for D2L using Markov Chain (MC) simulation during 2007-2008 to 2010-2011 is

$$P = \begin{bmatrix} 0.00000000 & 0.1875000 & 0.2500000 & 0.3750000 & 0.18750000 \\ 0.09259259 & 0.2777778 & 0.2962963 & 0.2407407 & 0.09259259 \\ 0.07865169 & 0.1685393 & 0.3483146 & 0.3146067 & 0.08988764 \\ 0.05405405 & 0.2297297 & 0.3513514 & 0.2567568 & 0.10810811 \\ 0.00000000 & 0.1538462 & 0.4615385 & 0.3076923 & 0.07692308 \end{bmatrix} \qquad (3.4.6)$$

From the TPM matrix we observed that the profit of M.net starts with 68356-120895 dollar in the first year will be retain in the same range in the next year with probability 0%, there is 19% chance in increase the profit to 120896-173434 dollar,25% likelihood for increasing the profit to 173435-225973 dollar, 38% possibility to increase in next state , and finally only 18 % Chance to increase the profit range 278513-331049 in the next year.

When the profit of M.net starts with 120896-173434 dollar in the first year it decreases to 68356-120895 taka in the next year with probability .09, there is 28% chance to remain the same profit range, 29% likelihood for increasing the profit to 173435-225973 dollar, 24% possibility to raise in 225974-278512 dollar and only 9% chance remaining to earn 278513-331049in the next year.

When the profit of M.net starts with 173435-225973 dollar in the first year it decreases 68356-120895 dollar in the next year with probability .07, there is 17% chance in decreases the profit to120896-173434, there is 35% chance to remain the same profit range, 31%

likelihood for increasing the profit to225974-278512 dollar, 9% possibility to raise in 278513-331049 dollar in the next year.

When the profit of M.net starts with 225974-278512 dollar in the first year it decreases to 68356-120895 dollar in the next year with probability .054, there is 23% chance to decrease profit to120896-173434 dollar, 35% likelihood for decreasing the profit to 173435-225973 dollar, 26% possibility to remain on the same range and only 11% chance to increase the profit range 278513-331049 dollar in the next year.

When the profit of M.net starts with more than 278513-331049 dollar in the first year it decreases to 68356-120895 dollar in the next year with probability 0, there is 15% chance to decrease profit range120896-173434 dollar, 46% likelihood for decreasing the profit to173435-225973 dollar, 30% chance to decrease in 225974-278512and there is 8% chance to remain the profit in the same range.

$$P^2 = \begin{bmatrix} 0.05729430 & 0.2092130 & 0.3609294 & 0.2777667 & 0.09479664 \\ 0.06203738 & 0.2140095 & 0.3359766 & 0.2851133 & 0.10286318 \\ 0.06000679 & 0.2063718 & 0.3429477 & 0.2880859 & 0.10258784 \\ 0.06278439 & 0.2087821 & 0.3440703 & 0.2853011 & 0.09906205 \\ 0.06717781 & 0.2030428 & 0.3499557 & 0.2489109 & 0.09491281 \end{bmatrix} \quad (3.4.7)$$

From the 2$^{nd}$ order TPM matrix we observed that the profit of M.net starts with 68356-120895 dollar in the first year will be retain in the same range in the next year with probability 0.057, there is 21% chance in increase the profit to 120896-173434 dollar, 36% likelihood for increasing the profit to 173435-225973 dollar, 28% possibility to increase in next state, and finally only 9 % Chance to increase the profit range 278513-331049 in the next year.

When the profit of M.net starts with 120896-173434 dollar in the first year it decreases to 68356-120895 taka in the next year with probability .062, there is 21% chance to remain the same profit range, 34% likelihood for increasing the profit to 173435-225973 dollar, 29% possibility to raise in 225974-278512 dollar and only 10% chance remaining to earn 278513-331049in the next year.

When the profit of M.net starts with 173435-225973 dollar in the first year it decreases 68356-120895 dollar in the next year with probability .06, there is 21% chance in decreases the profit to120896-173434, there is 34% chance to remain the same profit range, 29% likelihood for increasing the profit to225974-278512 dollar, 10% possibility to raise in 278513-331049 dollar in the next year.

When the profit of M.net starts with 225974-278512 dollar in the first year it decreases to 68356-120895 dollar in the next year with probability .062, there is 21% chance to decrease profit to120896-173434 dollar, 34% likelihood for decreasing the profit to 173435-225973 dollar, 29% possibility to remain on the same range and only 10% chance to increase the profit range 278513-331049 dollar in the next year.

When the profit of M.net starts with more than 278513-331049 dollar in the first year it decreases to 68356-120895 dollar in the next year with probability 0.067, there is 20% chance to decrease profit range120896-173434 dollar, 35% likelihood for decreasing the profit to173435-225973 dollar, 25% chance to decrease in 225974-278512and there is 9% chance to remain the profit in the same range.

Similarly we get $P^3, P^4, P^5, P^6, P^7, P^8\ P^9$, and finally $P^{10}$. Where $P^{10}$ is the convergent point.

$$P^{10} = \begin{bmatrix} 0.06177606 & 0.2084942 & 0.3436293 & 0.2857143 & 0.1003861 \\ 0.06177606 & 0.2084942 & 0.3436293 & 0.2857143 & 0.1003861 \\ 0.06177606 & 0.2084942 & 0.3436293 & 0.2857143 & 0.1003861 \\ 0.06177606 & 0.2084942 & 0.3436293 & 0.2857143 & 0.1003861 \\ 0.06177606 & 0.2084942 & 0.3436293 & 0.2857143 & 0.1003861 \end{bmatrix} \qquad (3.4.8)$$

From the 10th order TPM matrix we observed that the profit of M.net starts with 68356-120895 dollar in the first year will be retain in the same range in the next year with probability 0.061, there is 21% chance in increase the profit to 120896-173434 dollar, 34% likelihood for increasing the profit to 173435-225973 dollar, 29% possibility to increase in next state, and finally only 10 % Chance to increase the profit range 278513-331049 in the next year.

