AN INVESTIGATION ON THE ROLE OF GENERALIZED LOGISTIC DISTRIBUTION IN EXTREME VALUE THEORY

Thesis submitted to the University of Calicut
for the degree of
DOCTOR OF PHILOSOPHY
in
STATISTICS

By

NIDHIN, K.

Under the Supervision of

Prof. C. CHANDRAN

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT
KERALA, INDIA.

October 2009

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALICUT

Dr. C. Chandran
Professor

Calicut University, P.O.
Kerala, India

CERTIFICATE

This is to certify that the work reported in this thesis entitled **"An Investigation on the Role of Generalized Logistic Distribution in Extreme Value Theory"** submitted to the University of Calicut for the award of the degree of Doctor of Philosophy in Statistics, under the Faculty of Science, is a bonafide research work carried out by Mr. Nidhin, K., under my supervision and guidance in the Department of Statistics, University of Calicut. The results embodied in this thesis have not been included in any other thesis submitted previously for the award of any degree or diploma.

University of Calicut
23, October 2009

Prof. C. Chandran.

DECLARATION

I hereby declare that the matter embodied in this thesis is the result of investigations carried out by me in the Department of Statistics, University of Calicut, under the supervision and guidance of Dr. C. Chandran, Professor, Department of Statistics, University of Calicut. This thesis contains no material, which has been accepted for the award of any degree or diploma in any university or institute and to the best of my knowledge and belief; it contains no material previously published by any other person, except where the due references are made in the text of the thesis.

University of Calicut
23 October 2009

Nidhin. K.

ACKNOWLEDGEMENT

I am deeply indebted to my Supervisor, **Prof. C. Chandran** for his inspiring and invaluable guidance throughout my research work. I would like to thank him for all the advice, encouragement, help and everything that I learnt from him during this period. Without him this thesis would not have been possible. Besides my advisor, I would like to thank the rest of my teachers: **Dr. K. Jayakumar**, **Prof. N. Raju** and **Prof. M. Manoharan** for their encouragement, insightful comments, and hard questions. Also, I wish to thank **Dr. Haritha N Haridas** and **Ms. Juviriyya** for their support.

I express my sincere thanks to the University Grants Commission of India, for providing Research Fellowship during the entire period of this study.

I would like to specially thank **Prof. J.R.M. Hosking** (IBM Research Division, USA) for his kind interest in Chapters 4 and 5 of this thesis and his encouragement throughout the data analysis part of this work. It was a pleasure in learning the basics of L-moments from him especially for computation and applications of L-moments.

I would like to thank **Sri. S. Suthagaran** and **Sri. M.V. Shantakumar** Librarians of the department for their constant help during this work. I also thank all administrative staff of the department of statistics for providing help in many ways during my stay.

I am very lucky to have **Sri. K. Radakrishnan Nair**, **Sri. K.E. Rajasekaran**, **Sri. V. M. Chacko**, **Sri. V. Jilesh**, **Sri. Cyriac George** and **Sri. Joby K Jose** as my research colleagues. I would like to thank all of them for the support I have received.

I would like to thank the research scholars **Libeesh** and **Ratheesh** of the department of Mathematics and Noufal and Shinoj of the department of English for their help during my research period.

My deep gratitude for my parents and my brother for being with me all the time.

Nidhin, K.

Contents

Chapter 1

Classical Extreme Value Theory - An Introduction

1.1 Introduction

Asymptotic theory of functions of random variables plays a very important role in modern statistics. The objective of the asymptotic theory is to approximate distributions of large sample statistics with limiting distributions which are often much simpler to work than their exact distributions. For example, if $\{X_n, n \geq 1\}$ is a sequence of independent and identically distributed random variables (i.i.d.r.v.'s) with X_1 distributed like F and $E(X_1) = \mu$, exists, then the asymptotic theory of partial sum, $S_n = \sum_{i=1}^{n} X_i$ and partial maximum $M_n = \max(X_1, X_2, \ldots, X_n),\ n \geq 1$ play a fundamental role in statistics. From the well known theories of laws of large numbers, and the central limit theorem available in probability literature one can handle the statistics S_n, the partial sum or $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, the arithmetic mean, very effectively. That is, by weak law of large numbers $\bar{X}_n \xrightarrow{p} \mu$ and by strong law of large numbers $\bar{X}_n \stackrel{a.s.}{\to} \mu$. The central result in statistics, the central limit theorem, says that under proper normalization the partial sum sequence $\{S_n\}$ or $\bar{X}_n$ converges weakly to Levy-skew stable distributions. That is, if there exists real sequences of constants say $\{a_n > 0\}$ and $\{b_n\}$ such that

$$P(\frac{S_n - b_n}{a_n} \leq x) \quad \stackrel{w}{\to} \quad H(x) \tag{1.1.1}$$

for some non-degenerate d.f. H, then H can be identified as Levy-skew stable distribution for which normal distribution is a special case, see Laha Rohatgi (1979) and Billingsly (1968) for more details. Similarly, the limit theory of partial maxima $M_n = \max(X_1, X_2, \ldots, X_n)$ of i.i.d.r.v.'s is also available in the literature. Like the partial sum S_n, there is no unique asymptotic model for M_n, but several models exists. In the next section, we introduce various asymptotic models for M_n which are available in the statistics literature.

1.2 Classical Extreme Value Models

The study of probabilistic and statistical aspects of partial maxima M_n, partial minima $m_n = \min(X_1, X_2, \ldots, X_n)$, and other related order statistics of sequence $\{X_n\}$ of i.i.d.r.v.'s is known as the classical extreme value theory. The order statistics M_n or m_n are commonly referred as extremes in extreme value theory. Classical extreme value theory is well developed and a number of books are available in the area, see for example, Gumbel (1958), Galambos (1978), Leadbetter et al. (1983), Resnick (1987), Embrechts et al. (1997), Reiss et al. (2004), de Haan and Ferreira (2006) etc. In this section, we introduce various asymptotic models available in the classical extreme value theory.

1.2.1 The Generalized Extreme Value (GEV) Model

Let $\{X_n\}$ be a sequence of i.i.d.r.v.'s and X_1 is distributed like F. It can easily be verified that $M_n \xrightarrow{p} x_F$, where $x_F = sup\{x \in R, F(x) < 1\}$ is the right end point of the d.f. F. Further $\{M_n, n \geq 1\}$ is an increasing sequence, hence, $M_n \xrightarrow{a.s.} x_F$. This is similar to what we observed for partial sum sequence $\{S_n\}$ in the beginning of this chapter. Hence, under proper normalization, if a limit distribution can be derived for M_n as in the case of the central limit problem, that limit distribution can be used as an approximate model for M_n. Fisher and Tippet (1928) proved that if a limit distribution exists for maximum under proper normalization then it is one among the three classes of distributions known as the extreme value distributions (EVD). The theorem is known as the Extremal Types Theorem (ETT) or the Fisher-Tipett theorem in the statistics literature, which we state below. For the latest available detailed proof see Leadbetter et al. (1983).

Theorem 1.2.1 *(**Extremal Types Theorem**): Let $\{X_n,\ n \geq 1\}$ be a sequence of i.i.d.r.v.'s and $M_n = \max(X_1, X_2, \ldots, X_n)$. If there exist sequences of norming constants $\{a_n > 0\}, \{b_n\}$, and a non-degenerate d.f. G such that,*

$$P\{\frac{M_n - b_n}{a_n} \leq x\} \ \xrightarrow{w}\ G(x), \tag{1.2.2}$$

then G, in equation 1.2.2, is one among the following three classes of distributions:

$$\textit{Type I}: \qquad G_1(x) = \exp(-\exp(-x)), -\infty < x < \infty$$

$$\textit{Type II}: \qquad G_2(x) = \begin{cases} 0 & \textit{if} \quad x \leq 0, \\ \exp(-x^{-\alpha}) & \textit{if} \quad x > 0,\ \alpha > 0\ . \end{cases}$$

$$\textit{Type III}: \qquad G_3(x) = \begin{cases} \exp(-(-x)^{\alpha}) & \textit{if} \quad x \leq 0, \\ 0 & \textit{if} \quad x > 0,\ \alpha > 0\ . \end{cases}$$

which are the Type I, Type II, and Type III extreme value distributions.

The Type I, Type II, and Type III extreme value distributions are also known as the Gumbel distribution, the Frechet distribution and the reverse-Weibull distribution respectively in the statistics literature. The above three types of distributions are the standard extreme value distributions. The corresponding location-scale family can be derived by replacing x in Theorem 1.2.1 by $\frac{x-\mu}{\sigma}$, which are given below.

$$G_1(x) = \exp(-\exp(-[\frac{x-\mu}{\sigma}])) - \infty < x < \infty$$

$$G_2(x) = \begin{cases} 0 & \text{if} \quad \frac{x-\mu}{\sigma} \le 0, \\ \exp(-[\frac{x-\mu}{\sigma}]^{-\alpha}) & \text{if} \quad \frac{x-\mu}{\sigma} > 0,\ \alpha > 0\ . \end{cases}$$

$$G_3(x) = \begin{cases} \exp(-(-[\frac{x-\mu}{\sigma}])^{\alpha}) & \text{if} \quad \frac{x-\mu}{\sigma} \le 0, \\ 0 & \text{if} \quad \frac{x-\mu}{\sigma} > 0,\ \alpha > 0\ . \end{cases}$$

The Extremal Types Theorem suggests one of the three probable asymptotic models for maxima of large data set, and is extensively used in many real life situations (see for example Gumbel (1958), Kotz and Nadarajah(2000)). The disadvantage of this model is that it is not easy to identify which one of these three distributions is a good model for a given data. Hence, one need to fit these three models and find the best model, if it exists. To overcome this difficulty, von-Mises and Jenkinson suggested an alternative approach by incorporating the three extreme value distributions into a single class by introducing a shape parameter in the model. The new model is known as the generalized extreme value (GEV) distribution. That is, the extremal types theorem suggests GEV distribution as a model for the maximum of sequence $\{X_n\}$ of i.i.d.r.v.'s. In this thesis we denote this model as GEV(max), we define GEV(max) model below.

Definition 1.2.1 *A random variable X is said to follow the Generalized Extreme*

Value distribution (GEV(max)) if its d.f. $G_\eta(x)$ *is given by,*

$$G_\eta(x) = \begin{cases} exp\{-(1+\eta[\frac{x-\mu}{\sigma}])^{-1/\eta}\} & if \quad \eta \neq 0 \\ exp\{-exp(-[\frac{x-\mu}{\sigma}])\} & if \quad \eta = 0 \end{cases}$$

where $1+\eta[\frac{x-\mu}{\sigma}] > 0$. *So the supports are*

$$\begin{cases} \frac{x-\mu}{\sigma} > -\eta^{-1} & for \quad \eta > 0 \\ \frac{x-\mu}{\sigma} < -\eta^{-1} & for \quad \eta < 0 \\ \frac{x-\mu}{\sigma} \in R & for \quad \eta = 0. \end{cases}$$

The shape parameter η is known as the extreme value index. For $\eta = 0$, the GEV distribution becomes Type I extreme value distribution; for $\eta < 0$, it is the Type II extreme value distribution; and for $\eta > 0$, it is the Type III extreme value distribution given in Theorem 1.2.1. The proof of Theorem 1.2.1 consists of two important stages. First stage is to identify the limit distribution G of the normalized maxima in equation 1.2.2 as a class having a stability property called max-stability. The second stage consists of identifying this class exactly as the class of GEV(max) distributions. The property of max-stability will be used quite often in the forgoing chapters of this thesis, hence, we define it below.

Definition 1.2.2 *A non-degenerate random variable* X *is said to be max-stable if, for each* $n = 2, 3, \ldots$ *there are constants* $a_n > 0$ *and* b_n *such that*

$$\max(X_1, X_2, \ldots, X_n) \stackrel{d}{=} (a_n X + b_n).$$

where $X_1, X_2, \ldots, X_n$ *are same copies of* X.

Example 1.2.1 *Let* X *be a random variable which follows standard Gumbel distribution given by,*

$$G(x) = \exp\{-\exp(-x)\}, -\infty < x < \infty.$$

Then X is max-stable since there are constants $a_n = 1$ and $b_n = \log n$ such that, for every n,

$$\max(X_1, X_2, \ldots, X_n) \stackrel{d}{=} (a_n X + b_n).$$

where $X_1, X_2, \ldots, X_n$ are same copies of X.

The GEV(max) model can be fitted to a given data by first dividing data set in to number of blocks of appropriate sizes (for example daily, weekly, monthly, yearly, etc.) and calculate the maximum for each block. To the new sequence of maxima the GEV(max) distribution can be fitted. The fitted model can be used for prediction or any other statistical purposes. The model fitted has the disadvantage that most of the observation in the data are discarded, and only one observation is considered in each block. Hence, data is not properly used for fitting GEV(max) model. To overcome this disadvantage, another alternative model is suggested in the literature which is known as the rth largest order statistics model. In the next subsection we introduce this model.

1.2.2 The r Largest Order Statistics Model

This model is an extension of the maximum approach, idea is to use the r largest observations in each block, where $r > 1$. The mathematical result on which this lies is that, equation 1.2.2 can be easily extended to the joint distribution of the r largest order statistics, as $n \to \infty$ for a fixed $r > 1$, and this may therefore be used as a basis of statistical study. The following theorem describes the mathematical model.

Theorem 1.2.2 *Let $X_{1:n} \geq X_{2:n} \geq \cdots \geq X_{r:n} \geq \cdots \geq X_{n:n}$ be order statistics of i.i.d sample of size n, $a_n > 0$ and b_n be normalizing constants in equation 1.2.2, then*

$$(a_n[X_{1:n} - b_n], \ldots, a_n[X_{r:n} - b_n])$$

converges in distribution to a limiting random vector $(Y_1, Y_2, \ldots, Y_r)$, *whose density is*

$$h(y_1, \ldots, y_r) = \sigma^{-r}.\exp\{-(1+\eta\frac{y_r-\mu}{\sigma})^{-1/\eta} \; - (1+\frac{1}{\eta})\sum_{j=1} r\log(1+\eta\frac{y_j-\mu}{\sigma})\}.$$

For more details about this result and proof, see Leadbetter et al. (1983). To fit this model to a given data, first divide the data into blocks of appropriate sizes (for example daily, weekly, monthly, yearly, etc.) and calculate the maximum, 2^{nd} maximum, ..., r^{th} maximum for each block. To the new sequence of r dimensional vector maxima the distribution given in Theorem 1.2.2 can be fitted. The fitted model can be used for prediction or any other statistical purposes. Such models are extensively used in statistics. For example see Smith (1986) and Tawn (1988) on hydrological extremes, Robinson and Tawn (1995) and Smith (1997) for a novel application to the analysis of athletic records. The model for extremes suggested in Thereom 1.2.2 has the disadvantage that choosing r maximum in each block is not required or there are more large values in some blocks. Hence, data is not properly used for fitting this model as well. Another alternative model suggested in the litteratute for extreme to overcome this disadvantage is to fix a threshold 'u' and take all the observations above this threshold, which we describe in the next subsection.

1.2.3 The Peak Over Threshold (POT) Model

The theoretical aspects of this model have been suggested by two independent works namely Belkama and de Haan (1974) and Pickands (1975). The model can be explained in the following theorem.

Theorem 1.2.3 *let* $\{X_n, n \geq 1\}$ *be a sequence of i.i.d.r.v.'s with common continuous distribution function* $F(x)$. *Suppose there exists a pair of sequences* $\{a_n\}$ *and* $\{b_n\}$

with $a_n > 0$ for all n and a non-degenerate distribution function $G(x)$ such that

$$\lim_{n\to\infty} P\{(M_n - b_n)/a_n \le x\} = \lim_{n\to\infty} F^n(a_n x + b_n) = G(x) \tag{1.2.3}$$

for all x at which $G(x)$ is continuous. Let

$$P_u(x) = P\{X > u + x | X > u\} = \frac{1 - F(u+x)}{1 - F(u)}, \tag{1.2.4}$$

Then, under proper normalization, $P_u(x)$ can be approximated by generalized Pareto (GP) distribution, when $u \to x_F$. Also let $\{X_n, n \ge 1\}$ be a sequence of i.i.d.r.v.'s with common continuous distribution function $F(x)$. Suppose the observations above exceedances of a level u has generalized pareto as a limit distribution. Then there exist a pair of sequence $\{a_n\}$ and $\{b_n\}$ with $a_n > 0$ for all n such that $\frac{M_n - b_n}{a_n}$ converges to a random variable with GEV distribution.

For proof, see Pickands (1975). Hence, GEV (max) model and generalized Pareto model are two equivalent models for extremes. The generalized Pareto (GP) distribution is given below.

Definition 1.2.3 *A random variable X is said to follow the generalized Pareto (GP) distribution if its distribution function is given by,*

$$Q_\eta(x) = \begin{cases} 1 - \{(1 - \eta x)^{1/\eta}\} & \textit{if} \quad \eta \neq 0 \\ 1 - \exp\{(-x)\} & \textit{if} \quad \eta = 0. \end{cases}$$

and the supports are

$$\begin{cases} x > \eta^{-1} & \textit{for} \quad \eta < 0 \\ x < \eta^{-1} & \textit{for} \quad \eta > 0 \\ x \in R & \textit{for} \quad \eta = 0. \end{cases}$$

where η is the shape parameter.

One can extend this class by adding location and scale parameter. Like the GEV(max) distribution the GP distribution possesses a characterizing property of stability called POT stability w.r.t non-random sample size. Below we define POT stability for non-random sample sizes.

Definition 1.2.4 *A non-degenerate random variable X is said to be POT-stable if, for each $u > 0$ there are constants $a_u > 0$ and b_u such that*

$$(X - u)/X > u \stackrel{d}{=} (a_u X + b_u).$$

To fit the POT model or the GP model to a given data, first fix a threshold u and take all the observations which are greater than u. To the new sequence of observations, which are greater than u, the GP distribution can be fitted. The fitted model can be used for prediction or any other statistical purposes. A good application of GP model can be seen in Smith (2001) for air pollution data.

Since $\min(X_1, X_2, \cdots, X_n) = -\max(-X_1, -X_2, \cdots, -X_n)$ the asymptotic distribution for minimum can be derived from the limit distribution of maximum as $1 - G(-x)$, where G is the GEV(max) distribution. The following theorem describes the mathematical model.

Theorem 1.2.4 *Let $\{X_n,\ n \geq 1\}$ be a sequence of i.i.d.r.v.'s and $m_n = \min(X_1, X_2, \ldots, X_n)$. If there exist sequences of norming constants $\{a_n > 0\}, \{b_n\}$, and a non-degenerate d.f. G such that,*

$$P\{\frac{m_n - b_n}{a_n} \leq x\} \stackrel{w}{\to} G(x), \tag{1.2.5}$$

then G is one among the following three classes of distributions:

$$G_1(x) = 1 - \exp(-\exp([\frac{x - \mu}{\sigma}])), -\infty < x < \infty$$

$$G_2(x) = \begin{cases} 0 & \text{if} \quad \frac{x-\mu}{\sigma} \leq 0, \\ 1 - \exp([\frac{x-\mu}{\sigma}]^{-\alpha}) & \text{if} \quad \frac{x-\mu}{\sigma} > 0,\ \alpha > 0\ . \end{cases}$$

$$G_3(x) = \begin{cases} 1 - \exp(-([\frac{x-\mu}{\sigma}])^{\alpha}) & \text{if} \quad \frac{x-\mu}{\sigma} \leq 0, \\ 0 & \text{if} \quad \frac{x-\mu}{\sigma} > 0,\ \alpha > 0\ . \end{cases}$$

As in the case of maximum these three distributions can be incorporated into a single class namely the Generalized extreme value distribution for minima, which we define below.

Definition 1.2.5 *A random variable X is said to follow generalized extreme value distribution of minima (GEV(min)) if its d.f. $G_\eta(x)$ is given by*

$$H_\eta(x) = \begin{cases} 1 - exp\{-(1-\eta x)^{-1/\eta}\} & \text{if} \quad \eta \neq 0 \\ 1 - exp\{-exp(x)\} & \text{if} \quad \eta = 0. \end{cases}$$

where $1 - \eta x > 0$. So the supports are

$$\begin{cases} \eta^{-1} > x & \text{for} \quad \eta > 0 \\ \eta^{-1} < x & \text{for} \quad \eta < 0 \\ x \in R & \text{for} \quad \eta = 0. \end{cases}$$

One can introduce the related location-scale family H_η by replacing the argument x in Definition 1.2.5 by $\frac{x-\mu}{\sigma}$ for $\mu \in R$ and $\sigma > 0$. The support has to be adjusted accordingly. That is,

$$H_\eta(x) = \begin{cases} 1 - exp\{-(1-\eta[\frac{x+\mu}{\sigma}])^{-1/\eta}\} & \text{if} \quad \eta \neq 0 \\ 1 - exp\{-exp(-[\frac{x+\mu}{\sigma}])\} & \text{if} \quad \eta = 0. \end{cases}$$

where $1 - \eta[\frac{x+\mu}{\sigma}] > 0$. So the supports are

$$\begin{cases} \frac{x-\mu}{\sigma} < \eta^{-1} & \text{for} \quad \eta > 0 \\ \frac{x-\mu}{\sigma} > \eta^{-1} & \text{for} \quad \eta < 0 \\ \frac{x-\mu}{\sigma} \in R & \text{for} \quad \eta = 0. \end{cases}$$

The largest order statistics model and the peak over threshold model can also be extended to minima using similar argument. For details see Leadbetter et al. (1983).

