Quantitative risk management and data analytics with applications to finance

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Kapil Chandak

IISER PUNE

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

April, 2023

Supervisor: Prof. Arnab Kumar Laha

Certificate

This is to certify that this dissertation entitled Quantative risk management and data analytics with application to finance towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Kapil Chandak at Indian Institute of Management Ahmedabad under the supervision of Prof. Arnab Kumar Laha, Associate Professor, Departnment of Production and Quantative methods ,Indian Institute of Management, Ahmedabad during the academic year 2022-2023.

Arnab Kumar Laha

Prof. Arnab Kumar Laha

Committee:

Prof. Arnab Kumar Laha

Prof. Anindya Goswami

This thesis is dedicated to my family

Declaration

I hereby declare that the matter embodied in the report entitled Quantitative risk management and data analytics with applications to finance are the results of the work carried out by me at the Departnment of Production and Quantative methods, Indian Institute of Management, Ahmedabad, under the supervision of Prof. Arnab Kumar Laha and the same has not been submitted elsewhere for any other degree.

Kapil Chandak

Acknowledgments

I would like to express my sincere and deepest gratitude towards Prof. Arnab Kumar Laha for supervising the project. This thesis would not have been possible without his continuous support and guidance throughout the project. He has invested much of his precious time in this project and was always available whenever I faced any difficulty or wanted to clear any doubts. I would also like to thank Prof. Anindya Goswami for his valuable comments and guidance in the project. And I would also like to thank IIM Ahmedabad for providing the facilities to carry out this project without difficulty.

I want to thank Vikas Shukla as he supported me immensely throughout the journey; without his support, I would not have achieved my goals. I wish to thank Lubdhak Mondal, a great colleague during this journey. I would also thank ChatGPT, as its simplified response has helped me numerous times during the project not to waste my time on unnecessary stuff. I would also thank Mr. and Miss. Mundada for helping me with various difficulties during my stay in Ahmedabad. I would also like to thank other Ph.D. students, Huzaifa and Bharat, for helping me. I would also like to thank other Research and Academic Associates, including Prince, Satender, Vaishnav, Madhubanti, Vishnu, Muskan, Srija, Divesh, Jalshayin, Pulkita, and others, for supporting me. I would also thank Miteshbhai and Aditya for their help and support. I would also like to thank Infosys Foundation for giving me financial support in the form of an Infosys Foundation scholarship.

Abstract

Risk Management has important implications for many organizations, and quantifying those risks is essential. The financial impact of the extreme events which lead to these risks is huge, a part of which is discussed in this thesis. We have various risk measures and properties associated with them. We studied risk measures in the context of market risk associated with equities and stylized facts of financial time series. We focused on the mathematical aspects of some of these stylized facts, which have implications for financial risk. These mathematical discussions have huge implications when working with situations where we have such nasty data that arise in many cases and thus implications of data analysis of these kinds of problems. We also examined the impacts of these new statistics on quantifying risk and capturing other aspects of financial time series. We also examined the dependence of these risk events and does it make sense to predict these extreme events based on past data and quantify their effects. We also discussed the economic and data analytic implications of the work. We also had a deeper look at how to make sense of the use of machine learning, some limitations of it, and how it affects our analysis and looked at the behavioral finance area from a new perspective. These insights may help open new horizons for further research in Machine learning, quantitative finance, behavioral finance, and other regions.

Contents

Preliminaries

Written below are some of the standard symbols and their meanings used in the entire thesis

- If X is a random variable then $F(X)$ denotes the Cumulative distribution function(CDF) of the random variable X.
- If X is a random variable then $f(X)$ denotes the probability distribution/density function(PDF) of the random variable X.
- t denotes time.
- If X is a random variable then $E(X)$ denotes the Expectation of the random variable X.
- If X and Y is a random variable then $E(X|Y)$ denotes the Conditional Expectation of the random variable X given Y.
- Δt or Δ denotes the time scale.
- $S(t)$ or $s_t$denotes the price of stock at time t.
- $r(t, \Delta t)$ denotes the value of log returns at time t and time scale Δt.
- $\rightarrow$ denotes convergence
- $\xrightarrow{P}$ denotes converges in probability

Introduction

Risk is the potential for an activity, decision, or unforeseen changes in the situation that may adversely affect the objectives of an individual or an organization. Every individual, firm, organization, and country is exposed to risk. For instance, the operation of institutes such as IISER Pune is fraught with risks, such as experiments going awry and igniting a fire or water supply problems affecting the health of campus residents, etc. Risks may arise in various contexts, including but not limited to financial investments, health, safety, security, and more. It is impossible to remove all risks from every possible scenario altogether. Still, risks can be effectively managed with the aid of safety guidelines, which reduce the frequency and severity of the impacts of extreme events. Risk management involves identifying, analyzing, and developing strategies to reduce the impact of unforeseen circumstances that could negatively impact an individual's or an organization's objectives. It involves identifying possible risks, determining the frequency and severity of the risks, and developing strategies to reduce or avoid them. Effective risk management enables individuals and organizations to make well-informed decisions by understanding the potential risks and benefits of different options and achieving their objectives while minimizing potential losses and adverse impacts. It can be applied in various areas, including finance, health, cybersecurity, etc. Risks may originate from unforeseeable sources that nobody knows about, thus making it impossible to eliminate them. Numerous of these risks frequently entail enormous costs. As evidenced by the preceding discussion, risk management is a crucial aspect of decision-making for individuals, small/large organizations, and even nations. Assessing and managing risk is an important activity that ensures the long-term health of organizations in the present day. Quantitative risk management techniques supplement and enhance expert judgment in order to manage risk appropriately within reasonable constraints.

Many risks organizations and nations encounter may have substantial financial implications, directly and indirectly impacting financial institutions like banks and insurance companies.

The Nobel Prize in Economics 2022 was awarded to Ben S. Bernanke, Douglas W. Diamond, and Philip H. Dybvig for research on banks and financial crises. Their work highlighted the importance of financial institutions like banks. It emphasizes the social function of financial institutions like banks and why avoiding bank crises is crucial. Thus we focus our attention on financial risk management.

Financial risk management is essential for the smooth operation of day-to-day activities for accomplishing the organization's goals. Organizations like banks, insurance, etc., deal with financial risk management issues to manage clients' financial risk. For insurance, banking, and other financial industries, quantifying risk arising from multiple sources is vital. As insurance companies assume risk-related financial obligations, if the premium charged is not commensurate with the amount of risk they have assumed, they may encounter numerous problems. If the premium is overcharged, they may lose market competitiveness and customers, and if undercharged, they may run out of money and go bankrupt. If banking and mutual fund companies do not effectively quantify and manage risk, they will face several issues, particularly during financial crises.

Thus, appropriate quantification of risks from various sources is essential, especially for financial organizations. Appropriate quantification of risk is necessary to know how much money these financial institutions must hold as a buffer to safeguard against future unexpected losses. Therefore we focus on quantitative risk management techniques in this thesis. Risk modeling, stress testing, scenario analysis, etc., are some tools used in quantitative risk management. They all involve mathematical modeling and data analytical tools to help organizations access and quantify the risks they face. Technological advancements, changes in geopolitical situations, etc., create new risks and may also change ways to deal with existing risks. For example, the advancement of quantum computation poses a threat to cryptographic technologies, which still rely on the NP-hardness of factoring prime numbers or discrete log problems. Thus, when a new technology emerges, it brings unknown risks and may change the entire risk scenario. Studying data on past risk events and appropriately analyzing and modeling them helps us model risks properly. It helps us deal with existing risks and gives us indications while dealing with new risks. Thus, we focused on analyzing past risk events in this thesis. We focused on market risk problems, but the discussed techniques can also be applied to other forms of risk, like credit and liquidity risks.

Managing financial risk has a long history and has undergone significant development. Thus, before we proceed with financial risk management, let us briefly overlook the history of financial risk management. There are numerous instances in ancient and medieval India, Greece,

Rome, and other places where merchants frequently pooled their resources and shared the dangers of perilous, protracted travels. Ancient Indian text Arthashastra [20] also mentions systems similar to insurance and reserves used in modern times and other systems like Hundi in place to manage financial risk. The mathematical formulation of probability theory on which the foundations of modern risk management were laid in the early 18th century, and subsequent advancements made in probability theory by Bayes [5], Kolmogorov [21], and others made it possible to be helpful in the practical sense of modern risk management. Following works by Louis Bachelier [3], John Maynard Keynes, David Dodd, and Benjamin Graham laid further foundations. Subsequent works by Harry Markowitz [25] and William Sharpe [37] created the current portfolio theory to optimize investment portfolios based on risk and return. The Capital Asset Pricing Model (CAPM) model [38] has significantly influenced the development of contemporary portfolio theory and how investors and financial analysts see risk and return. Edward O. Thorp's 1961 article on the Kelly Criteria [40] has been used to design techniques for maximizing returns while lowering risk, which has significant implications for portfolio management.
Futures and options were also significant advancements in managing risk. The 1973 paper published by Fischer Black and Myron Scholes [6] gave investors a way to appropriately value options, leading to a development in the options market. John C. Hull's 1987 book "Options, Futures, and Other Derivatives" [18] has significantly influenced current financial market growth and how investors and institutions manage risk in sophisticated financial products. Subsequent developments on Value at Risk (VaR), black swan, and subsequent analysis of crises on the 2000 .com bubble and 2008 housing bubble by many researchers have led to a significant development in this field. After looking at the social consequences of financial crises and realizing their importance, the interest of researchers, practitioners, and even the general public has increased in financial risk management. These developments led to changes in how banks and insurance companies function.

The structure of the thesis is as follows. Chapter 1 discusses financial risk measurements, ways to calculate them empirically, and stylized facts about financial time series. Chapter 2 discusses two important stylized facts in detail and the mathematical issues associated with them. It discusses the financial and mathematical aspects related to them. Chapter 3 discusses the data on which we applied our data, a new correlation coefficient known as the Chatterjee correlation coefficient. Chapter 4 discusses the implications of our work and the scope for future research. The Graphs and codes are written in the Appendix.

Chapter 1

Market Risk Measurement and Stylized Facts

As discussed in [11][13][19], Financial risk is the potential for financial losses arising from changes in various risk factors. Financial risk can be classified into market, credit, liquidity, legal, operational, and model risks. Market risk is the risk from fluctuations in market prices and can affect the value of investments and portfolios. Credit risk is the risk of the borrower defaulting on a loan or debt, while liquidity risk arises from the inability to sell assets quickly enough to meet financial obligations. Operational risk is the risk of losses due to an inadequate or failed internal process. Legal risk is when an organization faces losses arising from legal constraints like lawsuits. Model risk is the risk due to using a model at an instance where the assumptions underlying the model are not satisfied. As discussed in the Introduction, risk quantification is essential, so this chapter discusses risk measurement. Here we are more focused on market risk measurement. However, some of the methods discussed, particularly the advanced one as discussed in [11][12][13], applies even to the other types of risk like credit and operational risk.
We will also look at some properties of financial time series known as stylized facts discussed in [9].

1.1 Risk measurement

The Markowitz theory laid some of the early foundations of modern financial risk management. The portfolio theory uses a risk-return analysis and optimizing the portfolio to get the best possible return for a given level of risk. In the portfolio optimization theory, the standard deviation of returns for a given time interval at appropriate steps was considered an appropriate measure of risk. All the subsequent works, like CAPM and others, also carried forward with that idea. The Redington-Fisher criterion for the optimality of bonds uses Net present value, Duration, and Convexity so that our bond portfolio is immunized to small changes in interest rates. Although many of these approaches consider risk, they are more suitable asset pricing models and not for extreme events. Although one may also incorporate skewness and kurtosis, which gives a better picture of the probability of extreme events, they are still insufficient. Thus the general asset management theories are not directly applicable when working on risk management because, in risk management, we are primarily concerned with extreme losses. Therefore, we place greater emphasis on the loss tail region than on the overall dynamics of the financial asset. Thus, although helpful, measures like standard deviation are insufficient and may even be misleading when working with extreme events. One needs to focus on the tail region for risk management. Thus before we move forward with the tail risk, let us briefly review some important stylized facts of financial time series.

1.2 Stylized facts of financial time series

Looking at financial time series statistically, the seemingly random variations of asset prices share some nontrivial statistical properties. Such properties that are common across various instruments, markets, and time periods are called empirical stylized facts. The reference used in this section is [9]. These facts have been observed in studies of different markets and instruments; thus, they are general properties. Still, these facts are so constraining that it is challenging to construct a stochastic process that possesses the same set of properties.
So let us first set up notation. Let $S(t)$ denote the price of a financial asset at time t. And Δt denotes a time scale. We define returns $r(t, \Delta t)$ as $r(t, \Delta t) = ln(\frac{S(t+\Delta t)}{S(t)})$. While dealing with Δt, one must also note that the statistical properties of time series strongly depend on Δt. So let us discuss some of the important Stylized facts.

1. **Absence of autocorrelations:** Linear autocorrelations of $r(t, \Delta t)$ are often insignificant, except for very small intraday Δ are in order of minutes for which market micro-structure effects come into play.

2. **Heavy tails:** The unconditional distribution of $r(t, \Delta t)$ seems to display a power-law or Pareto-like tail with a finite tail index, generally between 2 to 5. This excludes stable laws with infinite variance and the normal distribution for most cases. However, the method used to say about Pareto is based on the goodness of fitness test(discussed later), which generally rejects a hypothesis when there is sufficient evidence to reject the hypothesis. Thus although we cannot reject the hypothesis, that does not serve as proof of the exact distribution, which is difficult to determine.

3. **Aggregational Gaussianity:** By increasing Δt the distribution of $r(t, \Delta t)$ looks more and more like Gaussian distribution. Thus, the shape of the distribution changes with Δt.

4. **Volatility clustering:** Various volatility measures display a positive autocorrelation over several days, quantifying that high-volatility events tend to cluster in time.

5. **Conditional heavy tails:** Even after correcting $r(t, \Delta t)$ for volatility clustering, the residual time series exhibits heavy tails. Nevertheless, the tails are less heavy than in the unconditional distribution of $r(t, \Delta t)$.

6. **Slow decay of autocorrelation in absolute returns:** The autocorrelation function of $|r(t, \Delta t)|$ and $r(t, \Delta t)^2$ decays slowly as a function of Δt, roughly as a power law. This slow decay may be interpreted as a sign of long-range dependence.

7. **Volume/volatility correlation:** Trading volume is correlated with all Volatility measures.

Thus, the above discussed are some of the important stylized facts. The stylized facts of heavy tails, volatility clustering, and absence of autocorrelations are important and have considerable implications in financial risk management.

1.3 Risk measures

Thus, we can now appreciate the difference between asset management and risk management. [19] discusses some of the conventional methods for risk measurement. It discusses valuation problems, sensitivity analysis, and scenario analysis/stress testing. One of the significant issues in sensitivity analysis is that bonds, equities, options, and other financial instruments have different sensitivity measures. Thus comparisons across financial instruments and modeling for risk are complicated. While dealing with risk management, one needs to deal with all the financial instruments together and map for dependence structures, which becomes difficult to deal with if we use different risk measures. Thus, a standard risk measure that can measure extreme events for all kinds of financial instruments is essential. Another issue is that sensitivity measures are not additive and thus cannot be used to aggregate risk and do not directly translate easily to losses in exact monetary form. This issue was resolved using Value at Risk (VaR) and associated risk measures. VaR is a way to compare risk on bonds, derivatives, equities, etc., on one platform instead of using duration for bonds, the beta for stocks, the delta for derivatives, and other measures for other financial instruments. VaR and related measures are helpful in all market, credit, and operational risks, which is not the case with other risk measures. In later sections, we will discuss some risk measures related to VaR, its limitations of VaR and those risk measures. We will also discuss some valuable properties of risk estimators, like coherency.

1.3.1 Value at Risk (VaR) and Expected Tail Loss (ETL)

As discussed in [12] in market risk management, one is interested in the tail region of losses. So before moving ahead, let us define some variables and measures.

