2019-01-01

Inverse Gaussian Ornstein-Uhlenbeck Applied To Modeling High Frequency Data

Emmanuel Kofi Kusi
University of Texas at El Paso, ekkusi@miners.utep.edu

INVERSE GAUSSIAN ORNSTEIN - UHLENBECK APPLIED TO MODELING HIGH FREQUENCY DATA.

EMMANUEL KOFI KUSI

Master's Program in Mathematics

APPROVED:

Maria C. Mariani, Chair, Ph.D.

Joe Guthrie, Ph.D.

T. Sarkodie Gyan, Ph.D.

Stephen Crites, Ph.D.
Dean of the Graduate School

to my

MOTHER and FATHER

with love

INVERSE GAUSSIAN ORNSTEIN - UHLENBECK APPLIED TO MODELING HIGH FREQUENCY DATA.

by

EMMANUEL KOFI KUSI

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Mathematical Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

August 2019

Acknowledgements

My sincere gratitude goes to God Almighty for his guidance and confidence that has enabled me to complete this work successfully.

I would like to express my deep-felt gratitude to my advisor, Dr. Maria Christina Mariani of the Mathematical Sciences Department at The University of Texas at El Paso, for her advice, encouragement, enduring patience and constant support. She was never ceasing in her belief in me (though I was often doubting in my own abilities), always providing clear explanations when I was (hopelessly) lost, constantly driving me with energy I was tired, and *always* giving me her time, in spite of anything else that was going on. Her response to my verbal thanks one day was a very modest, “Excellent”. I wish all students the honor and opportunity to experience her ability to perform at that job.

I also wish to thank the other members of my committee, Dr. Joe Guthrie of the Mathematical Sciences Department and Dr. T. Sarkodie Gyan of the Electrical & Computing Engineering Department, both at The University of Texas at El Paso. Their suggestions, comments and additional guidance were invaluable to the completion of this work.

As a special note, Osei Kofi Tweneboah graciously volunteered to act as my academic mentor while He is also studying in PhD. Computational Sciences. He was extremely helpful in providing the additional guidance and expertise I needed in order to complete this work, especially with regard to the chapter on Inverse Gaussian Ornstein Uhlenbeck model and its application to Seismic and Market data.

Additionally, I want to thank The University of Texas at El Paso Mathematical Science Department professors and staff for all their hard work and dedication, providing me the means to complete my degree and prepare for a career as a mathematician. This includes (but certainly is not limited to) the following individuals:

Dr. Amy Wagler

She made it possible for me to have many wonderful experiences I enjoyed while a student, including the opportunity to teach beginning Mathematics students the basics of R and Python (something I wish I had been taught when I first started), and the ability to present

Dr. Joe Guthrie

His influence, though unbeknownst to him, was one of the main reasons for coming to UTEP. He taught me Topology in semester I and I obtain Grade A in His course.

And finally, I must thank my dear sister for putting up with me during the development of this work with continuing, loving support and no complaint. I do not have the words to express all my feelings here, only that I love you, Emelia!

Abstract

With about 226050 estimated deaths worldwide in 2010, earthquake is considered as one of the disasters that records a great number of deaths. This thesis develops a model for the estimation of magnitude of future seismic events.

We propose a stochastic differential equation arising on the Ornstein-Uhlenbeck processes driven by IG(a,b) process. IG(a,b) Ornstein-Uhlenbeck processes offers analytic flexibility and provides a class of continuous time processes capable of exhibiting long memory behavior. The stochastic differential equation is applied to geophysics and financial stock markets by fitting the superposed IG(a,b) Ornstein-Uhlenbeck model to earthquake and financial time series.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Overview

In this chapter I present the background of the study, the problem statement, the objectives of the study, the methodology, the justification (significance of the study) and finally the organization of the study.

1.2 Background of the Study

Seismic Hazard analyses involve the quantitative estimation of ground - shaking at a particular site. Seismic hazards may be analyzed deterministically, as when a specific earthquake scenario is assumed, or probabilistically, in which uncertainties in the size, location and time of occurrence of the earthquake are explicitly considered. There are a lot of researches and analyses currently going globally concerning the geophysical mechanisms that drive seismic events and so of today, a quite number of mathematical models which explore the time of occurrence and earthquake magnitude on a particular site is said to be stochastically dependent.

The motivation for studying this problem comes from continuous stochastic volatility models in financial mathematics. Barndorff - Nielsen and Shephard(2001, 2003) model stock price as a geometric Brownian motion and the diffusion coefficient of this motion as an Ornstein- Uhlenbeck (OU) process that is driven by the Lévy process which is (nonnegative and nondecreasing). In the field of Geophysics and Finance, distributions of returns from turbulent wind speed and financial assets can often be fitted very well by the Inverse Gaussian distribution. It is therefore of some interest, using the inverse Gaussian laws as building blocks to construct Ornstein-Uhlenbeck models on seismic and financial data.

It is therefore suitable to model phenomena where large values are more probable than any other cases. For Stock market returns and prices, a key characteristic is that it models that extremely large variations from typical (crashes) can occur even when almost all (normal) variations are small. With these areas of application in mind, mainly those of finance, we study in recent papers several kind of processes with the Inverse Gaussian marginals, in particular processes of Ornstein - Uhlenbeck type, superpositions of such processes and stochastic volatility models in one, two and several dimensions.

But for the purpose of this project, we will consider to find models with analytically and statistically correlation structure to the occurrences of seismic and financial events. And the feasibility of carrying out likelihood inference under the derived models will be discussed to some extent. The next three chapters discuss important ideas and theorems. In chapter 2, there are some pre - requisites for better understanding of the project which would be given under the preliminaries. This chapter contains various preliminaries, mostly well known, results concerning: Inverse Gaussian distributions; Infinite divisibility and exponential families; Self - decomposability; Ornstein - Uhlenbeck type processes and Background Driving Lévy processes; long range dependence and self - similarity; stochastic volatility and the processes of Inverse Gaussian type.

We would then move to the main body of the work; where we then define the Inverse Gaussian Ornstein - Uhlenbeck process and we characterize the associated BDLP as the sum of two homogeneous Levy processes, one being the inverse Gaussian Levy process and the second a Gamma Levy process.

Chapter 4 discusses the simulation methods and the analysis of the seismic data and discussion of findings. Chapter 5 investigate the model formulation and the analysis of the Inverse-Gaussian OU model and the Gamma OU model by comparison of these model using the seismic data and stock markets data. Finally we shall give conclusion of our work, summary and some recommendations in Chapter 6.

1.3 Introduction of Earthquakes

An earthquake is the sudden movement of the ground that releases elastic energy stored in rocks and generates seismic waves. The elastic energy can be built up and stored over a long time and then released in seconds or minutes. Strain on the rocks results in more elastic energy being stored which leads to far greater possibility of an earthquake event. The sudden release of energy during an earthquake causes low-frequency sound waves called seismic waves to propagate through the earth's crust or along its surface. These elastic waves radiate outward from the "source" and vibrate the ground.

Direct Shaking Hazards and Human-Made Structures

Most earthquake-related deaths are caused by the collapse of structures and the construction practices play a tremendous role in the death toll of an earthquake. In southern Italy in 1909 more than 100,000 people perished in an earthquake that struck the region. Almost half of the people living in the region of Messina were killed due to the easily collapsible structures that dominated the villages of the region. A larger earthquake that struck San Francisco three years earlier had killed fewer people (about 700) because building construction practices were different type (predominantly wood). Survival rates in the San Francisco earthquake was about 98%, that in the Messina earthquake was between 33% and 45% (Zebrowski, 1997). Building practices can make all the difference in earthquakes, even a moderate rupture beneath a city with structures unprepared for shaking can produce tens of thousands of casualties.

Although probably the most important, direct shaking effects are not the only hazard associated with earthquakes, other effects such as landslides, liquefaction, and tsunamis have also played important part in destruction produced by earthquakes.

In an earthquake, the initial movement that causes seismic vibrations occurs when two sides of a fault suddenly slide pass each other. A fault is a large fracture in rocks, across which the rocks have moved. There are three main types of faults, all of which may cause an interplate earthquake namely, normal, reverse (thrust) and strike-slip. An interplate earthquake is an earthquake that occurs at the boundary between two tectonic plates.

Earthquakes of this type account for more than 90% of the total seismic energy released around the world . Plate tectonics is the theory that earth's outer shell is divided into several plates that glide over the mantle, the rocky inner layer above the core. The plates act like a hard and rigid shell compared to the earth's mantle. The strong outer layer is called the lithosphere. If one plate is trying to move pass the other, they will be locked until sufficient stress builds up to cause the plates to slip relative to each other. The slipping process creates an earthquake with land deformations and resulting seismic waves which travel through the earth and along the earth's surface.

1.4 Problem Statement

It has always been a problem for people to estimate future seismic hazards of a particular region. There are quite a number of models which explore the occurrences of future seismic events where the prediction of time, location and magnitude focused on the magnitude of the future seismic events as well as the effects of the seismic hazards. There are few publications on the estimation of the magnitude hazards and its progression to earthquakes of a region. One of the few is by (**Maria et al**; **2015**) which analyzed the Stochastic Differential Equations of Earthquakes Series. There is the need to extend their work to include a new model and to determine the necessary ingredients for the simulation process.

1.5 Objectives of the study

The objectives of the study includes the following;

- To construct an Inverse Gaussian OU model following Maria et al(2016) of the stochastic differential equations applied to high frequency data in Geophysics.
- To perform the (magnitude of future seismic events) Data analysis of the model.
- To perform simulations to determine the effect of varying the model.
- Comparison of these 2 OU model (Inverse Gaussian OU model and Gamma OU model).

1.6 Methodology

In this study, an Ornstein - Uhlenbeck processes driven by IG(a,b) process is developed to investigate the magnitude of earthquake and its effects in a particular region.

We begin by developing the dynamics of Levy processes and Inverse Gaussian OU model adapted from **Mariani et al 2015**. The Self decomposability of the new stochastic differential equations which takes into account earthquake is well investigated. Finally we perform simulation on the solution of the proposed stochastic differential equation using MATLAB.

1.7 Significance of the Study

This study is a step towards understanding the relevance of properties such as Ornstein - Uhlenbeck processes that characterize the estimation of future stochastic events. A clear understanding of the impact of earthquakes as in the magnitude of earthquakes and making suggestion in promoting the prevention of impact of certain regions in the future.

1.8 Organization of Thesis

This study is organized into five chapters and outlined as follows:

1. **Introduction:** This chapter presents a general introduction to the study with a background to the study, the problem statement, objectives, methodology and the significance of the study.

2. **Literature Review:** In this chapter various literatures with relation to history, Ornstein-Uhlenbeck processes, mathematical models and application of control are presented.

3. **Methodology:** This chapter discusses various methods adopted for the study with a focus on definitions as well as a case study.

