VARIOUS PARAMETER ESTIMATION TECHNIQUES FOR STOCHASTIC DIFFERENTIAL EQUATIONS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

SEMİH ERGİŞİ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
STATISTICS

AUGUST 2019

Approval of the thesis:

VARIOUS PARAMETER ESTIMATION TECHNIQUES FOR STOCHASTIC DIFFERENTIAL EQUATIONS

submitted by **SEMİH ERGİŞİ** in partial fulfillment of the requirements for the degree of **Master of Science in Statistics Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Ayşen Dener Akkaya _____
Head of Department, **Statistics**

Assoc. Prof. Dr. Ceren Vardar Acar _____
Supervisor, **Statistics Department, METU**

Examining Committee Members:

Assoc. Prof. Dr. Ceylan Talu Yozgatlıgil
Department of Statistics, METU _____

Assist. Prof. Dr. Ceren Vardar Acar
Department of Statistics, METU _____

Assist. Prof. Dr. Ayşe Mutlu Derya
Department of Mathematics,TED _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: Semih Ergişi

Signature :

ABSTRACT

VARIOUS PARAMETER ESTIMATION TECHNIQUES FOR STOCHASTIC DIFFERENTIAL EQUATIONS

Ergişi, Semih
M.S., Department of Statistics
Supervisor : Assoc. Prof. Dr. Ceren Vardar Acar

August 2019, 71 pages

Dynamic systems appear in many fields from economics to physics, from biology to engineering include randomness. Therefore, stochastic differential equations are one of the necessary mathematical tools to model dynamic systems in these disciplines. In this study, we propose two parameter estimation methods when modelling with SDEs which are driven by Brownian motion. Maximum likelihood estimation and generalized method of moment techniques are used to estimate parameters and it is obtained that when the assumptions for Brownian motion satisfy, both techniques give the same result.

Keywords: Brownian Motion, Simulation, Parameter Estimation, Discretization, Euler-Maruyama

ÖZ

STOKASTİK DİFRENSİYEL DENKLEMLER İÇİN ÇEŞİTLİ PARAMETRE TAHMİN YÖNTEMLERİ

Ergişi, Semih
Yüksek Lisans, İstatistik Bölümü
Tez Yöneticisi : Doç. Dr. Ceren Vardar Acar

Austos 2019, 71 sayfa

Ekonomiden fiziğe, biyolojiden mühendisliğe kadar birçok alanda görünen dinamik sistemler rastgelelik içerir. Bu nedenle, stokastik diferansiyel denklemler, bu disiplinlerdeki dinamik sistemleri modellemek için yararlı matematiksel araçlardan biridir. Bu çalışmada Brownian hareketi ile temsil edilen stokastik diferansiyel denklemlerle modelleme yaparken iki parametre tahmin yöntemi önerilmektedir. Parametreleri tahmin etmek için en çok olabilirlik ve genelleştirilmiş moment teknikleri kullanılmış ve Brownian hareketi için yapılan varsayımlar yerine getirildiğinde, her iki tekniğin de aynı sonucu verdiği elde edilmiştir.

Anahtar Kelimeler: Brown Hareketi, Simulasyon, Parametre Tahmini, Kesikli hale Getirme, Euler-Maruyama

To my dear family and

for their continuous patience

ACKNOWLEDGMENTS

It is a debt to state my appreciation to my advisor Assist. Prof. Dr. Ceren Vardar for her continous assistance, tolerance and encouragement during my thesis research. If she did not encourage, help and motivate me, this thesis could not have appeared. Being her student and working with her make me feel happy. She is more than a supervisor due to her friendly, insightful and tolerant manner.

Finaly, My deepest heart feelings and my love are for my family Nermin, Hanifi, Mevlüt, Güler, Yusuf Deniz and Yasir Ali Ergişi for never-ending support, devotion, love and everything. I could not be able to do anything if they did not encourage me. And also thanks METU statistics department, Buket Coşkun, Berna Tuncer, Mustafa Aydoğan and Muhammet Aydın for their endless assistance, helpful suggestions and encouragements.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ADF	Augmented Dickey-Fuller
Bm	Brownian Motion
DE	Differential Equation
DMLE	Discrete Maximum Likelihood Estimation
E-M	Euler Maruyama
GMM	Generalized Method of Moment
GOF	Goodness of fit
KPSS	KwiatkowskiPhillipsSchmidtShin
MLE	Fractional Brownian Motion
ODE	Ordinary Differential Equation
R-S	Riemann-Stieltjes
SDE	Stochastic Differential Equation

CHAPTER 1

INTRODUCTION

Many of the dynamic systems that emerge in many fields from economics to physics, from biology to engineering include randomness [45]. Therefore, in order to model dynamics systems that emerge in these disciplines, stochastic differential equations (SDE) are one of the necessary mathematical tools. SDEs are useful mainly to predict the future with the help of the movements obtained by the variables belonging to an economic and social phenomenon. A lot of fundamentals, of laws or of reasons under the behaviour of events naturally happening are expressions or relatives including proportions in which things occur. When they are expressed by using mathematical terms, the connections are expressed by equations and the proportions are expressed by derivatives. They are called as differential equations (DE) when the equation relates some function with its derivatives. Thus, having information on analytic about DEs is helpful to understand and investigate problems such as population rise or population reduction among many others. When DEs are applied to any of the numerous fields in which they are convenient, firstly, formulating the suitable DE that defines or models the problem being examined is essential.

Predicting the world around us is noteworthy ability of DEs [44]. DEs are utilised in many fields such as from economics to physics, from biology to engineering [44]. Growing and decaying exponentially, the rise in population or the difference in investment return in the time are described by DEs [38]. They are useful not only for these but also for cancer growth modelling, disease spread, chemcial reactions modelling or best investment strategies in economics [38].

Here is a simple example for modelling a basic daily life experience. Suppose that hundred grams of sugar are dissolved in water. The speed of dissolving is proportional to the part that has not dissolved. Then the differential equation given below explains the speed of dissolving after t minutes where q is the dissolved part,

$$\frac{\partial q}{\partial t} = k(100 - q).$$

and k is constant rate which represents the proportion of speed of dissolving. By the equation, the change is multiplication of k (rate) by the remaining part of sugar after dissolving.

1.1 Ordinary Differential Equations

A differential equation (DE) is an equation, where the unknown quantity is a function, and the equation includes derivatives of the unknown function; i.e., DE is mathematical equation that relates the function and its derivatives. Suppose that t is an independent variable and $k = f(t)$ is an unknown function. The relation between the derivative of this unknown function and independent variable t is called a DE and it is generally written as below,

$$F(t, k, k', k'', ..., k^n) = 0.$$

DEs are classified by basing on the dependence of the unknown function on an independent single variable or on a few of independent variables [10]. If there are solely ordinary derivatives exist in the DE [10], it is called as an ODE. If the function depends on more than one variable and the differential equation involves partial derivatives, then the equation is called as partial differential equation (PDE) [10].

$$k'' + 5kk' = 0 \tag{1.1}$$

$$\frac{\partial g}{\partial q} = 2 + \frac{\partial p}{\partial q}, \tag{1.2}$$

where $g(p, q)$ is a function of two variables. The Equation (1.1) is called as an ODE while the Equation (1.2) is called as a PDE. The order of a DE is detected by its highest derivative. For example, the equation that includes first derivative and the first derivative as the highest derivative is called as a first order DE. If it includes the first and the second derivative, then it is called as second order DE. The Equation (1.3) is an example of third order DE.

$$k''' + 2e^t k'' + kk' = t^5 \tag{1.3}$$

The equation is called as nonlinear differential equation if the equation includes the product of dependent variable and its derivatives or the product of independent and dependent variables. Otherwise, it is called as linear differential equations.

$$k'' + 5k' + 6k = 0. \tag{1.4}$$

$$k'' + 5kk' = 0. \tag{1.5}$$

The first example given in the Equation (1.4) is linear while the second example given in the Equation (1.5) is not linear because there exist the product of dependent variable and its derivative. In order to give an example of ODE, let's consider the example of field mice population given in [5]. When there are no predators, it is assumed that increase of the population occurs proportionally to the existing population. When the time is denoted by t and the mouse population is denoted by $m(t)$, then it is assumed that population growth is stated by the equation

$$\frac{dm}{dt} = km \tag{1.6}$$

where k is called the constant rate or growth rate. To state the logic more clear, assume that time is measured in months and the k is equal to 0.6. Then each term in the Equation (1.6) is written in terms of mice/month unit. Now, it is supposed that few cats live in the same habitat and that 20 field mice are killed daily. By taking a month as 30 days, there will be 600 mice killed per month. When this is utilised in differential Equation (1.6), another term appears. Thus, it becomes as follows,

$$\frac{dm}{dt} = 0.6m - 600. \tag{1.7}$$

It is seen that -600 is used as the predation term rather than -20 due to time measurement in months and thus, the monthly predation rate should be written. The solution of the Equation (1.7) is found as below. First, the Equation (1.7) is rewritten in the form

$$\frac{dm}{dt} = \frac{3m - 3000}{5}, \tag{1.8}$$

or if $m \neq 1000$,

$$\frac{dm}{dt(m - 1000)} = 0.6 \tag{1.9}$$

For $m \neq 1000$, left hand side of the Equation (1.9) is same as taking derivative of $\ln|m - 1000|$ in terms of t. Thus, we have

$$\frac{d}{dt} \ln|m(t) - 1000| = 0.6 \tag{1.10}$$

Then, by taking integral of both sides of the Equation (1.10), we have

$$\ln|m(t) - 1000| = \frac{6t}{10} + C \tag{1.11}$$

where C is an arbitrary integration constant. By taking exponential of both sides of the Equation (1.11),

$$m(t) - 1000 = e^{\frac{6t}{10} + C} = e^C e^{\frac{6t}{10}}, \tag{1.12}$$

or

$$m(t) - 1000 = \pm e^C e^{(\frac{6t}{10})}, \tag{1.13}$$

and finally,

$$m(t) = 1000 + Ae^{\frac{6t}{10}}, \tag{1.14}$$

where $A = \pm e^C$ is also an arbitrary(nonzero) constant. If m is equal to 1000, the solution of the Equation (1.8) is also obtained, i.e., the change in the field mice population with respect to time is zero and that is seen in the Equation (1.14) if A is equal to zero.

In the example of field mice population, we obtain a set of solutions of the differential Equation (1.7), as the constant A in the Equation (1.14) varies. The Figure 1.1 shows different solutions of the equation given for the field mice population as A varies [5]. An arbitrary constant comes from the integration contained in the solution. Due to the fact that this constant takes different values, a lot of solutions come in view. In general, it is requested to get one single solution among all of plausible solutions by pointing out value for the arbitrary constant initially.

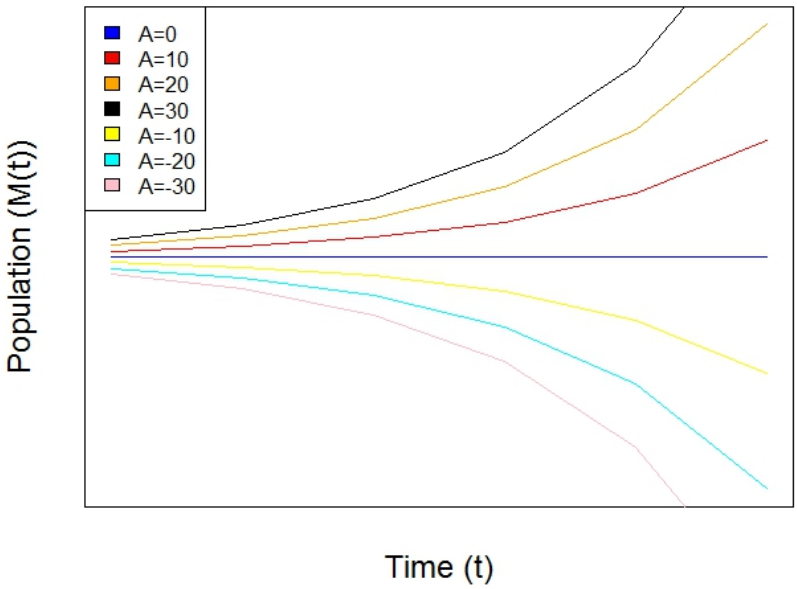

Figure 1.1: Graph of $m(t) = 1000 + Ae^{0.6t}$

Let us show this on the example of field mice population, by taking $m(0) = 1010$. Then, substituting $t = 0$, and $m = 1010$ into the Equation (1.14), we get

$$1010 = 1000 + A. \tag{1.15}$$

Thus $A = 10$, and by putting this value in the Equation (1.14), we get the solution,

$$m(t) = 1000 + 10e^{\frac{6t}{10}}. \tag{1.16}$$

The condition $m(0) = 1010$ is an example of initial condition. The Differential Equation (1.7) together with initial condition is an instance of an initial value problem. Fix a point $m(0) \in \mathbf{R}$ and the DE is stated as below,

$$\frac{dm}{dt} = H(m(t)) \tag{1.17}$$

where $H : \mathbf{R} \to \mathbf{R}$ is a given, smooth vector field and the solution is the path. In the Equation (1.17) $m(\cdot) = [0, \infty) \to \mathbf{R}$ and $m(t)$ denotes the value at time t. However, in reality, due to some environmental effects, the system trajectories does not behave as expected based on our measurements. Thus, the equation should be modified to represent these effects [16]. In order to build a more realistic model, we use stochastic differential equations (SDEs).

In this thesis, our main goal is to give a review of techniques for the estimation of parameters in SDEs of the form in the Equation (2.6). Two main estimation methods are given in this study for estimation of parameters; namely, maximum likelihood estimation (MLE) and generalized method of moment (GMM) estimation techniques.

The thesis is organized as follows: we present background information regarding stochastic process, Brownian motion, SDEs, analytic solutions and numerical solutions of SDEs and convergence of numerical solutions to the analytic solutions in Chapter 2. In Chapter 3, construction of the model, assumption checking for Brownian motion and goodness of fit for model are expressed. After the main model is completed, methods to asses the model prediction performance are presented. In Chapter 4, two techniques for estimation of parameters, MLE and GMM, are presented. In Chapter 5, two examples are shown for to apply the steps mentioned in the previous parts of the study. Chapter 6 is the conclusion of the thesis.

CHAPTER 2

BACKGROUND INFORMATION

It is aimed to give general information about stochastic processes, SDEs, stochastic integrals and solution of SDEs. These are vital to understand the SDEs. When the stochastic process is understood, it is easier to understand the solution of SDEs. In view of this, we start with providing definition of stochastic processes and their properties. Then, we give definition of stochastic integrals which are necessary to understand the modelling of SDEs. For the detailed information about the definitions and proofs of the theorems one can see [1, 2, 4, 21, 28, 36, 38, 43].

In the second part of this chapter, the solution of SDEs is presented. Firstly, we start with analytic solution of SDEs and give one example related to this. Secondly, we introduce Euler Maruyama solution technique which is the simplest and easily applicable technique among many other techniques and its convergence to the analytic solutions. For the detailed information about the solutions of SDEs, proofs of solutions and computational applications one can see [7, 26, 32, 41, 42].