1.3 Objective of this Study

The three models of extremes introduced in Section 1.2 are based on the assumption that the limit distribution of maxima exists. These three models have extensively been used in many contexts, for details see Smith (2001) and de Haan and Ferreira (2006). The extreme moment of share market price is another field of application of the three models of extremes described in this chapter. One should expect the generalized extreme value model or other equivalent models described above for the extremal behavior of prices of shares. However, Gettinby et al. (2004) shows that generalized logistic (GL) distribution (see Definition 3.2.7), introduced by Hosking, gives a better fit for extreme analysis of daily returns of the FT All share index (London Stock Exchange) than the generalized extreme value distribution. Tolikas and Brown (2006) and Tolikas et al. (2007) again propound that in majority of cases (especially minima) an adequate fit for extreme daily returns is generalized logistic distribution for Athens Stock Exchange and German Stock Exchange data respectively compared to generalized extreme value distribution. But no theoretical basis for this claim is established yet in the statistics literature. The main objective of this research work is to investigate whether there is any theoretical basis for the share market data evidences pointed out above. From the three studies of share market data analysis reported above, we point out some important conclusions; they all analyzed share market data up to moderately high time intervals (or blocks), such as two years. Hence, the rate of convergence in extremal types theorem is very poor (or unapplicable in modeling), compared to some other models, in financial sector.

As described above, the main objective of this study is to investigate whether there is any theoretical basis for the data evidences pointed out in Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al. (2007). We investigate the theoretical importance of the generalized logistic distribution in extreme value modeling. In this investigation we first identify the general class of functions where some sort of stability property holds. If so what is its asymptotic distribution? This is the question we answer in Chapter 2. Chapter 2 is completely based on Nidhin and Chandran (2009a). In Chapter 3, the main theorem of Chapter 2 is used to identify the statistic for which the asymptotic distribution is the generalized logistic distribution. The form of stability of maximum is also identified in Chapter 3. In Chapter 4, we study the extremal behavior of BSE sensex data. Among the various models available for extremes we found that the generalized logistic distribution is a good model for both the maximum and the minimum. The main tool used in Chapter 4 is the Anderson-Darling statistic. In Chapter 5, we introduce a goodness fit measure using L-moment ratio's and compare it with the existing goodness of fit statistics using simulation. We use this measure to identify that the generalized logistic distribution is the best fitted distribution to the BSE sensex data. Chapter 6 gives an over all summary of the thesis and some important problems for future work.

Chapter 2

Limit Theory of General Functions of i.i.d.r.v.'s

2.1 Introduction

In Chapter 1, we described stability property of maximum, minimum and sum with non-random sample sizes. Moreover, we saw that, under proper normalization, the statistic $M_n \xrightarrow{w} Z$, where Z is max-stable and $S_n \xrightarrow{w} Z^*$, where Z^* is sum-stable. In this Chapter, we investigate the convergence results for more general functions which includes M_n, m_n and S_n as special cases. Asymptotic distribution of such general functions under proper normalization are derived. The central limit theorem and extremal types theorem are derived as special cases. As a consequence of the main result of the chapter, the limit distribution of random sums, random maximum, and random minimum are also derived. The contents of this Chapter is based on Nidhin and Chandran (2009a).

Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed random variables defined on some probability space (Ω, Ψ, P). Let $g(X_1, X_2\ , \ldots, X_{n(p)})$ be a Borel-measurable function of $X_1, X_2\ , \ldots, X_{n(p)}$, where $n(p)$ is a positive integer valued random variable, independent of $\{X_i\}$, which depends on the parameter p or $n(p)$ is a fixed integer which belongs to the set of natural numbers. In particular, we denote n when we consider fixed sample size, N for random sample size and for general case we use $n(p)$, throughout this chapter.

Definition 2.1.1 *Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d.r.v.'s with $X_1 \sim F$ and g be a Borel function of $X_1, X_2\ , \ldots, X_{n(p)}$. For every $n(p) \geq 1$, define $d_{n(p)} : (-\infty, \infty) \times [0, 1] \to (-\infty, \infty) \times [0, 1]$ such that*

$$(x, F_{g(X_1,X_2,\cdots,X_{n(p)})}(x)) = d_{n(p)}[(x, F(x))]\ ,\ \forall\ x \in \Re. \tag{2.1.1}$$

The existence of such $d_{n(p)}$ is guaranteed because if A and B are two sets with same cardinality then there exist a bijective function from A on to B. Below we give an example to such a function $d_{n(p)}$ for fixed $n(p) = n$.

Example 2.1.1 *Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d.r.v.'s. Let $g(X_1, X_2, \cdots, X_{n(p)}) = \max(X_1, X_2, \cdots, X_n)$ so that*

$$(x, F_{\max(X_1,X_2,\cdots,X_n)}(x)) = d_n(x, F(x)) = (x, (F(x))^n).$$

Then $d_n(x, y) = (x, y^n)$.

Now, let $\mathbf{F}$ be the class of all graphs of distribution functions, G be a non-degenerate distribution function which belongs to $\mathbf{F}$ and define the class $\mathbf{F}^*$, w.r.t g, such that,

$$\mathbf{F}^* = \{(F, \{a_{n(p)} > 0\}, \{b_{n(p)}\})\ |\ \ d_{n(p)}(x, F(a_{n(p)}x + b_{n(p)})) \to (x, G(x)),\ \ F \in \mathbf{F}\}.$$

Instead of taking class of distribution functions we take the class of graph of distribution functions since there is a one to one relation between the class of d.f.s and the class of graph of d.f.s. We use the notation $F_{t,a,b}$ for a member $(F_t, \{a_{n(p)} > 0\}, \{b_{n(p)}\})$ in $\mathbf{F}^*$.

Example 2.1.2 *Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d.r.v.'s with d.f. F such that,*

$$F(x) = \begin{cases} 1 - \exp(-x), & \text{if} \quad x > 0, \\ 0, & \text{if} \quad x < 0. \end{cases}$$

Let $g(X_1, X_2, \ldots, X_{n(p)}) = \max(X_1, X_2, \cdots, X_n)$. Then for $a_n = 1$ and $b_n = \log n$, Leadbetter et al. (1983, page 20) showed that

$$d_{n(p)}[x, F(x + \log n)] = (x, F^n(x + \log n)) = (x, [1 - e^{-x+\log n}]^n)$$

converges to G, where G is the Type I extreme value distribution given by Gumbel with d.f.,

$$G(x) = \exp\{-\exp(-x)\}, -\infty < x < \infty.$$

Hence, $((x, 1 - e^{-x}), \{1\}, \{\log n\})$ is a member in $\boldsymbol{F}^*$.

Example 2.1.3 *Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d.r.v.'s which follows standard Bernoulli distribution with parameter p. Let $g(X_1, X_2, \ldots, X_{n(p)}) = \sum_{i=1}^n X_i$, that is, $g(X_1, X_2, \ldots, X_{n(p)})$ is the number of successes in n trials. Then for $a_n = \sqrt{npq}$ and $b_n = np$, by De Moivre-Laplace central limit theorem (see Billingsly, 1968, p. 1), $F_{\sum_{i=1}^n X_i}(\sqrt{npq}\ x + np)$ converges to standard normal distribution, which implies $((x, F_{\sum_{i=1}^n X_i}(x)), \{\sqrt{npq}\}, \{np\})$ is a member in $\boldsymbol{F}^*$.*

For $g(X_1, X_2, \ldots, X_{n(p)}) = S_n$ and $g(X_1, X_2, \ldots, X_{n(p)}) = M_n$, the limit d.f. of g is stable in some sense as described in Chapter 1. For general function g, we define a stability property in the next section.

2.2 Stability Property of g

In this section we introduce a stability property for the general function g. We also give some examples to illustrate the stability property when g takes special forms.

Definition 2.2.1 *Let $\{d_{n(p)}\}$ be a given one to one functions as described in Definition* 2.1.1. *We say that $F_{1,a,b} \in \boldsymbol{F}^*$ satisfies stability property for the given functions $\{d_{n(p)}\}$ if, for each $n(p)$,*

$$d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) = (x, F_1(x)) \ , \ \forall \ x \ \in \ \Re.$$

A member $F_{1,a,b} \in \mathbf{F}^*$ is a stable member in $\mathbf{F}^*$ if it satisfies the stability property as described in Definition 2.2.1. The following remark describes the stability property in Definition 2.2.1 in terms of the general function g. We also give some examples.

Remark 2.2.1 *Let $\{X_n\}$ be a sequence of i.i.d.r.v.'s and g be a Borel function of $X_1, X_2, \ldots, X_{n(p)}$. Then we say that $\{X_n, n \geq 1\}$ satisfies stability property for the given function g if, for every $n(p)$,*

$$F_{g(X_1,X_2,\cdots,X_{n(p)})}(a_{n(p)}x + b_{n(p)}) = F_{X_1}(x) \ , \ \forall \ x \ \in \ \Re.$$

Example 2.2.1 *(Feller, (2000)) Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d.r.v.'s with X_1 follows normal distribution with mean zero. Take $g(X_1, X_2, \ldots, X_{n(p)}) = \sum_{i=1}^{n} X_i$, $\forall n \geq 1$. The sequence $\{X_n, n \geq 1\}$ satisfies stability property as described in Remark* 2.2.1 *w.r.t the function g, since $\exists$ sequences $\{\sqrt{n}\}$ and $\{0\}$ such that, for each n,*

$$F_{\sum_{i=1}^{n} X_i}(\sqrt{n}x) = F_{X_1}(x) \ , \ \forall \ x \ \in \ \Re.$$

Example 2.2.2 *(Leadbetter et al., (1983)) Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d random variables with X_1 follows standard Gumbel distribution given by,*

$$G(x) = \exp\{-\exp(-x)\}, -\infty < x < \infty.$$

Let $g(X_1, X_2, \ldots, X_{n(p)}) = \max(X_1, X_2, \ldots, X_n)$*. Then* $\{X_n, n \geq 1\}$ *satisfies stability property as described in Remark* 2.2.1 *w.r.t the function* g*, since* $\exists$ *sequences* $\{1\}$ *and* $\{\log n\}$ *such that, for every* n*,*

$$F_{\max(X_1,X_2,\ldots,X_n)}(x + \log n) = F_{X_1}(x) \ , \ \forall \ x \ \in \ \Re.$$

Example 2.2.3 *(Voorn, (1987)) Let* $\{X_n, n \geq 1\}$ *be a sequence of i.i.d random variables and* X_1 *follows logistic distribution. Let* N *be a random variable, independent of* $\{X_i\}$*, and* N *follows geometric distribution. Take* $g(X_1, X_2, \ldots, X_{n(p)}) = \max(X_1, X_2, \ldots, X_N)$*. Then* $\{X_n, n \geq 1\}$ *satisfies stability property as described in Remark* 2.2.1 *w.r.t the function* g*, since* $\exists$ *sequences* $\{1\}$ *and* $\{b > 0\}$*, depends on* N*, such that*

$$F_{\max(X_1,X_2,\ldots,X_N)}(x + b) = F_{X_1}(x) \ , \ \forall \ x \ \in \ \Re.$$

2.3 Classes of Domain of Attraction

In this section, we identify the class of members converging to stable members discussed in Definition 2.2.1. This is identified by defining a relation on the class $\mathbf{F}^*$ and proving that this relation is an equivalence relation.

Consider the class $\mathbf{F}^*$ defined in Section 2.1, we define a relation $\mathbf{M}$ on $\mathbf{F}^*$ and show that $\mathbf{M}$ is an equivalence relation.

Definition 2.3.1 *Let* $F_{1,a,b}$*, and* $F_{2,l,m} \in \boldsymbol{F}^*$*. We say that* $F_{1,a,b}$ *and* $F_{2,l,m}$ *are* ***M*** *related (denoted by* $F_{1,a,b} \overset{M}{\sim} F_{2,l,m}$*) if for all* $\epsilon > 0$*, there exists a* $N_0 \in \boldsymbol{N}$ *(the set of natural numbers), independent of* x*, such that for every* $n(p) \geq N_0$*,*

$$\mid \ d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) - d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)})) \ \mid < \epsilon. \qquad (2.3.2)$$

The following lemma shows that the relation $\mathbf{M}$ is an equivalence relation on $\mathbf{F}^*$.

Lemma 2.3.1 *The relation* $\boldsymbol{M}$ *is an equivalence relation on* $\boldsymbol{F}^*$*.*

Proof: Let $F_{1,a,b} \in \mathbf{F}^*$. Since

$$d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) = d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})), \forall x \text{ and } n(p)$$

the relation is reflexive. Also let $F_{1,a,b}$, and $F_{2,l,m} \in \mathbf{F}^*$ and $F_{1,a,b} \overset{M}{\sim} F_{2,l,m}$. That is for $n(p) \geq N_0$,

$$\mid \ d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) - d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)})) \ \mid < \epsilon$$

implies

$$\mid \ d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)})) - d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) \ \mid < \epsilon,$$

for $n(p) \geq N_0$, or $F_{2,l,m} \overset{M}{\sim} F_{1,a,b}$ so that the relation is symmetric. Now let $F_{1,a,b}$, $F_{2,l,m}$ and $F_{3,e,f} \in \mathbf{F}^*$, such that $F_{1,a,b} \overset{M}{\sim} F_{2,l,m}$ and $F_{2,l,m} \overset{M}{\sim} F_{3,e,f}$. That is, for every $\epsilon > 0$ there exist $N_1, N_2 \in \mathbf{N}$ such that for every $n(p) \geq N_1$,

$$\mid \ d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) - d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)})) \ \mid < \frac{\epsilon}{2}$$

and for $n(p) \geq N_2$,

$$\mid \ d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)})) - d_{n(p)}(x, F_3(e_{n(p)}x + f_{n(p)})) \ \mid < \frac{\epsilon}{2}.$$

Then for $R = \max\{N_1, N_2\}$ and for every $n(p) \geq R$,

$$\mid \ d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) - d_{n(p)}(x, F_3(e_{n(p)}x + f_{n(p)})) \ \mid < \epsilon.$$

that is $\mathbf{M}$ is transitive. Hence $\mathbf{M}$ is an equivalence relation on $\mathbf{F}^*$. $\square$

Hence, relation $\mathbf{M}$ on $\mathbf{F}^*$ partitions $\mathbf{F}^*$ into equivalent classes in such a way that if $F_{1,a,b}$, $F_{2,l,m}$ belongs to same class then $F_{1,a,b}$, $F_{2,l,m}$ are $\mathbf{M}$ related and $F_{1,a,b}$, $F_{2,l,m}$

belongs to different classes are not **M** related. Next we prove that within each equivalent class all members of the class converge to a unique stable member, if a stable member exists within that class. We prove this in the following two lemmas.

Lemma 2.3.2 *Let $\boldsymbol{F}_0$ be one of the equivalent class introduced by the relation $\boldsymbol{M}$. Let $F_{1,a,b} \in \boldsymbol{F}_0$ satisfies the stability property for a given sequence of function $\{d_{n(p)}\}$. That is,*

$$d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) = (x, F_1(x)),\ \forall\ x\ \in\ \Re \text{ and } \forall\ n(p).$$

Then for all $F_{2,l,m} \in \boldsymbol{F}_0$,

$$d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)})) \to (x, F_1(x)),\ \text{ as } n(p) \xrightarrow{p} \infty.$$

Proof: Given that $F_{1,a,b},\ F_{2,l,m} \in \mathbf{F}_0$ where $\mathbf{F}_0$ is one of the class introduced by the relation **M**. Since $F_{2,l,m}\ \in\ \mathbf{F}_0$ then for all $\epsilon > 0,\ \exists\ N_3 \in \mathbf{N}$ such that for $n(p) \geq N_3$,

$$|d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) - d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)}))| < \epsilon.$$

But

$$d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) = (x, F_1(x)),\ \forall\ x\ \in\ \Re \text{ and } \forall\ n(p).$$

The above two equation implies $\forall\ \epsilon > 0\ \exists\ N_3$ such that for $n(p) \geq N_3$,

$$|d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)})) - (x, F_1(x))| < \epsilon.$$

Therefore,

$$d_{n(p)}(x, F_2(l_{n(p)}x + m_{n(p)})) \to (x, F_1(x)),\ \text{as } n(p) \xrightarrow{p} \infty \text{ and } \forall\ x.\ \square$$

Hence, if there exists a stable member in an equivalent class then all members of that class converges to the same stable member. The following lemma shows that the convergent member is unique in every class whenever there exists one such member.

Lemma 2.3.3 *Let $\boldsymbol{F}_0$ be a nonempty equivalent class introduced by relation $\boldsymbol{M}$ which contains a stable member. Then $\boldsymbol{F}_0$ has one and only one stable member.*

Proof: Given $\mathbf{F}_0$ is one of the class introduced by the relation $\mathbf{M}$. If possible let $F_{1,a,b} \in \mathbf{F}_0$ and $F_{2,e,f} \in \mathbf{F}_0$ and satisfy the stability property. Then

$$d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) = (x, F_1(x)), \ \forall \ x \ \in \ \Re \text{ and } \forall \ n(p).$$

Now assume $F_{2,e,f} \in \mathbf{F}_0$ satisfies stability property then

$$d_{n(p)}(x, F_2(e_{n(p)}x + f_{n(p)})) = (x, F_2(x)), \ \forall \ x \ \in \ \Re \text{ and } \forall \ n(p).$$

Since $F_{1,a,b}, F_{2,e,f} \in \mathbf{F}_0$, for all $\epsilon > 0, \ \exists \ N_0 \ \in \ \mathbf{N}$ such that for every $n(p) \geq N_0$,

$$|d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) - d_{n(p)}(x, F_2(e_{n(p)}x + f_{n(p)}))| < \epsilon,$$

which gives $\forall \ \epsilon > 0, \ \exists \ N_0 \ \in \ \mathbf{N}$ such that,

$$|F_1(x) - F_2(x)| < \epsilon \text{ when } n(p) > N_0.$$

That is,

$$|F_1(x) - F_2(x)| < \epsilon, \ \forall \epsilon > 0$$

which gives $F_1(x) = F_2(x)$ for all x. $\square$

Therefore for each equivalent class there exists one and only one stable member if it exists, and all other members converges to the stable member. This class we call as the domain of attraction of that stable member, if it exist. Motivated by this below we define domain of attraction of a stable member G.

Definition 2.3.2 *Let $d_{n(p)}$ be given one to one functions. For $G \in \boldsymbol{F}$ and $\boldsymbol{F}_0 \subseteq \boldsymbol{F}^*$, $\boldsymbol{F}_0$ is the domain of attraction of G if for every $F_{1,a,b} \in \boldsymbol{F}_0$,*

$$d_{n(p)}(x, F_1(a_{n(p)}x + b_{n(p)})) \to (x, G(x)) \ \text{as } n(p) \xrightarrow{p} \infty.$$

When $n(p) = n$, non-random, then $n(p) \xrightarrow{p} \infty$ may be interpreted as $n \to \infty$.

All the convergence results described in this section are under the assumption that a stable member exist in each class. In the next section we identify the condition on g such that each class has a stable member.

2.4 Existence of Stable Member

In this section we define a property **Q** of g. Then show that every equivalent class, introduced by relation **M**, contains a stable member, if g satisfies the **Q** property. A procedure followed in de Haan (1976) and Leadbetter et al. (1983) is used to prove results in this section. The following lemma shows that $d_{n(p)}^{-1}(x, F_1(x))$ always exists and is a graph of a distribution function.

Lemma 2.4.1 *Let* $\{d_{n(p)}\}$ *be given functions defined in equation* (2.1.1)*, and* $\{d_{n(p)}\}$ *has the property that* $d_{n(p)}$*'s are one to one, on to and monotone. Let* $F(x)$ *be a distribution function. Then* $d_{n(p)}^{-1}(x, F(x))$ *always exists and is a graph of a distribution function.*

Proof: Clearly, $d_{n(p)}^{-1}(-\infty, 0) = (-\infty, 0)$ and $d_{n(p)}^{-1}(\infty, 1] = (\infty, 1)$. Now $d_{n(p)}$ is monotone gives $d_{n(p)}^{-1}(x, F_1(x))$ is non-decreasing and right continuous. That is, $d_{n(p)}^{-1}(x, F_1(x))$ is a distribution function. $\square$

Next we define a property **Q** of g.

Definition 2.4.1 *Let $g(x_1, x_2 \ , \ldots, x_{n(p)})$ be a Borel-measurable function of $x_1, x_2 \ , \ldots, x_{n(p)}$, then g satisfies property* ***Q*** *if,*

$$\begin{aligned}g(x_1, x_2, \cdots, x_{n(p)}) = g\{&g(x_1, x_2, \cdots, x_{n_1(p)}),\\ &g(x_{n_1(p)+1}, x_{n_1(p)+2}, \cdots, x_{2n_1(p)}),\\ &\cdots, g(x_{(n_2(p)-1)n_1(p)+1}, x_{(n_2(p)-1)n_1(p)+2}, \cdots, x_{n_2(p)n_1(p)})\}\end{aligned}$$

where $n(p)$, $n_1(p)$ and $n_2(p)$ are integer valued random variables or integers such that $P\{n(p) = n_1(p)n_2(p)\} = 1$.

Below we give functions that are commonly used in statistics which satisfy the property **Q** defined above.

Example 2.4.1 *Let $\{X_i\}$ be a sequence of i.i.d.r.v. and*

$$g(X_1, X_2, ..., X_{n(p)}) = \max(X_1, X_2, \cdots, X_n),$$

where n is a natural number. For a special case take $n_1(p) = 2, n_2(p) = 2$ and for X_1, X_2, X_3, X_4, we get,

$$\max(X_1, X_2, X_3, X_4) = \max\{\max(X_1, X_2), \max(X_3, X_4)\}.$$

Hence, $g(X_1, X_2, ..., X_{n(p)}) = \max(X_1, X_2, \cdots, X_n)$ satisfies property ***Q****.*

Remark 2.4.1 *Let $\{X_i\}$ be a sequence of i.i.d.r.v.'s then $\sum_{i=1}^{n} X_i$, $\prod_{i=1}^{n} X_i$, are other examples of g satisfying property* ***Q****. For N integer valued r.v.'s, $\max(X_1, X_2, \cdots, X_N)$, and $\sum_{i=1}^{N} X_i$, also satisfy* ***Q****. Here $\max(X_1, X_2, \ldots, \ X_N)$ satisfies the property* ***Q****, since given that $P\{n(p) = n_1(p)n_2(p)\} = 1$ and w.l.o.g. if $n(p)$ take the value 10 then $n_1(p)$ and $n_2(p)$ should be 5 and 2 or 2 and 5 or 10 and 1 or 1 and 10 respectively and hence the property follows.*

Now the following result gives a property of $d_{n(p)}$ from property $\mathbf{Q}$ of g.

