AN ABSTRACT OF THE DISSERTATION OF

Arpita Mukherjee for the degree of Doctor of Philosophy in Statistics presented on June 5, 2020.

Title: Limiting Behavior of Stochastic Processes Involving Martingale Structures

Abstract approved: ______________________________

Mina E. Ossiander

A *stochastic process* is given by a family of random variables indexed by elements of a set. We have considered stochastic processes of three different types, each involving an associated martingale structure. *Martingale* is a sequence of random variables for which the conditional expectation at a certain time point given the entire past is given by the present value of the sequence. Martingales possess nice theoretical properties with wide applicability. We have exploited martingale tools and techniques to derive the limiting results related to the stochastic processes. The processes we have considered are given below.

- An evolutionary urn scheme based on the *rock-paper-scissors* game
- The *random multiplicative cascade* model for intermittent processes
- A *set-indexed partial sum process* with dependent increments

The evolutionary urn scheme based on the rock-paper-scissors game is known to model species interactions in ecological systems. Therefore its limiting behavior is of interest to ecologists to understand the long term species composition of a certain ecological system. We have considered a generalization of the process to accommodate more than three species. Simulations in the general set up suggest interesting phenomena that are counter-intuitive when compared to the three-player case.

The second chapter of this thesis is motivated by data sets with variable intermittency, which makes it difficult to use standard modeling tools. It has been observed that a special class of multiplicative models, namely the Random Multiplicative Cascade models reproduce some characteristics of the data. We have derived theoretical results related to the multiplicative cascade models under a missing data set up. We have applied the method to the daily stock volume data of Tesla. Also we have proposed a change point detection method for intermittent time series. This can possibly be extended to spatial processes as well.

The last chapter of the thesis is related to a set indexed partial sum process, with martingale differences as its increments. We have derived the weak limit of the system under the *Lindeberg* type condition and the *metric entropy integrability* condition.

In spite of a common martingale structure underlying each of these three processes, they are fundamentally different. Therefore the methods to derive the limiting properties are unique to each process. For example, in the case of the rock-paper-scissors urn scheme, the key idea behind the derivation of almost sure limit is noticing a connection between sub-martingale structures within the game and a well known convergence theorem of polynomial sequence. However, for the random multiplicative cascade model, the main challenge lies in deriving asymptotic theory on a tree structure. In the third chapter, we have used probabilistic tools and techniques like *generic chaining, symmetrization,* and *truncation* to derive weak limit of the set indexed partial sum process.

Limiting Behavior of Stochastic Processes Involving Martingale Structures

by

Arpita Mukherjee

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented June 5, 2020
Commencement June 2020

Doctor of Philosophy dissertation of Arpita Mukherjee presented on June 5, 2020.

APPROVED:

Major Professor, representing Statistics

Chair of the Department of Statistics

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Arpita Mukherjee, Author

ACKNOWLEDGEMENTS

Five years back, I left an industry job in India, moved to the US and joined the PhD program at Oregon State University. The transition wasn't easy. I would like to express thanks to all those who have paved the path for me.

Thanks to my mother for her unconditional love and the sacrifices she has made. Thanks to my father for being my role model. They together with my grand-parents form a great support system.

I am extremely fortunate to have Prof. Mina Ossiander as my advisor. Thanks for her patience and sincerity in listening and answering each of my questions. I am grateful to her for the constant support and encouragement.

I am pleased to have Prof. Elaine Cozzi as my minor adviser. Taking the Functional Analysis class with her has been one of the most profound experiences of my graduate student life. I would like to thank Prof. Lisa Madsen, Prof. Sarah Emerson, and Prof. Mike Rosulek for serving on my committee. I am thankful to Prof. Debashis Mondal for partially supporting this research under NSF DMS-1519890. My heartfelt thanks to Prof. Virginia Lesser for her support. My sincere thanks to Prof. Javier Rojo for all his advice.

Thanks to my husband, Sharmodeep for always being there by my side with all his generosity and kindness. Thanks to my younger brother, Arnab for making me smile during the most difficult times. Thanks to my teachers, cousins, extended family, in-laws, and neighbors in India for their love and support.

Thanks to my dear friend Si Liu for the wonderful time we shared together learning and struggling. Thanks to Chenxiao, Jiarui, Aaron, and Connor for being amazing office-mates. Thanks to Srila, Priyadarshini, Pritha, Nilika, Laya, Debashis, and Kashyap for making me feel at home in Corvallis.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

Figure Page

LIST OF ALGORITHMS

Limiting Behavior of Stochastic Processes Involving Martingale Structures

1 Introduction

1.1 Introduction

Stochastic processes are a family of random variables indexed by a set. Each random variable of a stochastic process takes value from a common space, termed as the *state space* of the process. Depending on the nature of the *index set* and the state spaces stochastic processes can be classified into various categories. We consider stochastic processes from three different classes and study their limiting properties.

The study of limiting behavior of stochastic processes is of interest for many reasons. Stochastic processes are often used to model the dynamics of a system or naturally occurring physical phenomena. In that context in order to understand the underlying process better it is crucial to answering all or some of the following questions:

- Does the system converge?
- If yes, where does it converge?
- If no, are there conditions under which the system converges?
- How fast does the system converge?

- Which factors contribute to determining the speed of convergence?
- What is the distributional limit of the process? i.e. is there a non-degenerate distribution to which the distribution converges?

Each of these questions motivates the study of the long-term behavior of a process. Also, various important test statistic can be represented as a supremum (or infimum) of a stochastic process.

Example 1: Consider the Kolmogorov-Smirnov test that is used for testing the equality of continuous distribution, i.e. for $X_1, \ldots, X_n \overset{iid}{\sim} F$ it tests the following hypothesis:

$$H_0 : F = F_0 \text{ vs } H_A : F \neq F_0$$

where F_0 is a fully specified cumulative distribution function. The test statistic in this case is given by

$$T_{\text{stat}} = \sqrt{n} \sup_x |\hat{F}_n(x) - F_0(x)|$$

where $\hat{F}_n(x)$ is the sample analogue of cumulative distribution function, also termed as the *empirical distribution function*. Thus the test statistic can be viewed as the supremum of a stochastic process given by $\{U_n(x) = \sqrt{n}\left(\hat{F}_n(x) - F_0(x)\right) : x \in \mathbb{R}\}$. In order to derive the null asymptotic distribution of T_{stat}, it is important to study the limiting distribution of the process $\{U_n(x)\}$.

Example 2: Another area of application includes bootstrap methodology. The bootstrap method is a universal tool for obtaining the distribution of any statistics. Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} P$ with the empirical distribution given by $P_n = \frac{1}{n}\sum_{i=1}^n \delta_{X_i}$.

Suppose we are interested in the distribution of the statistic given by $T(P_n)$, which is a functional of P_n. Depending on the form of P, it might get too complicated to derive the null distribution of such statistic. Thus bootstrap is used to derive the null distribution in such cases. The idea behind bootstrap lies in the fact that as P_n and P are close for sufficiently large n, therefore if we get a sample $X_1^*, \ldots, X_{m_n}^* \stackrel{iid}{\sim} P_n$, then conditional on P_n, $T(P_{m_n}^*)$ should approximate $T(P_n)$, where

$$P_{m_n}^* = \frac{1}{m_n} \sum_{i=1}^{m_n} \delta_{X_i^*}$$

Now in order to show that the bootstrap distribution converges weakly to the actual distribution of the test statistic, the scaled difference between the two is viewed as a stochastic process. Hence in order to justify bootstrap approximation, it is important to study the limiting distribution of the corresponding stochastic process. Further details on this can be found in Barbe and Bertail (2012).

Now that we have explained the motivation behind the study of stochastic processes and their limiting behavior, we shall briefly describe each of the stochastic processes that we have considered, followed by the commonality between them.

1.2 The rock-paper-scissors urn

The first process considered here is an evolutionary urn scheme based on the well-known rock-paper-scissors game. The counts of rock, paper, and scissors in the urn change according to a set of rules. The rules are governed by the basic principles of the rock-paper-scissors game, in which each object enjoys equal power but when it comes to pairwise comparison one object is always more powerful than the other. This is a discrete-time stochastic process with the state space $\mathbb{N}^3$, where $\mathbb{N}$ denotes the set of all natural numbers. The current state of the process, given the entire past, depends only on the recent past. Thus the process forms a vector-valued Markov chain, with a time-inhomogeneous transition matrix, i.e. the transition probabilities are dependent on the specific time point.

An ecological system is known to retain biodiversity only if the system is non-hierarchical. The inherent nature of the rock-paper-scissors game makes it non-hierarchical. Thus this game is natural to consider for modeling species interaction in ecological systems. In spite of broad usage, the mathematical foundation related to the long-term behavior of such a system has been unresolved. This thesis gives a rigorous mathematical framework for such a system and provides a derivation of limiting properties.

We have identified several martingale structures within the game and then combined those with algebraic theorems related to polynomial convergence to come up with a novel technique to prove almost sure limit convergence for such systems. A central limit theorem has also been derived in this context.

We have also considered a generalization of the game in a k-player situation. We have done extensive simulation studies to understand the dynamics of such complicated systems and have come up with an algorithm to identify the limiting proportions. However, the proof of convergence still remains unaddressed. There seem to be nice mathematical structures in the three-player case that disappear in more complex setups.

1.3 The random multiplicative cascades

The second chapter of the thesis is about the random multiplicative cascade models; these are used to model naturally occurring physical phenomena that are intermittent in nature. The model is based upon the fact that the total amount of a certain object (say, rainfall) divides randomly into smaller scales. These random splits are determined by the distribution of a set of independent and identically distributed random variables, known as cascade generating random variables. This splitting mechanism gives rise to a tree structure, each level of which can be thought of as a time point. Thus the entire process can be thought of as a stochastic process, where the level of granularity at every step is higher than the previous step, i.e. over the time the splitting process goes into finer and finer scale. Rainfall, turbulence, internet traffic, etc displays such cascading characteristics.

The real data is considered as the fine-scale limit of the cascade measure at a particular resolution. In order to simulate realizations from the system, it is essential to estimate the cascade generating distribution from the data. A crucial component of a multiplicative cascade model is the structure-function, which uniquely determines the cascade generating distribution. The goal is to come up with a sample statistic that converges to the structure-function of the cascade generating random variable almost surely.

There is a good volume of work dedicated to the estimation of the structure-function, which determines the cascade distribution. All of these works assume that the complete data is available, i.e. observations on an entire grid at a particular resolution is observed. We have considered the case where there are observations that are missing, which happens quite commonly in real datasets. We have proposed an estimator of structure-function that can handle missing values. We have also derived almost sure convergence of these estimators and a central limit theorem in this context. Here the key challenge is that the asymptotics is on a tree structure.

We have demonstrated the use of a random multiplicative cascade model on the stock volume data of Tesla. We have also demonstrated the missing data handling technique by applying this model to the data containing missing observations. The daily stock volume data over time seems to have changed patterns after a certain time point. We have modeled the data before and after that point separately and have proposed a naive approach for change point detection in intermittent time series. The method can be significantly improved by deriving a uniform limit theorem for the estimated structure-function, which still is an open problem in this field.

1.4 A set indexed partial sum process

In the third chapter of the thesis we have considered a set indexed partial sum process, which is defined on a 'suitable' collection of sets $\mathcal{A}$, defined as follows.

$$S_n(A) := \frac{1}{n}\sum_{i=1}^{n} X_i \mathbf{1}(V_i \in A) \text{ for } A \in \mathcal{A}$$

Here we have considered X_i's to be $\mathcal{F}_i$-measurable martingale differences and V_i's to be $\mathcal{F}_{i-1}$-measurable. We have derived a central limit theorem under a *Lindeberg* type condition. The central theorem holds uniformly on $\mathcal{A}$. In this case, the limiting process is a *set indexed Gaussian process.* The proof of weak convergence of the process relies on *finite-dimensional convergence* and *tightness*, which is implied by *asymptotic equicontinuity.*

Let's consider a motivating example to understand why this type of partial sum process is of interest to the statisticians. As can be seen from example 1, the Kolmogorov Smirnov statistics for testing equality of continuous distributions is given by

$$G_n = \sup_x |\sqrt{n}(\hat{F}_n(x) - F_0(x))|$$

The distribution of G_n is independent of F_0 when X_i's take values in $\mathbb{R}$. However, when X_i's are random vectors, the distribution of the test statistics is not distribution-free anymore. Depending on the form of F_0 it might get quite difficult to derive the null distribution of G_n in such cases. Thus one easy way to go about it is to use the bootstrap method. In a *generalized bootstrap* method we approximate the empirical distribution function $\hat{F}_n(x)$ by $\hat{F}_{W,n}(x)$ given as follows:

$$\hat{F}_{W,n}(x) = \sum_{i=1}^{n} W_i \mathbf{1}(X_i \leq x)$$

where W corresponds to the weights of Bootstrap. In the standard bootstrap procedure $(W_1, \ldots, W_n) \sim$ Multinomial $\left(1, \frac{1}{n}, \frac{1}{n} \ldots, \frac{1}{n}\right)$. Now in order to prove that the bootstrap distribution approximates the actual distribution we need to derive the weak limit of $\hat{F}_{W,n}(x)$. It is easy to note that $\hat{F}_{W,n}(x)$ can be viewed as a set-indexed partial sum process, where the indexing set is of the form $(-\infty, x]$. In our case, we have allowed dependence between and within the random variables X_i's and V_i's. Also, we have considered a more general collection of indexing sets. We have illustrated a possible application of our process in testing hypotheses related to multiplicative error models of non-negative time series.

1.5 Martingale: The unifying theme

All of the processes considered so far involve martingales, which is a common probabilistic structure to model dependence. A *martingale* is a sequence of random variables for which the conditional expectation of the present given the entire past is the same as the value of the sequence in the recent past. Let $\{X_n : n \geq 1\}$ be a martingale sequence and $\mathcal{F}_n$ the history of the process up to time n. Thus for each $n \geq 1$, X_n must satisfy

$$\mathbb{E}(X_n | \mathcal{F}_{n-1}) = X_{n-1}$$

$\{X_n : n \geq 1\}$ is said to be a sub(super)-martingale if

$$\mathbb{E}(X_n | \mathcal{F}_{n-1}) \lesseqgtr X_{n-1}$$

Convergence results such as the strong law and the central limit theorem that holds for a sequence of independent identically distributed random variables also have analogous versions for martingale sequences of random variables. Details on these can be found in Hall and Heyde (2014).

For the rock paper scissors urn problem, we have identified martingale and sub-martingale structures within the game. Almost sure convergence of these martingales plays a crucial role in the derivation of the almost sure convergence of the long term proportions of each individual object type, namely rock, paper, and scissors. Also, the related central limit theorem is primarily based on the martingale central limit theorem.

While constructing the random multiplicative cascade measure, the cascade generating random variables are chosen to be random variables with mean one. Thus the Radon-Nikodym derivative of the cascade measure with respect to the Lebesgue measure forms a martingale sequence. The convergence results related to the random multiplicative cascade process rely on the convergence of this martingale.

In the set-indexed partial sum process, the increments are considered to be martingale differences. Here the primary goal was to work with a model that allowed dependence between the increments; martingales seem to be the most natural choice. Established results then allow the derivation of the weak and strong convergence results related to the process.

Though each of the stochastic processes described above shares an underlying martingale structure, they are fundamentally different. In the subsequent chapters, we shall discuss each of the processes in detail, derive related results, present interesting findings, and mention possible directions for future work.

2 Limiting behavior of the rock-paper-scissors urn

2.1 Abstract

We consider a randomized urn scheme based on a stochastic version of the well-known rock-paper-scissors game. The cyclic and non-hierarchical nature of the scheme in conjunction with its simplicity yields an attractive model of competitive population dynamics with applications in scientific fields ranging from ecology and biology to physics and economics. The rock-paper-scissors game has a non-hierarchical and non-transitive structure that mimics, for example, strategies that help to maintain biodiversity in the long term in ecological settings. In this chapter, our key focus is to study the limiting behavior of the stochastic urn model and to derive a central limit theorem in this context. Techniques from both polynomial and martingale theory are used. We also explore several extensions of the game and present simulation studies suggesting future directions for research.

2.2 Introduction

In the well-known rock-paper-scissors game, rock crushes scissors, scissors cuts paper, and paper wraps rock. This means the rock is more powerful than scissors, scissors is more powerful than paper, and paper is more powerful than rock. Thus each of the three objects is more powerful than one of the other two and less powerful than the remaining object leading to no clear winner in the game. This unique cyclic and non-transitive nature resembles several naturally occurring scientific and physical phenomena, and is used to model or describe these scientific systems. This simple model thus provides a simple and flexible setting in which to study the dynamics of cyclic systems.

In ecology, non-hierarchical systems are believed to play a pivotal role in maintaining biodiversity. In other words, a system does not appear to retain its species richness in the long term unless it is non-hierarchical. Rock-paper-scissors is among the simplest model of a non-hierarchical system with applications in modeling species diversity; see for example Kerr et al. (2002), Sinervo and Lively (1996), Shi et al. (2010), Kirkup and Riley (2004), Frean and Abraham (2001), Mobilia (2010).

Structural properties of rock-paper-scissors games have also been of interest to game-theorists with analysis of theoretical and experimental aspects appearing in Semmann et al. (2003), Cook et al. (2011), Xu et al. (2013) and Cason et al. (2013).

Rock-paper-scissors games also appear in physics as models of interacting particle systems. In Tainaka (2000) the standard rock-paper-scissors is extended to a two-dimensional lattice structure with applications to the voter and biological systems. In Venkat and Pleimling (2010), Jiang et al. (2012), Hua et al. (2013), Peltomäki and Alava (2008), the structural properties of the rock-paper-scissors game is used to explore the Spatio-temporal dynamics of various physical systems.

The Lotka-Volterra model (Lotka (1926), Volterra (1927)) is widely used to analyze population systems. In Mao et al. (2003), Avelino et al. (2012), Bahar and Mao (2004), Zhu and Yin (2009b), Li and Mao (2009), Liu and Fan (2017), Zhu and Yin (2009a) several aspects of stochastic Lotka-Volterra system have been studied. Stochastic version of the game have intriguing properties not present in the deterministic counterpart. In Mao et al. (2003) and Dobrinevski and Frey (2012) the authors have shown how the asymptotic properties of a stochastic Lotka-Volterra model differ from its deterministic version. In Liu and Wang (2014), Du and Sam (2006), Cattiaux and Méléard (2010) different versions of stochastic Lotka-Volterra systems have been considered. The study of rock-paper-scissors games provides insights into such systems.

In this chapter, we consider a stochastic version of the rock-paper-scissors game by defining a competitive urn scheme. The urn is initially populated with one object of each type, namely rock, paper, and scissors. Over time, the urn evolves as follows: at each step, two objects of different types are drawn at random and then returned to the urn together with an additional copy of the more powerful object. Thus at each step, the urn size goes up exactly by one. This is a generalization of the classical Polya's urn model, as seen in Durrett (2019). We have studied the limiting properties of the urn as time increases. In Pemantle et al. (2007), Lasmar et al. (2018) and Laslier et al. (2017) various stochastic urn schemes have been considered. However, none accommodates the competitive aspects of the rock-paper-scissors urn considered here.

Contributions of our work include the following. (1) We show rigorously that the rock-paper-scissors urn remains almost surely balanced in the long term and derive related central limit theorems. These results rely on a novel combination of martingale methods with algebraic theorems related to polynomial convergence. (2)We also consider a rock-paper-scissors urn with random step-sizes, i.e. a Poisson number of additional balls are added to the urn at each time step. The limiting properties of the Poissonized games correspond to those with deterministic step sizes. (3)We propose an algorithm giving the limiting behavior of competitive urn models with more than three object types and present related simulations. Interestingly, the limiting properties of these generalized competitive urns do not always correspond to that of the standard rock-paper-scissors model.

In Section 2.3 we describe the problem, introducing required notation. This is followed by the derivation of our central results for the standard rock-paper-scissors urn scheme in Section 2.4. Section 2.6 presents extension of the model to urns containing 4 distinct objects. Simulation studies are included to illustrate limiting properties. In Section 2.7 a generalized urn with k distinct object types is considered. We present an algorithm that identifies the limiting proportions of individual object types. Finally, in Section 2.8, we return to the standard urn with three object types and introduce random population growth. At the end of each section, we present simulation studies for illustration. Section 2.9 contains some concluding remarks and suggestions for further research.

2.3 Set up and Notation

Let $n \geq n_0$ denotes the step number and $1 \leq i \leq k$ the object type. Clearly, for a standard rock-paper-scissors game, there are three distinct object types in the urn, i.e. $k = 3$. The number of objects of type i at step n is denoted by $X_{n,i}$ with corresponding vector $\mathbf{X}_n = (X_{n,1}, \ldots, X_{n,k})^T$. Let $\mathcal{F}_n$ be the σ-algebra generated by $\{\mathbf{X}_{n_0}, \ldots, \mathbf{X}_n\}$; i.e. the history of the game up to time n. The game is initialized at $n_0 = k$ with one object of each type in the urn, giving $\sum_{i=1}^k X_{n_0,i} = k$. At each step the total number of objects in the urn increases exactly by 1 so $\sum_{i=1}^k X_{n,i} = n$ for all $n \geq n_0$. Let $P_{n,i}$ denotes the conditional probability that $\mathbf{X}_{n+1} = \mathbf{X}_n + e_i$ given $\mathbf{X}_n$. Here the e_i's are the standard unit basis vectors. The relationship between the $P_{n,i}$'s and the $X_{n,i}$'s depends on the specific game structure. For example, in the standard three-type rock-paper-scissors game, the conditional probabilities are given by $P_{n,i} = p_i(\mathbf{X}_n)$ where

$$\begin{aligned}
p_1(\mathbf{x}) &= \frac{x_1 x_2}{x_1 x_2 + x_1 x_3 + x_3 x_2} \\
p_2(\mathbf{x}) &= \frac{x_2 x_3}{x_1 x_2 + x_1 x_3 + x_3 x_2} \\
p_3(\mathbf{x}) &= \frac{x_1 x_3}{x_1 x_2 + x_1 x_3 + x_3 x_2}.
\end{aligned}$$

Let $\mathbf{P}_n$ denote the vector $\{P_{n,i} : 1 \leq i \leq k\}$.

2.4 Standard rock-paper-scissors game

The urn-scheme based on the rock-paper-scissors game begins with an urn containing one rock, one paper, and one scissors. At each time point, two objects of different types are drawn at random. Then they are returned to the urn together with an additional copy of the more powerful object. In investigating the limiting behavior of the urn we obtain the almost sure limits of the proportions of rocks, papers, and scissors as the fixed point of a system of linear equations. In order to model the dependence between subsequent steps, we identify key sub-martingales. Properties of these will be exploited in combination with results from polynomial theory to derive the almost sure limit of the system. Key findings describing limiting properties of this urn scheme are summarized in Theorem 2.4.1 and Theorem 2.4.12.

Theorem 2.4.1 (Convergence of proportions)**.** *In the rock-paper-scissors urn process, the long term proportions of each individual object type converges to* $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ *almost surely; that is* $P\left(\lim_n \frac{\mathbf{X}_n}{n} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)\right) = 1$.

The complete proof of Theorem 2.4.1 is deferred until the end of the section. The following preliminary result provides an outline of how martingale methods and polynomial theory are used in tandem.

Theorem 2.4.2. $\frac{\mathbf{X_n}}{n}$ *converges almost surely.*

The proof relies on the following classical theorem which appears as proposition (5.2.1) of Artin (2010).

Theorem 2.4.3 (Convergence of roots of a real polynomial)**.** *Let $q_k(t)$ be a sequence of monic polynomials of degree $\leq m$, and let $q(t)$ be another monic polynomial of degree m. Let $\alpha_{k,1}, \dots \alpha_{k,m}$ and $\alpha_1, \alpha_2, \dots \alpha_m$ be the roots of these polynomials. If $q_k \to q$, the roots $\alpha_{k,v}$ of q_k can be numbered in such a way that $\alpha_{k,v} \to \alpha_v$ for each $v = 1, 2, \dots m$.*

Proof of Theorem 2.4.2. Define the functions t and u on $\mathbf{R}^3$ as $t(x) = x_1x_2 + x_1x_3 + x_2x_3$ and $u(x) = x_1x_2x_3$. Let $T_n = t\left(\frac{\mathbf{X}_n}{n}\right)$ and $U_n = u\left(\frac{\mathbf{X}_n}{n}\right)$. In the following section we use martingale methods to show in Lemmas 2.4.4 and 2.4.8 respectively that T_n and U_n converge almost surely to bounded r.v.'s. Thus the polynomial Q_n on $\mathbf{R}$ defined via $Q_n(r) = r^3 - r^2 + T_nr - U_n$ converges pointwise. It is easy to check that the roots of Q_n are given by the components of $\frac{\mathbf{X}_n}{n}$. From Theorem 2.4.3 we see that the almost sure convergence of Q_n implies the almost sure convergence of the proportions $\frac{\mathbf{X}_n}{n}$ of different objects in the urn. □

2.4.1 Sub-martingale convergence

In this section, some sub-martingale structures are derived that figure in the proof of Theorem 2.4.2.

Lemma 2.4.4. *$\widetilde{U}_n = (n(n+1)(n+2))^{-1} \prod_{i=1}^{3} X_{n,i}$ is a bounded sub-martingale and converges almost surely.*

Since U_n and $\widetilde{U}_n$ are asymptotically equivalent, an immediate consequence is the following.

Corollary 2.4.5. *U_n converges almost surely.*

The following algebraic result is used in the proof of Lemma 2.4.4.

Lemma 2.4.6. *If $x_i \geq 1$ for $i = 1, 2, 3$ with $\sum x_i = n$, then $\frac{2n-3}{n^2} \leq \frac{t(\mathbf{x})}{n^2} \leq \frac{1}{3}$.*

Proof. The upper bound is verified as follows. The usual Cauchy-Schwarz inequality gives $t(\mathbf{x}) \leq \sum x_i^2$. Then $\frac{t(\mathbf{x})}{(x_1+x_2+x_3)^2} = \frac{t(\mathbf{x})}{x_1^2+x_2^2+x_3^2+2t(\mathbf{x})} \leq \frac{t(\mathbf{x})}{3t(\mathbf{x})} = \frac{1}{3}$. It is straightforward to verify the lower bound via calculus. □

Returning to the previous lemma we have the following.

Proof of Lemma 2.4.4. Recall that $\sum_i X_{n,i} = n$.

$$\begin{aligned} E(\widetilde{U}_{n+1}|\mathbf{X}_n = \mathbf{x}) &= \frac{\prod_{i=1}^3 x_i + x_2x_3p_1(x) + x_1x_3p_2(\mathbf{x}) + x_1x_2p_3(\mathbf{x})}{(n+1)(n+2)(n+3)} \\ &= \frac{\prod_{i=1}^3 x_i(1 + n/t(\mathbf{x})}{(n+1)(n+2)(n+3)} \\ &\geq \frac{\prod_{i=1}^3 x_i(1 + 3/n)}{(n+1)(n+2)(n+3)} \\ &= \widetilde{U}_n. \end{aligned}$$

Thus $E(\widetilde{U}_{n+1}|\mathcal{F}_n) \geq \widetilde{U}_n$ and $\{\widetilde{U}_n : n \geq 1\}$ is a non-negative sub-martingale with upper bound

$$\widetilde{U}_n \leq (n^3)^{-1} \prod_{i=1}^{3} X_{n,i} \leq \frac{1}{3^3}. \tag{2.1}$$

Hence $\widetilde{U}_n$ is almost surely convergent. □

For $n \geq 3$ let $V_n = \frac{\prod_{i=1}^3 X_{n,i}}{(n+3)(n+4)\, t^{1/2}(\mathbf{X}_n)}$.

Lemma 2.4.7. *$\{V_n : n \geq 3\}$ is a bounded sub-martingale and converges almost surely.*

Proof. For $\sum_i x_i = n$ and e_i the usual unit basis vector, $t(x+e_i) = t(x) + n - x_i$, so for $\beta < 0$

$$\begin{aligned} t^\beta(x+e_i) - t^\beta(x) &= \beta \int_{y=0}^{1} (n-x_i)(t(x) + (n-x_i)y)^{\beta-1} dy \\ &\geq \beta(n-x_i)t^{\beta-1}(x) \end{aligned}$$

In particular,

$$\begin{aligned} \sum_{i\geq 1}(x_i+1)x_{i+1}t^{-1/2}(x+e_i) &\geq \sum_{i\geq 1}(x_i+1)x_{i+1}(t^{-1/2}(x) - (n-x_i)t^{-3/2}(x)/2) \\ &= \left(t^{-1/2}(x) - \frac{n}{2}t^{-3/2}(x)\right)(t(x)+n) \\ &+ \frac{1}{2}t^{-3/2}(x)\sum_i x_i(x_i+1)x_{i+1} \\ &\geq \left(t^{-1/2}(x) - \frac{n}{2}t^{-3/2}(x)\right)(t(x)+n) + \frac{1}{2}t^{-3/2}(x)(t(x)+t^2(x)/n) \\ &= t^{1/2}(x)\left(1+\frac{n}{t(x)}\right)\left(\frac{2n+1}{2n} - \frac{n}{2t(x)}\right) \end{aligned}$$

Above we have used the Cauchy-Schwarz inequality to derive the bound

$$\sum_{i\geq 1} x_i^2 x_{i+1} \geq \frac{t^2(\mathbf{x})}{n}$$

Notice that for $\sum_i x_i = n$ with $x_i \geq 1$, the function $\left(1+\frac{n}{t(x)}\right)\left(\frac{2n+1}{2n} - \frac{n}{2t(x)}\right)$ is minimized over the range of $n/t(x)$ at the minimal value of $n/t(x)$; $n/t(x) = 3/n$.