2.1 Stochastic Differential Equation

It is useful to continue with an example of an ODE, a simple field mice population growth model from previous part to perceive the SDE. Let $M(t)$ determines the population at time t, $k(t)$ is the deterministic relative growth rate at time t and $\frac{dM}{dt}$ is the rate of change of the population size and M_0 is the given initial data value. Then,

below ODE gives the change in the population at t,

$$\frac{dM}{dt} = k(t)M(t), \quad M(0) = M_0. \tag{2.1}$$

The Equation (2.1) shows that the rate of change of the field mice population at time t is equal to multiplication of the growth rate and the population at that time. However, usually $k(t)$ is not known completely, and it is exposed to some environmental impacts [34]. Consequently, it is written as,

$$k(t) = b(t) + \beta noise, \tag{2.2}$$

where $b(t)$ is deterministic term, β is a real valued constant number, and $noise$ corresponds to the random term whose behaviour is not definitely known, solely its probability distribution is known. The $noise$ term in the Equation (2.2) is generally taken as a white noise which is officially taken into consideration as the derivative of a Brownian motion (Bm), $\dot{W}(t)$, denoted by $\dot{W}(t) = \frac{dW(t)}{dt}$. Then, the Equation (2.1) is,

$$\frac{dM}{dt} = (b(t) + \beta noise)M(t), \tag{2.3}$$

or

$$\frac{dM}{dt} = (b(t) + \beta \dot{W}(t))M(t); \tag{2.4}$$

by rearranging the Equation (2.4), the following equation is obtained.

$$dM(t) = b(t)M(t)dt + \beta M(t)dW(t). \tag{2.5}$$

We call DE as SDE due to this $noise$ term. In general, SDEs are obtained by adding randomness to the coefficients in the DEs. In this case, randomness comes from the noise. Moreover, there is another source of noise due to measurement errors. For a detail information about this type of source, one can check [38]. In this study, our central interest is the SDEs given as in the Equation (2.5).

Our main purpose is to estimate parameters in the stochastic differential equation given in the Equation (2.6) and to make forecast by using the SDEs in this thesis. A DE is a DE whose several terms are stochastic processes and so the solution also results in stochastic process [42]. In general SDE is given by the following equation,

$$dX(t) = f(X(t), t; \theta)dt + g(X(t), t; \theta)dW(t). \tag{2.6}$$

The noise term is represented by $W(t)$ and it is Brownian motion in the Equation (2.6). The representation of noise term changes according to the type of stochastic process added to the equation. This means that the noise is added to the equation by other stochastic process such as Fractional Brownian motion, Levy processes etc. In this thesis, studies are based on the noise represented by Brownian motion. Before giving information about Brownian motion, the Equation (2.6) is written in the following form;

$$X(t) = X(0) + \int_0^t f(X(s), s; \theta)ds + \int_0^t g(X(s), s; \theta)dW(s). \tag{2.7}$$

Actually, the Equation (2.7) is not different from the Equation (2.6). The former is the integral form of the latter. To achieve the aim of the thesis, it is important to solve the integral equation given in the Equation (2.7). However, it is not easy to solve this integral due to the stochastic processes and their properties. In the Equation (2.7), f, g and $W(t)$ should be clearly analyzed before solving the integral. Also, reasonable mathematical interpretation of $W(t)$ should be found. Firstly, we begin with giving the definition of stochastic process and their properties.

2.2 Stochastic Process

Definition 1 *[18] For a given probability space $(\Omega, \mathcal{A}, \mathcal{P})$, and measurable space (V, Σ), a stochastic process is a collection of V-valued random variables, which are written as;*

$$\{Y(t) : t \in \tau\}.$$

A stochastic process is also written as $\{Y(t, \omega) : t \in \tau\}$. This means that a stochastic process is a function of two variables, $t \in \tau$ and $\omega \in \Omega$. If τ is an interval, then the stochastic process is continuous time stochastic process. If τ takes discrete values, then the stochastic process is called discrete time stochastic process. Consider that for fixed $t \in \tau$ we have random variable $\omega \to Y_t(\omega)$ where $\omega \in \Omega$ and for fixed $\omega \in \Omega$ the function which is a path of Y_t where $t \in \tau$ [38]. In our study, we are interested in the function that is a path of Y_t for fixed $\omega \in \Omega$.

2.2.1 Brownian Motion

In 1827, a motion of the pollen grains of some plants in the fluid was unearthed by Robert Brown, the Scottish Botanist. While investigating this movement, the place of the pollen particles changed in a random motion was considered. He described this phenomenon as Brownian motion (Bm) in his paper [6]. On the contrary, the correct explanation lied under the reason of pollen particles' motion could not be given until Einstein discovered the kinetic theory [14]. He realized that this motion emerged from collusion of the molecules in the liquid. After Brown and Einstein constructed physically the Bm, some mathematicians like Weiner, Donsker, Kolmogorov, Levy gave the mathematical foundations of it. In recent years, this stochastic process is extensively used in many disciplines such as mathematical statistics, physics and management science. Especially, nowadays, it is applied mainly for modelling the stock price and financial markets.

Definition 2 *[37] A stochastic process $\{W_t\}$ is called a standard Bm or Wiener process that satisfies the following properties.*

i. It has continuous sample paths.

ii. It has stationary and independent increments. For all $0 \leq t_1 \leq t_2.... \leq t_{i-1}$ and $i \geq 0$, $W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, ..., W_{t_i} - W_{t_{i-1}}$ are independent random variables, that is, $W(t)$ has independent increments. For any $0 \leq s \leq t$ and any $k > 0$ $W_{t+k} - W_{s+k} \sim W_t - W_s$, it has stationary increments.

iii. It has normal distribution, $W_t \sim \mathcal{N}(0, t)$ for all t.

The linear transformation of independent Gaussian $(W_{t_1} W_{t_2}.....W_{t_p})$ has joint normal distribution for all $0 \leq t_1 < t_2..... \leq t_p$ because Bm increments are distributed normally. Note that, if all vectors $(W_{t_1} W_{t_2}.....W_{t_p})$ are Gaussian, then the W_t is also Gaussian [37] having zero mean and covariance function (C) $C(W_t, W_s) = \sigma^2 min(s, t)$.

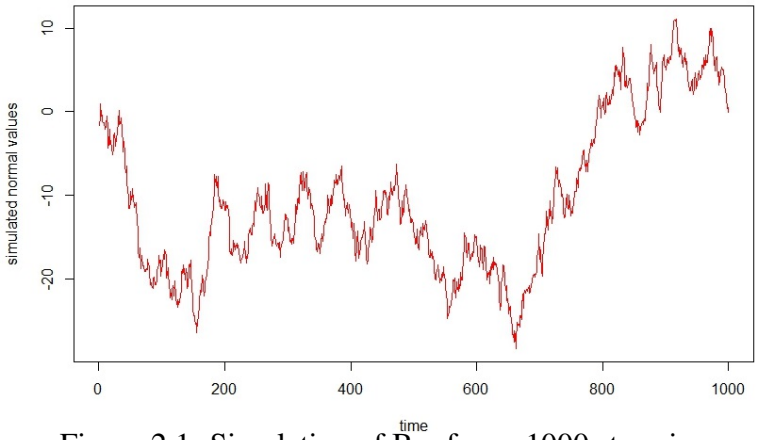

Figure 2.1: Simulation of Bm for n=1000 step size

For $s \leq t$ and by using the second and third properties of Bm, covariance is obtained in the following way;

$$C(W_t, W_s) = E(W_t W_s) - E(W_s)E(W_t) = E(W_t W_s) \tag{2.8}$$

Now, let us substitute $(W_t - W_s) + W_s$ for W_t,

$$\begin{aligned} C(W_t, W_s)) &= E(((W_t - W_s) + W_s)W_s) \\ &= E((W_t - W_s)W_s) + E({W_s}^2) \end{aligned} \tag{2.9}$$

Because the $W_t - W_s$ only depends on $t - s$, these increments are independent and thus zero

$$C(W_t, W_s) = E({W_s}^2) = Var(W_s) = \sigma^2 s. \tag{2.10}$$

In a similar manner, for $t \leq s$ covariance equals to $\sigma^2 t$. Thus, $C(W_t, W_s) = \sigma^2 min(s, t)$.

Moreover, Bm is a Markov process and so it shows the Markov property. In general, it briefly gives an idea that the future only depends on present value. Before mentioning more sophisticated properties of Bm, it is necessary to give the definition of filtration and martingale.

Definition 3 *[37] The collection $(\mathcal{F}_t, t \geq 0)$ of σ fields on Ω is called as filtration if $(\mathcal{F}_s \leq \mathcal{F}_t)$ for all $0 \leq s \leq t$.*

Let W_t on probability space with filtration $\mathcal{F}_t$, collection of σ algebra $(\Omega, \mathcal{F}, \mathcal{P})$, is called Markov process if for any time $s \geq t$.

$$\begin{aligned} P(W_{t+s}|\mathcal{F}_t) &= P(W_{t+s}|W_0, W_1, W_2, ..., W_t) \\ &= P(W_{t+s}|W_t) \end{aligned} \tag{2.11}$$

If W_t is a Bm, the process $W_{t+s} - W_s$ is independent of the process W_s. This comes from the independency of increments.

Definition 4 *[37] The stochastic process $X(t, t \geq 0)$ is called martingale with respect to the filtration $(\mathcal{F}_t, t \geq 0)$ if*

i. $E(X_t) \leq \infty$ for all $t \geq 0$

ii. X is adapted to the $\mathcal{F}_t$.

iii. For $\forall s \leq t$ $E[X_t|\mathcal{F}_s] = X_s$.

Bm is a Martingale that it has constant expectation function. For fix $0 \leq s \leq t$, by Definition 4, W_t with filtration $\mathcal{F}_t$ is a martingale if W_t is integrable and the statement $E(W_t|\mathcal{F}_s) = W_s$, is almost surely true. As known from the definition of Bm, $E(W_t) = 0$, it is clear that $E(W_t)$ is integrable because of $E(W_t) \leq \infty$, and then

$$E[W_t|\mathcal{F}_s] = E[(W_t - W_s) + W_s|\mathcal{F}_s] = E[W_t - W_s|\mathcal{F}_s] + E[W_s|\mathcal{F}_s] \tag{2.12}$$

Because of independence of $W_t - W_s$ from $\mathcal{F}_t$, $E(W_t - W_s|\mathcal{F}_s) = E(W_t - W_s) = 0$

$$E[W_t|\mathcal{F}_s] = W_s + E[W_t - W_s] = W_s \tag{2.13}$$

Therefore, W_t is a martingale.

Sample paths of the Bm is non-differentiable and it has unbounded variations. If a function is differentiable, it has a smooth graph. But the path of the Bm is not smooth. In fact,

$$\lim_{\Delta t \to 0} \sup\left(\frac{\mathrm{W}_{(\mathrm{t}_0+\Delta \mathrm{t})} - \mathrm{W}_{\mathrm{t}_0}}{\Delta \mathrm{t}}\right) = \infty. \tag{2.14}$$

This means that sample paths are non-differentiable with probability 1 [37]. It is said that the graph of the Bm does not change its shape in the neighborhood of any point in a predictable way. The Equation (2.14) is generally inferred based on the theorem given below.

Theorem 1 *[37] Brownian motion sample paths have unbounded variations on a finite interval* $[0, T]$*. Its mathematical expression is;*

$$\sup_{\tau} \sum_{j=1}^{n} |B_{t_j}(w) - B_{t_{j-1}}(w)| = \infty \tag{2.15}$$

where supremum is taken by the consideration of all over the possible partitions τ : $0 = t_0 < t_1 ... < t_n = T$ *for* $[0, T]$.

As a result of unbounded variations and non-differentiability of Brownian sample paths which classical integration methods fail to solve the integral equation, it is necessary to define stochastic integral and their properties to solve the integral equation.

2.3 Stochastic Integral

The Itô Stochastic integral concept is introduced due to non-differentiability and boundless variation of Brownian sample paths. Before defining the Itô integral, let us start with Riemann integral which is helpful to form the basis of Itô integral.

Definition 5 *Assume* f *is arbitrary function on* $[c, d]$ *and a partition of the interval* $[c, d]$ *is considered as* $\tau_h : c = t_0 < t_1 < ... < t_h = d$*. Let* $\Delta_j = t_j - t_{j-1}$ *for* $j = 1, 2, ..., h$*. The Riemann sum is defined as;*

$$K_h = \sum_{j=1}^{h} f(z_j)(t_j - t_{j-1}) = \sum_{j=1}^{h} f(z_j)\Delta_j.$$

From the Definition 5, it is said that Riemann sum is a weighted average of $f(z_j)$, where lengths Δ_j of the intervals $[t_{j-1}, t_j]$ correspond to the weights. In fact, the given mathematical definition of the Riemann sum is clearly understandable from the

Figure 2.2. In the Figure 2.2, the area is calculated approximately by summing the rectangular areas. That is, $\int_a^b f(x)dx$.

$$K = \lim_{h\to\infty} K_h = \lim_{h\to\infty} \sum_{j=1}^{h} f(z_j)\Delta_j$$

exists when $mesh(\tau_h) \to 0$ and K is not dependent to the selection of partitions τ_h. Thus, K is Riemann integral for f on given interval $[a, b]$.

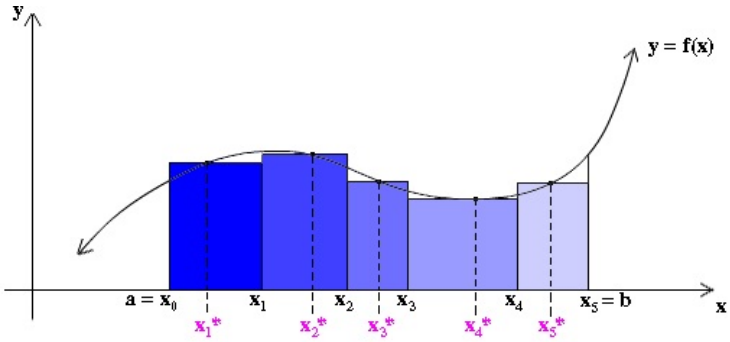

Figure 2.2: Calculation of the area under $y = f(x)$

Riemann integral is the basis of the integral that is necessary to achieve the aim of this study. However, $\int_a^b f(t)dB_t(w)$ is not calculated by Riemann integral when f is a stochastic process or function on $[a, b]$ and $B_t(w)$ is Brownian sample path. Path-wise integral which is the permission of evaluating the integral like $\int_a^b f(t)dB_t(w)$ is utilized. Therefore, we utilize the Riemann-Stieltjes (R-S) integral.