4. **Model Formulation and Analysis:** In this chapter, the model is formulated. The Inverse Gaussian OU model and Gamma OU model are also well investigated.

5. **Analysis and Simulations:** In this chapter, the magnitude of the model are computed using parameter values from the simulations of the earthquake time series and market crises which are also presented and the results discussed.

6. **Conclusion and Recommendations:** This chapter concludes the entire study and lays out some recommendations for future studies.

Chapter 2

Literature Review

In this chapter, a reviewed literature on the history of OU processes is presented. Some literatures are also presented on two types of OU processes as well as various works on stochastic and lévy processes. Literatures on mathematical models related to earthquakes and stock markets are also presented.

2.1 Stochastic and Lévy Processes

Definition 1. *A stochastic process is a family of random variables* $\{X(t) : t\epsilon T\}$*, where* t *usually denotes time. That is, at every time* t *in the set* T*, a random number* $X(t)$ *is observed.*

Stochastic processes are widely used as mathematical models of systems and phenomena that appear to vary in a random manner. They have applications in many disciplines including sciences such as biology, chemistry, ecology, neuroscience, and physics as well as technology and engineering fields such as image processing, signal processing, information theory, computer science, cryptography and telecommunications. Furthermore, seemingly random changes in financial markets have motivated the extensive use of stochastic processes in finance.

Definition 2. *A càdlàg stochastic process* $Y = \{Y_t\}_{t\geq 0}$ *with* $Y_0 = 0$ *is a Lévy process if and only if it has independent and (strictly) stationary increments .*

Lévy processes are types of stochastic processes that can be considered as generalizations of random walks in continuous time. These processes have many applications in fields such as finance, fluid mechanics, physics and biology. The main defining characteristics are their stationarity and independence properties with stationary and independent increments.

2.1.1 Examples of Stochastic and Lévy Processes

In this section, we will present some examples of Stochastic and Lévy Processes. Then, we look at processes that live on the real line. We shall pay attention also to their density function, their characteristic function, their Lévy triplets, together with some of their properties. We compute moments, variance, skewness and kurtosis, if possible.

The Weiner Process

A stochastic process $W = W_t, t \geq 0$ is a Weiner process (Brownian motion) if the following conditions holds:

i. $W_0 = 0$.

ii. W has stationary increments, i.e. the distribution of the increment $W_{t+s} - W_t$ over the interval $[t, t+s]$.

iii. W has independent increments; i.e. if $1 < s \leq t < u$, $W_u - W_t$ and $W_s - W_l$ are independent random variables. In other words, increments over non - overlapping time intervals are stochastically independent.

iv. For $0 \leq s < t$, $W_{t+s} - W_t$ follows normal distribution with mean 0 and variance $s > 0$: $W_{t+s} - W_t \sim$ Normal(0,s).

The Random Walk

Let $Y_k, k \geq 1$, be *i.i.d.* Then

$$S_n = \sum_{k=1}^{n} Y_k, \qquad n \in \mathcal{N}$$

is a random walk.

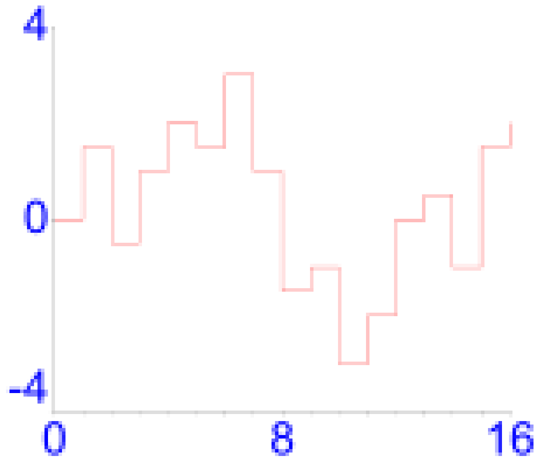

Figure 2.1: Random Walk

Random walks are stochastic processes that are usually defined as sums of independent and identically distributed(iid) random variables in Euclidean space so they are processes that change in discrete time. Random walks can also be referred processes that change in continuous time, particularly the Wiener process used in finance, which has led to some confusion, resulting in its criticism. There are other various types of random walks, defined so their state spaces can be other mathematical objects, such as lattices and groups, and in general they are highly studied and have many applications in different disciplines.

A classic example of a random walk is known as the simple random walk, which is a stochastic process in discrete time with the integers as the state space, and is based on a Bernoulli process, where each Bernoulli variable takes either ± 1.

The Bernoulli Process

One of the simplest stochastic processes is the Bernoulli process, which is a sequence of independent and identically distributed random variables, where each random variable takes either the value one or zero, say one with probability $1 - p$. This process can be linked to repeatedly flipping a coin, where the probability of obtaining a head is p and its value is one, while the value of a tail is zero.

The Poisson Process

The Poisson process is the simplest Lévy Processes we can think of. It is based on the Poisson distribution, which depends on a single parameter λ and has the following characteristic function:

$$\phi_{Poisson}(u; \lambda) = \exp(\lambda(\exp(iu) - 1))$$

The Poisson distribution lives on the nonnegative integers $j = \{0, 1, 2, ...\}$; the probability mass function at point j is given by;

$$f(j, \lambda) = \frac{\lambda^j e^{-\lambda}}{j!}$$

Since the Poisson(λ) distribution is infinitely divisible, we can define $N = \{N_t, t \geq 0\}$ with intensity parameter $\lambda > 0$ as the process which starts at 0, has independent and stationary increments and where the increment over a time interval of length $\delta > 0$ follows a Poisson(λs) distribution. The Poisson process turns out to be an increasing pure jump process, with jump sizes always equal to 1.

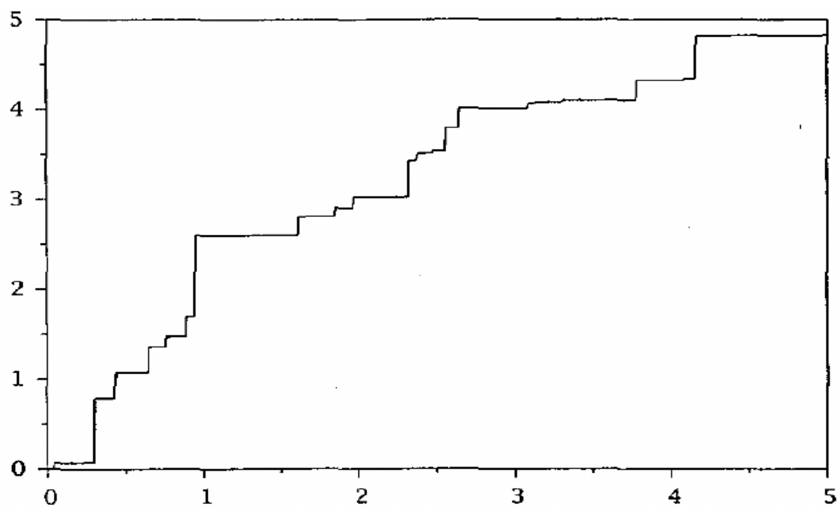

Figure 2.2: Simulation of Poisson Process

Table 2.1: Moments of the Poisson distribution with intensity λ

Properties	Poisson(λ)
Mean	λ
Variance	λ
Skewness	$\frac{1}{\sqrt{\lambda}}$
Kurtosis	$3 + \lambda^{-1}$

The Compound Poisson Process

Suppose $N = \{N_t, t \geq 0\}$ is a Poisson process with intensity parameter $\lambda > 0$ and that $Z_i, i = 1, 2, ...,$ is an independent of N and following a law L, with characteristic function $\phi_Z(u)$. Then

$$X_t = \sum_{k=1}^{N_t} Z_i, \quad t \geq 0$$

is a Compound Poisson Process.

The value of the process at time t, X_t, is the sum of N_t random numbers with law L. The ordinary Poisson Process corresponds to the case where $Z_i = 1, i = 1, 2, ...,$ i.e where the law L is degenerate at the point 1.

Then the characteristic function of X_t is given by;

$$E[\exp(iuX_t)] = \exp(t \int_{-\infty}^{+\infty} (exp(iux) - 1)\nu(dx)) \tag{2.1}$$

$$= \exp(t\lambda(\phi_Z(u) - 1)) \tag{2.2}$$

where ν is called the Lévy measure of process $Z = \{X_t, t \geq 0\}$. ν is called a positive measure on $\Re$ but not a probability measure since $\int \nu(dx) = \lambda \neq 1$.

From this we can easily obtain the Lévy triplets:

$$[\int_{-1}^{+1} x\nu(dx), 0, \nu(dx)]$$

.

The Gamma Process

The density function of the Gamma distribution $\Gamma(a,b)$ with parameters $a > 0$ and $b > 0$ is given by;

$$f_{Gamma}(x;a,b) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-xb), \quad x > 0$$

The density function clearly has a semi - heavy (right) tail. The characteristic function is given by;

$$\phi_{Gamma}(u;a,b) = (1 - iu/b)^{-a}$$

This characteristic function is infinitely divisible. The Gamma process $X^{(Gamma)} = \{X_t^{(Gamma)}, t \geq 0\}$ with the stochastic process which starts at zero and has stationary and independent Gamma distributed increments where $X_t^{(Gamma)}$ follows a $\Gamma(at, b)$ distribution.
The Lévy triplet of the Gamma process is given by;

$$[a(1 - \exp(-b))/b, 0, a\exp(-bx)x^{-1}1_{(x>0)}dx]$$

Table 2.2: Moments of the $\Gamma(a,b)$ distribution

Properties	$\Gamma(a,b)$
Mean	a/b
Variance	a/b^2
Skewness	$2a^{-1/2}$
Kurtosis	$3(1+2a^{-1})$

The Inverse Gaussian Process

Let $T^{(a,b)}$ be the first time a standard Brownian motion with drift $b > 0$, i.e. $\{W_s + b_s, s \geq 0\}$, reaches the positive level $a > 0$. It is well known that this random time follows the so - called Inverse - Gaussian $IG(a,b)$ law and has a characteristics function

$$\phi_{IG}(u;a,b) = \exp(-a\sqrt{-2iu + b^2} - b).$$

The IG distribution is infinitely divisible and we define the IG process $X^{(IG)} = \{X_t^{IG}, t \geq 0\}$, with parameters $a, b > 0$, as the process which starts at 0 and has independent and stationary increments such that

$$\begin{aligned} E[\exp(iuX_t^{(IG)})] &= \phi_{IG}(u; at, b) & (2.3) \\ &= \exp(-at(\sqrt{-2iu + b^2}) - b)). & (2.4) \end{aligned}$$

The density function of the $IG(a, b)$ is explicitly known:

$$f_{IG}(x; a, b) = \frac{a}{\sqrt{2\pi}} \exp(ab) x^{-3/2} \exp(-\frac{1}{2}(a^2 x^{-1} + b^2 x)), \quad x > 0$$

The Lévy measure of the $IG(a, b)$ law is given by;

$$\nu_{IG}(dx) = (2\pi)^{-1/2} a x^{-3/2} \exp(-\frac{1}{2} b^2 x) 1_{(x>0)} dx$$

and the first component of the Lévy triplet equals

$$\gamma = \frac{a}{b}(2N(b) - 1),$$

where the $N(x)$ is the Normal distribution function.