That is, $\left(1+\frac{n}{t(x)}\right)\left(\frac{2n+1}{2n}-\frac{n}{2t(x)}\right) \geq (n+3)(n-1)/n^2$. Thus

$$\begin{aligned} E(\prod_{i=1}^{3} X_{n+1,i}\, t^{-1/2}(\mathbf{X}_{n+1})|\mathbf{X}_n = x) &= (\prod_{1\leq j\leq 3} x_j)t^{-1}(x)\sum_i (x_i+1)x_{i+1}t^{-1/2}(x+e_i) \\ &\geq (\prod_{1\leq j\leq 3} x_j)t^{-1/2}(x)(n-1)(n+3)/n^2 \end{aligned}$$

and, after verifying that $\frac{(n+3)^2(n-1)}{n^2(n+5)} \geq 1$, we see that $E(V_{n+1}|\mathbf{X}_n) \geq \frac{(n-1)(n+3)^2}{n^2(n+5)}V_n \geq V_n$. Consequently $\{V_n : n \geq 1\}$ is a sub-martingale. Observing that $3\prod_{i=1}^{3} x_i \leq nt(\mathbf{x})$ we see that V_n is bounded and thus converges almost surely. □

Lemma 2.4.8. *$T_n = \frac{t(\mathbf{X_n})}{n^2}$ converges almost surely to a finite r.v. T.*

Proof. After noting that $T_n^{1/2} = \frac{(n+1)(n+2)\widetilde{U}_n}{(n+3)(n+4)V_n}$ this follows from the almost sure convergence of $\widetilde{U}_n$ and V_n. □

2.4.2 Martingale methods

The distribution of individual object types $\mathbf{X}_n$ at each step n depends only on the immediate past $\mathbf{X}_{n-1}$ rather the entire past history $\mathcal{F}_n$. Thus $\{\mathbf{X}_n : n \geq 1\}$ is a vector valued Markov chain with corresponding vector-valued martingale

$$\mathbf{M}_n = \mathbf{X}_n - \sum_{l=n_0+1}^{n} P_{l-1}. \tag{2.2}$$

Lemma 2.4.9. *$\{\mathbf{M}_n : n \geq n_0\}$ is a vector-valued martingale with respect to the filtration $\{\mathcal{F}_n : n \geq n_0\}$.*

Proof. Since

$$\mathbb{E}(\mathbf{X}_n - \mathbf{X}_{n-1}|\mathcal{F}_{n-1}) = \mathbf{P}_{n-1} \tag{2.3}$$

therefore $\mathbb{E}(\mathbf{M}_n|\mathcal{F}_{n-1}) = \mathbf{M}_{n-1}$ Hence $\{\mathbf{M}_n\}$ is a Martingale. □

The quadratic variation of the individual components of the martingale $\{\mathbf{M}_n : n \geq 1\}$ is given by

$$\begin{aligned}\sum_{l=n_0+1}^{n} \mathbb{E}[(M_{l,i} - M_{l-1,i})^2|\mathcal{F}_{l-1}] &= \sum_{l=n_0+1}^{n} \mathbb{E}[(\delta_{l,i} - P_{l-1,i})^2|\mathcal{F}_{l-1}] \\ &= \sum_{l=n_0+1}^{n} P_{l-1,i}(1 - P_{l-1,i})\end{aligned}$$

where $\delta_{l,i} = X_{l,i} - X_{l-1,i}$.

Lemma 2.4.10. *For every object type denoted by* i, $\frac{1}{n}\left(X_{n,i} - \sum_{l=n_0+1}^{n} P_{l-1,i}\right) \xrightarrow{a.s.} 0$

Proof. The term on left hand side can be rewritten as $\frac{1}{n}\sum_{l=n_0+1}^{n}(\delta_{l,i} - P_{l-1,i})$ where $\delta_{l,i}$ is one or zero depending on whether or not the object of type i increased at step l. Note that $\delta_{l,i} \sim \text{Bernoulli}(P_{l-1,i})$ Fix i and denote $S_n = \sum_{l=n_0+1}^{n}(\delta_{l,i} - P_{l-1,i})$. For fixed $\epsilon > 0$ and $N \geq 1$, Kolmogorov's inequality gives the following:

$$\begin{aligned}\mathbb{P}(\max_{n_0+1 \leq n \leq 2^N} |S_n| > 2^N\epsilon) &\leq \frac{1}{2^{2N}\epsilon^2}\sum_{l=n_0+1}^{2N} \mathbb{E}(\delta_{l,i}^2) = \frac{1}{2^{2^N}\epsilon^2}\sum_{l=n_0+1}^{2^N} EP_{l-1,i} \\ &\leq \frac{1}{2^{2N}\epsilon^2}2^N = \frac{1}{2^N\epsilon^2}\end{aligned}$$

By the Borel-Cantelli lemma it follows that $\frac{S_n}{n} \xrightarrow{a.s.} 0$. Hence $\frac{X_{n,i} - \sum_{l=n_0+1}^{n} P_{l-1,i}}{n} \xrightarrow{a.s.}$

0

□

2.4.3 Identifying the limiting proportions

It remains to identify the almost sure limit of $\mathbf{X_n}/n$ given in Theorem 2.4.1.

Lemma 2.4.11. $\mathbf{\Pi} = \lim_n \mathbf{X}_n/n$ *satisfies*

$$\mathbf{\Pi} = \left(\frac{\pi_1\pi_2}{t(\mathbf{\Pi})}, \frac{\pi_2\pi_3}{t(\mathbf{\Pi})}, \frac{\pi_1\pi_3}{t(\mathbf{\Pi})}\right) \tag{2.4}$$

Proof. Since $\mathbf{X}_n/n$ converges almost surely, the vector $\mathbf{P}_n$, which is scale invariant and depends continuously on $\mathbf{X}_n$ also converges almost surely. From Lemma 2.4.10 we see that $\mathbf{\Pi} = \lim_n \frac{\mathbf{X}_n}{n} = \lim \frac{\sum_{n_0}^{n-1} \mathbf{P}_l}{n} = \lim_n \frac{\mathbf{P}_n}{n} = \left(\frac{\pi_1\pi_2}{t(\mathbf{\Pi})}, \frac{\pi_2\pi_3}{t(\mathbf{\Pi})}, \frac{\pi_1\pi_3}{t(\mathbf{\Pi})}\right)$. □

Proof of Theorem 2.4.1. The only possible solution of this system of equations (2.4) is given by $\mathbf{\Pi} = (\pi_1, \pi_2, \pi_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. By using Theorem 2.4.3 and Lemma 2.4.11, we can conclude that in a rock-paper-scissors game, the proportions of each type of object converges almost surely to $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. □

2.4.3.1 A central limit theorem for the limiting proportions

The central limit theorem for the limiting proportions follows from the martingale central limit theorem in Billingsley (2008).

Theorem 2.4.12 (A central limit theorem for rock-paper-scissors urn process)**.** *For any $i \neq j$ with $i, j \in \{1, 2, 3\}$ we have*

$$\sqrt{n}\left(\left(\frac{X_{n,i}}{n}, \frac{X_{n,j}}{n}\right)^T - \left(\frac{1}{3}, \frac{1}{3}\right)^T\right) \xrightarrow{d} N_2(\mathbf{0}, \Sigma) \tag{2.5}$$

where $N_2(0, \Sigma)$ denotes a 2-dimensional Gaussian random vector with mean zero and covariance matrix $\Sigma = \begin{bmatrix} \frac{2}{9} & -\frac{1}{9} \\ -\frac{1}{9} & \frac{2}{9} \end{bmatrix}$.

Proof. For $i \in \{1, 2, 3\}$, let $\Lambda_i = \{\omega : \sum_{l \geq n_0} P_{l,i}(1 - P_{l,i}) = \infty\}$. Since the $P_{n,i}$'s converge almost surely to $1/3$ from Lemma 2.4.11, $P(\cap_{i=1}^3 \Lambda_i) = 1$. The martingale difference sequence $\mathbf{M}_n - \mathbf{M}_{n-1} = \delta_n - \mathbf{P}_{n-1}$, has uniformly bounded components. Now note that

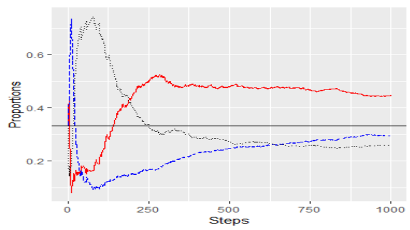
$$(\delta_{n,1}, \delta_{n,2}, \delta_{n,3} | \mathcal{F}_{n-1}) \sim \text{Multinomial}(1, P_{n,1}, P_{n,2}, P_{n,3}) \tag{2.6}$$

with the marginal distributions given by

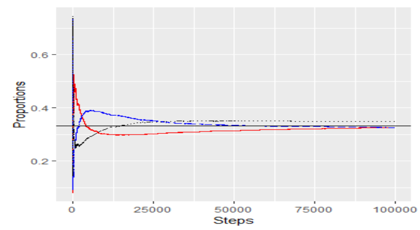
$$(\delta_{n,i} | \mathcal{F}_{n-1}) \sim \text{Binomial}(1, P_{n,i}) \text{ for all } i \in \{1, 2, 3\} \tag{2.7}$$

and

$$\begin{aligned} \text{Cov}\left(\delta_{n,i}, \delta_{n,j} | \mathcal{F}_{n-1}\right) &= -P_{n,i}P_{n,j} \text{ for any } i \neq j \\ \text{Var}(\delta_{n,i} | \mathcal{F}_{n-1}) &= P_{n,i}(1 - P_{n,i}) \text{ for all } i. \end{aligned}$$



(a) stepsize=1000, initial counts= $(1, 1, 1)$



(b) stepsize=100000, initial counts= $(1, 1, 1)$

Figure 2.1: Convergence of a standard rock-paper-scissors urn

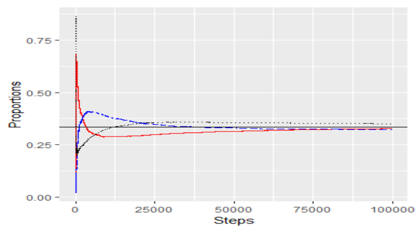
Clearly conditional on $\mathcal{F}_{n-1}$, $P_{n-1,i}$'s are constants, therefore

$$\begin{aligned} \mathrm{Cov}\left(M_{n,i} - M_{n-1,i}, M_{n,j} - M_{n-1,j} | \mathcal{F}_{n-1}\right) &= -P_{n,i}P_{n,j} \\ \mathrm{Var}(M_{n,i} - M_{n-1,i} | \mathcal{F}_{n-1}) &= P_{n,i}(1 - P_{n,i}) \text{ for all } i \end{aligned}$$

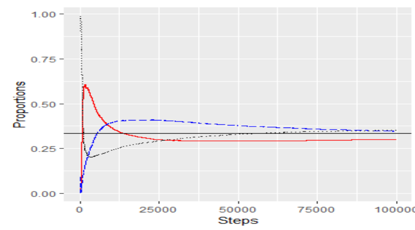
Since as $n \to \infty$ we have $P_{n,i} \xrightarrow{a.s.} \frac{1}{3}$ for all i, therefore the result follows from Billingsley (2008). □

2.5 Simulation studies

Figure 2.1 represents the evolution of a standard rock-paper-scissors urn over time. Object types are denoted by different colors. As shown previously, all three object types in this game are exchangeable. Thus we have not attached labels to trajectories of object proportions. The figure on the left includes 1000 steps and the figure on the right includes 10000 steps. The convergence of long term proportions to $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ is evident from both of the diagrams.



(a) stepsize=100000, initial counts= $(1, 10, 10)$



(b) stepsize=100000, initial counts= $(1, 10, 100)$

Figure 2.2: Proportions over time for different starting counts

Figure 2.2 shows the proportions of objects over time for different initial counts. Here we have considered two cases. For the first case, initial counts of two object types are the same and the other one is different. For the second case, the initial counts of all three objects are very different. As we can see, the long term proportions are independent of the initial counts. However, the speed of convergence seems to depend on the initial composition of the urn. The more similar initial counts are, the faster is the convergence to limiting proportions.

2.6 A four player extension

2.6.1 Introduction

Consider stochastic urn containing four distinct object types with cyclic and non-transitive game structure. In this case, the underlying game is not unique. Also, the limiting properties of the games are quite different. For example unlike the standard urn process involving three object types, simulation suggests that in this case one of the object types eventually dies out as time goes off to infinity.

For a game involving 4 object types, denote the 4 different types of objects by A, B, C, D. Unlike the rock-paper-scissors game, this game won't be unique because, for 4 types of objects, there are total $\binom{4}{2} = 6$ possible pairwise comparisons. If the cyclic structure of the game is kept fixed, i.e. $A > B > C > D > A$ then depending on whether $(B \lesseqgtr D)$ and $(A \lesseqgtr C)$, we shall end up having $2.2 = 4$ possible distinct games. We look at the limiting behavior of each such game. The object type that eventually dies out appears to be determined by the structure of the game.

2.6.2 Identifying the limiting Proportions

If the proportions of individual object types converge then depending on the structure of the game their limit can be identified using the following lemmas.

Lemma 2.6.1. *If $A > B > C > D > A; A > C; B > D$ and $\frac{\mathbf{X}_n}{n}$ converges almost surely then C eventually dies out.*

Proof. Suppose $\frac{\mathbf{X}_n}{n} \xrightarrow{a.s.} \mathbf{\Pi}$. In that case $\mathbf{P}_n \xrightarrow{a.s.} \mathbf{\Pi}$ as well. Then

$$\pi_A = \frac{\pi_A\pi_B + \pi_A\pi_C}{t(\mathbf{\Pi})} \quad ; \quad \pi_B = \frac{\pi_C\pi_B + \pi_B\pi_D}{t(\mathbf{\Pi})}$$
$$\pi_C = \frac{\pi_D\pi_C}{t(\mathbf{\Pi})} \quad ; \quad \pi_D = \frac{\pi_A\pi_D}{t(\mathbf{\Pi})}$$

where $t(\mathbf{\Pi}) = \pi_{\mathbf{A}}\pi_{\mathbf{B}} + \pi_{\mathbf{A}}\pi_{\mathbf{C}} + \pi_{\mathbf{A}}\pi_{\mathbf{D}} + \pi_{\mathbf{B}}\pi_{\mathbf{C}} + \pi_{\mathbf{B}}\pi_{\mathbf{D}} + \pi_{\mathbf{C}}\pi_{\mathbf{D}}$.

$$\begin{aligned}
\pi_A &= \frac{\pi_A\pi_B + \pi_A\pi_C}{t(\mathbf{\Pi})} &\implies& \quad t(\mathbf{\Pi}) = \pi_{\mathbf{B}} + \pi_{\mathbf{C}} \\
\pi_B &= \frac{\pi_C\pi_B + \pi_B\pi_D}{t(\mathbf{\Pi})} &\implies& \quad t(\mathbf{\Pi}) = \pi_{\mathbf{D}} + \pi_{\mathbf{C}} \\
\pi_C &= \frac{\pi_D\pi_C}{t(\mathbf{\Pi})} &\implies& \quad t(\mathbf{\Pi}) = \pi_{\mathbf{D}} \\
\pi_D &= \frac{\pi_A\pi_D}{t(\mathbf{\Pi})} &\implies& \quad t(\mathbf{\Pi}) = \pi_{\mathbf{A}}
\end{aligned}$$

Therefore $\pi_B = \pi_D = \pi_A = t(\mathbf{\Pi}); \pi_{\mathbf{C}} = \mathbf{0}$. Hence in this case the limiting proportions will be given by $(\frac{1}{3}, \frac{1}{3}, 0, \frac{1}{3})$. □

Lemma 2.6.2. *If $A > B > C > D > A; A < C; B < D$, then A eventually dies out.*

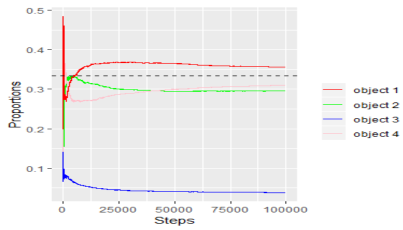
Lemma 2.6.3. *If $A > B > C > D > A; A > C; B < D$, then B eventually dies out.*

Lemma 2.6.4. *If $A > B > C > D > A; A < C; B > D$, then D eventually dies out.*

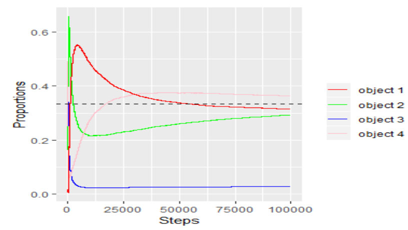
Remark 1. *The proofs of the Lemma 2.6.2, 2.6.3, 2.6.4 will be similar to that of Lemma 2.6.1.*

2.6.3 Simulation studies

Now we shall illustrate some of the interesting features of the cyclic urn processes involving four object types.



(a) stepsize=1000, initial counts= $(1, 1, 1, 1)$



(b) stepsize=100000, initial counts= $(1, 10, 20, 30)$

Figure 2.3: Proportions over time for different step sizes

Figure 2.3 shows the proportions of distinct object types over time for different starting counts. Note that proportion of one of the object types converges to zero and the rest converge to $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. The set up of Lemma 2.6.1 was considered to simulate the observations. Note that just as standard rock paper scissors game, in this case too the long term proportions are independent of the limiting counts. However, the speed of convergence seems to depend on the initial composition of the urn.

2.7 A generalization to the k-player case

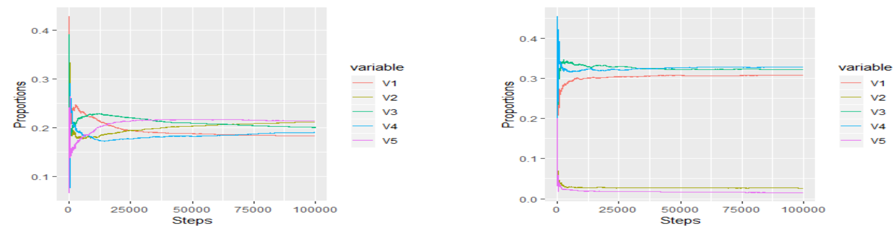
2.7.1 Set up and notation

In this section we shall consider a generalization of the rock-paper-scissors urn scheme to an urn, involving k distinct types of objects. The cyclic structure of the game is maintained by assuming the following order of power. object1 $>$ object2 $>$ object3 $> \ldots$ object$k >$ object1. For the remaining $\left(\binom{k}{2} - k\right)$ pairwise comparisons, there are 2 choices for the order of power between each pair. Hence altogether $2^{\binom{k}{2}-k} = 2^{\frac{k(k-3)}{2}}$ possible games are there. In this section we shall illustrate how the limiting behavior of the game changes depending on structure of the game. Let I_i denote the set of all object types that are less powerful than object i. Clearly if $j \in I_i$, then $i \notin I_j$. For $n \geq 1$ and $i = 1, 2, \ldots, k$, the probability of choosing the i-th object type at n-th step is given by $P_{n,i} = \frac{X_{n,i} \sum_{j \in I_i} X_{n,j}}{t(\mathbf{X}_n)}$ where $t(\mathbf{X}_n) = \sum_{1 \leq i < j \leq k} X_{n,i} X_{n,j} = \sum_{i=1}^{k} X_{n,i} \sum_{j \in I_i} X_j$.

2.7.2 Identifying the limiting proportions

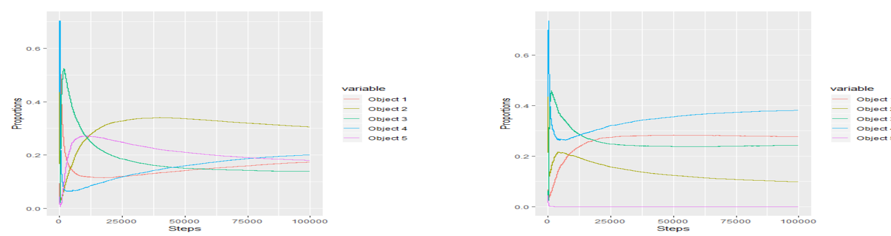
If for $i \in \{1, 2, 3, \ldots, k\}$, $P_{n,i}$ converges to $\mathbf{\Pi}_i$ almost surely we must have $\pi_i = \frac{\pi_i S_i(\mathbf{\Pi})}{t(\mathbf{\Pi})}$ where $S_i(\mathbf{\Pi}) = \sum_{j \in I_i} \pi_{i,j}$. This implies $t(\mathbf{\Pi}) = S_i(\mathbf{\Pi})$ for all i The limiting proportions can be identified by solving this system of linear equations. We propose the following algorithm 2.7.1 to calculate the limiting proportions if all the orderings in pairwise comparisons are known. The solution to this linear equation gives the limiting proportions of each object type.

Remark 2. *Unlike the games involving* 3 *or* 4 *objects types, it appears that the game with* $k (k \geq 5)$ *types of objects can result in totally different limiting proportions depending on the pairwise orderings between the distinct object types. We shall illustrate such differences through the simulation studies.*



(a) Same limiting proportions (b) Different limiting proportion

Figure 2.4: Game involving 5 object types, with same initial counts



(a) Same limiting proportions (b) Different limiting proportion

Figure 2.5: Game involving 5 object types, with different initial counts

2.7.3 Simulation study

Figure 2.4 shows the convergence of proportions of individual object types in case of a cyclic game involving five distinct object types. Two different game structures, each involving five object types were considered. Let's denote the individual object types by A, B, C, D, E. The figure on the left corresponds to the game, with an order of powers given as follows: $A > B > C > D > E, E > A, A < C, A > D, B < D, B > E, C < E$. Note that in this case, the proportions converge to

Algorithm 2.7.1 Algorithm to identify limiting proportions

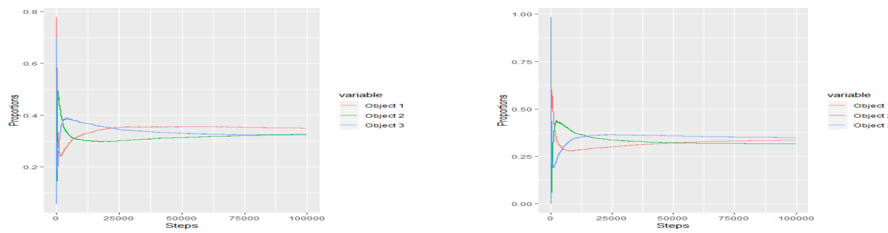
Input: A square matrix B containing k rows, with the (i, j)-th entry given by

$$\begin{aligned} B_{i,j} &= 1 \text{ if object } i > \text{object } j \\ &= 0 \text{ if object } j > \text{object } i \end{aligned}$$

for i=1,2,. . . , k-1 **do**
  for j=1,2,. . . ,k **do**
    $C_{i,j} = B(i, j) - B(i + 1, j)$
  end for
end for
for j=1,2,. . . ,k **do**
  $C_{k,j} = 1$
end for
Define a vector $\mathbf{b}$ of length k as: $\mathbf{b} = (0, 0, \ldots, 0, 1)^T$
Solve for $\mathbf{x}$ in $C\mathbf{x} = \mathbf{b}$

$(0.2, 0.2, 0.2, 0.2, 0.2)$. The figure on the right-hand side corresponds to the game given by $A > B > C > D > E, E > A, A > C, A < D, B < D, B > E, C > E$. In this case the proportions of individual object types converge to $(\frac{1}{3}, 0, \frac{1}{3}, \frac{1}{3}, 0)$. Note that a slight change in pairwise ordering results in completely different system dynamics. For both these games, we have considered the initial counts to be $(1, 1, 1, 1, 1)$, i.e. at the beginning counts of every object type are equal. In both cases, the number of steps shown in the plot is 100000.

Figure 2.5 shows the convergence of the same two games as shown in Figure 2.4. However, in this case, instead of considering the same initial counts of all five object types, the initial counts were chosen to be $(1, 10, 1, 10, 1)$. Note that the difference in initial counts results in a slower speed of convergence.



(a) Stepsizes are Poi(10) random variables

(b) Stepsizes are Poi(100) random variables

Figure 2.6: Rock-paper-scissors game with Poisson step sizes

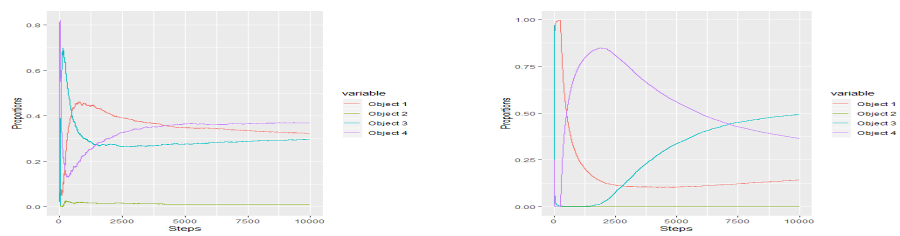
2.8 Urn processes with random step sizes

In this section a natural extension of the standard rock-paper-scissors game will be considered, where the amount of increment at each step is not fixed, i.e. the step sizes are random. We assume the random increment at step n is a Poisson(λ) random variable. Surprisingly enough, the limiting behavior of this game seems to have similar limiting properties as the standard rock-paper-scissors game.

2.8.1 Simulation results

In this section, we shall consider a wide variety of games, where the step-sizes are random, specifically these step sizes are random realizations from a Poisson distribution with a fixed mean. We shall begin with the rock-paper-scissors game with random step-sizes, and then we shall go into its extension to cyclic games involving more than three objects. Clearly the random step-sizes mimic the population dynamics of real-life data more closely.

Figure 2.6 shows the plots of proportions over time for each object type in



(a) Stepsizes are Poi(10) random variables

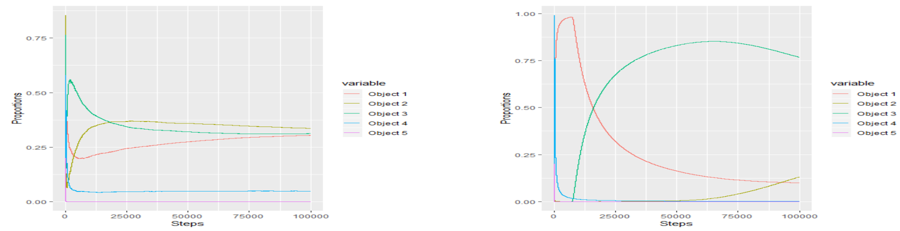
(b) Stepsizes are Poi(100) random variables

Figure 2.7: 4 player cyclic game with Poisson step sizes

a rock-paper-scissors game with random step sizes. Here the step sizes were assumed to be Poisson random variables. The figure on left corresponds to Poisson increments with a mean 10. Whereas the figure on right corresponds to the Poisson increments with mean 100. In both cases, the initial counts are chosen to be $(1, 1, 1)$, and a total of 10000 steps were observed.

Figure 2.7 shows the plots of proportions over time for a cyclic game involving four object types. Let's denote the object types by A, B, C, D. Thus the game considered here is specified as follows: $A > B > C > D > A, A > C, B < D$. Here the plots correspond to Poisson increments with mean 10 and 100 respectively. For both these cases, the initial counts were considered to be $(1, 1, 1, 1)$, and 10000 steps were observed. Note that the convergence is much faster when the step-sizes correspond to Poisson random variables with smaller mean. Note that in this case too the limiting proportions are the same as in case of deterministic step-sizes, i.e. when the increment at each step is fixed.

Figure 2.8 shows the plots of proportions over time in case of a cyclic game involving five object types and random step sizes. Here the simulations correspond



(a) Stepsizes are Poi(100) random variables

(b) Stepsizes are Poi(100) random variables

Figure 2.8: 5 player cyclic game with Poisson step sizes

to the Poisson random variables with means 10 (left) and 100 (right) respectively. For both cases the initial counts were considered to be $(1, 1, 1, 1, 1)$ and the number of steps$= 100000$. In this case too lower value of the Poisson parameter results in faster convergence of the proportions.

2.9 Conclusion

This chapter reviews some generalized stochastic versions of competitive games with underlying cyclic and non-transitive structures with the simplest example being the rock-paper-scissors urn model. Competition in these games is completely determined by pairwise orderings of distinct object types. Thus they can be characterized by a complete, directed graph, where the nodes correspond to distinct object types and the direction of edges indicate which of the connecting objects is more powerful. The extensions considered can be categorized as (1) extensions to cyclic games involving more than 3 distinct object types, and (2) extensions to games with random step sizes.

We have explicitly derived theoretical results for the stochastic rock-paper-scissors urn and presented extensive simulation studies illustrating the extensions. We have also given an algorithm that determines the long term proportions of distinct object types for all such games conditional on the fact that proportions of individual object types do converge. However, the theory behind convergence for games involving more than 3 object types still remains unaddressed.

Another interesting problem would be to find the rate of convergence for such processes. As mentioned at the beginning of the chapter, the proportions of each object type over time can be viewed as a time-dependent Markov process. It is known that Markov processes have exponential mixing times. Also, the simulation results support the fact that the speed of convergence is extremely slow for this type of urn processes and the rates are dependent on the urn composition at the beginning. For processes with random (Poisson) step sizes, the speed of convergence seems to depend on the parameter of the Poisson distribution. For scientific processes, this will answer questions related to how long will it take for a system to reach its steady-state.

3 Random multiplicative cascade models

3.1 Abstract

Random multiplicative cascade models are useful in modeling physical systems that exhibit intermittency in space and/or time. One of the key ingredients of this model is the structure-function, which uniquely identifies the cascade generating random variable. Thus the statistical estimation of the cascade model requires estimating the structure-function. Statistical estimation for Multiplicative cascade models is based on the assumption that the complete data is available at a particular resolution. However in most real-life applications, we have incomplete data, i.e. some of the data pointsat a particular resolution are missing. Thus we observe the random masses at resolution b^{-n} on a subset A of the whole space T. In the first part of this chapter, our goal is to come up with sample quantities under this setup that has the structure-function as their almost sure limit. We shall also derive a central limit theorem in this context. In the later part of this chapter, we shall apply the multiplicative cascade model to analyze the stock volume data of Tesla.

3.2 Introduction

Random multiplicative cascades are used to model physical systems for which the total amount of a certain object iteratively splits from a coarser to finer scale according to a set of rules. The limit of such processes results in a multiplicative cascade measure, which is the mathematical foundation for modeling highly intermittent behavior of naturally occurring physical phenomena. Processes involving such a splitting mechanism are non-linear in nature. Random multiplicative cascades are a special class of multiplicative models, that can incorporate this non-linearity of the data, which is otherwise difficult to handle. The self-symmetry or multifractal structure is inherent in this type of model.

The idea of Random Multiplicative cascades dates back to Kolmogorov in his seminal works (Kolmogorov (1941), Kolmogorov (1962)) on hypothesizing the local structure of turbulence and energy dissipation from larger to smaller scales in a highly turbulent fluid flow. Validation and further sophistication of Kolmogorov's hypotheses can be found in Mandelbrot (1974). Rigorous mathematical developments of Multiplicative cascade models are due to Kahane and Peyriere (1976), JOFFE et al. (1973) Frisch and Kolmogorov (1995).

There are a wide variety of scientific fields in which the data displays an intermittent property. The earliest of such fields that used Random multiplicative cascade models is the velocity field of fully developed turbulence. In Anselmet et al. (1984), She and Leveque (1994), She and Waymire (1995) the authors have explored theoretical properties of multiplicative cascade in this context. Applications of cascade model in real data of turbulence can be found in Arneodo et al. (1997), Arneodo et al. (1998). In Schmiegel et al. (2004) the authors have considered the multiplicative cascade model in the context of wind tunnel turbulent shear flow data.