Definition 6 *A partition of the interval $[c, d]$ is considered $\tau_h : c = t_0 < t_1 < .... < t_h = d$. let f and z be two real-valued function on $[c, d]$ and $\Delta_j z = z(t_j) - z(t_{j-1}), j = 1, 2, ..., h$. The Riemann-Stieltjes sum is;*

$$K_h = \sum_{j=1}^{h} f(y_j)\Delta_j z = \sum_{j=1}^{n} f(y_j)[z(t_j) - z(t_{j-1})].$$

It differs from the Riemann integral in terms of weighting the values $f(y_i)$ with increments $\Delta_j z$ on the $[t_{j-1}, t_j]$.

$$K = \lim_{h\to\infty} K_h = \lim_{h\to\infty} \sum_{j=1}^{h} f(y_j)\Delta_j z$$

exists when mesh$(\tau_h) \to 0$ *and K is independent from the partitions of* τ_h*, then K is defined as the Riemann-Stieltjes integral of* f *with respect to* z *on the given interval* $[c, d]$.

If we consider the integral that we try to solve, some questions may rise from the definition of R-S integral. One of the questions is that when the $\int_a^b f(t)dz(t)$ exists and is it possible to replace $dz(t)$ with $dB(t)$ on the interval $[a, b]$. For these questions, the answer is $\int_a^b f(t)dz(t)$ exists if for the functions f and z there have no discontinuities at the same points on $[a, b]$ and f is continuous and z has bounded variation. Due to the unbounded variations of $B(t)$, it is not possible to say that R-S integral for $\int_a^b f(t)dB(t)$ exists. The following theorem provides the conditions where the integral $\int_a^b f(t)dB(t)$ exists in terms of R-S sense.

Theorem 2 *[37] Assume* f *is a deterministic function or the sample path of a stochastic process. If* f *is differentiable and has bounded derivative on* $[c, d]$ *then the Riemann-Stieltjes integral* $\int_c^d f(t)dB(t)$ *exists for each Brownian sample path.*

Actually, the theorem given above is helpful to show the presence of following integrals

$$\int_a^b e^t dB(t), \int_a^b t^k dB(t), k \geq 0$$

because, derivatives are bounded on the interval $[a, b]$. Although, the answer of the existence of the integral is obtained in R-S, it does not mean that we obtain the solution of the integral in terms of Bm. The integral in terms of Bm is solved by Itô integral. Moreover, the theorem given above is not enough for the existence of the integrals in the Equation (2.16). Itô integral is defined in an intuitive manner.

$$\int_a^b B(t)dB(t) \tag{2.16}$$

Let $K_n = \sum_{i=1}^n B_{t_{i-1}} \Delta_i B$ be a R-S sum, $\tau_n : 0 = t_0 < t_1 < ... < t_n = t$ is any partition of the interval $[0, t]$, $\Delta_i B = B_{t_i} - B_{t_{i-1}}$ is a corresponding Brownian motion increments and $\Delta_i = t_i - t_{i-1}$. Let write the K_n in the following way;

$$K_n = \frac{1}{2}B_t^2 - \frac{1}{2}\sum_{i=1}^n (\Delta_i B)^2 = \frac{1}{2}B_t^2 - \frac{1}{2}U_n(t).$$

It is known that the Brownian motion has such important properties as independent and stationary increments. Therefore, using these properties

$$\mathrm{E}(\Delta_{\mathrm{i}}\mathrm{B}\Delta_{\mathrm{j}}\mathrm{B}) = \begin{cases} 0 & \text{if i=j,} \\ \mathrm{Var}(\Delta_{\mathrm{i}}\mathrm{B}) = t_i - t_{i-1} = \Delta_i & \text{if i}\neq j \end{cases} \tag{2.17}$$

is obtained and

$$\mathrm{E}(\mathrm{U_n(t)}) = \sum_{i=1}^{n} \mathrm{E}(\Delta_{\mathrm{i}}\mathrm{B})^2 =_{i=1}^{n} \Delta_i = t$$

Also, by using the variance formula,

$$\mathrm{Var}(\mathrm{U_n(t)}) = \sum_{i=1}^{n} \mathrm{Var}[(\Delta_{\mathrm{i}}\mathrm{B})^2] = \sum_{i=1}^{n} \mathrm{E}(\Delta_{\mathrm{i}})^4 - \Delta_{\mathrm{i}}^2.$$

If $Z \sim \mathcal{N}(0,\,1)$, then the result of the $\mathrm{E}(\Delta_{\mathrm{i}})^4$ is found by using the moments of standard normal distribution.

$$\mathrm{Var}(\mathrm{U_n(t)}) = \sum_{i=1}^{n}(3\Delta_i^2 - \Delta_i^2) = 2\sum_{i=1}^{n}\Delta_i^2.$$

Therefore, when $mesh(\tau_n) = \max_{i=1,2,\ldots,n}(\Delta_i) \to 0$

$$\mathrm{Var}(\mathrm{U_n(t)}) \leq 2\ \mathrm{mesh}(\tau_n)\sum_{i=1}^{n}\Delta_i = 2t\ \mathrm{mesh}(\tau_n) \to 0$$

is obtained. It is said that $U_n(t)$ converges to t in mean square sense because of $\mathrm{Var}(\mathrm{U_n(t)}) = \mathrm{E}(\mathrm{U_n(t)} - \mathrm{t})^2 \to 0$. As a result of these, $K_n = 0.5[B^2(t) - U_n(t)]$ converges to the $0.5[B^2(t) - t]$. Actually, the answer of the integral is $\int_0^t B(s)dB(s) = 0.5[B^2(t) - t]$. This result gives us information on how to obtain the Bm integral which is not obtained in the sense of R-S integral. After the intuitive definition, we now present the properties.

1. The stochastic process $I_t(X) = \int_0^t X_s dB_s, t \in [0,T]$, is a martingale with respect to Brownian filtration $(\mathcal{F}_t, t \in [0,T])$.

2. Its expected value is zero.

3. It has isometry property $E(\int_0^t X_s dB_s)^2 = \int_0^t EX_s^2 ds, t \in [0,T]$.

4. For constants a_1, a_2 and X^1, X^2 are processes on $[0,T]$, it has linearity $\int_0^T [a_1 X_s^1 + a_2 X_s^2] dB_s = \int_0^T a_1 X_s^1 dB_s + \int_0^T a_2 X_s^2 dB_s$

After obtaining $\int_0^t B_s dB_s = 0.5[B_t^2 - t]$ in mean square sense, the important step is obtaining the solution of this type of integral in terms of Brownian motion. However, Itô Lemma, the chain rule for Itô stochastic integral is proposed by the integral $\int_0^t B_s dB_s = 0.5[B_t^2 - t]$ due to the the effect of the second order term. This is not negligible due to the non-differentiability of the Brownian sample paths.

2.4 Solution of SDEs

In this section, the aim is to obtain the solution of SDEs specifically the Equation (2.6) and make predictions about the future. To solve the SDE in the Equation (2.6), it is necessary to solve the integral Equation (2.7). If the stochastic process X satisfies the Equation (2.6) where $W(t)$ is Bm, the solution of given SDE is X. There exists a filtration to which $W(t)$ is adapted. Considering the Equation (2.7) it is clear that adaptation to the same filtration for X exists. There are two type of solutions for an SDE. One of them is strong solution and the other one is weak solution. Before solving SDE, it is important to give stochastic chain rule.

Theorem 3 *[1] Consider an SDE which has the form in the Equation* (2.6) *and a function $F(t,x)$ has continuous derivatives. Itô Lemma is defined as*

$$dF(t,x) = (\frac{dF(t,x)}{dt} + f(t,x)\frac{dF(t,x)}{dX} + \frac{1}{2}g^2(t,x)d^2\frac{F(t,x)}{dx^2})dt \\ +g(t,x)\frac{dF(t,x)}{dx}dW(t).$$

By the above equation not all but some SDEs are solved analytically. Black-Scholes (B-S) is an SDE that is solved analytically and is very well known in finance.

$$dX(t) = aX(t)dt + sX(t)dW(t) \tag{2.18}$$

where $X(0) = x_0$, $f(t, X(t)) = aX(t)$ and $g(t, X(t)) = sX(t)$ are coefficient functions. Each side of the Equation (2.18) is divided by $X(t)$ term, left-hand side of the equation obtained is similar to the derivation of $\ln X(t)$. Therefore, Itô Lemma is

applied by letting $K(t,x) = \ln(X(t))$.

$$d\ln(X(t)) = aX(t)\frac{1}{X}dt + \frac{1}{2}s^2X^2(t)\frac{-1}{X^2}dt + sX(t)dW(t)$$

similarly,

$$\int_0^t d\ln(X(t)) = \int_0^t (a - 0.5s^2)dt + \int_0^t sdW(u)$$

when we solve the integral, following equation is obtained.

$$X(t) = X_0 \exp(a - \frac{s^2}{2})t + sW(t) \tag{2.19}$$

The Equation (2.19) is solution of the Equation (2.18). In the Equation (2.19) analytical solution or closed form solutions are obtained. However, many of SDEs' solutions is not obtained in closed form [42]. Therefore, utilising numerical approximation techniques to get the solution of an SDE is needed. In this study, Euler-Maruyama, one of the well known approximation techniques is discussed.

2.4.1 Euler-Maruyama Numerical Solution

The Equation (2.20) includes Riemann and Itô integral. In this part of the thesis, we are interested in solving the integral in the Equation (2.20). It is known that many of SDEs do not have closed-form solution [42], so numerical approximation techniques are applied to get the solution.

$$X(t) = X(0) + \int_0^t f(X(s), s; \theta)ds + \int_0^t g(X(s), s; \theta)dW(s). \tag{2.20}$$

In the Equation (2.20) f and g are numeric functions, $X(0)$ is an initial value and second integral is taken with respect to Bm. We do not obtain direct solution like in the ordinary differential equation. We introduce a numerical method in order to solve this SDE and the solution $X(t)$ that we obtain is the random variable that comes in view when we take zero step size limit in the numerical method. The Equation (2.20) is usually rewritten in the DE form as below;

$$dX(t) = f(X(t), t; \theta)dt + g(X(t), t; \theta)dW(t). \tag{2.21}$$

To show solution of the Equation (2.20) numerically, the Equation (2.21) is utilised. Consider that $\frac{\partial W(t)}{\partial t}$ is not permitted due to non differentiability of Bm. In order to apply numerical method for the Equation (2.21) over $[0, T]$, this interval is discretized firstly and it is as $0 < \tau_0 < \tau_1 < .... < \tau_j < \tau_{j+1} = T$ where $0 \leq j \leq N$. Let $\Delta t = T/N$ be the step size and it is defined as $\Delta t = \tau_{j+1} - \tau_j$ for $N > 0$, and $\tau_j = j\Delta t$. X_j denotes the numeric approximation to X_{τ_j}. Then, the E-M method takes the form;

$$X_j = X_{\tau_{j-1}} + f(\tau_{j-1}, X_{\tau_{j-1}})\Delta t + g(\tau_{j-1}, X_{\tau_{j-1}})(W(\tau_j) - W(\tau_{j-1})) \qquad j = 1, 2, ..., N. \quad (2.22)$$

To get the sense where the Equation (2.22) comes from, we will set $t = \tau_{j+a}$, $a = 0, 1$ in the Equation (2.20).

$$X(t) = X_0 + \int_0^t f(X(s), s)ds + \int_0^t g(X(s), s)dW(s)$$

then, for a=0,

$$X(\tau_j) = X_0 + \int_0^{\tau_j} f(X(s), s)ds + \int_0^{\tau_j} g(X(s), s)dW(s)$$

and for a=1,

$$X(\tau_{j+1}) = X_0 + \int_0^{\tau_{j+1}} f(X(s), s)ds + \int_0^{\tau_{j+1}} g(X(s), s)dW(s)$$

thus,

$$X(\tau_{j+1}) - X(\tau_j) = \int_{\tau_j}^{\tau_{j+1}} f(X(s), s)ds + \int_{\tau_j}^{\tau_{j+1}} g(X(s), s)dW(s)$$

that is,

$$X(\tau_{j+1}) = X(\tau_j) + \int_{\tau_j}^{\tau_{j+1}} f(X(s), s)ds + \int_{\tau_j}^{\tau_{j+1}} g(X(s), s)dW(s)$$

Both integrals are approximated as follows:

$$\begin{aligned} \int_{\tau_j}^{\tau_{j+1}} f(X(s), s)ds \approx \int_{\tau_j}^{\tau_{j+1}} f(X(\tau_j), \tau_j)ds &= f(X(\tau_j), \tau_j) \int_{\tau_j}^{\tau_{j+1}} ds \\ \int_{\tau_j}^{\tau_{j+1}} g(X(s), s)dW(s) &\approx g(X(\tau_j), \tau_j)(W(\tau_{j+1}) - W(\tau_j)) \end{aligned}$$

Hence, the integral form of the Equation (2.22) is

$$X(\tau_j) = X(\tau_{j-1}) + \int_{\tau_{j-1}}^{\tau_j} f(X(s), s)ds + \int_{\tau_{j-1}}^{\tau_j} g(X(s), s)dW(s) \tag{2.23}$$

The Equation (2.22) gives approximation of corresponding terms in the Equation (2.23). Increments $\Delta W(\tau_j) = W(\tau_j) - W(\tau_{j-1})$ which are determined by normal random number generator are utilised to model Bm [42]. The computation of random increments is $\Delta W(\tau_j) = m_j\sqrt{\Delta\tau_j}$ where $m_j \sim \mathcal{N}(0, 1)$.

Now, we evaluate the accuracy of numerical solution. This is done by using two different criterias, namely strong and weak convergence. These two criteria are based on the mean difference of numerical solution and exact solution. Strong convergence is achieved by using the mean of error whereas weak convergence is achieved by using error of means of numeric and analytic solution. By the way, the difference between exact and numeric solution is called as error. At time T, we obtain a random value from each approximate solution path because it is random variable. Error is a random variable due to this. A discrete time approximation strongly converges to the solution $X(t)$ at time T if the condition given below is satisfied;

$$\lim_{\Delta t \to 0} E\{|X(T) - x_{\Delta t}(T)|\} = 0 \tag{2.24}$$

where $x_{\Delta T}(T)$ is approximate solution at time T computed with constant stepsize Δt and E denotes the expected value [22].

When the condition below is satisfied

$$\lim_{\Delta t \to 0} E\{f(X(T))\} = E\{f(x_{\Delta t}(T))\} \tag{2.25}$$

where f is any continuously differentiable function, then it weakly converges to the solution [42].