Properties

The density is unimodal with a mode at $\frac{(\sqrt{4a^2b^2+9}-3)}{(2b^2)}$. All positive and negative moments exist. If X follows on $IG(a, b)$ law, we have that

$$E[X^{-\alpha}] = (\frac{b}{a})^{2\alpha+1} E[X^{\alpha+1}], \qquad \alpha \in \mathbb{R}$$

The Generalized Inverse Gaussian Process

The Inverse Gaussian $IG(a, b)$ law can be generalized to what is called the Generalized Inverse Gaussian distribution $GIG(\lambda, a, b)$. This distribution on the positive half line is given in terms of its density function.

$$f_{GIG}(x; \lambda, a, b) = \frac{(b/a)^\lambda}{2K_\lambda(ab)} x^{\lambda-1} \exp(-\frac{1}{2}(a^2 x^{-1} + b^2 x)), \quad x > 0$$

Table 2.3: Moments of the $IG(a, b)$ distribution

Properties	$IG(a, b)$
Mean	a/b
Variance	a/b^3
Skewness	$3(ab)^{-1/2}$
Kurtosis	$3(1 + 5(ab)^{-1})$

The parameters λ, a and b are such that $\lambda \in \mathbb{R}$ while a and b are both nonnegative and not simultaneously 0.

The characteristic function is given by;

$$\phi_{GIG}(u; \lambda, a, b) = \frac{1}{K_\lambda(ab)}(1 - 2iu/b^2)^{\lambda/2} K_\lambda(ab\sqrt{1 - 2iub^{-2}}),$$

where $K_\lambda(x)$ denotes the modified Bessel function of the third kind with index λ.

Barndorff - Nielsen and Halgreen(1977) showed that the distribution is infinitely divisible. We can thus define the GIG process as the Lévy process where the increment over the interval $[s, s + t], s, t \geq 0$ has characteristic function $(\phi_{GIG}(u; \lambda, a, b))^t$.

The Lévy measure has a density on the positive real line is given by;

$$u(x) = x^{-1} \exp(-\frac{1}{2}b^2 x)(a^2 \int_0^\infty \exp(-xz)g(z)dz + max\{0, \lambda\}),$$

where

$$g(z) = (\pi^2 a^2 z (J^2_{|\lambda|}(a\sqrt{2z}) + N^2_{|\lambda|}(a\sqrt{2z}))^{-1}$$

and where J_ν and N_ν are Bessel functions.

Properties

The moments of a random variable X $\frac{(\sqrt{4a^2b^2+9}-3)}{(2b^2)}$ following a $GIG(\lambda, a, b)$ distribution are given by;

$$E[X^k] = (\frac{a}{b})^k \frac{K_{\lambda+k}(ab)}{K_\lambda}, \qquad k \in \Re$$

Table 2.4: Moments of the $GIG(\lambda, a, b)$ distribution

Properties	$GIG(\lambda, a, b)$
Mean	$aK_{\lambda+1}(ab)/(bK_{\lambda}(ab))$
Variance	a/b^2

Martingale

In probability theory, a martingale is a stochastic process (i.e, a sequence of random variables) such that the conditional expected value of an observation at some time t, given all the observations up to some earlier time s, is equal to the observation at that earlier time s.

A discrete-time martingale is discrete-time stochastic process $X_1, X_2, X_3, ...$that satisfies for all n

$$E(\mid X_n \mid) < \infty$$
$$E(X_{n+1} \mid X_1, ..., X_n) = X_n$$

$i.e$, the conditional expected value of the next observation, given all the past observations, is equal to the last observation.

Somewhat more generally, a sequence $Y_1, Y_2, Y_3, ...$ is said to be a martingale with respect to another $X_1, X_2, X_3, ...$

$$E(\mid Y_n \mid) < \infty$$
$$E(\{Y_{n+1} \mid X_1, ..., X_n\}) = Y_n.$$

Similarly, a continuous-time martingale with respect to the stochastic process X_t is a stochastic process Y_t such that for all t.

$$E(\mid Y_t \mid) < \infty$$
$$E(\{Y_t \mid X_\tau, \tau \leq s\}) = Y_s, \forall s \leq t.$$

This expresses the property that the conditional expectation of an observation at time t, given all the observations up to time s, is equal to the observation at time s (provided that $s \leq t$).

Interlacing processes processes

Interlacing processes = Gaussian process + compound Poisson process

Stable Levy processes

• Stable probability distributions arise as the possible weak limit of normalized sums of i.i.d random variable in the central limit theorem.

• Example: Cauchy Process with density(index of stability is 1)

$$f_t(x) = \frac{t}{\pi(x^2 + t^2)}$$

Subordinators

• A subordinator $T(t)$ is a one-dimensional Lévy process that is non-decreasing.

• Important application: time change of Lévy process $X(t)$:

$Y(t) := X(T(t))$ is also a new Lévy process.

Relativistic processes

2.1.2 Properties of Stochastic and Lévy Processes

Property 1. ***Stationarity***

Given $Y_k, \quad k \geq 1$, *be i.i.d. Then*

$$S_n = \sum_{k=1}^{n} Y_k, \qquad n \in \mathbb{N}$$

is a random walk.

We also explained that random walks have stationary and independent increments.

$$Y_k = S_k - S_{k-1}, \qquad k \geq 1$$

Thus Stationarity means the Y_k have identical distribution.

Property 2. ***Independence***

Two stochastic processes X and Y defined on the same probability space $(\sigma, \mathcal{F}, P)$ with the same index set T are said to be independent if for all $n \in \mathcal{N}$ and for every choice of epochs $t_1, ..., t_n \in \mathcal{T}$, the random vectors $(X(t_1), ..., X(t_n))$ and $(Y(t_1), ..., Y(t_n))$ are independent.

Property 3. ***Uncorrelatedness***

Two stochastic processes X_t and Y_t are called uncorrelated if their cross - variance

$$K_{XY}(t_1, t_2) = E[(X(t_1) - \mu_X(t_1))(Y(t_2) - \mu_Y(t_2))]$$

is zero $\forall t$.

Thus, Independence implies uncorrelatedness. If two stochastic processes X and Y are independent, then they are also uncorrelated.

Property 4. ***Orthogonality***

*Two stochastic processes X_t and Y_t are called **orthogonal** if their cross-correlation*

$$R_{XY}(t_1, t_2) = E[X(t_1, \bar{Y_{t_2}})]$$

is zero $\forall t$.

Property 5. ***Separability of a stochastic process***

$X = R_k$ - a measurable space.

$I \subseteq (-\infty, \infty)$ - an interval of a real line.

$\mathcal{G}$ - the class of all subsets of R_k and $\mathcal{H}$ be the class of all open subsets of I.

$\mathcal{T}_1$ - the class of countable subsets of I dense in I.

The stochastic process $Y_t, t \in I$ is called separable if there exists set $\mathcal{J} \in \mathcal{T}_1$ such that for any set $G \in \mathcal{G}$ and $H \in \mathcal{H}$,

$$P\{Y_t \in G, t \in H = PY_t \in G, t \in H \cap J\} \tag{2.5}$$

Thus, relation (2.5) is equivalent to the relation

$$P\{A_{G,H,J} \backslash A_{G,H}\} = 0$$

The set J is called a set of separability for the process Y_t.

Lemma 1.1. *Let $Y_t, t \in I$ is a real - valued process. Then it is separable, if for any open interval $(a, b) \subseteq I$*

$$P\{\sup_{a<t<b} Y_t = \sup_{a<t<b, t\in I} Y_t\} = P\{\inf_{a<t<b} Y_t = \inf_{a<t<b, t\in J} Y_t\} = 1$$

Example 1. *The process $Y_t = I(t = \rho), t \in I$ is not separable while the process $Z_t \equiv I$ is separable and any set $J \in \mathcal{T}_1$ is the set of separability for the process.*

Example 2. *The process $\tilde{\omega}(t) = \omega(t)I(t \neq \tau), t > 0$ is not separable while the process $\omega(t), t \geq 0$ is separable and any set $J \in \mathcal{T}_{(0,\infty)}$ is the set of separability for this process.*
Let $X = R_k$ and $\tilde{R_k} = [-\infty, \infty] \times ... \times [-\infty, \infty]$ be an extended Euclidean space R_k.

Property 6. ***Stochastic continuity of a stochastic process***

A stochastic process $Y_t, t \in I$ with the phase space R_k is stochastically continuous in the point $t_0 \in I$,

$$P\{| Y_t - Y_{t_0} |> \delta\} \to 0 \quad as \quad t \to t_0, \delta > 0$$

Therefore, the process $Y_t, t \in I$ is stochastically continuous if it is stochastically continuous in every point $t_0 \in I$.

Theorem 2.1. *If the process $Y_t, t \in I$ with the phase space R_k is stochastically continuous, then there exists a separable modification $Z_t, t \in I$ for the process Y_t such that any set $J \in \mathcal{T}_1$ is a separability set for the process Z_t.*

Example 1. *The process $Y_t = I(t = \rho), t \in I$ is not separable but it is a stochastically continuous process. Any set $J \in \mathcal{T}_1$ is a separability set for the separable modification $Z_t \equiv 0, t \in I$ for the process Z_t.*

Example 2. *The process* $\tilde{\omega}(t) = \omega(t)I(t \neq \tau), t > 0$ *is not separable but it is a stochastically continuous process. Any set* $J \in \mathcal{T}_{[0,\infty)}$ *is the set of separability for the separable modification* $\omega(t), t \geq 0$ *for the process* $\tilde{\omega}(t)$.

Property 7. ***Indistinguishable***

Two stochastic processes X *and* Y *defined on the same probability space* $(\Omega, \mathcal{F}, P)$ *with the same index set* T *and set space* S *are said to indistinguishable if following*

$$P(X_t = Y_t \quad \forall t \in T) = 1,$$

holds. If two X *and* Y *are modifications of each other and are almost surely continuous, then* X *and* Y *are indistinguishable.*

Property 8. ***Modification***

A modification of a stochastic process is another stochastic process, which is closely related to the original stochastic process. More precisely, a stochastic process X *that has the same index set* T*, set space* S*, and probability space* $(\Omega, \mathcal{F}, P)$ *as another stochastic process* Y *is said to be a modification of* Y *if for all* $t \in T$ *the following*

$$P(X_t = Y_t) = 1,$$

holds. Two stochastic processes that are modifications of each other have the same law and they are said to be ***stochastically equivalent****.*

Property 9. ***Stochastic volatility***

Stochastic volatility is the main concept used in the fields of financial economics and mathematical finance to deal with endemic time-varying volatility and codependence found in financial markets. Such dependence has been known for a long time, early comments include Mandelbrot (1963) and Officer (1973).