In Gupta and Waymire (1993) the authors have built a theoretical framework to show that spatial rainfall and river flowdata share certain theoretical properties with turbulence velocities. The results developed in this paper provides a foundation to study spatial variability in a variety of hydrologic processes. Study of multiplicative cascade model in context of hydrologic data can also be found in Over and Gupta (1994), Molnar and Burlando (2005), Schertzer and Lovejoy (1997), Jothityangkoon et al. (2000).

The class of multiplicative cascade models in which the total mass is constant at each stage is known as the conservative cascade. Conservative cascades are building blocks of modeling internet traffic flow and understanding the local and global scaling behavior of such processes. Examples include Gilbert et al. (1999), Uhlig (2003), Resnick et al. (2003),Abry et al. (2002).

Another area of application of random multiplicative cascade models is financial time series. The stochastic volatility processes underlying financial time-series displays multiscaling behavior. Early works in this domain includes Mandelbrot et al. (1997), Mandelbrot (2001). There have been significant developments in this direction in later years. Some of the notable examples are Muzy et al. (2000), Bacry et al. (2001), Bacry et al. (2008).

The key statistical problem in the area of the random multiplicative cascade model is related to the estimation of the cascade parameters from the data. Parameters of the cascade generating distribution are uniquely identifiable from the structure-function of the cascade. Hence the results in this direction involve deriving sample quantities that can be used to estimate the structure-function of the cascade. The pioneering paper in this area, namely Ossiander et al. (2000) derived sample statistics that have the structure-function as their almost sure limits. The authors have also derived the central limit theorem in this context. In Resnick et al. (2003) the authors have come up with wavelet-based estimates of structure-function in case of conservative cascades, which has an inherent dependency that makes it mathematically more challenging. In Troutman et al. (1999) the authors have proposed estimates of Renyi exponents in the case of Random Multiplicative Cascade models. In Leövey and Lux (2012) the authors have obtained a generalized method of moments estimator for cascade parameters.

The contribution of this chapter in the study of multiplicative cascades is two-fold in nature. The first part of the chapter deals with the theory of multiplicative cascades in the context of missing data. The statistical theory of multiplicative cascade models assumes that the complete data is available at a particular resolution. However quite often in real datasets, some of the data points are missing. Thus it is very important to have a well-developed theory for the set up when the observed data is not complete. We have proposed sample quantities under this set up that converges almost surely to the actual structure-function of the cascade generating random variable. We have also derived a central limit theorem in this context. The proposed sample statistics are motivated by the ones in Ossiander et al. (2000). Towards the latter part of this chapter, we have used the random multiplicative cascade model to analyze the data on the daily stock volume of Tesla. The data analysis part can broadly be divided into the following three parts.

- The log-Normal, log-Poisson, and Beta distribution were used as cascade generating distribution to model the observed data.These cascade generating distributions were compared to assess their fits to the actual data.

- Datasets with different proportions of missing values were considered. Impact of the proportion of missing values in estimating the actual structure-function was evaluated.

- The plot of the daily stock volume of Tesla shows that there is a point, before and after which the data seems to have a different pattern. We have analyzed the data before and after the change point separately to see how different the estimated structure functions are for the two different time frames.

The organization of this chapter is as follows. We shall start by describing the framework and notations in section 3.3, followed by the problem description in section 3.4. Limiting results, as well as a central limit theorem in this context, will be derived in section 3.5. A comparison of different cascade generating distributions to fit the Tesla stock volume data will be presented in Section 3.8.The impact of missing values on the estimation of the structure-function of the cascade will be presented in Section 3.9. Analysis of the data before and after the changepoint will be presented in Section 3.10.

3.3 Notation and framework

Let $b \geq 2$ be a fixed and known natural number and T be the set of all infinite sequences, with its elements in the set $\{0, 1, 2, \ldots, b-1\}$, i.e $T = \{0, 1, 2, \ldots, b-1\}^{\mathbb{N}}$, equipped with the metric ρ given by

$$\rho(t_1, t_2) = b^{-|t_1 \wedge t_2|} \text{ for } t_1, t_2 \in T \tag{3.1}$$

Here $|t_1 \wedge t_2| = \inf\{i \geq 0 : t_1|i+1 \neq t_2|i+1\}$ where $t|n = (t_1, t_2, \ldots, t_n)$. Let's denote the Borel σ-field generated by T as $\mathcal{B}(T)$. For $n \geq 1$ and $t \in T$, denote $t = (t_1, t_2, \ldots)$. We can think of each $t \in T$ as a path in a b-ary tree. In that case $v_n = t|n \in \{0, 1, , \ldots, b-1\}$ represents the n-th stage vertex of the tree along the path t. We write $|v_n| = n$, i.e. $|v_n|$ is the level of the vertex in the b-ary tree. For $v = (v_1, v_2, \ldots, v_n)$ and $u = (u_1, u_2, \ldots, u_m)$ denote by $v * u$ the concatenation of v and u given by:

$$(v_1, v_2, \ldots, v_n) * (u_1, u_2, \ldots, u_m) = (v_1, v_2, \ldots, v_n, u_1, u_2, \ldots, u_m)$$

For $t \in T$ and $n \geq 1$, the closed ball around t of radius b^{-n} is given by:

$$\Delta_n(t) \equiv \Delta_n(t|n) = \{s \in T : \rho(s,t) \leq b^{-n}\} = \{s \in T : t_i = s_i \forall i \leq n\} \tag{3.2}$$

The normalized Haar measure on T is given by:

$$\lambda(\Delta_n(t)) = b^{-n} \text{ for all } t \in T \text{ and } n \geq 1 \tag{3.3}$$

There is a random cascade generator corresponding to each vertex of the b-ary tree. The cascade generators are non-negative iid random variables with mean 1, defined on a common probability space $(\Omega, \mathcal{F}, P)$. Those are denoted by:

$$\{W_v : v \in \{0, 1, 2, \ldots, b-1\}^n, n \geq 1\} \tag{3.4}$$

W_v corresponds to the random cascade generator corresponding to the vertex v. In practice, common families of cascade generating distributions are log-normal, log-Poisson and Gamma.

Define the increasing filtration $\{\mathcal{F}_n : n \geq 1\}$ as follows:

$$\mathcal{F}_n = \sigma(W_v : |v| \leq n) \tag{3.5}$$

i.e. $\mathcal{F}_n$ is the $\sigma-$field generated by all cascade generators up to the n-th stage. $\{\lambda_n : n \geq 1\}$ is a sequence of random measures on $(T, \mathcal{B}(T))$, defined as follows:

$$\frac{d\lambda_n}{d\lambda} = \prod_{i=0}^{n} W_{t|i} \text{ for } t \in T \tag{3.6}$$

Here the cascade generator W_ϕ is an a.s. positive random variable which is independent of $\mathcal{F}_n, n \geq 1$.

Lemma 3.3.1. *For any bounded Borel measurable function f, $\int_T f(t) d\lambda_n(t)$ is a L_1- martingale.*

This result follows because the cascade generators W_v have mean $\mathbb{E}(W_v) = 1$. Any L_1 martingale converges almost surely. This asserts existence of a limiting random measure λ_∞, with

$$\lambda_\infty(A) = \lim_{n\to\infty} \lambda_n(A)$$

where

$$\mathbb{P}(\lambda_n \xrightarrow{v} \lambda_\infty \text{ as } n \to \infty) = 1$$

Here '$\xrightarrow{v}$' denotes vague convergence of measures.

Definition 3.3.2. *A sequence of measures μ_n converges vaguely to a measure μ if*

$$\int f d\mu_n \to \int f d\mu \text{ as } n \to \infty \text{ for all continuous function } f \text{ with bounded support} \tag{3.7}$$

Now, consider the following new sequence of random variables $Z_\infty(v)$ corresponding to individual vertices v.

$$Z_\infty(v) = \lim_{N\to\infty} \sum_{u:|u|=N-n} \prod_{i=1}^{N-n} W_{v*(u_1,u_2,\ldots u_i)} b^{N-n} \tag{3.8}$$

These can be thought of as multiplicative error terms in the context of multiplicative cascades. To see that note for any $N \geq n+1$, $n = 1, 2, \ldots$ one has:

$$\lambda_N(\Delta_n(v) \cap A) = Z_N^{(n)}(v).\lambda_n(\Delta_n(v) \cap A) \tag{3.9}$$

where

$$Z_N^{(n)}(v) = \sum_{u:|u|=N-n} \prod_{i=1}^{N-n} W_{v*(u_1,u_2,\ldots u_i)} b^{N-n} \tag{3.10}$$

Clearly for each $n \geq 1$ and $|v| = n$, the sequence $\{Z_N^{(n)}(v) : N = n+1, n+2, \dots\}$ is a non-negative martingale. Therefore it converges almost surely, hence $Z_\infty^{(n)}(v)$ exists a.s. and is independent of $\mathcal{F}_n$. Thus we can get

$$\lambda_\infty(\Delta_n(v) \cap A) = Z_\infty(v|n)\lambda_n(\Delta_n(v) \cap A) \tag{3.11}$$

Note that the limiting random measure of a set is expressed as the product of random measure at a finite level and a multiplicative error term $Z_\infty(v|n)$.

Definition 3.3.3. *The structure function corresponding to a cascade generator W is given by*

$$\chi_b(h) = \log_b \mathbb{E}\left(W^h \mathbf{1}(W > 0)\right) - (h-1) \tag{3.12}$$

There is a one to one correspondence between the cascade generating distribution and the structure function. In Ossiander et al. (2000) the authors have shown that under certain regularity conditions, $\{\lambda_\infty(\Delta_n(v)) : v \in \{0, 1, \dots, b-1\}^n\}$ for a fixed $n \geq 1$, uniquely determines the distribution of the cascade generator W. We shall derive an analogous result in case of incomplete data, i.e. when observations of $\lambda_\infty(\Delta_n(v) \cap A)$ are available instead of $\lambda_\infty(\Delta_n(v))$ for some A. Before going into the problem description we shall state a theorem from Kahane and Peyriere (1976) to establish the connection between the properties of structure function and whether or not the cascade will eventually die out.

Theorem 3.3.4 (Kahane and Peyriere (1976))**.** *The following statements hold.*

1. $\mathbb{E}(\lambda_\infty(T)) > 0$ *iff* $\chi_b'(1-) < 0$.

2. $\mathbb{E}(\lambda_\infty^h(T)) < \infty$ *for* $0 \leq h \leq 1$, *and if* $h_c := \sup\{h \geq 1 : \chi_b(h) \leq 0\} > 1$, *then* $\mathbb{E}(\lambda_\infty^h(T)) < \infty$ *for* $1 \leq h \leq h_c$.

3. $\mathbb{E}(\lambda_\infty(T)) = 1$ *iff* $\mathbb{E}(\lambda_\infty(T)) > 0$.

Figure 3.1 shows the random multiplicative cascades at different levels denoted by n, where the cascade generating distribution is log-normal with mean 1 and variance 0.1.

3.4 Problem description

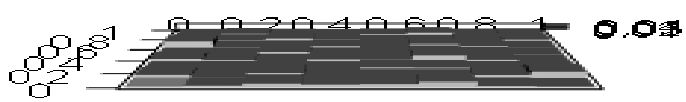
In real life usually we don't get to observe the complete data at a resolution b^{-n}, given by $\{\lambda_\infty(\Delta_n(v)) : v \in \{0, 1, \ldots, b-1\}^n; n \geq 1\}$. Almost always there are some observations that are missing. Here we assume, that the data is observed on a measurable proper subset A of T, i.e. we have observed $\{\lambda_\infty(\Delta_n(v) \cap A) : |v| = n\}$ where for each $v : |v| = n$ we have $\Delta_n(v) \cap A \neq \phi$.

Our goal is to derive large sample results under this set up. We shall start by defining sample quantities in this context that will have the actual structure function of the cascade generating distribution as their almost sure limit. Then we shall derive a central limit theorem. Here we shall assume $\chi_b'(1-) < 0$, so that $\lambda_\infty(T) > 0$. Then we shall consider a $A \subset T$ such that $\lambda_\infty(A) > 0$ and $\mathbb{E}(\lambda_\infty(A)) < \infty$. We shall also need the following convergence result:

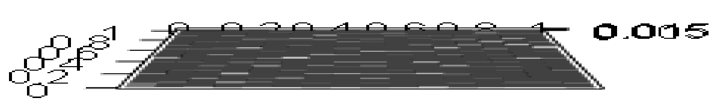
$$\lambda_n(A) \to \lambda_\infty(A) \ P\text{-a.s.} \tag{3.13}$$

(a) log-Normal cascade at level 2

(b) log-Normal cascade at level 3



(c) log-Normal cascade at level 4



(d) log-Normal cascade at level 5

Figure 3.1: log-Normal cascades at different levels

where λ_∞ is the almost sure vague limit of λ_n and

$$\lambda_n(\Delta_n(v)) = \prod_{i=1}^{n} \frac{W_{v|i}}{b} \tag{3.14}$$

3.5 h-cascades and its properties

The h-cascades $W_v(h)$ are obtained by doing a size bias transformation of the original cascade generating random variables. For any v, h-cascades are defined as follows:

$$W_v(h) = \frac{W_v^h}{\mathbb{E}(W_v^h)} \text{ for } h \in \mathbb{R} \tag{3.15}$$

Clearly these are also non-negative random variables with mean 1. Hence these can also be treated as cascade generating random variables. For $n = 1, 2, \ldots$ the measure corresponding to the n-th h-cascades given by $\lambda_n(h;.)$, are defined as follows.

$$\frac{d\lambda_n(h;.)}{d\lambda}(t) = \prod_{i=0}^{n} W_{t|i}(h) \text{ for } t \in T \tag{3.16}$$

Lemma 3.5.1. *For a suitable $A \in \mathcal{B}(T)$*

$$\frac{\lambda_n^h(\Delta_n(v) \cap A)}{b^{n\chi_b(h)} C_n(v,h;A)} = \lambda_n(h; \Delta_n(v) \cap A) \tag{3.17}$$

where $C_n(v,h;A) = (b^n \lambda(\Delta_n(v) \cap A))^{h-1}$ is deterministic.

Proof.

$$\begin{aligned}
\lambda_n^h(\Delta_n(v)\cap A) &= \prod_{i=1}^{n} W_{v|i}^h \lambda^h(\Delta_n(v)\cap A) \\
&= \left[\prod_{i=1}^{n} \frac{W_{v|i}^h}{\mathbb{E}(W^h)}\right] \lambda(\Delta_n(v)\cap A)(\mathbb{E}(W^h))^n \lambda^{h-1}(\Delta_n(v)\cap A) \\
&= \left[\prod_{i=1}^{n} \frac{W_{v|i}^h}{\mathbb{E}(W^h)}\right] \lambda(\Delta_n(v)\cap A) b^{n[\log_b \mathbb{E}(W^h)-(h-1)]} \left(\frac{\lambda^{h-1}(\Delta_n(v)\cap A)}{(1/b^n)^{h-1}}\right) \\
&= \lambda_n(h;\Delta_n(v)\cap A) b^{n\chi_b(h)} C_n(v,h;A)
\end{aligned}$$

□

Define the n-th scale sample moment corresponding to the observed data as follows:

$$M_n(h;A) = \sum_{v:|v|=n} \frac{\lambda_\infty^h(\Delta_n(v)\cap A)}{C_n(v,h;A)} \tag{3.18}$$

Correspondingly define

$$\hat{\tau}_n(h;A) = \frac{1}{n}\log_b M_n(h;A) \tag{3.19}$$

$$\tilde{\tau}_n(h;A) = \log_b \frac{M_{n+1}(h;A)}{M_n(h;A)} \tag{3.20}$$

The above definitions require $C_n(v,h;A)$ to be non-zero for all v such that $|v| = n$. We shall show that these sample quantities converge to some functionals of the actual structure-function of the cascade generating distribution. Moreover if the moment generating function of $\{\log W_v\}$ is uniquely defined at the origin, then the collection $\{\lambda_\infty(\Delta_n(v)) : n \geq 1\}$ uniquely determines the cascade generating distribution, i.e. the distribution of W_v's. Also, the central limit theorem for a suitably normalized version of these estimators is presented towards the end.

Lemma 3.5.2.

$$\frac{M_n(h;A)}{b^{n\chi_b(h)}} = \sum_{v:|v|=n} Z^h_\infty(v)\lambda_n(h;\Delta_n(v)\cap A) \tag{3.21}$$

Proof. First note that for any $N \geq n+1$, $n = 1, 2, \ldots$ one has:

$$\lambda_\infty(\Delta_n(v)\cap A) = Z_\infty(v|n)\lambda_n(\Delta_n(v)\cap A) \tag{3.22}$$

Hence by the previous lemma:

$$\begin{aligned}
\frac{\lambda^h_\infty(\Delta_n(v)\cap A)}{C_n(v,h;A)b^{n\chi_b(h)}} &\overset{\mathrm{d}}{=} Z^h_\infty(v|n)\frac{\lambda^h_n(\Delta_n(v)\cap A)}{C_n(v,h;A)b^{n\chi_b(h)}} \\
&= Z^h_\infty(v|n)\lambda_n(h;\Delta_n(v)\cap A) \\
\frac{M_n(h;A)}{b^{n\chi_b(h)}} &= \sum_{v:|v|=n} Z^h_\infty(v|n)\lambda_n(h;\Delta_n(v)\cap A)
\end{aligned}$$

□

Remark 3. *Note that* $b^n\lambda(\Delta_n(v)\cap A)$ *represents the ratio of the Lebesgue measure of* $\Delta_n(v)\cap A$ *to that of* $\Delta_n(v)$*. Thus whenever* $\lambda(\Delta_n(v)\cap A)$ *is small, more weight is given on the observed total mass that corresponds to the set* $\Delta_n(v)$*. Thus this constant acts as a normalizing constant. Also note that we can think of this ratio* *as*

$$b^n \lambda(\Delta_n(v) \cap A) = \int_{\Delta_n(v)} \frac{d\lambda_A}{d\lambda} d\lambda$$

where $\lambda_A(C) = \lambda(A \cap C)$*. In this sense* $M_n(h; A)$ *is a weighted mean, where the weights depend on the set* A *and the value of* h*.*

The following result about structure-function from Ossiander et al. (2000) will be used to derive the asymptotic results in the next sections.

$$\chi_{b,h}(r) = \chi_b(hr) - r\chi_b(h) \tag{3.23}$$

where $\chi_{b,h}(.)$ is the structure function of the h-cascade.

3.6 Strong Law of large number

In this section, we shall prove the consistency of the point estimate of structure-function under the missing data set up. We shall show that the proposed estimate converges almost surely to the structure-function of the cascade generating distribution. Before going into the theorem, first recall proposition 2.2 from Ossiander et al. (2000).

Proposition 3.6.1. *Assume that* $\chi_b'(1-) < 0$ *and let*

$$\begin{aligned} H_c^+ &= \sup\{h \geq 1 : h\chi_b'(h) - \chi_b(h) < 0\} \\ H_c^- &= \inf\{h \leq 0 : h\chi_b'(h) - \chi_b(h) < 0\} \end{aligned}$$

Then $H_c^- \leq 0 < 1 \leq H_c^+$, with $h\chi_b'(h) - \chi_b(h) < 0$ for all $H_c^- < h < H_c^+$. Furthermore for $h \in [0,1] \cup (H_c^-, H_c^+)$, $\lambda_n(h;T) \to \lambda_\infty(h;T)$ P-a.s., where $\mathbb{E}(\lambda_\infty(h;T)) = 1$.

Remark 4. *To see that the condition $\chi_b'(1-) < 0$ is natural, first note that $-\chi_b'(1-)$ is the Hausdorff dimension of the subset of T that supports λ_∞. See Waymire and Williams (1995) for reference. Also for a fractal structure, the Hausdorff dimension exceeds the topological dimension, which is 0 in this case. Hence we must have $\chi_b'(1-) < 0$.*

Theorem 3.6.2. *For $h \in [0,1] \cup (H_c^-, H_c^+)$,*

$$\frac{M_n(h;A)}{b^{n\chi_b(h)}} \to \lambda_\infty(h;A)\mathbb{E}(Z_\infty^h(v)) \tag{3.24}$$

P-a.s. as $n \to \infty$ uniformly for A that satisfies the desired conditions.

Proof. For any fixed $h \in (H_c^-, H_c^+)$, we can write :

$$\frac{M_n(h;A)}{b^{n\chi_b(h)}} = \sum_{v:|v|=n} Z_\infty^h(v)\lambda_n(h;\Delta_n(v) \cap A) \tag{3.25}$$

Choose $\epsilon > 0$ small enough so that we have both $\chi_{b,h}(1+\epsilon) < 0$ and $h(1+\epsilon) \in (H_c^-, H_c^+)$. Set $\alpha = b^{\chi_{b,h}(1+\epsilon)}$. Note $\alpha < 1$. Fix $n > 1$, and for $|v| = n$, let

$$\tilde{Z}(v) = Z_\infty(v)\mathbf{1}(Z_\infty^h(v)\lambda_n(h;\Delta_n(v)) < \alpha^{n/2(1+\epsilon)}) \tag{3.26}$$

Now we shall use this $\tilde{Z}$ and a conditional centering to decompose $M_n(h;A)$.

Write

$$\frac{M_n(h;A)}{b^{n\chi_b(h)}} - \mathbb{E}(Z^h_\infty(v)).\lambda_n(h;A) = \sum_{v:|v|=n} (Z^h_\infty(v) - \tilde{Z}^h(v))\lambda_n(h; A\cap\Delta_n(v))$$
$$+ \sum_{v:|v|=n} (\tilde{Z}^h(v) - \mathbb{E}(\tilde{Z}^h(v)|\mathcal{F}_n)\lambda_n(h; A\cap\Delta_n(v))$$
$$- \left[\mathbb{E}(Z^h_\infty(v))\lambda_n(h;A) - \sum_{v:|v|=n} \mathbb{E}[\tilde{Z}^h(v)|\mathcal{F}_n]\lambda_n(h; A\cap\Delta_n(v))\right] \tag{3.27}$$

We shall show that each of the term on the right hand side converges to 0 P-a.s. as $n \to \infty$. That will imply

$$\frac{M_n(h;A)}{b^{n\chi_b(h)}} - \mathbb{E}(Z^h_\infty(v)).\lambda_n(h;A) \to 0$$

Hence the asserted result will follow, because:

$$\lambda_n(h;A) \to \lambda_\infty(h;A) \ P\text{-a.s.} \tag{3.28}$$

Let $A_n = \bigcup_{v:|v|=n} [Z_\infty(v) \neq \tilde{Z}(v)]$ Using union-sum inequality and Markov's inequality yields the following:

$$
\begin{aligned}
P(A_n) &\leq \sum_{v:|v|=n} \mathbb{P}(Z_\infty(v) \neq \tilde{Z}(v)) \\
&= b^n \mathbb{P}(Z_\infty^h(v)\lambda_n(h;\Delta_n(v)) > \alpha^{n/2(1+\epsilon)}) \\
&\leq b^n \mathbb{E}\left[Z_\infty^{h(1+\epsilon)}(v)\lambda_n^{1+\epsilon}(h;\Delta_n(v))\alpha^{-\frac{n}{2}}\right] \\
&= b^n \mathbb{E}(Z_\infty^{h(1+\epsilon)}(v))\mathbb{E}(\lambda_n^{1+\epsilon}(h;\Delta_n(v)))\alpha^{-\frac{n}{2}} \\
&= b^n \mathbb{E}(Z_\infty^{h(1+\epsilon)}(v))\alpha^{-\frac{n}{2}}\left[\prod_{i=1}^{n}\frac{W_{v|i}^h}{\mathbb{E}(W^h)}\right]^{1+\epsilon} \lambda^{1+\epsilon}(\Delta_n(v)\cap A) \\
&= b^n \mathbb{E}(Z_\infty^{h(1+\epsilon)}(v))\alpha^{-\frac{n}{2}} b^{n(\chi_{b,h}(1+\epsilon)-1)}\left(\lambda(\Delta_n(v)\cap A).b^{-n}\right)^{1+\epsilon} \\
&= \alpha^{\frac{n}{2}}\mathbb{E}(Z_\infty^{h(1+\epsilon)}(v))\left(\lambda(\Delta_n(v)\cap A)b^{-n}\right)^{1+\epsilon} \\
&\leq \alpha^{n/2}\mathbb{E}(Z_\infty^{h(1+\epsilon)}(v))
\end{aligned}
$$

The following result was used to derive the above equalities:

$$
\begin{aligned}
b^{n(\chi_{b,h}(1+\epsilon))} &= b^{n\left[\log_b \mathbb{E}(W^{h(1+\epsilon)})-(h(1+\epsilon)-1)-(1+\epsilon)\log_b \mathbb{E}(W^h)+(1+\epsilon)(h-1)\right]} \\
&= \left[\frac{\mathbb{E}(W^{h(1+\epsilon)})}{(\mathbb{E}(W^h))^{1+\epsilon}}\right]^n b^{-n\epsilon}
\end{aligned}
$$

By our assumption $\mathbb{E}(\lambda_\infty^{h(1+\epsilon)}(A)) < \infty$ and we have $\alpha < 1$, therefore $\sum_n \mathbb{P}(A_n) < \infty$. Therefore by Borel-Cantelli lemma $\mathbb{P}(A_n \text{ i.o.}) = 0$. Thus as $n \to \infty$, one has P-a.s. $\sum_{v:|v|=n}(Z_\infty^h(v) - \tilde{Z}^h(v))\lambda_n(h; A \cap \Delta_n(v)) \to 0$.

Now let's consider the third term in equation (3.27). It can be written as:

$$\begin{aligned}
\mathbb{E}Z^h_\infty(v)\lambda_n(h;A) \quad &- \sum_{v:|v|=n} \mathbb{E}[\tilde{Z}^h(v)|\mathcal{F}_n]\lambda_n(h;A\cap\Delta_n(v)) = \sum_{v:|v|=n} \Big[\mathbb{E}(Z^h_\infty(v) - \mathbb{E}(\tilde{Z}^h(v)|\mathcal{F}_n))\Big] \\
&\quad \mathbb{E}(\lambda_n(h;\Delta_n(v)\cap A)) \\
&= \sum_{v:|v|=n} \mathbb{E}\left[Z^h_\infty(v)\mathbf{1}\{Z^h_\infty(v)\lambda_n(h;A\cap\Delta_n(v)) > \alpha^{n/2(1+\epsilon)}\}|\mathcal{F}_n\right] \\
&\quad \mathbb{E}(\lambda_n(h;\Delta_n(v)\cap A)) \\
&\leq \sum_{v:|v|=n} \mathbb{E}\left[Z^{h(1+\epsilon)}_\infty(v)\lambda_n^{1+\epsilon}(h;\Delta_n(v))\right]\alpha^{-\frac{n\epsilon}{2(1+\epsilon)}} \\
&= \mathbb{E}\left(Z^{h(1+\epsilon)}_\infty(v)\right)\alpha^{-\frac{n\epsilon}{2(1+\epsilon)}} \sum_{v:|v|=n} \lambda_n^{1+\epsilon}(h;\Delta_n(v)) \\
&= \mathbb{E}\left(Z^{h(1+\epsilon)}_\infty(v)\right)\alpha^{\frac{n(2+\epsilon)}{2(1+\epsilon)}} \\
&\quad \sum_{v:|v|=n} \lambda_n(h(1+\epsilon);\Delta_n(v))(\lambda(\Delta_n(v))b^n)^\epsilon \\
&\leq \mathbb{E}\left(Z^{h(1+\epsilon)}_\infty(v)\right)\alpha^{\frac{n(2+\epsilon)}{2(1+\epsilon)}} \\
&\quad \sum_{v:|v|=n} \lambda_n(h(1+\epsilon);\Delta_n(v)) \\
&= \mathbb{E}(Z^{h(1+\epsilon)}_\infty(v))\alpha^{\frac{n(2+\epsilon)}{2(1+\epsilon)}}\lambda_n(h(1+\epsilon),T)
\end{aligned}$$

Since $\mathbb{E}(Z^{h(1+\epsilon)}_\infty(v)) < \infty$ and $\lambda_n(h(1+\epsilon),T) \to \lambda_\infty(h(1+\epsilon),T)$ P-a.s. and $\alpha < 1$, therefore this sum converges to 0, P-a.s. as $n \to \infty$.

Finally we shall consider the middle term, given by

$$S_n(A) = \sum_{v:|v|=n} (\tilde{Z}^h(v) - \mathbb{E}(\tilde{Z}^h(v)|\mathcal{F}_n))\lambda_n(h;A\cap\Delta_n(v))$$

Note that for any $u \neq v$ with $|v| = |u| = n$, $\tilde{Z}(u)$ and $\tilde{Z}(v)$ are conditionally independent and thus conditionally uncorrelated given $\mathcal{F}_n$. Therefore

$$\begin{aligned}
\mathrm{Var}(S_n(A)) &= \mathbb{E}\left[\mathbb{E}(S_n(A)^2|\mathcal{F}_n)\right] \\
&\leq \mathbb{E}\sum_{v:|v|=n}\left[(\tilde{Z}^h(v) - \mathbb{E}(\tilde{Z}^h(v)|\mathcal{F}_n))^2|\mathcal{F}_n\right]\lambda_n^2(h; A\cap\Delta_n(v)) \\
&\leq \mathbb{E}\sum_{v:|v|=n}\mathbb{E}(\tilde{Z}^{2h}(v)\lambda_n^2(h; A\cap\Delta_n(v))|\mathcal{F}_n) \\
&\leq \alpha^{n/2(1+\epsilon)}\sum_{v:|v|=n}\mathbb{E}(\tilde{Z}^h(v)\lambda_n(h; A\cap\Delta_n(v))) \\
&\leq \alpha^{n/2(1+\epsilon)}\mathbb{E}(\lambda_\infty^h(T))
\end{aligned}$$

Since $\alpha < 1$ and $\mathbb{E}(\lambda_\infty^h(T)) < \infty$, therefore it follows $\sum_n \mathrm{Var}(S_n(A)) < \infty$ and hence $S_n \to 0$, P-a.s. as $n \to \infty$.

□

Corollary 3.6.3. *On the set* $[\lambda_\infty(A) > 0]$, *one has* P*-a.s. that*

$$\{\tilde{\tau}_n(h) : h \in [0,1] \cup (H_c^-, H_c^+)\} \to \{\chi_b(h) : h \in [0,1] \cup (H_c^-, H_c^+)\} \tag{3.29}$$

Proof. On the set $[\lambda_\infty(A) > 0]$,

$$b^{\tilde{\tau}_n(h;A)} = \frac{M_{n+1}(h;A)}{M_n(h;A)} = b^{\chi_b(h)}\frac{M_{n+1}(h;A)}{b^{(n+1)\chi_b(h)}}\left(\frac{M_n(h;A)}{b^{n\chi_b(h)}}\right)^{-1} \to b^{\chi_b(h)} \tag{3.30}$$

P-a.s. for $h \in [0,1] \cup (H_c^-, H_c^+)$. □

3.7 Pointwise Central limit theorem

For every $n \geq 1$, let $\{X_n(v) : |v| = n\}$ be a collection of independent random variables which are also independent of $\mathcal{F}_n$. Define

$$S_n(h; A) = \sum_{v:|v|=n} X_n(v)\lambda_n(h; \Delta_n(v) \cap A) \tag{3.31}$$

and

$$R_n(h; A) = \frac{S_n(h; A)}{\left(\sum_{v:|v|=n} \lambda_n^2(h; \Delta_n(v) \cap A)\right)^{1/2}} \tag{3.32}$$

We assume $R_n(h; A) = 0$ when $\lambda_n(A) = 0$.