2.4.2 Application of Euler-Maruyama to B-S SDE

In this part of the thesis, it is aimed to show the solution of an SDE obtained by using numerical approximation technique, Euler-Maruyama, is approaching to the exact solution. Recall that B-S is in the following form;

$$dX = \mu X dt + \sigma X dW_t, \;\; X(0) = X_0. \tag{2.26}$$

In the above equation μ and σ are fixed. We present the analytic solution previously and it is given in the Equation (2.27).

$$X(t) = X_0 \exp(\mu - \frac{\sigma^2}{2})t + \sigma W(t). \tag{2.27}$$

For the application of Euler-Maruyama solution of the Equation (2.26) over the interval $[0, 1]$, this interval is discretised and $0 < \tau_0 < \tau_1 < ..... < \tau_{j+1} = 1$. In this simulation study, we take $\Delta t = T/N = 1/1000$ as the step size and it is defined as $\Delta t = \tau_{j+1} - \tau_j = (j+1)/1000 - (j)/1000 = 1/1000$ for $N = 1000$ and $\tau_j = j\Delta t$. Then, the numerical approximation of this;

$$X_{\tau_j} = X_{\tau_{j-1}} + \mu X_{\tau_{j-1}} \Delta t + \sigma X_{\tau_{j-1}} \sqrt{\Delta \tau_j} \eta_j, \quad j = 1, 2, ..., 1000 \tag{2.28}$$

where $\mu = 0.2, X_0 = 1, \eta_j \sim \mathcal{N}(0, 1)$ and $\sigma = 1$. In the Figure 2.3 analytic solution and numeric solution are presented. In the Figure 2.3, red line represents the analytic solution given in the Equation (2.27) and black line represents the numeric solution given in the Equation (2.28). By examining the Figure 2.3, it is said that the analytic solution and numeric solution overlap. This means that in case of not obtaining the analytic solution which is common situation for many SDEs, numeric solution is used as an alternative approach. In the Figure 2.4, $\mu = -0.2$ and $\sigma = 0.3$ are used but all other parameters kept same. By the Figure 2.4, it is said that analytic solution and numeric solution overlap. Notice also that in both figures the fluctuations are affected by value of σ.

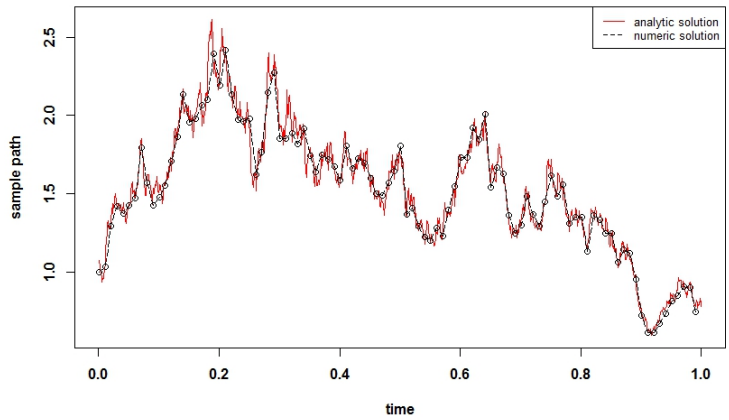

Figure 2.3: Application of Euler-Maruyama technique $\mu = 0.2, \sigma = 1$

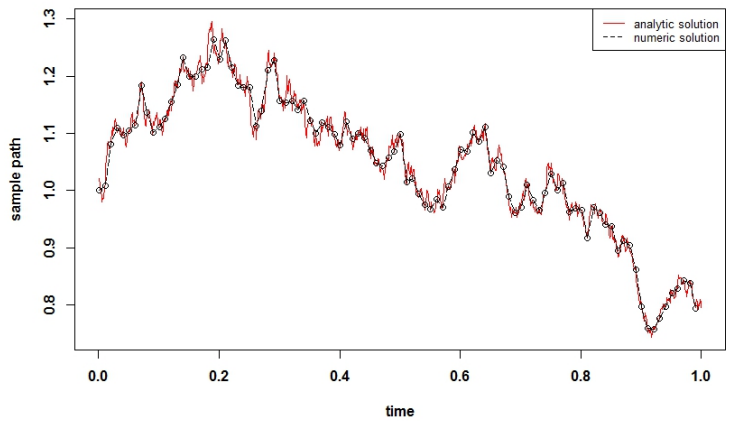

Figure 2.4: Application of Euler-Maruyama technique $\mu = -0.2, \sigma = 0.3$

For to get the logic of convergence, using different Δt values are needed. In the Table 2.1 different Δt values and their respective average errors which are mean of difference between analytic solution and numeric solution at time T are given. As we examine the Table 2.1, it is seen that the mean error is approaching to zero as the intervals become narrower. The Table 2.1 shows that analytic solutions obtained by Euler-Maruyama is applicable when closed-form of SDEs are not obtained since the mean error of solutions are approaching to zero.

Table 2.1: Mean error of analytic and numeric solutions for different values of N

N	Δt	Mean(Error)
2^5	$1/2^5$	0.0887
2^6	$1/2^6$	0.0197
2^7	$1/2^7$	0.0196
2^8	$1/2^8$	0.0396
2^9	$1/2^9$	0.0106
2^{10}	$1/2^{10}$	0.0087
2^{11}	$1/2^{11}$	0.0069
2^{12}	$1/2^{12}$	0.0029
2^{13}	$1/2^{13}$	0.00018

CHAPTER 3

MODEL CONSTRUCTION

In this section, it is aimed to model with a stochastic differential equation. To achieve this aim, some steps are necessary such as model building, assumption checking, goodness of fit and evaluation of model performance. Modelling with SDEs are similar to the modelling with ODEs. In general, most of SDE models are developed by the following way. Firstly, the potential changes of dependent variable with their transition probabilities are obtained in a small time interval Δt. Secondly, expectation and covariance of the change are obtained. Thirdly, with this information an SDE model is built for the dynamical system. Let S be a dynamic system, $S(t)$ be the value at t and ΔS be the change at Δt for the system S. Suppose that changes in the S are $\Delta S = \lambda$, $\Delta S = -\lambda$ and $\Delta S = 0$ respectively. Let $b_1(t, z)$ and $b_2(t, z)$ be transition densities. Possible changes and corresponding transition probabilities are given in the Table 3.1.

Table 3.1: Potential changes with respective probabilities

Changes	Probabilities
$\Delta S_1 = \lambda$	$p_1 = b_1(t, z)\Delta t$
$\Delta S_2 = -\lambda$	$p_2 = b_2(t, z)\Delta t$
$\Delta S_3 = 0$	$p_3 = 1 - (p_1 + p_2)$

By using information in the Table 3.1, expected change and second moment around

origin is calculated in the following way;

$$\begin{aligned} E(\Delta S) &= \sum_{i=1}^{3} p_i \Delta S_i = \lambda * b_1(t,z)\Delta t - \lambda * b_2(t,z)\Delta t \\ &= (b_1 - b_2)\lambda \Delta t \end{aligned} \tag{3.1}$$

and

$$\begin{aligned} E[(\Delta S)^2] &= \sum_{i=1}^{3} p_i (\Delta S_i)^2 = \lambda^2 * b_1(t,z)\Delta t + \lambda^2 * b_2(t,z)\Delta t \\ &= (b_1 + b_2)\lambda^2 \Delta t. \end{aligned} \tag{3.2}$$

Now, expectation, $\mu = \frac{E(\Delta S)}{\Delta t}$, and variance $\sigma^2 = \frac{E[(\Delta S)^2]}{\Delta t}$ are defined. Probability density of S at time t is $p(t,z)$ and by considering information in the Table 3.1,

$$\begin{aligned} p(t+\Delta t, z) = p(t,z)[1-(p_1+p_2)] + b_1(t, z-\lambda)p(t, z-\lambda) \\ \Delta t + b_2(t, z+\lambda)p(t, z+\lambda)\Delta t. \end{aligned} \tag{3.3}$$

In the Equation (3.3), when the second and third terms are expanded out about the point (t,z) by using Taylor series, following equations

$$b_1(t, z-\lambda)p(t, z-\lambda) \approx pb_1 - \frac{dpb_1}{dz}\lambda + \frac{1}{2}\frac{d^2pb_1}{dz^2}\lambda^2 \tag{3.4}$$

and

$$b_2(t, z+\lambda)p(t, z+\lambda) \approx pb_2 + \frac{dpb_2}{dz}\lambda + \frac{1}{2}\frac{d^2pb_2}{dz^2}\lambda^2 \tag{3.5}$$

are obtained. When we rewrite the Equation (3.3) by utilising these two equations

$$\begin{aligned} p(t+\Delta t, z) = p(t,z) - p(t,z)b_1\Delta t - p(t,z)d\Delta t + pb_1\Delta t - \frac{dpb_1}{dz}\lambda\Delta t \\ + \frac{1}{2}\frac{d^2pb_1}{dz^2}\lambda^2\Delta t + pb_2\Delta t + \frac{dpb_2}{dz}\lambda\Delta t + \frac{1}{2}\frac{d^2pb_2}{dz^2}\lambda^2\Delta t \end{aligned} \tag{3.6}$$

is obtained. It is seen that Fokker-Planck equation is solved approximately by $p(t,z)$ by assuming Δt and λ are small [1]. This means that rearranging the Equation (3.6) leads to the Equation (3.7)

$$\frac{dp}{dt}(t,z) = -\frac{d((b_1-b_2)\lambda p)}{dz}(t,z) + \frac{1}{2}\frac{d^2((b_1+b_2)\lambda^2 p)}{dz^2}(t,z). \tag{3.7}$$

In fact, the probability distribution $p(t, z)$ suffices the Equation (3.7) is similar to the distribution of solutions to the SDE system in the Equation (3.8) [1].

$$dz(t) = \mu(t)dt + \sigma(t)dB(t), \qquad z(t_0) = z_0. \tag{3.8}$$

By using the Equation (3.8), the dynamic system is modelled. Moreover, the drift and diffusion terms of an SDE are equal to the division of expected change by Δt and division of the square root of variance by Δt which are obtained from the discrete stochastic model. As a result, there are three necessary steps in order to build an SDE for a dynamic process.

In this study, we draw the graph of the process that we try to build stochastic model in order to determine the μ and σ. This means that if the graph shows that there exist high trend and fluctuations, these coefficients should be determined to represent these patterns. In the Figure 3.1, the population data given in [1] is considered to build model. In the Equation (3.9), the model constructed to this data is given. By looking at the Figure 3.1, it is said that there exist trend and small fluctuations in the data. Therefore, In the Equation (3.9), the deterministic part is representative for this trend. For the fluctuations, the second term in the Equation (3.9) with Bm is representative.

$$dY(t) = \alpha Y(t)dt + \sqrt{\beta Y(t)}dW(t) \tag{3.9}$$

In the Equation (3.9), the stochastic process to construct model is Bm. This means that the noise term in the SDE is represented by Bm. Before estimating parameters, it is necessary to check whether Bm assumptions are satisfied or not. In the Equation (3.9), it is not known that whether the closed form solution of this SDE exists or not. From the Chapter 2, it is known that E-M solution is used to obtain solution when the closed form solution of an SDE is not obtained. Therefore, it is necessary to obtain the difference $Y_i - Y_{i-1}$ where $i = 1, 2, .., 47$ in order to check the Bm assumption.

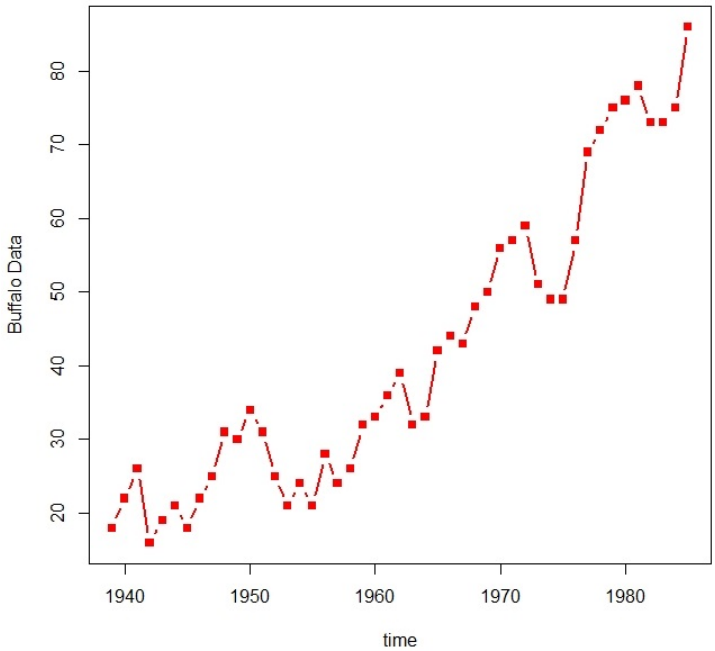

Figure 3.1: Buffalo population from year 1939 to 1985

After model construction and model assumption checking are mentioned, model validation and model performance for forecast are mentioned. Goodness of fit test procedure for the model validation is mentioned. For the model performance mean absolute percentage error(MAPE) method is mentioned.

3.1 Assumption Checking

Before the data is tested for Bm, any seasonal variation should be removed from the data. One can see [35] for the detail information about removing seasonal variation. Considering the properties of Brownian motion, there are three properties to be checked [40]:

1. Stationarity
2. Normality
3. Independence of increments

3.1.1 Stationarity

In order to test the correlational structure of time series, stationarity is the vital property. For the stationary process, the properties of stochastic process are constant over time [23]. "When the joint distribution of any part of the series of $Z_{t_1}, Z_{t_2}, ..., Z_{t_n}$ have same distribution with any other part of the series $Z_{t_{1+k}}, Z_{t_{2+k}}, ..., Z_{t_{n+k}}$, where k is any integer, time series $(Z_1, Z_2, ...,)$ is stationary" is stated by Kadılar (2005) in his book [30]. It is meant that series distribution is not affected by shifting the time series.

In our study, we are looking for the stationarity of $(Y_t - Y_{t-1})$. When this difference is not stationary, variance of the process is not constant and it changes in time t. In our study, two well-known tests are used. One of them is KPSS test [33] and the other one is ADF test [12].

KPSS test:

Non-stationarity is shown by unit root presence in the series [25]. One of the most powerful tools to test unit root is KPSS test. Hypothesis for this test is constructed as

1. H_0: Y_t is stationary
2. H_1: Y_t is a unit root process.

If test statistics is large, the rejection of H_0 occurs . For a detailed information about the test statistics and its calculation, [8] is applicable.

ADF test:

Another commonly known unit root test is ADF test. For this test, the construction of the hypothesis is vice versa of the construction of the hypothesis of the KPSS test. Hypothesis of this test is constructed as

1. H_0: Y_t is a unit root process
2. H_1: Y_t is stationary.

When the critical value is bigger than the test statistics, the rejection of H_0 does not occur.

3.1.2 Normality

Plotting the histogram of difference of data obtained by E-M and comparing it to a normal distribution plot is an easiest way to check normality. But this only supports visiual detection. Therefore, statistical test for normality should be conducted. In this study two of the normal tests are utilised. One of them is Shapiro-Wilk and tho other one is Jarque-Bera test. The hypothesis for these tests are same. That is,

1. H_0: Distribution is normal
2. H_1: Distribution is not normal

and decision is taken by considering the test p-value. When it is bigger than the specified level of significance (p-value$> \alpha$), then the H_0 is not rejected.

3.1.3 Independence of increments

Investigating whether the increments are independent or not is necessary to build an appropriate model. Plotting the auto-correlation is an easiest way to detect the independence. However, it supports visual results and it may not be reliable. Therefore, Box-Ljung or Box-Pierce is used to test. For the detail information about the test procedure, [30] can be applicable. For this test, hypothesis is constructed as

1. H_0: Auto-correlation coefficients for all lags are equal to zero
2. H_1: One is different from zero

and if p-value is greater than specified level of significance (p-value$> \alpha$), the H_0 is not rejected and is said that independence of increments is satisfied.