Lemma 2.4.2 *Let g satisfies property $\boldsymbol{Q}$ and $\{d_{n(p)}\}$ be defined by equation (2.1.1). Then*

$$d_{n(p)m(p)}(x, F(x)) = d_{n(p)}d_{m(p)}(x, F(x)) \ \ \forall x \text{ and } \forall\ m(p),\ n(p).$$

Proof: Let $\{X_n,\ n \geq 1\}$ be i.i.d r.v's and $g(X_1, X_2, \ldots, X_{n(p)})$ be a Borel measurable function. Now for $n \geq 1$, define

$$Y_n = g(X_{(n-1)m(p)+1}, X_{(n-1)m(p)+2}, \cdots, X_{nm(p)})$$

then $\{Y_n,\ n \geq 1\}$ is also a sequence of i.i.d r.v's. Let $P\{X_1 \leq x\} = F(x)$ and $P\{Y_1 \leq x\} = G(x)$. Then

$$(x, P\{g(Y_1, Y_2, \cdots, Y_{n(p)}) \leq x\}) = d_{n(p)}(x, G(x)).$$

Which is same as,

$$(x, P\{g[g(X_1, X_2, \cdots, X_{m(p)}), g(X_{m(p)+1}, X_{m(p)+2}, \cdots, X_{2m(p)}), \cdots,$$

$$g(X_{(n(p)-1)m(p)+1}, X_{(n(p)-1)m(p)+2}, \cdots, X_{n(p)m(p)})] \leq x\}) = d_{n(p)}(x, G(x)).$$

But we know that

$$\begin{aligned}(x, G(x)) = (x, P\{Y_1 \leq x\}) &= (x, P\{g(X_1, X_2, \cdots, X_{m(p)}) \leq x\}) \\ &= d_{m(p)}(x, F(x))\end{aligned}$$

Hence,

$$\begin{aligned}&(x, P\{g[g(X_1, X_2, \cdots, X_{m(p)}), g(X_{m(p)+1}, X_{m(p)+2}, \cdots, X_{2m(p)}), \\ &\qquad \cdots, g(X_{(n(p)-1)m(p)+1}, X_{(n(p)-1)m(p)+2}, \cdots, X_{n(p)m(p)})] \leq x\})\end{aligned}$$

$$= d_{n(p)}d_{m(p)}(x, F(x)).$$

But $(x, P\{g(X_1, X_2, \cdots, X_{n(p)m(p)}) \le x\}) = d_{n(p)m(p)}(x, F(x))$. So by property **Q**

$$d_{n(p)m(p)}(x, F(x)) = d_{n(p)} d_{m(p)}(x, F(x)) \;\; \forall x \text{ and } \forall\ m(p),\ n(p). \;\; \square$$

The next two lemmas establish the existence of stable member in each equivalent class when g satisfies property **Q**.

Lemma 2.4.3 *If there is a sequence $\{F_n\}$ of d.f's and constants $\{a_{n(p)} > 0\}$ and $\{b_{n(p)}\}$ such that*

$$(x, F_{n(p)}(a_{n(p)k(p)}x + b_{n(p)k(p)})) \xrightarrow{w} d^{-1}_{k(p)}(x, F_{t_1}(x)). \tag{2.4.3}$$

for any $k(p)$. Then there exist some sequence $\{e_{n(p)} > 0\}$ and $\{f_{n(p)}\}$ such that $F_{t_1,e,f}$ satisfies stability property.

Proof: If equation (2.4.3) holds for each $k(p)$ then by Khintchine's theorem,

$$d^{-1}_{k(p)}(x, F_{t_1}(x)) = (x, F_{t_1}(e_{n(p)}x + f_{n(p)})) \;\; \forall \;\; x$$

for some values $\{e_{n(p)} > 0\}$ and $\{f_{n(p)}\}$. Which implies

$$(x, F_{t_1}(x)) = d_k(x, F_{t_1}(e_{n(p)}x + f_{n(p)})).$$

So that $F_{t_1,e,f}$ satisfies stability property. $\square$

From this point onwards we consider weak convergence since we have used Khintchine's theorem.

Lemma 2.4.4 *Let $\boldsymbol{F}_0$ be one of the equivalent class introduced by the relation $\boldsymbol{M}$. If for $F_{t_2,a,b} \in \boldsymbol{F}_0$ has the following convergence ,*

$$d_{n(p)}(x, F_{t_2}(a_{n(p)}x + b_{n(p)})) \xrightarrow{w} (x, F_{t_1}(x)).$$

Then there exist sequences $\{e_{n(p)} > 0\}$ and $\{f_{n(p)}\}$ such that $F_{t_1,e,f}$ satisfies stability property.

Proof: Given that there exist atleast one $F_{t_2,a,b}$ such that

$$d_{n(p)}(x, F_{t_2}(a_{n(p)}x + b_{n(p)})) \xrightarrow{w} (x, F_{t_1}(x)).$$

Which implies

$$d_{n(p)k(p)}(x, F_{t_2}(a_{n(p)k(p)}x + b_{n(p)k(p)})) \xrightarrow{w} (x, F_{t_1}(x)).$$

for every $k = 1, 2, \cdots$. By using Lemma 2.4.2 we get

$$d_{k(p)}d_{n(p)}(x, F_{t_2}(a_{n(p)k(p)}x + b_{n(p)k(p)})) \xrightarrow{w} (x, F_{t_1}(x)).$$

That is,

$$d_{n(p)}(x, F_{t_2}(a_{n(p)k(p)}x + b_{n(p)k(p)})) \xrightarrow{w} d^{-1}_{k(p)}(x, F_{t_1}(x)).$$

Putting $d_{n(p)}(x, F_{t_2}) = (x, F_{n(p)})$ we get

$$(x, F_{n(p)}(a_{n(p)k(p)}x + b_{n(p)k(p)})) \xrightarrow{w} d_k^{-1}(x, F_{t_1}(x)).$$

and by Lemma 2.4.3 there exist sequence $\{e_{n(p)} > 0\}$ and $\{f_{n(p)}\}$ such that $F_{t_1,e,f}$ satisfies stability property. $\square$

Hence for every equivalent class of $\mathbf{F}^*$ there exists a unique stable member, if g satisfies property **Q**.

2.5 Main Theorem

In this section we prove the limit theorem of general functions of i.i.d.r.v.'s which posses property **Q**.

Theorem 2.5.1 *Let $g(x_1, x_2, \ldots, x_{n(p)})$ be a Boral-measurable function of $x_1, x_2, \ldots,$ $x_{n(p)}$ such that g satisfies property* $\boldsymbol{Q}$*, and $n(p)$ is independent of $\{X_i\}$. Let* $\boldsymbol{X}$ *be a* *class of random variables such that for all $X \in \boldsymbol{X}$ there exist a sequences $\{a_{n(p)} > 0\}$ and $\{b_{n(p)}\}$ such that*

$$F_{g(X_1, X_2, \cdots, X_{n(p)})}(a_{n(p)}x + b_{n(p)}) = F_{X_1}(x), \ \forall\ x \ \in \ \Re \ \textit{and} \ \forall\ n \ \in \ \boldsymbol{N},$$

where X_i are independent copies of X. Let $\{Y_n, n \geq 1\}$ be any i.i.d sequence of r.v's such that

$$F_{g(Y_1, Y_2, \cdots, Y_{n(p)})}(e_{n(p)}x + f_{n(p)}) \xrightarrow{w} F_Z(x)$$

for some $\{e_{n(p)} > 0\}$ and $\{f_{n(p)}\}$. Then Z is either degenerate r.v or $Z \in \boldsymbol{X}$.

Proof: By Lemma 2.4.4 we get if $F_{g(Y_1, Y_2, \cdots, Y_{n(p)})}(e_{n(p)}x + f_{n(p)})$ converges to a non-degenerate function $F_Z(x)$ then there exist a sequence $\{a_{n(p)} > 0\}$ and $\{b_{n(p)}\}$ such that

$$F_{g(Z_1, Z_2, \cdots, Z_n)}(a_{n(p)}x + b_{n(p)}) = F_Z(x), \ \forall\ x \ \in \ \Re \text{ and } \forall\ n(p).$$

where $Z_1, Z_2, \cdots$ are same copies of Z. Then by Lemma 2.3.2 and Lemma 2.3.3 if $F_{g(Y_1, Y_2, \cdots, Y_{n(p)})}(e_{n(p)}x + f_{n(p)})$ converges to a function $F_Z(x)$ then Z is degenerate or $Z \in \mathbf{X}$. $\square$

2.6 Applications of the Main Theorem

In this section we describe the applications of Theorem 2.5.1. The Central limit theorem and extremal types theorem are derived as special cases. We also derive the asymptotic distribiton of sum, maximum and minimum of a random number of r.v.'s.

2.6.1 Central Limit Theorem as a Special Case

Let $n(p) = n$ and $g(X_1, X_2, \ldots, X_{n(p)}) = \sum_{i=1}^{n} X_i$, then g satisfies property $\mathbf{Q}$ as described in Definition 2.4.1. Let $\mathbf{X}$ be the class of random variables which follows

Levy-skew stable distributions. Then for all $X \in \mathbf{X}$ there exist a sequence $\{a_n > 0\}$ and $\{b_n\}$ such that, for each n,

$$F_{\sum_{i=1}^{n} X_i}(a_n x + b_n) = F_{X_1}(x), \ \forall \ x \ \in \ \Re$$

where X_i are same copies of X. Now $\{Y_n, n \geq 1\}$ be an i.i.d sequence of r.v's such that

$$F_{\sum_{i=1}^{n} Y_i}(c_n x + d_n) \xrightarrow{w} F_Z(x)$$

for some $\{c_n > 0\}$ and $\{d_n\}$. Then Z is either degenerate r.v or $Z \in \mathbf{X}$.

2.6.2 Extremal Types Theorem as a Special Case

Let $n(p) = n$ and $g(X_1, X_2, \ldots, X_{n(p)}) = \max(X_1, X_2, ..., X_n)$, then g satisfies property $\mathbf{Q}$ by Example 2.4.1. Let $\mathbf{X}$ be the class of random variables follows Generalized Extreme Value distributions given by,

$$F_1(x) = \begin{cases} exp\{-(1-kx)^{1/k}\} & \text{if} \quad k \neq 0 \\ exp\{-exp(-x)\} & \text{if} \quad k = 0. \end{cases}$$

and the supports are

$$\begin{cases} x > k^{-1} & \text{for} \quad k < 0 \\ x < k^{-1} & \text{for} \quad k > 0 \\ x \in R & \text{for} \quad k = 0. \end{cases}$$

Then for all $X \in \mathbf{X}$ there exist a sequence $\{a_n > 0\}$ and $\{b_n\}$ such that

$$F_{\max(X_1, X_2, \cdots, X_n)}(a_n x + b_n) = F_{X_1}(x) \ \forall \ x \ \in \ \Re \text{ and } \forall \ n \ \in \ \mathbf{N}$$

where X_i are same copies of X. Now $\{Y_n, n \geq 1\}$ be a sequence of i.i.d.r.v.'s and if

$$F_{\max(Y_1, Y_2, \cdots, Y_n)}(c_n x + d_n) \xrightarrow{w} F_Z(x)$$

for some $\{c_n > 0\}$ and $\{d_n\}$. Then Z is either degenerate r.v or $Z \in \mathbf{X}$.

2.6.3 Random Sum Convergence as a Special Case

Let $n(p) = N$, a r.v follows geometric distribution and $g(X_1, X_2, \ldots, X_{n(p)}) = \sum_{i=1}^{N} X_i$, then g satisfies property $\mathbf{Q}$. Let $\mathbf{X}$ be the class of random variables which follows geometric stable distribution. Then for all $X \in \mathbf{X}$ there exist constants $\{a_{n(p)} > 0\}$ and $\{b_{n(p)}\}$ such that, for each n,

$$F_{\sum_{i=1}^{N} X_i}(a_{n(p)}x + b_{n(p)}) = F_{X_1}(x) \;\forall\; x \;\in\; \Re$$

where X_i are same copies of X. Now $\{Y_n, n \geq 1\}$ be an i.i.d sequence of r.v's such that

$$F_{\sum_{i=1}^{n} Y_i}(l_{n(p)}x + m_{n(p)}) \xrightarrow{w} F_Z(x)$$

for some $\{l_{n(p)} > 0\}$ and $\{m_{n(p)}\}$. Then Z is either degenerate r.v or $Z \in \mathbf{X}$.

The results discussed in subsections 2.6.1, 2.6.2 and 2.6.3 are known results that can be seen in literature (see Feller, (2000), Leadbetter et al., (1983), and Kozubowski and Rachev, (1999)). In the next subsection we use Theorem (2.5.1) to derive the limit theorems of $\max(X_1, X_2, \ldots, X_N)$ and $\min(X_1, X_2, \ldots, X_N)$, where N is an integer valued variable which follows geometric or extended geometric distribution.

2.6.4 Limit Theory of Maxima and Minima when Sample Size is Random

Let $N, X_1, X_2, \ldots$ be independent random variables, $X_1, X_2, \ldots$ with common distribution function F. Voorn (1987) showed that if N follows geometric distribution then X_i is max-stable w.r.t N (see Section 3.4) if and only if $X_i, i = 1, 2, \ldots$ follows either Logistic, loglogistic or backward loglogistic distribution. Using Theorem 2.5.1, one can derive the limit distribution of $\max(Y_1, Y_2, \ldots, Y_N)$, for any sequence $\{Y_i\}$ of i.i.d.r.v.'s when the limit distribution exists and N follows geometric distribution.

Corollary 2.6.1 *Let $Y_1, Y_2, \ldots$ be independent and identically distributed random variables with distribution function F. Let $Z_N = \max(Y_1, Y_2, \ldots, Y_N)$ where N is an integer valued random variable which follows geometric distribution and N is independent of each Y_i. Then the limit distribution of Z_N after proper normalization, if it exists, is one of the following class of distributions:*

$$\textit{Logistic}: \qquad G(x) = \frac{1}{1+e^{-x}}, -\infty < x < \infty$$

$$\textit{Loglogistic}: \qquad G(x) = \begin{cases} 0 & \textit{if} \quad x \leq \alpha \ , \\ \frac{1}{1+(x-\alpha/\beta)^{-\gamma}} & \textit{if} \quad x > \alpha \ . \end{cases}$$

$$\textit{Backward loglogistic}: \qquad G(x) = \begin{cases} \frac{1}{1+(\alpha-x/\beta)^{-\gamma}} & \textit{if} \quad x \leq \alpha \ , \\ 0 & \textit{if} \quad x > \alpha \ . \end{cases}$$

Proof: Voorn (1987) has shown that the class of max-stable distribution is either Logistic, Loglogistic and Backward loglogistic when N is geometric random variable. Using Theorem 2.5.1, the limit distribution of Z_N under proper normalization is one of the three distributions mentioned above. □

Similar results hold when N is an extended geometric random variable where the limit distribution will be extended logistic, extended loglogistic and extended backward loglogistic distribution respectively, provided N follows extended geometric distribution.

Similarly, from Theorem 2.5.1, the limit theorem of the random minima can easily be derived for the same N. So using Theorem 2.5.1, to derive the limit distribution of $\max(X_1, X_2, \ldots, X_N)$, when N follows a particular distribution, we need the distributions with characterizing stability property w.r.t. the same N. To find the distributions with stability property is simple see for example (Voorn (1987), Sreehari (1995), Satheesh and Nair (2002), etc.).

2.7 Conclusion

The limit theory of general functions of i.i.d.r.v.'s presented in this chapter is a unified approach of existing important asymptotic theorems in statistics. From the main result, it is verified that Central limit theorem, Extreme value theorem and similar important convergence theorems are special cases of a more general convergence result. Moreover the results implies a comparatively simple way to derive limit distribution of a statistic g. That is, first derive the class of stable distributions for the statistic g, then the limit distribution is identified as the same class. We use this idea to prove a limit theorem of a statistic to the generalized logistic distribution in the next chapter.

Chapter 3

Generalized Logistic Distribution in Extreme Value Modeling

3.1 Introduction

In Chapters 1 and 2, we saw that asymptotic or approximated model for large values is the GEV distribution or GP distribution. However, in a number of recent share market data analysis, various authors have empirically shown that the extreme movement of share market data can adequately be modeled by the generalized logistic distribution than GEV or GP distribution (see for example Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al. (2007) to mention a few). No theoretical basis of this claim has yet been established. In this Chapter, we try to investigate this problem at three different angles. First note that large values always arise from the right tail of distributions. Hence, if GEV, GP and GL distributions are good asymptotic models for extreme data then one should expect tail equivalence of these distributions asymptotically. In Section 3.3, we prove that the tails of GL, GEV and

GP distributions are equivalent asymptotically. In Chapters 1 and 2, we described that both the asymptotic models of extremes, namely, GEV and GP are stable distributions through extreme order statistics in some sense. If GL distribution is also a model for extremes then one should expect some sort of stability property for GL through the extreme order statistics M_n or m_n. This problem is investigated in Section 3.4. In Section 3.5, the asymptotic probability law of the extreme order statistic, to be discussed in Section 3.4, which is stable is identified as the GL distribution. This chapter is based on Nidhin and Chandran (2010a). Before describing the theoretical issues of this chapter, we introduce some basic concepts in Section 3.2, which are used to prove the results in the forthcoming sections of this chapter.

3.2 Basic Concepts

First, we introduce the max-stability w.r.t. a discrete distribution. This is the stability of maximum of a random number of random variables.

Definition 3.2.1 *Let F be a non-degenerate distribution function of a random variable X and N be a discrete random variable defined on set of positive integers with probability mass function $\{p_n\}$. That is,*

$$P\{N = n\} = p_n, n = 1, 2, 3, \ldots.$$

Then F is said to be max-stable w.r.t. $\{p_n\}$ (or r. v. X is said to be maximum stable w.r.t. a r. v. N) if there exist real numbers $a_N > 0$ and b_N such that,

$$\frac{\max(X_1, X_2, \ldots, X_N) - b_N}{a_N} \stackrel{d}{=} X$$

where $X_n,\ n \geq 1$ are same copies of X.

Similarly, we can define min-stability w.r.t. a discrete distribution as follows.

Definition 3.2.2 *Let F be a non-degenerate distribution function of a random variable X and N be a discrete random variable defined on set of positive integers with probability mass function $\{p_n\}$. That is,*

$$P\{N = n\} = p_n, \quad n = 1, 2, 3, \ldots.$$

Then F is said to be min-stable w.r.t. $\{p_n\}$ (or r. v. X is said to be minimum stable w.r.t. a r. v. N) if there exist real numbers $c_N > 0$ and d_N such that,

$$\frac{\min(X_1, X_2, \ldots, X_N) - d_N}{c_N} \stackrel{d}{=} X$$

where $X_n,\ n \geq 1$ are same copies of X.

To prove tail equivalence of two d.f.s we need the concepts of right end point and left end point of distributions, which are introduced below.

Definition 3.2.3 *The right end point and left end point of a d.f. F denoted by x_F and y_F respectively are,*

$$x_F = \sup\{x : F(x) < 1\},$$

$$y_F = \inf\{x : F(x) > 0\}.$$

Another important concept used in this chapter is the tail equivalence of two distributions, which we define below.

Definition 3.2.4 *Two distributions F and G are equivalent in their right tail if they have the same right end point, i.e. if $x_F = x_G$, and*

$$\lim_{x \uparrow x_F} \frac{1 - F(x)}{1 - G(x)} = 1.$$

Definition 3.2.5 *Two distributions F and G are equivalent in their left tail if they have the same left end point, i.e. if $y_F = y_G$, and*

$$\lim_{x \downarrow y_F} \frac{F(x)}{G(x)} = 1.$$

Logistic distribution plays some important role in modeling of extremes of data. Below we define Logistic distribution and then discuss its importance in Extreme value theory.

Definition 3.2.6 *A random variable X is said to follow standard logistic distribution if its d.f is given by*

$$F_X(x) = \frac{1}{1 + e^{(-x)}}, \qquad -\infty < x < \infty. \tag{3.2.1}$$

The following are some important results which connects logistic distribution and extreme value theory.

Result 3.2.1 *(Voorn (1987)): Let $\{X_n, n \geq 1\}$ be a sequence of i.i.d.r.v.s with symmetric distribution function $F(.)$. Let N be an integer valued random variable independent of $\{X_n\}$ and*

$$P(N = k) = p(1-p)^{k-1}, 0 < p < 1, \quad k \geq 1. \tag{3.2.2}$$

Let $M_N = max(X_1, X_2, \ldots, X_N)$ and there exist constants a_N and b_N such that

$$\frac{M_N - b_N}{a_N} \stackrel{d}{=} X_1$$

iff F is the logistic distribution function.

Using Definition 3.2.1, Result 3.2.1 tells us that the geometric random maximum M_n of sequence of i.i.d.r.v.s $\{X_n\}$ is max-stable iff the i.i.d. sequence $\{X_n\}$ is such that

X_1 follows Logistic distribution. For more details about including proof see Voorn (1987). The logistic distribution appears as a limiting distributions of maximum M_n as described in the following result

Result 3.2.2 *(Balakrishnan (1996)): Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed random variables with distribution function $F(.)$. Let N be an integer valued random variable which are not necessarily independent of $\{X_n\}$. Assume F is in the domain of attraction of Type 1 extreme value distribution and*

$$\lim_{N\to\infty} P(N \leq nz) = A(z), \qquad z > 0,$$

where $A(z)$ is a proper distribution function. Then the limit distribution of M_N is logistic iff $A(z) = 1 - \exp\{-az\}, a > 0$.