Definition 1.3.1. *The Profit and Loss(P&L) Random Variable is defined as* $L_{t+1} = -(S_{t+1} - s_t)$*. Let us define* L *as the distribution of* L_{t+1} *where* $0 < t < \infty$*.*

Definition 1.3.2. *Let* $0 < \alpha < 1$*. The Value-at-Risk of* L *at confidence level* α *is defined as*

$$VaR_\alpha(L) = inf\{x \in \mathbb{R} : F_L(x) \geq \alpha\}$$

Definition 1.3.3. *The Expected Tail Loss(ETL) of L at confidence level α is defined as*

$$ETL_\alpha(L) = E[L|L > VaR_\alpha(L)]$$

ETL is also known by alternative names like ES, CVaR, etc. These measures are frequently used risk measures in modern risk management.

1.3.2 Coherent Risk Measures

Before discussing the advantages and disadvantages of VaR and ETL, we would like first to understand what coherent risk measures are. This is discussed in[11].
Let $\mu(.)$ be a risk measure, and X and Y be the future values of two risky positions.

Definition 1.3.4. *A risk measure $\mu(.)$ is said to be coherent if it satisfies the following properties:*

1. $\mu(X) + \mu(Y) \geq \mu(X + Y)$ *(sub-additivity)*
2. *For a positive number t if $\mu(tX) = t\mu(X)$ (homogeneity)*
3. $\mu(X) \geq \mu(Y)$ *if $X \leq Y$ (monotonicity)*
4. $\mu(X + n) = \mu(X) - n$ *(risk-free condition)*

$X \leq Y$ means Returns from portfolio Y are greater than that in X almost surely.

The sub-additivity property is one of the vital properties. The sub-additivity condition can be interpreted as the total risk of the entire portfolio is less than the sum of the risk of individual components. This property has vast applications because if one uses this as the risk measure to determine the risk of the entire portfolio, then they can always calculate it for small individual components of the portfolio and add it up to be sure that the total risk of the whole portfolio is surely less than that. This property is essential for a stock exchange or a broking firm if it desires to set the security amount needed for a particular individual. So if this sub-additivity property is not satisfied, then the sum of the security amount from individuals will not offset the risk due to fluctuations in share prices the exchange or the

broking firm faces, which is undesirable.
If a risk measure is sub-additive, then a large firm may use the measure to value risks at, let us say, department level and the addition of these risks will undoubtedly be less than the total risk the firm faces, thus significant implications in terms of practice. Therefore as discussed, sub-additivity is an essential property that has enormous implications.

1.3.3 Advantages and limitations of risk measures

As discussed above, the risk measures VaR and ETL have many advantages and limitations.[11][12] discusses those issues in detail. In this section, we will highlight some of those important points. VaR and ETL get around many of the issues we face in traditional risk measures like now we can compare the risk of a portfolio with bonds, stocks, derivatives, and other financial instruments on a common platform which directly translates into monetary loss, thus making it easier to interpret. We can use it to measure all three kinds of market, credit, and operational risks. Although this is a big advantage, these measures still have some advantages and limitations, and the following are some of them:

- ETL is an expectation and may not even be defined depending upon the parametric model assumed for the P&L distribution function, which VaR does not face. Also, one must note that for computing ETL, one needs to compute VaR, but it is not the other way around.
- From a statistical estimation point of view, the confidence interval of the estimation of VaR is much narrower than ETL. Thus, we may get much more reliable VaR estimates than ETL.
- VaR seems to be a more robust risk measure as per some intuition of robustness, although the conception of robustness is vague, and one may design conceptions where ETL is robust. Thus this is a grey area.
- VaR does not give an idea of the worst of the worst. Any huge value below the significant value in the tail region does not change the value of VaR, which is not the case with ETL.
- VaR is not sub-additive and thus not a coherent risk measure which ETL is. This non-subadditivity is one of the major issues with VaR.

- Blindly using VaR may be misleading and [12] gives examples of that. One of the examples demonstrates how using VaR maybe lead us to select a less diversified portfolio due to non-subadditivity highlighting the limitations of using VaR, which is not the case with ETL.

Thus although VaR and ETL are very important and useful risk measures, we must keep the limitations of VaR and ETL in mind while using them. Furthermore, this also presents the further scope of working on better risk measures.

1.4 Calculating VaR and ETL

Until now, we have seen the definitions of VaR and ETL and their advantages and limitations. Now let us proceed on how we calculate them. [11][19] are the main references used in this section.There are three major ways used to calculate VaR and ETL. They are:

- Historical methods
- Parametric method
- Monte-Carlo method

Proceeding ahead, we will discuss each method briefly.

1.4.1 Historical method

The historical method for calculating VaR at α significance for the L distribution involves finding the $(1-\alpha)^{th}$ quantile value of L. There is no assumption on the distribution of L, and L is just the distribution of historical data. ETL is computed as the average of all the values above the VaR value in the L historical data.

1.4.2 Parametric method

The parametric method for calculating VaR at α significance for the L distribution involves assuming some parametric distribution for our L distribution, like normal, t, inverse Gaussian, etc., and estimating the parameters from the historical data. Then $(1-\alpha)^{th}$ quantile value of the fitted L distribution function is computed. ETL is computed as the conditional expectation using the formula $ETL_\alpha(L) = E[L|L > VaR_\alpha(L)]$. Moreover, given that we have assumed and calibrated the distribution of L, we can compute this conditional expectation, thus calculating the ETL value.

1.4.3 Monte Carlo method

A statistical technique called Monte Carlo simulation involves creating random samples of inputs for a particular model and then modeling the model's output for each input value set. The Monte Carlo method for calculating VaR and ETL at α significance for the L distribution involves simulating many values of L distribution and following the same steps as done in historical on these simulated data instead of historical data.

Thus, we had a brief overview of financial risk measurement techniques and stylized facts like heavy tails and volatility clustering in finance. One interesting thing to note is that volatility clustering and the absence of autocorrelations are slightly incompatible. Volatility clustering implies that high volatility events tend to cluster, thereby indicating some dependence property. One should also note that the absence of autocorrelations means no linear dependency among returns but does not imply any non-linear dependencies. Also, some long-range dependencies might have some implications for risk, and we will discuss these in more detail.

Chapter 2

Heavy tails and Volatility Clustering

As discussed earlier chapter on financial risk management, we are more interested in tail risk. Thus a detailed investigation of the tail region, particularly in light of the stylized facts that financial time series display heavy tails, is essential. Also, one must note that due to volatility clustering, an extreme value of returns is followed by another extreme value, and smaller return values follow a smaller value which has implications for financial risk management.

2.1 Asset pricing models

Before proceeding further, let us look at some popular Asset pricing models. We will also see them in the context of stylized facts and risk management. The references used for this section are [7][17][24].

2.1.1 Bond pricing

Let $B_t \geq 0$ denote the price of a deterministic financial asset(bonds) at time t, then the differential equation used for pricing bond is given by

$$B_t = B_0 e^{rt} \tag{2.1}$$

where r is the interest rate, and $B_0 \geq 0$ is the bond price at the initial price. This is a case of deterministic pricing where no inherent stochasticity is involved.

2.1.2 Geometric Brownian Motion

Let $S_t \geq 0$ denote the stock price at time $t \geq 0$, then the differential equation in the Geometric Brownian Motion model is given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.2}$$

which gives

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \tag{2.3}$$

where μ and $\sigma > 0$ are called the drift and volatility coefficients, respectively and $S_0 > 0$ is the stock price at initial time, and W_t is the standard Wiener process.
However, this model is built on the assumption of normality of log prices which is at odds with stylized facts like heavy tails and volatility clustering, which are important from a risk management perspective. However, still, this is a fundamental model and serves as a starting point.

2.1.3 Jump diffusion

The reference used for this section is [24].In Jump diffusion, the differential equation is given by

$$dS_t = \mu S_t dt + \sigma S_t dW_t + S_t dM_t \tag{2.4}$$

where M_t is a compound Poisson process defined as $M_t = \sum_{i=1}^{N_t} \xi(i)$ where $\mathcal{N} = \{N_t\}$ is a Poisson process with intensity $\lambda > 0$ and $\xi(i)$ is a sequence of independent random variables with identical CDF with zero mean and finite variance. One must note that although with appropriate distribution of ξ, one may generate heavy tails. Nevertheless, we still cannot capture volatility clustering due to the independent nature of jumps due to the memoryless property of the inter-arrival times of Poisson processes.

2.1.4 Hawkes process

Hawkes process is an advancement to Jump diffusion as it can also capture volatility clustering. The reference used for the section is [17]. It is a self-exciting and mutually self-exciting process. Let us define the univariate self-exciting process. Let $\lambda(t) > 0$ denote the intensity of the Hawkes process. For univariate self-exciting process

$$\lambda(t) = \mu + \sum_{T_i < t} \gamma(t - T_i) \tag{2.5}$$

where $0 < T_1 < T_2 < .... < T_n < ...$ are the time when an event occurs which is jump(or some extreme value) in this case, and $\gamma(i) \geq 0, i > 0$ is the exciting kernel, and this changes the intensity of the jump events which in turn gives rise to volatility clustering due to the increased intensity.
Some of the commonly used excitation kernels are: -

Exponential Kernel Here

$$\gamma(i) = \alpha\beta e^{-\beta i}, i > 0$$

and $\alpha, \beta > 0$

Powerlaw Kernel Here

$$\gamma(i) = \frac{\alpha\beta}{(1 + \beta i)^{1+p}}, i > 0$$

and $\alpha, \beta, p > 0$

One may also incorporate mutual dependencies among the stocks, which will capture mutual self-excitement of the stocks.

There are many other models used for asset pricing, like ARCH, GARCH, Fractional Brownian motion, etc. Still, we focused on some fundamental models and discussed stylized facts important for risk management. There are also some important things to remember while thinking about risk management is about dependence. In light of volatility clustering, all extreme events will cluster together, which has an important bearing on risk management. Also, properties like long-range dependence, if present, may have some vital bearing. We will discuss this further, but let us focus first on heavy tails, their general implications, and finance and risk management.

2.2 Heavy tails

Heavy tails are any distribution in which the probability of observing extreme events decays more slowly than the exponential/normal distribution. In essence, the distribution has the probability of extreme events or outliers significantly higher than expected from a normal distribution. Heavy-tailed distributions are more likely to generate extreme values far from the distribution's mean or median. The presence of heavy tails can have important implications for risk management and decision-making, as it can result in unexpected and potentially catastrophic events occurring with higher frequency than expected.

2.2.1 Background Mathematics

Much of the statistics deals with the non-heavy-tailed distribution, which decays at an exponential rate as by normal distribution and thus quickly decays to zero. Let us have a brief view of the background mathematics. The reference used in this section is [33].

Regular variation A heavy tail possessed by Pareto tail with index $\alpha > 0$ and $x > 0$ given by

$$P[X > x] = x^{-\alpha}, x > 1 \tag{2.6}$$

Generalized version is

$$P[X > x] = x^{-\alpha} L(x) \tag{2.7}$$

where L is slowly varying such that

$$\lim_{t \to \infty} \frac{L(tx)}{L(t)} = 1$$

Which is equivalent to

$$\lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)} = x^{-\alpha}, x > 0 \tag{2.8}$$

Where $1-F$ is a regular variation of order $-\alpha$ denoted as $1-F \in RV_{-\alpha}$

2nd order Regular variation Similar as in case of Regular variation, 2nd order regular variation is denoted by $1-F \in 2RV(-\alpha,\rho)$ where ρ is a 2nd order parameter.For 2nd order regular variation there should exist an $B(t) \to 0, t \to \infty$ such that the expression holds

$$\lim_{t\to\infty}\frac{\frac{1-F(tx)}{1-F(t)} - x^{-\alpha}}{B(t)} = H(x) := cx^{-\alpha}\int_1^x u^{\rho-1}\,du\,, x > 0 \tag{2.9}$$

Theorem 2.2.1. *Weak Laws of Large numbers: Let $X_1, X_2, ..., X_n$ be an infinite series which is independent and identically distributed Lebesgue integrable random variables such that $E(X_n) = \mu$ and sample average $\bar{X_n} = \frac{X_1+X_2+...+X_n}{n}$. Then for $n \to \infty$, $\bar{X_n} \to \mu$*

2.3 Test for the fit of distribution

Before discussing more on heavy tails, let us briefly overview how we say from data that a particular seems to follow a specific distribution. Support of the data, summary statistics like mean, median, mode, standard deviation, kurtosis, skewness, and other measures, the process generating the data, etc. are the few ways which help say on which distribution a data comes from. We may also use some quantile-based estimates, which gives us more idea of whether data comes from a given distribution. Some sophisticated methods we use to comment on distribution are Quantile-Quantile(QQ)-plots and the goodness of fitness test. The basic idea behind a goodness of fit test is to compare the observed data with the expected values from a theoretical model. If the observed data is a good fit for the model, the test will indicate that the model is valid. If not, the test will indicate otherwise. The chi-squared test, Kolmogorov-Smirnov test, and Anderson-Darling test are some of the most commonly used tests.

2.3.1 Quantile Quantile plot

The Quantile Quantile plot is a prevalent particle approach to check the fit of the distribution .[29] is the main reference for this section. Quantile is defined as the fraction(or percentage) of points below a certain value. For example, 0.70(70%) quantile is the point at which 70% of data is below that value. Thus a Quantile Quantile plot(QQ plot) is a plot of the quantiles of two data sets against each other. So in, this may be used to compare two data sets of equal length(although nonequal length variants also exist). We can consider one of the data sides coming from the theoretical quantiles from a known distribution. Thus on one side, we have sample quantiles from our data and theoretical quantiles from known distribution and compare them. The distributions are the same if both are on the 45^0 reference line.
From the QQ plot of normal distribution, we can understand many aspects of the distribution. Based on how the quantiles behave with respect to the reference line, one may know whether data is normally distributed, thin-tailed, or heavy-tailed.

2.3.2 Kolmogorov-Smirnov test

The Kolmogorov-Smirnov (KS) test is a non-parametric test used to compare the observed data with a theoretical continuous distribution. The reference for this section is[26]. The test statistic is based on the maximum absolute difference between the empirical distribution function (EDF) of the observed data and the cumulative distribution function (CDF) of the theoretical distribution.
In the Kolmogorov-Smirnov test, we test for

$$H_0 : \text{Data follows a given distribution}$$
$$\text{vs}$$
$$H_a : \text{Data does not follow a given distribution}$$

Let $X_1, X_2, ..., X_n$ be the data points arranged in ascending order. We define EDF as $F_n(x) = \frac{k}{n}$ where $k = \{1, 2, ..., n\}$. Let $F(x)$ be the CDF of our known distribution on which were are testing our hypothesis. Let $D_n = \sup_x |F_n(x) - F(x)|$. This test statistic is used for testing the hypothesis and commenting on the fit of the distribution.
These test statistics can be used to test hypotheses about the shape of the underlying

distribution, such as whether the data belongs to a particular distribution or not. Note that for the Kolmogorov-Smirnov test to work, we are X_k should be distinct and thus not appropriate if there are ties in our data.

2.3.3 Anderson-Darling test

The Anderson-Darling (AD) test is a non-parametric test that compares the observed data with a theoretical distribution. The reference used for this section is [28]. The test statistic is based on a weighted sum of the differences between the EDF of the observed data and the CDF of the theoretical distribution. In the Anderson-Darling test, we test for

$$H_0 : \text{Data follows a given distribution}$$
$$\text{vs}$$
$$H_a : \text{Data does not follow a given distribution}$$

The weights are chosen to give more weight to the tails of the distribution, where deviations from the theoretical distribution are more important and thus better for our study as we focus more on the extreme values. This can be used even if we have ties in our data. The test statistics is $AD^2 = -n - P$ where

$$P = \sum_{i=1}^{n} \frac{2i-1}{n} [\ln F(X_i) + \ln (1 - F(n+1-i))]$$

.

These test statistics are used to test the hypothesis of the fit of a distribution.