When the profit of M.net starts with 120896-173434 dollar in the first year it decreases to 68356-120895 taka in the next year with probability .062, there is 21% chance to remain the same profit range, 34% likelihood for increasing the profit to 173435-225973 dollar, 29%

possibility to raise in 225974-278512 dollar and only 10% chance remaining to earn 278513-331049in the next year.

When the profit of M.net starts with 173435-225973 dollar in the first year it decreases 68356-120895 dollar in the next year with probability .062, there is 21% chance in decreases the profit to120896-173434, there is 34% chance to remain the same profit range, 29% likelihood for increasing the profit to225974-278512 dollar, 10% possibility to raise in 278513-331049 dollar in the next year.

When the profit of M.net starts with 225974-278512 dollar in the first year it decreases to 68356-120895 dollar in the next year with probability .062, there is 21% chance to decrease profit to120896-173434 dollar, 34% likelihood for decreasing the profit to 173435-225973 dollar, 29% possibility to remain on the same range and only 10% chance to increase the profit range 278513-331049 dollar in the next year.

When the profit of M.net starts with more than 278513-331049 dollar in the first year it decreases to 68356-120895 dollar in the next year with probability 0.062, there is 21% chance to decrease profit range120896-173434 dollar, 34% likelihood for decreasing the profit to173435-225973 dollar, 29% chance to decrease in 225974-278512and there is 10% chance to remain the profit in the same range.

### 3.4.2: Markov Chain Estimation using MCMC Simulation (Metropolish Hasting Algorithm)

Before simulating the profit of two companies, we need to classify different states of this variable. Since the minimum profit of D2L is 26642054 taka and maximum is 373638195 taka, whereas the minimum profit is 68356 dollar and maximum profit is 331049 dollar for M.net. So the states ranges are different for two companies.

**Table 3.15** Classification of States in the profit of D2L for Markov Chain Estimation using MCMC Simulation during year2007-2008 to 2010-2011 Markov Chain Estimation using MCMC Simulation

| Generated Profit Range(in taka) | state |
|---|---|
| 44568310-103476151 | 1 |
| 103476152-162383993 | 2 |
| 162383994-221291834 | 3 |
| 221291835-280199675 | 4 |
| 280199676-339107514 | 5 |

**Table 3.16** Classification of States in the profit of M.net for Markov Chain Estimation using MCMC Simulation during year 2006 to 2010 Markov Chain Estimation using MCMC Simulation.

| Generated Profit Range (in dollar) | state |
|---|---|
| 101510.8-143364.36 | 1 |
| 143364.37-185217.92 | 2 |
| 185217.93-227071.48 | 3 |
| 227071.49-268925.04 | 4 |
| 268925.05-310778.6 | 5 |

Using R-program (Appendix) we get Transition matrix T for the generated profit of D2L in the second phase simulation (MCMC) is given below:

$$T = \begin{bmatrix} 0 & 1 & 2 & 2 & 0 \\ 2 & 6 & 13 & 19 & 5 \\ 3 & 17 & 36 & 16 & 7 \\ 0 & 15 & 19 & 15 & 7 \\ 0 & 6 & 9 & 4 & 3 \end{bmatrix} \qquad (3.4.9)$$

As a result, Transition probability matrix (TPM) for D2L using Markov Chain Monte Carlo (MCMC) simulation during 2007-2008 to 2010-2011 is

$$P = \begin{bmatrix} 0.00000000 & 0.2000000 & 0.4000000 & 0.4000000 & 0.00000000 \\ 0.04444444 & 0.1333333 & 0.2888889 & 0.4222222 & 0.11111111 \\ 0.03797468 & 0.2151899 & 0.4556962 & 0.2025316 & 0.08860759 \\ 0.00000000 & 0.2678571 & 0.3392857 & 0.2678571 & 0.12500000 \\ 0.00000000 & 0.2727273 & 0.4090909 & 0.1818182 & 0.13636364 \end{bmatrix} \qquad (3.4.10)$$

From the TPM matrix we observed that the profit of D2L starts with 44568310-103476151taka in the first year will be retain in the same range in the next year with probability 0, there is 20% chance in increase the profit to 103476152-162383993 taka, 40% likelihood for increasing the profit to162383994-221291834, 40% possibility to increase in next state and there is no chance remaining to earn more 280199676-339107514 in the next year.

When the profit of D2Lstarts with 103476152-162383993 taka in the first year it decreases to 44568310-103476151taka in the next year with probability .04, there is 13% chance to remain the same profit range, 29% likelihood for increasing the profit to 162383994 taka, 42% possibility to raise in 221291835-280199675 and only 11% chance remaining to earn 280199676-339107514 taka in the next year.

When the profit of D2L starts with 162383994-221291834 taka in the first year it decreases to 44568310-103476151 taka in the next year with probability .037, there is 22% chance in decreases the profit to 103476152-162383993 taka, there is 45% chance to remain the same profit range, 20% likelihood for increasing the profit to 221291835-280199675 taka, 9% possibility to raise in 280199676-339107514 in the next year.