Theorem 3.7.1. *If $\mathbb{E}(X_n^2(v)) = 1$ and $\mathbb{E}(X_n(v)) = 0$ for each v with $|v| = n$, and there exists $\delta > 0$ such that*

$$\sup_n \sup_{|v|=n} \mathbb{E}|X_n(v)|^{2(1+\delta)} < \infty$$

then for $h \in (H_c^-/2, H_c^+/2)$ we have:

$$\lim_{n\to\infty} \mathbb{E}\left[e^{izR_n(h;A)}|\mathcal{F}_n\right] = \mathbf{1}[\lambda_\infty(A) = 0] + \mathbf{1}[\lambda_\infty(A) > 0]e^{-(1/2)z^2} \tag{3.33}$$

Proof. Fix $h \in (H_c^-/2, H_c^+/2)$ and choose sufficiently small δ such that $h(1+\delta) \in (H_c^-/2, H_c^+/2)$, $\sup_n \sup_{|v|=n} \mathbb{E}|X_n(v)|^{2(1+\delta)} < \infty$ and $\chi'_{b,2h}(1+\delta) < 0$. The last condition together with the convexity of $\chi_{b,2h}$ implies

$$\chi_b(2h(1+\delta)) - (1+\delta)\chi_b(2h) < 0$$

For $|v| = n$, let

$$Y_n(v;A) = \frac{X_n(v)\lambda_n(h;\Delta_n(v)\cap A)}{\left(\sum_{u:|u|=n} \lambda_n^2(h;\Delta_n(v)\cap A)\right)^{1/2}} \tag{3.34}$$

Take $Y_n(v;A) = 0$ when $\lambda_n(A) = 0$. Given $\mathcal{F}_n$, $Y_n(v)$'s are all conditionally independent mean zero random variables. Therefore

$$\mathbb{E}\left[\left(\sum_{v:|v|=n} Y_n(v;A)\right)^2 |\mathcal{F}_n\right] = \sum_{v:|v|=n} \mathbb{E}(Y_n^2(v;A)|\mathcal{F}_n) = \mathbf{1}(\lambda_n(A) > 0) \tag{3.35}$$

Clearly

$$B_n := [\lambda_n(A) > 0] \downarrow [\lambda_\infty(A) > 0] := B \text{ a.s. } P. \tag{3.36}$$

We shall show that Lindeberg's condition holds conditionally on $\mathcal{F}_n$. The following relation is useful for the subsequent part of the proof:

$$
\begin{aligned}
\sum_{v:|v|=n} \lambda_n^r(h;\Delta_n(v)\cap A) &= \sum_{v:|v|=n} \frac{\lambda_n^{hr}(\Delta_n(v)\cap A)}{b^{nr\chi_b(h)}C_n^r(v,h;A)} \\
&= \sum_{v:|v|=n} \lambda_n^{hr}(\Delta_n(v)\cap A)\mathbb{E}^{-nr}(W^h)b^{-nr(h-1)}C_n^{-r}(v,h;A) \\
&= \sum_{v:|v|=n} \lambda_n(hr;\Delta_n(v)\cap A)b^{n\chi_b(hr)}C_n(v,hr;A)\mathbb{E}^{-nr}(W^h)b^{-nr(h-1)} \\
&\quad C_n^{-r}(v,h;A) \\
&= b^{n\chi_{b,h}(r)} \sum_{v:|v|=n} \lambda_n(hr;\Delta_n(v)\cap A)C_n^{-r}(v,h;A)C_n(v,hr;A) \\
&= b^{n\chi_{b,h}(r)} \sum_{v:|v|=n} \lambda_n(hr;\Delta_n(v)\cap A)(b^n\lambda(\Delta_n(v)\cap A))^{hr-1-r(h-1)} \\
&= b^{n\chi_{b,h}(r)} \sum_{v:|v|=n} \lambda_n(hr;\Delta_n(v)\cap A)(b^n\lambda(\Delta_n(v)\cap A))^{r-1}
\end{aligned}
$$

Now

$$
\begin{aligned}
\mathbb{E}\left[\sum_{v:|v|=n} Y_n^2(v)\mathbf{1}[|Y_n(v)|>\epsilon]|\mathcal{F}_n\right] &\leq \sum_{v:|v|=n} \mathbb{E}\left[Y_n^2(v)\mathbf{1}(|Y_n(v)|>\epsilon)|\mathcal{F}_n\right]\mathbf{1}[B]+\mathbf{1}[B_n-B] \\
&\leq \epsilon^{-2\delta}\sum_{v:|v|=n} \mathbb{E}\left[|Y_n(v)|^{2(1+\delta)}|\mathcal{F}_n\right]\mathbf{1}[B]+\mathbf{1}[B_n-B] \\
&\leq \epsilon^{-2\delta}\sup_{|v|=n}\mathbb{E}|X_n(v)|^{2(1+\delta)} \\
&\quad \frac{\sum_{v:|v|=n}\lambda_n^{2(1+\delta)}(h;\Delta_n(v)\cap A)}{\left(\sum_{v:|v|=n}\lambda_n^2(h;\Delta_n(v)\cap A)\right)^{1+\delta}}\mathbf{1}[B]+\mathbf{1}[B_n-B] \\
&\leq C_n\epsilon^{-2\delta}b^{n(\chi_b(2h(a+\delta)-(1+\delta)\chi_b(2h))} \\
&\quad \frac{\sum_{v:|v|=n}\lambda_n(2h(1+\delta);\Delta_n(v)\cap A)(b^n\lambda(\Delta_n(v)\cap A))^{2(1+\delta)-}}{(\sum_{v:|v|=n}\lambda_n(2h;\Delta_n(v)\cap A)(b^n\lambda(\Delta_n(v)\cap A)))^{1+\delta}} \\
&\quad \mathbf{1}[B]+\mathbf{1}[B_n-B]
\end{aligned}
$$

Note that a sufficient condition to have $(\sum_{v:|v|=n} \lambda_n(2h; \Delta_n(v) \cap A)(b^n \lambda(\Delta_n(v) \cap A)) > 0$ is given by $\lambda_n(2h; A) > 0$. Both $\sum_{v:|v|=n} \lambda_n(2h; \Delta_n(v) \cap A) \to \sum_{v:|v|=n} \lambda_\infty(2h; \Delta_n(v) \cap A)$ and $\lambda_n(2h; A) > 0$ P-a.s. on the set B. By equation (3.36) and $(\chi_b(2h(a+\delta) - (1+\delta)\chi_b(2h)) < 0$, therefore the right hand side of the above equation converges to 0 P-a.s. as $n \to \infty$. Thus

$$\lim_n \mathbb{E}\left[\sum_{v:|v|=n} Y_n^2(v)\mathbf{1}[|Y_n(v)| > \epsilon]|\mathcal{F}_n\right] = 0$$

The rest of the proof is same as the usual proof of Lindeberg's central limit theorem as can be found in page 369 of Billingsley (1986).

□

Corollary 3.7.2. *For* $h \in (\frac{H_c^-}{2}, \frac{H_c^+}{2})$,

$$\lim_n \mathbb{E}(e^{izR_n(h;A)}) = \mathbb{P}(\lambda_\infty(A) = 0) + e^{-(1/2)z^2}\mathbb{P}(\lambda_\infty(A) > 0) \tag{3.37}$$

The proof of this corollary follows by simply applying the dominated convergence theorem to the above result.

Corollary 3.7.3. *For* $h \in (\frac{H_c^-}{2}, \frac{H_c^+}{2})$,

$$R_n(h; A) \xrightarrow{d} \eta N_h \tag{3.38}$$

where N_h *has a standard normal distribution and* η *has the same distribution as* $\mathbf{1}[\lambda_\infty(A) > 0]$.

Proof. First note that $\mathbb{E}(e^{iz\eta N_h}) = \mathbb{P}(\lambda_\infty(A) = 0) + e^{-(1/2)z^2}\mathbb{P}(\lambda_\infty(A) > 0)$. The rest follows from the continuity theorem for characteristic functions. □

Corollary 3.7.4. *For $h \in (\frac{H_c^-}{2}, \frac{H_c^+}{2})$,*

$$\frac{M_n(h;A)/b^{n\chi_b(h)} - \mathbb{E}(\lambda_\infty^h(T))\lambda_n(h;A)}{[Var(\lambda_\infty^h(T))]^{1/2}\left[\sum_{v:|v|=n}\lambda_n^2(h;\Delta_n(v)\cap A)\right]^{1/2}} \xrightarrow{d} \eta N_h \tag{3.39}$$

where η and N_h are as mentioned before.

Proof. In theorem 3.7.1, put $X_n(v) = (Z_\infty^h(v) - \mathbb{E}\lambda_\infty^h(T))(\mathrm{Var}\lambda_\infty^h(T))^{-1/2}$ in the expression of $R_n(h;A)$ to get the desired result. □

Corollary 3.7.5. *For $h \in (\frac{H_c^-}{2}, \frac{H_c^+}{2})$,*

$$\frac{M_n(h;A)/b^{n\chi_b(h)} - \mathbb{E}Z_\infty^h(v).\lambda_n(h;A)}{(VarZ_\infty^h(v))^{1/2}\left(\sum_{v:|v|=n}\lambda_n^2(h;\Delta_n(v)\cap A)\right)^{1/2}} \xrightarrow{d} \eta N_h \tag{3.40}$$

Proof. Put $X_n(v) = \frac{Z_\infty^h(v) - \mathbb{E}(Z_\infty^h(v))}{\mathrm{Var}(Z_\infty^h(v))^{1/2}}$ in equation 3.32 to obtain the desired result. □

Corollary 3.7.6. *For $h \in (\frac{H_c^-}{2}, \frac{H_c^+}{2})$,*

$$\frac{M_n(h;A)/b^{n\chi_b(h)} - M_{n+1}(h;A)/b^{(n+1)\chi_b(h)}}{\left(\sum_{v:|v|=n}\lambda_n^2(h;\Delta_n(v)\cap A)\right)^{1/2}} \xrightarrow{d} c_h\eta N_h \tag{3.41}$$

where η and N_h are independent with $\eta \stackrel{d}{=} \mathbf{1}[\lambda_\infty(A) > 0]$, N_h has a standard normal distribution and $c_h^2 = Var(X_n(v)$.

Proof. The numerator can be written as:$\sum_{v:|v|=n} X_n(v)\lambda_n(h;\Delta_n(v)\cap A)$ with

$$X_n(v) = Z^h_\infty(v) - \sum_{i=0}^{b-1} \frac{W^h_{v*i}}{b\mathbb{E}(W^h)} Z^h_\infty(v*i) \text{ for } v \in A$$

For each n, these are iid mean zero random variables that are independent of $\mathcal{F}_n$. □

Now define an estimator of $\chi_b(h)$, given by $\hat{\tau}_n(h;A) = \log_b(M_{n+1}(h;A)/M_n(h;A)$

We already have the asymptotic consistency of the estimator given by

$$\log_b(M_{n+1}(h)/M_n(h)) \text{ for } h \in (H_c^-, H_c^+)$$

Now our aim is to develop an observable normalization for the estimator. This will give an estimate of the variance, which is computable from the data. Define

$$D_n^2(h;A) = \sum_{v:|v|=n} \left(\frac{\lambda^h_\infty(\Delta_n(v)\cap A)}{M_n(h;A)C_n^{1/2}(v,2h;A)} - \frac{1}{M_{n+1}(h;A)} \sum_{i=0}^{b-1} \frac{\lambda^h_\infty(\Delta_{n+1}(v*i)\cap A)}{C_n^{1/2}(v,2h;A)} \right)^2 \tag{3.42}$$

Corollary 3.7.7. *For $h \in (H_c^-/2, H_c^+/2)$,*

$$\frac{(M_{n+1}(h;A)/M_n(h;A)b - \chi_b(h) - 1)}{D_n(h;A)} \xrightarrow{d} \eta N_h \tag{3.43}$$

Proof. We shall use the following equality to prove the corollary.

$$\begin{aligned}
b^{-n\chi_{b,h}(2)} &= b^{-n\chi_b(2h)+2n\chi_b(h)} \\
&= b^{-n\{\log_b \mathbb{E}(W^{2h})-(2h-1)\}+2n\{\log_b \mathbb{E}(W^h)-(h-1)\}} \\
&= (\mathbb{E}(W^{2h}))^{-n} b^{n(2h-1)} (\mathbb{E}(W^h))^{2n} b^{-2n(h-1)} \\
&= \frac{(\mathbb{E}(W^h))^{2n}}{(\mathbb{E}(W^{2h}))^n} b^{2nh-n-2nh+2n} \\
&= \frac{(\mathbb{E}(W^h))^{2n}}{(\mathbb{E}(W^{2h}))^n} b^n
\end{aligned}$$

Now note that,

$$\begin{aligned}
\sum_{v:|v|=n} \lambda_n^2(h;\Delta_n(v)\cap A) b^{-n\chi_{b,h}(2)} &= \sum_{v:|v|=n} \left[\prod_{i=1}^n \frac{W_{v|i}^h}{\mathbb{E}(W^h)} \lambda(\Delta_n(v)\cap A)\right]^2 \frac{(\mathbb{E}(W^h))^{2n}}{(\mathbb{E}(W^{2h}))^n} b^n \\
\sum_{v:|v|=n} \prod_{i=1}^n \left(\frac{W_{v|i}^{2h}}{\mathbb{E}(W^{2h})}\right) b^n \lambda^2(A\cap\Delta_n(v)) &= \sum_{v:|v|=n} \lambda_n(2h;\Delta_n(v)\cap A)(b^n \lambda(\Delta_n(v)\cap A)) \\
&= \sum_{v:|v|=n} \lambda_n(2h;\Delta_n(v)\cap A) C_n(v,2;A)
\end{aligned}$$

So,

$$\begin{aligned}
\sum_{v:|v|=n} \lambda_n^2(h;\Delta_n(v)\cap A) \frac{b^{-n\chi_{b,h}(2)}}{C_n(v,2;A)} &= \sum_{v:|v|=n} \lambda_n(2h;\Delta_n(v)\cap A) \\
&= \lambda_n(2h;A) \to \lambda_\infty(2h;A)
\end{aligned}$$

Define,

$$\tilde{X}_n(v;A) = Z^h_\infty(v) - \frac{M_n(h;A)}{M_{n+1}(h;A)} \sum_{i=0}^{b-1} b^{-h} W^h_{v*i} Z_\infty(v*i) \tag{3.44}$$

Also note that by 3.5.1

$$\begin{aligned} \sqrt{\lambda_n(2h,\Delta_n(v)\cap A)} &= \sqrt{\frac{\lambda_n^{2h}(\Delta_n(v)\cap A)}{b^{n\chi_b(2h)} C_n(v,2h;A)}} \\ &= \frac{\lambda_n^h(\Delta_n(v)\cap A)}{b^{n/2\chi_b(2h)} C_n^{1/2}(v,2h;A)} \end{aligned}$$

Let

$$\tilde{X}_n(v;A) = Z^h_\infty(v) - \frac{M_n(h;A)}{M_{n+1}(h;A)} \sum_{i=0}^{b-1} b^{-h} W^h_{v*i} Z^h_\infty(v*i) \tag{3.45}$$

Now note that

$$M_n^2(h;A) D_n^2(h;A) b^{-n\chi_b(2h)} = \sum_{v:|v|=n} \tilde{X}_n^2(v) \lambda_n(2h;\Delta_n(v)\cap A) \to \tilde{c}_h^2 \lambda_\infty(2h,A) \tag{3.46}$$

where $\tilde{c}_h^2 = \mathrm{Var}(\tilde{X}_n(v))$

To see that the above equality holds, first note that, it would be enough to show that the square root of the individual summands are the same. To start with firs consider the square root of the summands of the second term in the above equation. This is given by:

$$\begin{aligned}
\tilde{X}_n(v)\sqrt{\lambda_n(2h;\Delta_n(v)\cap A)} &= \tilde{X}_n(v)\frac{\lambda_n^h(\Delta_n(v)\cap A)}{b^{n/2\chi_b(2h)}C_n^{1/2}(v,2h;A)} \\
\frac{Z_\infty^h(v)\lambda_n^h(\Delta_n(v)\cap A)}{b^{n/2\chi_b(2h)}C_n^{1/2}(v,2h;A)} &- \frac{M_n(h;A)}{M_{n+1}(h;A)}\sum_{i=0}^{b-1} b^{-h}W_{v*i}^h Z_\infty^h(v*i)\frac{\lambda_n^h(\Delta_n(v)\cap A)}{b^{n/2\chi_b(2h)}C_n^{1/2}(v,2h;A)} \\
&= \frac{1}{b^{n/2\chi_b(2h)}}\left[\frac{\lambda_\infty^h(\Delta_n(v)\cap A)}{C_n^{1/2}(v,2h;A)} - \frac{M_n(h;A)}{M_{n+1}(h;A)}\sum_{i=0}^{b-1}\frac{\lambda_\infty^h(\Delta_{n+1}(v*i)\cap A)}{\sqrt{C_n(v,2h;A)}}\right.
\end{aligned}$$

Now consider the square root of the summands of the first term $M_n(h;A)D_n(h;A)b^{-n/2\chi_b(2h)}$, given by :

$$b^{-n/2\chi_b(2h)}\left[\frac{\lambda_\infty^h(\Delta_n(v)\cap A)}{C_n^{1/2}(v,2h;A)} - \frac{M_n(h;A)}{M_{n+1}(h;A)}\sum_{i=0}^{b-1}\frac{\lambda_\infty^h(\Delta_{n+1}(v*i)\cap A)}{\sqrt{C_n(v,2h;A)}}\right] \tag{3.47}$$

Hence the above equality holds. Note that

$$\frac{M_n^2(h;A)D_n^2(h;A)C_n(v,2;A)}{b^{2n\chi_b(h)}\sum_{v:|v|=n}\lambda_n^2(h;\Delta_n(v)} \to \tilde{c}_h^2\mathbf{1}[\lambda_\infty(A)>0] \tag{3.48}$$

Now rewrite

$$\begin{aligned}
&\frac{M_n(h;A)/b^{n\chi_b(h)} - M_{n+1}(h;A)/b^{(n+1)\chi_b(h)}}{\tilde{c}_h\left(\sum_{v:|v|=n}\lambda_n^2(h;\Delta_n(v)\cap A)\right)^{1/2}} \\
= \quad & -\frac{M_n(h;A)D_n(h;A)}{b^{n\chi_b(h)}\tilde{c}_h\left(\sum_{v:|v|=n}\lambda_n^2(h;\Delta_n(v)\cap A)\right)^{1/2}}\left[\frac{(M_{n+1}(h;A)/M_n(h;A))b^{-\chi_b(h)}-1}{D_n(h;A)}\right]
\end{aligned}$$

to get the desired result. □

Remark 5. *The independence of the N_h's as h varies implies that the errors in the estimation of $\chi_b(h)$ by $\tilde{\tau}_n(h)$ are asymptotically independent for different values of h.*

Corollary 3.7.8. *For $h \in (H_c^-/2, H_c^+/2)$,*

$$\frac{\tilde{\tau}_n(h; A) - \chi_b(h)}{D_n(h; A)} \xrightarrow{d} ((\log b)^{-1} \eta N_h) \tag{3.49}$$

The proof follows by taking a logarithm in the previous corollary and Taylor approximation to the distributional limit.

Remark 6. *The above corollary gives the asymptotic distribution of a completely observable test statistic. Also note that the asymptotic distribution is completely independent of the cascade generating distribution, which is unobservable.*

3.8 Tesla Stock Volume: Data Analysis

3.8.1 Data description

The data on [Tesla stock market](https://www.kaggle.com/rpaguirre/tesla-stock-price) was obtained from Kaggle. The data contains the daily opening, closing, maximum, and minimum stock prices along with the stock volumes between June 29-th,2010, and March 17-th 2017. Note that not all days of the year are present in the dataset, for example, the data on weekends or holidays are not available. The total number of datapoints is 1692 in the dataset. We have focused on the stock

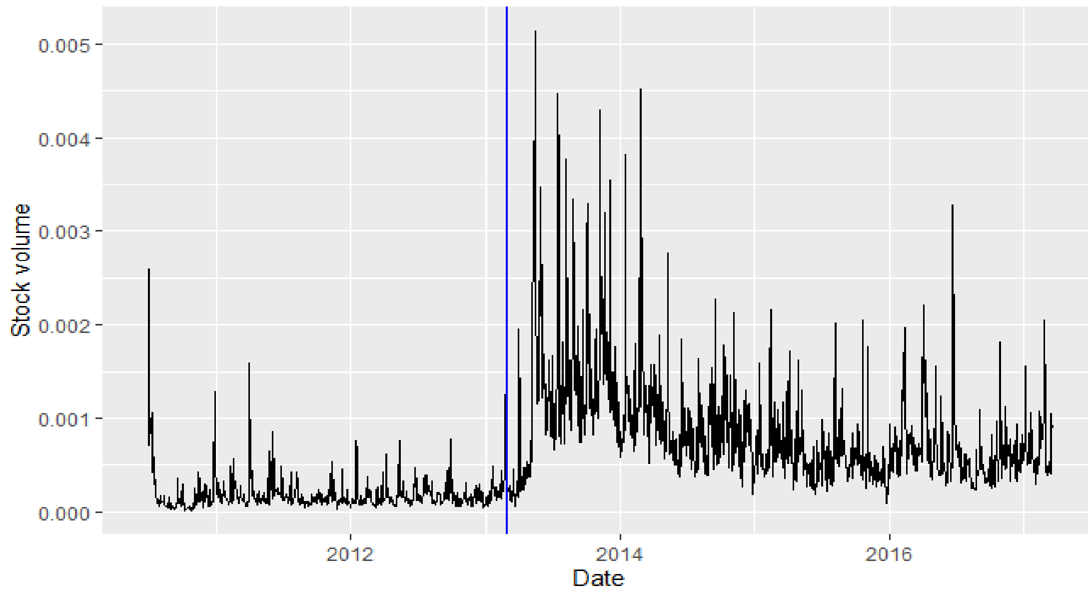


Figure 3.2: Tesla stock volume over time

volume data only and the absolute stock volumes were converted into fractional stock volumes so that the total amount of stock at any given time period is exactly 1.

Figure 3.2 shows the plot of the fractional stock volume of Tesla over time. We shall consider the last $2^{10} = 1024$ datapoints of this data. We shall start by estimating the structure-function from the data. This will requires us to calculate $M_n(h)$ and $M_{n+1}(h)$ for $n = 9$. However this enables us to simulate data even for $n \geq 9$, i.e. we can simulate on a coarser as well as a finer grid. This means even though the data is available at a daily level we can get predictions at the hourly level for example. We shall combine every consecutive 2 points and form a coarse grid to compute $M_9(h)$. Following is the structure-function estimated from the observed data. The estimated structure-function is denoted by $\tilde{\tau}_9(h)$.

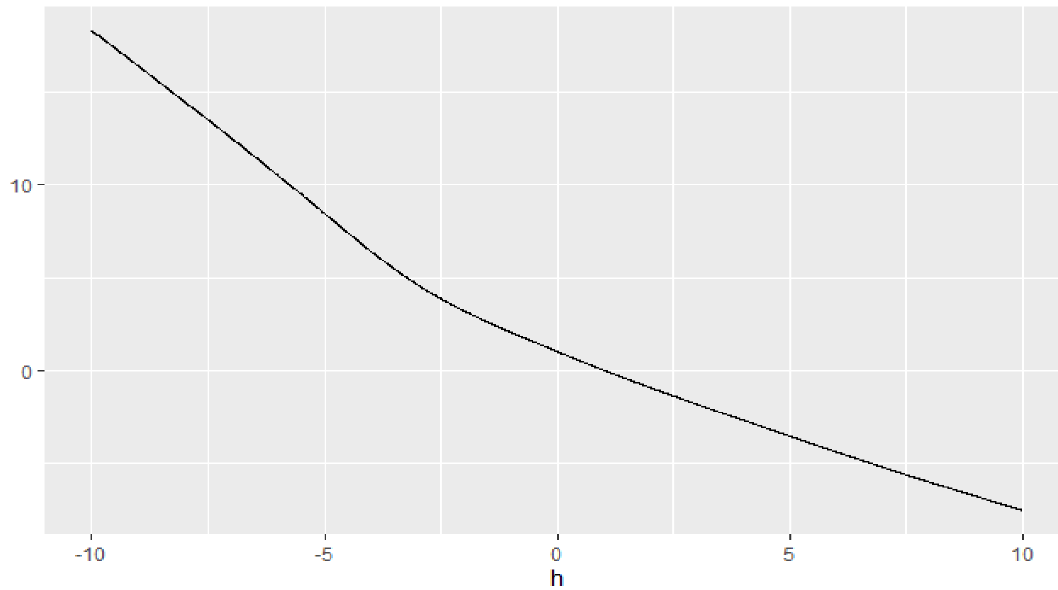


Figure 3.3: Tesla stock volume: Estimated structure function

Figure 3.3 shows the plot of the estimated structure-function $\tilde{\tau}_9(h)$.

3.8.2 log-Normal Cascade generating distribution

First, we are attempting to find the parameters of a log-Normal distribution for which the structure-function of the corresponding log-Normal random variable closely matches this estimated structure-function. To begin with, we shall derive the structure-function of a log-Normal random variable.

Suppose $Z \sim \text{Normal}(\mu, \sigma^2)$, with $Z = \log W$, i.e. W is the log-Normal random variable corresponding to the standard normal random variable Z. Then,

$$\begin{aligned}\mathbb{E}(W^h) &= \mathbb{E}e^{hZ} \\ &= \mathbb{E}\left(e^{h(Z-\mu)+h\mu}\right) \\ &= e^{h\mu+\frac{h^2\sigma^2}{2}}\end{aligned}$$

We want the cascade generating random variable to have mean one, i.e. $\mathbb{E}(W) = 1$. Thus we must have $\mu + \frac{\sigma^2}{2} = 0$, i.e. $\mu = -\frac{\sigma^2}{2}$. Hence the structure function of the log-Normal random variable will be given by:

$$\begin{aligned}\chi_b(h) &= \log_b(e^{\frac{h^2\sigma^2}{2}-\frac{h\sigma^2}{2}}) - (h-1) \\ &= \frac{\sigma^2 h(h-1)}{2\ln b} - (h-1) \\ &= (h-1)\left(\frac{\sigma^2 h}{2\ln b} - 1\right)\end{aligned}$$

Let's denote $k = \frac{\sigma^2}{2\ln b}$.

After trying multiple values of k, we have come to the conclusion that $k = .055$ corresponds to the log-Normal random variable the structure-function of which matches the estimated one most closely on (H_c^-, H_c^+). The next plot shows the structure-function estimated from the data and the structure-function of a log-Normal random variable with $k = .055$.

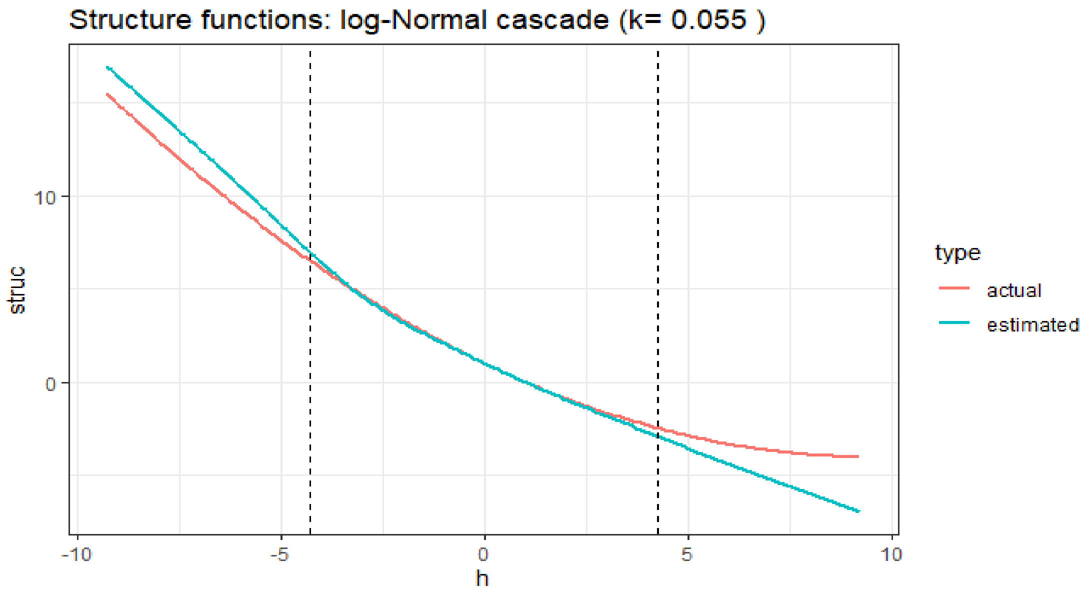


Figure 3.4: Estimated and actual(log-Normal with $k = .055$) structure function

From Figure 3.4, we can see that these two structure functions don't match very well towards the ends of the critical interval given by (H_c^-, H_c^+). This suggests that log-Normal distribution is not a good choice of a cascade generating distribution for this data.

Also as the standard practice in physics literature suggests, we have tried to find out a k for which the structure-function of the log-Normal variable is closest to the estimated structure-function on $(0, H_c^+)$. $k = 0.02$ seems to be the best fit in this case. Figure 3.5 shows the structure-function of a log-Normal random variable with $k = 0.02$ along with the estimated structure-function.