2.4 Implications of heavy tails

Saying a process is from a heavy-tailed distribution has numerous implications beyond changing the distribution function. The reference used for this section is [39]. The implications are listed below: -

1. Weak laws of large numbers as stated in **theorem 2.2.1** work, but our sense of large numbers coming from general statistical notions does not apply directly. The mean converges very slowly, and how large is enough for convergence to happen is a big issue. The mean of the distribution will rarely correspond to the sample mean; it will have a persistent small sample effect.

2. Commonly used measures of central tendencies like mean, standard deviation, and many other estimators may not be that informative and, in some cases, may not even exist.

3. Some other measures, apart from the commonly used measures, will be required to have a complete picture. For Pareto, the value of the tail index α is helpful to look further into parts of the tail that are not observed.

4. The extreme values, which are much farther from the center of the distribution, are much more informative and valuable and thus need special attention.

5. In layperson's terms, from the perspective of information, the data point that is much farther from the mean, median, or central part of the distribution is much more informative. A disproportional amount of information is present in the tails, which is very important and needs to be taken care of.

6. Given that the financial tail distribution does seem to follow the Pareto distribution, the value of the tail index α may be helpful in addition to the mean, median, and standard deviation.

7. Many commonly used financial metrics, like Beta and the Sharpe Ratio, are not informative. Although some solutions like [23] exist.

8. Dynamic hedging never reduces the risks associated with financial options.

9. Parameter estimating techniques like the method of moments (MoM) fail to work as higher moments are uninformative or do not exist. However, techniques like Maximum likelihood methods may work for some estimating parameters of some distributions.

10. The linear least-square regression does not work due to the failure of the Gauss-Markov theorem, and so do many popularly used tools based on that.

2.4.1 Behavioural finance theories in the context of heavy tails

A large amount of literature in behavioral finance is centered around some sense of rationality. Generally, the theories use mean and standard deviation and argue around the irrationality of human beings and the fact that human beings are utterly sensitive to risk. Many theories, like the Prospect theory and other advancements, as discussed in [41], help us understand that. Now we know that many phenomena in the real world may come from processes following a heavy-tailed distribution; thus, the mean, median, and standard deviation are insufficient. Other measures also need to be taken care of to understand the situation. Therefore, this may help us understand to some extent why human beings generally follow this kind of behavior and the shape of the prospect curve. This also poses the prospect of further research in this direction.

2.5 Heavy tails in finance

As discussed in section 1.2, financial time series of log returns are heavy-tailed and seem to follow Pareto distribution, and thus, based on the discussion in section 2.2.1 and section 2.4, some estimates like tail index α may be helpful. Thus we will have a detailed look at the estimation of parameters of the Pareto distribution and the value of the tail index α. Before discussing more on the tail index α, let us have a brief overview of Pareto distribution.

2.5.1 Pareto distribution

The references used in this section are [1][33][36]. Generalized Pareto distribution, as discussed in [36], is commonly used to model other distributions' tails. The Generalized Pareto distribution (GDP) has a probability density function: -

$$gPa(y|k,\sigma) = \frac{1}{\sigma}(1 - ky/\sigma)^{1/(k-1)} \ k \neq 0 \tag{2.10}$$

$$gPa(y|k,\sigma) = \frac{1}{\sigma}e^{-y/\sigma} \ k = 0 \tag{2.11}$$

where $\sigma > 0$ is the scale parameter and k is the shape parameter. If $k \leq 0$ support of the distribution is $\mathbb{R}^+$ else it is $(0, \sigma/k]$ if $k > 0$.

2.5.2 Estimating tail index

As discussed in section 2.5, we are interested in the value of the tail index. [36] discusses methods like the Maximum Likelihood estimator, Probability weighted moments, Jeffreys prior, Bayesian Reference Intrinsic (BRI) estimator, and Hill estimator. Based on the results in the paper, the BRI and hill estimators seem to be working best. They are thus reasonable to use due to low relative MSE and the small absolute value of bias for estimating the shape parameter, which corresponds with our tail index. Therefore, let us discuss this further in detail.

2.5.3 BRI and Hill estimator

The Bayesian Reference Intrinsic (BRI) estimator is discussed in [36]. The Bayesian Reference Intrinsic (BRI) estimator value of shape parameter k is based on a loss function given by

$$d(k_0|\hat{k}) = \int_0^\infty \delta(k, k_0) Ga(k|n-1, n/\hat{k}) dk \tag{2.12}$$

where $d(k_0|\hat{k})$ denotes the value of the loss function on which we have to find minima based on which we get the estimate of the shape parameter k^* that is the value of tail index α. Ga denotes Gamma distribution and k, n as defined in section 2.5.1 and $\hat{k} = (\log \frac{max(X_1, X_2, ..., X_n)}{\prod_{i=1}^n X_i^{1/n}})^{-1}$ and

$$\delta(k, k_0) = -n \log\theta + n\theta - n \ if \ \theta < 1$$
$$\delta(k, k_0) = n \log\theta + n\theta^{-1} - n \ if \ \theta \geq 1$$

where $\theta = \frac{k}{k_0}$. The loss function optimal point has to be computed numerically, but we can approximate it as $k^* = \hat{k}(1 - \frac{3}{2n})$ which the BRI estimators value of the tail index α. One of the advantages of the BRI estimator is that it is invariant under the linear transformation of the data.

The Hill estimator, as discussed in [33][36], is a way to calculate the tail index α, thereby detecting heavy tails. Although this estimator is not linear invariant, it is consistent for Pareto and even a larger class of distribution known as Regular Variation as discussed section 2.2.1. The reference for the definition, theorem, and discussion below is [33]. Let $X_1, X_2, ..., X_n$

be independent and identically distributed(iid) from a heavy-tailed distribution of which we wish to estimate the α. Let $X_{(1)} > X_{(2)} > .... > X_{(n)}$ be the order statistics.

Definition 2.5.1. *The Hill Estimator is defined as* $H_{p,n} := \frac{1}{p}\sum_{i=1}^{p} \log \frac{X_{(i)}}{X_{(p+1)}}$

Theorem 2.5.1. *Suppose* $\{X_t\}$ *is stationary and the marginal distribution satisfies*

$$P[X_1 > x] = x^{-\alpha}L(x), x \to \infty \tag{2.13}$$

If either

1. $\{X_t\}$ *is iid or*
2. $\{X_t\}$ *is weakly dependent or*
3. $\{X_t\}$ *is an* $MA(\infty)$ *process*

then if $n \to \infty$ *and* $p \to \infty$ *but* $\frac{p}{n} \to 0$*, we have*

$$H_{p,n} \xrightarrow{P} \alpha^{-1} \tag{2.14}$$

This result assumed first-order regular variation as discussed in eq. (2.8). If we assume second-order regular variation as discussed in eq. (2.9) and some further restrictions on p, we have.

$$\sqrt{p}(H_{p,n} - \alpha^{-1}) \to N(0, \alpha^{-2}) \tag{2.15}$$

Based on eq. (2.14) and eq. (2.15), one can asymptotically calculate the Hill estimator. So for calculating the hill estimator, we plot the Hill coefficient vs. p, and we get a stable region around the middle part of the plot where conditions of **theorem 2.5.1** are satisfied. Sometimes due to some volatility in the plots, it becomes a bit difficult to calculate the value of α. Then in such situations, we use the SmooHill estimator.

$$SmooH_{p,n} := \frac{1}{(u-1)p}\sum_{j=p+1}^{up} H_{j,n} \tag{2.16}$$

The asymptotic variance of $SmooH_{k,n} = \frac{2}{\alpha^2 u}(1 - \frac{logu}{u})$ which is less than $1/\alpha^2$ of Hill and larger u we choose the better results we get. Generally, $u = 2, 3$ works. In Alt plotting we plot $H^{-1}_{\lceil n^\theta \rceil,n}$ vs θ for $0 \leq \theta \leq 1$. So this is basically a change in scale at which we are looking at the graph and helps in certain cases.

Being discussed these estimators [33] also discuss QQ plots to calculate the value of α. In that, we first pick k upper order statistics and plot $(-log(1 - \frac{j}{k+1})), logX_{(j)}), 1 \leq j \leq k$. If conditions or regular variations are satisfied, the data will closely follow a straight line whose slope will be $1/\alpha$.We can also use the QQ estimator given by

$$\hat{\alpha^{-1}}_{k,n} = [\frac{1}{k}\sum_{i=1}^{k}(-\log(\frac{i}{k+1}))\log(\frac{X_{(i)}}{X_{(k+1)}}) - \frac{1}{k}\sum_{i=1}^{k}(-\log(\frac{i}{k+1}))H_{k,n}]\times$$
$$[\frac{1}{k}\sum_{i=1}^{k}(-\log(\frac{i}{k+1}))^2 - \frac{1}{k}\sum_{i=1}^{k}(-\log(\frac{i}{k+1}))^2]^{-1}$$

.We can use the above formula and plot the dynamic qq plot by plotting (k,$\frac{1}{\hat{\alpha^{-1}}_{k,n}}$) for $1 \leq k \leq n$ and work similarly as we did in Hill plot.

Thus all of these methods gave us the value of α, which can provide us with way more insights into risk aspects apart from the standard summary statistics and is applicable even if our data belongs to the class of First order regular variation. [36] used this tail index for accurately mapping risk in financial risk management. Thus, this is a useful measure applicable to multiple risk scenarios in financial data.

2.5.4 Dependence of stock return

As one knows from section 1.2, the linear autocorrelation function of stock returns is insignificant, but that of absolute value or square returns is significant and decays slowly. Also, the phenomena of volatility clustering imply some form of dependence among stock returns which may be non-linear or have a long range. So, before moving forward with this, we first need to define some important definitions. The main reference used in this section is [10]

Definition 2.5.2. *Long range dependence: A stationary process X_t(with finite variance) is said to have long-range dependence if its autocorrelation function $C(\tau) = corr(X_t, X_{t+\tau})$ decays as power of lag τ: $C(\tau) = corr(X_t, X_{t+\tau}) \underset{\tau\to\infty}{\sim} \frac{L(\tau)}{\tau^{1-2d}} 0 < d < 1/2$ where L satisfies $\lim_{x\to\infty} \frac{L(at)}{L(t)} \to$*

$1 \forall a > 0$

Definition 2.5.3. *Short range dependence: If autocorrelation function decreases at a geometric rate:* $\exists K > 0, c \in [0,1], |C(\tau)| \leq Kc^{\tau}$

Definition 2.5.4. *Self similarity: A stochastic process* $(X_t)_{t\geq 0}$ *is self similar if* $\exists H > 0$ *such that for any scaling factor* $c > 0$ *such that* $(X_{ct})_{t\geq 0}$ *and* $(C^H X_t)_{t\geq 0}$ *have the same law(This process can't be stationary)*

Definition 2.5.5. *Arbitrage: An arbitrage strategy is a strategy where with no money, there is a positive probability of earning something positive without the possibility of losing anything.*

Models like GARCH are discussed in [10] and how GARCH successfully captures volatility clustering but gives rise to tails decaying at an exponential rate. Detecting Long range dependence as defined in **definition 2.5.2** depends on the behavior of autocorrelations at large lags, which is challenging to estimate empirically. Thus we try to formulate models with long-range dependence on self-similar processes. Let us briefly discuss self-similarity and the consequences if stock prices are self-similar. Some asset pricing models, like Fractional Brownian, are self-similar processes whose increments display long-range dependence. Self-similarity does not imply long-range dependence, nor does long-range dependence imply it. Asset returns do not seem to hold self-similarity in the strict sense. Let us have a look at the consequences if asset prices are self-similar.

Self Similarity from **definition 2.5.4**

$$\implies \forall t > 0, X_t \stackrel{d}{=} t^H X_1 \implies F_t(x) = F_1(\frac{x}{t^H}) \tag{2.17}$$

Differentiating (2.17)

$$\rho_t(x) = \frac{1}{t^H}\rho_1(\frac{x}{t^H}) \tag{2.18}$$

Substituting $x = 0$ in (2.18)

$$\implies \rho_t(0) = \frac{1}{t^H}\rho_1(0) \tag{2.19}$$

$$\implies \ln \hat{\rho}_t(0) = -H \ln \frac{t}{\Delta} + \ln \hat{\rho}_\Delta(0) + \epsilon \tag{2.20}$$

Note that just satisfying only (2.20) does not imply the process is self-similar. It is a necessary condition but not sufficient.(2.20) May be used for calculating and testing for self-similarity. (Note (2.20) as appears in [10] is without a negative H sign which, when worked out by us, was found to have a negative sign which has been written in the thesis)

Examples of literature on models following self-similarity or long-range dependence and their issues are discussed in [10]. Let $\mathcal{P}$ denote a real-world measure. [16] gives a very fundamental theorem of arbitrage pricing theory which states that price evolution of $S(t)$ is arbitrage-free iff $\exists \mathcal{Q} \sim \mathcal{P}$ which is a probability measure such that $S(t) = E^{\mathcal{Q}}[S(T)|\mathcal{F}_t]$. Thus, under a measure $\mathcal{Q}$, it is a martingale; in $\mathcal{P}$ is a semi-martingale. Models like Fractional Brownian motion give results that fail to satisfy semi-martingale conditions. Thus $\exists$ strategies that contain arbitrages if we assume our model follows fractional Brownian motion. [10] also gives examples of models that capture long-range dependence and are arbitrage-free. One must also note that if the asset price follows the Fractional Brownian motion model, strategies lead to arbitrage. As given in [35], one will understand that these strategies cannot be carried out in the real world as they involve an arbitrarily large amount of trading in finite time. Some practical constraints, like an arbitrarily small amount of trading time, can rule out those strategies. Thus one must not simply rule out models like Fractional Brownian Motion which gives self-similarity and long-range dependence together, or any other model like that just due to considerations of arbitrage which are much more theoretical and not easy to be followed in practice.[10] also discusses some factors like Heterogeneity of time horizons of economic agents, Switching between trading strategies, Investor inertia which may lead to long-range dependence and some other modeling perspectives like Evolutionary models that may be helpful for the purpose.

In light of long-range dependence and heavy tail, we will have a re-look at section 2.1.4. section 2.1.4 discusses two kernels. Let's look at it from the perspective of the time effect that the effect of the jump vanishes. We will understand from previous discussions that the exponential kernel's decay rate will be way higher than that of the power law kernel. Thus the effect of jump on a process with an exponential kernel will not be as long-lived as that with a power law kernel. Thus, the exponential kernel is appropriate if we wish to have short-range dependence in returns data, and the power law kernel is appropriate for long-range dependence. Thus, these kernels, as discussed in section 2.1.4, relate to the duration of dependence among stock returns.

2.5.5 Autocorrelation value in the context of heavy tails

Let $X_1, X_2, ..., X_n$ be a time series and $\hat{\rho}(h)$ denote sample autocorrelation function(acf) values as given by

$$\hat{\rho}(h) = \frac{\sum_{i=1}^{n-|h|}(X_i - \bar{X})(X_{i+h} - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{2.21}$$

In non-heavy-tailed scenarios, this sample acf vale $\hat{\rho}(h) \xrightarrow{P} \rho(h)$.But, if our data comes from a heavy-tailed distribution mean, variances may not even exist. A more appropriate modification of the autocorrelation function for heavy-tailed distribution is

$$\hat{\rho}(h) = \frac{\sum_{i=1}^{n-|h|} X_i X_{i+h}}{\sum_{i=1}^{n} X_i^2} \tag{2.22}$$

More insights into the Autocorrelation values given by eq. (2.21) and eq. (2.22) are discussed in [33] [34]. Limitations of Auto correlations values in heavy tails are discussed in [34].