The next result brings the connection between logistic distribution and the extreme value theory through midrange. For proof see Balakrishnan (1996).

Result 3.2.3 *(Balakrishnan (1996)): Let $\{X_n, n = 1, 2, \ldots\}$ be a sequence of independent and identically distributed random variables with X_1 follows $F(.)$. Let $M_n = \max(X_1, X_2, \ldots, X_n)$ and $m_n = \min(X_1, X_2, \ldots, X_n)$. Assume F is symmetric distribution and belongs to the domain of attraction of Type 1 extreme value distribution then under proper normalization the midrange*

$$\eta_n = \frac{m_n + M_n}{2}$$

converges to a random variable Z which follows standard logistic distribution.

In the statistics literature, there are several ways of generalizing the logistic distribution given in Definition 3.2.6 (for more details see Johnson et. al. (1995)). Among the various generalization of logistic distribution the one given by Hosking (see Hosking and Wallis (1997)) is called the 5th generalized logistic distribution (Johnson et.

al. (1995)). This form of generalized logistic distribution is used for modeling share market data by various authors (see for example Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al. (2007)). Motivated by this we introduce the 5th generalized logistic distribution and call this generalization of logistic distribution as generalized logistic (GL) distribution in this thesis.

Definition 3.2.7 *A random variable is said to follow generalized logistic distribution if its d.f.* $F_\eta(x)$ *is given by*

$$F_\eta(x) = \begin{cases} \frac{1}{1+(1-\eta x)^{1/\eta}} & \text{if} \quad \eta \neq 0 \\ \frac{1}{1+x} & \text{if} \quad \eta = 0. \end{cases}$$

The support is

$$\begin{cases} \eta^{-1} > x & \text{for} \quad \eta > 0 \\ \eta^{-1} < x & \text{for} \quad \eta < 0 \\ x \in R & \text{for} \quad \eta = 0. \end{cases}$$

One can introduce the related location-scale family by replacing the argument x in Definition 3.2.7 by $\frac{x-\mu}{\sigma}$ for $\mu \in R$ and $\sigma > 0$. The support has to be adjusted accordingly. That is,

$$F_\eta(x) = \begin{cases} \frac{1}{1+(1-\eta(\frac{x-\mu}{\sigma}))^{1/\eta}} & \text{if} \quad \eta \neq 0 \\ \frac{1}{1+\frac{x-\mu}{\sigma}} & \text{if} \quad \eta = 0. \end{cases}$$

The support is

$$\begin{cases} \eta^{-1} > \frac{x-\mu}{\sigma} & \text{for} \quad \eta > 0 \\ \eta^{-1} < \frac{x-\mu}{\sigma} & \text{for} \quad \eta < 0 \\ x \in R & \text{for} \quad \eta = 0. \end{cases}$$

The parameter η is known as the shape parameter of the distribution and $\eta \in \Re$. For $\eta = 0$, F_0 can be identified as the Logistic distribution given in Definition 3.2.6. For

$\eta > 0$, and some scale transform in Defintion 3.2.7, we get Loglogistic distribution given in Corollary 2.6.1. Similarly for $\eta < 0$, through some scale transform we get Backward loglogistic distribution given in Corollary 2.6.1. The statistical properties and estimation issues of the GL distribution are discussed in Johnson et. al. (1995).

3.3 Tail Equivalence of GEV, GP and GL Distributions

Two important theoritical asymptotic probability models for maximum are the GEV and GP distributions discussed in Chapter 1. However, in a number of recent papers on share market data analysis (see Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al. (2007))), it is verified through data that GL is also a probable model for the analysis of fluctuation of maximum. Figure 3.1 gives the probability density functions of standard Generalized Extreme Value distribution for maxima, standard generalized Pareto distribution, standard generalized extreme value distribution for minima and standard Generalized Logistic distribution. Figure 3.2 gives the distribution functions of standard generalized extreme value distribution for maxima, standard generalized Pareto distribution, standard generalized extreme value distribution for minima and standard generalized logistic distribution. Figures 3.1 and 3.2 indicate that the right tails of the GEV(max), GP and GL distributions are similar. This clearly indicates that these three distributions are asymptotically equivalent in their tails, which we prove in the following theorems. We use notations G, H, F, and Q for distribution functions of GEV(max), GEV(min), GL, and GP respectively.

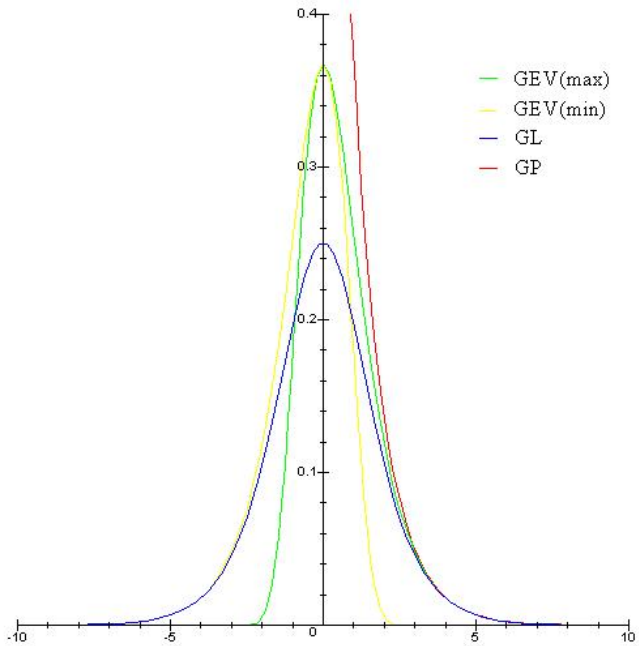


Figure 3.1

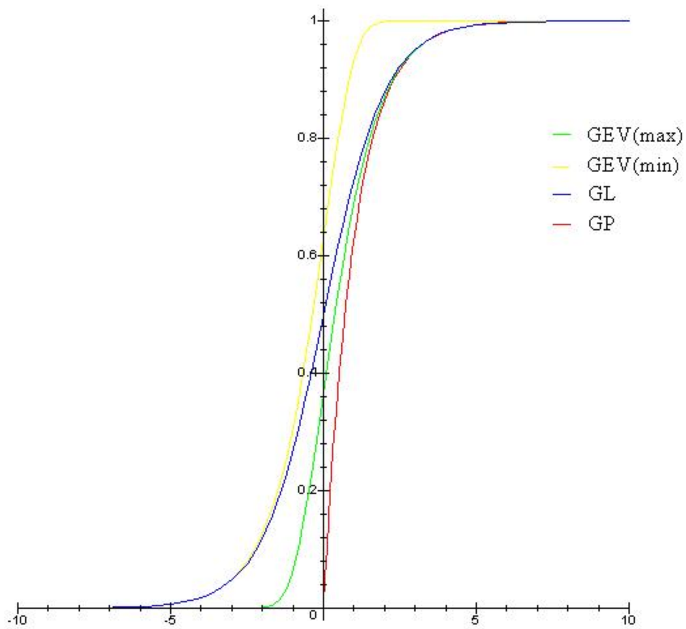


Figure 3.2

Theorem 3.3.1 *The distributions GEV(max) and GL are equivalent at their right tails.*

Proof: As per the assumed notation for the distributions of GEV(max) and GL, to prove the asymptotic equivalence of their right tails, by Definition 3.2.4, it is enough to prove that $x_G = x_F$ and

$$\lim_{x \to x_F} \frac{1 - G(x)}{1 - F(x)} = 1, \quad \forall \eta \in \Re.$$

We prove this in the following three cases.

Case 1 $[\eta = 0]$. Here $x_G = x_F = \infty$,

$$\lim_{x \to \infty} \frac{1 - G(x)}{1 - F(x)} = \lim_{x \to \infty} \frac{1 - \exp\{-\exp(-x)\}}{1 - \frac{1}{1+\exp(-x)}}.$$

This is $\frac{0}{0}$ form. So applying L- Hospital's rule we get,

$$\lim_{x \to \infty} \frac{d/dx[1 - G(x)]}{d/dx[1 - F(x)]} = \lim_{x \to \infty} [\exp(-x). \exp\{-\exp(-x)\}] \frac{(1 + \exp(-x))^2}{\exp(-x)}$$

$$= \lim_{x \to \infty} \exp\{-\exp(-x)\}[1 + \exp(-x)]^2 = 1.$$

Case 2. $[\eta < 0]$. Here also $x_G = x_F = \infty$,

$$\lim_{x \to \infty} \frac{1 - F_\eta(x)}{1 - G_\eta(x)} = \lim_{x \to \infty} \frac{1 - \frac{1}{1+(1-\eta x)^{1/k}}}{1 - \exp\{-[1 - \eta x]^{1/\eta}\}}$$

which is again $\frac{0}{0}$ form. So applying L-Hospital's rule,

$$\lim_{x \to \infty} \frac{d/dx[1 - F_\eta(x)]}{d/dx[1 - G_\eta(x)]}$$

$$= \lim_{x \to \infty} \Big\{ \frac{(1 - \eta x)^{1/\eta}}{(1 - \eta x)[1 + (1 - \eta x)^{1/\eta}]} \frac{1 - \eta x}{(1 - \eta x)^{1/\eta} \exp\{-(1 - \eta x)^{1/\eta}\}}$$

$$- \frac{[(1 - \eta x)^{1/\eta}]^2}{[1 + (1 - \eta x)^{1/\eta}]^2.(1 - \eta x)} \frac{1 - \eta x}{(1 - \eta x)^{1/\eta} \exp\{-(1 - \eta x)^{1/\eta}\}} \Big\}$$

$$= 1 - \lim_{x \to \infty} [\frac{(1 - \eta x)^{1/\eta}}{1 + 2(1 - \eta x)^{1/\eta} + (1 - \eta x)^{2/\eta}}]. \frac{1}{\exp\{-(1 - \eta x)^{1/\eta}\}}$$

$$= 1 - \lim_{x \to \infty} [\frac{1}{\frac{1}{(1-\eta x)^{1/\eta}} + 2 + (1 - \eta x)^{1/\eta}}]. \frac{1}{\exp\{-(1 - \eta x)^{1/\eta}\}}$$

$$= 1 - \frac{1}{\infty + 2 + 0} = 1.$$

Case 3. $[\eta > 0]$. In this case $x_G = x_F = \frac{1}{\eta}$,

$$\lim_{x \to 1/\eta} \frac{1 - F_\eta(x)}{1 - G_\eta(x)} = 1 - \frac{0}{1} = 1. \quad \Box$$

The following theorem proves the right tail equivalence of GP and GL distributions.

Theorem 3.3.2 *The distributions GP and GL are equivalent at their right tails.*

Proof: Using the same notation as described above, it is enough to prove that

$$\lim_{x \to x_F} \frac{1 - Q(x)}{1 - F(x)} = 1, \quad \forall \eta \in \Re.$$

Case 1 $[\eta = 0]$. Here $x_Q = x_F = \infty$,

$$\lim_{x \to \infty} \frac{1 - Q_\eta(x)}{1 - F_\eta(x)} = \lim_{x \to \infty} \frac{\exp(-x)}{1 - \frac{1}{1+\exp(-x)}}$$

This is $\frac{0}{0}$ form. So applying L- Hospital's rule we get,

$$\lim_{x \to \infty} \frac{d/dx[1 - Q_\eta(x)]}{d/dx[1 - F_\eta(x)]} = \lim_{x \to \infty} [\exp(-x)] \frac{(1 + \exp(-x))^2}{\exp(-x)}$$

$$= \lim_{x \to \infty} [1 + \exp(-x)]^2 = 1.$$

Case 2. $[\eta < 0]$. Here also $x_Q = x_F = \infty$,

$$\lim_{x \to \infty} \frac{1 - Q_\eta(x)}{1 - F_\eta(x)}$$

$$= \lim_{x \to \infty} \frac{[1 - \eta x]^{1/\eta}}{1 - \frac{1}{1+(1-\eta x)^{1/\eta}}}$$

$$= \lim_{x \to \infty} 1 + (1 - \eta x)^{1/\eta} = 1.$$

Case 3. $[\eta > 0]$ In this case $x_G = x_F = \frac{1}{\eta}$,

$$\lim_{x \to 1/\eta} \frac{1 - Q_\eta(x)}{1 - F_\eta(x)}$$

$$= \lim_{x \to 1/\eta} \frac{[1 - \eta x]^{1/\eta}}{1 - \frac{1}{1+(1-\eta x)^{1/\eta}}}$$

$$= \lim_{x \to 1/\eta} 1 + (1 - \eta x)^{1/\eta} = 1. \quad \Box$$

Similar to the behavior of the maximum, the left tails of GEV(min) and GL are similar. We prove their asymptotic equivalence in the left tail in the following theorem. We prove this using the same notations described above.

Theorem 3.3.3 *The distributions GEV(min) and GL are equivalent at their left tails.*

Proof: It is enough to prove that

$$\lim_{x \to y_F} \frac{H(x)}{F(x)} = 1, \quad \forall \eta \in \Re.$$

Case 1 $[\eta = 0]$. Here $y_H = y_F = -\infty$,

$$\begin{aligned}\lim_{x \to -\infty} \frac{H(x)}{F(x)} &= \lim_{x \to -\infty} [1 - \exp\{-\exp(-x)\}].[1 + \exp(-x)] \\ &= 0.\infty\end{aligned}$$

Applying L-Hospital's rule,

$$\begin{aligned}\lim_{x \to -\infty} \frac{d/dx[H(x)]}{d/dx[F(x)]} &= \lim_{x \to -\infty} [\exp(x).\exp\{-\exp(x)\}] \frac{(1 + \exp(-x))^2}{\exp(-x)} \\ &= \lim_{x \to -\infty} [\exp(x).\exp\{-\exp(x)\}][1 + 2\exp(-x) + \exp(-2x)] \\ &= \lim_{x \to -\infty} \exp\{-\exp(-x)\}[\exp(2x) + 2\exp(x) + 1] = 1.[0 + 0 + 1] = 1.\end{aligned}$$

Case 2. $[\eta > 0]$ Here also $y_Q = y_F = -\infty$,

$$\begin{aligned}\lim_{x \to -\infty} \frac{H_\eta(x)}{F_\eta(x)} &= \lim_{x \to -\infty} [1 - \exp\{-(1 - \eta x)^{1/\eta}\}].[1 + (1 - \eta x)^{1/\eta}] \\ &= 0.\infty\end{aligned}$$

$$
\begin{aligned}
&\lim_{x\to-\infty}\frac{d/dx[H_\eta(x)]}{d/dx[F_\eta(x)]}\\
&=\lim_{x\to-\infty}\Big[\frac{(1-\eta x)^{1/\eta}\exp\{-(1-\eta x)^{-1/\eta}\}}{1-\eta x}\Big]\Big[\frac{\{1+(1-\eta x)^{1/\eta}\}^2(1-\eta x)}{(1-\eta x)^{1/\eta}}\Big]\\
&=\lim_{x\to-\infty}\Big[\frac{\exp\{-(1-\eta x)^{-1/\eta}\}}{(1-\eta x)^{2/\eta}}\Big][1+2(1-\eta x)^{1/\eta}+(1-\eta x)^{2/\eta}]\\
&=\lim_{x\to-\infty}\exp\{-(1-\eta x)^{-1/\eta}\}\Big[\frac{1}{(1-\eta x)^{2/\eta}}+\frac{2}{(1-\eta x)^{1/\eta}}+1\Big]\\
&=1.[0+0+1]=1.
\end{aligned}
$$

Case 3. $[\eta < 0]$ In this case $y_Q = y_F = \frac{1}{\eta}$,

Again we get the limit is

$$\lim_{x\to 1/\eta}\frac{H_\eta(x)}{F_\eta(x)} = 0.\infty.$$

Applying L-Hospital's rule,

$$
\begin{aligned}
&\lim_{x\to1/\eta}\exp\{-(1-\eta x)^{-1/\eta}\}\Big[\frac{1}{(1-\eta x)^{2/\eta}}+\frac{2}{(1-\eta x)^{1/\eta}}+1\Big]\\
&=1.[0+0+1]=1.\ \square
\end{aligned}
$$

The distributions GEV(max) and GP are the asymptotic models for maximum. The distribution GEV(min) is the model for minima. Since GL distribution has a right tail equivalent to both GEV(max) and GP distributions and GL distribution has a left tail equivalence to GEV(min) distribution, GL provides an alternative model for both maximum and minimum. This is an advantage of the GL model. The three models for maxima namely, the GEV(max), GP and GEV(min) distributions are stable in terms of extreme order statistics as we see in Chapter 1. The asymptotic tail equivalences proved in Theorems 3.3.1 - 3.3.3 also justifies the application of GL distribution in extreme value theory. That is GL can also be used to model extremes of data in some sense. Another consequence of Theorems 3.3.1 - 3.3.3 is that unlike GEV(max), GEV(min), and GP, the two tails of GL may be used to model the extremal behavior of data. Since, the right tail of GL has tail equivalence with GEV(max) and GP, and GEV(max) and GP are used for the analysis of maximum of data, GL can be used to model maximum of data in some sense. Also the left tail of GL has tail equivalence with GEV(min), and GEV(min) is used for the analysis of minimum of data, GL can be used to model minimum of data in some sense. In the next two sections we identify some important property of GL, which answers the sense in which GL can be used as a model for maximum and minimum of data.

3.4 Stability Property of GL Distribution

In this section we identify another notable characteristic of GL distribution in extreme value theory which is in terms of random sample size. We know that the distribution, GEV(max) includes three well known family of distributions namely Gumbel, Frechet, and reverse-Weibull. This is the only family of distributions which possess a characterizing property of max-stability w.r.t fixed sample size (see Definition 1.2.2). Similarly, GEV(min) incorporates three family of distributions namely reverse-Gumbel, reverse-Frechet, and Weibull which is the class of distributions possessing the characterizing property of min-stability w.r.t fixed sample size (see Definition 1.2.5). Same is the situation with GP distribution. GP includes Exponential, Pareto, and Beta distributions and is the class having the characterizing property of POT-stability w.r.t fixed sample size (see Definition 1.2.2). In Section 3.2, we have seen that Logistic, Loglogistic and Backward loglogistic distributions are members of the GL family of distributions. In this section, we prove that GL distribution characterizes the max-stability and min-stability w.r.t geometric distribution. To prove the max-stability of GL, we use the following lemma from Voorn (1987),

Lemma 3.4.1 *Let F be a non-degenerate distribution function which is maximum* *stable w.r.t* $\{p_n,\ n \geq 1\}$*. If* $y_F = -\infty$*,* $x_F = \infty$*, and for any real constant* c*, and d.f's* G *and* H *defined by* $F(x) = G(c + \exp(x))$ *and* $F(x) = H(c - \exp(-x))$ *for all real* x*, then* G *and* H *are non-degenerate and maximum stable w.r.t* $\{p_n,\ n \geq 1\}$ *with* $y_G = c$ *and* $x_H = c$ *respectively.*

For detailed discussion and proof of this result see Voorn (1987). In the next theorem, we prove max-stability and min-stability of GL distribution using Lemma 3.4.1.

Theorem 3.4.1 *The Generalized Logistic distribution given in Definition* 3.2.7 *characterizes max-stability and min-stability w.r.t geometric distribution.*

Proof: Let F be the distribution function of Generalized Logistic distribution with $\eta < 0$, and location and scale parameters γ and ξ respectively. That is,

$$F(x) = \frac{1}{1 + (1 - \eta(\frac{x-\gamma}{\xi}))^{1/\eta}}, \quad \gamma + \frac{\xi}{\eta} < x < \infty.$$

Now take the transformation $G[x_G + \exp(x)]$, we get,

$$F[x_G + \exp(x)] = \frac{1}{1 + (\frac{-\eta}{\xi}\exp(x))^{1/\eta}}, \quad -\infty < x < \infty$$

which is the distribution function of Logistic distribution given in Definition 3.2.6 with $\sigma = -\eta$ and $\mu = \log(-\frac{\xi}{\eta})$. That is, $G[x_G + \exp(x)]$ is logistic distribution function and by Lemma (3.4.1), the distribution function F is max-stable w.r.t geometric distribution. Similarly, for $\eta > 0$, the same proof works. That is, GL distribution satisfies max-stability w.r.t geometric distribution. To prove that GL is the only distribution with property of max-stability, let X be a random variable which follows GL distribution with shape parameter η. Now take the transformation

$$Y = \begin{cases} X - \frac{\xi}{k} & \text{if} \quad k \neq 0 \\ X & \text{if} \quad k = 0 \end{cases} \tag{3.4.3}$$

then for $\eta > 0$, Y follows loglogistic distribution with distribution function

$$F_Y(y) = \begin{cases} 0 & \text{if} \quad y \leq \gamma \\ \{1 + (\frac{x-\gamma}{b})^{-c}\}^{-1} & \text{if} \quad y > \gamma \end{cases}$$

where $b = \frac{\xi}{\eta}$ and $c = \frac{1}{\eta}$. For $\eta < 0$, Y follows backward loglogistic distribution. By Voorn (1987), we know that the class containing logistic, loglogistic, and backward loglogistic posses the characterizing property of max-stability w.r.t to geometric distribution. The relation (3.4.3) is one to one implies the max-stability property w.r.t to geometric distribution characterize Generalized logistic distribution. Similarly we can prove min-stability characterizes GL distribution w.r.t geometric distribution. □

In the next section we use max-stability and min-stability of GL distribution to derive limit distribution of random maximum and random minimum of i.i.d.r.v.'s w.r.t geometric distribution.