Our motivation for studying this research comes from continuous stochastic volatility models in financial mathematics. Stochastic volatility models provide a basis for realistic modeling *of option prices. Thus, It is the time-change by an integrated stationary volatility process, e.g. OU processes driven by subordinators X_t:*

$$\begin{aligned} Y_t &= exp\{-\lambda t\}Y_0 + \int_0^t exp\{-\lambda(t-s)\}dX_{\lambda s} \\ I_t &= \int_0^t Y_s ds \\ Z_t &= B_{I_t} \end{aligned}$$

This model is by Barndorff-Nielsen and Shephard. This and others can be simulated and used for option pricing.

Property 10. ***Infinite divisibility***

Suppose $Y := \{Y_t\}_{t\geq 0}$ is a Lévy process on $\mathbb{R}^d$, then for all $t > 0$, Y must be divisible into $n \geq 2$ i.i.d random variables

$$Y_t = \sum_{k=1}^{n} (Y_{tk/n} - Y_{t(k-1)/n}),$$

since Y_t since these are successive increments (independent) of equal length(stationary).

Property 11. ***Self - decomposability***

Given λ as a positive number. Then, an infinitely divisible distribution F_Y is called self-decomposable, if there exists a random variable $X = X_{t,\lambda}$, such that, for each $t \in \mathbb{R}_+$;

$$\phi_{F_Y}(u) = \phi_{F_Y}(e^{-\lambda t}u)\phi_{F_X}u, \quad u \in \mathbb{R}$$

where $\phi_{F_Y}(u)$ and $\phi_{F_X}(u)$ are the characteristic functions corresponding to F_Y and F_X corresponding. If $\int_{|x|>1} \log(|x|)F_L(dx) < \infty$, then the class of all possible invariant distributions of Y forms the class of all self-decomposable distributions F_Y with the background driven Lévy driven process L.

Theorem 2.2. *Let $v(dx)$ denote the Lévy measure of an infinitely divisible measure F on $\mathbb{R}$. Then the following statements are equivalent:*

1. *F is self-decomposable.*

2. *The functions on the positive half-line given by $v((-\infty, -e^s])$ and $v([e^s, \infty))$ are both convex.*

3. *$v(dx)$ is of the form $v(dx) = u(x)dx$ with $|x|u(x)$ increasing on $(-\infty, 0)$ and decreasing on $(0, \infty)$.*

2.2 Subordination of Lévy process

Subordination was first considered by Bochner (1949) and introduced into finance by Clark (1973). Many Lévy process models can be represented as Brownian motions with an appropriate independent subordinator (see Geman, Madan, and Yor (2001)). Subordination is a transformation of a stochastic process to a new stochastic through random time change by increasing Lévy process (subordinator) independent of the original process. The new process is called subordinate to the original one.

Definition 3. *A subordinator* $T := \{T_t\}_{t\geq 0}$ *is a one-dimensional Lévy process such that* $t \to T_t$ *is non-decreasing. Since* $T_0 = 0$ *all subordinators take nonnegative values only.*

Proposition 1. *A Lévy process* T *on* $\mathbb{R}$ *is a subordinator iff its Lévy triple has the form* $(a, 0, m)$*, where* $m((-\infty, 0)) = 0$ *and* $\int_0^\infty (1 \wedge x)m(dx) < \infty$.

2.3 The Lévy - Ito Decomposition: Structure of the Sample Paths of Lévy Processes

$$X(t) - bt + B_a(t) + \int_{|x|<1} x\tilde{N}(t, dx) + \int_{|x|\geq 1} xN(t, dx)$$

where $\tilde{N}(t, dx) = N(t, dx) - tv(dx)$ is a compensation Poisson random measure with intensity $v(dx)$.

Application to Finance:

• Replace Brownian motion in BSM model with a more general Levy process (P. Carr, H. Geman, D. Madan and M. Yor) . Note the following:

1. small jumps terms describes the day-to-day jitter that causes minor fluctuations in stock prices.

2. big jumps term describes large stock price movements caused by major market upsets arising from, e.g., earthquakes, etc.

Main problems with Lévy Processes in Finance

• Market is incomplete, i.e. there may be more than one possible pricing formula.

• One of the methods to overcome the problems associated with Lévy Processes in Finance is the entropy minimization.

Example; hyperbolic Lévy process(E. Eberlain) (with no Brownian motion part); a pricing formula have been developed that has minimum entropy.

2.4 Deterministic Differential Equations

A differential equation is a mathematical equation that relates some function of one or more variables with its derivatives. Differential equations arise whenever a deterministic relation involving some continuously varying quantities (modeled by functions) and their rates of change in space and/or time (expressed as derivatives) is known or postulated. Because such relations are extremely common, differential equations play a prominent role in many disciplines including engineering, physics, economics, and biology.

Differential equations are mathematically studied from several different perspectives, mostly concerned with their solutions - the set of functions that satisfy the equation. Only the simplest differential equations are solvable by explicit formulas; however, some properties of solutions of a given differential equation may be determined without finding their exact form. If a self-contained formula for the solution is not available, the solution may be numerically approximated using computers. The theory of dynamical systems puts emphasis on qualitative analysis of systems described by differential equations, while many numerical methods have been developed to determine solutions with a given degree of accuracy.

A deterministic (ordinary) differential equation is an equation involving a function (of one variable) and its derivatives. They are essential for a mathematical description of nature - they lie at the core of many physical theories. For example, let us just Newton's and Lagrange's equations for classical mechanics, Maxwell's equations for classical electromagnetism, Schrodinger's equation for quantum mechanics, and Einstein's equation for the general theory of gravitation.

Examples:

(a) **Newton's law:** Mass times acceleration equals force, $ma = f$, where m is the particle mass, $a = d^2x/dt^2$ is the particle acceleration, and f is the force acting on the particle. Hence Newton's law is the differential equation

$$m\frac{d^2x}{dt^2}(t) = f(t, x(t), \frac{dx}{dt}(t)),$$

where the unknown is $\mathbf{x}(t)-$ the position of the particle in space at the time t. As we see above, the force may depend on time, on the particle position in space, and on the particle velocity.

(b) **Radioactive Decay:** The amount u of a radioactive material changes in time as follows,

$$\frac{du}{dt}(t) = -ku(t), \quad k > 0,$$

where k is a positive constant representing radioactive properties of the material.

(c) **The Heat Equation:** The temperature T in a solid material changes in time and in three space dimensions-labeled by $\mathbf{x} = (x, y, z)-$ according to the equation

$$\frac{\partial T}{\partial t}(t, x) = k(\frac{\partial^2 T}{\partial x^2}(t, x) + \frac{\partial^2 T}{\partial y^2}(t, x) + \frac{\partial^2 T}{\partial z^2}(t, x)), \quad k > 0,$$

2.5 Stochastic Differential Equations

A stochastic differential equation(SDE) is a differential equation in which one or more of the terms is a stochastic process resulting in a solution which is itself a stochastic process. Consider the deterministic diffferential equation:

$$dx(t) = k(t, x(t))dt, \quad x(0) = x_0. \tag{2.6}$$

The easiest way to introduce randomness in the equation is to randomize the initial condition. The solution $x(t)$ then becomes a stochastic process $\{X_t, t \in [0, T]\}$:

$$dX_t = k(t, X_t)dt, \quad X_0(\omega) = Y(\omega). \tag{2.7}$$

Such an equation is called a random differential equation. Random differential equations can be considered as a deterministic equation with a perturbed initial condition. The investigation is of great interest if we study the robustness of the solution to a differential equation under a small change of the initial condition.

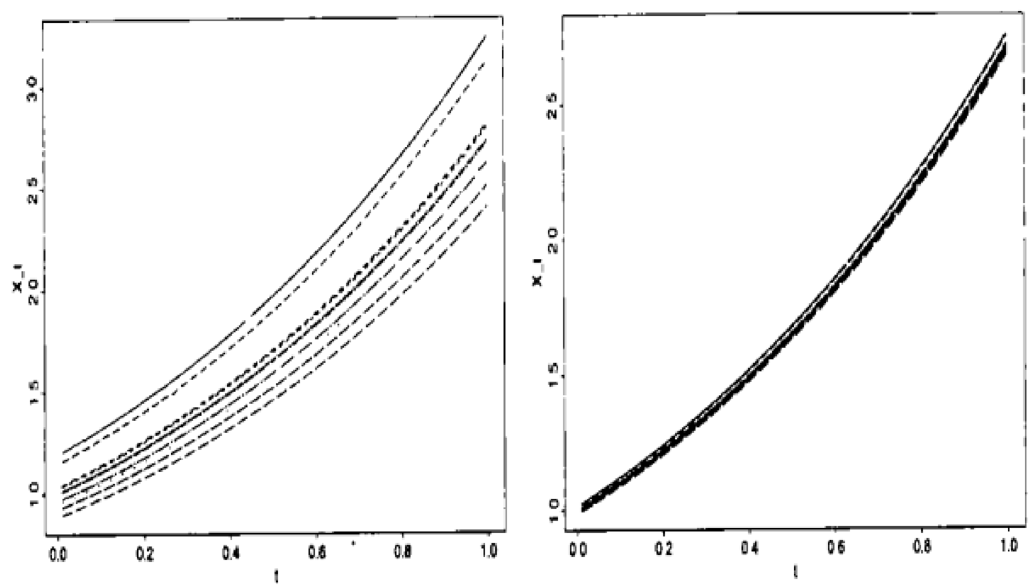

Figure 2.3: The solutions $X_t = X_0 e^t$ to the random differential equation $dX_t = X_t dt$ with initial condition $X_0 = e^N$, where N has an $N(0, \sigma^2)$ distribution with $\sigma^2 = 0.01$ and 0.0001 respectively

In this section, the randomness in the differential equation is introduced via an additional random noise term:

$$dX_t = k(t, X_t)dt + b(t, X_t)dB_t, \quad X_0(\omega) = Y(\omega) \tag{2.8}$$

$B = (B_t, t \geq 0)$ denotes Brownian motion, and $k(t, x)$ and $l(t, x)$ are deterministic functions. The process X, if it exists, is then a stochastic process. The randomness of $X = (X_t, t \in [0, T])$ is as a result of the initial condition and the noise generated by Brownian motion.