2.6 Change Point

The main reference used in this section is [14][15]. Let us describe the general formulation of Change point analysis. Let $X_1, X_2, .., X_q, .., X_n$ be our time series values, and we have distribution functions $F_1, F_2, .., F_q, .., F_n$ be the distribution functions from the parametric family $F(\phi)$, where $\phi \in \mathbb{R}^k$ and $k \in \mathbb{N}$. Our problem is determining q and locations $1 < k_1 < k_2, ... < k_q < n$ such that $\phi_1 = ... = \phi_k \neq \phi_{k_1+1} = ... = \phi_{k_2} \neq ... \neq \phi_{k_q-1} = ... = \phi_{k_q} \neq \phi_{k_q+1} = ... = \phi_n$. This is done by testing hypothesis

$$H_0 : \phi_1 = \phi_2 = ... = \phi_n = \phi$$
$$\text{vs}$$
$$H_a : \phi_1 = ... = \phi_{k1} \neq \phi_{k_1+1} = ... = \phi_{k_2} \neq ... \neq \phi_{k_q-1} = ... = \phi_{k_q} \neq \phi_{k_q+1} = ... = \phi_n$$

This was the general case. In our case, we will study change points in the context of changes in the mean and variance of log returns of stocks. Assuming our distribution is $N(\mu, \sigma^2)$ and analyzing some results, we do not observe much difference in the number of change points as well as the locations of change points in the context of stock market log-returns if we detect changes in both μ, σ^2 or just σ^2. Thus we focus our attention mostly on changes in variance in stock market returns time series. This kind of detection has implications for regime changes in the context of changes in variance. As the variance is closely tied with volatility, we are detecting changes in volatility.

2.6.1 Detecting changes in variance

Ways to detect changes in variance are discussed in [14]. For our study, we focus our attention on [15]. In variance change point detection, we test for the hypothesis.

$$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_n^2 = \sigma^2$$
$$\text{vs}$$
$$H_a : \sigma_1^2 = ... = \sigma_{k1}^2 \neq \sigma_{k_1+1}^2 = ... = \sigma_{k_2}^2 \neq ... \neq \sigma_{k_q-1}^2 = ... = \sigma_{k_q}^2 \neq \sigma_{k_{q+1}}^2 = ... = \sigma_n^2$$

The rest being the same as defined in the previous section. [15] uses a Binary segmentation procedure for detecting change points. We can detect only one change point using binary segmentation in the original form. Using the procedure again on the separate regions, we obtain more points and continue until no further change points are found. Thus, we check for

$$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_n^2 = \sigma^2$$
$$\text{vs}$$

$$H_a : \sigma_1^2 = ... = \sigma_p^2 \neq \sigma_{p+1}^2 = ... = \sigma_n^2 \tag{2.23}$$

. In binary segmentation procedure, the paper uses Schwarz information criterion (SIC) defined as $-2\log L(\hat{\theta}) + k\log n$ where p is the number of free parameters in the model and $L(\hat{\theta})$ is maximum likelihood function of the model. For the context of the original hypothesis, we check if $SIC(n) \leq min_k SIC(k)$. If it is true, we will accept H_0. Else we will reject H_0. Based on some calculations, piratical conditions, and derivation assumptions, we get the condition as equivalent to $SIC(n) \leq min_{2\leq k\leq n-2} SIC(k)$.We estimate the position of change point $\hat{k}$ as

$$SIC(\hat{k}) = min_{2\leq k\leq n-2} SIC(k) \tag{2.24}$$

The value of $\hat{k}$ obtained is strongly consistent with the true position, as proved in [15].The statement of theorem is

Theorem 2.6.1. *Let p be the true position of the change point as in (2.23). Let $\hat{k}$ be the estimate of p from eq. (2.24). Then $\hat{k}$ is strongly consistent for p.*

Due to **theorem 2.6.1**, we can follow the procedure of the Binary segmentation method with SIC criterion and get the locations of the change points with strong consistency.

2.7 Change point in the context of Volatility clustering and Risk

Thus, as discussed, volatility clustering and risk are related and important, as volatility clustering worsens a bad event. As discussed in section 2.6.1, we are detecting a change in variance. So if there is some volatility clustering event happening, then surely a change point will be there where there is a change in variance. Due to these volatility clusters, there will be a sudden jump in the variance value. One should also note that though the change point detects changes in variance and hence detects all points where volatility clustering happens, that does not mean all points where there is a change point associated with the volatility cluster.

But said that this change point method, as discussed in section 2.6.1, is a perfect tool for detecting regime changes and performing analysis on changes in volatility as well as volatility clusters as points where these phenomena occur are subsets of these.

Thus, in this chapter, we looked at some important stylized facts like heavy tails and volatility clusters, which are important from the risk management perspective. We have a look at some asset pricing models and some properties like long-range dependence and looked at them in the light of the important stylized facts. We also looked at QQ plots and goodness of fit tests for detecting the shape of a distribution. We also looked at a change point, how to detect changes in variance, and how they all connect with risk.

Chapter 3

Methodology

3.1 Chatterjee correlation

Before proceeding further, let us discuss a new non-parametric way to calculate the correlation. This new correlation coefficient is [8]. This is a simple correlation coefficient that consistently estimates some simple and interpretable measure of the degree of dependence between the variables and has a simple asymptotic theory under the hypothesis of independence. The new correlation coefficient if no ties are present is defined as: -

$$\xi_n(X,Y) := 1 - \frac{3\sum_{i=1}^{n-1} |r_{i+1} - r_i|}{n^2 - 1} \tag{3.1}$$

where X, Y are random variables and we have rearranged data $(X_{(1)}, Y_{(1)}), ..., (X_{(n)}, Y_{(n)})$ such that $X_{(1)} \leq ... \leq X_{(n)}$ and r_i is the rank of Y_i that is the number of j such that $Y_{(j)} \leq Y_{(i)}$. And if ties are present, we define it as -

$$\xi_n(X,Y) := 1 - \frac{n\sum_{i=1}^{n-1} |r_{i+1} - r_i|}{2\sum_{i=1}^{n} l_i(n - l_i)} \tag{3.2}$$

where l_i is a permutation of 1,...,n so the denominator expression becomes $n(n^2 - 1)/3$.
This new correlation coefficient is non-parametric, capturing the relationships among data, whether linear or non-linear. Thus, we plan to use this to study the relationships among data, as the usual correlation coefficient is suitable for capturing only linear trends among

data.

3.2 Data and methods used

Based on the above-discussed theories, we did our analysis. The data used in the analysis was downloaded from the Bloomberg Terminal. We worked with stock market data (daily close price data). As discussed in section 1.2, we will first look at the stylized facts of financial time series. So we took data from 2 stocks, TESLA and APPLE, and four index data, S&P 500, NIFTY 50, SENSEX, and KOSPI. The Bloomberg ticker names were TSLA US Equity, AAPL US Equity, SPX Index, SENSEX Index, NIFTY Index, and KOSPI Index. We took the max data on a daily time scale, that is, the data from the date stock or index started trading till 30/3/2023 as at 10:00 pm IST. This boils down to data from 28-06-2010 to 30-03-2023 for TESLA, 12-12-1980 to 30-03-2023 for AAPL, 30-12-1927 to 30-03-2023 for S& P 500, 03-04-1979 to 29-03-2023 for SENSEX, 03-07-1990 to 29-03-2023 for NIFTY 50 and 04-01-1980 to 30-03-2023 for KOSPI.

One must note that we have included KOSPI, the South Korean index. This was included explicitly because South Korea was a developing nation before 1996, when it was included in OCEED, recognizing it to be developed. Due to this shift, I expect some change in the way markets should work after that due to a part of an advanced group. This will be useful in gaining more insights into developing vs. developed counties' market risk. Thus for KOSPI, we did three analyses: KOSPI entirely, KOSPI (developing) (till 1996), and KOSPI (developed) (post-1996).

From discussions in section 2.1, we can understand why to use log returns of data while analyzing returns time series. When one assumes section 2.1.2, one will expect log returns to be iid and normally distributed.fig. 4.1fig. 4.2fig. 4.4fig. 4.5fig. 4.3fig. 4.6fig. 4.7fig. 4.8 gives QQ pots (as discussed section 2.3.1) of the log returns of APPLE, TESLA, S& P 500, SENSEX, NIFTY, KOSPI, KOSPI(developing) and KOSPI (developed) respectively. As we can see consistently in all QQ plots in lower theoretical quantiles, the points are below the blue reference line; after that, in higher ones, they are above. This shows that the stock market log returns show heavy-tailed behavior in all scenarios as discussed in section 1.2.

We further did an ACF analysis. The ACF analysis was done with the modified method(as discussed in section 2.5.5 as we already know that our data is from a heavy-tailed distribution. For R implementation (demean=FALSE) in the function, does the implementation of the heavy-tailed modified version we have done for our analysis. We further focused on the loss tail region and checked whether the condition for regular variation as discussed in section 2.2.1 are satisfied. The package "ptsuite" as discussed in [27] was used to check the hypothesis of Pareto fit. First, we analyzed the 5% quantile. This threshold worked for the stocks and some indices; we went up to the 1% quantile for the rest. We found that stocks and indices that did not give positive results for Pareto at 5% gave positive results at some value greater than 1%. I took the max quantile, which gave a positive result for the Pareto distribution, as estimated from more data is better if underlying assumptions of using (2.14) are valid. The point after which the tail follows conditions of Regular variation as discussed section 2.2.1 is known as Karamata Point. The package gives a p-value for the data set, and I set the significance level as our standard 5%. If the p-value is greater than that, we accept that the data is coming from a Pareto distribution; else not. This procedure is as given in [27]. From the part where we can detect Pareto tails, we found the Hill estimator as discussed in section 2.5.3. Package "evir" was used to get the hill plot details found on [30]. Thus we computed the hill coefficient. One can look for ACF and Hill plots at fig. 4.1fig. 4.2fig. 4.4fig. 4.5fig. 4.3fig. 4.6fig. 4.7fig. 4.8 of the log returns of APPLE, TESLA, S& P 500, SENSEX, NIFTY, KOSPI, KOSPI(till 1996) and KOSPI (after 1996) respectively. Further, we implement changepoint as discussed in section 2.6.1 using R package "changepoint" as discussed in [31][32]. Based on changepoint results, we considered each segment as a regime. Below are the important summary measures like Min, Median, Max, Mean, Standard deviation, Skewness, and Kurtosis. We also computed the value of VaR and ETL using the historical method at 5 % significance, the Value of the Hill estimator, and the density of changepoints, which is the number of change points over the total number of days. The tables below provide the results of the above summary measures for the stocks discussed.

Stock	Min	Median	Max	Mean	Standard deviation	Skewness
APPLE	-0.73	0	0.289	0.0067	0.0287	-1.7383
TESLA	-0.2365	0.0013	0.3408	0.0016	0.0365	0.2005
NIFTY	-0.1390	0.0007	0.1633	0.0005	0.0157	-0.2245
S& P 500	-0.2290	0.0005	0.1537	0.0002	0.0119	-0.4670
SENSEX	-0.1410	0.0007	0.1599	0.0006	0.0156	-0.091
KOSPI	-0.1280	0.0003	0.1128	0.0003	0.0144	-0.2341
KOSPI(till 1996)	-8.77E-02	-8.19E-05	5.25E-02	3.77E-04	0.0114	0.1305
KOSPI(after 1996)	-0.1281	0.0007	0.1128	0.0002	0.0163	-0.3083

Stock	Kurtosis	VaR	ETL	Hill estimator	Changepoint density
APPLE	46.36897	-0.0419642	-0.0644849	3.04	0.002906977
TESLA	6.864349	-0.05265287	-0.0828028	3.15	0.004360012
NIFTY	8.420805	-0.02334204	-0.03719444	3.16	0.004154602
S& P 500	18.61262	-0.01697337	-0.02884243	3.12	0.003552621
SENSEX	7.17256	-0.02306856	-0.03628711	3.59	0.003236881
KOSPI	5.863361	-0.02196457	-0.03479543	3.7	0.003982339
KOSPI(till 1996)	2.776423	-0.01698626	-0.02441292	3.49	0.004417671
KOSPI(after 1996)	5.38741	-0.02621505	-0.04070001	3.63	0.003956178

Based on the table above, changepoint density positively correlates with skewness and negatively with kurtosis and Mean. We verified the proposition using linear regression, which gives significance to skewness, kurtosis, and Mean coefficients when run individually. We also identified our risk measures VaR and CVaR as closely associated with standard deviation. We carried out changepoint statistics of detecting regime change using both mean and variance vs. only a variance. We can see that detecting regime change only using variance suffices as almost the same number of changepoints detected, and locations are also nearly identical. As changepoint detects regime change using standard deviation, which is associated with risk measures and changes in variance are closely related to regime change whose frequency is closely associated with skewness and kurtosis. [36] discusses using the Hill coefficient to calculate risk in the financial instruments discussed there.

Thus, I have summarized different summary statistics and explored their relation using the Chatterjee correlation as discussed in section 3.1. So based on the changepoint, I divided the time series into different regimes, calculated summary statistics in those individual regimes, and studied their dependence using the Chatterjee correlation. Based on the results, we found that the frequency of crashes in a particular regime depends on the standard deviation, 25th, 50th, and 75th quantile of log returns. The skewness and kurtosis have the most negligible dependence on them. These relationships have important implications for understanding which variables our model should take as input when trying to model these extreme values. This also helps us overcome some intuitions and enhances which factors are essential for extreme importance in impact and frequency that is data-driven and has empirical evidence.

We also looked at inter-arrival regime change and crash times, defined as a particular percentage fall and dynamic historical VaR, where crash changes on VaR value of the past few days(100 in our analysis). On inter-arrival data, we did many analyses like trying to fit exponential distribution based on Kolmogorov-Smirnov(KS) test and Anderson-Darling(AD) test as discussed in section 2.3.2 and section 2.3.3 respectively. We also did an ACF analysis of the inter-arrival crash times. We can find the ACF plots are the inter-arrival crash timings at fig. 4.9afig. 4.9fig. 4.10afig. 4.10fig. 4.11afig. 4.11fig. 4.12afig. 4.12 for APPLE, TESLA, S & P 500, Sensex, Nifty, KOSPI, KOSPI(developing), KOSPI(developed) respectively. The results of all of this analysis are discussed in the conclusion. One can see in the graphs that for Static VaR as our definition of Crash, we get significant values for even considerable lags, which is not the case in dynamic VaR. However, we can reject exponential fit in both cases, implying dependence and not following the memoryless property. We can obtain an exponential distribution fit for all indices for regime length, which is the interarrival time of changepoint locations. These findings have important implications for which one should base predictions, like just calibrating the value of lambda and the last changepoint location is sufficient. Going for any advanced machine learning algorithm will not yield any better results. Still, the same is not valid for crash timing prediction, implying scope for it as we can reject the hypothesis for exponential fit. Thus, these statistical analyses are better done before using any advanced methods, as this helps us understand whether these advanced methods will give better results or not. This also has essential bearings when designing models. Suppose we are trying to have some regime-switching models. In that case, we may have the regime switch timings as Markov, but having the same for jumps that are extreme events is not appropriate. These insights are critical before deciding on a model suitable for the problem. One should also have a look at the distribution of parameters and predicted values of the models because if there is some model where one of them is heavy-tailed, or the error term after correcting for other variables is still heavy-tailed, one should be very careful of the results as the mean and other summary measures may not exist. Thus there might be no bias or variance, or our number of data points may not be sufficient to reach the asymptotic values. The R code and graphs of the implementation are provided in the Appendix.

Chapter 4

Conclusion

Based on the results of log returns, we can see that the heavy-tailed aspect is displayed in the log returns data for all stocks. We also found no significant autocorrelation values in most assets, and for some, there is one lag; the value is relatively small. Also, as seen in our analysis, this is observed in developing economies but not followed, or the extent is less in developed economies. Although more extensive and detailed studies need to be carried out before commenting on anything at this stage, I conjecture that there is some change in mechanisms in financial markets in developing and developed economies due to factors like a better flow of information, easy access to exchanges, and greater public participation.

We also observed in changepoint analysis that there is a sudden jump in the volatility at specific locations that are not sustained for a very long time. Thus, this may be attributed to volatility clustering. We also extensively analyzed inter-arrival crash timings and regime change and found them to be memoryless for regime change and not for crashes defined by both static VaR and dynamic VaR; thus, although regime change is exponentially distributed, the same is not suitable for crash timings, and therefore, each crash has some past dependencies, shown by ACF plots. We get a very long time dependence for static VaR. We do not get many significant autocorrelation values for dynamic VaR; thus, one needs to explore the direction of non-linear dependence, as there is some dependence given a non-memoryless distribution. Granted, there might be a highly complex distribution whose analytical analysis may be difficult, but some machine-learning techniques for stock market analysis may be helpful as they can do it very effectively .[4] is a good read for the same.