3.2 Goodness of fit of SDE

After developing an SDE model, checking whether the proposed model is fitting to the data or not is necessary. For this aim, we use a method called as GOF test which is easy to apply to test the model fit to the data in this part. For this test, the core idea is to obtain Monte-Carlo simulations of trajectories between consecutive observations and compare these observations with original data at each point. Based on comparing the original data and simulated observations at each point, ranks are obtained and depending on these ranks test statistics is observed [3].

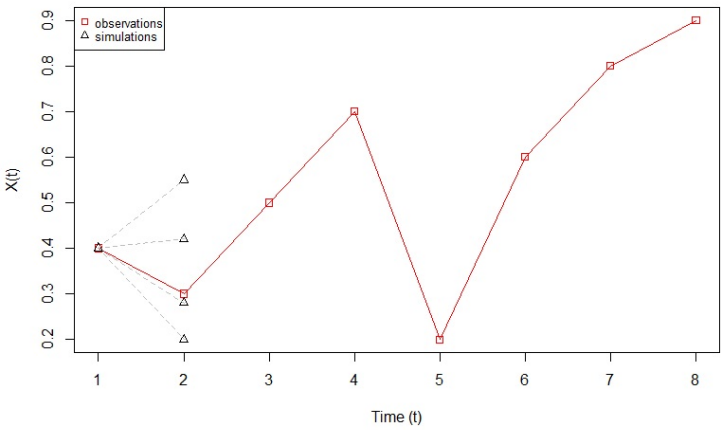

Figure 3.2: Simulated trajectories for $t \in [1, 2]$ and original data values

Suppose that the process $Y(t), t > 0$ is observed at time points $t_1, t_2, .., t_N$ and $Y_1, Y_2, ...Y_N$ be observations. M Simulated values for each t_i where $i = 2, ..., N$ are obtained between t_{i-1} and t_i beginning from Y_{i-1}. After this, let r_t be the rank values of each Y_t when compare to the M simulated values. In the Figure 3.2 for $t = 2$ there exist 4 simulated values represented by black triangle box and also original data value represented by red rectangular box. When we compare these in the Figure 3.2, we observe that the original data value is bigger than two of the Mth simulated values and smaller than two of these. That is, $r_2 = 3$. Under H_0 it is expected that by considering r_t is a stochastic observations $P(r_t = k) = p_{tk} = \frac{1}{M+1}$ where $k = 1, 2, ..., M + 1$ and $t = 2, 3, ...N$. Usually, it is of interest that p_{tk} is dependent on t and k or not. On the other hand, only one observation is available and therefore, it should be assumed

that p_{tk} is not dependent of time. Actually, the core idea of the test is that ranks are taking values between 1 to $M+1$ with equal probability and under the null hypothesis χ^2 GOF test is used [31]. In order to apply this, it is necessary to find observed and expected frequencies and $1 \leq r_t \leq M+1$ for each $t = 2, ..., N$. Let $I_{t,k}$ be an indicator function that $I_{t,k} = 1$ if $r_t = k$ $I_{t,k} = 0$ otherwise. Observed frequency is Ω_k and $\Omega(k) = \sum_{t=2}^{N} I_{t,k}$. Under the null hypothesis the expected frequency is $\frac{N-1}{M+1}$ due to $\sum_{k=1}^{M+1} \Omega(k) = N-1$. Thus, test statistics is

$$\chi^2(M) = \sum_{k=1}^{M} \frac{(\Omega(k) - \frac{N-1}{M+1})^2}{\frac{N-1}{M+1}}. \tag{3.10}$$

When large value of test statistics is obtained, it is the signal of lack of fit. Small expected frequency under the null hypothesis leads to failure of χ^2 approximation [31]. Also, this is useful for determining the number of simulations, M. In fact, expected frequency should be at least 5, and by applying $\frac{N-1}{M+1} \geq 5$ $M \leq \frac{(N-6)}{5}$. This means that the number of simulations should be at most equal to $\frac{(N-6)}{5}$. In addition to test whether there exists lack of fit or not, this test is useful to select the best model amongst others. The test value obtained by using the Equation (3.10) should be compared for each model in order to select the best model. Model which gives the smallest test value is the best model that fits to the data. In other words, the smallest test value means that there is compatibility between the model and the data, because there is not much deviance of observed frequency from the expected frequency.

3.3 Model Performance Evaluation

After construction of the model, it is important to measure prediction accuracy of the model. In this part of the study, MAPE is presented to achieve this aim.

Mean Absolute Percentage Error: It is used to measure the accuracy of prediction for different models or for different time. This is calculated by using

$$MAPE = \frac{1}{k}(\sum_{j=1}^{k} \frac{|A_j - F_j|}{A_j}).100 \tag{3.11}$$

where k is the observation number, A_j is observed values and F_j is the forecast value from the model. Interpretations of different MAPE values are given in the Table 3.2 in [19].

Table 3.2: Interpretation of MAPE values

Value	Interpretation
$MAPE < 0.1$	Perfect forecast
$0.1 \leq MAPE < 0.2$	Good forecast
$0.2 \leq MAPE < 0.5$	Reasonabe forecast
$MAPE \geq 0.5$	Imperfect forecasting

CHAPTER 4

PARAMETER ESTIMATION

It is aimed to present the parameter estimation techniques which are vital for SDE in this part. SDE having form of $dY(t) = f(t, Y(t); \theta)dt + g(t, Y(t); \theta)dW(t)$ is used to model the problem in our study. That is, we try to model the change in our variable with respect to time by using this equation. In this equation $\theta \in R^m$ is unknown parameter vectors. There is an assumption that $y_0, y_1, ..., y_N$ are observance of $Y(t)$ and they are observed in the equally distributed times $t_j = j\Delta t$ where $j = 0, 1, ..., N$ and $\Delta t = \frac{T-0}{N}$. In this study, the core idea is to estimate the θ vector by using data to model the problem in which we are interested. In this section, two estimation methods are presented. One is maximum likelihood estimation (MLE) method, and the other is generalized method of moments (GMM) which is a non-parametric estimation technique.

4.1 Maximum Likelihood Estimation Technique

An SDE having the following form

$$dY(t) = f(t, Y(t); \theta)dt + g(t, Y(t); \theta)dW(t) \tag{4.1}$$

is taken into consideration. In the Equation (4.1) we try to estimate the $\theta \in R^m$ vector by using maximum likelihood approximation technique. For a detailed mathematical explanation and proof of this technique, please refer to [11], [15], [17], [27], and [29]. For to estimate parameters in the Equation (4.1) by using MLE technique, construction of the likelihood function is essential. For this reason, knowing the transition

probability density function (pdf) is necessary. When the Equation (4.1) is solved, Y is obtained. Let $y_0, y_1, , , y_N$ be observed values obtained at uniform time intervals. Let $p(t_j, y_j|t_{j-1}, y_{j-1}; \theta)$ be the transition pdf of (t_j, y_j) starting from (t_{j-1}, y_{j-1}), given θ and $p_0(y_0|\theta)$ as the initial state density. For the MLE of parameters, the joint density

$$H(\theta) = p_0(y_0|\theta) \prod_{j=1}^{N} p(t_j, x_j|t_{j-1}, x_{j-1}; \theta) \tag{4.2}$$

is maximized over $\theta \in R^m$. However, it is easy to minimize the function $-ln(H(\theta))$, that is the logarithm of $H(\theta)$. If we write the likelihood function by using $-ln(H(\theta))$ instead of $H(\theta)$, we have the following form

$$-ln(H(\theta)) = -ln(p_0(y_0|\theta)) - \sum_{j=1}^{N} ln(p(t_j, x_j|t_{j-1}, x_{j-1}; \theta)) \tag{4.3}$$

which is minimized over $\theta \in R^m$. The basic step to write the Equation (4.3) is to obtain the transition pdf of the process, but it is not easy to obtain the transition pdf due to the lack of closed form solutions for many SDEs. Therefore, one of the difficulties of this technique is obtaining transition pdf of process. In order to deal with this difficulty, one may use discrete maximum likelihood (DML) estimation rather than exact maximum likelihood estimation. In the DML approach, the core idea is to approximate the transitional pdf by using the Euler-Maryuama approach with one step of increment $\Delta = (t_j - t_{j-1})$. In order to generate approximate solution for the transition pdf with E-M technique, it is

$$Y_j \approx y_{j-1} + f(t_{j-1}, y_{j-1}; \theta)\Delta t + g(t_{j-1}, y_{j-1}; \theta)\sqrt{\Delta t}\eta_j \tag{4.4}$$

where $\eta_j \sim \mathcal{N}(0, 1)$. From the Equation (4.4), it is concluded that the transition pdf of Y is approximately equal to normal distribution with the following parameters

$$\mu_j = y_{j-1} + f(t_{j-1}, y_{j-1}; \theta)\Delta t$$
$$\sigma_j = g(t_{j-1}, y_{j-1}; \theta)\sqrt{\Delta t}.$$

This means that the transition pdf $p(t_j, y_j|t_{j-1}, y_{j-1}; \theta) \sim \mathcal{N}(\mu_j, \sigma_j)$ has the following form

$$p(t_j, y_j; \theta) \approx \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp(\frac{-(y_j - \mu_j)^2}{2\sigma_j^2}). \tag{4.5}$$

In this method, another difficulty is to compute optimal θ obtained by minimizing the likelihood function $-ln(H(\theta))$. To deal with this difficulty, Nelder-Mead numerical optimization algorithm is used. For a detailed information for this method, [39] could be followed.

An example given in [1] for fitting population data to an SDE is useful in order to understand the underlying mathematics of MLE method. In this example, for Buffalo population from 1939 to 1985, following SDE model to the data

$$dY(t) = \alpha Y(t)dt + \sqrt{\beta Y(t)}dW(t) \tag{4.6}$$

is fitted. In this model, $Y(t)$ is the population size, the initial value is $Y(0) = 18$, and α, β are unknown parameters. Firstly, E-M schema should be formed to find the transition pdf by the following way;

$$Y_j \approx Y_{j-1} + \alpha Y_{j-1}\Delta t + \sqrt{\beta Y_{j-1}\Delta t}\eta_j \tag{4.7}$$

where $\eta_j \sim \mathcal{N}(0,1)$ and $\Delta t = 1$. This implies that $p(t_j, Y_j|t_{j-1}, Y_{j-1}; \alpha, \beta) \sim \mathcal{N}(\mu_j, \sigma_j)$ where $\mu_j = Y_{j-1} + \alpha Y_{j-1}\Delta t$ and $\sigma_j = \sqrt{\beta Y_{j-1}\Delta t}$. If this is utilised in the likelihood function, then

$$\begin{aligned} H(\alpha,\beta) &= p_0(y_0|\alpha,\beta)\prod_{j=1}^{N}\frac{1}{\sqrt{2\pi\sigma_j^2}}\exp(\frac{-(y_j-\mu_j)^2}{2\sigma_j^2}) \\ -ln(H(\alpha,\beta)) &= -\ln(p_0(Y_0|\alpha,\beta)) + \sum_{j=1}^{N}\frac{1}{2}\ln(2\pi) + \sum_{j=1}^{N}\frac{1}{2}(\ln\beta Y_{j-1}) \\ &\quad + \sum_{j=1}^{N}(\frac{(Y_j - Y_{j-1} - \alpha Y_{j-1})^2}{2\beta Y_{j-1}}) \end{aligned} \tag{4.8}$$

the Equation (4.8) is obtained. Minimizing the likelihood function is obtained by taking the derivative of the Equation (4.8) in terms of α, β respectively and setting them equal to zero as follows.

$$\begin{aligned} \frac{dH(\alpha,\beta)}{d\alpha} &= \sum_{j=1}^{N}((2(-Y_{j-1}\Delta t))\frac{(Y_j - Y_{j-1} - \alpha Y_{j-1}\Delta t)}{2\beta Y_{j-1}\Delta t}) = 0 \\ \hat{\alpha} &= \frac{\sum_{j=1}^{47} Y_{j-1}(Y_j - Y_{j-1})}{\sum_{j=1}^{47}(Y_{j-1}Y_{j-1}\Delta t)} = \frac{\sum_{j=1}^{47}(Y_j - Y_{j-1})}{\sum_{j=1}^{47}(Y_{j-1}\Delta t)}. \end{aligned} \tag{4.9}$$

$$\frac{dH(\alpha,\beta)}{d\beta} = \sum_{k=1}^{N}(\frac{Y_{j-1}}{2Y_{j-1}\beta} - \frac{1}{2\beta^2}\frac{(Y_j - Y_{j-1} - \alpha Y_{j-1}\Delta t)^2}{Y_{j-1}\Delta t})) = 0$$

$$\hat{\beta} = \frac{\sum_{j=1}^{47}(Y_j - Y_{j-1} - \hat{\alpha}Y_{j-1}\Delta t)^2}{\sum_{j=1}^{47}(Y_{j-1}\Delta t)}. \tag{4.10}$$

4.2 Generalized Method of Moments

In the GMM, the core idea is to specify the moment conditions. When the GMM is utilised to estimate parameters , there are many ways to obtain moment conditions. In this part of thesis, first order E-M discretization of SDE is used to derive a set of approximate moment conditions. This method was used by Chan, Karolyi, Longstaff and Sanders (1992) in [9]. Another method is to derive moment conditions based on using the infinitesimal generator proposed by Hansen and Scheinkman [24]. For detailed information and methods about deriving moment conditions [20] and [13] are applicable.

Firstly, E-M schema of the process is obtained to obtain GMM estimator by using the E-M method. In the Equation (4.11) the schema is

$$Y_i = Y_{i-1} + f(t_{i-1}, Y_{i-1}; \theta)\Delta t + g(t_{i-1}, Y_{i-1}; \theta)\sqrt{\Delta t}\eta_i \tag{4.11}$$

where $\Delta t = \frac{T}{N}$ and $i = 1, 2, ..N$. From the Equation (4.11), $Y_i \sim \mathcal{N}(Y_{i-1} + f(t_{i-1}, Y_{i-1}; \theta)\Delta t, g^2(t_{i-1}, Y_{i-1}; \theta)\Delta t)$. According to the Equation (4.11) two moment conditions are derived as,

$$\begin{aligned} &\phi_1(Y;\theta) : E[Y_i - Y_{i-1} - f(t_{i-1}, Y_{i-1}; \theta)\Delta t] = 0 \\ &\phi_2(Y;\theta) : E[(Y_i - Y_{i-1} - f(t_{i-1}, Y_{i-1}; \theta)\Delta t)^2 - g^2(t_{i-1}, Y_{i-1}; \theta)\Delta t] = 0. \end{aligned} \tag{4.12}$$

The core idea is to specify a number of moment conditions' number which should be equal to parameter numbers [9]. At least, these should be equal to number of parameters which one tries to estimate. When parameters of an SDE are obtained from the two dimensional parameter space ($\theta \in \mathbf{R}^2$), the Equation (4.12) is used for estimation of parameters. For higher dimensional cases ($\theta \in \mathbf{R}^p$ where $p > 2$) some extra terms should be added to the Equation (4.12). The following equations

for higher dimensional cases

$$\begin{aligned}\phi_{2j+1}(Y;\theta) &= \phi_1(Y;\theta)Y^j \\ \phi_{2j+2}(Y;\theta) &= \phi_2(Y;\theta)Y^j\end{aligned} \tag{4.13}$$

where $j = 1, 2, ...$

$$\begin{aligned}\phi_3(Y;\theta) &: E[(Y_i - Y_{i-1} - f(t_{i-1}, Y_{i-1};\theta)\Delta t)Y_i] = 0 \\ \phi_4(Y;\theta) &: E[(((Y_i - Y_{i-1} - f(t_{i-1}, Y_{i-1};\theta)\Delta t)^2)Y_i \\ -g^2(t_{i-1}, Y_{i-1};\theta)\Delta t)Y_i] &= 0.\end{aligned} \tag{4.14}$$

are used.