When the profit of D2L starts with 221291835-280199675 taka in the first year it decreases to 44568310-103476151 in the next year with probability 0, there is 27% chance to decrease profit to103476152-162383993, 34% likelihood for decreasing the profit to 162383994-221291834 taka, 26% possibility to remain on the same range and only 13% chance to increase the profit range 280199676-339107514 taka in the next year.

When the profit of D2Lstarts with more than 280199676-339107514 taka in the first year it decreases to 44568310-103476151 taka in the next year with probability 0, there is 27% chance to decrease profit range103476152-162383993, 41% likelihood for decreasing the profit to162383994-221291834 taka, 18% chance to decrease in 221291835-280199675 and there is 14% chance to remain the profit in the same range.

$$P^2 = \begin{bmatrix} 0.02407876 & 0.2198855 & 0.3757705 & 0.2726000 & 0.1076653 \\ 0.01689639 & 0.2322309 & 0.3766504 & 0.2658805 & 0.1083419 \\ 0.02686891 & 0.2127634 & 0.3899795 & 0.2687007 & 0.1016874 \\ 0.02478903 & 0.2145635 & 0.3740086 & 0.2762861 & 0.1103528 \\ 0.02765631 & 0.2102872 & 0.3826825 & 0.2715001 & 0.1078739 \end{bmatrix} \qquad (3.4.11)$$

From the 2nd order TPM matrix we observed that the profit of D2L starts with 44568310-103476151taka in the first year will be retain in the same range in the next year with probability 0.024, there is 22% chance in increase the profit to 103476152-162383993 taka, 38% likelihood for increasing the profit to162383994-221291834, 27% possibility to increase in next state and there is 11% chance remaining to earn more 280199676-339107514 in the next year.

When the profit of D2Lstarts with 103476152-162383993 taka in the first year it decreases to 44568310-103476151taka in the next year with probability .016, there is 23% chance to remain the same profit range, 38% likelihood for increasing the profit to 162383994 taka, 27% possibility to raise in 221291835-280199675 and only 10% chance remaining to earn 280199676-339107514 taka in the next year.

When the profit of D2L starts with 162383994-221291834 taka in the first year it decreases to 44568310-103476151 taka in the next year with probability .027, there is 21% chance in decreases the profit to 103476152-162383993 taka, there is 39% chance to remain the same profit range, 27% likelihood for increasing the profit to 221291835-280199675 taka, 10% possibility to raise in 280199676-339107514 in the next year.

When the profit of D2L starts with 221291835-280199675 taka in the first year it decreases to 44568310-103476151 in the next year with probability 0.025, there is 21% chance to decrease profit to103476152-162383993, 37% likelihood for decreasing the profit to 162383994-221291834 taka, 28% possibility to remain on the same range and only 11% chance to increase the profit range 280199676-339107514 taka in the next year.

When the profit of D2Lstarts with more than 280199676-339107514 taka in the first year it decreases to 44568310-103476151 taka in the next year with probability 0.028, there is 21% chance to decrease profit range103476152-162383993, 38% likelihood for decreasing the profit to 162383994-221291834 taka, 27% chance to decrease in 221291835-280199675 and there is 10% chance to remain the profit in the same range.

Similarly we get $P^3, P^4, P^5, P^6, P^7, P^8$ $P^9$, and finally $P^{10}$ . Where $P^{10}$ is the convergent point

$$P^{10} = \begin{bmatrix} 0.02415459 & 0.2173913 & 0.3816425 & 0.2705314 & 0.1062802 \\ 0.02415459 & 0.2173913 & 0.3816425 & 0.2705314 & 0.1062802 \\ 0.02415459 & 0.2173913 & 0.3816425 & 0.2705314 & 0.1062802 \\ 0.02415459 & 0.2173913 & 0.3816425 & 0.2705314 & 0.1062802 \\ 0.02415459 & 0.2173913 & 0.3816425 & 0.2705314 & 0.1062802 \end{bmatrix} \quad (3.4.12)$$

From the 10th order TPM matrix we observed that the profit of D2L starts with 44568310-103476151taka in the first year will be retain in the same range in the next year with probability 0.024, there is 22% chance in increase the profit to 103476152-162383993 taka, 38% likelihood for increasing the profit to162383994-221291834, 27% possibility to increase in next state and there is 11% chance remaining to earn more 280199676-339107514 in the next year.

When the profit of D2Lstarts with 103476152-162383993 taka in the first year it decreases to 44568310-103476151taka in the next year with probability .024, there is 22% chance to remain the same profit range, 38% likelihood for increasing the profit to 162383994 taka, 27% possibility to raise in 221291835-280199675 and only 11% chance remaining to earn 280199676-339107514 taka in the next year.

When the profit of D2L starts with 162383994-221291834 taka in the first year it decreases to 44568310-103476151 taka in the next year with probability .024, there is 22% chance in decreases the profit to 103476152-162383993 taka, there is 22% chance to remain the same profit range, 27% likelihood for increasing the profit to 221291835-280199675 taka, 11% possibility to raise in 280199676-339107514 in the next year.

When the profit of D2L starts with 221291835-280199675 taka in the first year it decreases to 44568310-103476151 in the next year with probability 0.024, there is 22% chance to decrease profit to103476152-162383993, 38% likelihood for decreasing the profit to 162383994-221291834 taka, 27% possibility to remain on the same range and only 11% chance to increase the profit range 280199676-339107514 taka in the next year.

When the profit of D2Lstarts with more than 280199676-339107514 taka in the first year it decreases to 44568310-103476151 taka in the next year with probability 0.024, there is 22%

chance to decrease profit range103476152-162383993, 38% likelihood for decreasing the profit to 162383994-221291834 taka, 27% chance to decrease in 221291835-280199675 and there is 11% chance to remain the profit in the same range.