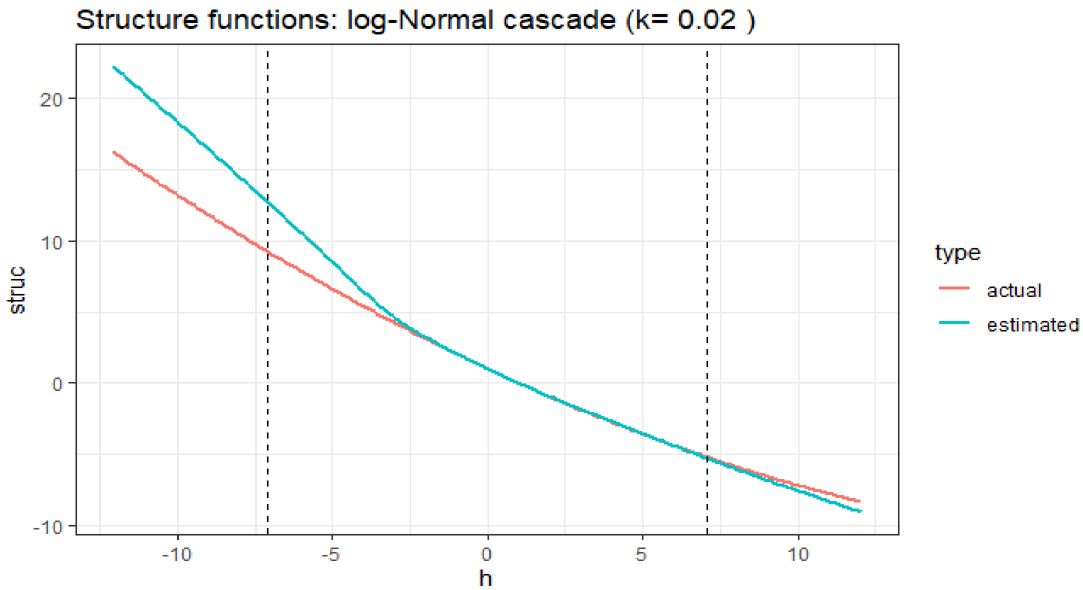


Figure 3.5: Estimated and actual(log-Normal with $k = .02$) structure function

Remark 7. *Note that the main theorem suggests that we should look for the random variable, for which the structure-function matches the estimated structure-function on the entire critical interval. But since there are multiple instances in the physics literature where the comparison was made only on the positive real line, therefore I decided to consider both and try to see if there is any significant difference in these two models and/or to figure out the reason behind such discrepancies.*

3.8.2.1 Observations from log-Normal cascade

Figure 3.6 shows the actual data and a single sample realization from the random multiplicative cascade model where the parameter of the log-Normal random variable is $k = 0.05$ and $k = 0.02$ respectively.

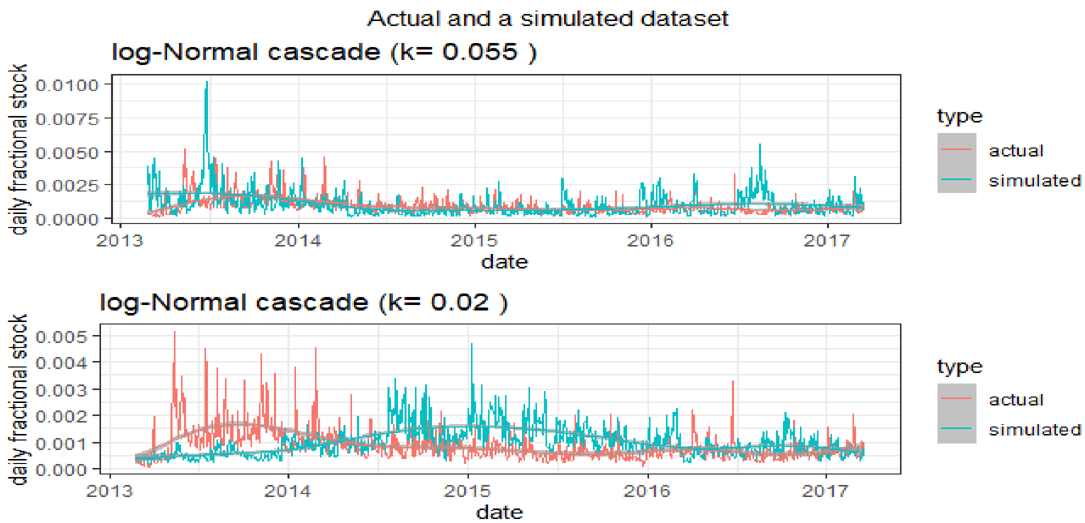


Figure 3.6: Actual data and a single sample realizations from log-Normal cascade

Figure 3.7 shows the respective histograms.

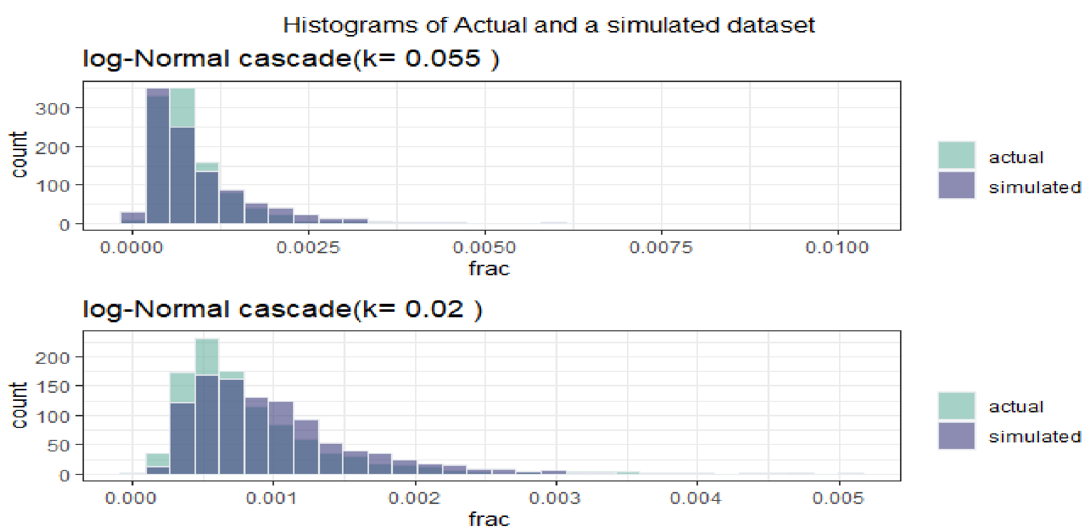


Figure 3.7: Histogram of Actual data and a single sample realizations from log-Normal cascade

At this point, we have pulled together the data from multiple (1000) realizations from a random multiplicative cascade model and the histogram of the combined data for both log-Normal models are shown in Figure 3.8.

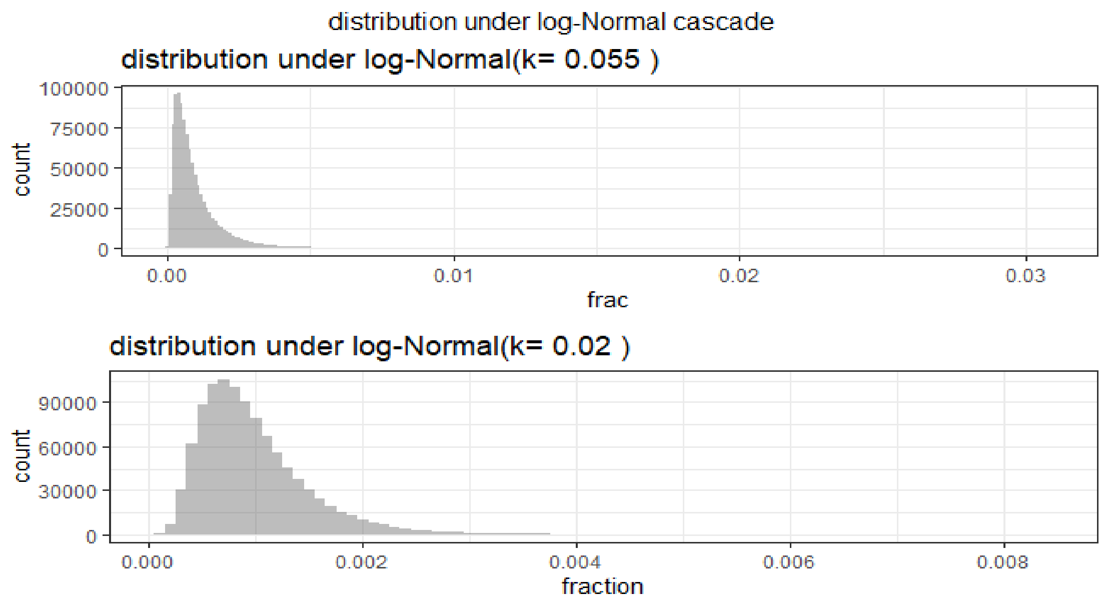


Figure 3.8: Histograms of 1000 realizations from log-Normal cascade

3.8.3 log-Poisson cascade generating distribution

Now we shall try to see if a log-Poisson cascade generating distribution gives a better fit to the data as compared to log-Normal. First note that W follows a log-Poisson distribution if it is distributionally identical to $b^{2/3}\beta^{Y}$ where $Y \sim \text{Poi}(2\ln b/3(1-\beta)$. The structure-function of a log-Poisson random variable is given by

$$\chi_b(h) = (2\beta^h - (1-\beta)h + 1 - 3\beta)/3(1-\beta)$$

We have found that $\beta = .85$ gives the closest match to the estimated structure-function. Figure 3.9 is a plot of the structure-function estimated from the data and the actual structure function corresponding to a log-Poisson random variable with $\beta = .85$.

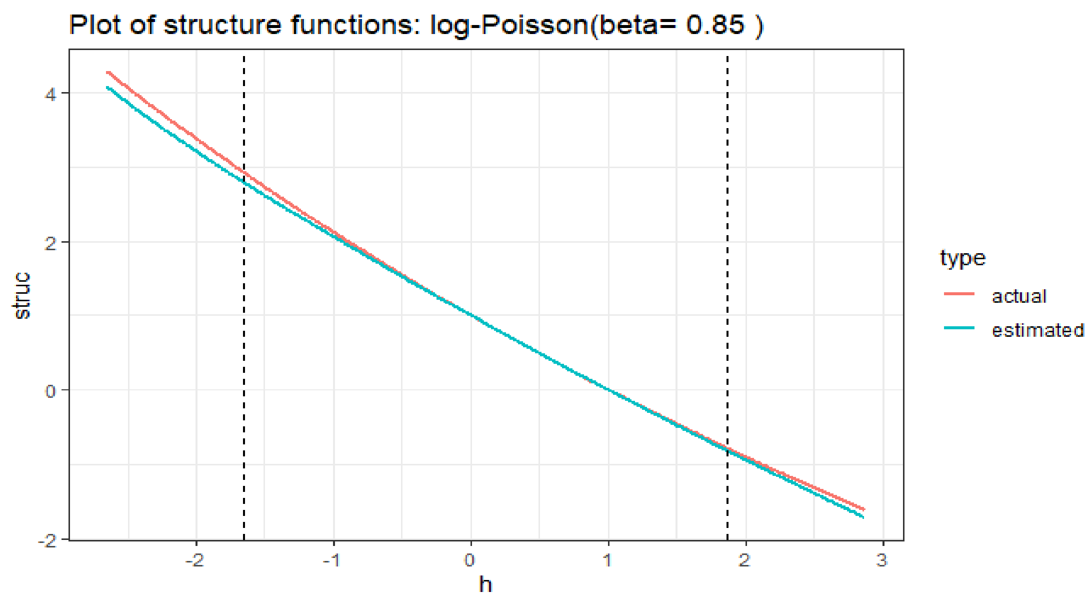


Figure 3.9: Estimated and actual log-Poisson($k = 0.85$) structure function

From the visual comparison, log-Poisson seems to be a better fit to the data as compared to log-Normal.

3.8.4 Observations from log-Poisson cascade

The plots of an actual and a single sample realization from a log-Poisson cascade model are shown in Figure 3.10

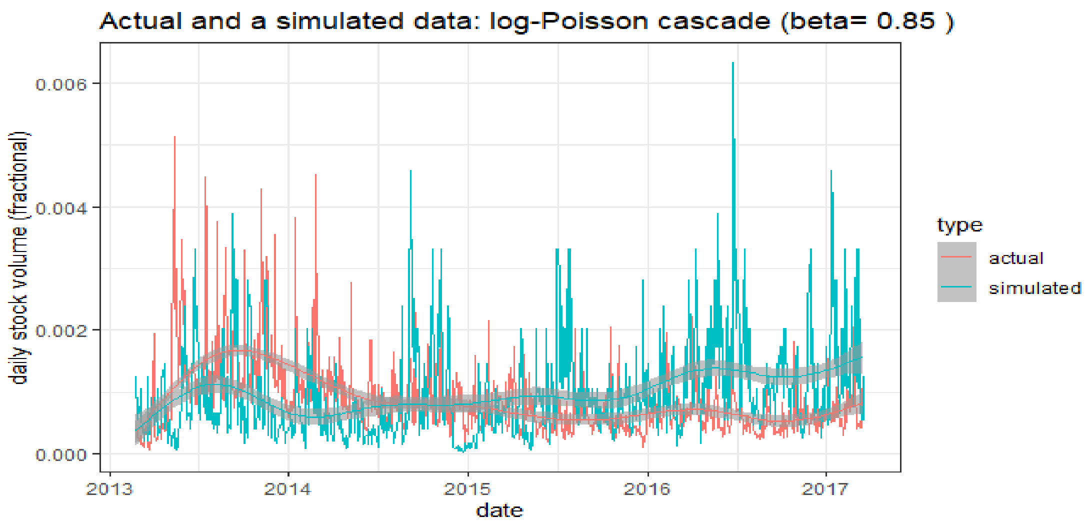


Figure 3.10: Actual data and a single realization from the log-Poisson cascade

Figure 3.11 shows the histograms for the actual and a single realization from the log-Poisson cascade model.

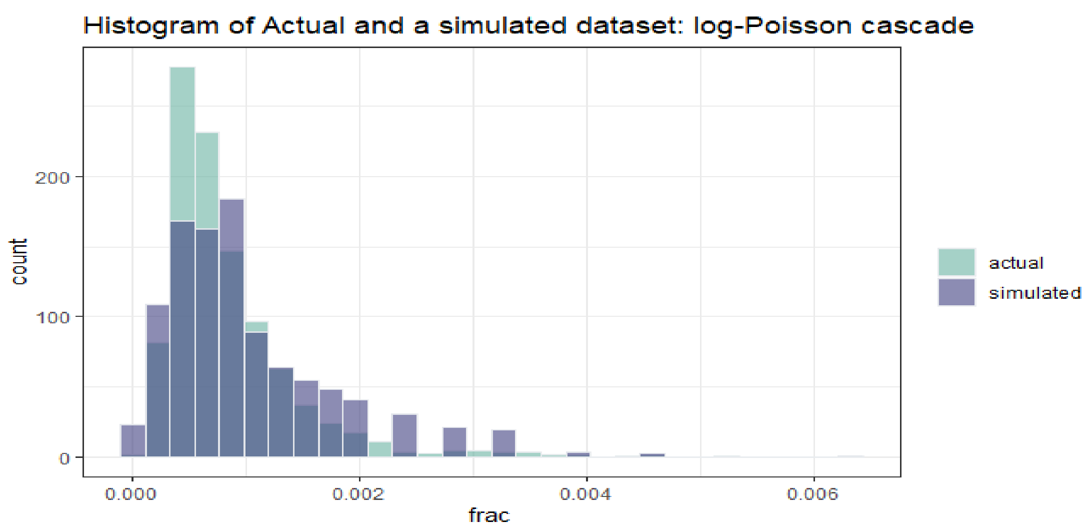


Figure 3.11: Histogram of actual data and a single realization from the log-Poisson cascade

We have simulated 1000 datasets using the above-mentioned cascade generating random variable. The histogram of the combined data is shown in Figure 3.12.

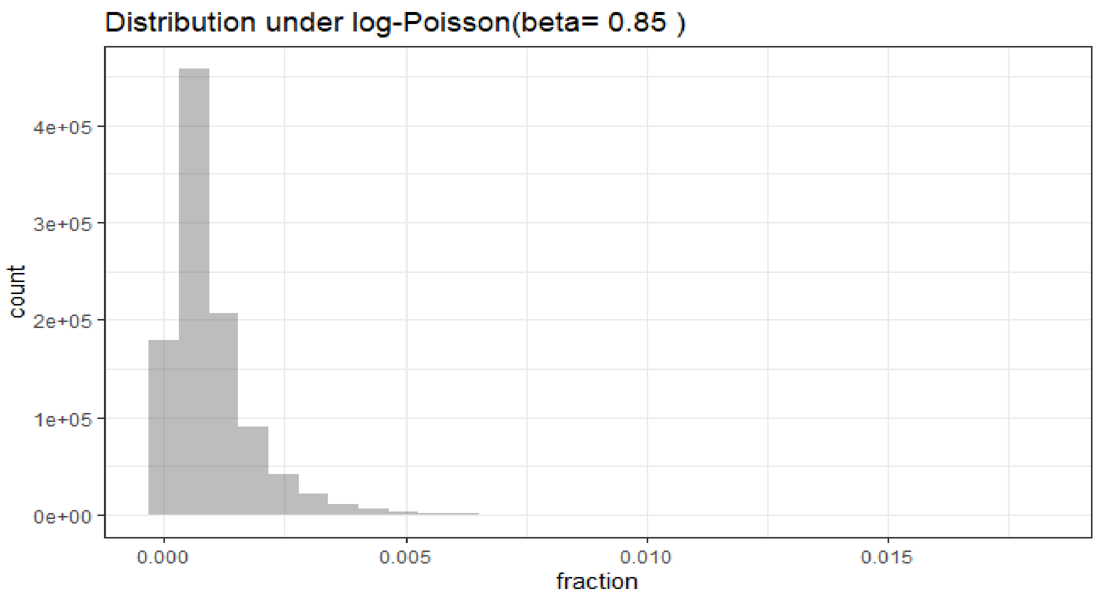


Figure 3.12: Histogram of 1000 realizations from the log-Poisson cascade

3.8.5 Beta cascade generating distribution

In this section, we shall consider the cascade generating distribution to be a Beta random variable with parameters p and $(b-1)p$, so that the mean of the beta distribution is given by $\frac{1}{b} = 0.5$. In this case, $b = 2$, therefore the cascade generating distribution will be given by $W = 2 * \text{Beta}(p, p)$, so that the mean of the cascade generating random variable is 1. All we need to do is to find the beta parameter p for which the structure-function of W closely matches the estimated structure-function. We shall start by deriving the structure-function for W.

$$\begin{aligned}
\chi_b(h) &= \log_b\left(\mathbb{E}((2W)^h)\right) - (h-1) \\
&= \log_b\left(2^h \frac{\Gamma(p+h)}{\Gamma(2p+h)} \frac{\Gamma(2p)}{\Gamma(p)}\right) - (h-1) \\
&= h + \log_b\left(\frac{\Gamma(p+h)}{\Gamma(2p+h)} \frac{\Gamma(2p)}{\Gamma(p)}\right) - (h-1) \\
&= \log_b\left(\frac{\Gamma(p+h)}{\Gamma(2p+h)} \frac{\Gamma(2p)}{\Gamma(p)}\right) + 1
\end{aligned}$$

Now in order to find the limits of the critical interval given by (H_c^-, H_c^+) we need to solve the following equation:

$$h\chi_b'(h) - \chi_b(h) = 0$$

In this case

$$\chi_b'(h) = (\log_2 e) * (\psi(p+h) - \psi(2p+h)) \tag{3.50}$$

where $\psi(.)$ is the digamma function. This equation reduces to the following:

$$h \log_2 e * (\psi(p+h) - \psi(2p+h)) = \log_2\left(\frac{\Gamma(p+h)}{\Gamma(2p+h)} \frac{\Gamma(2p)}{\Gamma(p)}\right) + 1 \tag{3.51}$$

For a fixed value of p, we need to solve this equation to get H_c^- and H_c^+. This can't be solved analytically. We have used the 'uniroot' function in R to solve it numerically. Figure 3.13 shows the plot of estimated and actual structure-function within the critical interval for different values of p.

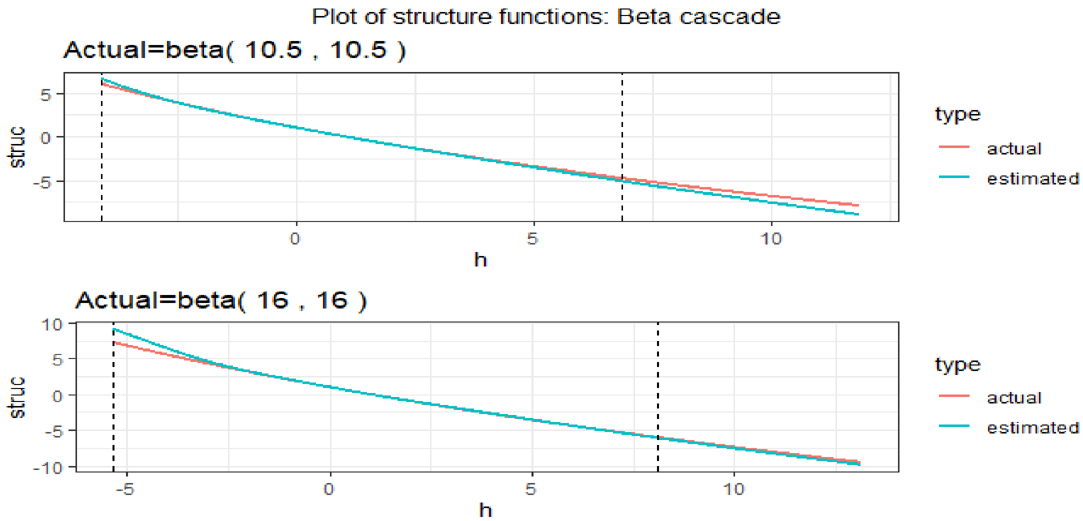


Figure 3.13: Estimated and actual (Beta) structure function

The histogram of 1000 simulations from both of the Beta cascades are shown in Figure 3.14.

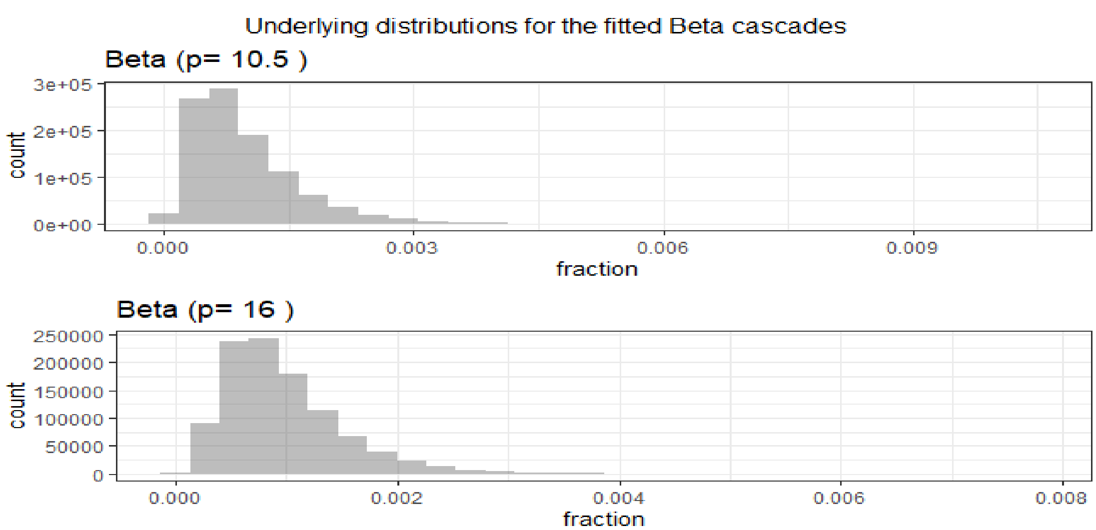


Figure 3.14: Histogram of multiple realizations from Beta cascades

3.9 Impact of Missing data in estimation

Now we shall consider the case when the entire data is not observed at a particular resolution. Rather there are missing values in the dataset. We shall try to see how much does the proportion of missing data impact the model, i.e. the change in the value of the parameter of the cascade distribution for different amounts of missing values. Here we are assuming that the data is missing systematically, i.e. the missing data pointsare a systematic sample from the original data points, namely from 1 : 1024. Note that if the missing data pointsare not a systematic sample of the original dataset, then we have to consider the resolution b^n, where n is such that $C_n(v, h; A) > 0$ for all v such that $|v| = n$.

To begin with, we shall assume that a certain percentage of total observations are missing. Figure 3.15 shows the estimated structure functions for different proportions of missing data.

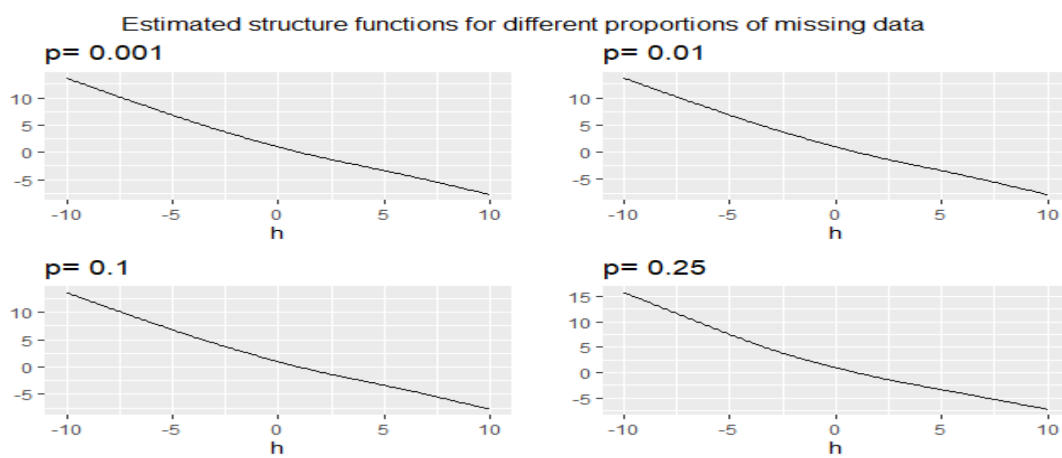


Figure 3.15: Estimated structure functions for data with missing values

3.9.1 log-Normal cascade

Figure 3.16 shows the estimated structure functions and the structure functions of the nearest log-Normal random variables for different proportions of missing data.

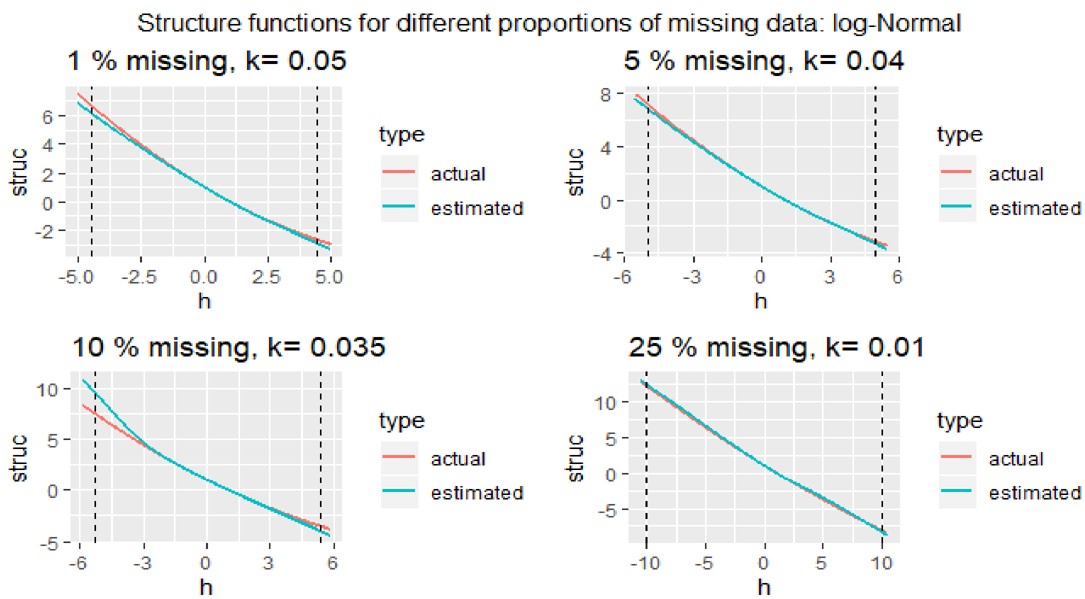


Figure 3.16: log-Normal structure functions for data with missing values

Note that as the proportion of missing observations has increased, the value of k has gone further away from the true value of k, under complete observations.

3.9.2 log-Poisson cascade

Now we shall try to figure out the value of the log-Poisson parameter β, for which the actual and estimated structure functions are closest to each other. The value of structure-function along with the actual log-Poisson structure functions are shown in Figure 3.17.

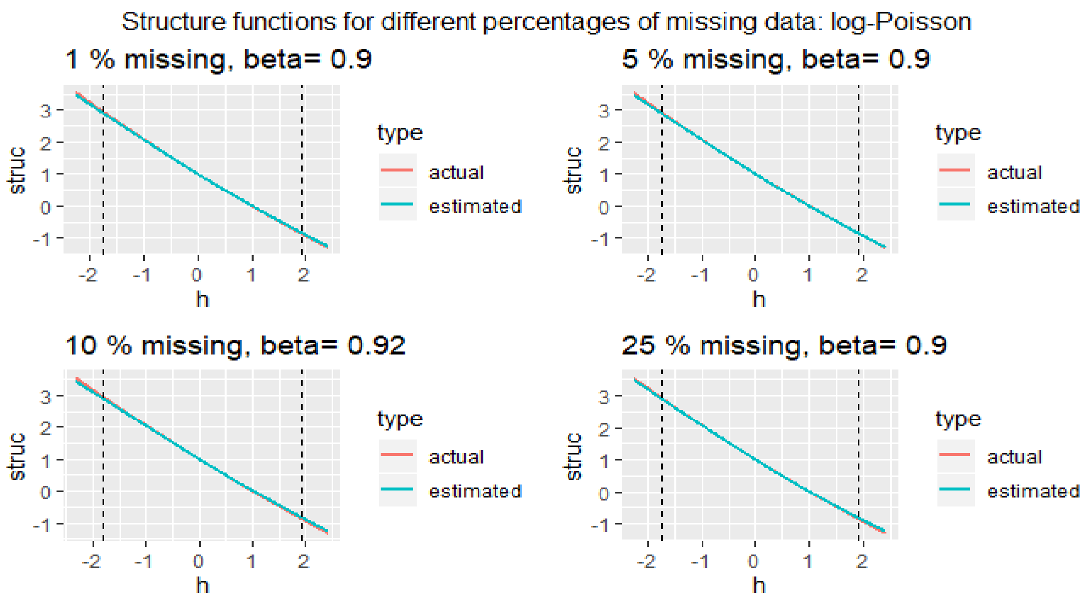


Figure 3.17: log-Poisson structure functions for data with missing values

Note that for different proportions of missing data the estimated value of the cascade parameter, i.e. β remains more or less the same. Thus the approximation works really well in this case.