Unlike the parametric approach, this method is easier to implement and there is no distributional. In other words, there is no distributional assumption such as normality about the changes. The only requirement is that the change should be stationary [9]. Previous example given in [1] for fitting population data to an SDE is useful to understand the mathematics underlying GMM method. In this example, for Buffalo population from 1939 to 1985, following the SDE model

$$dY(t) = \alpha Y(t)dt + \sqrt{\beta Y(t)}dW(t) \tag{4.15}$$

is fitted to the data. Firstly, E-M schema should be formed to derive moment conditions by;

$$Y_j \approx Y_{j-1} + \alpha Y_{j-1}\Delta t + \sqrt{\beta Y_{j-1}\Delta t}\eta_j \tag{4.16}$$

where $\eta_j \sim \mathcal{N}(0,1)$ and $\Delta t = 1$. From the Equation (4.16) two moment conditions are derived and estimator of parameters are obtained as follows;

$$\begin{aligned}\phi_1(Y;\alpha) &: E[(Y_i - Y_{i-1} - f(t_{i-1}, Y_{i-1};\alpha)\Delta t)] = 0 \\ \phi_2(Y;\beta) &: E[((Y_i - Y_{i-1} - f(t_{i-1}, Y_{i-1};\alpha)\Delta t)^2 \\ -g^2(t_{i-1}, Y_{i-1};\beta)\Delta t)] &= 0.\end{aligned} \tag{4.17}$$

When f and g functions are placed in the Equation (4.17), we observe

$$\begin{aligned}\hat{\alpha} &= \frac{\sum_{i=1}^{46}(Y_i - Y_{i-1})}{\sum_{i=1}^{46} Y_{i-1}} \\ \hat{\beta} &= \frac{\sum_{i=1}^{46}(Y_i - Y_{i-1} - \hat{\alpha}Y_{i-1})^2}{\sum_{i=1}^{46} Y_{i-1}}.\end{aligned} \tag{4.18}$$

CHAPTER 5

APPLICATION

In this section, it is aimed to show how to apply modelling with SDEs step by step. For this purpose, firstly, example given in Chapter 3 is used. For this first example, we show how to check assumptions, obtaining estimation of parameters both using MLE and GMM methods and how to check model lack of fit by using Goodness of Fit. Also, selecting the best model among the constructed models is shown. To show the model selection, goodness of fit is used.

5.1 Buffalo Population Example

An example given in [1] for fitting population data to an SDE is useful to understand the steps for modelling. In this example, for Buffalo population from 1939 to 1985, following SDE model

$$dY(t) = \alpha Y(t)dt + \sqrt{\beta Y(t)}dW(t) \tag{5.1}$$

is fitted. $Y(t)$ represents the population size with initial value $Y(0) = 18$ and α, β are unknown parameters. Firstly, E-M schema should be formed;

$$Y_j \approx Y_{j-1} + \alpha Y_{j-1}\Delta t + \sqrt{\beta Y_{j-1}\Delta t}\eta_j \tag{5.2}$$

where $\eta_j \sim \mathcal{N}(0,1)$ and $\Delta t = 1$. This implies that $p(t_j, Y_j|t_{j-1}, Y_{j-1}; \alpha, \beta) \sim \mathcal{N}(\mu_j, \sigma_j)$ where $\mu_j = Y_{j-1} + \alpha Y_{j-1}\Delta t$ and $\sigma_j = \sqrt{\beta Y_{j-1}\Delta t}$.

Before estimating the parameters, assumption checking should be done. Assumption checking is applied to $Y_j - Y_{j-1}$ series. In other words, we test that $\eta_j =$

$\frac{Y_j - Y_{j-1} - \alpha Y_{j-1}\Delta t}{\sqrt{\beta Y_{j-1}\Delta t}} \sim \mathcal{N}(0,1).$

Stationarity

Table 5.1: Result of KPSS and ADF Test

Tests	P-value	Result
KPSS	0.1	0.1 > 0.05
ADF	0.01	0.01 < 0.05

From the Table 5.1, it is observed that the series are stationary. P-value of KPSS test is greater than specified level of significance and thus the hypothesis of stationary series is not rejected. The same result that is the reverse hypothesis in ADF test is obtained. This means that the hypothesis of the series have a unit root that leads to break stationarity is rejected because ADF test p-value is smaller than the specified level of significance. As a result series are stationary.

Normality

Table 5.2: Result of Shapiro-Wilk and Jarque-Bera Test

Tests	P-value	Result
Shapiro-Wilk	0.44	0.44 > 0.05
Jarque-Bera	0.81	0.81 > 0.05

From the Table 5.2, it is observed that there is no evidence against that the series are not normal. P-value of S-W test is greater than the specified level of significance, and thus, the null hypothesis that the series are normal is not rejected. The same result is seen in the Jarque-Bera test and so the hypothesis of normality is not rejected.

From the Table 5.3, it is observed that there is no auto-correlation. That is, auto-correlation coefficients for all lags are not different from zero. P-value of Box-Pierce test is bigger than the specified level of significance and thus, the hypothesis of auto-

correlation coefficients for all lags are zero is not rejected.

Independence of Increments

Table 5.3: Result of Box-Pierce Test

Test	P-value	Result
Box-Pierce	0.808	p-value > α value(0.05)

In Chapter 3, MLE and GMM techniques are mentioned and an example is given to understand the mathematics underlying these techniques. When we use MLE and GMM techniques to estimate the parameters of SDE given in the Equation (5.1), estimators are found in the Equation (5.3) and the Equation (5.4) for both techniques, respectively.

$$\begin{aligned} \hat{\alpha} &= \frac{\sum_{j=1}^{47}(Y_j - Y_{j-1})}{\sum_{j=1}^{47}(Y_{j-1}\Delta t)} = 0.0361 \\ \hat{\beta} &= \frac{\sum_{j=1}^{47}(Y_j - Y_{j-1} - \alpha Y_{j-1}\Delta t)^2}{\sum_{j=1}^{47}(Y_{j-1}\Delta t)} = 0.529. \end{aligned} \tag{5.3}$$

$$\begin{aligned} \hat{\alpha} &= \frac{\sum_{i=1}^{46}(Y_i - Y_{i-1})}{\sum_{i=1}^{46} Y_{i-1}} = 0.0361 \\ \hat{\beta} &= \frac{\sum_{i=1}^{46}(Y_i - Y_{i-1} - \alpha Y_{i-1})^2}{\sum_{i=1}^{46} Y_{i-1}} = 0.529. \end{aligned} \tag{5.4}$$

As we examine the estimators of parameters, it is observed that estimators are same if the assumptions are satisfied. This implies that series $(Y_i - Y_{i-1})$ satisfy the assumptions which are stationary, normal and independent increments, estimators of parameters are same for both techniques. As we utilise the estimate of the parameters in the Equation (5.1), the following Equation (5.5) is obtained.

$$dY(t) = 0.0361Y(t)dt + \sqrt{0.529Y(t)}dW(t) \tag{5.5}$$

For this problem, another model is given in the following form in [1]

$$dY(t) = \frac{\alpha}{Y^2(t)}dt + \frac{\beta}{Y^2(t)}dW(t) \tag{5.6}$$

where $Y(t)$ is the population size, α, β are unknown parameters and $Y(0) = 18$ is the initial value. The parameters are estimated using MLE. Before estimating parameters, E-M schema is obtained to get transition pdf for the likelihood function. The E-M schema of the Equation 5.6 is obtained as follows

$$Y_j \approx Y_{j-1} + \frac{\alpha}{Y_{j-1}^2}\Delta t + \frac{\beta}{Y_{j-1}^2}\sqrt{\Delta t}\eta_j \tag{5.7}$$

where $\eta_j \sim \mathcal{N}(0,1)$. From the Equation (5.7) it is inferred that the transitional pdf of Y is approximately normal distribution with the following parameters

$$\mu_j = y_{j-1} + \frac{\alpha}{Y_{j-1}^2}\Delta t$$
$$\sigma_j = \frac{\beta}{Y_{j-1}^2}\sqrt{\Delta t}.$$

When we utilise these in the likelihood function,

$$H(\alpha,\beta) = p_0(y_0|\alpha,\beta)\prod_{j=1}^{N}\frac{1}{\sqrt{2\pi\sigma_j^2}}\exp(\frac{-(y_j-\mu_j)^2}{2\sigma_j^2})$$
$$-ln(H(\alpha,\beta)) = -\ln(p_0(Y_0|\alpha,\beta)) + \sum_{j=1}^{N}\frac{1}{2}\ln(2\pi) + \sum_{j=1}^{N}\frac{1}{2}(\ln(\frac{\beta^2}{Y_{j-1}^4}))$$
$$+\sum_{j=1}^{N}(\frac{(Y_j - Y_{j-1} - \frac{\alpha}{Y_{j-1}^2})^2}{2\frac{\beta^2}{Y_{j-1}^4}}) \tag{5.8}$$

The Equation (5.8) is obtained. Minimizing the likelihood function is obtained by taking the derivative of the Equation (5.8) in terms of α, β respectively and equate the obtained equations to zero results in

$$\frac{\partial H(\alpha,\beta)}{\partial\alpha} = \sum_{j=1}^{N}((2(\frac{\alpha}{Y_{j-1}^2}))\Delta t)\frac{(Y_j - Y_{j-1} - \frac{\alpha}{Y_{j-1}^2}\Delta t)}{2\frac{\beta^2}{Y_{j-1}^4}\Delta t}) = 0$$
$$\hat{\alpha} = \frac{\sum_{j=1}^{47}(Y_jY_{j-1}^2 - Y_{j-1}^3)}{46} = 3510. \tag{5.9}$$

and

$$\frac{\partial H(\alpha,\beta)}{\partial \beta} = \sum_{j=1}^{N}(\frac{1}{\beta} - \frac{Y_{j-1}^{4}(Y_j - Y_{j-1} - \frac{\alpha}{Y_{j-1}^{2}})^2}{\beta^3}) = 0$$

$$\hat{\beta} = \sqrt{\sum_{j=1}^{N} \frac{Y_{j-1}^{4}(Y_j - Y_{j-1} - \frac{\alpha}{Y_{j-1}^{2}})^2}{46}} = 13500 \tag{5.10}$$

As we utilise the estimate of parameters in the Equation 5.6, the following Equation 5.11 is obtained.

$$\partial Y(t) = \frac{3510}{Y^2(t)}\partial t + \frac{13500}{Y^2(t)}\partial W(t) \tag{5.11}$$

After developing an SDE model, it is necessary to check whether the proposed model is fitting to the data or not. In order to achieve this aim, goodness of fit test in Chapter 3 is applied. For application of this test, Monte-Carlo simulations of trajectories between consecutive observations are obtained, and these observations are compared with original data at each point.

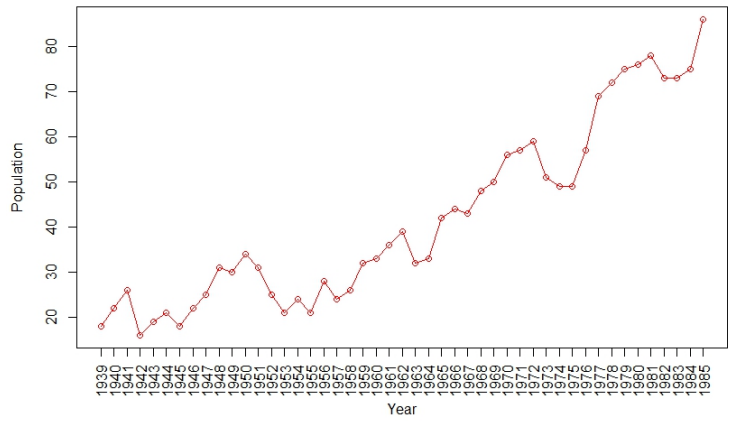

Figure 5.1: Original Data Values from 1939 to 1985

In Figure 5.1 the original data values in example are drawn. To apply GOF test it is important to make Monte-Carlo simulations for each data. In the simulation, observation obtained in 1939 is initial value and simulated values are obtained from 1940 to 1985. For this data at most 8 simulated values for each points are done and ranks are obtained according to this.

In Figure 5.2, simulated and original data values are drawn. Black stars represent the simulated values and red circles represent the original data values. Using the Figure 5.2, it is obtained that there is no lack of fit due to the ranks which are obtained by comparing original and simulated values that are close to each other. This means that if there exists lack of fit, then it could have deviated from the expected ranks, and it may have led to increase in the test statistics which results in rejection of null hypothesis.

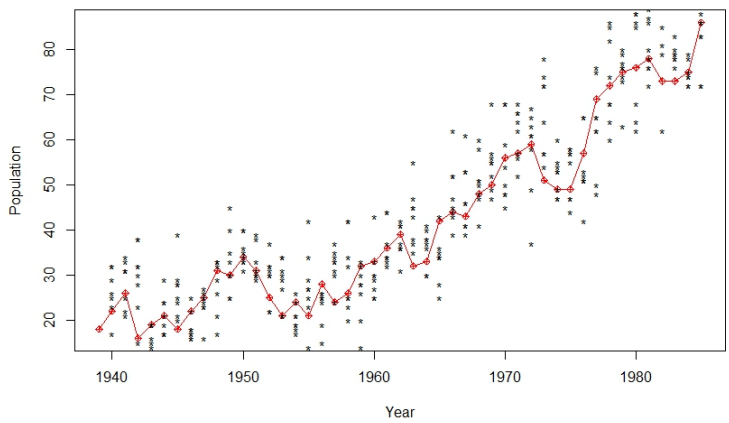

Figure 5.2: Simulated and Original Data

Actually, the Figure 5.2 is visual detection of lack of fit. To understand clearly the lack of fit, GOF is used and the Table 5.4 presents the rank values. The Table 5.4 suggests that only one of the original data is bigger than all the simulated values, and this is seen in the Figure 5.2. This refers to the observation obtained in 1943. To explain the Table 5.4 clearly, it is said that ranks values do not deviate from the expected values ($\frac{N-1}{M+1} = 46/9 = 5.11$) and test statistics given in Chapter 3 is obtained 7.218. Because of the estimation of (α, β) from the data, the degrees of freedom is equal to $(M-2)$. The p-value is $P(\chi^2(M-2) > 7.218) = 0.69$ and is greater than $\alpha = 0.05$. According to the test result, the null hypothesis that there is no lack of fit between model and the data is not rejected due to $p - value > \alpha$.