Interpretation of Equation 2.8 tells us that the change $dX_t = X_{t+dt} - X_t$ is caused by a change dt of time, with factor $b(t, X_t)$. Since the Brownian motion, with factor $k(t, X_t)$ in a combination with a change $dB_t = B_{t+dt} - B_t$ of Brownian motion, with factor $l(t, X_t)$. Since the Brownian motion does not have differentiable sample paths, the question that normally arises is in which sense can we interpret Equation (2.8). Applying integration on both sides of Equation (2.8), we obtain

$$X_t = X_0 + \int_0^t k(s, X_s)ds + \int_0^t l(s, X_s)dB_s, \quad 0 \leq t \leq T, \tag{2.9}$$

Where the first integral on the right-hand side is a Riemann integral, and the second one is a Ito stochastic process satisfying Equation (2.9).
We will proceed to prove the existence, in a certain sense of

$$\int_0^t f(s, \omega)dB_s(\omega) \tag{2.10}$$

where $B_t(\omega)$ is a 1-dimensional Brownian motion starting at the origin, for a wide class of functions $f : [0, \infty] \times \Omega \to \mathbb{R}$.

2.5.1 Ito Integral

We continue with the construction of Ito integral using the following propositions.

Proposition 2. *Suppose a sequence of simple processes X_n satisfies (1). There exists a process $Z_t \in M_{2,c}$ satisfying $\lim_n E[(Z_t - I_t(X^n))^2] = 0$ for all $0 \leq t \leq T$. This process is unique a.s. in the following sense: if X_t^n is another process satisfying (1) and $\acute{Z}$ is the corresponding limit, then $P(\acute{Z}_t = Z_t, \forall t \in [0, T]) = 1$.*

Now we can formally state the definition of Ito integral.

Definition 1 (Ito integral). *Given a stochastic process $X_t \in \mathcal{L}_2$ and $T > 0$, its Ito integral $I_t(X), t \in [0, T]$ is defined to be the unique process Z_t constructed in proposition 2.*

We have defined Ito integral as a process which is defined for all $t \geq 0$, by taking $T \to \infty$ and taking approximate limits.

2.5.2 Properties of the Ito integral

Theorem 2.5.2. *Let $f, g \in V(S, T)$ and let $0 \leq S < U < T$. Then*

1. $\int_S^T f dB_t = \int_S^U f dB_t + \int_U^T f dB_t.$
2. $\int_S^T (cf + g) dB_t = c \int_S^T f dB_t + \int_S^T g dB_t, \quad for \quad c \in \mathbb{R}.$
3. $E[(\int_S^T f dB_t)] = 0.$
4. $\int_S^T f dB_t \quad is \quad \mathcal{F}_T - measurable.$

2.6 Ito process, Ito formula

An Ito process or stochastic integral is a stochastic process on $(\Omega, \mathcal{F}, P)$ adopted to $\mathcal{F}_t$ which can be written in the form

$$X_t = X_0 + \int_0^t U_s ds + \int_0^t V_s dB_s \tag{2.11}$$

where $U, V \in \mathcal{L}_2$. As a shorthand notation, we will write (equation 11) as

$$dX_t = U_t dt + V_t dB_t$$

Ito formula

In the previous sections, we have observed that a sample Brownian path is nowhere differentiable with probability 1. In other words, the differentiation

$$\frac{dB_t}{dt}$$

does not exist. However, while studying Brownian motions, or when using Brownian motion as a model, the situation of estimating the difference of a function of the type

$$f(B_t)$$

over an infinitesimal time difference occurs quite frequently (suppose that f is a smooth function). To be more precise, we are considering a function $f(t, B_t)$ which depends only on the second variable. Hence there exists an implicit dependence on time since the Brownian motion depends on time.

We now introduce the most important formula of Ito calculus:

Theorem 1 (Ito formula). *Let $X_t$be an Ito process $dX_t = U_t dt + V_t dB_t$. Suppose $g(x) \in C^2(\mathbb{R})$ is a twice continuously differentiable function (in particular all second partial derivatives are continuous functions). Suppose $g(X_t) \in \mathcal{L}_2$. Then $Y_t = g(X_t)$ is again an Ito process and*

$$dY_t = \frac{\partial g}{\partial x}(X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_t)(dX_t)^2$$

Using the notational convention for $dX_t = U_t dt + V_t dB_t$ and $(dX_t)^2$, we can rewrite the Ito formula as

$$dY_t = (\frac{\partial g}{\partial x}(X_t)U_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(X_t)V_t^2)dt + \frac{\partial g}{\partial x}(X_t)V_t dB_t.$$

Thus, we see that the space of Ito processes is closed under twice - continuously differentiable transformations.

2.6.1 Multidimensional Ito formula

There is a very useful analogue of Ito formula in many dimensions. Before turning to the formula we need to extend our discussion to the case of Ito processes with respect to many dimensions, as so far we have considered Ito integrals and Ito processes with respect to just one Brownian motion. Thus suppose we have a vector of d independent Brownian motions $B_t = (B_{i,t}, 1 \leq i \leq d, t \in \mathbb{R}_+)$. A stochastic process X_t is defined to be an Ito process with respect to B_t if there exists $U_t \in \mathcal{L}_2$ and $V_{i,t} \in \mathcal{L}_2, 1 \leq i \leq d$ such that $X_t = U_t dt + \sum_i V_{i,t} dB_{i,t}$, in the sense explained above.

Theorem 2.. *Suppose $dX_t = U_t dt + V_t dB_t$, where vector $U = (U_1, ..., U_d)$ and matrix $V = (V_{11}, ..., V_{dd})$ have $\mathcal{L}_2$ components and B is the vector of d independent Brownian motions. Let $g(x)$ be twice continuously differentiable function from $\mathbb{R}^d$ into $\mathbb{R}$. Then $Y_t = g(X_t)$ is also an Ito process and*

$$dY_t = \sum_{i=1}^{d} \frac{\partial g}{\partial x_i}(X_t) dX_{i,t} + \frac{1}{2} \sum_{i,j=1}^{d} \frac{\partial^2 g}{\partial x_i x_j}(X_t) dX_{i,t} \cdot dX_{j,t},$$

where $dX_{i,t} \cdot dX_{j,t}$ is computed using the rules $dtdt = dtdB_t = dB_i dt = 0, dB_i dB_j$ for all $i \neq j$ and $(dB_i)^2 = dt$.

2.7 Seismic Events and Financial Markets : The Source of High Frequency Data

A famous climber, when asked why he was willing to put his life in danger to climb dangerous summits, answered : "Because they are there". We would be tempted to give the same answer when people ask us why we take so much pain in dealing the high - frequency data. The reason is very simple: Financial markets and Seismic occurrences are the source of high frequency data.

By its very nature it is irregularly spaced in time, however, and with the sheer volume being reported by liquid markets can only be understood using continuous dynamics (Hanif and Protopapas, 2013). Financial data providers usually report hundreds of thousands of prices for a single market a day.

Dacorogna et al.(2001) argue, correctly, that high frequency data should be primary object of research for those who are interested in understanding seismic events such as (Floods, explosions, earthquakes), and financial markets, especially given the effect of market dynamics on everyday investors.

Yet, most of the published papers and studies in the field of geophysics and finance deal with a lower frequency, regularly spaced data because of these two reasons:
Firstly, it is costly and resource-intensive to collect, store, manipulate and curate high frequency data.This is precisely why most available data is either daily or lower frequencies. The second season is somehow more subtle but still quite important: most of the statistical apparatus has been developed and thought for homogeneous (equally spaced in time) time series. Nowadays with the development of computer technology, data availability is becoming less and less of as problem.Little work has been done to look into irregular data with (Hanif and Protopapas, 2013), described above, working towards bridging the gap between regularized and irregular time series.

Chapter 3

Methodology

3.1 Introduction

The aim of this chapter is to discuss the methods adopted for this study. Therefore, this chapter focuses on the necessary ingredients for the simulation process which entails the Ornstein - Uhlenbeck process introduced by [] for modeling the volatility coefficient in a stock price process that follows a geometric Brownian motion. We will discuss modeling issues of Levy driven OU processes and present some simulation results.

Definition 1. *A stochastic process is a family of random variables* $\{X(t) : t\epsilon T\}$, *where* t *usually denotes time. That is, at every time* t *in the set* T, *a random number* $X(t)$ *is observed.*

3.1.1 Arithmetic Brownian motion

A Brownian motion with drift is called arithmetic Brownian motion(ABM). The actual model of ABM is a stochastic differential equation (SDE) of this form;

$$dX = \mu dt + \sigma dz$$

. The model has two parameters;

1. Drift

2. Volatility (sometime also known as the diffusion coefficient)

3.1.2 Geometric Brownian motion

In the modeling of seismic data and financial market, especially earthquake series and stock market, Brownian Motion play a significant role in building a statistical model. In this section, we will explore some of the technique to build a seismic and financial model called the Inverse Gaussian Ornstein-Uhlenbeck model using Brownian Motion and write our Matlab code for simulation and model building.

Before the building our code in Matlab, we will introduce some concepts here in order to understand the in depth process of Brownian Motion in the Ornstein-Uhlenbeck model.

Definition 2. *A stochastic process S_t is said to follow a Geometric Brownian Motion if it satisfies the following stochastic differential equation:*

$$dS_t = \mu S_t dt + \sigma S_t W_t \tag{3.1}$$

Where W_t is a Wiener process(Brownian Motion) and μ, σ are constants.

Normally, μ is called the percentage drift and σ is called the percentage volatility. So, consider Brownian motion trajectory that satisfy the differential equation, the right hand side term $\mu S_t dt$ controls the "trend" of this trajectory and the term $\sigma S_t dW_t$ controls the "random noise" effect in the trajectory.

Since it is a differential equation, we want to find a solution by applying the separation of variables technique, then the equation (3.1) becomes:

$$\frac{dS_t}{S_t} = \mu dt + \sigma W_t \tag{3.2}$$

Then take the integration of both side

$$\int \frac{dS_t}{S_t} = \int (\mu dt + \sigma dW_t) dt \tag{3.3}$$

Since $\frac{dS_t}{S_t}$ relates to derivative of $\ln(S_t)$, then the equation (3.3) involving the $It\bar{o}$ calculus gives us equation (3.4);

$$\ln(\frac{dS_t}{S_t}) = (\mu - \frac{1}{2}\sigma^2)t + \sigma W_t \tag{3.4}$$

Taking the exponential of both side of equation (3.4) and plugging the initial condition S_0, we obtain the solution. The analytical solution of this geometric brownian motion (equation (3.1)) is given by:

$$S_t = S_0 exp((\mu - \frac{\sigma^2}{2})t + \sigma W_t)$$

Thus, the process above is of solving a stochastic differential equation, and in fact, geometric brownian motion is defined as a stochastic differential equation.

3.2 Definition and existence of Ornstein-Uhlenbeck Process

A continuous time stationary and non-negative stochastic process $X = \{X(t)\}$ is said to be an Ornstein - Uhlenbeck process if it satisfies the stochastic differentiation

$$\begin{cases} dX(t) = & -\lambda X(t) + \sigma dL(t) \\ X(0) = & X_0 \end{cases}$$

where $L = \{L(t)\}$ is a Gaussian process with mean 0 and variance $t\sigma^2$, λ is a strictly positive intensity parameter and X_0 is an independent random variable. One possible generalization of this process emerges from allowing L to be a Levy process. Such a model is called a process of Ornstein - Uhlenbeck type. The L is termed as the background driving Levy process (BDLP).