One can also see from the results of the Chatterjee correlation that Volatility is related to Crash events. Thus, the changepoint, which helps us detect regime changes, is a vital statistic. The frequency of regime change is connected with volatility and heavy tails. APPLE, KOSPI post-developed, and S&P 500 seem to have a lower density of changepoints than NIFTY, SENSEX, and KOSPI pre-developed. This all connects with volatility, heavy tails, and, thus, risk. So one must comprehensively look at the issue in more detail, as this is critical for all researchers, practitioners, and governments. All of it has enormous implications and poses many challenges, which also warrants the use of some use of new statistics that behave differently than classical statistics and are thus not convenient but essential. One must also develop optimal strategies for allocating portfolios under the heavy-tailed distribution of returns, and [22] is a good read for the same.
Thus, as seen empirically, we know that the memoryless regime's length can only be a little help other than knowing the average expected time for the following regime change. However, the average time still depends on the degree of heavy tailless and volatility. But still, the rate has some important implications. But crash does have a dependence structure, and whether it is non-linear or long-range is to be determined. Both of these have significant implications. One must look at data with a substantial time delay if dependence is long-range to understand and predict currently, and just training machine learning models on a small local data set will not help. Thus, these questions warrant a much more detailed investigation and present many interesting mathematical and data analysis questions with a vast scope for further research. Also, looking at the distribution of error in the model, parameter, and predicted variables is important, as machine learning is based on the bias-variance tradeoff, which may not exist or be calibrated appropriately given the amount of data, which is something we need to think about. There are some scopes of researching expanding ML beyond the Bias Variance tradeoff for which summary statistics like the Hill estimator may be useful. We also looked at new perspectives, which may help us look at behavioral finance differently and yield new insights.

Working on financial risk management problems requires many advanced tools of mathematics and warrants development in mathematics to work rigorously on these issues we currently need to get. Looking at these works also helps us understand the limitations of our current data analytic tools. We also understood the new challenges to our current understanding of data science and statistics that the problems in finance give rise to. Although deep learning models may help predict these aspects to some extent, given that they are not interpretable, they are of little use without that. We cannot understand the underlying economic mechanisms underlying it, and thus, it is not much help if we wish to devise ways to prevent it. Therefore, this field has tremendous social impacts and many interesting questions that may interest almost all disciplines, as risk comes from all sources. Quantification, prediction, and prevention of risk help us in our long-run growth.

Appendix

4.1 R code

```
library(readxl)
library(ptsuite)
library(evir)
library(goftest)
library(changepoint)
library(PerformanceAnalytics)
library(vcd)
KOSPI <- read_excel(# file path, sheet = "KOSPI")
SPX <- read_excel(# file path, sheet = "SPX")
SENSEX <- read_excel(# file path, sheet = "SENSEX")
NIFTY <- read_excel(# file path, sheet = "NIFTY")
AAPL <- read_excel(# file path, sheet = "AAPL")
TSLA <- read_excel(# file path, sheet = "TSLA")
reta<-rep(0,length(AAPL$`Last Price`)-1)
for (i in 1:length(reta)) {reta[i]<-log(AAPL$`Last Price`[i+1]/
                                  +AAPL$`Last Price`[i])}
rett<-rep(0,length(TSLA$`Last Price`)-1)
for (i in 1:length(rett)) {rett[i]<-log(TSLA$`Last Price`[i+1]/
                                  +TSLA$`Last Price`[i])}
retn<-rep(0,length(NIFTY$`Last Price`)-1)
for (i in 1:length(retn)) {retn[i]<-log(NIFTY$`Last Price`[i+1]/
                                  +NIFTY$`Last Price`[i])}
rets<-rep(0,length(SENSEX$`Last Price`)-1)
```

```
for (i in 1:length(rets)) {rets[i]<-log(SENSEX$`Last Price`[i+1]/
                                  +SENSEX$`Last Price`[i])}
retsp<-rep(0,length(SPX$`Last Price`)-1)
for (i in 1:length(retsp)) {retsp[i]<-log(SPX$`Last Price`[i+1]/
                                  +SPX$`Last Price`[i])}
retk<-rep(0,length(KOSPI$`Last Price`)-1)
for(i in 1:length(retk)){retk[i]<-log(KOSPI$`Last Price`[i+1]/
                                  +KOSPI$`Last Price`[i])}
qreta<-reta[reta<quantile(reta,0.05)]
pareto_test(abs(qreta))
par(mfrow=c(1,2))
qqnorm(reta,main="QQ plot")
qqline(reta,col="blue")
acf(reta,demean=FALSE,main="ACF plot")
fita<cpt.var(reta,penalty="SIC",method="BinSeg",Q=100,test.stat="Normal",
    +class=TRUE,param.estimates=TRUE,minseglen=2)
hill(abs(qreta))
plot(fita, main="Changepoint")
qrett<-rett[rett<quantile(rett,0.05)]
pareto_test(abs(qrett))
qqnorm(rett,main="QQ plot")
qqline(rett,col="blue")
acf(rett,demean=FALSE,main="ACF plot")
fitt<cpt.var(rett,penalty="SIC",method="BinSeg",Q=100,test.stat="Normal",
    +class=TRUE,param.estimates=TRUE,minseglen=2)
hill(abs(qrett))
plot(fitt, main="Changepoint")
qretn<-retn[retn<quantile(retn,0.03)]
pareto_test(abs(qretn))
qqnorm(retn,main="QQ plot")
qqline(retn,col="blue")
acf(retn,demean=FALSE,main="ACF plot")
fitn<cpt.var(retn,penalty="SIC",method="BinSeg",Q=100,test.stat="Normal",
    +class=TRUE,param.estimates=TRUE,minseglen=2)
```

```
hill(abs(qretn))
plot(fitn, main="Changepoint")
qretsp<-retsp[retsp<quantile(retsp,0.02)]
pareto_test(abs(qretsp))
par(mfrow=c(1,2))
qqnorm(retsp,main="QQ plot")
qqline(retsp,col="blue")
acf(retsp,demean=FALSE,main="ACF plot")
fitsp<cpt.var(retsp,penalty="SIC",method="BinSeg",Q=100,test.stat="Normal",
    +class=TRUE,param.estimates=TRUE,minseglen=2)
hill(abs(qretsp))
plot(fitsp, main="Changepoint")
qrets<-rets[rets<quantile(rets,0.02)]
pareto_test(abs(qrets))
qqnorm(rets,main="QQ plot")
qqline(rets,col="blue")
acf(rets,demean=FALSE,main="ACF plot")
fits<cpt.var(rets,penalty="SIC",method="BinSeg",Q=100,test.stat="Normal",
    +class=TRUE,param.estimates=TRUE,minseglen=2)
hill(abs(qrets))
plot(fits, main="Changepoint")
qretk<-retk[retk<quantile(retk,0.02)]
pareto_test(abs(qretk))
qqnorm(retk,main="QQ plot")
qqline(retk,col="blue")
acf(retk,demean=FALSE,main="ACF plot")
fitk<cpt.var(retk,penalty="SIC",method="BinSeg",Q=100,test.stat="Normal",
    +class=TRUE,param.estimates=TRUE,minseglen=2)
hill(abs(qretk))
plot(fitk, main="Changepoint")
l<-retk[1:4980]
qretkpd<-l[l<quantile(l,0.05)]
pareto_test(abs(qretkpd))
qqnorm(l,main="QQ plot")
```

```
qqline(l,col="blue")
acf(l,demean=FALSE,main="ACF plot")
fitkpd<cpt.var(l,penalty="SIC",method="BinSeg",Q=100,test.stat="Normal",
    +class=TRUE,param.estimates=TRUE,minseglen=2)
hill(abs(qretkpd))
plot(fitkpd, main="Changepoint")
l<-retk[4980:length(retk)]
qretkd<-l[l<quantile(l,0.03)]
pareto_test(abs(qretkd))
qqnorm(l,main="QQ plot")
qqline(l,col="blue")
acf(l,demean=FALSE,main="ACF plot")
fitkd<cpt.var(l,penalty="SIC",method="BinSeg",Q=100,test.stat="Normal",
    +class=TRUE,param.estimates=TRUE,minseglen=2)
hill(abs(qretkd))
plot(fitkd, main="Changepoint")
par(mfrow=c(1,1))
a<-cpts(fita)
b<-rep(0,length(a)-1)
for (i in 1:length(b)) {b[i]<-a[i+1]-a[i]}
ad.test(b,"pexp",1/mean(b))
var <- VaR(reta, p = 0.95, method = "historical")
d<-reta[reta<var[,1]]
j=1
for (i in 1:length(reta)) {if(reta[i]<var[,1]){d[j]<-i
j=j+1}}
e<-rep(0,length(d)-1)
for (i in 1:length(e)){e[i]<-d[i+1]-d[i]}
ad.test(e,"pexp",1/mean(e))
acf(e,lag.max = 200)
a<-c(1,a,length(reta))
y<-rep(0,12*length(a)-12)
k<-rep(0,6)
for (i in 1:(length(a)-1)) {p<-reta[a[i]:a[i+1]]
```

```
k<-summary(p)
y[12*i-11]<-k[1]
y[12*i-10]<-k[2]
y[12*i-9]<-k[3]
y[12*i-8]<-k[4]
y[12*i-7]<-k[5]
y[12*i-6]<-k[6]
y[12*i-5]<-sd(p)
y[12*i-4]<-length(p)
y[12*i-3]<-length(p[p<var[,1]])
y[12*i-2]<-y[12*i-3]*100/y[12*i-4]
y[12*i-1]<-skewness(p)
y[12*i]<-kurtosis(p)}
mat<-matrix(y,ncol=12,byrow = TRUE)
colnames(mat)<-c("Min","1stQu","Median","Mean","3rd Qu","Max","SD","length","
q<-xicor(mat)
crash<-rep(0,length(reta))
for (j in 1:(length(reta)-100)) {log_returns <- reta[j:(99+j)]
  var<-VaR(reta[j:(99+j)],p=0.95,method="historical")
  if(reta[100+j]<var){crash[j]<-1}}
m<-rep(0,sum(crash))
j=1
for (i in 1:length(crash)) {if (crash[i]==1){m[j]=i
j=j+1}}
f<-rep(0,length(m)-1)
for (i in 1:length(f)) {f[i]<-m[i+1]-m[i]}
w<-seq(0,5,0.0001)
for (i in 1:length(w)) {y<-ad.test(f,"pexp",w)
if (y$p.value>0.05) {print(i,w)}}
acf(f,lag.max = 200)
```