3.5 Limit Distribution of Random Maxima and Minima

The limit theory of random maxima and random minima has been studied by many authors (see for example Galambos (1978)). In this section we derive Generalized logistic distribution as the limit of random maxima and random minima when sample size follows geometric distribution using the results in Section 3.4 and the main theorem in Chapter 2.

Theorem 3.5.1 *Let $\{X_n,\ n \geq 1\}$ be a sequence of i.i.d.r.v.'s and $X_1 \sim F$. Let $M_N = \max(X_1,\ X_2, \ldots,\ X_N)$, $m_N = \min(X_1,\ X_2, \ldots,\ X_N)$, and N is an integer* *valued random variable follows geometric distribution, independent of* $\{X_i\}$. *Also let* Y *and* Z *are random variables with non-degenerate distribution functions* G *and* H *respectively,* $a_N > 0$, b_N, $c_N > 0$ *and* d_N *such that,*

$$\frac{M_N - b_N}{a_N} \xrightarrow{d} Y$$

and

$$\frac{m_N - d_N}{c_N} \xrightarrow{d} Z.$$

Then

$$G(x) = \begin{cases} \frac{1}{1+(1-k(\frac{x-\mu}{\sigma}))^{1/k}} & \text{if} \quad k \neq 0 \\ \frac{1}{1+exp(-(\frac{x-\mu}{\sigma}))} & \text{if} \quad k = 0. \end{cases}$$

with supports are

$$\begin{cases} x > \mu + \sigma k^{-1} & \text{for} \quad k < 0 \\ x < \mu + \sigma k^{-1} & \text{for} \quad k > 0 \\ x \in R & \text{for} \quad k = 0. \end{cases}$$

also H *has the same distributional form of* G *with different parameters.*

Proof: The proof consist of two parts. First, from Theorem 3.4.1 , it is verified that GL posses characterizing property of max-stability and min-stability w.r.t geometric distribution. By using Theorem 2.5.1, the random maxima and random minimum, when sample size follows geometric distribution, converge to the GL distribution, if it exist. □

Theorem 3.5.1 identifies the limit distribution of geometric random maxima and geometric random minima, if it exists, as the GL distribution given in 3.2.7. That is, GL distribution can be used as an asymptotic model for maximum and minimum of a random number (N) of random variables when N follows geometric distribution.

3.6 Conclusion

In this chapter we proved the asymptotic equivalence of right tails of GL distribution with the existing models for extremes namely GEV(max) and GP. We also established the asymptotic equivalence of left tails of GL with the existing model for minima namely GEV(min). We characterized the GL distribution with a stability property called max-stability and min-stability w.r.t. geometric distribution. The asymptotic distribution of random maximum and random minimum are identified as the GL distribution with appropriate parameters. We also saw that this class includes three well known distributions namely logistic, loglogistic and backward loglogistic. These findings are parallel to the existing asymptotic models for extremes for non-random sample sizes. These results justified the theoretical basis for the empirical findings of GL as a suitable model for extremes of share market data. For example Gettinby et al. (2004) verified that generalized logistic distribution as a better model for extreme analysis of daily returns of the FT All share index (London Stock Exchange) than the generalized extreme value distribution. Tolikas and Brown (2006) and Tolikas et al. (2007) concluded that in majority of extreme cases an adequate fit for extreme daily returns is GL in Athens Stock Exchange and German Stock Exchange data respectively compared to Generalized Extreme Value distribution. They are also conformed that GL distribution provide an adequate fit for European Stock Exchange data. Also Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al. (2007) show that GEV(max), GEV(min) and GP models are unapplicable in some financial extreme situations like highly volatile share market.

Chapter 4

Analysis on the Extremal Behavior of BSE Sensex Data

4.1 Introduction

In Chapter 3, we established the theoretical significance of GL distribution in the modeling of extremes. This chapter is an empirical study about the performance of GL distribution in Extreme value modeling in the Indian context. Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al. (2007) illustrate some empirical applications of GL distribution in Extreme value modeling in the financial context. Here, we could see the advantage of GL distribution in modeling extreme movements of BSE sensex data. This chapter is based on Nidhin and Chandran (2009b).

In financial literature, the choice of using an appropriate probability model for financial returns are clearly exemplified rather than selecting a conventional model (see for example Kearns and Pagan (1997)). The fitted distribution is crucially important in financial studies. For Value at Risk (VaR) estimation, one requires appropriate probability distribution of extremes as input. A vast number of literature show the importance of best fitted distribution in VaR analysis (see Danielson et al. (1998)). Another important area of application of probability models of extremes is in hedging procedure. Hedging procedure is clearly based on the probability of fitted distribution (see Cotter John (2001)). To measure the risk of a share price which critically depends on the shape of the distribution, especially the shape of the tail of the distribution, the choice of appropriate model for the tails are crucially important for this measurement of risk. Gettinby et al. (2004) largely illustrates the importance of appropriate probability models for extremes of financial returns data.

In history, a majority of literature on share market data analysis are using normal and lognormal distributions. In the last five decades, different authors have been showing empirical evidence that the distribution of extreme daily share returns are far from normal (See for example Mandelbrot (1963), Fama (1965), Grey and French (1990), Piero (1994), Mc Donald and Xu (1995), Theodossian (1998), Harris and Kucukozmen (2001)). They all analyze daily share returns by using non-normal distributions viz. t-distribution, α-stable distribution etc. However, in 1990's interest in modeling the extremal behavior of finance data have been diverted to extreme value theory. Longin (1996) argues that US daily share returns are far from normal. His analysis of US daily share returns shows that the data is negatively skewed with a skewness of -0.506 and a high kurtosis of 22.057 which indicates the adequate distribution is far from normal. He has concluded that the appropriate distribution for extreme daily share returns is Frechet distribution, which is an extreme value distribution. Similarly, a number of literature on the analysis of US daily share returns show that the appropriate distribution for extreme daily share returns is extreme value distribution (see for example Longin, (1996, 2000)).

However, Gettinby et al. (2004) shows generalized logistic distribution gives a better fit for extreme analysis of daily returns of the FT All share index (London Stock Exchange) than the Generalized Extreme Value distribution. Tolikas and Brown (2006) and Tolikas et al. (2007) again propound that in majority of cases (especially weekly minima) an adequate fit for extreme daily returns is GL in Athens Stock Exchange and German Stock Exchange data respectively compared to Generalized Extreme Value distribution . They are also conformed that GL distribution provide an adequate fit for European Stock Exchange data. Gettinby et al. (2004) concluded with a suggestion that academics and practitioners should be careful of the choice of which distribution to employ in the analysis of stock exchange data. Also Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al. (2007) show that extreme value theory is unapplicable in financial extreme data.

Till now, very little work has been done on the analysis of extreme share returns in Indian stock exchange data and some exceptions are Sarma (2001, 2005). Sarma (2001, 2005) analyze NIFTY extreme returns based on extreme value theory. If a distribution fitted is better than extreme value distribution to the share market data, then one can estimate more accurate result of extreme behavior in data using the best fitted distribution in many applications like VaR, hedging, etc. This chapter makes an attempts to fit the GL distribution to the Bombay Stock Exchange data. This verifies the findings of Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al. (2007) in the Indian context. The rest of this chapter is arranged as follows. Section 4.2 introduces the basic concepts required for the data analysis in this chapter. Section 4.3 introduces the data used for analysis in the chapter. A preliminary analysis of the data is also included in Section 4.3. Section 4.4 identifies the appropriate models for testing goodness fit of the data by using L-moment ratio diagram. In Section 4.5 we use Anderson Darling test to find the adequate distribution for extreme share data. The Section 4.6 concludes the finding of the Chapter.

4.2 Basic Concepts

To identify the appropriate model for the data used in the chapter, we use the concepts of L-moments introduced by Hosking (1986). Motivated by this, we describe L-moment and related concepts in this section. As a tool of goodness of fit we use Jerque-Bera test, Shapiro-Wilk test and Anderson-Darling test. They are also included in this section.

L-moments are linear combination of expectation of order statistics. The L-moments have the following advantages over conventional moments. First, a distribution may be specified by its L-moments, when its first order moment alone exists. Second, the L-moment estimators are more robust to outliers of the data and gives more reliable estimates of parameters than the moment estimators. Third, L-moments can characterize a wider class of distributions. Moreover, parameter estimator obtained from L-moments are sometimes more accurate in small samples than the maximum likelihood estimates. Below we introduce L-moments, for a detailed treatment of the theory of L-moments, see Hosking (1986).

Definition 4.2.1 *Let $X_{1:n},\ X_{2:n}, \ldots,\ X_{n:n}$ be ordered samples taken from a population with distribution function $F(.)$. The k^{th} population L-moment of $F(.)$ is defined by:*

$$\lambda_k = k^{-1} \sum_{r=1}^{k-1} (-1)^r \binom{k-1}{r} EX_{k-r:k}, \quad k = 1, 2, \ldots$$

Using the definition of expectation of order statistics and some calculations, the first four L-moments can be obtained as follows:

$$\lambda_1 = \int_0^1 X(F) dF$$

$$\lambda_2 = \int_0^1 X(F)(2F-1) dF$$

$$\lambda_3 = \int_0^1 X(F)(6F^2 - 6F + 1)dF$$

$$\lambda_4 = \int_0^1 X(F)(20F^3 - 30F^2 + 12F - 1)dF$$

where $X(F)$ is the quantile function defined by $X(F) = \inf\{x : F \leq F(x)\},\ 0 < F < 1$. Standardized L-moments of higher order ($k \geq 3$), which are independent of the units of measurement of X, are known as the L-moment ratios, which are sometimes quite useful. We define them below.

Definition 4.2.2 *Let $\lambda_k,\ k \geq 1$ be the L-moments given in Definition* 4.2.1. *Then for $k \geq 3$, the k^{th} L-moment ratio τ_k is,*

$$\tau_k = \frac{\lambda_k}{\lambda_2}, \quad k \geq 3.$$

Remark 4.2.1 *The first two L-moment ratio's of a d.f. F, that is, τ_3 and τ_4 are known as L-skewness and L-kurtosis of F respectively.*

Below we evaluate the L-moments and L-moment ratios for two known distributions.

Example 4.2.1 *(Exponential Distribution): Let X be a random variable with distribution function F given by,*

$$F(x) = 1 - e^{-(x-\mu)/\sigma}, \quad x > 0.$$

Then the first four L-moments of X are,

$$\lambda_1 = \mu + \sigma$$

$$\lambda_2 = \sigma/2$$

$$\lambda_3 = \sigma/6$$

$$\lambda_4 = \sigma/12.$$

Also L-skewness and L-kurtosis of X are

$$\tau_3 = \frac{1}{3}$$

$$\tau_4 = \frac{1}{6}.$$

Example 4.2.2 *(Logistic Distribution): Let X be a random variable with distribution function F given by,*

$$F(x) = \frac{1}{1 + e^{-(x-\mu)/\sigma}}, \quad x \in \Re.$$

Then the first four L-moments of X are,

$$\lambda_1 = \mu$$

$$\lambda_2 = \sigma$$

$$\lambda_3 = 0$$

$$\lambda_4 = \sigma/62.$$

Also L-skewness and L-kurtosis of X are

$$\tau_3 = 0$$

$$\tau_4 = \frac{1}{6}.$$

From Examples 4.2.1 and 4.2.2, it is clear that the populations L-moments are functions of the unknown parameters of the distributions. Hence L-moments are population parameters. The L-moments can be estimated using sample functions, such estimates we call as the sample L-moments.

Definition 4.2.3 *Let* $x_{1:n},\ x_{2:n}, \ldots,\ x_{n:n}$ *be the ordered observed samples, and define the* r^{th} *sample L-moments as*

$$l_k = \binom{n}{k} \underbrace{\sum\sum\cdots\sum}_{1\le i_1<i_2<\cdots<i_k\le n} k^{-1}\sum_{r=0}^{k-1}(-1)^r\binom{k-1}{r}x_{i_{k-r}:n}, \quad k = 1, 2, \ldots, n;$$

in particular,

$$l_1 = n^{-1}\sum_{i=1}^{n} x_i,$$

$$l_2 = \frac{1}{2}\binom{n}{2}^{-1}\underbrace{\sum\sum}_{i>j}(x_{i:n} - x_{j:n}),$$

$$l_3 = \frac{1}{3}\binom{n}{3}^{-1}\underbrace{\sum\sum\sum}_{i>j>k}(x_{i:n} - 2x_{j:n} + x_{k:n}),$$

$$l_4 = \frac{1}{4}\binom{n}{4}^{-1}\underbrace{\sum\sum\sum\sum}_{i>j>k>m}(x_{i:n} - 3x_{j:n} + 3x_{k:n} - x_{m:n}),$$

The sample L-moments defined above are exactly the U-statistics introduced by Hoeffding (1948) and they satisfy the important properties like the unbiasedness and the asymptotic normality. Wang (1996) gives a direct method to compute the sample L-moments and Hosking (1999) gives Fortran codes to compute sample L-moments for given data set by identifying the parameters. The sample L-moment ratios are denoted by $t_k = \frac{l_k}{l_2},\ k = 3, 4, \ldots$. For a complete treatment of L-moment estimation see Hosking (1986).

In applied statistics, to identify the possible distributions for a data L-moment ratio diagram is used by many authers (see Vogel and Fennessey, 1993; Sankarasubramanian and Srinivasan, 1999; and Hosking and Wallis, 1997; Tolikas and Brown, (2006), Tolikas et. al., (2007), etc.). The following remarks facilitate the use of L-moment ratio diagram in distribution identification problem. First, a distribution whose mean exists is characterized by its L-moments $\lambda_r : r = 1, 2, \ldots$. Second, the main features of a probability distribution can be well summarized by the first four L-moments, especially L-skewness and L-kurtosis. Moreover for skewed distributions, Hosking (1990) exemplified the advantage of using L-moments rather than using conventional moments.

In L-moment ratio diagram, the L-skewness is plotted in the X-axis and L-kurtosis in the Y-axis. The possible distributions of the data sets are those distributions whose theoretical L-moment curve passes through close to the sample L-moments data sets. For a two parameter probability distribution it makes a point on the graph. A curve represents a three parameter distribution and an area represents a four parameter distribution in the L-moment ratio diagram. For more details on this issue see Hosking (1990).

To verify the non-normality of the data in the chapter, we use the Jerque-Bera test. Jerque-Bera test, introduced by Bera et al. (1980, 1981), is a commonly used test for testing normality of the sample in both applied and theoretical statistics. The Jarque-Bera test is a goodness of fit statistic which measures departure from normality, based on the sample kurtosis and skewness. The goodness of fit statistic is defined as

$$JB = \frac{n}{2}\left\{s^2 + \frac{(k-3)^2}{4}\right\} \tag{4.2.1}$$

where n is the number of observations, s is the sample skewness, k is the sample kurtosis. The statistic JB has an asymptotic chi-square distribution with two degrees of freedom and can be used to test the null hypothesis that the data are from a normal distribution.

Shapiro-Wilk test, introduced by Shapiro and Wilk (1965), is a very popular test for testing normality of a given sample. Let $x_1,\ x_2, \ldots,\ x_n$ be a given sample and $x_{1:n},\ x_{2:n}, \ldots,\ x_{n:n}$ be its order statistic. Then the Shapiro-Wilk statistic is defined as,

$$W = \frac{(\sum_{i=1}^{n} a_i x_{i:n})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

where $\bar{x}$ is the sample mean and the constants a_i are given by,

$$(a_1,\ a_2, \ldots,\ a_n) = \frac{\mu' V}{\mu' V^{-1} V^{-1} \mu}$$

where $\mu = (\mu_1,\ \mu_2, \ldots,\ \mu_n)'$ and $\mu_1,\ \mu_2, \ldots,\ \mu_n$ are expected values of order statistics of independent and identically-distributed random variables sampled from the standard normal distribution, and V is the covariance matrix of those order statistics. The test reject the null hypothsis (data are sampled from a normal distribution) if W is too small. For more details about Shapiro-Wilk test, see Shapiro and Wilk (1965).

The Anderson-Darling test (Stephens, (1974)) is a non-parametric test based on empirical distribution function, which is used to test if a sample of data came from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov (K-S) test and gives more weight to the tails than does the K-S test. Let $X_{1:n},\ X_{2:n}, \ldots,\ X_{n:n}$ be ordered samples and F be the test distribution. Then the Anderson-Darling statistic is defined as,

$$A^2 = -n - S$$

where $S = \sum_{i=1}^{n} \frac{2i-1}{n}[\ln F(X_{i:n}) + \ln(1 - F(X_{n+1-i:n}))]$. The critical values for the Anderson-Darling test are dependent on the specific distribution that is being tested. For more details see Stephens (1974).

4.3 The Data and its Descriptive Analysis

The data used for analysis in this Chapter is the closing time Bombay stock exchange sensex index in every working day from 1/1/1982 to 12/31/2007. Figure 4.1 gives a time series representation of original data. To eliminate trend we use a transformation as in Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et. al.(2007), which is the logarithmic successive ratio of daily sensex returns from 1/1/1982 to 12/31/2007. That is, our data points are $t_i = log\frac{s_{i+1}}{s_i}$, where s_i is the sensex index at closing time of i^{th} working day from 1/1/1982. Figure 4.2 provides the time series representation of t_i and Table 4.3 is a descriptive analysis on data, which gives a number of interesting findings. An overall minimum of the data is -.137 which is at 4/13/1992 and a maximum of the data is .123 at 3/18/1992 due to several extraneous variables like Harshad mehtas scam, Ketan Parekh's securities scam, Pre economic liberalization effect (India witnessed new economic liberalization policy in the year 1995). We also use a boxplot test to detect the outliers, which is illustrated in Figure 4.3 - 4.5.

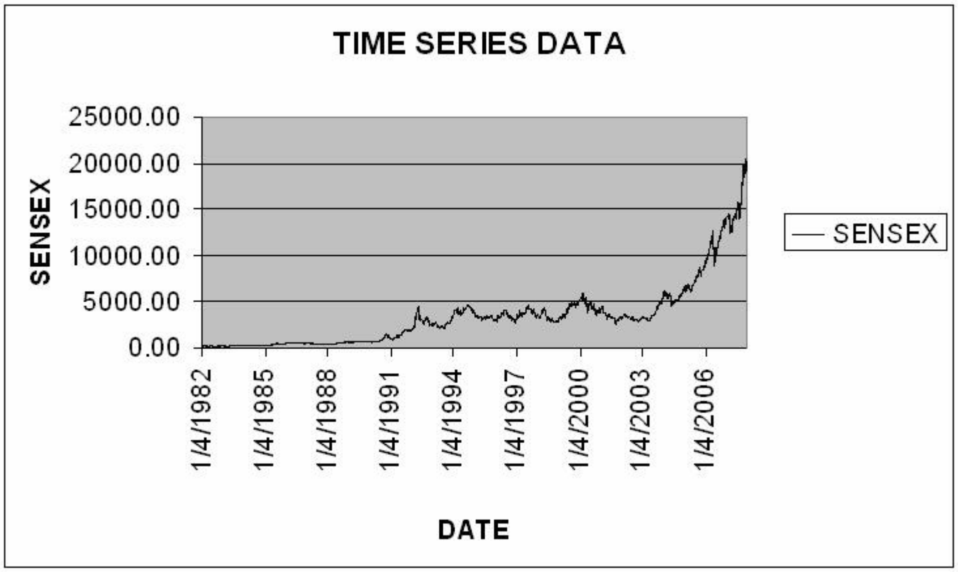

TIME SERIES DATA
SENSEX
25000.00
20000.00
15000.00
10000.00
5000.00
0.00
1/4/1982
1/4/1985
1/4/1988
1/4/1991
1/4/1994
1/4/1997
1/4/2000
1/4/2003
1/4/2006
DATE
SENSEX

Figure 4.1

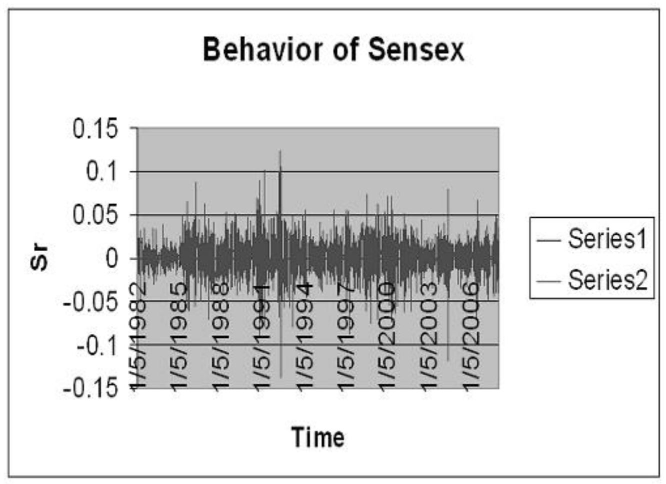
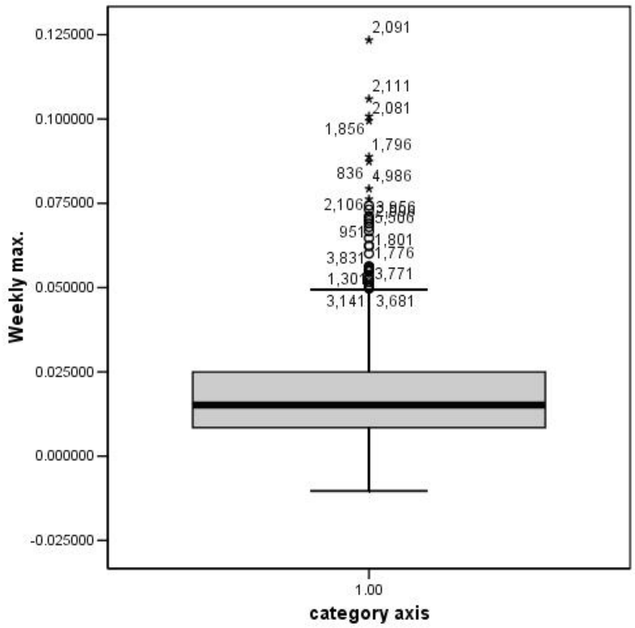

Behavior of Sensex
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
Sr
1/5/1982
1/5/1985
1/5/1988
1/5/1991
1/5/1994
1/5/1997
1/5/2000
1/5/2003
1/5/2006
Time
Series1
Series2

Figure 4.2

From Figure 4.2, the inconsistency of volatility indicates a number of factors in the Indian economy. In the early 1990s there was no apex body to control and regulate the share market activities in India and which resulted in the uneven fluctuation and ambiguous volatility. But later when the government launched SEBI (Securities Exchange Board of India) in April 1992 as a governing body to control, regulate and for the smooth functioning of stock exchange and capital market in Indian economy. After full fledged functioning of SEBI in 1993, we can see a balanced fluctuation from the figure. But from 1996 to 2002 we can observe an uneven fluctuation from the figure. The main reasons could be, in 1998, India experimented Pockran-2 and followed by economic sanction from the western countries and then India went through Cargil war and also India has faced political instability during this period. Three Prime ministers named H. P. Devagowda, I. K. Gujaral, and A. B. Vajpayee came to power and the respective governments collapsed within short period. Later again A.B. Vajpayee reelected as prime minster of India and thereafter India has gone to political stability and the figure directly shows political stability has a direct link with sensitivity of stock market activities. In 2002, MRTA (Monopoly Restricted Trade Practice), and FERA (Foreign Exchange Regulation Act) was abolished and FEMA (Foreign Exchange Management Act) was established. Also SEBI amended their internal policies. In 2004, India also underwent an election which resulted a volatility in the sensex which is clear from the picture.