In our work, we will be mainly interested in stationary processes of this type, which can be generated by imposing certain conditions to the BDLP L or more interestingly by designing a stationary self-decomposable law, say D, and finding then a BDLP L that exactly matches this distribution.

3.3 Solution of the Ornstein-Uhlenbeck Processes

As the OU process is used to model mean reverting behavior, we have $\lambda > 0$ and $t > 0$; The solution of the stochastic differential equation (SDE) can be found by applying Ito's lemma to $V(t, X_t) = e^{\lambda t} X_t$

The Ito's Lemma

$$\begin{aligned} d(v(t, X_t)) &= [\frac{\partial V}{\partial t}(t, X_t) + a(t, X_t)\frac{\partial V}{\partial X_t}(t, X_t) + \frac{1}{2}(b(t, X_t))^2 \frac{\partial^2 V}{\partial X_t^2} d(t, X_t)]dt + b(t, X_t)\frac{\partial V}{\partial X_t} dL_t \\ d(e^{\lambda t} X_t) &= [\lambda X_t e^{\lambda t} - \lambda X_t e^{\lambda t}]dt + e^{\lambda t} L_t \\ &= e^{\lambda t} L_t \end{aligned}$$

Time integration gives then

$$e^{\lambda t} X_t - X_t = \int_0^t e^{\lambda s} L_s$$

Dividing through by $e^{\lambda t}$ yields,

$$X_t = e^{\lambda t} X_0 + \int_0^t e^{\lambda(t-s)} dL_s$$

for the initial condition $X_0 = x_0$.

3.4 Inverse Gaussian Ornstein Uhlenbeck - Model and Parameter Estimation

3.4.1 Estimation of the shape parameter a and rate parameter b of the IG(a,b) Ornstein - Uhlenbeck Model

We also need to know how the parameters μ and σ^2 relate to a and b.

Proposition 3. *Suppose that $\{L_t\}_{t\geq 0}$ is a Levy process such that $E(L(1)) = \mu < \infty$ and $Var(L(1)) = \sigma^2 < \infty$. Let M be the largest constant satisfying Equation ... and assume that $\lambda > 0$. Then the following holds:*

1. $E(X_0) = \mu$
2. $Var(X_0) = \frac{\sigma^2}{2}$

We have that

$$a = \mu\sqrt{\frac{2\mu}{\sigma^2}} \quad and \quad b = \sqrt{\frac{2\mu}{\sigma^2}} \tag{3.5}$$

3.4.2 Estimation the mean reverting parameter λ of the IG(a,b) Ornstein - Uhlenbeck Model

Regarding the λ parameter, we recall our estimators: In order to estimate the mean reverting parameter λ, we begin with the definition of the autocorrelation function for the process given in Equation 3.4 which is of the form,

$$\begin{aligned}\rho(h) &= corr(X_{t+h}, X_t) = e^{-\lambda h}, \quad for \quad h \in \mathcal{N} \\ \lambda &= -\log_e \rho(h)\end{aligned}$$

without loss of generality, we take $h = 1$. The our estimated intensity parameter is

$$\lambda = -log(\hat{\rho}(1)) \tag{3.6}$$

3.4.3 Estimation of the arrival times of a Poisson Process $N = (N_s)_{s\geq 0}$ of rate λ for the IG(a,b) - OU simulation

Theorem 3.1. *For a Poisson process of rate λ, and for any $t > 0$, the PMF for $N(t)$ (i.e the number of arrivals in $(0, t]$) is given by the Poisson PMF,*

$$P_{N(t)}(n) = \frac{(\lambda t)^n \exp(-\lambda t)}{n!} \tag{3.7}$$

Therefore the arrival times of a Poisson Process for our simulation is given by;

$$c_i = \frac{(\lambda t)^i \exp(-\lambda t)}{i!}$$

where $c_1 < c_2 < c_3 < ...$ with intensity parameter $\frac{ab}{2}$.

Chapter 4

Model Formulation and Numerical Simulation

4.1 Introduction

In this section we discuss modeling issues of Levy driven Ornstein -Uhlenbeck processes and present some simulation results for an inverse Gaussian OU process. We use the simulated data to test the performance for the simulation process.

A very important ingredient in modeling of Levy driven OU processes is the connection between the Levy density of the stationary distribution of X to the Levy density of the probability law of L_1. In particular we have the following proposition.

Proposition 4. *Assume that the Levy density of X, $v_X(z)$, is differentiable and denote the Levy density of the probability law of L_1 by $v_X(z)$. Then the following relation holds.*

$$v_L(x) = -v_X(x) - xv_X^{'}(x). \tag{4.1}$$

Proof. It follows directly by the fact that the stationary solution, X, to (3.1) satisfies

$$X =^D \int_0^\infty e^{-\lambda s} dL(\lambda s).$$

Hence, given $v_L(x)$ we can find $v_Y(x)$ and vice versa.

4.2 The Inverse Gaussian OU model

An **Ornstein-Uhlenbeck** process $X = (X_t)_{t\geq 0}$ is a decision of the stochastic differential equation

$$dX_t = -\lambda X_t dt + dL_{\lambda t}, \quad X_0 > 0$$

We denote the Lévy measure of L_1 by $W(dz)$ such that Lévy density $u(z)$ of the marginal law D be differentiable, then the Lévy measure W has a density $w(z)$, and $u(z)$ and $w(z)$ are related by

$$w(z) = -u(z) - zu'(z)$$

We define the tail mass function of W(dz) as:

$$W^{+}(z) = \int_{z}^{\infty} w(y)dy$$

The inverse function of $W^{+}(z)$ is of the form:

$$W^{-1}(z) = inf\{y > 0 : W^{+}(y) \leq z\}$$

We develop the Inverse Gaussian OU model using the inverse tail mass function. Thus the Invere Gaussian OU process by the series representation via the inverse tail mass function can be represented using the following approximation.

4.3 Simulation techniques

In this section, the inverse Gaussian OU model developed earlier using the inverse tail mass function. Simulations are performed using Matlab based on the solution of our proposed stochastic differential equation.

$$\begin{aligned} X(t) &= X_0 e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)} dL_{\lambda s} && (4.2) \\ &= X_0 e^{-\lambda t} + e^{-\lambda t} \int_0^t e^{s} dL_{\lambda} && (4.3) \end{aligned}$$

4.4 Simulation via the inverse tail mass function

An Inverse - OU process by the series representation via the inverse tail mass function can be presented using the following approximation.

$$W^{-1}(z) \sim \frac{a^2}{2\pi z^2}$$

Algorithm

1. Simulate a Poisson process $N = (N_s)_{s \geq 0}$ with intensity parameter $\frac{ab}{2}$.

2. Simulate the independent uniform random numbers $u_i \sim Uniform(0, 1), i = 1, 2, ...$

3. Sample the path of the IG - OU process $X = (X_t)_{t \geq 0}$

$$X_{n\Delta t} = e^{-\lambda \Delta t} X_0 + e^{(-\lambda \Delta t)} \sum_{i=1}^{N_t} \frac{a^2 (\Delta t)^2}{2\pi c_i^2} e^{(\lambda t u_i)}, \quad X_0 > 0$$

where c_i are the arrival times of a poisson process $N = (N_s)_{s \geq 0}$ with intensity parameter $\frac{ab}{2}$.

Table 4.1: Parameter Descriptions for Chile Earthquake Data

Regions	**Number of Observations**	λ	**a**	**b**	μ	σ^2
Region 1	389	2.1766	27.0669	5.8736	4.6082	0.2671
Region 2	1167	3.0955	29.8684	6.5759	4.5421	0.2101
Region 3	892	1.6348	26.3897	5.7509	4.5888	0.2775
Region 4	5000	1.6344	30.3151	6.7865	4.4670	0.1940

Table 4.2: Chile Data Results for Superposed $\Gamma(a, b)$ Ornstein - Uhlenbeck model with $\Delta t = 0.0001$

Regions	λ_1	λ_2	X_1	X_2	w_1	w_2	**RMSE**
Region 1	2.1766	2.9766	4.3000	4.5000	0.5000	0.5000	0.1066
Region 2	3.0955	4.6150	4.0000	4.5000	0.4000	0.6000	0.0948
Region 3	1.6348	2.3480	4.7000	4.6000	0.4500	0.5500	0.1017
Region 4	1.6344	2.6344	4.6000	4.5000	0.4000	0.6000	0.0941