Similarly, Using R-program (Appendix) we get Transition matrix T for the generated profit of M.net in the second phase simulation (MCMC) is given below:

$$T = \begin{bmatrix} 2 & 0 & 7 & 8 & 4 \\ 2 & 8 & 10 & 7 & 2 \\ 5 & 10 & 16 & 16 & 9 \\ 8 & 7 & 17 & 19 & 14 \\ 4 & 5 & 11 & 9 & 2 \end{bmatrix} \tag{3.4.13}$$

As a result, Transition probability matrix (TPM) for M.net using Markov Chain Monte Carlo (MCMC) simulation during 2006t to 2010 is

$$P = \begin{bmatrix} 0.09523810 & 0.0000000 & 0.3333333 & 0.3809524 & 0.19047619 \\ 0.06896552 & 0.2758621 & 0.3448276 & 0.2413793 & 0.06896552 \\ 0.08196721 & 0.1639344 & 0.2622951 & 0.3442623 & 0.14754098 \\ 0.12307692 & 0.1076923 & 0.2615385 & 0.2923077 & 0.21538462 \\ 0.12903226 & 0.1612903 & 0.3548387 & 0.2903226 & 0.06451613 \end{bmatrix} \tag{3.4.14}$$

From the TPM matrix we observed that the profit of M.net starts with 101510.8-143364.36 dollar in the first year will be retain in the same range in the next year with probability .095, there is no chance in increase the profit to 143364.37-185217.92 dollar, 33% likelihood for increasing the profit to 185217.93-227071.48 dollar, 39% possibility to increase in next state, and finally only 19 % Chance to increase the profit range 268925.05-310778 dollar6in the next year.

When the profit of M.net starts with 143364.37-185217.92 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability .068, there is 28% chance to remain the same profit range, 34% likelihood for increasing the profit to 185217.93-227071.48 dollar, 24% possibility to raise in 227071.49-268925.04 dollar and only 7% chance remaining to earn 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with 185217.93-227071.48 dollar in the first year it decreases 101510.8-143364.36 dollar in the next year with probability .08, there is 16% chance in decreases the profit 143364.37-185217.92, there is 26% chance to remain the same profit range, 34% likelihood for increasing the profit 227071.49-268925.04 dollar, 15% possibility to raise in 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with 227071.49-268925.04 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability .123, there is 11% chance to decrease profit to 143364.37-18521.92 dollar, 26% likelihood for decreasing the profit to 185217.93-227071.48 dollar, 29% possibility to remain on the same range and only 21% chance to increase the profit range 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with more than 268925.05-310778 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability 0.129, there is 16% chance to decrease profit range 143364.37-185217.92 dollar, 35% likelihood for decreasing the profit to 185217.93-227071.48 dollar, 29 % chance to decrease in 227071.49-268925.04 and there is 6% chance to remain the profit in the same range.

$$P^2 = \begin{bmatrix} 0.10785672 & 0.1263924 & 0.2863998 & 0.3176901 & 0.1616610 \\ 0.09246467 & 0.1697472 & 0.2961615 & 0.3021504 & 0.1394762 \\ 0.10202011 & 0.1490938 & 0.2950413 & 0.3045593 & 0.1492855 \\ 0.10435414 & 0.1388021 & 0.2996377 & 0.3108938 & 0.1492855 \\ 0.09655407 & 0.1443355 & 0.2905239 & 0.3138389 & 0.1547476 \end{bmatrix} \qquad (3.4.15)$$

From the 2nd order TPM matrix we observed that the profit of M.net starts with 101510.8-143364.36 dollar in the first year will be retain in the same range in the next year with probability .107, there is 13%chance in increase the profit to 143364.37-185217.92 dollar, 29% likelihood for increasing the profit to 185217.93-227071.48 dollar, 32% possibility to increase in next state, and finally only 16 % Chance to increase the profit range 268925.05-310778 dollar6in the next year.

When the profit of M.net starts with 143364.37-185217.92 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability .092, there is 17% chance to remain the same profit range, 30% likelihood for increasing the profit to 185217.93-227071.48 dollar, 30% possibility to raise in 227071.49-268925.04 dollar and only 14% chance remaining to earn 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with 185217.93-227071.48 dollar in the first year it decreases 101510.8-143364.36 dollar in the next year with probability .102, there is 15% chance in decreases the profit 143364.37-185217.92, there is 30% chance to remain the same profit range, 30% likelihood for increasing the profit 227071.49-268925.04 dollar, 15% possibility to raise in 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with 227071.49-268925.04 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability .104, there is 14% chance to decrease profit to 143364.37-18521.92 dollar, 30% likelihood for decreasing the profit to 185217.93-227071.48 dollar, 31% possibility to remain on the same range and only 15% chance to increase the profit range 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with more than 268925.05-310778 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability 0.096, there is 14% chance to decrease profit range 143364.37-185217.92 dollar, 29% likelihood for decreasing the profit to 185217.93-227071.48 dollar, 31% chance to decrease in 227071.49-268925.04 and there is 15% chance to remain the profit in the same range.