Moreover, Table 4.3 shows various descriptive statistics for the daily returns and different time intervals of weekly, monthly, quarterly, half yearly, yearly minima as well as maxima (all intervals are approximated by number, that is, the size of the intervals are 5, 22, 66, 131, and 261 days respectively). Specifically, the number of observation, minimum, maximum, mean, standard deviation, skewness and kurtosis along with their standard errors, and normality test statistics Jeraque-Bera and Shapiro-Wilk are also shown in the table.



0.125000
0.100000
0.075000
0.050000
0.025000
0.000000
-0.025000
Weekly max.
1.00
category axis
2,091
2,111
2,081
1,856
1,796
836
4,986
951
1,801
3,831
1,776
1,301
3,771
3,141
3,681

Figure 4.3

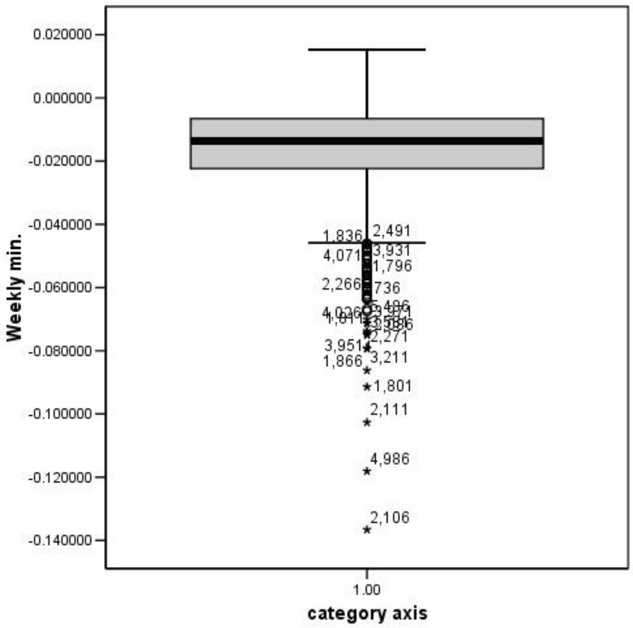

0.020000
0.000000
-0.020000
-0.040000
-0.060000
-0.080000
-0.100000
-0.120000
-0.140000
Weekly min.
1,836
2,491
4,071
3,931
1,796
2,266
736
2,271
3,951
1,866
3,211
1,801
2,111
4,986
2,106
1.00
category axis

Figure 4.4

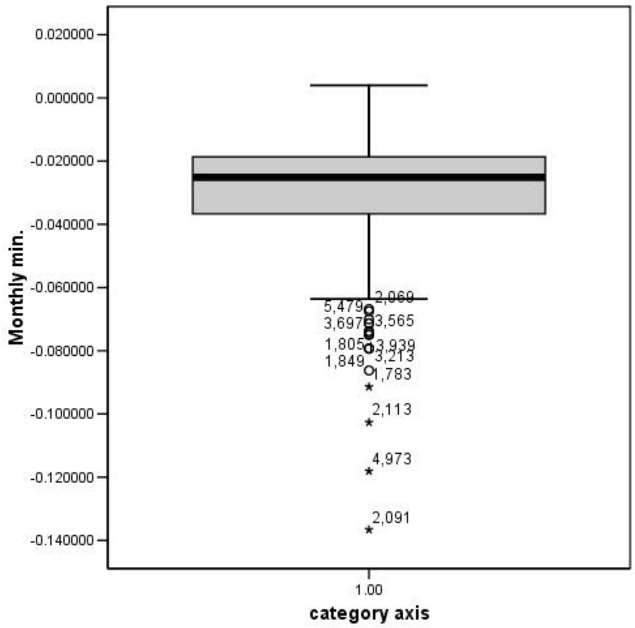

0.020000
0.000000
-0.020000
-0.040000
-0.060000
-0.080000
-0.100000
-0.120000
-0.140000
Monthly min.
2,069
5,479
3,697
3,565
1,805
3,939
3,213
1,849
1,783
2,113
4,973
2,091
1.00
category axis

Figure 4.5

Interval	N	Range	Min.	Max.	Mean	S. D	Sk. (S.E)	Kt. (S.E)	J-B
Daily	5897	0.260	-0.137	0.123	0.001	0.017	-0.144 (0.032)	4.614 (0.064)	1300.518
Weekly max	118	0.134	-0.01	0.123	0.018	0.014	1.868 (0.071)	6.793 (0.142)	2100.96
Monthly max	269	0.123	0	0.123	0.031	0.017	1.866 (0.149)	5.539 (0.296)	300.617
Quart. max.	90	0.11	0.01	0.12	0.042	0.021	1.434 (0.254)	2.514 (0.503)	32.616
Half-Yearly max	46	0.119	0.004	0.123	0.051	0.024	1.054 (0.350)	1.369 (0.688))	18.714
Yearly max.	23	0.093	0.031	0.123	0.062	0.023	0.998 (0.481)	0.987 (0.935)	11.587
Weekly min	1180	0.152	-0.137	0.015	-0.017	0.015	-2.019 (0.071)	7.873 (0.142)	3135.926
Monthly min	269	0.141	-0.137	0.004	-0.030	0.019	-2.016 (0.149))	6.239 (0.296)	417.39
Quart. min.	90	0.127	-0.137	-0.01	-0.042	0.023	-1.574 (0.254)	3.136 (0.503)	37.301
Half-Yearly min	46	0.136	-0.137	-0.001	-0.051	0.026	-1.091 (0.350)	1.915 (0.688)	13.638
Yearly min.	23	0.112	-0.137	-0.024	-0.063	0.028	-0.914 (0.481)	0.808 (0.935)	12.412

Table 4.1: Descriptive Statistics of maxima and minima in different time interval: N for number of samples, S. D for standard Deviation, Sk.(S.E) is for Skewness with its standard error, Kt.(S.E) is for kurtosis with its standard error, J-B for Jeraque-Bera statistic and S-W for Shapiro-Wilk statistic

Again, the table shows the mean of different time intervals of maximum lies between 0.018 and 0.062 and which is positive and its counterpart of minimum intervals shows a range of -0.017 to -.063, which is negative. A large kurtosis of different time intervals indicates a non-normality behavior of the data. For all intervals, Jeraque-Bera statistic shows large numbers which are clearly greater than 10 and which suggests a non-normality behavior of the data. But Brys et al.(2004) have shown that the Jeraque- Bera statistic is highly sensitive to the outliers. Figure 4.3 - 4.5 shows the presents of outliers. So we use another test namely, Shapiro-Wilk test to confirm non-normality and the p-values clearly indicates the non-normality of the data.

4.4 Finding Appropriate Distributions for BSE Share Market Data Using L-moments

In Section 4.3, we confirmed the non-normality of the BSE sensex data. In this section, we use L-moment ratio diagram to find appropriate distributions for testing goodness of fit of the data. In Chapter 3, we described GEV, GP, and GL distributions as probable models for extremal behavior of any data. For the BSE sensex data, after eliminating trend, among the above three models we try to find out that models which are more appropriate for testing goodness of fit. We use Fortran codes from Hosking (1999) for the computational works.

Since GEV, GP, and GL are three parameter distributions, a curve represents each distributions in the L-moment ratio diagram. We divided the weekly maxima and minima of the share market data point in to ten intervals and took the sample L-moment ratio's t_3 and t_4 (Hosking (1990)) and these sample L-moment ratios are plotted in the L-moment ratio diagram in Figure 4.6. The curves in the figure consist of theocratical L-moment ratio curve of Generalized extreme value distribution, Generalized logistic distribution, and Generalized Pareto distribution. The plotted points in the figure corresponds to sample L-moment ratios. Figure 4.6 visually gives the sample points are comparatively near to GL and GEV distributions, .

For weekly maximum sample points, the diagram shows that both GL and GEV distributions are always equal fit but for weekly minimum gives an approximate idea that GL is better than GEV. So we need to analyze more on minimum data to know whether GL is an adequate fit for minimum than GEV. In the next section we use Anderson darling test to find which is the better distribution for minimum data among GL and GEV.

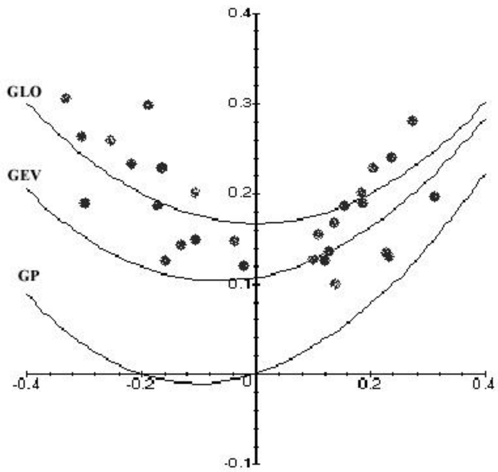


Figure 4.6

4.5 Goodness of Fit Test for the BSE Data

The L-moment ratio diagram in Section 4.4 revealed that GL and GEV are appropriate models for the extremal behavior of the BSE data. In this section, we use Anderson-Darling test for comparing goodness of fit of GEV and GL distributions in the extremal behavior of BSE data. Since the test is based on empirical distribution function (EDF) and among all the well known tests based on EDF, Anderson-Darling test has the highest power in testing normality against a number of alternatives when the parameters are unknown (Stephens, 1974). Table 4.2 shows a comparison between GEV(max) and GL in weekly maximum data. Out of fifteen intervals, GL is adequate in eight intervals and seven for GEV(max). Here GEV(max) and GL perform almost equally. However for weekly minimum data, the performance is seriously different.

To verify the findings of Gettinby et al.(2004), Tolikas and Brown (2006) and Tolikas et al.(2007) in Indian context, we use to test GEV(max) in minimum data. From Table 5.3, GL is adequate for weekly minimum data. Out of fifteen, GEV(max) performs better only in one interval than GL which clearly agrees the findings of Gettinby et al.(2004), Tolikas and Brown (2006) and Tolikas et al.(2007). In their papers, on minimum returns, Generalized extreme value distribution performs better only on one interval than GL distribution.

Table 4.4 shows a comparison between performance of GEV(min) and GL on minima returns. For fifteen intervals, GL provides an adequate fit in eleven intervals than GEV(min), which indicates that GL distribution provides a better fit than GEV(min). However, GEV(min) provides an adequate fit in 11 out of 15 intervals and in which most of the intervals p-values are large, which shows that GEV(min) also provides a model for minimum returns. From Table 5.3, GEV(max) provides an adequate fit for only 1 out of 15 intervals. So GEV(max) cannot be used as a model

	.	GEV(max)				GL				.
Intervals	n	loca	scale	shape	p	loca	scale	shape	p	Best fit
1	78	0.006	0.006	0.116	.412	0.009	0.004	-0.247	.347	GEV(max)
2	78	0.007	0.007	0.228	.319	0.010	0.006	-0.325	.229	GEV(max)
3	78	0.013	0.012	-0.018	.374	0.018	0.008	-0.158	.574	GL
4	78	0.013	0.01	-0.089	.625	0.017	0.007	-0.117	.548	GEV(max)
5	78	0.015	0.012	0.109	.068	0.021	0.009	-0.242	.356	GL
6	78	0.020	0.012	0.187	.024	0.025	0.009	-0.296	.000	GEV(max)*
7	78	0.011	0.008	0.118	.054	0.015	0.006	-0.248	.014	GEV(max)
8	78	0.009	0.008	-0.005	.318	0.013	0.006	-0.166	.089	GEV(max)
9	78	0.012	0.010	0.031	.236	0.016	0.007	-0.190	.709	GL
10	78	0.016	0.012	-0.053	.106	0.021	0.008	-0.136	.073	GEV(max)
11	78	0.012	0.011	-0.047	.149	0.017	0.007	-0.14	.155	GL
12	78	0.008	0.008	-0.035	.524	0.012	0.005	-0.147	.791	GL
13	78	0.011	0.006	0.092	.252	0.015	0.005	-0.230	.913	GL
14	78	0.010	0.005	-0.094	.238	0.013	0.003	-0.111	.498	GL
15	88	0.012	0.009	0.07	.039	0.016	0.006	-0.216	.115	GL

Table 4.2: Performance of GEV(max) and GL distribution for Weekly maximum data

	.	GEV				GL				.
Intervals	n	loca	scale	shape	p	loca	scale	shape	p	Best fit
1	78	-0.011	0.011	-0.976	.000	-0.009	0.005	0.324	.465	GL
2	78	-0.01	0.013	-1.059	.000	-0.007	0.005	0.356	.000	-
3	78	-0.024	0.014	-0.478	.000	-0.019	0.007	0.103	.869	GL
4	78	-0.019	0.010	-0.369	.46	-0.016	0.006	0.044	.95	GL
5	78	-0.021	0.021	-0.935	.000	-0.015	0.009	0.308	.225	GL
6	78	-0.029	0.023	-0.817	.000	-0.023	0.010	0.26	.938	GL
7	78	-0.017	0.010	-0.485	.034	-0.014	0.005	0.106	.152	GL
8	78	-0.018	0.01	-0.334	.172	-0.015	0.005	0.027	.219	GL
9	78	-0.02	0.016	-0.718	.000	-0.015	0.007	0.217	.932	GL
10	78	-0.021	0.015	-0.632	.000	-0.017	0.007	0.178	0.58	GL
11	78	-0.026	0.019	-0.631	.744	-0.021	0.009	0.177	.254	GEV(max)
12	78	-0.014	0.010	-0.746	.000	-0.011	0.005	0.229	0.125	GL
13	78	-0.016	0.015	-0.899	.000	-0.012	0.006	0.293	.348	GL
14	78	-0.011	0.009	-0.567	.000	-0.019	0.006	0.293	.494	GL
15	88	-0.019	0.017	-0.775	.000	-0.013	0.008	0.242	.00	-

Table 4.3: Performance of GEV(max) and GL distribution for Weekly minimum data

	.	GEVmin				GL				.
Intervals	n	loca	scale	shape	p	loca	scale	shape	p	Best fit
1	78	0.006	0.006	0.226	.396	-0.009	0.005	0.324	.465	GL
2	78	0.004	0.007	0.270	.000	-0.007	0.005	0.356	.000	-
3	78	0.014	0.011	-0.107	.112	-0.019	0.007	0.103	.869	GL
4	78	0.012	0.009	-0.207	.672	-0.016	0.006	0.044	.95	GL
5	78	0.010	0.012	0.204	.395	-0.015	0.009	0.308	.225	GEVmin
6	78	0.017	0.014	0.135	.599	-0.023	0.010	0.26	.938	GL
7	78	0.010	0.0082	-0.101	.031	-0.014	0.005	0.106	.152	GL
8	78	0.011	0.009	-0.235	.176	-0.015	0.005	0.027	.219	GL
9	78	0.010	0.011	0.071	.701	-0.015	0.007	0.217	.932	GL
10	78	0.013	0.011	0.012	.052	-0.017	0.007	0.178	0.58	GL
11	78	0.015	0.014	0.0113	.541	-0.021	0.009	0.177	.254	GEVmin
12	78	0.008	0.007	0.090	.009	-0.011	0.005	0.229	0.125	GL
13	78	0.008	0.008	0.183	.107	-0.012	0.006	0.293	.348	GL
14	78	0.006	0.007	-0.036	.401	-0.019	0.006	0.293	.494	GL
15	88	0.009	0.0117	0.117	.003	-0.013	0.008	0.242	.00	GEVmin

Table 4.4: Performance of GEVmin and GL distribution for Weekly minimum data

for minimum returns but GL and GEV(min) provides a good fit and comparing GL and GEV(min), GL provides a better fit.

Table 4.5 represents an Anderson-Darling test for GEV(max) and GL in monthly minimum returns data. Out of ten, GL provides a better fit in seven intervals. Also in five intervals GEV(max) shows a p-value of 0.0 suggests GEV(max) is an inadequate model for monthly minimum data.

Table 4.6 shows that GL provides a better fit in six interval comparing to GEV(min).

	.	GEV(max)				GL				.
Intervals	n	loca	scale	shape	p	loca	scale	shape	p	Best fit
1	26	-0.021	0.013	-1.17	.000	-0.018	0.005	0.396	.206	GL
2	26	-0.036	0.02	-0.536	.72	-0.03	0.01	0.132	.511	GEV(max)
3	26	-0.030	0.008	-0.426	.129	-0.027	0.004	0.075	.525	GL
4	26	-0.052	0.033	-0.821	.000	-0.043	0.015	0.261	.517	GL
5	26	-0.027	0.011	-0.544	.542	-0.023	0.005	0.135	.350	GEV(max)
6	26	-0.031	0.0143	-0.827	.000	-0.027	0.006	0.264	.339	GL
7	26	-0.038	0.021	-0.851	.42	-0.032	0.009	0.274	.258	GEV(max)
8	26	-0.033	0.018	-0.689	.217	-0.028	0.008	0.204	.625	GL
9	26	-0.021	0.015	-1.520	.000	-0.018	0.005	0.507	.122	GL
10	35	-0.029	0.016	-0.509	.000	-0.023	0.007	0.256	.886	GL

Table 4.5: Performance of GEV(max) and GL distribution for monthly minimum data

	.	GEVmin				GL				.
Intervals	n	loca	scale	shape	p	loca	scale	shape	p	Best fit
1	26	0.015	0.006	0.323	0.324	-0.018	0.005	0.396	.206	GEVmin
2	26	0.024	0.015	-0.060	.708	-0.03	0.01	0.132	.511	GEVmin
3	26	0.025	0.007	-0.153	.255	-0.027	0.004	0.075	.525	GL
4	26	0.034	0.020	0.137	.551	-0.043	0.015	0.261	.517	GEVmin
5	26	0.020	0.008	-0.054	.419	-0.023	0.005	0.135	.350	GEVmin
6	26	0.024	0.009	0.141	.208	-0.027	0.006	0.264	.339	GL
7	26	0.027	0.012	0.156	.257	-0.032	0.009	0.274	.258	GL
8	26	0.023	0.012	0.051	.593	-0.028	0.008	0.204	.625	GL
9	26	0.015	0.006	0.467	.049	-0.018	0.005	0.507	.122	GL
10	35	0.019	0.013	-0.081	.695	-0.023	0.007	0.256	.886	GL

Table 4.6: Performance of GEVmin and GL distribution for monthly minimum data

While GEV(min) provides an adequate fit in nine out of ten intervals shows GEV(min) is also a model for monthly minima.

From the above analysis of weekly and monthly minimum data, we conclude that GEV(max) is not a fit for the minimum data. So in the following analysis we consider GL and GEV(min) distribution only.

Comparing GL and GEV(min) in quarterly minimum data, Table 4.7 indicates the same result, that is, GL provides a better fit than GEV(min). Further GEV(min) and GL gives a good model in three out of four intervals, which shows both distribution can use to model quarterly data and comparing these distributions GL is a better model.

On analysis of half yearly data, Table 4.8 clearly shows GL provides a good fit

	.	GEVmin				GL				.
Intervals	n	loca	scale	shape	p	loca	scale	shape	p	Best fit
1	22	0.029	0.015	-0.074	.419	-0.035	0.01	0.123	.529	GL
2	22	0.034	0.015	0.3072	.103	-0.04	0.012	0.383	.186	GL
3	22	0.039	0.018	-0.073	.054	-0.046	0.011	0.124	.052	GEVmin
4	24	0.0252	0.010	0.263	.01	-0.029	0.008	0.350	.028	GL*

Table 4.7: Performance of GEV(min) and GL distribution for quarterly minimum data. * indicates neither distribution fit the data.

	.	GEVmin				GL				.
Intervals	n	loca	scale	shape	p	loca	scale	shape	p	Best fit
1	23	0.038	0.019	0.130	.452	-0.046	0.013	0.256	.569	GL
2	23	0.040	0.019	-0.008	.041	-0.047	0.013	0.165	.195	GL

Table 4.8: Performance of GEV(min) and GL distribution for half yearly minimum data

comparing to GEV(min). Finally, Table 4.9 illustrates GL provides a better for yearly minimum data compared to GEV(min). Both distribution fit for the data at 0.05 level, but small p-values shows a poor fit.

The above analysis reveal that, the minimum distribution empirically converges to GEV(min), but the rate of convergence is very poor. Also GL shows a superiority for fitting minimum returns than GEV(min) distribution.