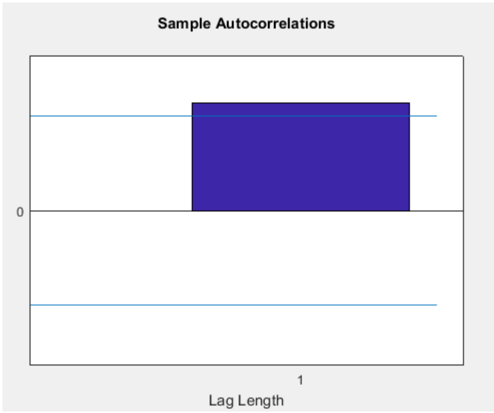

Figure 4.1: $\rho(1) = 0.1133$

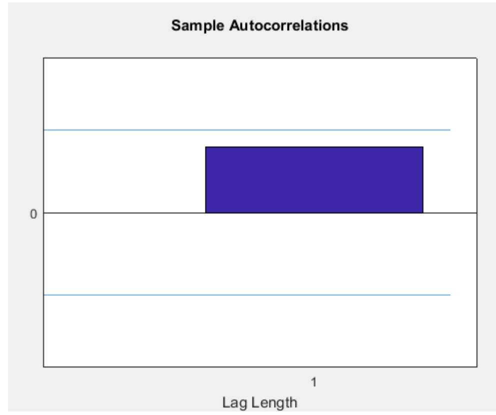

Figure 4.2: $\rho(1) = 0.0453$

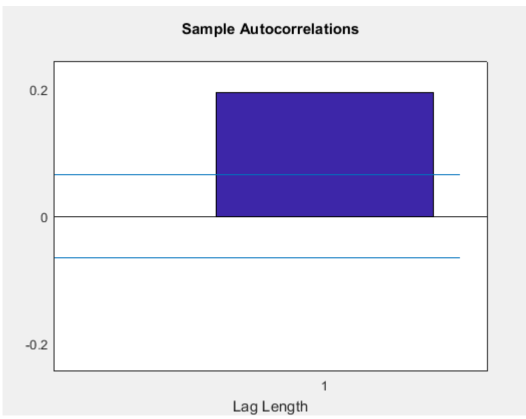

Figure 4.3: $\rho(1) = 0.1950$

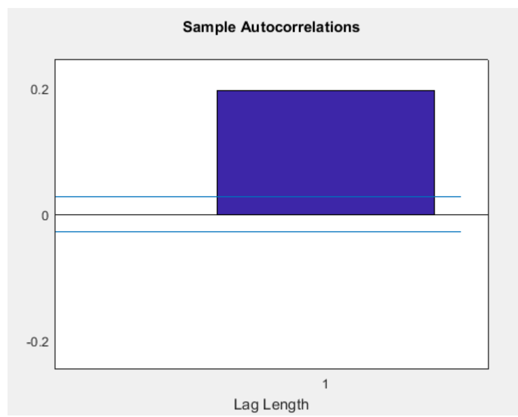

Figure 4.4: $\rho(1) = 0.1951$

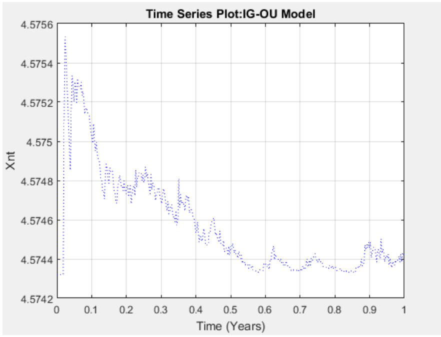

Figure 4.5: $\Delta t = 0.0001$

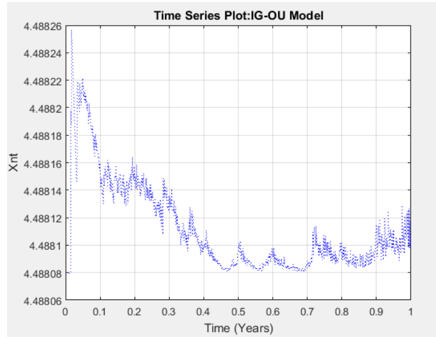

Figure 4.6: $\Delta t = 0.0001$

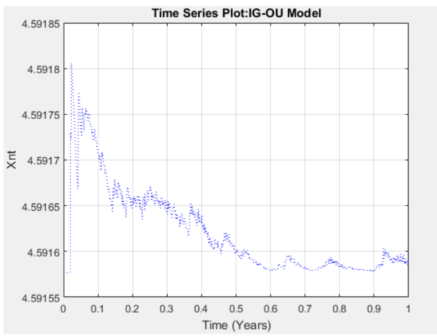

Figure 4.7: $\Delta t = 0.0001$

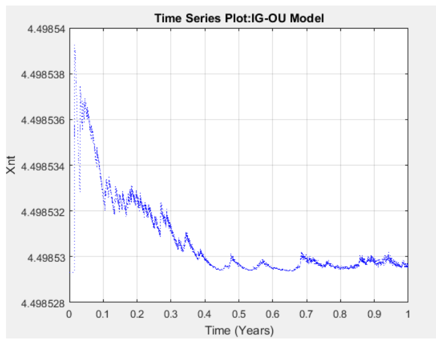

Figure 4.8: $\Delta t = 0.0001$

Table 4.3: Chile Data Results for Inverse Gaussian IG (a,b) Ornstein - Uhlenbeck model with $\Delta t = 0.0001$

Regions	$\lambda = -\log(\rho(1))$	$X_0 = \bar{X}_{n\Delta t}$	Root Mean Square Error
Region 1	2.1766	4.6000	0.1041
Region 2	3.0955	4.5000	0.0952
Region 3	1.6348	4.6000	0.1077
Region 4	1.6344	4.5000	0.0935

Table 4.4: Chile Data Results for Superposed $\Gamma(a, b)$ OU - model with $X_0 = \overline{X_1 + X_2}$

Regions	λ_1	λ_2	X_1	X_2	w_1	w_2	**RMSE**
Region 1	2.1766	2.9766	4.2000	5.0000	0.5000	0.5000	0.1054
Region 2	3.0955	4.6150	4.0000	5.0000	0.4000	0.6000	0.1149
Region 3	1.6348	2.3480	4.5000	4.7000	0.4500	0.5500	0.1201
Region 4	1.6344	2.6344	4.4000	4.6000	0.4000	0.6000	0.3499

Table 4.5: Parameter Descriptions for Emergent / Developed Market Asian Crises Data

Indices	**No. of Observations**	λ	**a**	**b**	μ	σ^2
IGPA	806	0.0021	700.7189	0.1513	4.6304e+03	4.0493e+05
NASDAQ	2727	7.7324e-04	83.2849	0.1047	795.1966	1.4498e+05
SETI	1059	0.0074	108.6003	0.2887	376.1272	9.0234e+03
XU100	1014	0.0032	196.6306	0.0253	7.7622e+03	2.4193e+07
MXX	2476	0.0015	182.0223	0.0486	3.7452e+03	3.1711e+06
MERVAL	1250	0.2285	120.0396	0.2184	549.5960	2.3042e+04
BOVESPA	2100	0.0017	213.9919	0.0254	8.4349e+03	2.6211e+07
HSI	2674	0.0021	384.9144	0.0383	1.0052e+04	1.3712e+07

Table 4.6: Emergent and Developed Stock Markets Data Results for Inverse Gaussian IG (a,b) Ornstein-Uhlenbeck model with $\Delta t = 0.0001$

Financial Indices	$\lambda = -\log(\rho(1))$	$X_0 = \bar{X}_{n\Delta t}$	Root Mean Square Error
IGPA	0.0021	4.6304e+03	0.1475
NASDAQ	7.7327e-04	795.1966	0.4965
SETI	0.0074	376.1272	0.2450 ‘
XU100	0.0032	7.7622e+03	0.7357
MXX	0.0015	3.7452e+03	0.5134
MERVAL	0.2285	549.5960	0.2701
BOVERPA	0.0017	8.4349e+03	1.4658
HSI	0.0021	1.0052e+04	0.4319

Chapter 5

Model Simulations and Analysis

5.1 Analysis of geophysical time series

In this chapter, the Inverse Gaussian OU model developed earlier in chapter 4 is analyzed by using the fitted model applied to real seismic data series from Chile. We normalized the data sets, by taking the logarithm of the time series data point. Then we also simulated independent paths of our model using different time steps for the data sets. Simulations are performed using Matlab to vary model variables and assess the estimated parameters.

We compared our model, that is the $IG(a, b)$ Ornstein - Uhlenbeck model to $\Gamma(a, b)$ Ornstein- Uhlenbeck model to check which of them best fits the data. In order to investigate our model fit, we computed the root mean square error for each region. The root mean square error indicates how well fitted is our model with respect to the given data set. The solution to our stochastic differential equation is a Lévy model, so the very good fitting obtained strengthens the previous results.

5.2 Chile earthquake time series

The high frequency geophysical data was obtained from 4 regions in Chile from the year 2000 to 2014. The data contain information about the location of events, date, and magnitudes of each 42 recorded earthquake in the region. The 4 regions under study were selected based on the fact that it recently generated macro- earthquakes ($M_w \geq 8$). The sizes of the regions were selected according to the distribution of aftershocks after the largest mega-earthquakes. Region 3 is of much interest since it shows clear fore-shocks. Figures 4.1,4.2,4.3 and 4.4 shows the map of the 4 regions under study. The earthquakes magnitude is the recorded data used in our analysis.

Figure 5.1: Region 1

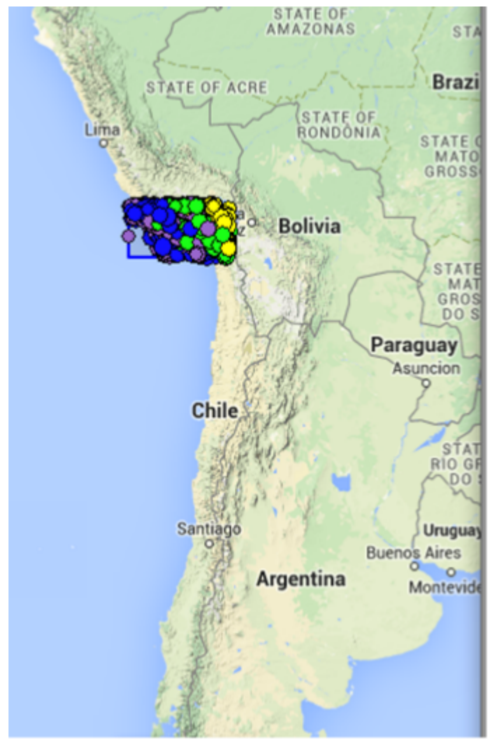

Figure 5.2: Region 2

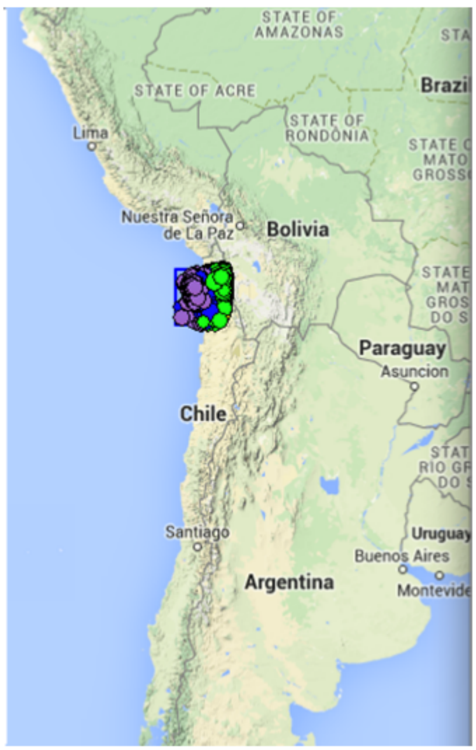

Figure 5.3: Region 3

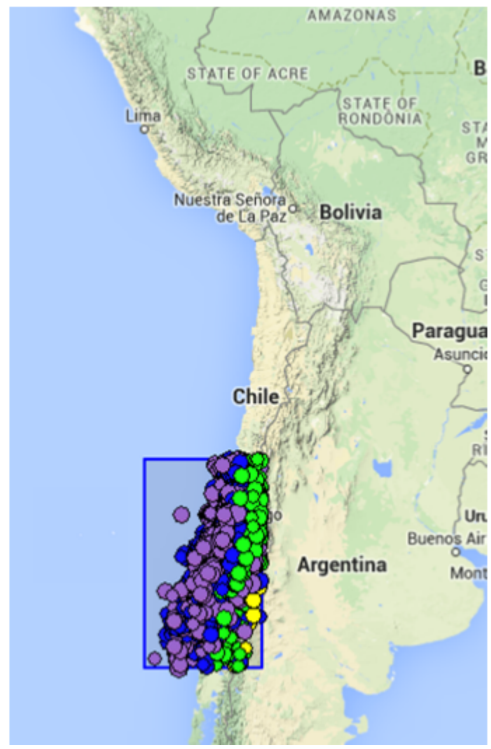

Figure 5.4: Region 4

5.3 Financial time series

We studied emergent market indices corresponding to eight countries: Brazil (BOVESPA), from 04-27-1993 to 10-22-2001; Argentina (MERVAL), from 10-8-1996 to 10-22-2001 and Hong Kong (HSI), from 01-2-1991 to 10-25-2001, USA, Thailand, Turkey, Mexico, Argentina. The number of data points for BOVESPA, MERVAL and 2100,1250 and 2675 respectively. We also analyzed the Standard and Poor's 500 (S & P 500), a major index of the New York Stock Exchange. In the latter case, the data corresponds to a period from 01-3-1950 to 06-14-2005 with 13,951 data points. The daily close values were used in our analyses.