4.2 Plots

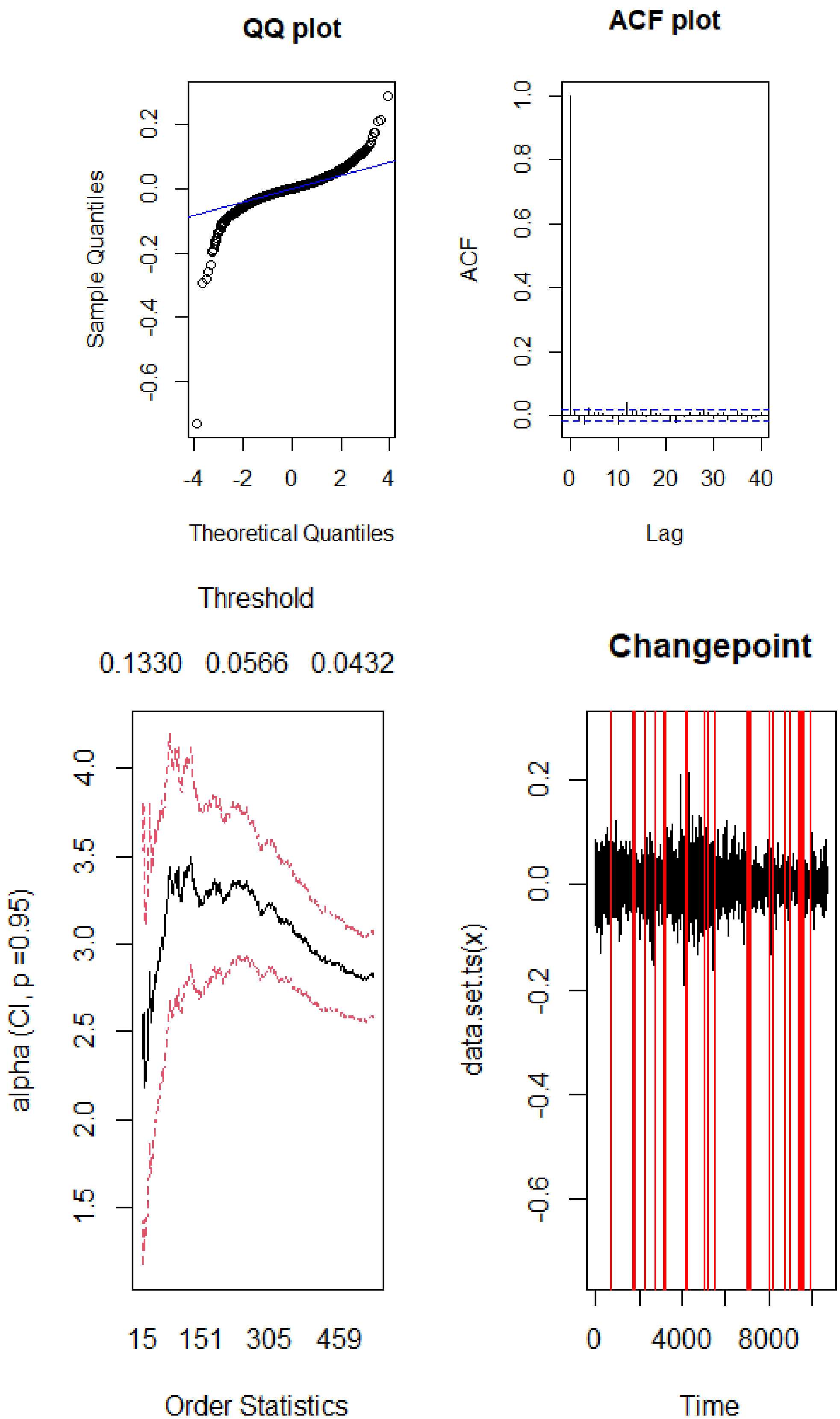


Figure 4.1: Graphs of analysis of APPLE stocks

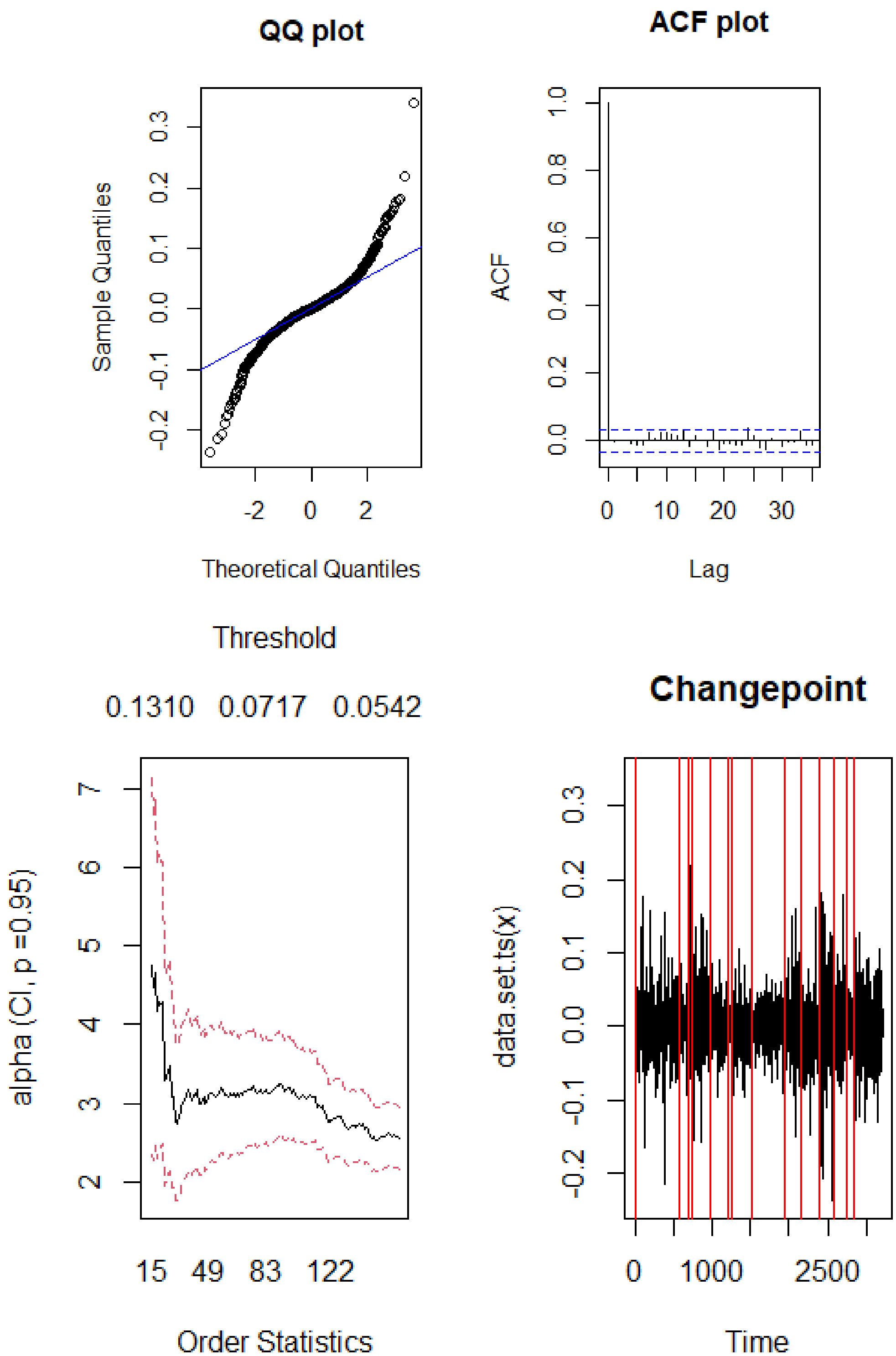


Figure 4.2: Graphs of analysis of TESLA stocks

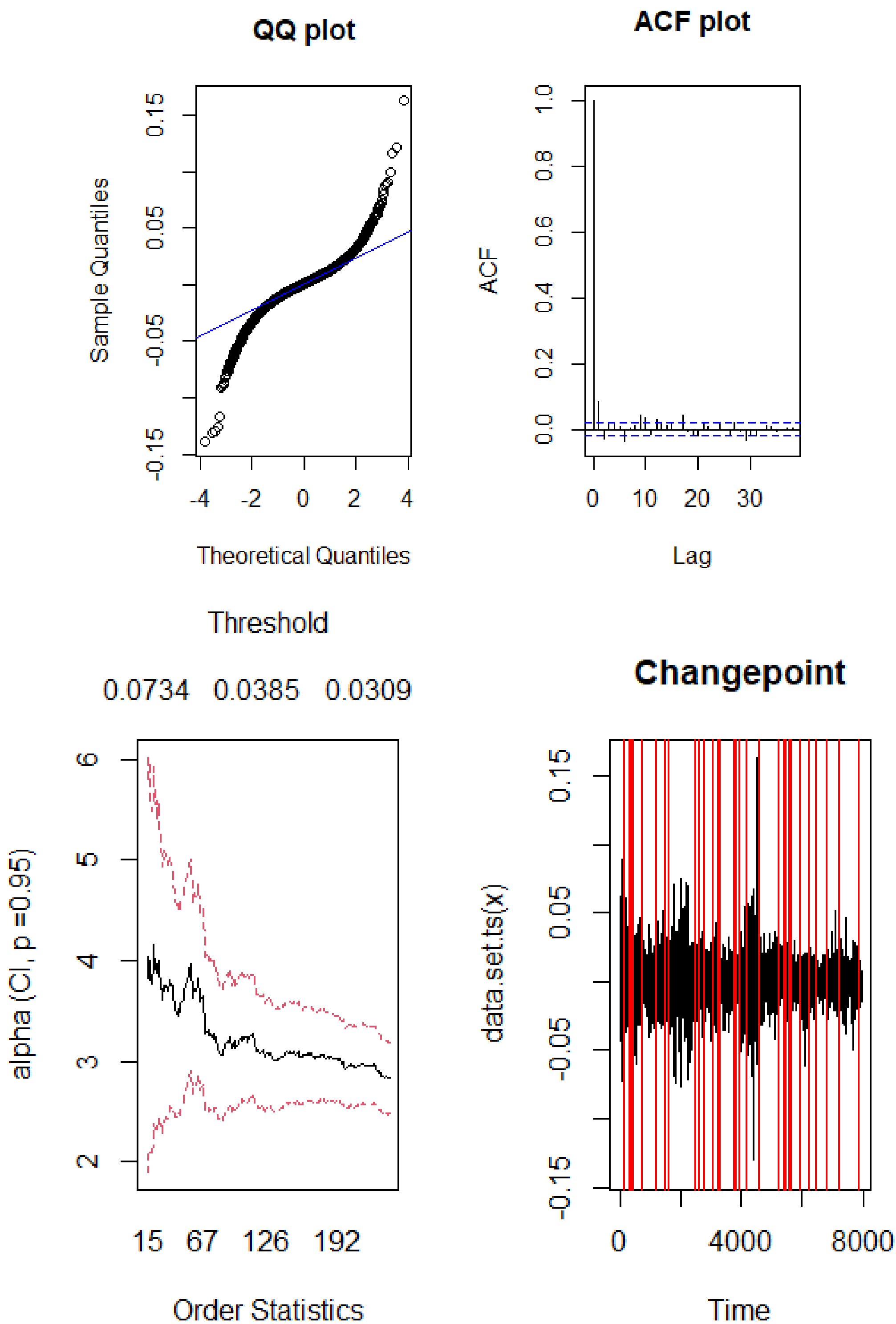


Figure 4.3: Graphs of analysis of NIFTY index

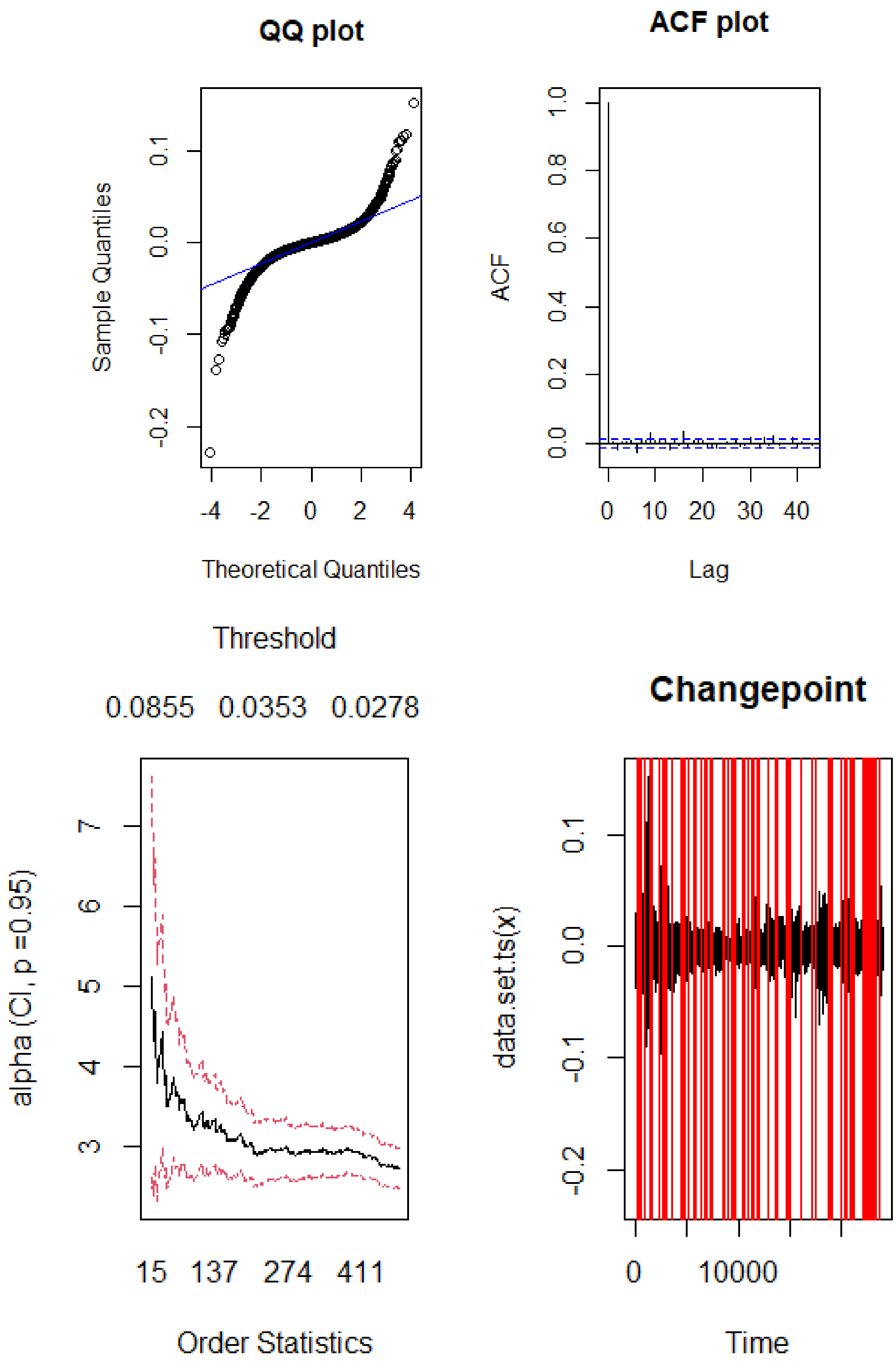


Figure 4.4: Graphs of analysis of S&P 500 index

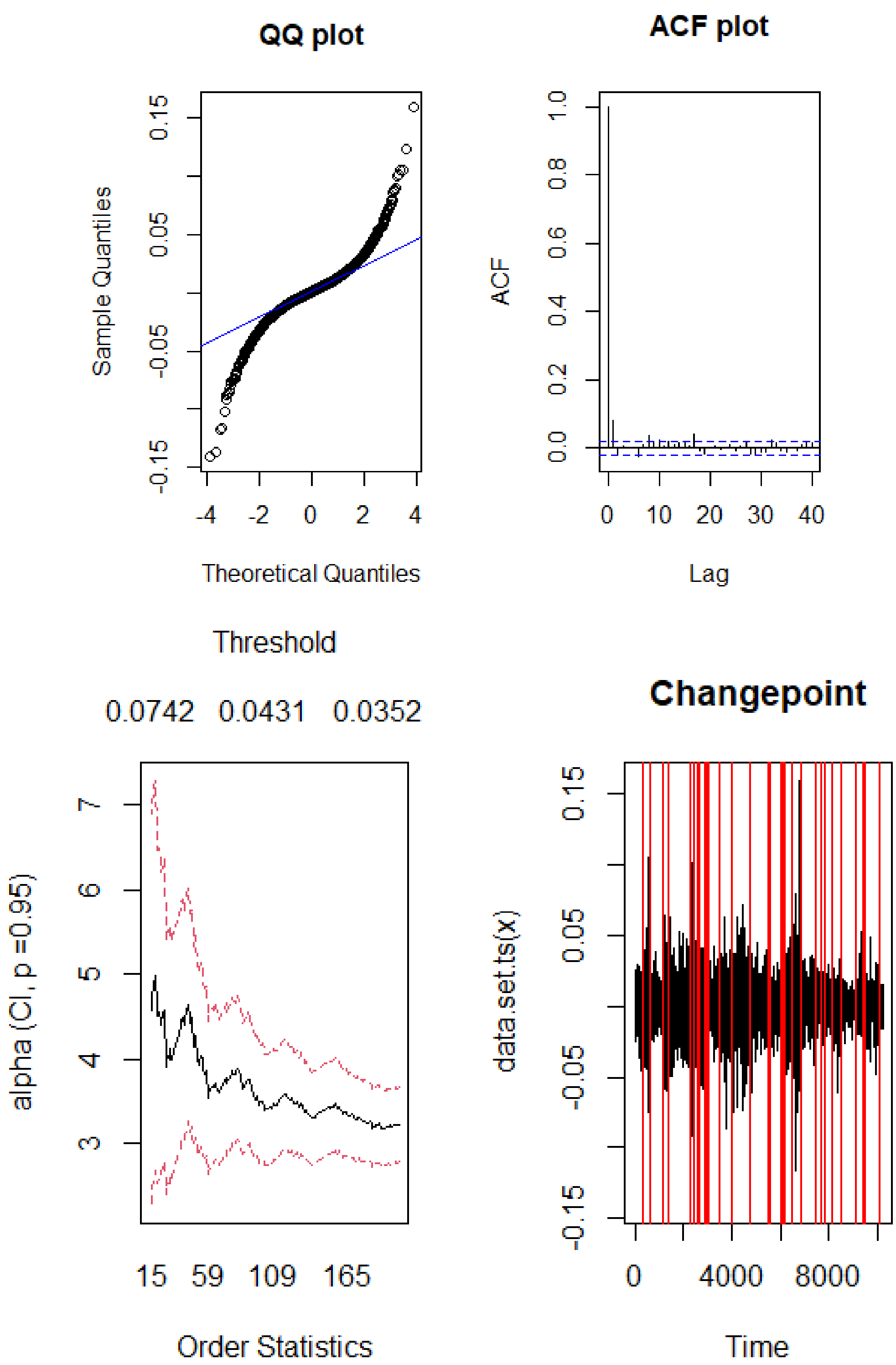


Figure 4.5: Graphs of analysis of Sensex index