Similarly we get $P^3, P^4, P^5, P^6, P^7, P^8\ P^9, P^{10}$ and finally $P^{11}$ . Where $P^{11}$ is the convergent point.

$$P^{11} = \begin{bmatrix} 0.1011224 & 0.1459242 & 0.2950775 & 0.3088748 & 0.149001 \\ 0.1011224 & 0.1459242 & 0.2950775 & 0.3088748 & 0.149001 \\ 0.1011224 & 0.1459242 & 0.2950775 & 0.3088748 & 0.149001 \\ 0.1011224 & 0.1459242 & 0.2950775 & 0.3088748 & 0.149001 \\ 0.1011224 & 0.1459242 & 0.2950775 & 0.3088748 & 0.149001 \end{bmatrix} \qquad (3.4.16)$$

From the $11^{th}$ order TPM matrix we observed that the profit of M.net starts with 101510.8-143364.36 dollar in the first year will be retain in the same range in the next year with probability .10, there is 15%chance in increase the profit to 143364.37-185217.92 dollar, 30% likelihood for increasing the profit to 185217.93-227071.48 dollar, 31% possibility to increase in next state, and finally only 15 % Chance to increase the profit range 268925.05-310778 dollar6in the next year.

When the profit of M.net starts with 143364.37-185217.92 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability .10, there is 15% chance to remain the same profit range, 30% likelihood for increasing the profit to 185217.93-227071.48 dollar, 31% possibility to raise in 227071.49-268925.04 dollar and only 15% chance remaining to earn 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with 185217.93-227071.48 dollar in the first year it decreases 101510.8-143364.36 dollar in the next year with probability .10, there is 15% chance in decreases the profit 143364.37-185217.92, there is 30% chance to remain the same profit range, 31% likelihood for increasing the profit 227071.49-268925.04 dollar, 15% possibility to raise in 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with 227071.49-268925.04 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability .10, there is 15% chance to decrease profit to 143364.37-18521.92 dollar, 30% likelihood for decreasing the profit to 185217.93-227071.48 dollar, 31% possibility to remain on the same range and only 15% chance to increase the profit range 268925.05-310778.6 dollar in the next year.

When the profit of M.net starts with more than 268925.05-310778 dollar in the first year it decreases to 101510.8-143364.36 dollar in the next year with probability 0.10, there is 15% chance to decrease profit range 143364.37-185217.92 dollar, 30% likelihood for decreasing the profit to 185217.93-227071.48 dollar, 31 % chance to decrease in 227071.49-268925.04 and there is 15% chance to remain the profit in the same range.

### 3.4.3 Long Run Probabilities

From section 3.4.1 and 3.4.3 we get the long run probabilities for D2L and M.net using both Mc and MCMC simulation. Which are as follows

**Table 3.17** Long run probabilities by Markov Chain Estimation using Monte Carlo (MC) Simulation for MLM Company D2L

| state | Probability |
|---|---|
| $\pi_1$ | 0.04825019 |
| $\pi_2$ | 0.1834928 |
| $\pi_3$ | 0.47813 |
| $\pi_4$ | 0.2511354 |
| $\pi_5$ | 0.038505865 |

The highest long run probability 47% is associated with the profit range 165440511-234839738 taka then 25% is associated with the profit range 234839739-304238966 taka, followed by 18% with profit 96441283-165440510 taka , 5 % with 26642054-96041282 taka and the minimum probability 4% associated with profit 304238967-373638195 taka.

**Table 3.18** Long run probabilities by Markov Chain Estimation using Monte Carlo Simulation (MC) for MLM Company M.net

| state | Probability |
|---|---|
| $\pi_1$ | 0.06177606 |
| $\pi_2$ | 0.2084942 |
| $\pi_3$ | 0.3436293 |
| $\pi_4$ | 0.2857143 |
| $\pi_5$ | 0.1003861 |

The maximum long run probability 34% is associated with the profit range 173435-225973 dollar, then 28% is associated with the profit range 225974-278512, followed by 21% with profit 120896-173434 dollar, 10% with 278513-331049 dollar and the minimum probability 6% associated with profit 68356-120895 dollar.

**Table 3.19** Long run probabilities by Markov Chain Estimation using MCMC Simulation (Metropolish Hasting Algorithm) for MLM Company D2L

| state | Probability |
|---|---|
| $\pi_1$ | 0.02415459 |
| $\pi_2$ | 0.2173913 |
| $\pi_3$ | 0.3816425 |
| $\pi_4$ | 0.2705314 |
| $\pi_5$ | 0.1062802 |

The highest long run probability 38% is associated with the profit range 162383994-221291834 taka then 27% is associated with the profit range 221291835-280199675 taka, followed by 22% with profit 103476152 -162383993 taka, 10 % with 280199676-339107514 taka and the minimum probability 2% associated with profit 44568310-103476151 taka

**Table 3.20** Long run probabilities by Markov Chain Estimation using MCMC Simulation (Metropolish Hasting Algorithm) for MLM Company D2L for MLM Company M.net

| state | Probability |
|---|---|
| $\pi_1$ | 0.1011224 |
| $\pi_2$ | 0.1459242 |
| $\pi_3$ | 0.2950775 |
| $\pi_4$ | 0.3088748 |
| $\pi_5$ | 0.149001 |

The maximum long run probability 31% is associated with the profit range 227071.49-268925.04 dollar, then 30% is associated with the profit range 185217.93-227071.48dollar, followed by 15% with profit 143364.37-185217.92 and 268925.05-310778.6dollar, and the minimum probability 10% associated with profit 101510.8-143364.36 dollar.

### 3.4.4 Kolmogorov Smirnov goodness of fit test

To ensure about the generated (by M-H algorithm) profit's distribution we use Kolmogorov-Smirnov test.