Table 5.4: Result of GOF test

Ranks	Values
$k = 1$	4
$k = 2$	6
$k = 3$	4
$k = 4$	9
$k = 5$	6
$k = 6$	6
$k = 7$	5
$k = 8$	5
$k = 9$	1

Another model in the Equation (5.6) is proposed for this data. After estimating parameters in this model, it is necessary to check whether the model is fitting the data or not. In order to achieve this aim, it is necessary to use GOF test. Before applying to this test, the lack of fit is visually determined from the Figure 5.3. As we examine the Figure (5.3), it is said that the rank values are deviated from the expected values $\frac{N-1}{M+1} = 5.11$.

Also, the lack of fit is determined by using this test. When we use this test, test statistics given in Chapter 3 is obtained 14.63342. The test p-value is $P(\chi^2(M-2) > 14.63342) = 0.023$ and is less than $\alpha = 0.05$. According to the result of the test, the null hypothesis that there is no lack of fit between model and the data is rejected due to $p - value < \alpha$.

If we did not reject the model given in the Equation (5.6) at 0.05 significance level, we could have selected the best models with having smallest test values. However, since the model given in the Equation (5.6) shows the lack of fit, first model given in the Equation (5.1) is selected to this data.

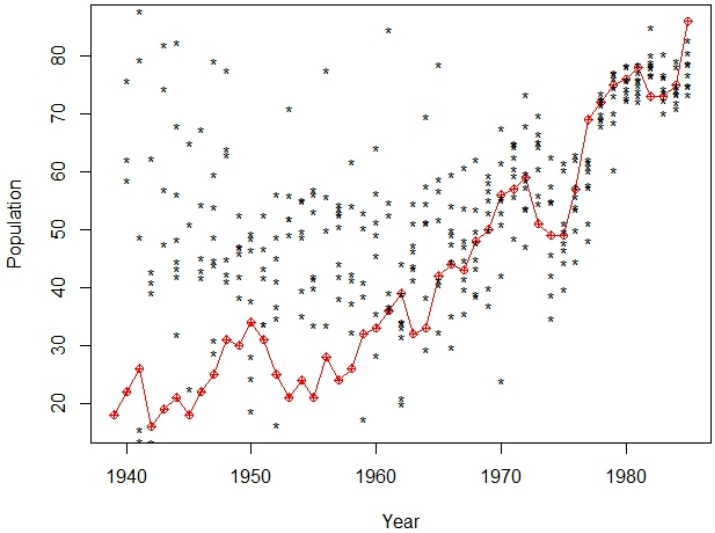

Figure 5.3: Simulated and Original Data

5.2 Real Data example

It is aimed to model the mandarin quantity between 1980 and 2014 in Turkey. This data is collected by Turkish Statistical Institute. We draw the graph of process that we try to build stochastic model in order to determine trend and diffusion. In the Figure 5.4, there exist trend and small fluctuations in the data. To represent the trend, the deterministic part in the following equations is representative. For the diffusion part, two different terms with Bm is representative. After model construction is completed, model assumptions are checked and result is given in tables from the Table 5.5 to Table 5.7. When we look at the tables, all assumptions are satisfied and thus, we use only MLE techniques to estimate parameters. Note that under the satisfied assumption, both estimation techniques give same estimators for parameters. In the Equation (5.12) and the Equation (5.13), $Q(t)$ represents mandarin production at time t, and where α and β are unknown parameters. To estimate parameters and to apply GOF procedure observations, data between 1980 and 2014 are used. This means that 35 observations are utilised for estimation and GOF procedure. Observations obtained from 2015 to 2019 are utilised for the model performance evaluation.

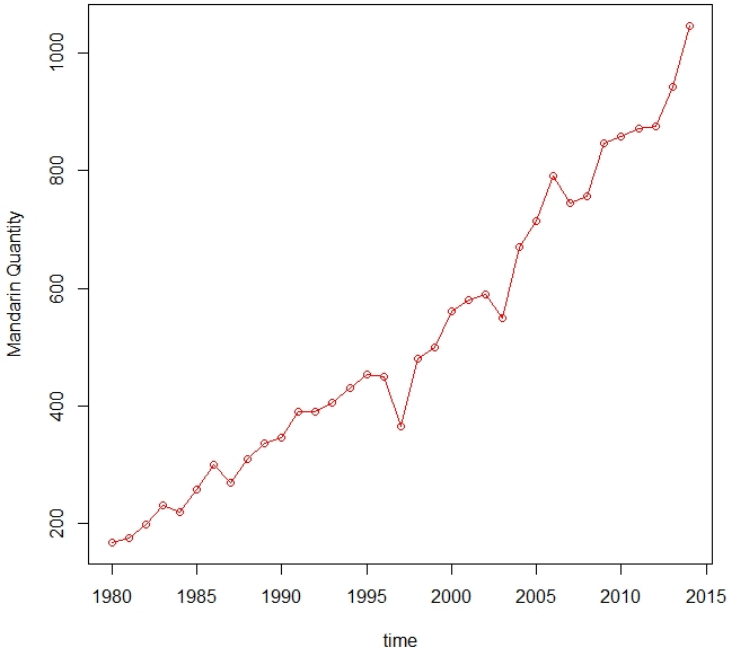

Figure 5.4: Mandarin quantity between 1980 and 2014

$$dQ(t) = \alpha Q(t)dt + \sqrt{\beta Q(t)}dW(t) \tag{5.12}$$

and

$$dQ(t) = \alpha Q(t)dt + Q(t)^{\beta}dW(t) \tag{5.13}$$

E-M schema is obtained as in the following equations for the Equation (5.12) and the Equation (5.13),

$$Q_j \approx Q_{j-1} + \alpha Q_{j-1}\Delta t + \sqrt{\beta Q_{j-1}\Delta t}\eta_j \tag{5.14}$$

and

$$Q_j \approx Q_{j-1} + \alpha Q_{j-1}\Delta t + Q_{j-1}^{\beta}\sqrt{\Delta t}\eta_j \tag{5.15}$$

where $\eta_j \sim \mathcal{N}(0,1)$ and $\Delta t = 1$.

The Equation (5.14) implies that $p(t_j, Q_j|t_{j-1}, Q_{j-1}; \alpha, \beta) \sim \mathcal{N}(\mu_j, \sigma_j)$ where $\mu_j = Q_{j-1} + \alpha Q_{j-1}\Delta t$ and $\sigma_j = \sqrt{\beta Q_{j-1}\Delta t}$.

The Equation (5.15) implies that $p(t_j, Q_j|t_{j-1}, Q_{j-1}; \alpha, \beta) \sim \mathcal{N}(\mu_j, \sigma_j)$ where $\mu_j = Q_{j-1} + \alpha Q_{j-1}\Delta t$ and $\sigma_j = Q_{j-1}^{\beta}\sqrt{\Delta t}$.

After forming E-M schema, it is necessary to check $Q_j - Q_{j-1}$ series are satisfied with assumptions.

Stationarity

Table 5.5: Result of KPSS and ADF Test

Tests	P-value	Result
KPSS	0.083	0.083 > 0.05
ADF	0.01	0.01 < 0.05

From the Table 5.5, it is obtained that the series are stationary. In the Table 5.5, it is seen that the null hypothesis for KPSS test is not rejected since p-value is greater than specific significance level. The similar result is seen in the ADF test which has reverse hypothesis. As a result, both tests provide that the stationarity of the series $Q_j - Q_{j-1}$.

Normality

Table 5.6: Result of Shapiro-Wilk and Jarque-Bera Test

Tests	P-value	Result
Shapiro-Wilk	0.2796	0.2796 > 0.05
Jarque-Bera	0.8303	0.8303 > 0.05

From the Table 5.6, the hypothesis of normality is not rejected according to both test results. In the Table 5.6, the normality of the series $Q_j - Q_{j-1}$ is not rejected, because p-value is greater than specific significance level. As a result, both tests provide the normality of series $Q_j - Q_{j-1}$.

Independence of Increments

Table 5.7: Result of Box-Pierce Test

Test	P-value	Result
Box-Pierce	0.1966	p-value > 0.05

From the Table 5.7, the hypothesis of auto-correlation coefficients for all lags are zero is not rejected, because p-value is greater than specific significance level.

After estimating parameters using the procedure given in Appendix A, estimates of parameters are put into equations and following equations are obtained.

$$dQ(t) = 0.049Q(t)dt + \sqrt{3752Q(t)}dW(t) \tag{5.16}$$

and

$$dQ(t) = 0.055Q(t)dt + Q^{0.82}(t)dW(t). \tag{5.17}$$

After developing SDE models, it is necessary to check whether the proposed models are fitting to the data or not. In two models, simulations of trajectories between consecutive observations are obtained and these are compared with original data at each point. In the simulation, observation in 1980 is initial value and simulated values are obtained from 1981 to 2014. 5 simulated values are obtained for each point and ranks are obtained according to these. It is expected that as a result of comparison between simulated and original values, 6 situations come in view. In fact, 6 ranks appear for each data point. Expectation of each ranks is equal to 5.67.

In the Figure 5.5, it is seen that there does not exist lack of fit between model and data. This means that ranks obtained by comparing original and simulated values do not deviate from expected ranks. This is visual detection for lack of fit. If we look at the GOF test result, then test statistics is obtained 3.763. The p-value is $P(\chi^2(M-2) > 3.763) = 0.28$ and is greater than $\alpha = 0.05$. According to the result of the test, the hypothesis that there is no lack of fit between model and the data is not rejected due to $p - value > \alpha$.

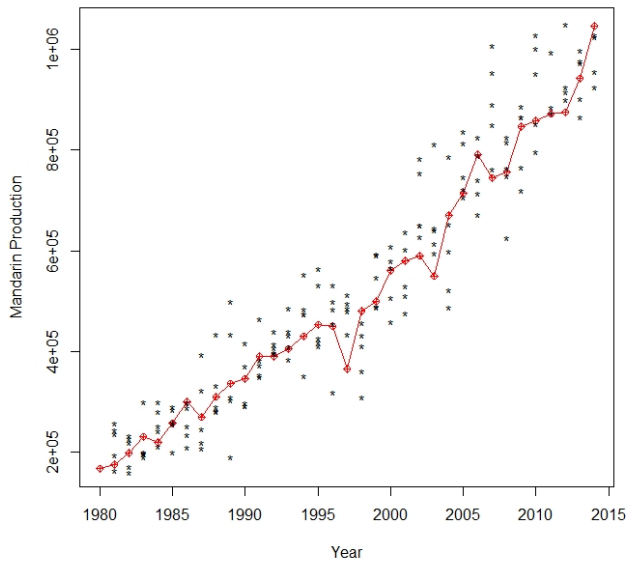

Figure 5.5: Simulated and Original Data for model in the Equation (5.16)

In the Figure 5.6, it is observed that there exists lack of fit between model and data. If we look at the GOF test result, test statistics is obtained as 9.406. The p-value is $P(\chi^2(M-2) > 9.406) = 0.024$ and is less than $\alpha = 0.05$. Therefore, the hypothesis that there is no lack of fit between model and the data is rejected due to $p-value < \alpha$.

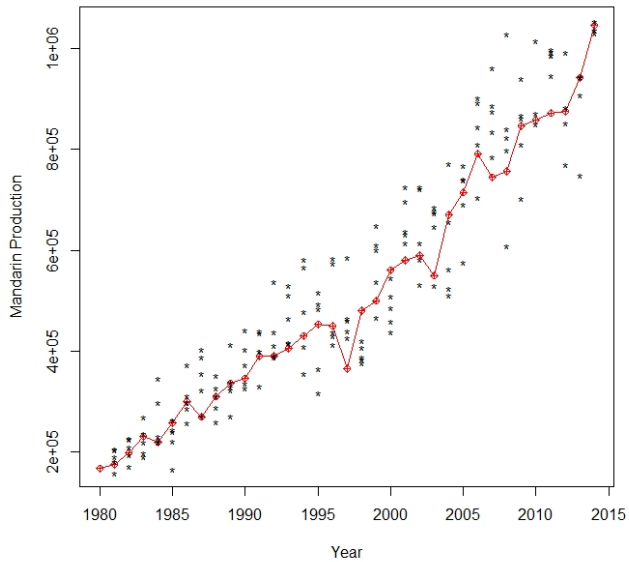

Figure 5.6: Simulated and Original Data for model in the Equation (5.17)

Actually, proposed model in the Equation (5.16) is compatible with data. When we look at the test statistics, it is seen that model in the Equation (5.16) gives smallest test statistics. Obtaining small test statistics means that expected ranks and simulated ranks are close to each other and thus, the null hypothesis of no lack of fit between model and data is not rejected. Model in the Equation (5.17) does not fit to the data. Large test statistics is obtained since expected ranks and simulated ranks are not close to each other. Therefore, the null hypothesis that no lack of fit is rejected. In the Table (5.8), rank values obtained by comparing models are given. In this table, the source for low and high test statistics for models in the Equation (5.16) and the Equation (5.17) may be clearly detected. In fact, ranks of simulations obtained from proposed model in the Equation (5.16) are close to expected ranks, while those from proposed model in the Equation (5.17) are deviated from expected ranks.

Table 5.8: Comparison ranks from proposed models

Ranks	Values for Model1	Values for Model2
$k = 1$	7	7
$k = 2$	6	8
$k = 3$	8	10
$k = 4$	6	5
$k = 5$	5	2
$k = 6$	2	2

After construct model, check assumptions, estimate parameters and check model fit, we forecast future values from 2015 to 2019. In the Table 5.9, we forecast future values by using 100000 simulations from the proposed model. In the Table 5, real values and expected values from the simulations for each year from 2015 to 2019 and their 99% confidence level are given. The proposed model performance for prediction is evaluated by utilising MAPE value. When we calculate MAPE value, it is equal to 17.62% and according to the Table 3.2 in Chapter 3, the proposed model in Equation (5.16) provides good forecast performance.

Table 5.9: Forecast values from 2015 to 2019

Year	Observed values	Forecast	Confidence Level (99 %)
2015	1156365	1098201	(935913.9, 1259962.7)
2016	1337037	1152040	(918365, 1392632)
2017	1550469	1208525	(917613.8, 1519234.5)
2018	1650000	1267483	(924330.8, 1644972.9)
2019	1750000	1329869	(934341.6, 1771296.2)

In the Figure 5.7, we draw the graph of forecast values with their 99% confidence level from the model in the Equation (5.16). From the Figure 5.7, it is said that observed values are close to the upper limit of our forecast and they deviate from forecast values. Deviations of observed values from forecast values are calculated as 17.62% by using the MAPE value and it provides good forecast. In fact, the Table 5.9 and the Figure 5.7 provide good result for this model which is fitted to the data. Also, they are supported by MAPE value.