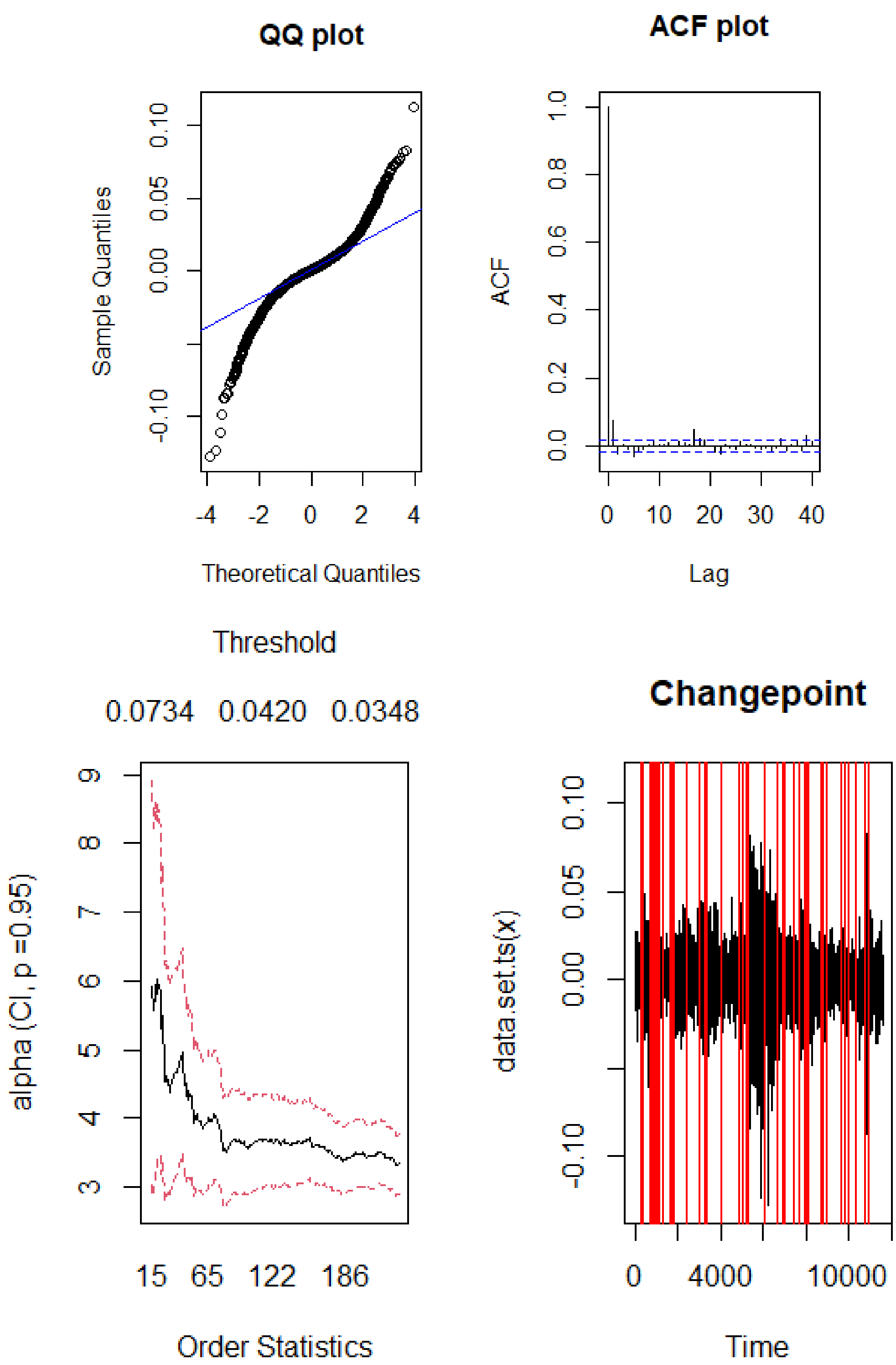


Figure 4.6: Graphs of analysis of KOSPI index

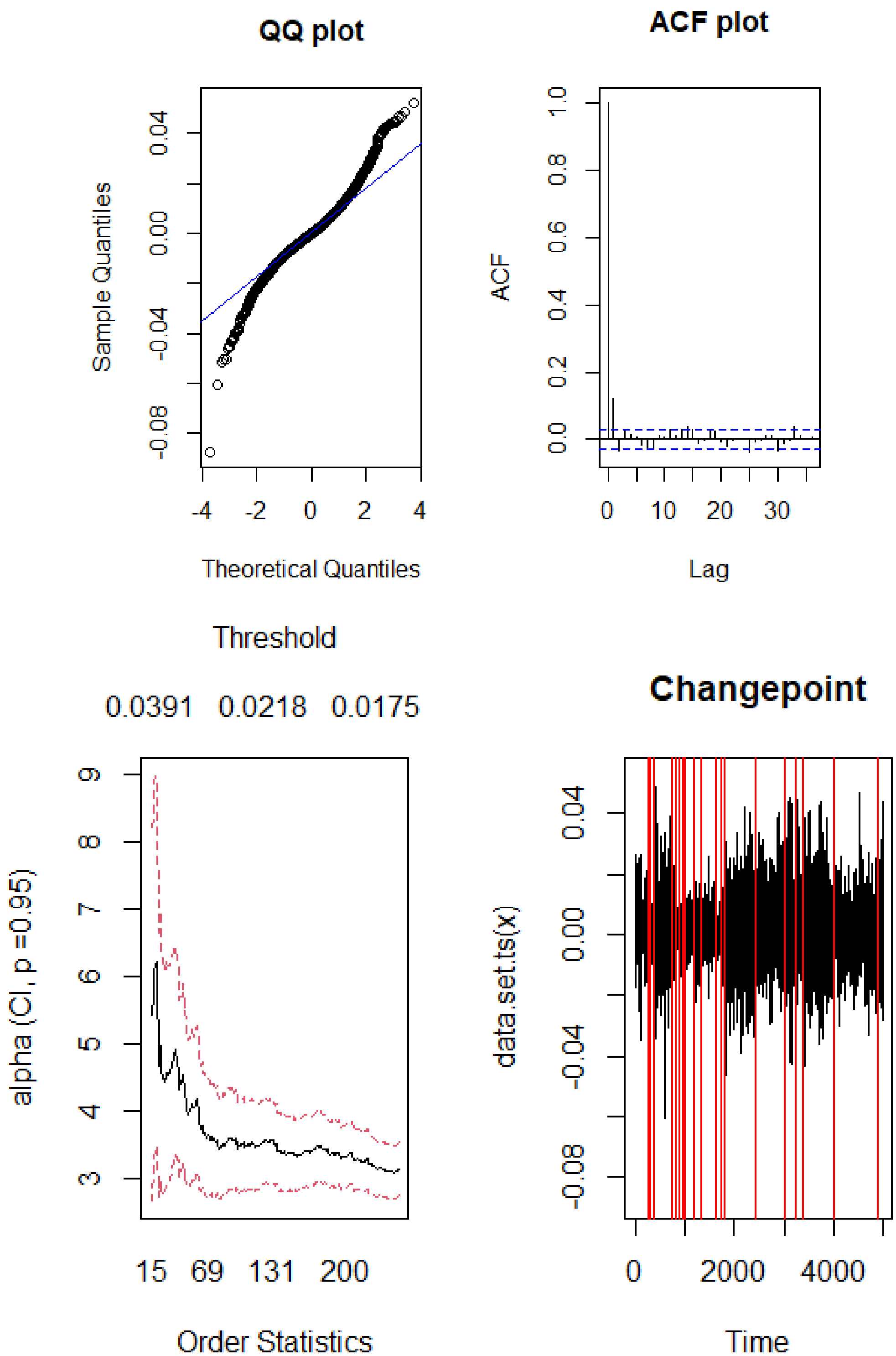


Figure 4.7: Graphs of analysis of KOSPI(developing) index

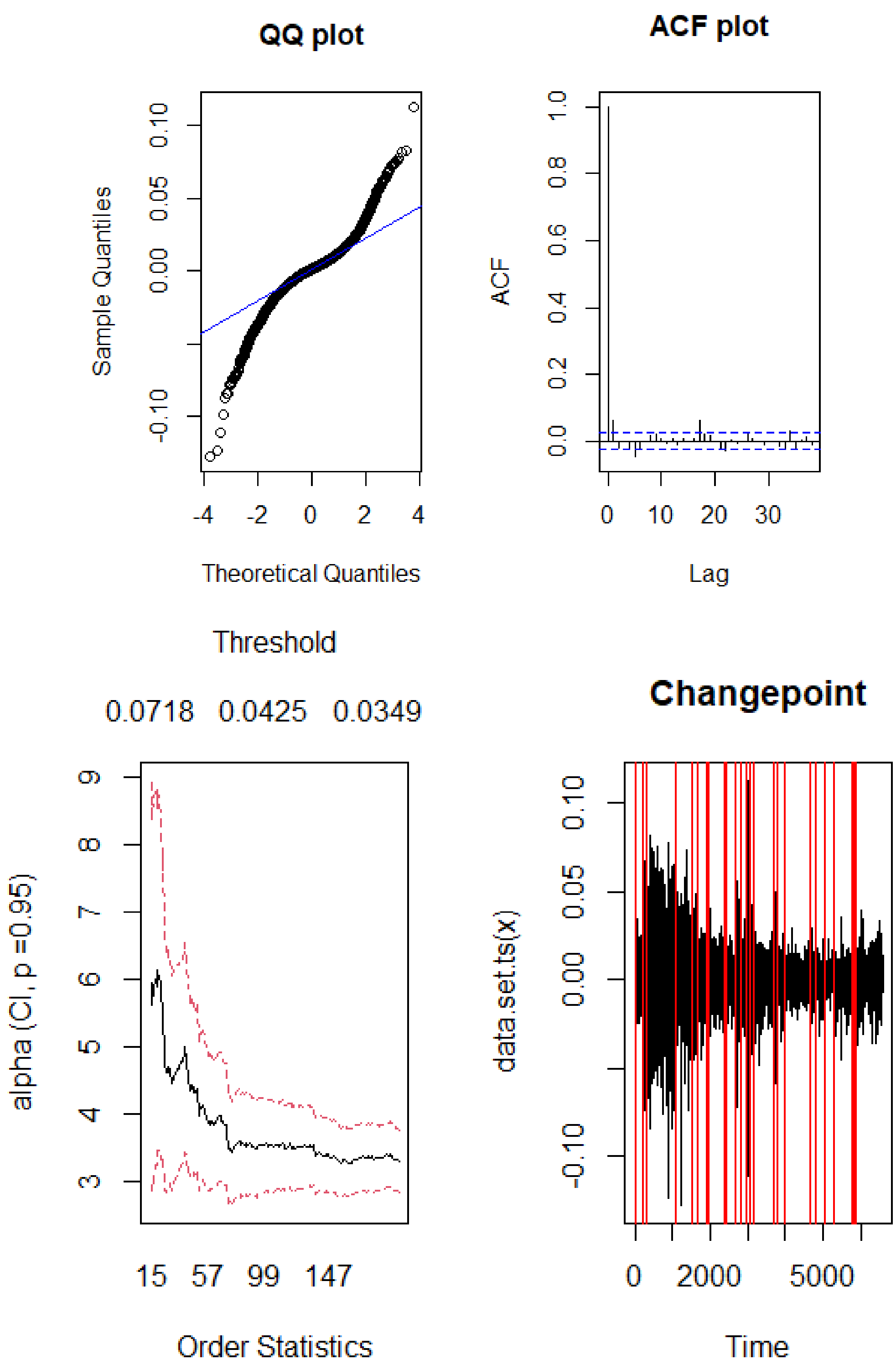
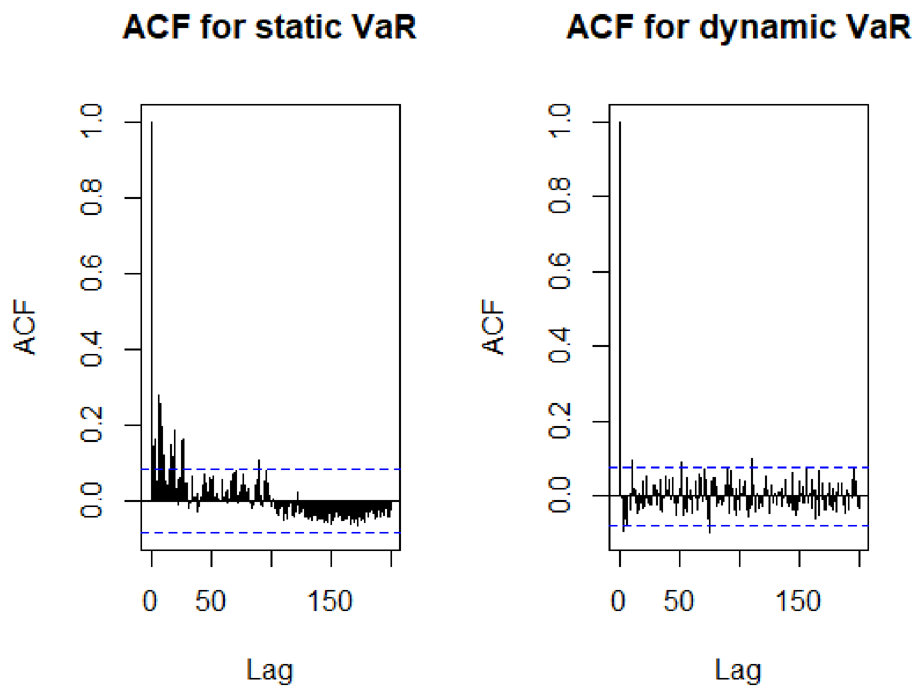


Figure 4.8: Graphs of analysis of KOSPI(developed) index



(a) ACF of Crash Inter arrival times of APPLE

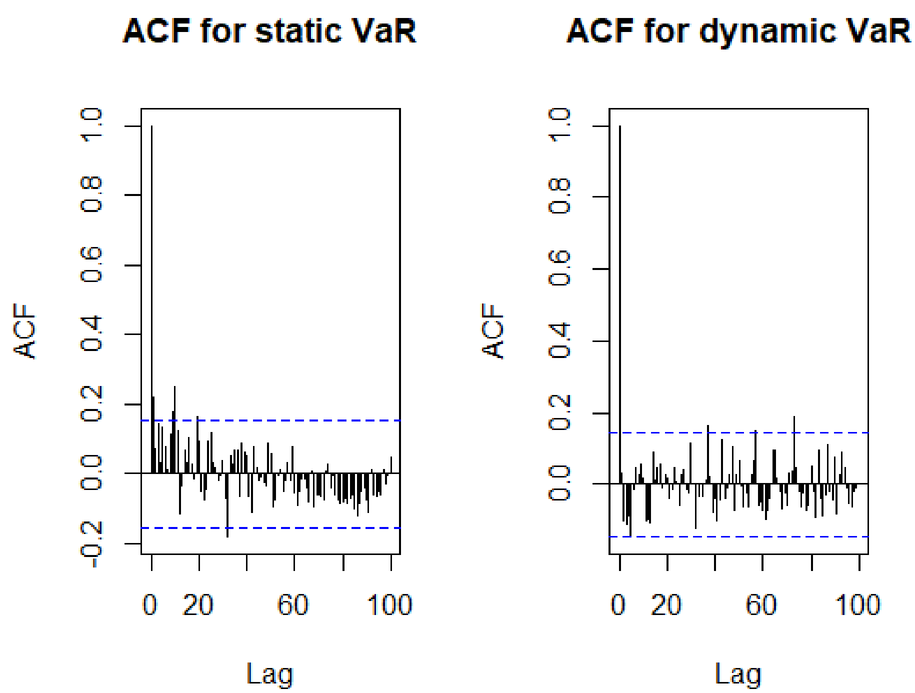
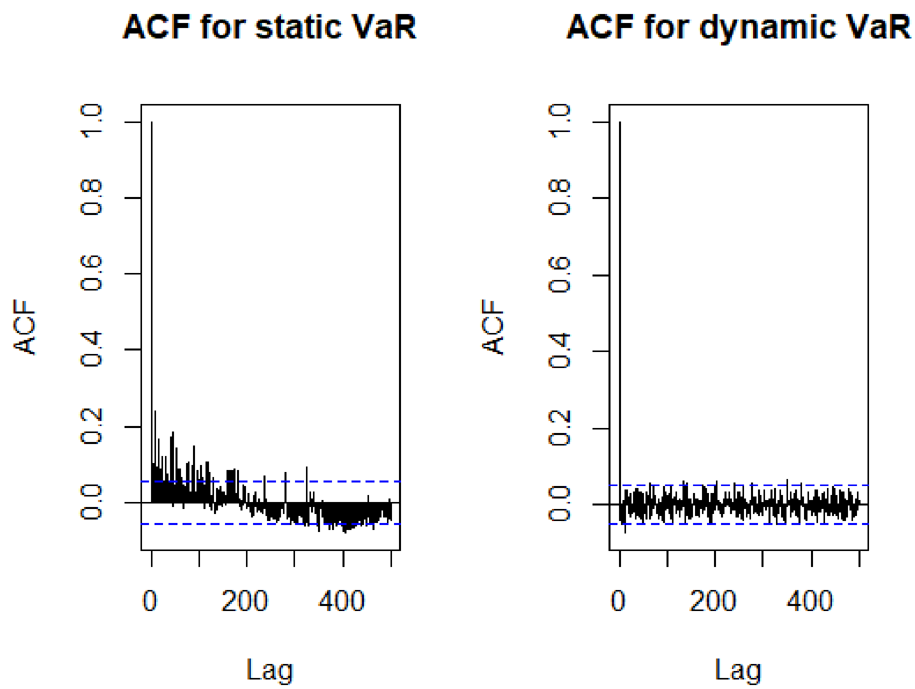