	.	GEVmin				GL				.
Intervals	n	loca	scale	shape	p	loca	scale	shape	p	Best fit
1	23	0.049	0.023	0.001	.109	-0.058	0.015	0.171	.131	GL

Table 4.9: Performance of GEV(min) and GL distribution for yearly minimum data

4.6 Conclusion

The analysis of BSE sensex data in this chapter reveals that GL gives a better fit for different blocks of minimum than GEV. From the analysis we recommend that instead of GEV, GL provides better estimate in VaR estimation, hedging, etc. But the analysis also suggest that GEV is also a good model for modeling minima and comparing these two distributions GL is better than GEV. This finding is different from the findings of Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et. al.(2007), where they found that GEV is not a model for minimum of stock exchange data. Also extreme value theory holds for finance data but the rate of convergence is very poor.

Another interesting empirical finding is that the distribution of minima, can be approximated by the GL distribution. That is, GL is a good asymptotic approximate distribution for minimum of the data in financial context. This is a data verification of the convergence of geometric minima to GL described in Chapter 3. The approximation is significantly better than GEV(min) distribution, this is verified using Anderson-Darling test. But we found that the data contains outliers and in some time intervals the sample size are very small. So we have to find whether this is the reason for such a result. In other words we have to check the performance of Anderson-Darling test when data contains outliers and sample size is small. In Chapter 5, we introduce a new goodness of fit measure which is not sensitive to outliers and small sample size. We also analyzed the same BSE data using the new measure.

Chapter 5

A Goodness of Fit Measure Using L-moment Ratio Diagram

5.1 Introduction

In Chapter 4, we empirically verified the results of Chapter 3 for the BSE sensex data using Anderson-Darling test. During the analysis we found that the data contains outliers and there are situations where sample sizes are very small. In this chapter, we introduce a new goodness of fit test using L-moment ratio diagram, which performs better when sample size is small. We empirically verify that the test is robust to outliers. We use this test to analyze the extrmal behavior of BSE sensex data. The content of this chapter is Nidhin and Chandran (2010b).

In parameter estimation, the method of moments is a commonly used method. It is a simple method of estimation however is very sensitive to outliers, especially when the sample size is small (see Pearson(1963), Hosking (1990)). The reason for this limitation is that moments are calculated by power functions. An alternative approach to tackle this difficulty is given by Hosking (1990), with the introduction of method of L-moments. Hosking has shown that first four L-moments of a distribution measures respectively the average, dispersion, symmetry and tail weight (or peakedness) of the distribution. In the last two decades, L-moments have turn out to be a popular tool in parametric estimation and distribution identification problems in different scientific areas (see e.g. Hosking (1990), Stedinger et al., (1993), Hosking and Wallis, (1997), Tolikas et al., (2006)). Section 5.2 introduces the basic concepts required in this chapter. This section also describes the existing goodness of fit theory based on L-moments in the statistical literature. Section 5.3 introduces the new goodness of fit statistic for two parameter distribution. A comparative study of this new statistic w.r.t. other existing well known goodness of fit statistics using monte carlo simulation is done in this section. In Section 5.4, we extend the idea to distributions involving more than two parameters. Section 5.5, applies this new test to the BSE sensex data described in Chapter 4 to identify the performance of GL distribution in modeling extremes. Finally, Section 5.6 concludes the finding of the Chapter.

5.2 Basic Concepts

The main objective of this chapter is to introduce a new goodness of fit measure and study its performance w.r.t. existing goodness of fit statistics. The commonly used goodness of fit statistics for testing normality are Jerque-Bera statistic, Anderson-Darling statistic and Robust statistic introduced by Brys et al (2004). The Jerque-Bera and Anderson-Darling test statistics are introduced in Chapter 4. The robust statistic is introduced in this section. The robust statistic for testing goodness of fit is based on three basic measures of data, which are medcouple (MC), left medcouple (LMC) and right medcouple (RMC) for a given data set. Below we define these concepts.

Definition 5.2.1 *Let $x_1, x_2, \ldots, x_n$ be n independent observations from a continuous univariate distribution F and let $m_F = F^{-1}(.5)$. Then the sample Medcouple (MC) of the sample $(x_1, x_2, \ldots, x_n)$ is,*

$$MC(F) = med_{y_1 \leq m_F \leq y_2} h(y_1, y_2)$$

where y_1 and y_2 are sampled from F and the kernel function is defined as

$$h(y_i, y_j) = \frac{(y_j - m_F) - (m_F - y_i)}{y_j - y_i}.$$

Definition 5.2.2 *Let $x_1, x_2, \ldots, x_n$ be n independent observations from a continuous univariate distribution F, let $m_F = F^{-1}(.5)$ and MC be as defined in Definition (5.2.1). Then the Left Medcouple (LMC) of the sample $(x_1, x_2, \ldots, x_n)$ is*

$$LMC(F) = -MC(x < m_F)$$

and the Right Medcouple (RMC) of the sample $(x_1, x_2, \ldots, x_n)$ is

$$RMC(F) = MC(x > m_F).$$

The population medcouples can be defined in a similar way. The robust statistic for goodness of fit introduced by Brys et al. (2004) is defined below.

Definition 5.2.3 *Let $x_1, x_2, \ldots, x_n$ be n independent observations from a continuous univariate distribution F and let $m_F = F^{-1}(.5)$, MC, LMC and RMC are as in Definitions (5.2.1) and (5.2.2) respectively. Then the robust statistic for goodness of fit is*

$$T = n(w - W)'\Sigma^{-1}(w - W)$$

for $w = (MC, LMC, RMC)'$, W is its population parameters and Σ is its asymptotic variance.

For more details about the above concepts see Brys et al.(2004). The remaining part of this section reviews existing goodness of fit theory based on L-moments available in the literature.

Gail and Gastwirth (1978) proposed a goodness of fit measure based on first and second L-moment for exponential distribution. They concluded that the statistic has very good power and the advantage of computation easiness. This also has the scale property, ease extension to a test free of scale and location parameters, availability of exact percentiles and insensitivity to small rounding or truncation measurement errors. Hosking (1986) proposes a number of goodness of fit test statistic using L-moments. For testing goodness of fit of Uniform distribution for a given data, he suggests the statistics $(35n/6)^{1/2}t_3$, $35n(t_3^2/6+t_4^2/2)$, where t_3 and t_4 are sample L-skewness and sample L-kurtosis defined in Section 4.2. Also he proposes probability weighted moment-based variant of Neyman's smooth test statistic, which is defined as:

Definition 5.2.4 *let X be a random variable with distribution function $F(x;\theta)$ and quantile function $G(u;\theta)$, where $\theta=(\theta_1,\theta_2,\cdots,\theta_p)'$ is a vector of real parameters with true (but unknown) value θ_0. Let $\kappa_m(u)$, $m=1,2,\ldots,p$ be a set of linearly independent polynomials. Let*

$$\gamma_m(\theta)=\int_0^1 \kappa_m(u)G(u;\theta)du,\ m=1,2,\ldots,p$$

be probability weighted moments. The goodness of fit test statistic for the distribution F is

$$n(d-\mu)'\hat{\Sigma}^{-1}(d-\mu)$$

where n is the sample size, $d=n^{-1}\sum_{i=1}^{n}\delta_j^{(n)}\hat{U}_{j:n}$, $\delta_j^{(n)}=\delta(\frac{j}{n+1})+O(n^{-1})$, $\delta:[0,1]\rightarrow\Re^m$ is a continuous bounded function, $\hat{U}_{j:n}=F(X_{j:n};\hat{\theta})$ and $\hat{\theta}$ is the probability weighted moment estimator of the parameter θ,

$$\mu=\int_0^1 \delta(u)\ u\ du,\quad \Sigma=2\int\int_{u<v}\eta(u)\{\eta(v)\}'u(1-v)dudv,$$

$$\eta(u) = \delta(u) - \sum_{i=1}^{p}\sum_{j=1}^{p}\{\int_0^1 \delta(u)\frac{\delta G(u;\theta_0)}{\delta\theta_i}\frac{du}{G'(u;\theta_0)}\}H^{ij}(\theta_0)\kappa_j(u)G'(u;\theta_0)$$

here $H^{ij}(\theta)$ is the $(i,j)^{th}$ element of the $m * m$ matrix $H(\theta)$ whose $(i,j)^{th}$ element is $H_{ij}(\theta) \equiv \delta\gamma_i(\theta)/\delta\theta_j$.

The statistic has some disadvantages. The main disadvantage is the slow convergence of the mean of $d - \mu$ to zero, so it requires large sample data for asymptotic approximation. Another disadvantage is that calculation of the elements of Σ usually requires some numerical iteration. Another goodness-of-fit measure suggested by Hosking and Wallis (1993, 1997), using the L-kurtosis, is defined as:

Definition 5.2.5 *For each distribution the goodness of fit statistic is*

$$Z = (\tau_4 - t_4 + \beta_4)/\sigma_4 \tag{5.2.1}$$

where τ_4 is the L-kurtosis of the fitted distribution to the data using the candidate distribution, and $\beta_4 = \sum_{i=1}^{N_{sim}} \bar{t_4}^{(m)}/N_{sim}$ is the bias of $\bar{t_4}$ estimated using the simulation technique as before with $\bar{t_4}^{(m)}$ being the sample L-kurtosis of the mth simulation, and

$$\sigma = [(N_{sim} - 1)^{-1}\{\sum_{i=1}^{N_{sim}}(\bar{t_4}^{(m)} - \bar{t_4})^2 - N_{sim}\beta_4^2\}]^{1/2}.$$

The L-moment ratio diagram has commonly been used for visual identification of distributions to be considered for the goodness of fit of the given data. In L-moment ratio diagram, the L-skewness is plotted in the X-axis and L-kurtosis in the Y-axis. The possible distributions of the data sets are those distributions whose theoretical L-moment curve passes close to the sample L-moments data sets. For a two parameter probability distribution it makes a point on the graph. A curve represents a three parameter distribution and an area represents a four parameter distribution in the L-moment ratio diagram. For more details see Hosking (1990). This idea of distribution identification to a given data set has been used by many authors, see for example, Vogel and Fennessey, (1993); Sankarasubramanian and Srinivasan, (1999); Hosking and Wallis, (1997); Gettinby et al. (2004), Tolikas and Brown, (2006), Tolikas et al. (2007), etc.). We try to use there ideas to define a new goodness of fit measure.

5.3 Goodness of Fit Measure for Two Parameter Distribution

In Section 5.2, we discussed the use of L-moment ratio diagram, as a visual identification tool to find the best distribution for a given data. In this section, we define a distance measure using L-moment ratio diagram and that is used to test the goodness of fit of a two parameter distribution for a given data. We also show the advantage of the test statistic in testing goodness of fit problems.

Let $x_1,\ x_2, \ldots,\ x_n$ be a given sample and it is decided to test whether the sample is from a given distribution F, where F involve only two parameters. We assume that the mean exists for the distribution F, hence all L-moments of F exist as discussed in Section 4.2. Now to test,

$$\begin{cases} H_0: & \text{The data are sampled from distribution } F \\ V/s & \\ H_1: & \text{The data are not sampled from distribution } F. \end{cases}$$

We define a new goodness of fit statistic using L-moments (denoted by T^L) below,

$$T^L = n(w - W)'A^{-1}(w - W) \tag{5.3.2}$$

for $w = (t_1, t_2)'$ be the estimators of $W = (\tau_1, \tau_2)'$ where t_1 and t_2 are the sample estimators of L-Skewness and L-Kurtosis respectively and τ_1 and τ_2 are the population L-Skewness and population L-Kurtosis respectively. Here A is the asymptotic covariance matrix of (t_3, t_4). That is, $A = ((A_{rs})), (r, s = 3, 4)$(we use 3 and 4 for notational similarity) which is defined as

$$A_{rs} = (B_{rs} - \tau_r B_{2s} - \tau_s B_{2r} + \tau_r \tau_s B_{rs})/\lambda_2^2$$

and B_{rs} is defined as,

$$B_{rs} = \int \int_{x<y} \{P^*_{r-1}(F(x))P^*_{s-1}(F(y)) + P^*_{s-1}(F(x))P^*_{r-1}(F(y))\}F(x)\{1 - F(y)\}dxdy$$

where $P^*_r(x)$ being the r_{th} shifted Legendre polynomial defined as,

$$P^*_r(x) = \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} x^k$$

Theorem 5.3.1 *Let t_3 and t_4 are sample L-skewness and L-kurtosis respectively also τ_3 and τ_4 are population L-skewness and L-kurtosis respectively. Then the statistic*

$$T^L = n(w - W)' A^{-1} (w - W)$$

follows chi-square distribution with two degrees of freedom asymptotically.

Proof: Given that t_3, t_4, τ_3 and τ_4 are sample l-skewness, sample l-kurtosis, population l-skewness and population l-kurtosis respectively. Then by Hosking (1990),

$$n^{1/2}[(t_3 - \tau_3)(t_4 - \tau_4)]' \to N(0, A) \tag{5.3.3}$$

where A is defined as above. Simple argument tells us that,

$$T^L = n(w - W)' A^{-1} (w - W) \approx \chi^2(2). \ \square$$

From Theorem (5.3.1), the statistic T^L has an asymptotic chi-square distribution and can be used to test the null hypothesis that the data are from any particular

Distribution	T^L
Uniform	$35n(\frac{t_3^2}{6}+\frac{t_4^2}{2})$
Exponential	$n(\frac{40}{3}t_3^2+\frac{224}{15}t_4^2-\frac{70}{3}t_3t_4-5t_3+\frac{14}{5}t_4+\frac{3}{5})$
Gumbel	$n(6.791776t_3^2+11.497577t_4^2-10.70669t_3t_4-0.6975599t_3+1.639404t_4+0.182541)$
Laplace	$n(\frac{t_3^2}{0.4593}+\frac{(t_4-0.2361)^2}{0.1565})$
Logistic	$n(\frac{t_3^2}{0.2899}+\frac{(t_4-1/6)^2}{0.1228})$
Normal	$n(\frac{t_3^2}{0.1866}+\frac{(t_4-0.1226)^2}{0.0883})$

Table 5.1: Goodness of fit statistic of different two parameter distributions.

distribution. By Hosking (1986), the asymptotic theory usually provides a good approximation to the exact distribution for samples of size $n \geq 50$, and is often adequate even for $n = 20$. This fast convergence is one of the important advantage of the statistics T^L, especially in extreme value theory.

Table 5.1 presents the goodness of fit statistic of some of the important two parameter distributions. The distributions are Uniform, Exponential, Gumbel, Laplace, Logistic and Normal. We can derive T^L statistic for any two parameter distribution, for which mean exists.

5.3.1 Goodness of Fit Statistic for Normal Distribution

In this section we study the performance of the statistic T^L, defined in equation 5.3.2, as a goodness of fit test statistic for testing Normality of the given data. We compare T^L with three statistics namely the Jerque- Bera statistic, Robust estimator for testing normality and Anderson-Darling statistic. This is done using Monte-Carlo simulation. Among tests using empirical distribution function, we consider Anderson-Darling test for comparison because among all the well known tests based on empirical distribution functions, Anderson-Darling test has the highest power for testing normality against a number of alternatives when the parameters are unknown (Stephens, 1974). Using simulation, we verify that the power of the test statistic T^L is better than the power of the other three widely used alternatives such as Jarque-Bera, Robust Statistic and Anderson-Darling statistic. We also verify that T^L is moderately less sensitive to outliers. We simulate 1000 samples each having 100 observations from a Tukey distribution, which includes both normal and non-normal distributions. For $g = h = 0$ the distribution is normal and for all other values of g and h the distribution is non-normal. We define the Tukey distribution (Hoaglin et al. (1985)) below:

Definition 5.3.1 *A random variable Y is said to follow Tukey's gh-distribution if Y is defined as,*

$$Y = \begin{cases} \frac{e^{gX}-1}{g} e^{\frac{hX^2}{2}} & \text{if} \quad g \neq 0, \\ X e^{\frac{hX^2}{2}} & \text{if} \quad g = 0. \end{cases}$$

where X follows standard normal distribution.

The Tukey's distribution has several properties. The parameters g and h determine its skewness and kurtosis respectively. Also it is easy to simulate from this distribution even though it is hard to explicitly state the probability density function. Given this flexibility, the gh family may be extremely useful in modelimg multiple imputations.

Now consider a sample of size n from the Tukey's gh distribution and we want to test the hypothesis that

$$\begin{cases} H_0: & \text{The data are sampled from normal population} \\ V/s & \\ H_1: & \text{The data are not sampled from normal population.} \end{cases}$$

Then T^L statistic is

$$T^L = n\left\{\frac{t_3^2}{0.1866} + \frac{(t_4 - 0.1226)^2}{0.0883}\right\}. \tag{5.3.4}$$

Using different values of g and h we simulate 1000 samples each of size 100 and calulate the power of T^L, Jerque- Bera and Robust estimator and Anderson-Darling. In Table 5.2, G g,h represent Tukey's gh distribution with g and h as parameters. The values represents the fraction of 1000 samples rejected normality assumption under 0.05 significance level. For all values of g and h, other than g=0 and h=0, representing non-normality, T^L take higher values. The values represent (other than G 0,0) exactly the power of the tests and for all values of T^L performs significantly better. Notice that the power of robust estimator is very low compared to T^L statistic.

Table 5.3 represents the performance of the tests in contaminated samples. The contamination is done by adding $100(1-\epsilon)\%$ of the contaminated data. We have taken $\epsilon = 5\%$. In Table 5.3 N, μ, σ represents the 5% of the standard normal samples replaced by samples from $N(\mu, \sigma)$. The table also represents the fraction of 1000 samples rejecting normality assumption under 0.05 significance level. From the table, it is clear that T^L shows a moderate performance to outliers. So for small sample testing T^L shows a better performance compared to other tests for testing normality.

The simulation study in this section reveals that the statistic posses computational higher power and a moderate performance to the outliers. Also the statistic has a asymptotic chi-square distribution and has a fast convergence. These important properties make the statistic is highly useful in statistical inference, especially in the

Tukey's gh distribution	JB	T^L	MC	AD
G 0,0.5	0.996	0.999	0.578	0.987
G 0,0.75	1.000	1.000	0.852	1.000
G 0,1	1.000	1.000	0.957	1.000
G 0.5,0	0.981	0.992	0.790	0.676
G 0.75,0	1.000	1.000	0.997	0.989
G 1,0	1.000	1.000	1.000	1.000
G 0.5,0.5	1.000	1.000	0.796	0.995

Table 5.2: Performance of fraction of 1000 samples of Turky distribution for different parameters. JB-Jerque Bera Statistic, T^L - Statistic Using L-moments, MC - Robust Estimator and AD - Anderson-Darling statistic

Normal (contaminated)	JB	T^L	MC	A-D
N,2,1	0. 195	0. 148	0.088	0.120
N,-2,1	0.197	0. 165	0.085	0.112

Table 5.3: Performance of fraction of 1000 contaminated samples of normal distribution. JB-Jerque Bera Statistic, T^L- Statistic Using L-moments, and MC- Robust Estimator and AD - Anderson-Darling statistic

extreme value theory.

5.4 Goodness of Fit Measure for Three Parameter Distributions

In this section we try to generalize the goodness of fit statistic to d.f. involving three parameter distribution. A two parameter probability distribution makes a point on the L-moment ratio diagram. That is we have only one L-skewness and L-kurtosis, also from the previous section we can use the statistic for goodness of fit tests. However, in L-moment ratio diagram, a curve represents for a three parameter distribution and an area represents for a four parameter distribution. To overcome this difficulty, we test the hypothesis in two steps. First, we identify the best point of population L-skewness and population L-kurtosis (from test population) for the goodness of fit test. Next, we use this population L-skewness and population L-kurtosis to test goodness of fit by the statistic from L-moments. Here we discuss only for three parameter distribution, the same argument we can be extended to four or more parameter distributions.

Again our main testing problem is,

$$\begin{cases} H_0: & \text{The data are sampled from a population with distribution function } F \\ V/s & \\ H_1: & \text{The data are not sampled from a population with distribution function } F. \end{cases}$$

where F is a three parameter distribution. The first step consists of finding best point of population L-skewness and population L-kurtosis (on the l-moment ratio curve) that the data can choose. We denote these point as (τ_3^F, τ_4^F). Below we discuss some methods to find (τ_3^F, τ_4^F).

5.4.1 Graphical Method

This method is a very simple method to find (τ_3^F, τ_4^F). From the L-moment ratio diagram, we can calculate (τ_3^F, τ_4^F) as a visual identification which L-moment ratio's are closer to (t_3, t_4), the observed L-skewness and L-kurtosis point. For example, see Figure 5.4.1, here we are testing goodness of fit statistic for a data to generalized logistic distribution. The point (t_3, t_4) represents the observed pair of L-skewness and L-kurtosis. From the figure, the point (.23,.25075), as a visual identification, is the nearest point on the theoretical curve than other points on the curve. Here $(\tau_3^{GL}, \tau_4^{GL})$ can be taken as (.23,.25075).

5.4.2 Under same L-skewness

Cong et al. (1993), Hosking and Wallis (1997) discussed a goodness of fit measure based on the assumption that candidate distribution has the same L-skewness as the sample data. Using this method we can calculate the pair (τ_3^F, τ_4^F). For example, suppose our observed (t_3, t_4) is $(0.3, 0.15)$. Then take $\tau_3^{GL} = 0.3$ and now we can calculate from the L-moment ratio diagram corresponding $\tau_4^{GL} = 0.24167$(GL). So the pair (τ_3^F, τ_4^F) is $(0.3, 0.24167)$.