3.9.3 Beta cascade

Now we shall consider the case when the underlying distribution is Beta. In that case, we shall try to find the parameter of the beta distribution, for which the structure-function is closest to the ones estimated from the data with different percentages of missing data. Figure 3.18 show the plots of estimated structure functions for different proportions of missing data, along with the structure functions of the beta distributions that are closest to the estimated ones.

Figure 3.18: Beta structure functions for data with missing values

The plots suggest that when the proportion of missing data increase, the estimated value of p, the parameters of beta distribution becomes more different from the one estimated in the complete data case.

3.9.4 Comparison

In this part, we shall compare the histograms of overall stock volumes in case of complete data and in case of data with 25% missing values. Figure **??** shows the side by side plots in all three cases, namely- log-Normal, log-Poisson and Beta.

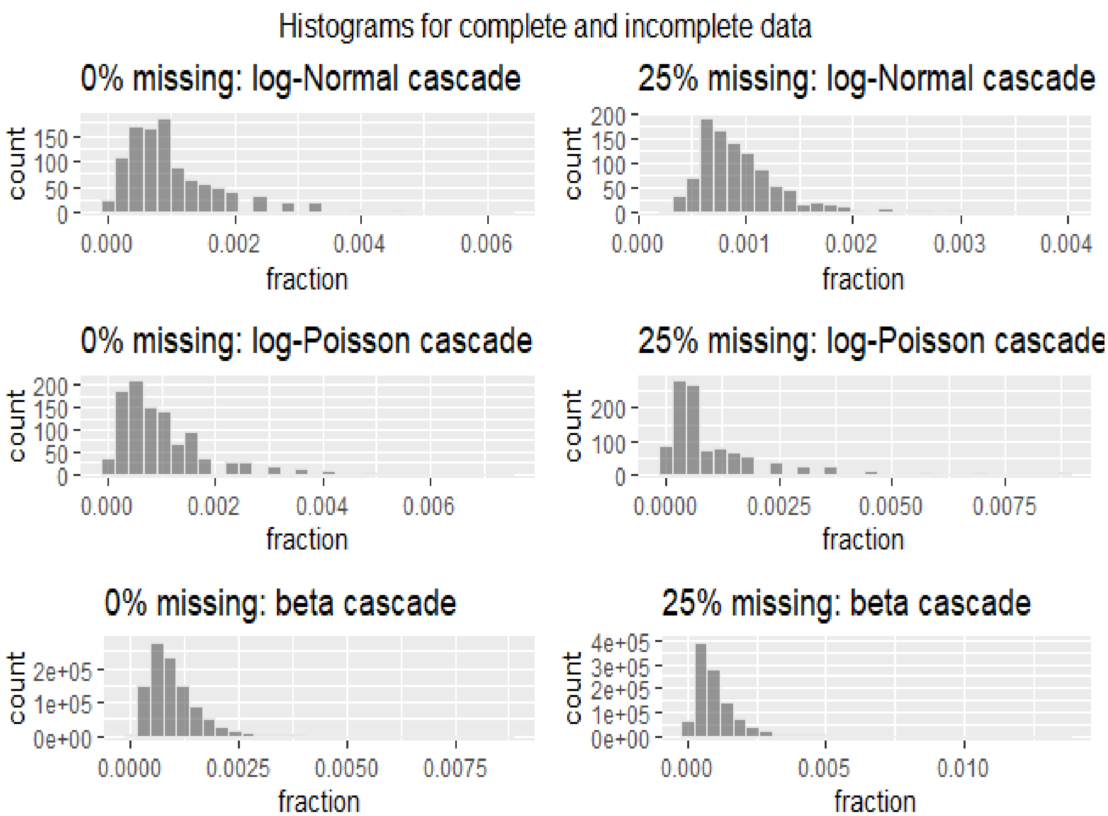


Figure 3.19: Histograms for data with and without missing values

Note that change between complete and incomplete data is less in the case of log-Poisson and Beta, as compared to log-Normal. However, in all three cases, the histograms in the case of complete and incomplete datasets are not very different. Thus our suggested estimate of structure-function in case of missing data performs fairly well in general.

3.10 Distribution of stock volumes before and after the change point

As can be seen from Figure 3.2, there is a change point in the stock-volume data, before and after which the pattern seems to have changed. Our goal is to see how different the underlying distributions are before and after this change point. The following plot shows the plot of the entire data, with the change point marked by a blue vertical line.

We shall analyze 512 data points before the change point and 512 datapoints after the change point separately. Then we shall obtain the estimates of the underlying cascade generating distribution in both cases. Figure 3.20 shows the stock volume for the 512 datapoints both before and after the change point, i.e. 2/25/2013.

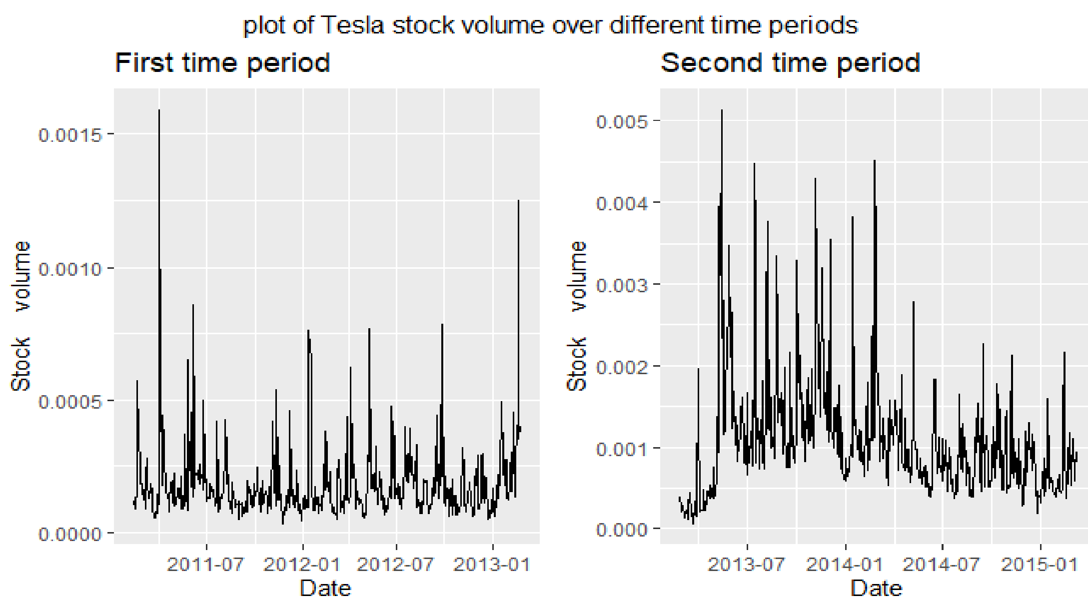


Figure 3.20: Stock volume of Tesla before and after the changepoint

Clearly, these two series are inherently different. The series before the changepoint has no specific overall trend, whereas the series after the changepoint has a quadratic change. Also, the mean value of the series on the left is lower compared to the one on the right.

3.10.1 Data before changepoint

In this part of the analysis, we shall consider the data between 2/10/2011 and 2/25/2013. The data consists of 512 data points. Figure 3.21 shows the structure-function as estimated from the data.

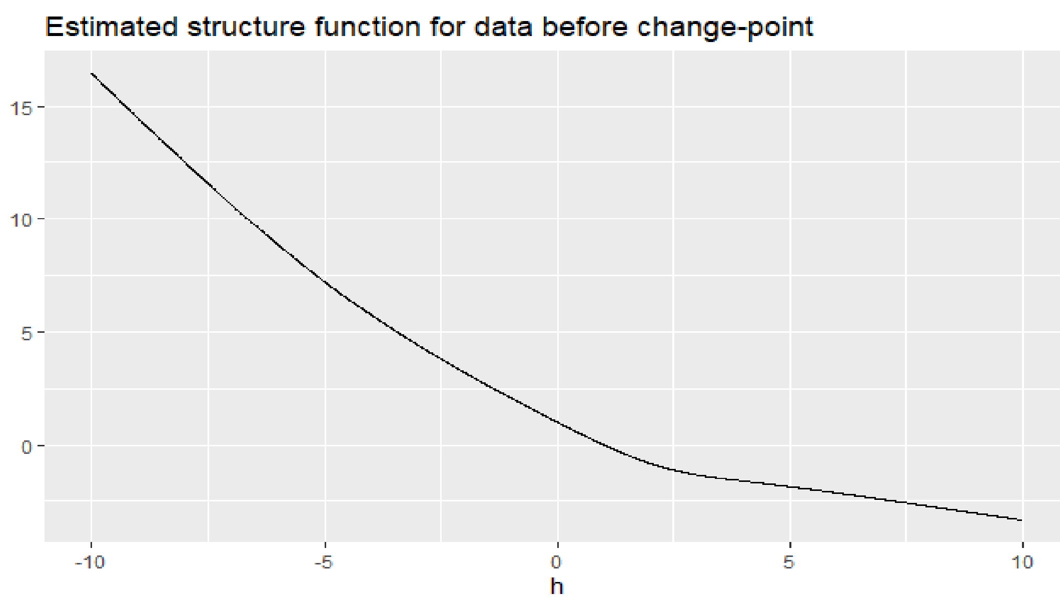


Figure 3.21: Estimated structure function from the data before the changepoint

The structure functions corresponding to the different cascade generating distribution along with the estimated structure function are shown in Figure 3.22

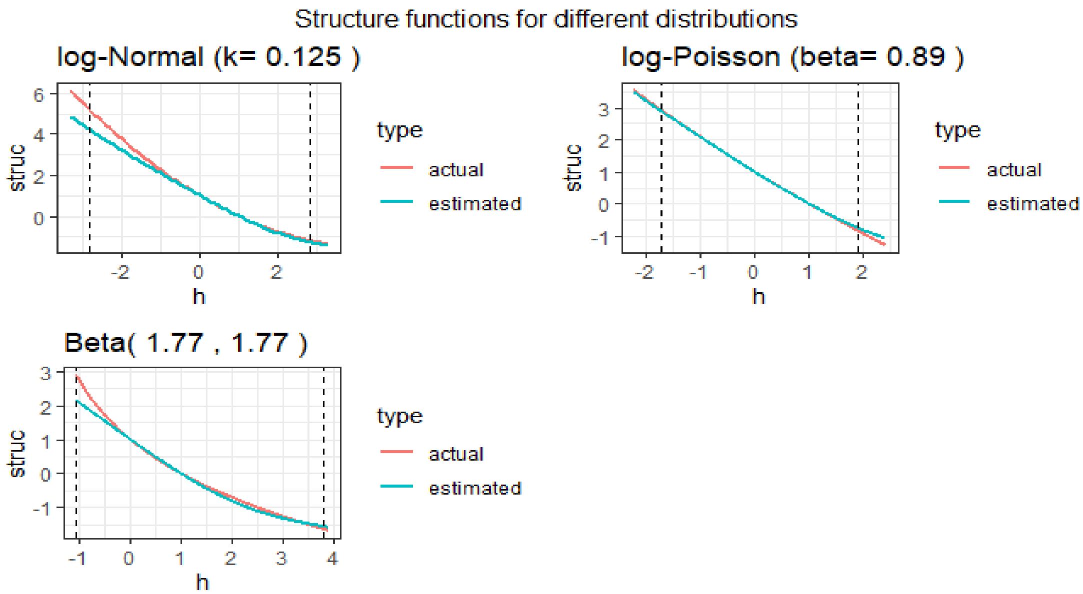


Figure 3.22: Different distributions and structure functions for the data before the changepoint

The log-Poisson distribution with $\beta = 0.89$ seems to be the best fit to the data.

3.10.2 Data after change-point

In this part of the analysis, we shall consider the data between 2/26/2013 and 3/9/2015. The data consists of 512 data points. Figure 3.23 shows the structure-function as estimated from the data.

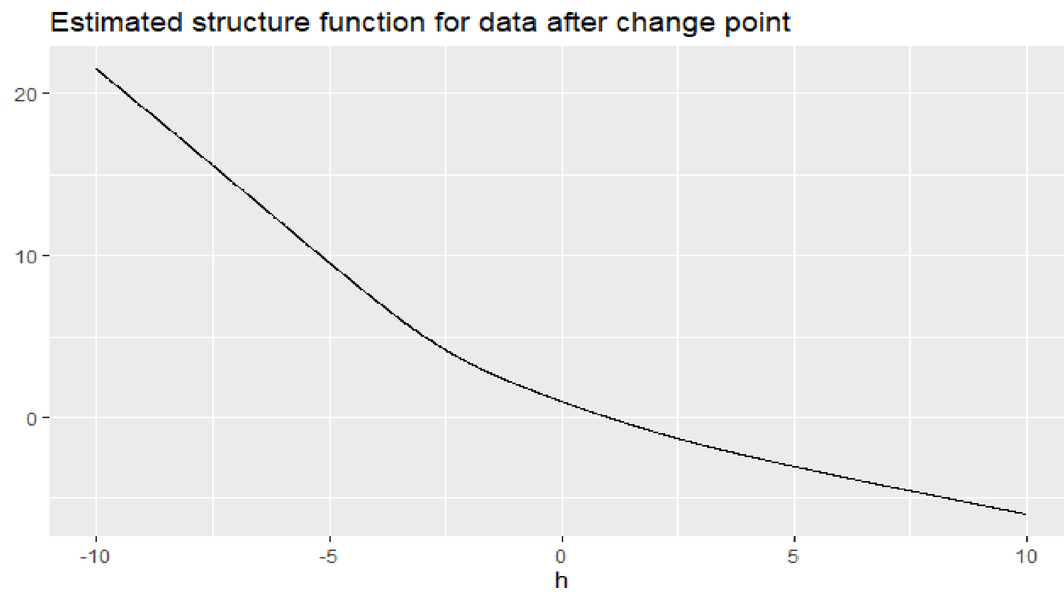


Figure 3.23: Estimated structure function for the data after changepoint

The structure functions corresponding to the different cascade generating distribution along with the estimated structure function are shown in Figure 3.24

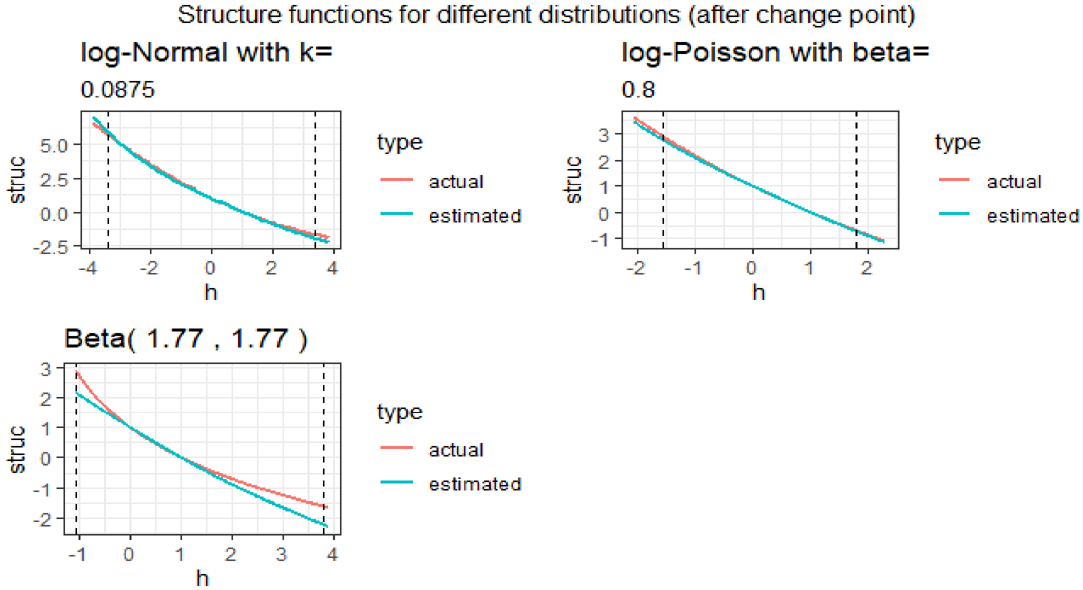


Figure 3.24: Different distributions and structure functions for data after change-point

In this case, the log-Normal with $k = .0875$ seems to be the best fit for the data. Figure 3.25 shows the histograms of 1000 simulations for the data before and after the changepoint.

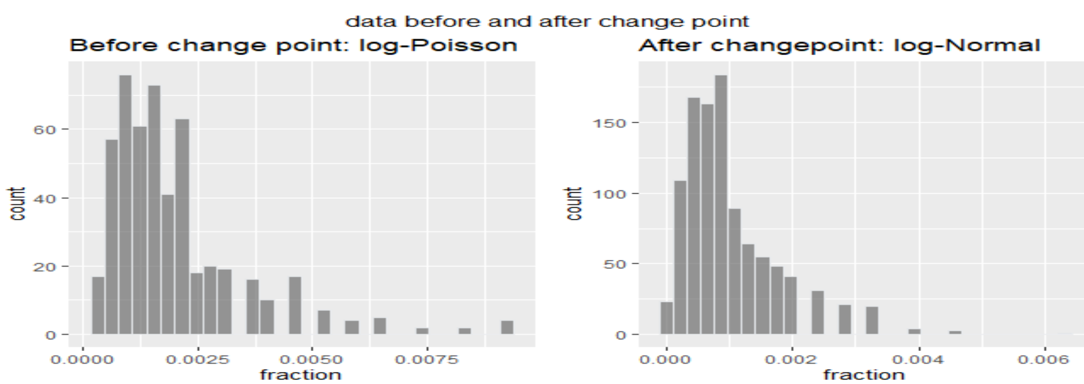


Figure 3.25: Histograms corresponding to data before and after the changepoint

Clearly these two distributions are different. Unlike most changepoint detection methods, in this case, we can see the difference in the overall distribution and not just the difference in mean or variance. To have an idea about how different the underlying distributions are, we have shown the histograms corresponding to the two cascade generating distributions in Figure 3.26.

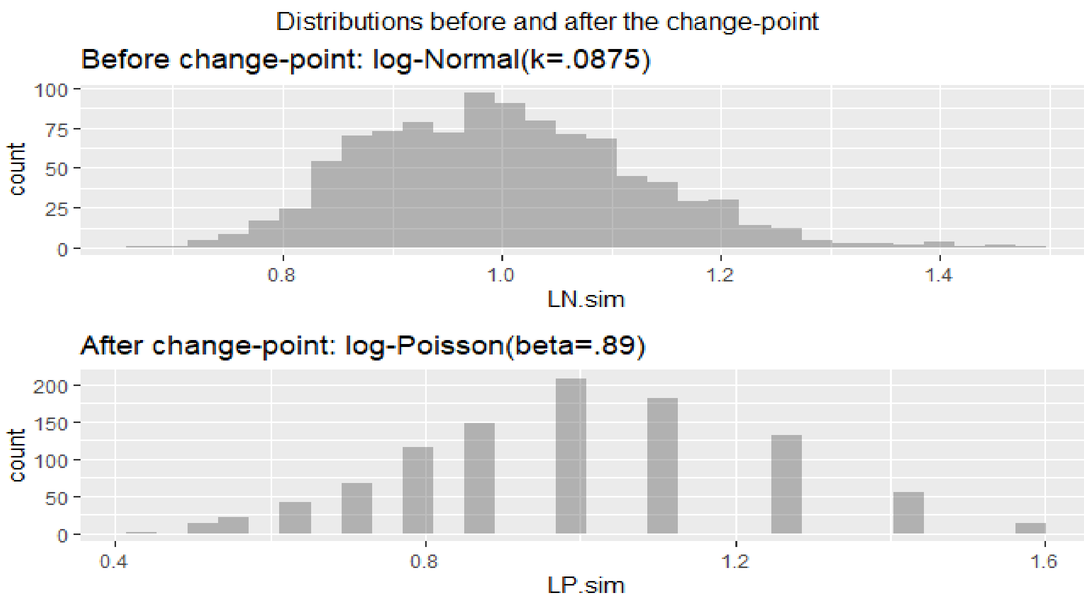


Figure 3.26: Distributions of the cascade generating random variables before and after the changepoint

3.10.3 A changepoint detection method

The random multiplicative cascade models can be used to build a change point detection method. That requires a uniform central limit theorem for the estimated structure-function. This will enable testing the equality of structure-function before and after the changepoint. However, the following naive approach can also be used to accomplish the goal. We can start by considering the estimated structure functions before and after the changepoint on a fine grid. Then for each point on the grid, we can do multiple hypothesis tests. And based on the test result we can decide whether or not the changepoint is truly a changepoint.

To make this idea concrete, let's consider the case, where the data before change point are denoted as $\lambda_\infty(\Delta_n(v))$ and data after change point are denoted as $\lambda_\infty(\Delta_n(u))$. Let the estimated structure functions before and after change point are given by $\tilde{\tau}_{n,1}(h)$ and $\tilde{\tau}_{n,2}(h)$ respectively. In order to check if the change point is truly a change point, we need to test

$$H_{0,h} : \chi_{b,1}(h) = \chi_{b,2}(h) \text{ vs } H_{1,h} : \chi_{b,1}(h) \neq \chi_{b,2}(h) \text{ for } h \in H$$

where H is a grid and $\chi_{b,1}(h)$ and $\chi_{b,2}(h)$ correspond to the true structure functions before and after the changepoint. Note that under the null both must have the same mean, i.e. $\chi_b(h)$. The null distribution of the test statistics given by $T_{\text{stat}}(h) = \tilde{\tau}_{n,1}(h) - \tilde{\tau}_{n,2}(h)$ will be normal with mean zero and variance, that can be determined from Corollary (3.7.8). Thus for each h on a grid we can perform the test. Suppose the p-values obtained in this way are given by $\alpha_1, \ldots, \alpha_n$, then after doing a level correction (for example: Boneferroni, Holm, Scheffe) we can obtain the test results, i.e. for each value of h on the grid whether or not the test results in a rejection. If any one of these tests result in a rejection, we shall conclude that the structure functions before and after the chanegpoint are different.

By Corollary (3.7.8)

$$\frac{\tilde{\tau}_{n,1}(h) - \chi_{b,1}(h)}{D_{n,1}(h)} \xrightarrow{d} (\log b)^{-1} \eta N_h$$

and

$$\frac{\tilde{\tau}_{n,2}(h) - \chi_{b,2}(h)}{D_{n,2}(h)} \xrightarrow{d} (\log b)^{-1} \eta N_h$$

Under the null,

$$\chi_{b,1}(h) = \chi_{b,2}(h)$$

and we assume that under null, the data before and after the change poimt are independent. Thus the distribuion of the test statistics under the null will be given by:

$$\tilde{\tau}_{n,1}(h) - \tilde{\tau}_{n,1}(h) \xrightarrow{d} N\left(0, (\log b)^{-2} \eta N_h \left[D_{n,1}^2(h) + D_{n,2}^2(h)\right]\right)$$

Thus under standard set up the test statistics can be writen as:

$$T_{stat}(h) := \frac{\tilde{\tau}_{n,1}(h) - \tilde{\tau}_{n,1}(h)}{(\log b)^{-1} \sqrt{D_{n,1}^2(h) + D_{n,2}^2(h)}} \xrightarrow{d} N(0,1)$$

for h in the intersection of the two critical intervals Thus the test results in a z-test.

3.11 Discussion

The methodology that we have developed here can be used for modeling datasets that include missing values. In this case, we have considered that the locations of missing datapoints are a systematic sample from the original data. However, this is not a required condition. In general the proposed method holds as long as the normalizing constant given by $C_n(v, h; A)$ is non-zero for every pixel $\Delta_n(v)$. In the above data analysis, we have used the Random Multiplicative Cascade model to analyze the variable intermittency in the daily stock volume data of Tesla. In each of the scenarios, the plot of structure-function was obtained from the data. Then the estimated structure-function was compared to the structure functions of candidate cascade generating distributions. The one that is closest to the estimated one, is considered to be the actual cascade generating random variable, that controls the data generating mechanism. Here the observed data is considered to be a single realization from the data generating procedure. Therefore in order to understand the underlying distribution of the data, it is important to simulate a good number of datasets and then aggregate those. Some of the shortcomings of this modeling approach are as follows:

- Here we are using binary cascade. Therefore the number of data points has to be of the form 2^n for $n \geq 1$.
- The Random Multiplicative Cascade model facilitates the understanding of the physical system. However, the model is not meant to be used for prediction. There have been some recent developments in this direction though.

- The choice of the actual cascade generating distribution is based on a visual check and thus subjective. Having a clear guideline for choosing the write distribution family will be helpful for applied scientists.

4 A central limit theorem for a set-indexed partial sum process

4.1 Abstract

In this chapter, we shall consider a partial sum process, indexed by sets. The indexing set is assumed to belong to a suitable collection of sets, say $\mathcal{A}$. We assume that the summands or increments of the process form a martingale difference sequence with respect to a filtration. We shall study the weak limit of the set indexed partial sum process, derive a central limit theorem in this context, and mention an application of the process in the context of Markovian multiplicative error models for non-negative valued time series.

4.2 Introduction

The study of empirical processes theory dates back to 1930's when the study of the empirical distribution function began. For a set of independent and identically distributed real-valued random variables $X_1, X_2, \ldots X_n$ with distribution function $F(.)$, the empirical distribution function is given by

$$\hat{F}_n(x) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \leq x)$$

The study of the limiting behavior of empirical distribution function is of great importance because of its vast applicability in proving asymptotic results of various types. Two key results in the theory of empirical processes are the Glivenko-Cantelli lemma and Donsker's theorem. The first one gives the uniform convergence of empirical distribution function to the true distribution function and the second one derives a distributional limit for the process $\{\hat{F}_n(x) : x \in \mathbb{R}\}$.

The generalization of these two fundamental results is the primary motivation behind the study of empirical processes. When the random observations X_i's take value in a more general space (than $\mathbb{R}$), then the idea of empirical distribution function is not so natural anymore. Also the definition of in distribution convergence is not well-defined in this context. So it makes more sense to define an empirical measure $\mathbb{P}_n$ defined on a collection of functions $\mathcal{F}$ defined on a general space, say χ and taking values in $\mathbb{R}$. Let $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathbb{P}$, taking values in χ. Then the empirical measure $\mathbb{P}_n$ is defined as:

$$\mathbb{P}_n := \frac{1}{n}\delta_{X_i}$$

where $\delta_x(y) = 1$ if $x = y$ and zero otherwise. Thus $\mathbb{P}_n$ puts equal mass at each observed data point. For any Borel set $A \subset \chi$,

$$\mathbb{P}_n(A) := \frac{1}{n}\sum \mathbf{1}(X_i \in A)$$

Let $\mathcal{F}$ is a collection of real-valued functions defined on χ. Then for any $f \in \mathcal{F}$,

$$\mathbb{P}_n f := \int f(x) d\mathbb{P}_n(x) = \frac{1}{n} \sum_{i=1}^{n} f(X_i)$$

Assuming $\mathbb{P}_f := \int f(x) d\mathbb{P}(x)$ exists for each $f \in \mathcal{F}$, let's define the empirical process $\{\mathbb{G}_n(f) : f \in \mathcal{F}\}$ as follows

$$\mathbb{G}_n(f) := \sqrt{n}(\mathbb{P}_n f - \mathbb{P} f)$$

The goal is to figure out conditions under which the uniform convergence holds, i.e.

$$\sup_{f \in \mathcal{F}} |\mathbb{P}_n f - \mathbb{P} f| \xrightarrow{a.s.} 0$$

and a distributional limit exists for the process given by $\{\mathbb{G}_n(f) : f \in \mathcal{F}\}$. It has been found that the conditions under which the Glivenko-Cantelli or Donsker's theorem hold depend on the complexity or size of the set $\mathcal{F}$. For detailed theory on this see Sen (2018), Wellner (2005).

Note that Donsker's theorem can be viewed as a generalization of the central limit theorem(CLT). There are several versions of the central limit theorem that corresponds to different moment conditions under which a sum of independent and identically distributed random variables converges to a normal random variable. Examples include the classical CLT, Lyapunov's CLT and Lindeberg's CLT. Details on these can be found in Billingsley (2008). There exist generalizations of these in multiple directions in the context of a dependent sequence of random variables, the most common one being in the case of martingales. Related results can be found in Brown et al. (1971) and Hall and Heyde (2014).

In Slonowsky (1998) the author has considered a vectorized version of our process, under Lyapunov type assumption and different types of complexity assumption on the indexing sets. In Koul et al. (2012) the authors have constructed a partial sum process with dependent increments which serves as a test statistic of interest. However, the process is indexed by sets of the form $(-\infty, x)$. We have considered the partial sum process indexed by a more general collection of sets. We have derived the Donsker's theorem in this context based on the metric entropy integrability condition, which quantifies the complexity of the underlying collection of sets. The central limit theorem is obtained under a Lindeberg type condition, which is more general than the Lyapunov type conditions. We have also derived a Glivenko-Cantelli type result in this context. We have used exponential inequalities for martingales that can be found in de la Pena et al. (1999) and Freedman (1975).

The organization of this chapter is as follows. Section 4.3 describes the setup and introduces the notation related to the problem. The assumptions are stated in Section 4.4. The finite-dimensional convergence of the process is derived in Section 4.5. Section 4.6 defines an underlying metric for the process. Section 4.7 introduces the notion of tightness in this context. Section 4.8 derives a symmetrization lemma for the partial sum process. The symmetrization lemma is used to derive tightness via asymptotic equicontinuity in Section 4.9. Section 4.10 derives a Glivenko-Cantelli type result in this context. An application of the main results in testing hypotheses related to Markovian multiplicative error models for non-negative time series is mentioned in Section 4.11. Conclusion and possible directions for future research are in 4.12.

4.3 Set up and notation

Let $(\mathcal{A}, d)$ be a separable metric space and $(\Omega, \mathcal{B}, P)$ be a probability space. $\{X_i\}$ and $\{V_i\}$ be two sequences of random variables defined on $(\Omega, \mathcal{B}, P)$. Let X_i's form a martingale difference sequence, with respect to a filtration $\mathcal{F}_i$, i.e. X_i is $\mathcal{F}_i$-measurable for each $i \geq 1$ and V_i's are $\mathcal{F}_{i-1}$-measurable. Also let V_i and V_j be independent $\forall i \neq j$. Note that we have not assumed the independence of X_i's and V_i's. We shall consider the partial sum process given by:

$$S_n(A) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \mathbf{1}_A(V_i) \text{ for } A \in \mathcal{A}$$

where $\mathcal{A}$ is a 'suitable' collection of sets. Thus note that here the increments are given by X_i's, which form a martingale difference sequence and whether an increment takes place or not is controlled by another random variable V_i. We shall derive a pointwise central limit theorem as well as a uniform central limit theorem on $\mathcal{A}$.

The main result of this paper is based on the uniform weak limit of the partial sum process and the limit will be given by a Gaussian process indexed by sets.