Let us consider that , our testing hypothesis are as follows

$H_0$:The data follow the normal distribution

$H_a$:The data do not follow the normal distribution

At 5% level of Significance we will test the above hypothesis. We get test statistics value using R-program(Appendix)

**Table 3.21** Two-sample Kolmogorov-Smirnov test for generated profit of D2L and M.net using Metropolis Hasting algorithm

| Company | Test Statistic | p-value | Decision |
|---|---|---|---|
| D2L | .084 | .05873 | The data follow the normal distribution |
| M.net | 0.056 | 0.4131 | The data follow the normal distribution |

### 3.5 Conclusion

In this chapter we see that, mean and variance of both of MLM Companies offspring is 1.29, 1.30 and .4859, .49. Then the ultimate extinction for Destiny 2000 Ltd is .325 and for medsit.net is .31818. We try to represent the graphical representation of down line Clients offspring distribution and hence calculate mean and variance of down line distributors generation to generation. At the next Markov Chain model is also conducted for the annual profit of two MLM companies. We found only 4 years annual profit for Destiny-2000 Ltd and 5 years for Medsit.net. This information is not adequate for estimating Markov chain. So simulation is carried out in two phase. The first phase is Markov Chain simulation and 2$^{nd}$ phase is Markov Chain Monte Carlo (MCMC) simulation. At lastly kolmogorov goodness of fit test have used to check the normality of generated profit by Markov Chain Monte Carlo (MCMC) simulation.

# Chapter-4

# Summary and Conclusion

Multi Level Marketing (MLM) is embracing more and more arenas today. MLM route provides employment opportunities to lakhs of people and enhances their social status. The multiple role of MLM companies can be looked at as a social contribution and these companies or cooperatives are emerging as a development oriented social movement. Most of the people of our country are extremely poor and illiterate. They have not too much knowledge about the investment such as fixed deposit, MLM related business. But they want to build up their bright future. Now a day, for the fulfillment of their desire and dream many of them join with different MLM companies in exchange of some capital. As a developing country, the risk of investment to this type of business is very high for any people of Bangladesh. To find the social and economical impact of such kind of business we can use stochastic models such as Branching Process and Markov chain. For Modeling of MLM Companies activities such as future profit and number of Clients involvement stochastic models help us. To explain our theoretical knowledge of stochastic modeling we consider two MLM Companies data: Destiny-2000 Ltd and Medsit.net.

In the Univariate and Graphical representation part of this book we observed that for MLM Company D2L of generated profit by MCMC is less compare to generated Profit by MC and original profit but standard deviation is greater than compare to others. Similarly for MLM Company M.net mean and standard deviation of generated profit by MCMC is less compare to generated Profit by MC and original profit. In this correspondence generated Profit by MCMC

is more suitable to construct Stochastic models.

In this study, branching process is introduced and evaluated to get idea of the distribution of down line distributors. This model is also used in characterizing the distribution of the size of the population for different generations and the probability of extinction of the population. The down line distributor's mean is 1.29 and variance is 0.4859 of Destiny-2000 Ltd. whereas the down line distributor's mean is 1.30 and variance is .49 of Medsit.net. So the population is supercritical for destiny-2000 as well as Medsit.net. The ultimate probability of extinction is .325 for destiny-2000 and .3181818 for Medsit.net. As a result, people's investment in destiny-2000 has less risk than that of Medsit.net. The first generation distributor's have a mean 1.29 with variance .4859 for Destiny 2000 Ltd. Then second generation distributor's mean 1.67 with variance 1.435 for destiny 2000 Ltd. In this way up to 20$^{th}$ generation distributor's mean 162.85 with variance 34235.17. Similarly The first generation distributor's have a mean 1.30 with variance .49 for Medsit.net. Then second generation distributor's mean 1.69 with variance 1.4651 for Medsit.net. In this way up to 20$^{th}$ generation distributor's mean 190.05 with variance 45141.33. Comparing both mean and variance of generation to generation Destiny 2000 Ltd has less mean and variance. So this company is more preferable than others for next generation.

The Markov Chain model is also conducted for the annual profit of two MLM companies. We found only 4 years annual profit for Destiny-2000 Ltd and 5 years for Medsit.net. This information is not adequate for estimating Markov chain. So simulation is carried out in two phase. The first phase is Markov Chain simulation and 2$^{nd}$ phase is Markov Chain Monte Carlo (MCMC) simulation.

In the first phase for company D2L, we found that the maximum probability (60%) associated with the profit range 165440511-234839738 taka whereas the minimum probability is .000 to remaining the profit with 26642054-96041282 taka in one year transition. The highest long run probability 47% is associated with the profit range 165440511-234839738 taka then 25% is associated with the profit range 234839739-304238966 taka, followed by 18% with profit 96441283-165440510 taka , 5 % with 26642054-96041282 taka and the minimum probability 4% associated with profit 304238967-373638195 taka. For company Medsit.net we found that the maximum probability (46%) associated with the profit range 173435-225973 dollar whereas the minimum probability is .000 to remaining the profit with 68356-120895 dollar in one year transition. The maximum long run probability 34% is associated with the profit range 173435-225973 dollar, then 28% is associated with the profit range 225974-278512, followed by 21% with profit 120896-173434 dollar, 10% with 278513-331049 dollar and the minimum probability 6% associated with profit 68356-120895 dollar.

In the 2$^{nd}$ phase for MLM Company D2L, we found that the maximum probability (45%) associated with the profit range 162383994-221291834 taka whereas the minimum probability is .000 to remaining the profit with 221291835-280199675 or more taka in one year transition. The highest long run probability 38% is associated with the profit range

162383994-221291834 taka then 27% is associated with the profit range 221291835-280199675 taka, followed by 22% with profit 103476152 -162383993 taka, 10 % with 280199676-339107514 taka and the minimum probability 2% associated with profit 44568310-103476151 taka. Similarly for MLM Company M.net, We see that that the maximum probability (38%) associated with the profit range 227071.49-268925.04 dollar whereas the minimum probability is .000 to remaining the profit with143364.37-185217.92 in one year transition. The maximum long run probability 31% is associated with the profit range 227071.49-268925.04 dollar, then 30% is associated with the profit range 185217.93-227071.48dollar, followed by 15% with profit 143364.37-185217.92 and 268925.05-310778.6dollar, and the minimum probability 10% associated with profit 101510.8-143364.36 dollar. For surety the distribution pattern of generated profit of MLM companies into 2$^{nd}$ phase shows that, distribution of both is normal.