5.4.3 Minimum Distance Method

This method is a most appropriate method to find (τ_3^F, τ_4^F). The method has some similarity with method of minimum chi square. Here we try to find the point (τ_3^F, τ_4^F) which minimize T^L. But it is hard to find out the point (τ_3^F, τ_4^F) which minimize T^L, so we approximate this value by minimizing the metric $d(\tau_3^F, \tau_4^F)$. The metric we can

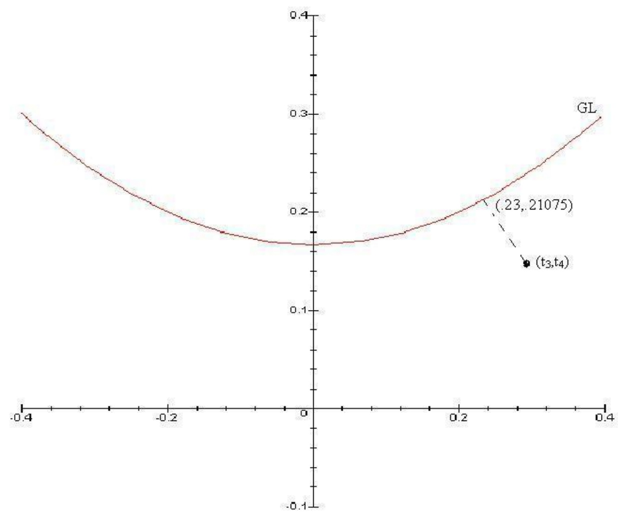


Figure 5.1

consider are the Euclidean distance or the absolute distance. That is,

$$d(\tau_3^F, \tau_4^F) = (t_3 - \tau_3)^2 + (t_4 - \tau_4)^2$$

or

$$d(\tau_3^F, \tau_4^F) = |t_3 - \tau_3| + |t_4 - \tau_4|.$$

Now the next step is to test contain to test the goodness of fit test using this (τ_3^F, τ_4^F). That is, our testing problem becomes,

$$\begin{cases} H_0: & \text{The data are sampled from a population with distribution function } F \\ & \text{with population L-skewness and population L-kurtosis } (\tau_3^F, \tau_4^F) \\ V/s & \\ H_1: & \text{The data are not sampled from a population with distribution function } F \\ & \text{with population L-skewness and population L-kurtosis } (\tau_3^F, \tau_4^F). \end{cases}$$

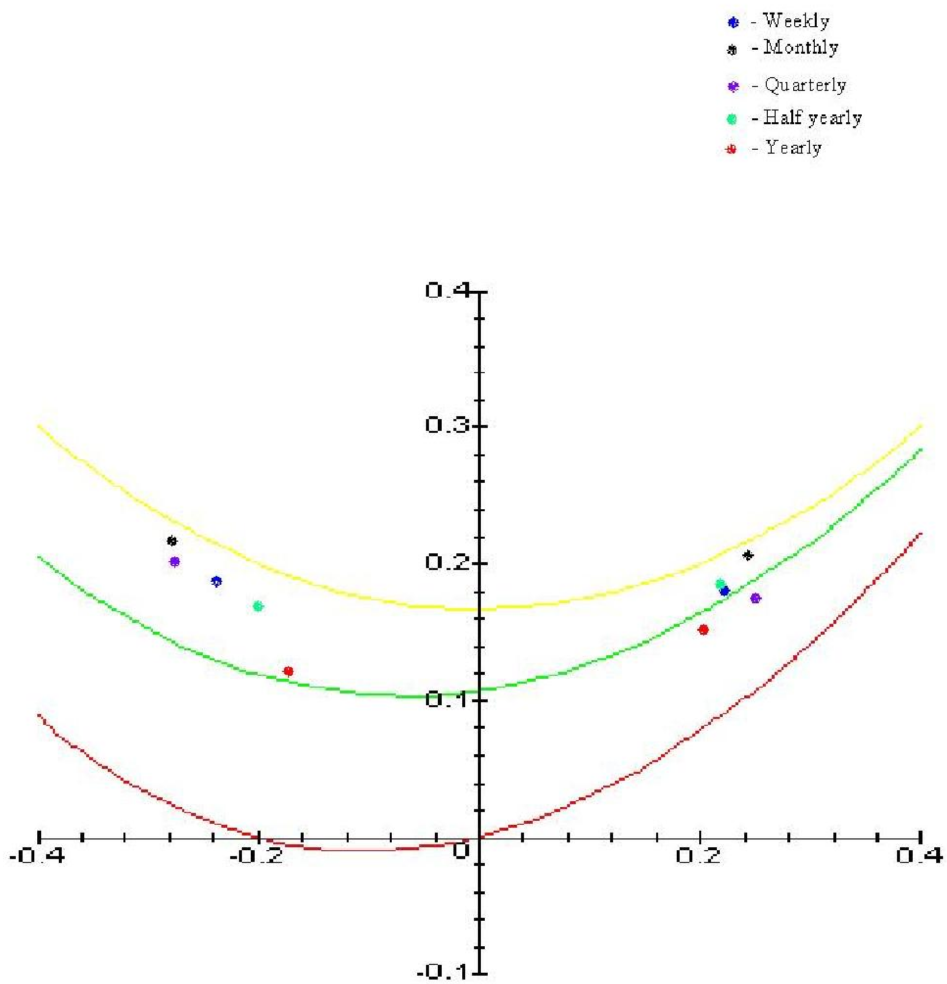
This is a testing problem discussed in Section 5.3. The statistic to be used is,

$$T^L = n(w - W)' A^{-1}(w - W)$$

for $w = (t_1, t_2)'$ be the estimators of $W = (\tau_1, \tau_2)'$ where t_1 and t_2 are the sample estimators of L-Skewness and L-Kurtosis respectively and τ_1 and τ_2 are the population L-Skewness and population L-Kurtosis respectively. Here A is the asymptotic covariance matrix of (t_3, t_4). Also statistic T^L follows asymptotic χ^2 distribution with two degrees of freedom. From the conclusion of this test, we reject or accept of our main hypothesis, that is, data from population with distribution function F.

5.5 Applications

In this section, we use the statistic T^L defined in Section 5.4 to study the performance of GL distribution in extremal behavior Bombay stock exchange data. The data used for analysis is the the same data used in Section 4.3. We consider GL distribution for goodness of fit test to conform the results in Chapter 3. That is, Gettinby et al. (2004), Tolikas and Brown (2006) and Tolikas et al.(2007) show that in majority of cases (especially minima) an adequate fit for extreme daily returns is GL for data from different nations Stock Exchanges. Also these results shows that GL distribution is an adequate model for extreme behavior of any time interval, which contradict the well known extremal types theorem. Their studies use Anderson-Darling test for goodness of fit even for small sample sizes. Here we reconsider the goodness of fit test using generalized logistic distribution for extreme behavior of Bombay stock exchange data using T^L.



- Weekly
- Monthly
- Quarterly
- Half yearly
- Yearly
0.4
0.3
0.2
0.1
-0.1
-0.4
-0.2
0
0.2
0.4

Time Interval	GL(t_3, t_4)
Weekly maxima	0.225 0.179
Monthly maxima	0.245, 0.208
Quarterly maxima	0.258, 0.175
Half yearly maxima	0.222, 0.185
Yearly maxima	0.203, 0.153
Weekly minima	-0.243 0.194
Monthly minima	-0.284, 0.219
Quarterly minima	-0.281, 0.204
Half yearly minima	-0.202 0.170
Yearly minima	-0.171, 0.120

Table 5.4: Estimated value of (t_3, t_4) on different time intervals

In Figure 5.2, data points in the first quadrant represent (t_3, t_4) of maximum of different time interval and in the second quadrant represent (t_3, t_4) of minimum of different time interval. Which is also given in Table 5.4. Time intervals include weekly, monthly, quarterly, half yearly, yearly of the data. Now we have to find (τ_3^F, τ_4^F) for each time interval. We use minimum distance method (by Euclidean metric) to find (τ_3^F, τ_4^F) since the other two methods, graphical method and same L-skewness method, are very simple to calculate (τ_3^F, τ_4^F). The Table 5.5 gives (τ_3^F, τ_4^F) for maxima and minima for different time intervals of GL. We have calculated (τ_3^F, τ_4^F) by using Fortran code given in the Appendix. The asymptotic variance obtained is given in Table 5.6. Also the values of T^L for different time intervals are given in Table 5.7.

From Table 5.7, we can conclude that for most of the time intervals GL is an adequate fit. For yearly time intervals, T^L shows a high value, but from the simulated study for small sample size like n=25, in case of three parameter distribution, we have

Time Interval	GL(τ_3^F, τ_4^F)
Weekly maxima	0.215, 0.205
Monthly maxima	0.242, 0.215
Quarterly maxima	0.242, 0.215
Half yearly maxima	0.215, 0.205
Yearly maxima	0.189, 0.196
Weekly minima	-0.235, 0.213
Monthly minima	-0.279, 0.231
Quarterly minima	-0.270, 0.227
Half yearly minima	-0.193, 0.198
Yearly minima	-0.154, 0.186

Table 5.5: Estimated value of (τ_3^F, τ_4^F) for different distribution on different time intervals

found that the critical value has a deviation of 2 points. So it can be concluded that GL distribution is also an adequate fit in yearly time intervals. Which agrees the result in Chapter 4.

5.6 Conclusion

In this chapter we introduced a goodness of fit test statistic based on L-moments for distributions whose mean exists. We have shown that the statistic has a asymptotic Chi-square distribution and this convergence is fast, that is, we can approximate the statistic to Chi-square distribution even for sample size is 20. For a particular case, when the test distribution is Normal, the statistic has computational higher power and moderate performance to outliers compared to the three well-known used tests

Time Interval		t_3	t_4
Weekly maxima	t_3	0.488034363	0.329522875
	t_4		0.210597926
Monthly maxima	t_3	0.556183701	0.412080163
	t_4		0.254962338
Quarterly maxima	t_3	0.55618586	0.4120892
	t_4		0.255006603
Half yearly maxima	t_3	0.488035411	0.329513358
	t_4		0.333706773
Yearly maxima	t_3	0.260255	0.188271
	t_4		0.1455574
Weekly minima	t_3	0.5287056	-0.4394237
	t_4		0.3804288
Monthly minima	t_3	0.342497095	-0.279963707
	t_4		0.343572398
Quarterly minima	t_3	0.647330388	-0.515478347
	t_4		0.548024138
Half yearly minima	t_3	0.44625932	-0.2718477
	t_4		0.285473296
Yearly minima	t_3	0.1843003	-0.11621096
	t_4		0.08287285

Table 5.6: The asymptotic variance matrix of sample L-skewness and sample L-kurtosis of maxima and minima for different time interval.

Time Interval	T^L
Weekly maxima	105.80
Monthly maxima	0.54
Quarterly maxima	4.92
Half yearly maxima	0.27
Yearly maxima	7.03
Weekly minima	6.98
Monthly minima	0.71
Quarterly minima	0.68
Half yearly minima	0.44
Yearly minima	7.59

Table 5.7: Estimated values of T^L of GL on different time intervals

viz. Jerque- Bera statistic, the statistic based on Robust estimator and Anderson-Darling statistic. We also found that the GL distribution is a good model for extremes of Bombay Stock exchange data.

Chapter 6

Conclusion and Ideas for Future Work

In this research work we found that the generalized logistic distribution can also be used as a model for extremes in addition to the existing models of extremes described in Chapter 1. We have shown in Chapter 3 that the maximum M_N of a random number N of random variables, where N is a geometric random variable, follows asymptotically the generalized logistic distribution. We saw that the generalized logistic distribution has a stability property in terms of random sample size N. Like the generalized extreme value model described in Chapter 1, the generalized logistic distribution also includes three well known families of distributions namely, the logistic, the loglogistic and the backward loglogistic distributions. We also verified that the generalized logistic model is a good fit for the Bombay stock exchange data. We introduced a goodness of fit measure in Chapter 5 and empirically verified that this has more power compared to other existing goodness of fit measures.

In Chapter 3, to prove that the generalized logistic distribution is the asymptotic distribution of the random maximum M_N, we used the main theorem of Chapter 2. The main theorem of Chapter 2 is proved for those function 'g' which satisfies a property called Q property. There may be bigger class of functions where the main theorem in Chapter 2 holds. One of the immediate interest is to identify such class of functions. The main theorem in Chapter 2 is also proved under the assumption that the general function g has an asymptotic distribution. The condition under which such an assumption holds can be investigated. This is another problem of interest. Another aspect regarding the convergence results in Chapters 2 and 3 to be looked into is the rate of convergence. All the results in Chapters 2 and 3 are under the assumption that the sequences of random variables under consideration are independent and identically distributed. One can try to relax this assumption and look for convergence. This may lead to investigating whether the results in Chapters 2 and 3 hold for independent non identically distributed random variables and dependent sequences of random variables like stationary sequences.

Appendix: Fortran code to calculate T_{GL}

```
implicit double precision (a-h,o-z)
DOUBLE PRECISION ss,kk,glq,s,sk,ku,n,gl,gp,gev
open(unit=1,file='10.DAT',status='old')
open(unit=2,file='11.txt',status='old')
read(1,*)sk,ku
n=1000.0
do 82 s=-1.0,1.0,0.001
gl=0.0
glq=0.0
glq=glq+(5.0*((s)**2))
glq=(glq+1)/6.0
gl=gl+((sk-s)**2)+((ku-glq)**2)
if(gl.gt.n) go to 23
n=gl
ss=0.0
kk=0.0
ss=ss+s
kk=kk+glq
23 d=b
82 continue
write(2,*)ss,kk
close(1)
stop
end
```

Bibliography

[1] Balakrishnan, N. (1992). *Handbook of the Logistic Distribution.* Marcel Dekker, New York.

[2] Barndorff-Nielsen, O. (1964). On the limit distribution of the maximum of a random number of independent random variables. *Acta Math. Acad. Sci. Hungar.* **15**, 399-403.

[3] Bera, A., and Jarque, C. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters.* **6**(3), 255-259.

[4] Bera, A., and Jarque, C. (1981). Efficient tests for normality, homoscedasticity and serial independence of regression residuals: monte Carlo evidence. *Economics Letters.* **7**(4), 313-318.

[5] Billingsly, P. (1968). *Convergence of Probability Measures.* John Wiley and Sons, New-York.

[6] Brys, G., Hubert, M., and Struyf, A. (2004). A robustification of the Jarque-Bera test of normality, in COMPSTAT 2004, Proceedings in Computational Statistics, edited by J. Antoch, Springer, Physica Verlag, 753-760.

[7] Cotter, J. (2001). Margin exceedences for european stock index futures using extreme value theory. *Journal of Banking and Finance.* **25**, 1475-1502.

[8] Danielson J., Hartmann, P., and de Vries, C. (1998). The cost of conservatism. *Risk*. **11**(1), 103107.

[9] de Haan, L., and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Verlag, New York.

[10] Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer-Verlang Berling Heidelberg, Germany.

[11] Fama, E. (1965). The behaviour of stock market price. *Journal of Business*. **38**, 34-105.

[12] Feller, W. (2000). *An Introduction to Probability Theory and Its Applications*. **2**, 2nd ed. John Wiley and Sons.

[13] Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*. Wiley, Newyork.

[14] Gale, M. H., and Gastwirth, J. L., (1978). A scale-free goodness-of-fit test for the exponential distribution based on the Gini statistic. *Journal of the Royal Statistical Socity-B*. **40**, 350-357.

[15] Gettinby,G.D., Sinclair, C. D., Power, D. M., and Brown, R. A. (2004). An analysis of the distribution of extreme share returns in the UK from 1975 to 2000. *Journal of Business Finance and Accounting*. **31**, 607-645.

[16] Gnedenko, B. V. (1982). On limit theorems for a random number of random variables, *Probability Theory and Mathematical Statistics. Proc.* 4^{th} *USSR-Japan Symp., Tbilisis, USSR*. ed. K. Ito and J. V. Prokhorov, Lecture Notes in Mathematics, Springer-Verlang, Berlin, 167-176.

[17] Gray, J. B., and French, D. W. (1990). Emperical omparisons of distributional models for stock index returns, *Journal of Business, Finance and Accounting.* **17**(3), 451-459.

[18] Gumbel, E. J. (1958). *Statistics of Extremes.* Columbia University Press, New York.

[19] Harris, R. D. F., and Kucukozmen, C. C. (2001). The emperical distribution of UK and US stock returns. *Journal of Business Finance and Accounting.* **28**(5) and (6), 715-740.

[20] Hoaglin, D.C., Mosteller, F., and Tukey, J.W. (1985). *Exploring data tables, trends and shapes.* John Wiley and Sons, New York.

[21] Hoeffding, W., (1948). A class of statistic with asymptotically normal distribution. *Ann. Math. Statist.* **19**, 293-325.

[22] Hosking, J. R. M., (1986). *The theory of probability weighted moments.* Research Report RC12210. IBM Research, Yorktown Heights.

[23] Hosking, J. R. M. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Socity.* **3**, 105-124.

[24] Hosking, J. R. M., (1999). The L-Moments page. (http://www.research.ibm.com/people/h/hosking/lmoments.html).

[25] Hosking, J. R. M. and Wallis, J. R. (1997) *Regional Frequency Analysis: An Approach Based on L-Moments.* Cambridge University Press, Cambridge, UK.

[26] Kalashnikov, V. (1997) *Geometric sums: bounds for rare events with applications: risk analysis, reliability, queueing.* Kluwer Academic Publishers, USA.

[27] Kearns, P., and Pagan, A. (1997). Estimating the density tail index for financial time series. *Review of Economic Statistics.***79**, 171-175.

[28] Kotz, S., and Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*. Imperial College Press. London.

[29] Kozubowski, T.J., and Rachev, S.T. (1999). Multivariate geometric stable laws. *J. Comput. Anal. Appl.*. **1**(4), 349-385.

[30] Laha, R. G., and Rohatgi, V. K. (1979). *Probability theory*. Wiley, New York.

[31] Leadbetter, M. R., Lindgren, G., and Rootzen, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York.

[32] Longin, F. M. (1996). The asymptotic distribution of extreme stock market returns. *Journal of Business*. **69**, 383-408.

[33] Longin, F. M. (2000). From value at risk to stress testing: the extreme value approch. *Journal of Banking and Finance*. **24**, 1097-1130.

[34] Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*. **36**, 394-419.

[35] McDonald, J., and Xu, Y. (1995). A generalization of Beta distribution with applications. *Journal of Econometrics*. **66**, 133-152.

[36] Nidhin, K. and Chandran, C.(2009a). Limit theorems of general functions of independent and identically distributed random variables. *Statistics and Probability Letters*. **79**, 2397-2404

[37] Nidhin, K. and Chandran, C. (2009b). An analysis of the extremal behavior of Bombay stock exchange data. *Vikalpa: The Journal of Decision Makers*. Communicated.

[38] Nidhin, K. and Chandran, C. (2010a). Some notable characteristic of generalized logistic distribution in extreme value theory. *Statistics and Probability Letters*. Communicated.

[39] Nidhin, K. and Chandran, C. (2010b). A goodness of fit measure using L-moment ratio diagram. Working paper.

[40] Pearson, E. S. (1963). Some problems arising in approximating to probability distributions, using moments, *Biomatrika*. **50**, 95-111.

[41] Peiro, A. (1994). The distribution of stock returns: internatinal evidence, *Applied Financial Economics*. **4**, 431-439.

[42] Pickands, J. (1975) Statistical inference using extreme order statistics. *The Annals of Statistics*. **3**(1), 119-131.

[43] Robinson, M.E. and Tawn, J. A. (1995). Statistics for exceptional athletics records.*Applied Statistics*. **44**, 499-511.

[44] Sankarasubramanian, A. and Srinivasan, K. (1999). Investigation and comparison of sampling properties of L-moments and conventional moments. *Journal of Hydrology*. **218**, 13-34.

[45] Satheesh, S. and Nair, N. U. (2002). A note on maximum and minimum stability of certain distributions. *Calcutta Statist. Assoc. Bulletin*. **53**, 249-252.

[46] Shapiro, S. S. and Wilk, M. B. (1965). "An analysis of variance test for normality (complete samples)", *Biometrika*. **52**(3/4), 591611.

[47] Smith, R. L. (1986). Extreme value theory based on the r largest annual events.*J. Hydrology*. **86**, 27-43.

[48] Smith, R. L. (1997). Statistics for exceptional athletics records: Letter to the editor *Applied Statistics*. **46**, 123-127.

[49] Smith, R. L. (2001) *Statistics of Extremes: With Applications in Environment, Insurance and Finance.* Lecture Notes SEMSTAT Conference on Extreme Value Theory, Gothenberg, Sweden.

[50] Sreehari, M. (1995). Max stability and a generalization. *Statist. Prob. Letters.* **23**, 339-342.

[51] Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. *Handbook of Hydrology.* (ed. by D. R. Maidment), **18**, 1-66.

[52] Stephens, M. A. (1974). ' EDF statistics for goodness of fit and some comparisons', *Journal of American Statistical Association.* **69**, 730-737.

[53] Tawn, J. A. (1988). An extreme value model for dependent observations.*J. Hydrology.* **101**, 227-250.

[54] Theodossiou (1998). 'Financial data and the skewed generalised-t distribution', *Management Science.* **44**, 1650-61.

[55] Tolikas, K., and Brown, R. A. (2006). The distribution of extreme daily share returns in the Athens stock exchange. *The European Hournal of Finance.* **12**(1), 1-12.

[56] Tolikas, K., Koulakiotis, A., and Brown, R. A. (2007). Extreme risk and value-at-risk in the German stock market. *The European Journal of Finance.***13**(4), 373-95.

[57] Vogel, R. M., and Fennessey, N. M. (1993). L-moment diagrams should replace product moment diagrams. *Water Resources Research.* **29**, 1745-1752.

[58] Voorn, W. J. (1987). Characterization of the logistic and loglogistic distributions by extreme value related stability with random sample size. *J. Appl. Prob.***24**, 838-851.

[59] Wang, Q. J. (1996). Direct sample estimators of L-moments. *Water Resources Research*. **12**, 3617-3619.