4.4 Assumptions

Before going into the details of the proof, we shall define the basic concepts and state the necessary assumptions, upon which the proof will be built.

Assumption A1. *There exists a non-negative measure $\mu(.)$ defined on $\mathcal{A}$ such that*

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i^2|\mathcal{F}_{i-1}\right)\mathbf{1}_A(V_i)\xrightarrow{a.s.}\sigma^2\mu(A) \tag{4.1}$$

where $\xrightarrow{a.s.}$ denotes almost surely convergence of a sequence of random variables to a limiting random variable. Here σ^2 is a scaling factor. Without loss of generality we shall assume $\sigma^2 = 1$.

Definition 4.4.1 (Covering number)**.** *For $\delta > 0$, the covering number of $\mathcal{A}$ with respect to the metric $d(.,.)$ is the minimum number of spherical balls of radius δ, required to cover $(\mathcal{A}, d)$ and it is denoted by $\mathcal{N}(\delta, \mathcal{A}, d)$. Therefore*

$$\mathcal{N}(\delta,\mathcal{A},d)=\min\{|\mathcal{A}(\delta)|:\mathcal{A}(\delta)\ \textit{is a}\ \delta-\textit{net for}\ \mathcal{A}\ \textit{with respect to the metric}\ d\}$$

Definition 4.4.2 (Packing number)**.** *For $\delta > 0$, the packing number of $\mathcal{A}$ with respect to the metric $d(.,.)$ is the maximum number of spherical balls of radius δ, that covers $(\mathcal{A}, d)$ and the centers of the balls are at least δ distance apart. It is denoted by $\mathcal{D}(\delta, \mathcal{A}, d)$. Therefore*

$$\mathcal{D}(\delta,\mathcal{A},d)=\max\{k:\mathcal{A}\subseteq\cup_{i=1}^{k}Ai\ \textit{and}\ d(A_i,A_j)\geq\delta\forall i\neq j\}$$

Definition 4.4.3 (Metric entropy)**.** *The metric entropy of $\mathcal{A}$ with respect to d is given by*

$$H(\delta, \mathcal{A}, d) = \ln \mathcal{N}(\delta, \mathcal{A}, d)$$

where $\mathcal{N}(\delta, \mathcal{A}, d)$ denotes the covering number.

Assumption A2. *For any fixed $\delta > 0$, we have*

$$\int_0^\delta \sqrt{H(u, \mathcal{A}, d)} du < \infty$$

This condition is termed the *metric entropy integrability* condition. The following result from equation (1.5) of Talagrand (2014) is crucial to understand the significance of this condition.

Proposition 4.4.4 (Dudley's bound)**.** *If $\{X_t : t \in T\}$ is a Gaussian process defined on a metric space (T, d), and $\mathcal{N}(\epsilon, T, d)$ is associated the covering number then for some constant L,*

$$\mathbb{E}\left[\sup_{d(s,t)\leq\delta} |X_s - X_t|\right] \leq L \int_0^\delta \sqrt{H(u)} du.$$

It is clear than in order to ensure the regularity of the Gaussian process the right-hand side needs to be finite.

Assumption A3. *Given any set $A \in \mathcal{A}$ and any real number $\delta > 0$, there exist A_δ and A^δ, with $A^\delta \supset A \supset A_\delta$, such that*

$$d(A^\delta, A_\delta) < \delta$$

A_δ and A^δ are called the lower and upper δ-approximations of the set A.

Assumption A4. *There exist a fixed $a_0 \in \mathbb{R}^+$ such that for all i*

$$\lim_{n\to\infty} \mathbb{P}(\sup_{i\le n} |X_i| \ge a_0\sqrt{n}) \to 0$$

First see that

$$\begin{aligned}\mathbb{P}(\sup_{i\le n} |X_i| \ge a_0\sqrt{n}) &\le \sum_{i=1}^{n} \mathbb{P}\left(|X_i| \ge a_0\sqrt{n}\right) \\ &\le \sum_{i=1}^{n} \mathbb{P}(|X_i|\mathbf{1}(|X_i| \ge a_0\sqrt{n})) \\ &\le \frac{1}{na_0^2}\sum_{i=1}^{n} \mathbb{E}\left(X_i^2\mathbf{1}[|X_i| \ge a_0\sqrt{n}]\right)\end{aligned}$$

The last step is a direct consequence of Chebyshev's inequality. Note that in order for A4 to hold we need the right-hand side of the above equation to go to zero as n goes off to infinity. Thus we require

$$\sum_{i=1}^{n} \mathbb{E}\left(X_i^2\mathbf{1}[|X_i| \ge a_0\sqrt{n}]\right) = o(n)$$

The choice of a_0 will be specified in the latter part of the chapter.

4.5 Finite Dimensional Convergence

Finite dimensional convergence is defined as follows.

Definition 4.5.1. $\{S_n(A) : n \ge 1, A \in \mathcal{A}\}$ *is said to converge to* $\{Z(A) : A \in \mathcal{A}\}$

in finite dimension if for any $k \geq 1$ and $A_1, \dots A_k \in \mathcal{A}$ we have

$$(S_n(A_1), \dots, S_n(A_k)) \xrightarrow{d} (Z(A_1), \dots, Z(A_k))$$

This means any finite-dimensional projection of the sequence of random variables S_n converges in distribution to those of the limiting random variable Z.

4.5.1 Mean function

Since the X_i's form a martingale difference sequence, therefore

$$\mathbb{E}(X_i) = \mathbb{E}\mathbb{E}(X_i|\mathcal{F}_{i-1}) = 0$$

For any $A \in \mathcal{A}$, mean function of the process is:

$$\begin{aligned} \mu(A) &= \mathbb{E}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i \mathbf{1}_A(V_i)\right) \\ &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \mathbb{E}\left(\mathbb{E}(X_i|\mathcal{F}_{i-1})\mathbf{1}_A(V_i)\right) \\ &= 0 \end{aligned}$$

4.5.2 Covariance Function

For any $A, B \in \mathcal{A}$, let's denote the covariance function by

$$C_n(A, B) = \mathrm{Cov}(S_n(A), S_n(B)).$$

This covariance function can be written as:

$$\begin{aligned} C_n(A,B) &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}\left(X_iX_j\mathbf{1}_A(V_i)\mathbf{1}_B(V_j)\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i^2\mathbf{1}_{A\cap B}(V_i)\right) + \frac{1}{n}\sum\sum_{i\neq j}\mathbb{E}\left(X_iX_j\mathbf{1}_A(V_i)\mathbf{1}_B(V_j)\right) \end{aligned}$$

We shall show that the summands of the second term on the RHS are 0. Without loss of generality take $i < j$. Conditioning on $\mathcal{F}_{j-1}$ we can get

$$\begin{aligned} \mathbb{E}\left(X_iX_j\mathbf{1}_A(V_i)\mathbf{1}_B(V_j)\right) &= \mathbb{E}\left(\mathbb{E}\left(X_iX_j\mathbf{1}_A(V_i)\mathbf{1}_B(V_j)|\mathcal{F}_{j-1}\right)\right) \\ &= \mathbb{E}\left(X_i\mathbf{1}_A(V_i)\mathbf{1}_B(V_j)\mathbb{E}(X_j|\mathcal{F}_{j-1})\right) = 0 \end{aligned}$$

Therefore we can write,

$$\begin{aligned} C_n(A,B) &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i^2\mathbf{1}_{A\cap B}(V_i)\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(\mathbb{E}\left(X_i^2\mathbf{1}_{A\cap B}(V_i)|\mathcal{F}_{i-1}\right)\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(\mathbf{1}_{A\cap B}(V_i)\mathbb{E}(X_i^2|\mathcal{F}_{i-1})\right) \end{aligned}$$

By A1 the following holds true:

$$C_n(A,B) \xrightarrow{a.s.} \sigma^2\mu(A\cap B)$$

The conditional covariance function can be written as:

$$\begin{aligned} C'_n(A,B) &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(X_i^2\mathbf{1}_{A\cap B}(V_i)|\mathcal{F}_{i-1}\right) \\ &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(X_i^2|\mathcal{F}_{i-1})\mathbf{1}_{A\cap B}(V_i) \end{aligned}$$

By A1

$$C'_n(A,B) \xrightarrow{a.s.} \sigma^2\mu(A\cap B)$$

Theorem 4.5.2 (Finite dimensional convergence)**.** *The finite-dimensional distributions of S_n converge to those of Z, where Z is a mean zero set-indexed Gaussian process with*

$$\lim_{n\to\infty} Cov(S_n(A), S_n(B)) = Cov(Z(A), Z(B))$$

Proof. This follows from the classical pointwise central limit theorem and the fact that a Gaussian process is fully characterized by its mean function and covariance function. □

4.6 Canonical L_2 Metric on $\mathcal{A}$

Now we shall define a pseudo-metric in $\mathcal{A}$, given by

$$d(A,B) = \lim_{n\to\infty}\left(\mathbb{E}(S_n(A)-S_n(B))^2\right)^{\frac{1}{2}} \tag{4.2}$$

This is the canonical L_2-metric for such partial sum processes. For reference see Ossiander (1987) for example.

Also define a distance metric at level n given by $d_n(.,.)$, such that $d(.,.)$ is obtained as the almost sure limit of this pseudo-metric. Note that

$$d_n^2(A,B) := \mathbb{E}(S_n(A) - S_n(B))^2 = \mathbb{E}(S_n(A)^2) + \mathbb{E}(S_n(B)^2) - 2\mathbb{E}(S_n(A)S_n(B))$$

We have already noticed that the mean function is zero for any fixed $A \in \mathcal{A}$. Therefore the above term reduces to the following:

$$\mathbb{E}(S_n(A)^2) + \mathbb{E}(S_n(B)^2) - 2\text{Cov}(S_n(A), S_n(B)) \tag{4.3}$$

Since

$$C_n(A,A) = \mathbb{E}(S_n(A)^2) \xrightarrow{a.s.} \sigma^2\mu(A)$$

and

$$C_n(A,B) = \mathbb{E}(S_n(A)S_n(B)) \xrightarrow{a.s.} \sigma^2\mu(A \cap B)$$

therefore the canonical L_2-metric will be given by

$$d(A,B) = \sigma\mu^{1/2}(A \triangle B)$$

If we had considered the conditional version of the covariance function this L_2

metric would have reduced to the same.

$$\lim_{n\to\infty} [C_n'(A,A) + C_n'(B,B) - 2C_n'(A,B)] = \sigma\mu(A\triangle B) \tag{4.4}$$

4.7 Tightness

Definition 4.7.1. *A sequence of probability measures $\{P_n : n \geq 1\}$ is said to be tight if for every $\epsilon > 0$, there exists a compact set K_ϵ such that*

$$\mathbb{P}_n(K_\epsilon) > 1 - \epsilon \text{ for all } n$$

Thus tightness prevents the mass from escaping to infinity.

Following is the statement of Theorem (8.1) of Billingsley (2013), which shows the connection between weak convergence and tightness.

Theorem 4.7.2. *Let $\{P_n : n \geq 1\}$ and P be probability measures on $(C, \mathcal{C})$. If finite dimensional distributions of P_n converges to that of P and if P_n is tight, then $P_n \xrightarrow{w} P$.*

Since we have already proven finite-dimensional convergence, therefore all we need now is to prove tightness in order to derive weak convergence of the set indexed partial sum process.

The following theorem from Wellner (2005) plays a crucial role in the derivation of our result. It shows the equivalence of tightness and asymptotic equicontinuity.

Theorem 4.7.3. *The followings are equivalent:*

1. *All the finite dimensional distributions of the sample bounded process S_n converges in distribution and there exists a pseudo metric ρ on $\mathcal{A}$ such that*

   - $(\mathcal{A}, \rho)$ *is totally bounded.*
   - X_n *is asymptotically equicontinuous in probability with respect to d, i.e. for any* $\epsilon > 0$

     $$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}\left[\sup_{d(A,B) \leq \delta} |S_n(A) - S_n(B)| > \epsilon\right] = 0$$

     This can be interpreted as: For every $\epsilon > 0$ and $\eta > 0$, there exists $\delta > 0$ such that

     $$\mathbb{P}^*(\sup_{A,B: d(A,B) < \delta} |S_n(A) - S_n(B)| > \eta) < \epsilon \text{ for sufficiently large } n$$

     This condition is known as asymptotic equicontinuity.

2. *There exists a process with tight Borel probability distribution on $l^\infty(T)$ such that*

   $$X_n \xrightarrow{\mathcal{L}} X \text{ in } l^\infty(T)$$

Thus we can conclude that in order to prove tightness, it will be enough to show asymptotic equicontinuity.

4.8 A symmetrization Lemma

In this section, we shall derive a symmetrization lemma to show that the tightness of the original process is implied by the tightness of the symmetrized version of the process. Now we shall go through the steps of construction of the symmetrized process.

Suppose X_i's be as mentioned before, i.e. $\{X_i : i \geq 1\}$ forms a martingale difference sequence with respect to the filtration $\{\mathcal{F}_i : i \geq 1\}$. Consider $\{X_i' : i \geq 1\}$ such that, for each i, X_i and X_i' are conditionally independent given $\mathcal{F}_{i-1}$. Also assume $[(X_1, V_1), \ldots (X_i, V_i)]$ and $[(X_1', V_1), \ldots (X_i', V_i)]$ are distributionally equal conditional on $\mathcal{F}_{i-1}$. Thus X_i's can be thought of as conditionally independent copies of X_i's. Denote the corresponding partial sum process by $S_n'(A)$. For $A \in \mathcal{A}$

$$S_n'(A) := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i' \mathbf{1}_A(V_i)$$

Thus the symmetrized process is given by $S_n - S_n'$. And we shall show that the tightness of S_n is implied by tightness of $(S_n - S_n')$.

To begin with, let's define a decreasing sequence of summable real numbers given by $\{\delta_k : k \geq 0\}$. Let $\mathcal{A}_k$ be a δ_k-net of $\mathcal{A}$ with $\mathcal{A}_k \subseteq \mathcal{A}_{k+1}$. Define $\mathcal{D} = \cup_{k \geq 1} \mathcal{A}_k$. Thus $\mathcal{D}$ is a countable dense subset of $\mathcal{A}$. Therefore, given any $\epsilon > 0$ and $A \in \mathcal{A}$, there exists $\tilde{A} \in \mathcal{D}$ such that $d(A, \tilde{A}) < \epsilon$. Define

$$G_n(A) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i^2 | \mathcal{F}_{i-1}) \mathbf{1}_A(V_i) \tag{4.5}$$

$$G(A) = \lim_{n\to\infty} G_n(A) \tag{4.6}$$

For $A, B \in \mathcal{D}$, define

$$\rho(A,B) = |G(A) - G(B)| \tag{4.7}$$

and for $f : \mathcal{D} \to \mathbb{R}$, define

$$||f||_\delta = \sup_{A,B\in\mathcal{D}:\rho(A,B)<\delta} |f(A) - f(B)| \tag{4.8}$$

Since $\mathcal{D}$ is a countable, dense subset of $\mathcal{A}$, therefore tightness of S_n on $\mathcal{D}$ will imply tightness of S_n on $\mathcal{A}$.

Note that here we are **not** assuming the independence of X_i's and V_i's. As before we shall just assume that X_i's are $\{\mathcal{F}_i\}$-measurable martingale differences and V_i are $\mathcal{F}_{i-1}$-measurable.

Lemma 4.8.1. *For fixed $\epsilon > 0, \delta > 0, \eta > 0$; $\epsilon < \eta$ and $\delta < \frac{\epsilon}{\sqrt{5}}$, the following holds true.*

$$\begin{aligned}\mathbb{P}(||S_n||_\delta > \eta) \quad &\leq \quad \left(1 - \frac{3\delta}{\epsilon^2}\right)^{-1} \mathbb{P}(||S_n - S'_n||_\delta > \eta - \epsilon) \\ &+ \quad \mathbb{P}(\sup_{k\geq n}\sup_{A\in\mathcal{D}} |\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i^2|\mathcal{F}_{i-1}) - \sigma^2\mu(A)| > \delta)\end{aligned}$$

Proof. Define,

$$\begin{aligned}
\mathcal{F} &:= \sigma\{(X_i, V_i : i \geq 1)\} \\
\rho_n(A,B) &:= |G_n(A) - G_n(B)| \\
&= |\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(X_i^2|\mathcal{F}_{i-1})(\mathbf{1}_A(V_i) - \mathbf{1}_B(V_i))| \\
&= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left(X_i^2|\mathcal{F}_{i-1}\right)\mathbf{1}_{A\triangle B}(V_i) \\
&= G_n(A\triangle B) \\
B_n(\eta; A, B) &:= \{\omega : |S_n(A) - S_n(B)| > \eta\} \\
B_n(\eta, \delta) &:= \{\omega : ||S_n||_\delta > \eta\} = \cup_{A,B\in\mathcal{D}:d(A,B)<\delta} B_n(\eta, A, B)
\end{aligned}$$

For fixed $A, B \in \mathcal{D}$, with $d(A,B) < \delta$,

$$B_n(\eta; A, B) \cap [||S_n - S_n'||_\delta > \eta - \epsilon] \supseteq B_n(\eta; A, B) \cap [|S_n'(A) - S_n'(B)| < \epsilon]$$

From Chebyshev's inequality

$$\mathbb{P}(|S_n'(A) - S_n'(B)| < \epsilon|\mathcal{F}) \geq 1 - \frac{\rho_n(A,B)}{\epsilon^2}$$

Therefore,

$$I(B_n(\eta; A, B))\mathbb{P}(||S_n - S_n'||_\delta > \eta - \epsilon|\mathcal{F}) \geq I(B_n(\eta; A, B))\left(1 - \frac{\rho_n(A,B)}{\epsilon^2}\right)$$

Let,

$$A_n(\delta) = \cup_{k\geq n}[\sup_{A\in\mathcal{D}} |G_k(A) - G(A)| > \delta]$$

Using the triangle inequality on $A_n^c(\delta)$ we can get for all $A, B \in \mathcal{D}$ and $k \geq n$,

$$\rho_k(A,B) \leq 2\delta + \rho(A,B) \tag{4.9}$$

Thus for $A, B \in \mathcal{D}$ with $\rho(A,B) < \delta$ we have

$$\rho_k(A,B) < 3\delta$$

Hence,

$$\begin{aligned} I(B_n(\eta;\delta)\cap A_n^c(\delta))\left(1-\frac{3\delta}{\epsilon^2}\right) &\leq I(B_n(\eta;\delta)\cap A_n^c(\delta))\mathbb{P}(||S_n - S_n'||_\delta > \eta - \epsilon|\mathcal{F}) \\ &\leq \mathbb{P}(||S_n - S_n'||_\delta > \eta - \epsilon|\mathcal{F}) \end{aligned}$$

Note that

$$\begin{aligned} I(B_n(\eta,\delta)) &\leq I(B_n(\eta,\delta)\cap A_n^c(\delta)) + I(A_n(\delta)) \\ &\leq \left(1-\frac{3\delta^2}{\epsilon^2}\right)^{-1}\mathbb{P}(||S_n - S_n'||_\delta > \eta - \epsilon|\mathcal{F}) + I(A_n(\delta)) \end{aligned}$$

Taking expectation on both sides,

$$\mathbb{P}(||S_n||_\delta > \eta) \leq \left(1 - \frac{3\delta}{\epsilon^2}\right)^{-1} \mathbb{P}(||S_n - S_n'||_\delta > \eta - \epsilon) + \mathbb{P}(A_n(\delta))$$

□

4.9 Asymptotic Equicontinuity

In order to prove tightness we shall show asymptotic equicontinuity of the stochastic process given by $\{S_n(A) : A \in \mathcal{A}\}$. As the symmetrization lemma (4.8.1) suggests, this requires us to prove the following theorem for the symmetrized process defined as:

$$S_n(A) - S_n'(A) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i X_i \mathbf{1}_A(V_i)$$

But for notational convenience now on we shall denote the symmetrized process $(S_n(A) - S_n'(A))$ by $S_n(A)$.

Theorem 4.9.1 (Tightness)**.** *If the metric entropy condition (A2) holds, then given any* $\epsilon > 0$ *and* $\eta > 0$*, there exists* $\delta > 0$*, such that* $\mathbb{P}^*(\sup_{A,B:d(A,B)<\delta} |S_n(A) - S_n(B)| > \eta) < \epsilon$ *for sufficiently large* n*.*

For fixed $\delta > 0$, let $\delta_k = \delta.\beta^k$, where $\beta < 1$ is a scaling factor. Fix a_0 such that assumption (A4) holds, i.e.

$$\mathbb{P}(\sup_i |X_i| > a_0\sqrt{n}) \to 0 \text{ as } n \to \infty$$

Now divide the interval $[0, a_0]$ into multiple disjoint subintervals, namely $I_1, I_2, \ldots I_{k_n+1}$ as follows:

$$
\begin{aligned}
I_k &= (a_k, a_{k-1}] \text{ for } k \le k_n \\
I_{k_n+1} &= [0, a_{k_n}]
\end{aligned}
$$

This will allow k_n to go to infinity as $n \to \infty$ at a rate that will be determined later. Clearly $\cup_{k=0}^{k_n+1} I_k = [0, a_0]$. Let $\bar{I}_k = [0, a_{k-1}] = \cup_{j=k}^{k_n+1} I_j$. Denote $A_0 = A^{\delta}$, $A_k = A^{\delta_k}$, $\bar{A}_k = A_{\delta_k}$. Thus A_k and $\bar{A}_k$ denotes the lower and upper δ_k-approximations of A.

Lemma 4.9.2. *For any $A \in \mathcal{A}$ and $A_0 = A^{\delta}$, $S_n(A_0) - S_n(A) = \sum_{j=1}^{k_n+1} S_{n,j}(A_j \setminus A_{j+1}) + R_n(A)$, where*

$$
\begin{aligned}
S_{n,j}(B) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i X_i \mathbf{1}_{\bar{I}_j}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_B(V_i) \\
R_{n,k}(A) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i X_i \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_{A_k \setminus A}(V_i) \\
R_n(A) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i \left[\sum_{k=1}^{k_n+1} X_i \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_{A_k \setminus A}(V_i)\right] \\
R_n(A) &= \sum_{k=1}^{k_n+1} R_{n,k}(A)
\end{aligned}
$$

Proof. Using a stratification based on $\cup_{k=0}^{k_n+1} I_k = [0, a_0]$ and generic chaining to a level depending on the stratification, namely $(k_n + 1)$, we can get

$$
\begin{aligned}
S_n(A_0) - S_n(A) &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i X_i \left(\mathbf{1}_{A_0}(V_i) - \mathbf{1}_A(V_i)\right) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i X_i \sum_{k=1}^{k_n+1} \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \left(\mathbf{1}_{A_0}(V_i) - \mathbf{1}_A(V_i)\right) \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i \sum_{k=1}^{k_n+1} X_i \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \left[\sum_{j=0}^{k-1} \mathbf{1}_{A_j \setminus A_{j+1}}(V_i) + \mathbf{1}_{A_k \setminus A}(V_i)\right] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i \sum_{j=0}^{k_n} \left[\sum_{k=j+1}^{k_n+1} X_i \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_{A_j \setminus A_{j+1}}(V_i)\right] \\
&+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i \left[\sum_{k=1}^{k_n+1} X_i \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_{A_k \setminus A}(V_i)\right] \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \epsilon_i \sum_{j=0}^{k_n} X_i \mathbf{1}_{\bar{I}_j}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_{A_j \setminus A_{j+1}}(V_i) + \sum_{k=1}^{k_n+1} R_{n,k}(A) \\
&= \sum_{j=0}^{k_n} S_{n,j}(A_j \setminus A_{j+1}) + \sum_{k=1}^{k_n+1} R_{n,k}(A)
\end{aligned}
$$

There is an exchange of summation above.

□

Suppose $\mathcal{A}_\delta$ is the **maximal** subset of $\mathcal{A}$ such that

$$d(A_0, B_0) \geq \delta \text{ for } A_0 \neq B_0 \in \mathcal{A}_\delta$$

Let A_0 and B_0 are sets in $\mathcal{A}_\delta$ with

$$d(A, A_0) < \delta \text{ and } d(B, B_0) < \delta$$

Thus we can write

$$\begin{aligned}
\{\omega : \sup_{A,B:d(A,B)<\delta} |S_n(A) - S_n(B)| > \eta\} \;\subset\; & \{\omega : \sup_{A,A_0:d(A,A_0)<\frac{\delta}{3}} |S_n(A_0) - S_n(A)| > \eta/3\} \\
\cup\; & \{\omega : \sup_{B,B_0:d(A,B)<\frac{\delta}{3}} |S_n(B_0) - S_n(B)| > \eta/3\} \\
\cup\; & \{\omega : \sup_{A_0,B_0:d(A_0,B_0)<3\delta} |S_n(A_0) - S_n(B_0)| > \eta/3\}
\end{aligned}$$

Note that the first and the second set in the right-hand side of the above equation are equivalent. Hence the above inclusion will reduce to the following:

$$\begin{aligned}
\{\omega : \sup_{A,B:d(A,B)<\delta} |S_n(A) - S_n(B)| > \eta\} \;\subset\; & \{\omega : \sup_{A,A_0:d(A,A_0)<\frac{\delta}{3}} |S_n(A_0) - S_n(A)| > \eta/3\} \\
\cup\; & \{\omega : \sup_{A_0,B_0:d(A_0,B_0)<3\delta} |S_n(A_0) - S_n(B_0)| > \eta/3\}
\end{aligned}$$

Union-sum inequality yields:

$$\begin{aligned}
\mathbb{P}\left(\sup_{A,B:d(A,B)<\delta} |S_n(A) - S_n(B)| > \eta \right) \le\; & \mathbb{P}\left(\sup_{A,A_0:d(A,A_0)<\frac{\delta}{3}} |S_n(A_0) - S_n(A)| > \eta/3 \right) \\
& + \mathbb{P}\left(\sup_{A_0,B_0:d(A_0,B_0)<3\delta} |S_n(A_0) - S_n(B_0)| > \eta/3 \right)
\end{aligned} \tag{4.10}$$

We can construct a triangular array by defining:

$$X_{ni} = \frac{X_i}{\sqrt{n}}$$

Denote the corresponding σ-fields by $\mathcal{F}_{n,i} = \sigma(X_{n1}, \ldots, X_{ni})$. let $\tilde{\tilde{\eta}} > 0$ and $\tilde{\eta} > 0$ are such that $\tilde{\eta} + \tilde{\tilde{\eta}} < \eta/3$.

We shall use the following Bernstein's inequality for Martingales that can be found as theorem $(1.2A)$ of de la Pena et al. (1999).

Theorem 4.9.3 (Bernstein Inequality for martingales)**.** *Suppose $\{X_i : i \geq 1\}$ is a martingale difference sequence with respect to the filtration $\{\mathcal{F}_i : i \geq 1\}$, i.e. $\mathbb{E}(X_i|\mathcal{F}_i) = 0$. Let $\sigma_i^2 = \mathbb{E}(X_i^2|\mathcal{F}_{i-1})$ and $V_n^2 = \sum_{i=1}^n \sigma_i^2$. Assume that there exist some finite c such that $\mathbb{P}(|X_i| \leq c|\mathcal{F}_{i-1}) = 1$. Then for all $x \in \mathbb{R}$ and $y \in \mathbb{R}$ we have the following:*

$$\mathbb{P}\left(\sum_{i=1}^n X_i \geq x, V_n^2 \leq y \text{ for some } n\right) \leq \exp\{-\frac{x^2}{2(y+cx)}\}$$

4.9.1 Exponential Bound on $R_n(A)$

Let's start by obtaining a bound on

$$\mathbb{P}(\sup_A |R_{nk}(A)| > \tilde{\eta}_k) \text{ where } \sum_{k=1}^{k_n+1} \tilde{\eta}_k \leq \tilde{\eta} \tag{4.11}$$

Notice that

$$\begin{aligned} V_{nk}^2 &:= \sum_{i=1}^{n} \mathbb{E}(X_{ni}^2 \mathbf{1}_{I_k}(X_{ni})|\mathcal{F}_{n,i-1}).\mathbf{1}_{V_i}(A_k \setminus A) \\ &\leq \bar{V}_{nk}^2 := \sum_{i=1}^{n} \mathbb{E}(X_{ni}^2|\mathcal{F}_{n,i-1}).\mathbf{1}_{V_i}(A_k \setminus \bar{A}_k) \end{aligned}$$

By assumption(A1),

$$\bar{V}_{nk}^2 \xrightarrow{a.s.} \mu(A_k \setminus \bar{A}_k)\sigma^2 \leq \delta_k^2 \tag{4.12}$$

This follows because $\mu(A_k \ \bar{A}_k) = d^2(A_k, \bar{A}_k)/\sigma^2$ and $\sigma^2 = 1$ (Recall: $A_k = A^{\delta_k}, \bar{A}_k = A_{\delta_k}$ and we have $\bar{A}_k \subseteq A \subseteq A_k$).

We can get $\epsilon_{nk} > 0$ such that $V_{nk}^2 < \delta_k^2(1 + \epsilon_{nk})$ with probability 1. Denote $c_{nk} := 1 + \epsilon_{nk}$.