5.3.1 Real Data Analysis of Chile Data

Table 5.1 and Table 5.2 summarize the results of the estimation of parameters for the Inverse Gaussian Ornstein - Uhlenbeck model and $\Gamma(a, b)$ Ornstein - Uhlenbeck model. We obtained the following parameters: intensity parameter(λ), shape parameter(a), scale parameter(b), mean(μ) and variance (σ^2) of the data sets for each of the four regions.

Table 5.1: Parameter Descriptions for Chile Data

Regions	**Number of Observations**	λ	**a**	**b**	μ	σ^2
Region 1	389	2.1766	27.0669	5.8736	4.6082	0.2671
Region 2	1167	3.0955	29.8684	6.5759	4.5421	0.2101
Region 3	892	1.6348	26.3897	5.7509	4.5888	0.2775
Region 4	5000	1.6344	30.3151	6.7865	4.4670	0.1940

Table 5.2: Numerical results for the $IG(\lambda, a, b)$ OU model

Regions	$\lambda = -\log(\rho(1))$	$X_0 = \bar{X}_{n\Delta t}$	Root Mean Square Error
Region 1	2.1766	4.6000	0.1041
Region 2	3.0955	4.5000	0.0952
Region 3	1.6348	4.6000	0.1077
Region 4	1.6344	4.5000	0.0935

5.3.2 Real Data Analysis of Financial indices

Table 4.5 and Table 4.6 summarize the results of the estimation of parameters for the Inverse Gaussian Ornstein - Uhlenbeck model. We obtained the following parameters: intensity parameter(λ), shape parameter(a), scale parameter(b), mean(μ) and variance (σ^2) of the data sets for the various countries.

5.3.3 Discussion of Numerical Results

In this project, we applied our model to time series arising in geophysics. This section describes the source of our data sets and also present the numerical simulation results when our model is applied to the data sets.

- The data set on earthquakes was obtained from the four regions in Chile from the year 2000 to 2014

- The Inverse-Gaussian Ornstein Uhlenbeck model performed better than the Gamma Ornstein Uhlenbeck model due to the compared results of the simulation.

Chapter 6

Concluding Remarks

6.1 Introduction

In this chapter, conclusions are drawn from the study and various recommendations are made.

6.2 Conclusion

We implemented flexible classes of processes that incorporate long - range dependence, i.e they have a slowly polynomially decaying autocovariance function and self - similarity like properties and that are capable of describing some of the key distributional features of typical geophysical and financial time series.

We constructed an independent $IG(a, b)$ Ornstein-Uhlenbeck processes and simulated data from our proposed model: Inverse Gaussian Ornstein-Uhlenbeck processes to estimate the magnitude of the earthquake for the four regions in Chile. We compared our numerical results to the $\Gamma(a, b)$ Ornstein-Uhlenbeck process.

Based on the numerical results we obtained; the root mean square error, the $IG(a, b)$ Ornstein-Uhlenbeck process performed better than the $\Gamma(a, b)$ Ornstein-Uhlenbeck process. Likewise, for the time series using on financial indices near a crash for both well developed and emergent markets, we estimated the daily closing values which is good to capture the dynamics of the stock market.

In the previous works [14, 15], the authors concluded that the generalized Lévy models were very suitable to describe critical events including financial crashes and earthquakes.

The solution to our stochastic differential equation is a Lévy model, so the very model fitting obtained reinforces the previous conclusions.

6.3 Recommendation

Based on the study, the following recommendations are made;

1. The Inverse - Gaussian Ornstein Uhlenbeck model can be a basis for extending the work to include spatial and structural aspects which could involve the use of stochastic differential equations.

2. An optimal control analyses of the model can also be formulated to optimize the costs and effectiveness of the control measures.

3. We conclude that the generalized Lévy models are very suitable to describe critical events including financial crashes and earthquakes.

6.4 Significance of the Result

Inverse Gaussian Ornstein-Uhlenbeck processes provide a class of continuous tine processes which exhibit a long memory behavior. The presence of long memory suggests that current information is highly correlated with past information at different levels, what may facilitate prediction. The methodology used in this research can be applied to other disciplines such as biology, bioinformatics, medicine and social sciences.

6.5 Future Work

1. Further work can be done to determine the the time an awareness should be raised for a high magnitude earthquake in the regions and also study the superposition of Inverse-Gaussian Ornstein-Uhlenbeck processes.

2. We have a plan to work on the return values of the emergent and developed stock market.
3. Future work can be done to describe cases in financial crashes and explosion.

References

[1] M.C. Mariani, I. Florescu, I. Sengupta, M. P. Beccar - Varela, P. Bezdek, L. Serpa, *Levy models and scale invariance properties applied to Geophysics*, Physica A: **392** (2013), pp. 824–839.

[2] Ole E. Barndorff - Nielsen, Neil Shepherd, *Non - Gaussian Ornstein - Uhlenbeck based models and some of their uses in financial economics*, J.R. Stat. Soc. Ser. B. 3 pp. 167–241, (2001).

[3] M.C. Mariani, O.K. Tweneboah, H.Gonzalez-Huizar and L.F.Serpa *Stochastic Differential Equation of Earthquake Series*, Pure and Applied Geophysics, volume 173 pp. 2359–2364, (2016).

[4] M.C. Mariani, M.A.M. Bhuiyan, O.K. Tweneboah, H. Gonzalez-Huizar, I. Florescu, Volatility Model Applied to Geophysics and High Frequency Financial Market Data, Physica A: Statistical Mechanics and its Applications, 503 (2018), 304-321.

[5] M.C. Mariani, P. Bezdek, L. Serpa, I. Florescu, *Ising type models applied to Geophysics and high frequency market data*, Physica A: 390 (23–24) (2011), pp. 4396–4402.

[6] M.C. Mariani, M.A.M. Bhuiyan, O.K. Tweneboah, Estimating Stochastic Volatility by using Ornestein-Uhlenbeck Type Models, Physica A: Statistical Mechanics and its Applications, 491 (2018), 167-176.

[7] M. P. Beccar-Varela, H. Gonzalez-Huizar, M. C. Mariani, L. F. Serpa and O. K. Tweneboah, “ Levy Flight and Long-range correlation Analysis of Earthquake magnitudes in Chile ”, Pure and Applied Geophysics, Volume 173, Issue 7, pp 2257-2266, 2016.

[8] A. Kuzmina, Levy processes simulation. *Proceedings of the Workshop* "International Congress on Computer Science on Computer Science: Information Systems and Technologies (CSIST 2011)," Minsk, Belarus 31 October - 3 November 2011(1.), pp. 86–91.

[9] S. Bochner, *Diffusion equation and stochastic processes*, Proceedings of the National Academy of Sciences, U.S.A, 35(7), Pages 368–370 (1949).

[10] R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer Academic Publishers, Norwell, MA, 1996.

[11] G.E. Uhlenbeck, L.S. Ornstein, On the theory of the Brownian motion, Physical Review, 36, 823-842 (1930).

[12] P. Carr, H. Geman, D. H. Madan and M. Yor. Stochastic volatility for Levy processes, *Mathematical Finance* 13 (2003), p. 345–382.

[13] W. Schoutens, *Levy processes in finance.* John Wiley & Sons Ltd. Bognor Regis, 2003. p. 196 .

[14] MIT Open Course Ware *15.070J / 6.265 J Advanced Stochastic Processes* "http://ocw.mit.edu " Fall 2013, Lecture 16 , pp. 2–3.

[15] B. ϕksendal, *Stochastic differential equation* Springer, 1991, p. 196 .

[16] Schoutens, W.(2003), Lévy Process in Finance: Pricing Financial Derivatives, *Wiley Series in Probability and Statistics.*

[17] Schoutens, W., and Cariboni J.(2009), Lévy Processes in Credit Risk, *John Wiley & Sons.*

[18] Spiliopoulos K.,(2009), Method of Moments Estimation of Ornstein-Uhlenbeck Processes Driven by General Lévy Process, *ANNALES de l'l.S.U.P.,* **53**, Pages 3-19.

[19] Bertoin, J.(1996), Lévy Processes and Stochastic Calculus, *Cambridge University Press*, Cambridge, UK.

[20] Cont, S. and Tankov, P.(2004), Financial modeling with jump processes, *Chapman and Hall/ CRC*.

[21] Bar-Lev, S., Bshouty, D. and Letac, G.(1992), Natural exponential families and self-decomposability *Statistics and Probability Letters*,**13**,147–152.

[22] Jurek, Z. J. and Vervaat, W.(1983), An integral representation for self-decomposable Banach space valued random variables *Zeitschrift fur Wahrscheinlichkeisheorie and verwandte Gebiete*,**62**,247–262.

[23] Mariani, M.C., Liu Y., (2006), Normalized truncated Lévy walks applied to the study of financial indices *Physica A*, **377**, pages 590–598.

[24] Barany, E., Beccar Varela M.P., Florescu, I., Sengupta I. (2011), Detecting Market crashes by analyzing long memory effects using high frequency data*Quantitative Finance*, **12(4)**, pages 623–634.

[25] Mikosch, T.(1998), Elementary stochastic calculus with Finance in view, *World Scientific Publishing Co. Pte*, Singapore.

Curriculum Vitae

Emmanuel Kofi Kusi was born in Ghana, the son of Mary and Paul Kusi. June 2016, He was awarded with a First Class Honors in Mathematics and employed by the Kwame Nkrumah University of Science and Technology as a Teaching and Research Assistant in the Department of Mathematics with Associate Professor Francis Tabi Oduro for the academic year 2016/17.

In the Fall of 2017, he gained both admission and assistantship offer in the Mathematical Sciences department and currently a Graduate Teaching Assistant at the University of Texas at El Paso. He worked as an intern at the University of Texas of El Paso who worked under the supervision of Dr. Osvaldo Mendez during the Summer 2018. Spring 2019, he also participated in a Biostatistics workshop, ENAR workshop at Philadelphia, Pennsylvania.

After graduation, Emmanuel will pursue his doctoral degree in Mathematical modeling at the Rochester Institute of Technology .

His hobbies include reading, listening music and watching football.

Present address: 803 Upson Drive

El Paso, Texas 79902