.

This research contributes significantly in both understanding of economic aspects of MLM companies and analytical capability of stochastic modeling.

## References:

Athanasios Papoulis, S. Unnikrishna Pillai (2002), "*Probability, Random Variables and StochasticProcesses*", 4th ed., Tata McGraw-Hill Publishing company Ltd, New Delhi.

*Bangladeshimlmnews.blogspot.com/.../list-of-best-mlm-companies-in*

Barbu, V.S, Limnios, N. (2008) Semi-Markov Chains and Hidden Semi-Markov Models toward Applications: Their Use in Reliability and DNA Analysis. ISBN 978-0387731711

Bhat, B.R. (2000), "*Stochastic Models*", New Age International (p) Ltd. Publishers, New Delhi.

Bhat, U.N.(1984), "*Elements of Applied Stochastic Processes*", 2nd edition, John Wiley and Sons, New York.

Bruss, F. Thomas (1984): "A Note on Extinction Criteria for Bisexual Galton– Watson Processes". *Journal of Applied Probability* 21: 915–919.

Chakravart, Laha, and Roy 1967, A note on Kolmogorov-Smirnov Goodness-of-Fit Test. Journal of Statistics.

Daneil, Wayne W. (1978): *Applied Non-parametric Statistics,* Houghton Mifflin Company, Boston Dallas, Geneva.

Dodge, Y. The Oxford Dictionary of Statistical Terms, OUP. ISBN 0-19- 920613-9.

*En.wikipedia.org/wiki/Multi-level_marketing*

Everitt, B.S.(2002): The Cambridge Dictionary of Statistics. CUP. ISBN 0-521-8109 -x.

F. Galton, *Problem 4001, Educational Times*, 1 April ,1873 P.17.

Gujarati, D.N. (2003), "*Basic Econometrics*", 4th ed., McGraw Hill, New York.

Hamdy A. Taha(1999), "*Operations Research An Introduction*", 6th ed. Prentice Hall of India (p) Ltd, New Delhi.

Hines, William W.; Douglas C. Montgomery; David M. Goldsman and Connine M.Borror (2003): *Probability and Statistics in Engineering,* John Wiley and Sons Inc.

http://www.multi-level-marketing-website.com/destiny-2000.html

Jagers, P. (1975), "*Branching Process with Biological Applications*" , Wiley.

Latouche, G. and V. Ramaswami (1999): *Introduction to Matrix Analytic Methods in Modelling, 1st edition*.

M.K.Roy (2004), "Fundamentals of Probability & Probability Distributions", 3rd ed.

Markov process, (mathematics)-Britannica online Encyclopedia.

Mason, R.D & Lind, D.A, "*Statistical Techniques in Business & Economics*", 12th ed., McGraw Hill, New York.

Medhi, j.(1994), " *Stochastic Processes*",2nd edition, New Age International (P) Ltd. Publisher, New Delhi.

*Mlminbangladesh.blogspot.com*.

Peskun , P. (1973) Optimum Monte Carlo sampling using Markov Chains, Biometrica 60.

Ross, S.M. (1995) Stochastic Processes. Wiley. ISBN 978-0471120629.

Ross, S. (2003), " *Probability Models*", 8th edition, Academic Press, USA.

S.C. Gupta, V.K. Kapoor (2008), "*Fundamentals of Mathematical Statistics*", 12th ed., Sultan Chand & Sons, New Delhi.

Schlechter, Kira (Monday March 02, 2009). "Hershey Medical Center to open redesigned . *The Patriot-News*.

Suddhendu Biswas (2004), “*Applied Stochastic Process*”, 2nd ed. New central book agency (p) Ltd, New Delhi.

Virtamo, J.,"Birth-death processes "[1], 38.3143 Queueing Theory.

William W.Hines, Douglas C. Montgomery, David M. Goldsman, Connie M.Borror (2004), “*Probability and Statistics in Engineering*”, 4th ed., John Wiley & Sons, Inc.

# Appendix

**Program 1:**

The R- Program of Transition probability matrix and long run probabilities by Markov Chain for MLM Company D2L are as follows:

```
Set.seed=558
m =208
d=5
ww=0
l=10
mu=207523188
sdd=63898878.33
z=rnorm(m)
u=z*sdd+mu
xx=matrix(round(u),m,1)
x= matrix(0,m,1)
for (jj in 1: m) {
x[jj,]=ifelse(xx[jj,]<96041282,1,ifelse(xx[jj,]<165440510,2,ifelse(xx[jj,]<234839738,
3,ifelse(xx[jj,]<304238966,4,5))))
}
t= matrix(0,d,d)
pp=matrix(0,d,d)
for (i in 1:m-1) {
t[x[i,],x[i+1,]]=t[x[i,],x[i+1,]]+1
}
w=rowSums(t)
for (j in 1:d) {
  for(k in 1:d) {
   pp[j,k]=t[j,k]/w[j]
}}print(pp)
r=diag(1,d,d)
for (kk in 1:l){
```

```
q=r%*%pp
ww=ifelse(identical(q,r), kk,0)
print(ww)
r=r%*%pp
print(r)
}
```