In this case $Y_i = \epsilon_i X_{ni} \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_{A_k \setminus A}(V_i)$ forms a martingale difference sequence with respect to the filtration $\{\mathcal{F}_{n,i}\}$. To see this note that

$$\begin{aligned} \mathbb{E}\left(\epsilon_i X_{ni} \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_{A_k \setminus A}(V_i)|\mathcal{F}_{n,i-1}\right) &= \epsilon_i \mathbf{1}_{A_k \setminus A}(V_i) \mathbb{E}\left[X_{ni} \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right)|\mathcal{F}_{n,i-1}\right] \\ &= \epsilon_i \mathbf{1}_{A_k \setminus A}(V_i) \mathbb{E}\left[X_{ni}|\mathcal{F}_{n,i-1}\right] \mathbf{P}\left(\frac{X_i}{\sqrt{n}} \in I_k\right) \\ &\quad + \epsilon_i \mathbf{1}_{A_k \setminus A}(V_i) \mathbb{E}\left[X_{ni}.0|\mathcal{F}_{n,i-1}\right] \mathbf{P}\left(\frac{X_i}{\sqrt{n}} \notin I_k\right) \\ &= 0 \end{aligned}$$

We have

$$\begin{aligned} Y_i &= \epsilon_i X_{ni} \mathbf{1}_{I_k}\left(\frac{X_i}{\sqrt{n}}\right) \mathbf{1}_{A_k \setminus A}(U_i) \leq a_{k-1} \\ V_{nk}^2 &\leq c_{nk}\delta_k^2 \end{aligned}$$

Theorem 4.9.3 gives

$$\begin{aligned} \mathbb{P}(\sup_A |R_n(A)| > \tilde{\eta}) &\leq \sum_{k\geq 1} \mathbb{P}(\sup_A |R_{nk}(A)| > \tilde{\eta}_k) \\ &\leq \sum_{k\geq 1} 2\mathcal{N}(\delta_k \mathcal{A}, d) \exp\left[-\frac{\tilde{\eta}_k^2}{2(c_{nk}\delta_k^2 + a_{k-1}\tilde{\eta}_k)}\right] \end{aligned}$$

4.9.2 Choice of Parameters to bound $R_n(A)$

To ensure the summability of the series we choose the parameters as follows:

$$\begin{aligned} \tilde{\eta}_k &= \delta_{k-1}\sqrt{H(\delta_k)} \\ a_{k-1} &= \frac{\delta_k}{\sqrt{H(\delta_k)}} \end{aligned}$$

After plugging in the values,

$$k - \text{th term } = \mathcal{N}(\delta_k, \mathcal{A}, d)^{1 - \frac{\delta_{k-1}^2}{2(c_{nk}\delta_k^2 + \delta_k\delta_{k-1})}} \tag{4.13}$$

The exponent of this term is :$1-\frac{n}{2(c_{nk}\beta^2+\beta)}$. To make the series summable we need:

$$1-\frac{n}{2(c_{nk}\beta^2+\beta)}<0 \tag{4.14}$$

Now choose $0<\beta<\sqrt{1+c_{nk}}-1$ so that the series is summable. Once the summability is assured, given any $\tilde{\epsilon}>0$, we can choose β and hence δ so that the sum is less than the fixed $\tilde{\epsilon}$. Denote this β by $\tilde{\beta}$ and the δ by $\tilde{\delta}$.

4.9.3 Exponential bound on $\sum_{j=0}^{k_n} S_{n,j}(A_j \setminus A_{j+1})$

Individual terms of $S_{n,j}(A_j \setminus A_{j+1})$ are given by $\epsilon_i X_{ni}\mathbf{1}_{\bar{I}_j}(X_{ni})\,\mathbf{1}_{A_j\setminus A_{j+1}}(V_i)$. Since

$$\mathbb{E}\left(\epsilon_i X_{ni}\mathbf{1}_{\bar{I}_j}(X_{ni})\,\mathbf{1}_{A_j\setminus A_{j+1}}(V_i)|\mathcal{F}_{n,i-1}\right)=0$$

therefore these individual terms form a martingale difference sequence. Now consider,

$$\begin{aligned} V_{nj}^2 &= \sum_{i=1}^{n}\mathbb{E}\left(X_{ni}^2\mathbf{1}_{\bar{I}_j}(X_{ni})\,|\mathcal{F}_{n,i-1}\right)\mathbf{1}_{A_j\setminus A_{j+1}}(V_i)\le \bar{V}_{nj}^2 \\ &= \sum_{i=1}^{n}\mathbb{E}\left(X_{ni}^2|\mathcal{F}_{n,i-1}\right)\mathbf{1}_{A_j\setminus A_{j+1}}(V_i) \end{aligned}$$

Assumption(A1) and Triangle inequality gives

$$V_{nj}^2\le \bar{V}_{nj}^2 \xrightarrow{a.s.} \mu(A_j\setminus A_{j+1})\le \delta_j^2+\delta_{j+1}^2 \tag{4.15}$$

Therefore there exist $c_{nj} > 1$ such that $c_{nj} \to 1$ and $V_{nj}^2 \le c_{nj}(\delta_j^2 + \delta_{j+1}^2)$. Bernstein's inequality for martingales (4.9.3) gives:

$$\begin{aligned}
&\mathbb{P}(\sup_{A,A_0\in\mathcal{A}} \sum_{j=0}^{k_n} S_{n,j}(A_j \setminus A_{j+1}) > \tilde{\tilde{\eta}}) \\
\le & \sum_{j=0}^{k_n} \mathbb{P}(\sup_{A_j,A_{j+1}\in\mathcal{A}} S_{n,j}(A_j \setminus A_{j+1}) > \tilde{\tilde{\eta}}_j) \\
\le & \sum_{j=0}^{k_n} (\mathcal{N}(\delta_j,\mathcal{A},d))^2\, 2.\exp\left[-\frac{\tilde{\tilde{\eta}}_j^2}{2((\delta_j^2+\delta_{j+1}^2)c_{nj} + a_{j-1}\tilde{\tilde{\eta}}_j)}\right]
\end{aligned}$$

with $\sum_{j=0}^{k_n} \tilde{\tilde{\eta}}_j < \tilde{\tilde{\eta}}$.

4.9.4 Choice of Parameters to bound $\sum_{j=0}^{k_n} S_{n,j}(A_j \setminus A_{j+1})$

The following parameter choices ensure summability of the series.

$$\begin{aligned}
a_{j-1} &= \frac{\delta_j}{\sqrt{H(\delta_j)}} \\
\tilde{\tilde{\eta}}_j &= \lambda\delta_{j+1}\sqrt{H(\delta_j)}
\end{aligned}$$

λ is suitably chosen positive real number.

$$\begin{aligned}
\text{exponent of } j\text{-th term} &= \left(2 - \frac{\lambda^2\delta_j^2}{2((\delta_j^2+\delta_{j+1}^2)c_{nj} + \delta_j\delta_{j+1}\lambda)}\right) \\
&= 2 - \frac{\lambda^2}{2(c_{nj}(1+\beta^2)+\beta\lambda)}
\end{aligned}$$

Summability of the series requires the exponent to be negative.

$$\begin{aligned} 2 - \frac{\lambda^2}{2(c_{nj}(1+\beta^2)+\beta\lambda)} &< 0 \\ 4(1+\beta^2)c_{nj} + 4\beta\lambda - \lambda^2 &< 0 \\ 4c\beta^2 + 4\beta\lambda + (4c_{nj} - \lambda^2) &< 0 \end{aligned}$$

The roots of the quadratic are :

$$\begin{aligned} \beta_1 &= \frac{-4\lambda - \sqrt{16\lambda^2 - 16c_{nj}(4c_{nj} - \lambda^2)}}{8\lambda} \\ \beta_2 &= \frac{-4\lambda + \sqrt{16\lambda^2 - 16c_{nj}(4c_{nj} - \lambda^2)}}{8\lambda} \end{aligned}$$

Clearly $\beta_1 < 0$. But we require β to be positive. Hence we shall go for β_2. But we need to make sure that $\beta_2 > 0$, which follows under the following condition:

$$\lambda > 2c_{nj}$$

This condition also ensures that the discriminant of the quadratic is positive.

If

$$0 < \beta < \frac{-\lambda + \sqrt{\lambda^2 - c_{nj}(4c_{nj} - \lambda^2)}}{2\beta}$$

then the exponent of the j-th term in the series will be negative. The quadratic is increasing in β. Hence it can be made arbitrarily small by choosing β sufficiently small. Denote the chosen β by $\tilde{\tilde{\beta}}$ and the corresponding δ by $\tilde{\tilde{\delta}}$.

Now that we have shown that the first term in equation (4.10) can be made arbitrarily small, all we need to show is that the second term can be made arbitrarily small as well. To begin with notice that in $\max_{A_0,B_0\in\mathcal{A}_\delta:d(A_0,B_0)\leq 3\delta}|S_n(A_0)-S_n(B_0)|$, the maximum is over a finite set. To see this first note that the cardinality of $\mathcal{A}_\delta$ is given by

$$|\mathcal{A}_\delta| = \mathcal{D}(\mathcal{A}, \delta, d)$$

with $\mathcal{D}(\mathcal{A}; \delta, d)$ denoting the packing number. Also since the covering number and packing number are equivalent in the following sense:

$$\mathcal{D}(\mathcal{A}; 2\delta, d) \leq \mathcal{N}(\mathcal{A}; \delta, d) \leq \mathcal{D}(\mathcal{A}; \delta, d)$$

therefore by metric entropy integrability condition $\mathcal{D}(\mathcal{A}; 2\delta, d)$ has to be finite. Hence there must exist a finite C for which

$$|S_n(A_0) - S_n(B_0)| \leq Cd(A_0, B_0)\forall A_0, B_0 \in \mathcal{A}_\delta : d(A_0, B_0) \leq 3\delta$$

Thus given $\eta > 0$ and $\epsilon > 0$, we can choose δ' such that

$$\mathbb{P}\left(\sup_{A_0,B_0:d(A_0,B_0)<3\delta'} |S_n(A_0) - S_n(B_0)| > \eta/3\right) \leq \epsilon/2$$

Proof of Theorem 4.9.1. Choose $\delta = \min\{\tilde{\delta}, \tilde{\tilde{\delta}}, \delta'\}$ so that asymptotic equicontinuity holds. Thus tightness will follow automatically. □

Theorem 4.9.4 (Uniform central limit theorem)**.** *If* $\{Z(A) : A \in \mathcal{A}\}$ *is a mean* *zero Gaussian process with*

$$Cov(S_n(A), S_n(B)) = Cov(Z(A), Z(B))$$

then under the metric entropy integrability condition as mentioned in A2, the following holds true.

$$S_n \xrightarrow{w} Z$$

Proof. Finite-dimensional convergence together with tightness implies convergence of $\{S_n(A) : A \in \mathcal{A}\}$ to the Gaussian process Z defined above. □

4.10 A Glivenko-Cantelli theorem

By using the Dudley-Yechura-Skorohod representation theorem in page 47 (theorem 4) of Shorack and Wellner (2009) we can claim that there exists a version of the partial sum process and the limiting Gaussian process in some space where

$$\sup_{A \in \mathcal{A}} |S_n(A) - Z(A)| \xrightarrow{a.s.} 0$$

4.11 An application

In Koul et al. (2012), the authors have constructed a lack of fit test of a given parametric form of multiplicative error model involving Markov structure. Multiplicative error models(MEM) are used for modeling non-negative time series. A multiplicative error model has the following form.

$$Y_i = \mathbb{E}(Y_i|\mathcal{H}_{i-1})\epsilon_i$$

where ϵ_i denotes the multiplicative error terms and $\mathcal{H}_{i-1}$ denotes history of the process upto time $(i-1)$. Conditional on $\mathcal{H}_{i-1}$, ϵ_i's are independent and identically distributed random variables with mean one and variance σ^2. An MEM is said to possess the Markov property if

$$\mathbb{E}(Y_i|\mathcal{H}_{i-1}) = \tau(Y_{i-1})$$

where $\tau(.)$ is a non-negative function defined on the positive real line. Let q be a known natural number and $\Theta \in \mathbb{R}^q$. We want to test whether $\tau(.)$ has a specific parametric form or not i.e.

$$H_0 : \tau(y) = \psi(y, \theta) \text{ for some } \theta \in \Theta \text{ vs } H_A : \text{Not } H_0$$

In Koul et al. (2012), the authors have considered the following test statistic:

$$T(y, v) = \int_0^y \left[\frac{\tau(x)}{\psi(x, v)} - 1\right] dG(x)$$

where $v \in \Theta$. Clearly under H_0 and when the true parameter value is known (given by θ) we must have

$$T(y,\theta) = \mathbb{E}\mathbf{1}(Y_0 \leq y)\left[\frac{Y_1}{\psi(Y_0,\theta)} - 1\right] = 0 \text{ for all } y \geq 0$$

The authors claim that an unbiased estimator of $T(y,\theta)$ is given by

$$n^{-\frac{1}{2}} U_n(y,v)$$

where

$$U_n(y,v) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{Y_i}{\psi(Y_{i-1},v)} - 1\right] \mathbf{1}(Y_{i-1} \leq y) \text{ for } y > 0, v \in \Theta$$

Clearly a more generalized test statistic in this case will be given by

$$\tilde{T}(A,v) = \int_A \left[\frac{\tau(x)}{\psi(x,v)} - 1\right] dG(x) \text{ for } A \in \mathcal{A}$$

where $\mathcal{A}$ is a suitable collection of sets that satisfy the metric entropy integrability condition. An unbiased estimate of this statistic will be given by

$$\tilde{U}_n(A,v) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[\frac{Y_i}{\psi(Y_{i-1},v)} - 1\right] \mathbf{1}(Y_{i-1} \in A) \text{ for } A \in \mathcal{A}$$

Under H_0, when the true parameter value is known

$$\tilde{T}(A,v) = \mathbb{E}\mathbf{1}(Y_0 \in A)\left[\frac{Y_1}{\psi(Y_0,\theta)} - 1\right] = 0 \text{ for all } A \in \mathcal{A}$$

It is easy to see that in this case $\frac{Y_i}{\psi(Y_{i-1},v)} - 1$ forms a martingale difference sequence that is $\mathcal{F}_i$ measurable and Y_{i-1} is $\mathcal{F}_{i-1}$ measurable. Thus this is a partial sum process of the form we have considered. As the regularity condition $(C1)$ in section 2.2 of Koul et al. (2012) suggests, the authors had assumed a fourth moment condition on the Y_0, which is a stronger condition than that of assumption A1. Thus our method can derive the null distribution of the test statistics under milder conditions.

4.12 Conclusion

In this chapter, we have considered a partial sum process that is indexed by set, where the set belongs to a suitable collection of sets $\mathcal{A}$. The increments of the process form a martingale difference sequence. We have derived a Donsker type result, i.e. a central limit theorem (Lindeberg type) that holds uniformly on the sets of $\mathcal{A}$. Also, a Glivenko-Cantelli type result was derived that gives uniform almost sure convergence of the process to a set-indexed Gaussian process. The results were derived under the metric entropy integrability condition.

A natural extension of the process would be to consider the case when V_i and X_i's are random vectors instead of being random variables. In that case, this process can be used as a test statistic to perform hypothesis tests for spatial processes. Another extension of the process would be to consider a more general class of indexing functions $g \in \mathcal{G}$, instead of the indicator functions indexed by sets.

5 Conclusion

This thesis contains a study of the limiting behavior of three different stochastic processes. Each of the processes has an inherent dependent structure that poses a challenge in deriving their long-term behaviors. Also, each of the processes is defined in a different state space with each requiring different techniques to derive limiting properties.

The first process considered here is an evolutionary urn scheme based on the well-known rock-paper-scissors game. The urn starts with one rock, one paper, and one scissor. Over time the urn evolves based on the rules of this game. At every step, the number of objects in the urn goes up exactly by one. We have found the long term proportions of each individual object type by identifying martingale structures within the game and connecting those with roots of cubic polynomials. We have also considered a natural extension of the game, where instead of three, there are k objects in the urn, which evolves according to a set of rules which can be thought of as a generalization of the rock-paper-scissors game. The following describes differences between the standard(involving three different object types) and the generalized version of the problem.

- In the generalized set up involving k object types the game is not unique, i.e. by maintaining the cyclic structure of the game, it is possible to construct different versions of the game by considering different dominance rules of the remaining pairs. For example: Consider the game involving 4 different object types, that are denoted by A, B, C, D. The cyclic structure of this game is governed by the rule $A > B > C > D > A$. However, there are two remaining pairwise comparisons, namely A, C and B, D. Based on whether $A \lesseqgtr C$ and $B \lesseqgtr D$, there exist four distinct versions of the game. The limiting behavior of the urn depends on the particular version of the game.

- As we have already seen, in the case of the three-player game the limiting proportions are uniform , i.e. each proportion converges to the same number $\frac{1}{3}$. As the simulation study suggests if the number of object types is an even number then there is no chance that the proportions will converge to the same number. Indeed every $(2n)$-player game gets reduced to a $(2n-1)$ player game, and the proportion of the remaining object type converges to zero.

The limiting behavior in the general set up (involving k object types) needs further investigation. Though we have suggested an algorithm for finding the limiting proportions, when it exists, the proof of almost sure convergence in the general set up still remains unaddressed. Also, the speed of convergence is of interest. The rate of convergence of the limiting proportions in a three-player and a general k-player set up should depend on the initial composition of the urn and the rules of the game. One promising technique to investigate the speed of convergence is the use of *Stein's method*, which involves the construction of exchangeable pair of random variables.

In the second chapter of the thesis, we have considered a special class of multiplicative models, namely the random multiplicative cascade models, that are used for modeling data with variable intermittency. The key component of such a model is the cascade generating distribution, which is uniquely identifiable from the structure-function of the cascade. The structure-function is estimated from the complete finite sample realization at a particular resolution. We have proposed an estimator under the set up when the incomplete data is observed at a particular resolution. Large sample properties, including a pointwise central limit theorem of the estimators, were derived under this setup. Also, we have shown how the pointwise limiting results extend to uniform limiting theorems, where the uniformity is on set on which the data was observed. We have illustrated the method using the daily stock volume data of Tesla. We have proposed a naive change point detection method for intermittent data sets. Further refinement of the method is possible by derivation of uniform central limit theorems for estimates of structure-function, denoted by $\tau_n(h; A)$. A uniform central limit theorem of $\{\tau_n(h; A) : h \in H\}$ will enable us to perform the hypothesis test of equality of two structure functions on their overlapping domain set H. The null distribution of the test statistic can be obtained from the uniform central limit theorem. Derivation of the uniform central limit theorem is an important open problem in this field. The connections between random multiplicative models and other spatio-temporal processes are also of interest.

The last chapter of the thesis gives the derivation of limiting properties of a set-indexed partial sum process. We have derived a Donsker type result and a

Glivenko-Cantelli type result in this context under the standard metric entropy integrability condition together with a Lindeberg type condition. This result can be used to inform hypothesis tests in the context of multiplicative error models of non-negative time series. The partial sum process is given by

$$S_n(A) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \mathbf{1}_A(V_i) \text{ for } A \in \mathcal{A}$$

where $\mathbf{1}_A$ denotes an indicator function defined on $A \in \mathcal{A}$. One possible extension of the process would be in the direction where the class of indicator functions is generalized to a broader class of functions defined on the respective sets. Extensions of this process from random variables to random vectors and then to random functions are also of interest as they would allow wider applicability of the process for testing hypotheses related to multivariate time series and spatio-temporal processes.

As pointed out earlier, each of these stochastic processes has an inherent martingale structure. Thus martingale tools and techniques have played a crucial role in deriving the limiting properties of these processes. The key was to identify the martingale structure and to understand how the martingales gradually evolve with the process. Another key ingredient of this thesis is the application of a wide variety of concentration inequalities that were instrumental in deriving the probabilistic bounds.

Bibliography

Patrice Abry, Richard Baraniuk, Patrick Flandrin, Rudolf Riedi, and Darryl Veitch. Multiscale nature of network traffic. *IEEE Signal Processing Magazine*, 19(3):28–46, 2002.

F Anselmet, Yl Gagne, EJ Hopfinger, and RA Antonia. High-order velocity structure functions in turbulent shear flows. *Journal of Fluid Mechanics*, 140:63–89, 1984.

A Arneodo, JF Muzy, and SG Roux. Experimental analysis of self-similarity and random cascade processes: Application to fully developed turbulence data. *Journal de Physique II*, 7(2):363–370, 1997.

A Arneodo, S Manneville, and JF Muzy. Towards log-normal statistics in high reynolds number turbulence. *The European Physical Journal B-Condensed Matter and Complex Systems*, 1(1):129–140, 1998.

Michael Artin. *Algebra, 2nd Edition*. Pearson, 2010.

PP Avelino, D Bazeia, L Losano, J Menezes, and BF Oliveira. Junctions and spiral patterns in generalized rock-paper-scissors models. *Physical Review E*, 86 (3):036112, 2012.

E Bacry, A Kozhemyak, and J-F Muzy. Log-normal continuous cascades: aggregation properties and estimation. application to financial time-series. *arXiv preprint arXiv:0804.0185*, 2008.

Emmanuel Bacry, Jean Delour, and Jean-François Muzy. Modelling financial time series using multifractal random walks. *Physica A: statistical mechanics and its applications*, 299(1-2):84–92, 2001.

Arifah Bahar and Xuerong Mao. Stochastic delay lotka–volterra model. *Journal of Mathematical Analysis and Applications*, 292(2):364–380, 2004.

Philippe Barbe and Patrice Bertail. *The weighted bootstrap*, volume 98. Springer Science & Business Media, 2012.

Patrick Billingsley. Probability and measure.,(john wiley and sons: New york). *Billingsley2Probability and Measure1986*, 1986.

Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

Bruce M Brown et al. Martingale central limit theorems. *The Annals of Mathematical Statistics*, 42(1):59–66, 1971.

Timothy N Cason, Daniel Friedman, and Ed Hopkins. Cycles and instability in a rock–paper–scissors population game: A continuous time experiment. *Review of Economic Studies*, 81(1):112–136, 2013.

Patrick Cattiaux and Sylvie Méléard. Competitive or weak cooperative stochastic lotka–volterra systems conditioned on non-extinction. *Journal of mathematical biology*, 60(6):797–829, 2010.

Richard Cook, Geoffrey Bird, Gabriele Lünser, Steffen Huck, and Cecilia Heyes. Automatic imitation in a strategic context: players of rock–paper–scissors imitate opponents' gestures. *Proceedings of the Royal Society B: Biological Sciences*, 279(1729):780–786, 2011.

Victor de la Pena et al. A general class of exponential inequalities for martingales and ratios. *The Annals of Probability*, 27(1):537–564, 1999.

Alexander Dobrinevski and Erwin Frey. Extinction in neutrally stable stochastic lotka-volterra models. *Physical Review E*, 85(5):051903, 2012.

Nguyen Huu Du and Vu Hai Sam. Dynamics of a stochastic lotka–volterra model perturbed by white noise. *Journal of mathematical analysis and applications*, 324(1):82–97, 2006.

Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

Marcus Frean and Edward R Abraham. Rock–scissors–paper and the survival of the weakest. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1474):1323–1327, 2001.

David A Freedman. On tail probabilities for martingales. *the Annals of Probability*, pages 100–118, 1975.

Uriel Frisch and Andreĭ Nikolaevich Kolmogorov. *Turbulence: the legacy of AN Kolmogorov*. Cambridge university press, 1995.

Anna C Gilbert, Walter Willinger, and Anja Feldmann. Scaling analysis of conservative cascades, with applications to network traffic. *IEEE Transactions on Information Theory*, 45(3):971–991, 1999.

Vijay K Gupta and Edward C Waymire. A statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, 32(2):251–267, 1993.

Peter Hall and Christopher C Heyde. *Martingale limit theory and its application*. Academic press, 2014.

Da-yin Hua, Lou-cheng Dai, and Cao Lin. Four-and three-state rock-paper-scissors games with long-range selection. *EPL (Europhysics Letters)*, 101(3):38004, 2013.

Luo-Luo Jiang, Wen-Xu Wang, Ying-Cheng Lai, and Xuan Ni. Multi-armed spirals and multi-pairs antispirals in spatial rock–paper–scissors games. *Physics Letters A*, 376(34):2292–2297, 2012.

A JOFFE, L LECAM, and J NEVEU. Law of large numbers for random bernoulli variables attached to dyadic tree. *COMPTES RENDUS HEBDOMADAIRES DES SEANCES DE L ACADEMIE DES SCIENCES SERIE A*, 277(19):963–964, 1973.

Chatchai Jothityangkoon, Murugesu Sivapalan, and Neil R Viney. Tests of a space-time model of daily rainfall in southwestern australia based on nonhomogeneous random cascades. *Water resources research*, 36(1):267–284, 2000.

J-P Kahane and Jacques Peyriere. Sur certaines martingales de benoit mandelbrot. *Advances in mathematics*, 22(2):131–145, 1976.

Benjamin Kerr, Margaret A Riley, Marcus W Feldman, and Brendan JM Bohannan. Local dispersal promotes biodiversity in a real-life game of rock–paper–scissors. *Nature*, 418(6894):171, 2002.

Benjamin C Kirkup and Margaret A Riley. Antibiotic-mediated antagonism leads to a bacterial game of rock–paper–scissors in vivo. *Nature*, 428(6981):412, 2004.

Andrey Nikolaevich Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. *Cr Acad. Sci. URSS*, 30: 301–305, 1941.

Andrey Nikolaevich Kolmogorov. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high reynolds number. *Journal of Fluid Mechanics*, 13(1):82–85, 1962.

Hira L Koul, Indeewara Perera, and Mervyn J Silvapulle. Lack-of-fit testing of the conditional mean function in a class of markov multiplicative error models. *Econometric Theory*, 28(6):1283–1312, 2012.

Benoit Laslier, Jean-François Laslier, et al. Reinforcement learning from comparisons: Three alternatives are enough, two are not. *The Annals of Applied Probability*, 27(5):2907–2925, 2017.

Nabil Lasmar, Cécile Mailler, and Olfa Selmi. Multiple drawing multi-colour urns by stochastic approximation. *Journal of Applied Probability*, 55(1):254–281, 2018.

Andrés E Leövey and Thomas Lux. Parameter estimation and forecasting for multiplicative log-normal cascades. *Physical Review E*, 85(4):046114, 2012.

Xiaoyue Li and Xuerong Mao. Population dynamical behavior of non-autonomous lotka-volterra competitive system with random perturbation. *Discrete and Continuous Dynamical Systems-Series A*, 24(2):523–593, 2009.

Meng Liu and Meng Fan. Permanence of stochastic lotka–volterra systems. *Journal of Nonlinear Science*, 27(2):425–452, 2017.

Meng Liu and Ke Wang. Stochastic lotka–volterra systems with lévy noise. *Journal of Mathematical Analysis and Applications*, 410(2):750–763, 2014.

Alfred J Lotka. Elements of physical biology. *Science Progress in the Twentieth Century (1919-1933)*, 21(82):341–343, 1926.

Benoit B Mandelbrot. Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *Journal of fluid Mechanics*, 62 (2):331–358, 1974.

Benoit B Mandelbrot. Scaling in financial prices: Ii. multifractals and the star equation. 2001.

Benoit B Mandelbrot, Adlai J Fisher, and Laurent E Calvet. A multifractal model of asset returns. 1997.

Xuerong Mao, Sotirios Sabanis, and Eric Renshaw. Asymptotic behaviour of the stochastic lotka–volterra model. *Journal of Mathematical Analysis and Applications*, 287(1):141–156, 2003.

Mauro Mobilia. Oscillatory dynamics in rock–paper–scissors games with mutations. *Journal of Theoretical Biology*, 264(1):1–10, 2010.

Peter Molnar and Paolo Burlando. Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model. *Atmospheric Research*, 77 (1-4):137–151, 2005.

Jean-François Muzy, Jean Delour, and Emmanuel Bacry. Modelling fluctuations of financial time series: from cascade process to stochastic volatility model. *The European Physical Journal B-Condensed Matter and Complex Systems*, 17(3): 537–548, 2000.

Mina Ossiander. A central limit theorem under metric entropy with l2 bracketing. *The Annals of Probability*, pages 897–919, 1987.

Mina Ossiander, Edward C Waymire, et al. Statistical estimation for multiplicative cascades. *The Annals of Statistics*, 28(6):1533–1560, 2000.

Thomas M Over and Vijay K Gupta. Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing. *Journal of Applied Meteorology*, 33(12):1526–1542, 1994.

Matti Peltomäki and Mikko Alava. Three-and four-state rock-paper-scissors games with diffusion. *Physical Review E*, 78(3):031906, 2008.

Robin Pemantle et al. A survey of random processes with reinforcement. *Probability surveys*, 4:1–79, 2007.

Sidney Resnick, Gennady Samorodnitsky, Anna Gilbert, Walter Willinger, et al. Wavelet analysis of conservative cascades. *Bernoulli*, 9(1):97–135, 2003.

D Schertzer and S Lovejoy. Universal multifractals do exist!: Comments on a statistical analysis of mesoscale rainfall as a random cascade. *Journal of Applied Meteorology*, 36(9):1296–1303, 1997.

Jürgen Schmiegel, Jochen Cleve, Hans C Eggers, Bruce R Pearson, and Martin Greiner. Stochastic energy-cascade model for (1+ 1)-dimensional fully developed turbulence. *Physics Letters A*, 320(4):247–253, 2004.

Dirk Semmann, Hans-Jürgen Krambeck, and Manfred Milinski. Volunteering leads to rock–paper–scissors dynamics in a public goods game. *Nature*, 425(6956):390, 2003.

Bodhisattva Sen. A gentle introduction to empirical process theory and applications. 2018.

Zhen-Su She and Emmanuel Leveque. Universal scaling laws in fully developed turbulence. *Physical review letters*, 72(3):336, 1994.

Zhen-Su She and Edward C Waymire. Quantized energy cascade and log-poisson statistics in fully developed turbulence. *Physical Review Letters*, 74(2):262, 1995.

Hongjing Shi, Wen-Xu Wang, Rui Yang, and Ying-Cheng Lai. Basins of attraction for species extinction and coexistence in spatial rock-paper-scissors games. *Physical Review E*, 81(3):030901, 2010.

Galen R Shorack and Jon A Wellner. *Empirical processes with applications to statistics*. SIAM, 2009.

Barry Sinervo and Curt M Lively. The rock–paper–scissors game and the evolution of alternative male strategies. *Nature*, 380(6571):240, 1996.

Dean Slonowsky. *Central limit theorems for set-indexed strong martingales.* University of Ottawa (Canada), 1998.

Kei-ichi Tainaka. Physics and ecology of rock-paper-scissors game. In *International Conference on Computers and Games*, pages 384–395. Springer, 2000.

Michel Talagrand. *Upper and lower bounds for stochastic processes: modern methods and classical problems*, volume 60. Springer Science & Business Media, 2014.

Brent M Troutman, Aldo V Vecchia, et al. Estimation of renyi exponents in random cascades. *Bernoulli*, 5(2):191–207, 1999.

Steve Uhlig. Conservative cascades: an invariant of internet traffic. In *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (IEEE Cat. No. 03EX795)*, pages 54–57. IEEE, 2003.

Siddharth Venkat and Michel Pleimling. Mobility and asymmetry effects in one-dimensional rock-paper-scissors games. *Physical Review E*, 81(2):021917, 2010.

Vito Volterra. *Variazioni e fluttuazioni del numero d'individui in specie animali conviventi*. C. Ferrari, 1927.

EC Waymire and SC Williams. Multiplicative cascades: dimension spectra and dependence. *Journal of Fourier Analysis and Applications*, 1:589–609, 1995.

Jon A Wellner. Empirical processes: Theory and applications. 2005.

Bin Xu, Hai-Jun Zhou, and Zhijian Wang. Cycle frequency in standard rock–paper–scissors games: evidence from experimental economics. *Physica A: Statistical Mechanics and its Applications*, 392(20):4997–5005, 2013.

C Zhu and G Yin. On hybrid competitive lotka–volterra ecosystems. *Nonlinear Analysis: Theory, Methods & Applications*, 71(12):e1370–e1379, 2009a.

Chao Zhu and G Yin. On competitive lotka–volterra model in random environments. *Journal of Mathematical Analysis and Applications*, 357(1):154–170, 2009b.