Figure 4.9: ACF of Crash Inter arrival times of TESLA



(a) ACF of Crash Inter arrival times of S&P 500

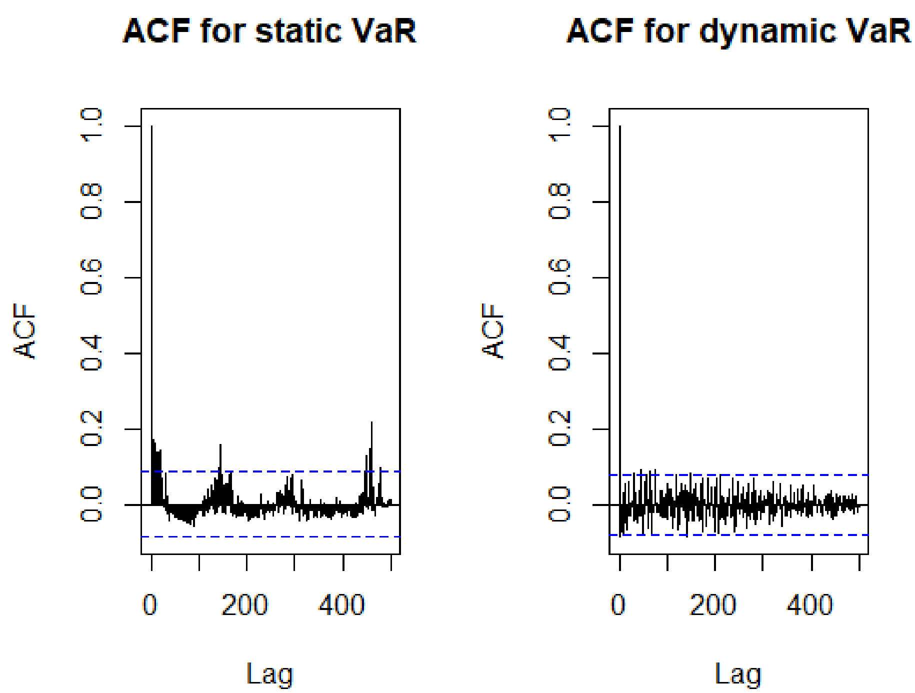
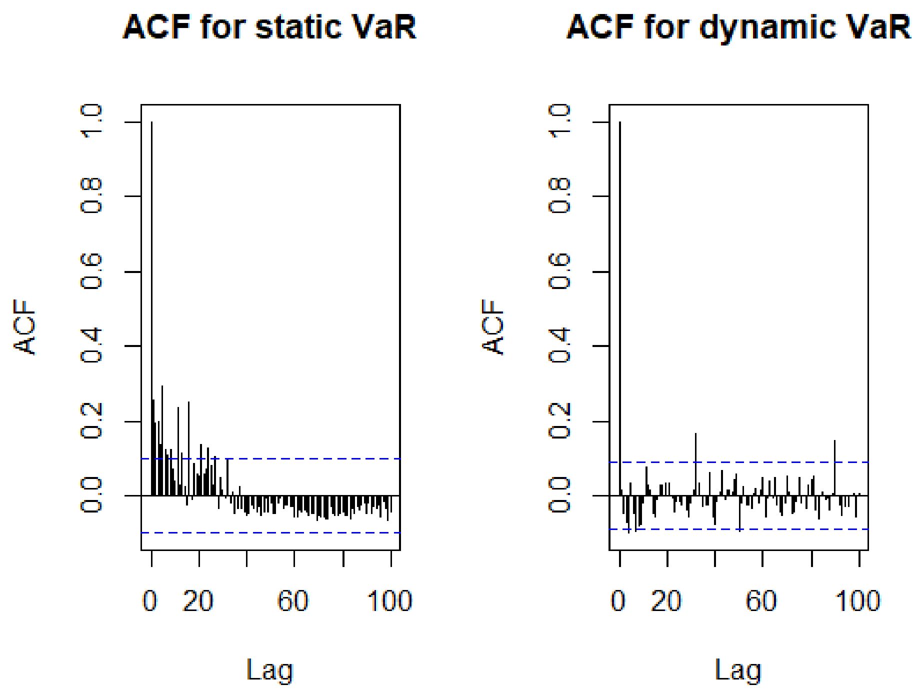


Figure 4.10: ACF of Crash Inter arrival times of Sensex



(a) ACF of Crash Inter arrival times of NIFTY

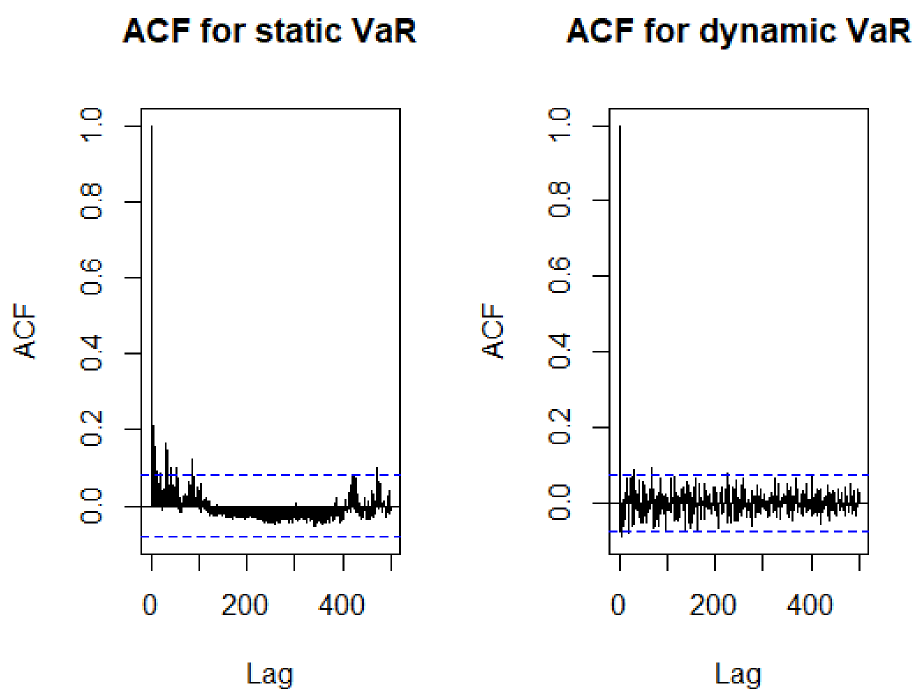
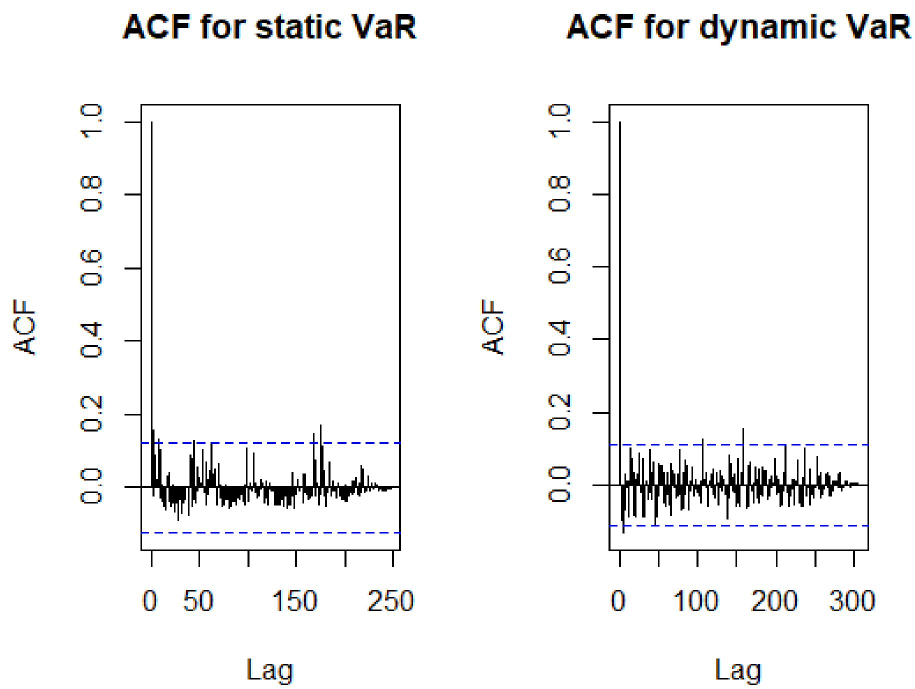


Figure 4.11: ACF of Crash Inter arrival times of KOSPI



(a) ACF of Crash Inter arrival times of KOSPI(developing)

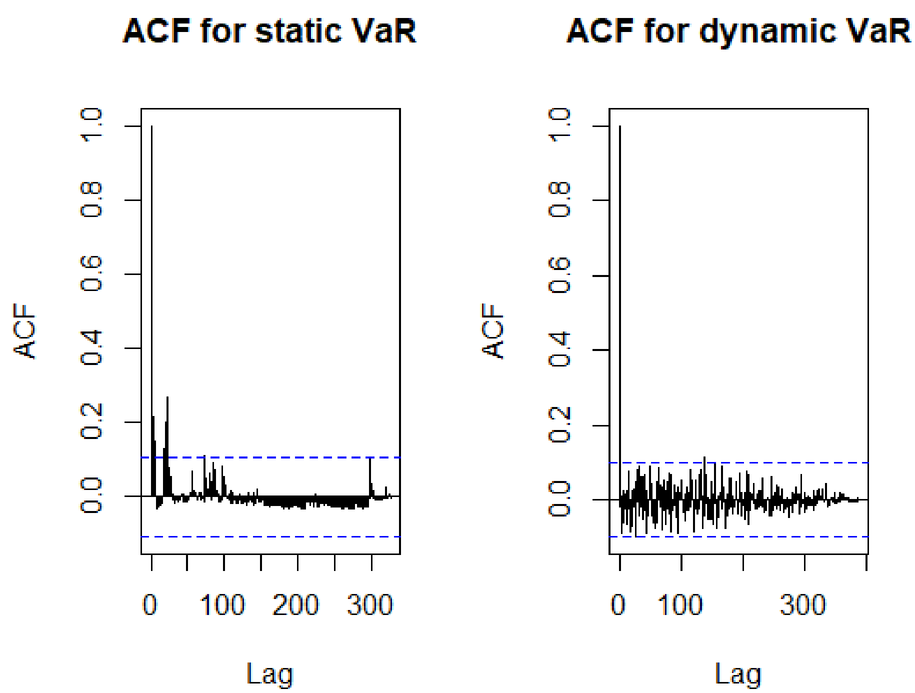


Figure 4.12: ACF of Crash Inter arrival times of KOSPI(developed)

Bibliography

[1] Robert J. Adler, Murad S. Taqqu, and Raisa E. Feldman. *A Practical Guide to Heavy Tails*. Statistical Techniques and Applications. Birkhäuser, 1998.

[2] Carol Alexander. *Value-At-Risk*, volume 4 of *Market Risk Analysis*. John Wiley & Sons, 2008.

[3] L. Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure*, 17:21–86, 1900.

[4] Treena Basu, Olaf Menzer, Joshua Ward, and Indranil SenGupt. A novel implementation of siamese type neural networks in predicting rare fluctuations in financial time series. *Risks*, 10(39), 2022.

[5] Mr. Bayes and Mr. Price. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions (1683-1775)*, 53:370–418, 1763.

[6] Fischer Black and Myron Scholes. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[7] Bihan Chatterjee. Path-dependent options in semi-markov regime switching models. Master's thesis, 2022.

[8] Sourav Chatterjee. A new coefficient of correlation. *JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION*, 116(536):2009–2022, 2021.

[9] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. *QUANTITATIVE FINANCE*, 1:223–236, 2001.

[10] Rama Cont. Long range dependence in financial markets. *Fractals in Engineering, Springer*, pages 159–179, 2005.

[11] KEVIN DOWD. *An Introduction to Market Risk Measurement*. John Wiley & Sons, 2002.

[12] Paul Embrechts and Marius Hofert. Statistics and quantitative risk management for banking and insurance. *Annual Review of Statistics and Its Application*, 1:493–514, 2014.

[13] Paul Embrechts, Alexander J. McNeil, and Rüdiger Frey. *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton Series in Finance. Princeton University Press, 2005.

[14] A.K. Gupta and Jie Chen. *Parametric Statistical change point analysis*. Birkhäuser, 2000.

[15] A.K. Gupta and Jie Chen. Testing and Locating Variance Changepoints with Application to Stock Prices. *Journal of the American Statistical Association*, 92(438):739–747, Jun. 1997.

[16] J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Process. Appl.*, 11:215–260, 1981.

[17] Alan G. Hawkes. Hawkes processes and their applications to finance: a review. *Quantitative Finance*, 18(2):193–198, 2018.

[18] John Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, 1946.

[19] Philippe Jorion. *Value-At-Risk The New Benchmark for Managing Financial Risk*. Market Risk Analysis. McGraw-Hill, 3 edition, 2007.

[20] Kauṭalya. *The Arthashastra*. Penguin Books New Delhi and New York, 1992.

[21] A. Kolmogoroff. *Basic concepts of probability theory*. Springer Verlag Berlin Heidelberg, 1933.

[22] Arnab Kumar Laha, Divyajyoti Bhowmick, and Bharathy Subramaniam. Portfolio allocation with heavy-tailed returns. *Applied Financial Economics Letters*, 3(4):237–242, 2007.

[23] Arnab Kumar Laha and K.C. Mahesh. A robust sharpe ratio. *Sankhya B:The Indian Journal of Statistics*, 83:444–465, 2021.

[24] Cecilia Mancini. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36:270–296, 2009.

[25] Harry Markowitz. Portfolio Selection. *The Journal of Finance*, 7(1):77–91, 1952.

[26] Frank J. Massey. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46(253):68–78, Mar. 1951.

[27] Ranjiva Munasinghe, Pathum Kossinna, Dovini Jayasinghe, and Dilanka Wijeratne. *ptsuite: Tail Index Estimation for Power Law Distributions*, 2019. R package version 1.0.0.

[28] Lloyd S. Nelson. The anderson-darling test for normality. *Journal of Quality Technology,*, 30(3):298–299, July. 1998.

[29] R. Wayne Oldford. Self-calibrating quantile–quantile plots. *The American Statistician*, 70(1):74–90, February 2016.

[30] Bernhard Pfaff and Alexander McNeil. *evir: Extreme Values in R*, 2018. R package version 1.7-4.

[31] Killick R and Eckley IA. changepoint: An R Package for Changepoint Analysis. *Journal of Statistical Software*, 58(3):1–19, June 2014.

[32] Killick R, Haynes K, and Eckley IA. *https://CRAN.R-project.org/package=changepoint*, 2022. R package version 2.2-4.

[33] Sidney I. Resnick. Heavy tail modeling and teletraffic data. *The Annals of Statistics*, 25(5):1805–1849, Oct. 1997.

[34] Sidney I. Resnick, Fang Xue, and Gennady Samorodnitsky. How misleading can sample ACFs of stable MAs be? (Very!). *The Annals of Applied Probability*, 9(3):797–817, August 1999.

[35] L. C. G. Rogers. Arbitrage with fractional brownian motion. *Mathematical Finance*, 7:95–105, 1997.

[36] James Sharpe and Miguel A. Juárez. Estimation of the pareto and related distributions – a reference-intrinsic approach. *Communications in Statistics-Theory and Methods*, 52(3):523–542, 2023.

[37] William F. Sharpe. The Sharpe Ratio. *Journal of Portfolio Management*, 21:49–58, 1994.

[38] William F. Sharpe. CAPITAL ASSET PRICES: A THEORY OF MARKET EQUILIBRIUM UNDER CONDITIONS OF RISK. *The Journal of Finance*, 19(3):425–442, September 1964.

[39] NASSIM NICHOLAS TALEB. *STATISTICAL CONSEQUENCES OF FAT TAILS: Real World Preasymptotics, Epistemology, and Applications Papers and Commentary.* The Technical Incerto Collection. STEM Academic Press, 2020.

[40] E. Thorp. The kelly criterion in blackjack, sports betting and the stock market. *Handbook of Asset and Liability Management*, 1:385–428, 2006.

[41] Xun Yu Zhou and Hanqing Jin. Greed, leverage, and potential losses: A prospect theory perspective. *Mathematical Finance*, 23(1):122–142, January 2013.