## Program 2:

The R- Program of Transition probability matrix and long run probabilities by Markov Chain (MC) for MLM Company M.net are as follows:

```
Set.seed=558
m =260
d=5
ww=0
l=10
mu=209931.8
sdd=53547.86
z=rnorm(m)
u=z*sdd+mu
xx=matrix(round(u),m,1)
x=matrix(0,m,1)
for (jj in 1: m) {
x[jj,]=ifelse(xx[jj,]<120895,1,ifelse(xx[jj,]<17434,2,ifelse(xx[jj,]<225973,3,ifelse(xx[
jj,]<278512,4,5))))
}
t=matrix(0,d,d)
pp=matrix(0,d,d)
for (i in 1:m-1) {
t[x[i,],x[i+1,]]=t[x[i,],x[i+1,]]+1
}
w=rowSums(t)
for (j in 1:d) {
  for(k in 1:d) {
```

```
  pp[j,k]=t[j,k]/w[j]
}}
print(pp)
r=diag(1,d,d)
for (kk in 1:l){
q=r%*%pp
ww=ifelse(identical(q,r), kk,0)
print(ww)
r=r%*%pp
print(r)
}
```

## Program-3:

The R- Program of Transition probability matrix and long run probabilities by MCMC (Metropolish Hasting algorithm) for MLM Company D2L are as follows:

```
Set.seed=558
sig=1
n=208
x=rep(rnorm(1),n)
for(i in 2:n){
y=rnorm(1)*sig+x[i-1]
u=runif(1)
alpha=dnorm(y,0,1)/dnorm(x[i-1],0,1)

x[i]=x[i-1]+(y-x[i-1])*(u<alpha)
}
ks.test(jitter(x),rnorm(n,0,1))
ss=63898878.33
mu=207523188

w=x*ss+mu
ks.test(jitter(w),rnorm(n,mu,ss))
```

```
###tpm from hasting##
m=208
d=5
ww=0
l=15
mu=207523188
sdd=63898878.33
z=rnorm(m)
u=z*sdd+mu
xx=matrix(round(u),m,1)
x=matrix(0,m,1)
for (jj in 1: m) {
x[jj,]=ifelse(xx[jj,]<103476151,1,ifelse(xx[jj,]<162383992,2,ifelse(xx[jj,]<22129188
3,3,ifelse(xx[jj,]<280199674,4,5))))
}
t=matrix(0,d,d)
pp=matrix(0,d,d)
for (i in 1:m-1) {
t[x[i,],x[i+1,]]=t[x[i,],x[i+1,]]+1
}
w=rowSums(t)
for (j in 1:d) {
  for(k in 1:d) {
   pp[j,k]=t[j,k]/w[j]
}}
print(pp)
r=diag(1,d,d)
for (kk in 1:l){
q=r%*%pp
ww=ifelse(identical(q,r), kk,0)
print(ww)
r=r%*%pp
```

**Program 4:**

The R- Program of Transition probability matrix and long run probabilities by MCMC (Metropolish-Hasting) for MLM Company M.net are as follows:

```
###hasting for company M.net##
set.seed=558
sig=1
n=260
x=rep(rnorm(1),n)
for(i in 2:n){
y=rnorm(1)*sig+x[i-1]
u=runif(1)
alpha=dnorm(y,0,1)/dnorm(x[i-1],0,1)
x[i]=x[i-1]+(y-x[i-1])*(u<alpha)
}
ks.test(jitter(x),rnorm(n,0,1))
ss=53547.85563
mu=209931.8
w=x*ss+mu
ks.test(jitter(w),rnorm(n,mu,ss))

###tpm from hasting##
m=260
d=5
ww=0
l=15
mu=209931.8
sdd=53547.85563
z=rnorm(m)
u=z*sdd+mu
xx=matrix(round(u),m,1)
x=matrix(0,m,1)
for (jj in 1: m) {
```

```
x[jj,]=ifelse(xx[jj,]<143364.36,1,ifelse(xx[jj,]<185217.92,2,ifelse(xx[jj,]<227071.48,
3,ifelse(xx[jj,]<268925.04,4,5))))
}
t=matrix(0,d,d)
pp=matrix(0,d,d)
for (i in 1:m-1) {
t[x[i,],x[i+1,]]=t[x[i,],x[i+1,]]+1
}
w=rowSums(t)
for (j in 1:d) {
  for(k in 1:d) {
   pp[j,k]=t[j,k]/w[j]
}}
print(pp)
r=diag(1,d,d)
for (kk in 1:l){
q=r%*%pp
ww=ifelse(identical(q,r), kk,0)
print(ww)
r=r%*%pp
print(r)
}
```

**Program 5:**

Kolmogorov smirnov goodness of fit test for MLM Company D2L:

```
set.seed=558
 sig=1
 n=500
 x=rep(rnorm(1),n)
 for(i in 2:n){
 y=rnorm(1)*sig+x[i-1]
 u=runif(1)
 alpha=dnorm(y,0,1)/dnorm(x[i-1],0,1)
```

```
x[i]=x[i-1]+(y-x[i-1])*(u<alpha)
}
ks.test(jitter(x),rnorm(n,0,1))
```

## Program 6:

Kolmogorov smirnov goodness of fit test for MLM company M.net:

```
set.seed=558
sig=1
n=500
x=rep(rnorm(1),n)
for(i in 2:n){
y=rnorm(1)*sig+x[i-1]
u=runif(1)
alpha=dnorm(y,0,1)/dnorm(x[i-1],0,1)

x[i]=x[i-1]+(y-x[i-1])*(u<alpha)
}
ks.test(jitter(x),rnorm(n,0,1))
```