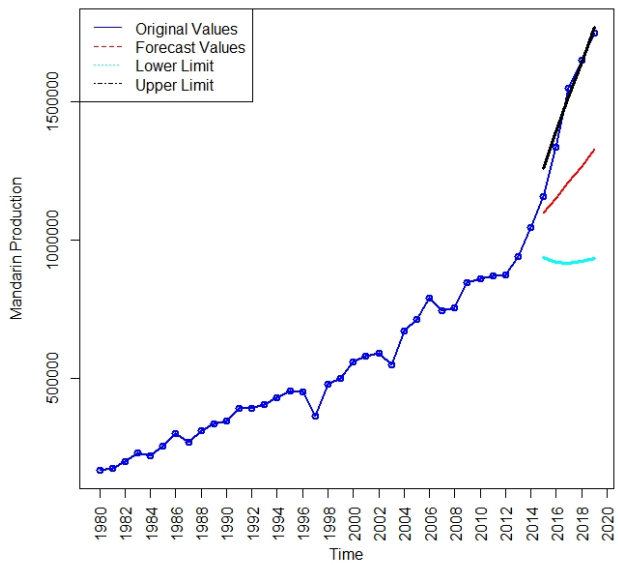

Figure 5.7: Mandarin production original and forecast values

CHAPTER 6

CONCLUSION

We concentrate on proposing two estimation methods for the unknown parameters in SDEs which are driven by Bm. These estimation techniques are MLE and GMM estimation techniques. When assumptions for Bm are satisfied, both techniques give the same estimators. Knowing the distribution is necessary to use MLE technique while it is not necessary to use GMM technique. In our study, we aim to give the basis of these techniques to estimate the unknown parameters in SDEs.

In the first section of this thesis, we introduce the SDE by starting with ODE. In the scope of ODE, the field mice population in the absence of predators is considered, and its increase is modelled proportional to current population. However, in reality, the system such as in the example of field mice population in Chapter 1 may not support our expectations due to some environmental effects. Thus, it is necessary to modify the ODE to represent these bothering effects. In fact, SDEs are used to represent more realistic models. Generally, SDEs are obtained by adding randomness to the coefficients in DEs. Randomness comes from adding noise and it is usually taken as a White noise which is officially regarded as the derivative of a Bm. The source of randomness is not only adding noise to the coefficients but also due to measurement errors [38].

After introducing SDEs, we mention stochastic processes, stochastic integrals and solution of SDEs in the second section. After giving definition of stochastic process, Bm is given with its properties. Due to unbounded variations and non-differentiable

sample paths of Bm, the solution of SDE is not obtained. To deal with this problem, the stochastic integral is given. In this concept, $\int_0^t B_s dB_s$ integral is intuitively solved and $0.5[B_t^2 - t]$ is obtained. Itô Lemma is presented due to the second order effect that is not negligible because of non-differentiability of Bm sample paths. To present how to solve SDEs, we give B-S equation as an example in Chapter 2. We obtain the analytic solution of B-S in the Equation (2.19). On the contrary, many of SDEs' solutions are not obtained explicitly [42]. Thus, we utilise E-M which is one of numeric solution techniques that is common and easy to apply. The Figure 2.3 and the Figure 2.4 show that analytic solution and numeric solution overlap. Also, these figures are supported by the Table 2.1. As we examine the Table 2.1, it is observed that the mean error, mean of difference between analytic and numeric solution at time T, is approaching to zero when the step size Δt decreases.

In the third section, we start with model construction and present how to determine the μ and σ functions in the Equation (3.8). For to show the determination of these functions, we propose to draw the graph of data. In the Figure 3.1, the population data given in [1] is considered to model. In the Equation (3.9), the model constructed to this data is given. Afterwards, we examine the assumption checking for Bm. Stationarity, normality and independence of increments are checked. Then, we propose a method to check whether the proposed model is fitting to the data or not. For this aim, we utilise goodness of fit test which is based on comparing the rank values that are obtained from the simulations and data. In this part, we draw the Figure 3.2 and obtain the test results for the model in the Equation (3.9) given in [1]. According to both visiual detection and test results, it is observed that there is compatibility between the model and data. Also, this test is used to select best model among other models. In other words, best model among others is determined by comparing test values, and the smallest one is the best for the data.

In the fourth section, we present two estimation techniques. For the first technique, it is necessary to know the transition pdf of process for each transition to construct likelihood function. However, many SDEs' solutions are not obtained explicitly, thus,

obtaining transition pdf to construct likelihood function is not easy. We propose to deal with this problem by using E-M method due to supporting approximation when exact solution does not exist. After mathematical explanation, we utilise the Equation (3.9) given in [1] to obtain estimators for α and β by using MLE. We also propose GMM techniques. Unlike MLE technique, GMM technique does not require information of distribution. The only requirement is stationarity of change. This technique is based on specifying moment conditions and is driven by using E-M. Again, we utilise the Equation (3.9) given in [1] to obtain estimators for α and β by using GMM. Both estimators are same when assumptions for Bm are satisfied and the results are given in Chapter 4.

In the application part, two examples are presented. In the first example, parameters are estimated by using MLE and MME techniques. Under the assumptions, it is obtained that estimators found by utilising MLE and MME are same. After model is constructed, the model fit is evaluated by utilising GOF procedure. Test result shows that there is no lack of fit between proposed model in the Equation (5.5) and the data. In the second example, we try to build model for the mandarin quantity in Turkey between 1980 and 2014. Only, MLE methods are used to estimate parameters since all assumptions are satisfied. Test result shows that the proposed model in the Equation (5.16) is compatible with data. Also, we obtain that simulated ranks do not deviate from the expected ranks and thus, small test statistic is obtained. Alternative model is proposed for this data. After estimating parameters and applying GOF procedure, it is obtained that the proposed model in the Equation (5.17) is not compatible with data. We obtain that simulated ranks deviate from the expected ranks and thus, it leads to lack of fit between model and data. After fitting model to the data, we forecast values from 2015 to 2019 to evaluate performance of the model. MAPE value is calculated as 17.62% by utilising forecast values and observed values. It shows that the proposed model in the Equation (5.16) provides good forecast.

As a conclusion, if we model the data with SDE which do not have closed form solutions, we get a numeric solution by using E-M technique. E-M is also useful to

obtain transition density which is necessary for likelihood function and drive moment conditions for GMM. Estimators obtained by MLE and GMM techniques are same under the assumptions for Bm. When only stationary assumption is satisfied, GMM technique may be useful. For future study, other numeric solution techniques, such as Milstein method, Runge Kutta method, etc., my be useful rather than E-M technique. Also, SDEs driven by other stochastic processes may be explored for future analysis.

REFERENCES

[1] ALLEN, E. *Modeling with Itô Stochastic Differential Equations*, vol. 22. Springer Science & Business Media, 2007.

[2] ALLEN, L. J. *An Introduction to Stochastic Processes with Applications to Biology*. Chapman and Hall/CRC, 2010.

[3] BAK, J., MADSEN, H., AND NIELSEN, H. A. Goodness of fit of stochastic differential equations. *Symposium i Anvendt Statistik* (1999), 341–346.

[4] BILLINGSLEY, P. *Probability and Measure*. John Wiley & Sons, 2008.

[5] BOYCE, W. E., DIPRIMA, R. C., AND MEADE, D. B. *Elementary Differential Equations and Boundary Value Problems*, vol. 9. Wiley New York, 1992.

[6] BROWN, R. *A Brief Account of Microscopical Observations Made on the Particles Contained in the Pollen of Plants, and on the General Existence of Active Molecules in Organic and Inorganic Bodies*. 1828.

[7] CARLETTI, M. Numerical solution of stochastic differential problems in the biosciences. *Journal of Computational and Applied Mathematics 185*, 2 (2006), 422–440.

[8] CARRION-I SILVESTRE, J. L., AND SANSÓ, A. A guide to the computation of stationarity tests. *Empirical Economics 31*, 2 (2006), 433.

[9] CHAN, K. C., KAROLYI, G. A., LONGSTAFF, F. A., AND SANDERS, A. B. An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance 47*, 3 (1992), 1209–1227.

[10] CODDINGTON, E. A. *An Introduction to Ordinary Differential Equations*. Courier Corporation, 2012.

[11] DENNIS, B., MUNHOLLAND, P. L., AND SCOTT, J. M. Estimation of growth and extinction parameters for endangered species. *Ecological Monographs 61*, 2 (1991), 115–143.

[12] DICKEY, D. A., AND FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association 74*, 366a (1979), 427–431.

[13] DUFFIE, D., AND SINGLETON, K. J. Simulated moments estimation of markov models of asset prices, 1990.

[14] EINSTEIN, A. *Investigations on the Theory of the Brownian Movement*. Courier Corporation, 1956.

[15] ELERIAN, O., ET AL. A note on the existence of a closed form conditional transition density for the milstein scheme. *Economics Discussion Paper* (1998), W18.

[16] EVANS, L. C. *An Introduction to Stochastic Differential Equations*, vol. 82. American Mathematical Soc., 2012.

[17] FLORENS-ZMIROU, D. Approximate discrete-time schemes for statistics of diffusion processes. *Statistics: A Journal of Theoretical and Applied Statistics 20*, 4 (1989), 547–557.

[18] FLORESCU, I. *Probability and Stochastic Processes*. John Wiley & Sons, 2014.

[19] FRECHTLING, D. *Forecasting Tourism Demand*. Routledge, 2012.

[20] GALLANT, A. R., AND TAUCHEN, G. Which moments to match? *Econometric Theory 12*, 4 (1996), 657–681.

[21] GARD, T. C. *Introduction to Stochastic Differential Equations*. M. Dekker, 1988.

[22] GENTLE, J. E., HÄRDLE, W. K., AND MORI, Y. *Handbook of Computational Statistics: Concepts and Methods*. Springer Science & Business Media, 2012.

[23] HAMILTON, J. D. Time series analysis. *Economic Theory. II, Princeton University Press, USA* (1995), 625–630.

[24] HANSEN, L. P., AND SCHEINKMAN, J. A. Back to the future: Generating moment implications for continuous-time markov processes. *Econometrica: Journal of the Econometric Society* (1995), 767–804.

[25] HARRIS, R. I. Testing for unit roots using the augmented dickey-fuller test: Some issues relating to the size, power and the lag structure of the test. *Economics Letters 38*, 4 (1992), 381–386.

[26] HIGHAM, D. J. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review 43*, 3 (2001), 525–546.

[27] HOGG, R. V., MCKEAN, J., AND CRAIG, A. T. *Introduction to Mathematical Statistics*. Pearson Education, 2005.

[28] HSU, H. P. *Schaum's Outline of Theory and Problems of Probability, Random Variables, and Random Processes*. McGraw-Hill, 1997.

[29] HURN, A. S., LINDSAY, K. A., AND MARTIN, V. L. On the efficacy of simulated maximum likelihood for estimating the parameters of stochastic differential equations. *Journal of Time Series Analysis 24*, 1 (2003), 45–63.

[30] KADILAR, C. *SPSS Uygulamalı Zaman Serileri Analizine Giriş*. 2005.

[31] KENDALL, M. G., AND STUART, A. The advanced theory of statistics (vol. 1,), london: Charles w. *Griffin and Co., Ltd* (1961), 1959–1963.

[32] KLOEDEN, P. E., AND PLATEN, E. *Numerical Solution of Stochastic Differential Equations*, vol. 23. Springer Science & Business Media, 2013.

[33] KWIATKOWSKI, D., PHILLIPS, P. C., SCHMIDT, P., AND SHIN, Y. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics 54*, 1-3 (1992), 159–178.

[34] MAO, X. *Stochastic Differential Equations and Applications*. Elsevier, 2007.

[35] MARATHE, R. R., AND RYAN, S. M. On the validity of the geometric brownian motion assumption. *The Engineering Economist 50*, 2 (2005), 159–192.

[36] MELSA, J. L., AND SAGE, A. P. *An Introduction to Probability and Stochastic Processes*. Courier Corporation, 2013.

[37] MIKOSCH, T. *Elementary Stochastic Calculus, with Finance in View*, vol. 6. World Scientific Publishing Company, 1998.

[38] OKSENDAL, B. *Stochastic Differential Equations: An Introduction with Applications*. Springer Science & Business Media, 2013.

[39] OLSSON, D. M., AND NELSON, L. S. The nelder-mead simplex procedure for function minimization. *Technometrics 17*, 1 (1975), 45–51.

[40] ROSS, S. M. *An Elementary Introduction to Mathematical Finance*. Cambridge University Press, 2011.

[41] SAITO, Y., AND MITSUI, T. Simulation of stochastic differential equations. *Annals of the Institute of Statistical Mathematics 45*, 3 (1993), 419–432.

[42] SAUER, T. Computational solution of stochastic differential equations. *Wiley Interdisciplinary Reviews: Computational Statistics 5*, 5 (2013), 362–371.

[43] SHIRYAEV, A. N. Probability. *Graduate Texts in Mathematics 95,* (1996), 81.

[44] SIMMONS, G. F. *Differential Equations with Applications and Historical Notes*. CRC Press, 2016.

[45] USLU, E. E. Türkiyede seçilmiş bitkisel üretim serilerinin zaman serileri analizi ile modellenmesi ve öngörüsü. Master's thesis, Atatürk University, 2010.

APPENDIX A

APPENDIX A

Here the R codes for given figures, table results, estimation and GOF are presented.
% The following codes generate Figures and and Tables in Chapter 2. When mu=-0.2 and sig=0.3 are given in the below code, Figure 2.4 can be obtained.

```
set.seed(007)
T=1
mu=0.2
sig=1
xzero=1
S=1000
ds=1/S
M=10
da=matrix(0,S,M)
a=matrix(0,S,M)
for(m in 1:M){
da[,m]=sqrt(ds)*rnorm(S)}
for(m in 1:M){
a[,m]=cumsum(da[,i]) }
k=matrix(seq(da,T,da),S,1)
xtrue=matrix(0,S,M)
for(s in 1:S){
for(m in 1:M){
xtrue[s,m]=xzero*exp((mu-0.5*sig^2)*(k[s,1])+sig*w[s,m])}}
```

```
U=10
Da=U*ds
F=S/U
winc=matrix(0,F,M)
for(m in 1:M){
for(l in 1:F){
winc[l,m]=sum(dw[(U*(l-1)+1):(U*l),m])}}
xem=matrix(0,F,M)
for(m in 1:M) {
xem[1,m]=xzero }
for(m in 1:M) {
for(l in 2:F){
xem[m,l]=xem[m-1,l]+Da*mu*xem[m-1,l]+sig*xem[m-1,l]*winc[m-1,l]}}
plot(seq(ds,T,ds),xtrue[,1],type="l",col="red",ylab="sample
path",xlab="time",font.lab=2,font.axis=2,font.sub=1.5,cex.lab=1)
points(seq(ds,T,Da),xem[,1],type="o",lty=5,col="black",ylab="sample path",xlab="time")
legend("topright", legend=c("analytic solution", "numeric solution"),col=c("red", "black"),
lty=1:2, cex=0.8,bg='white')

set.seed(007)
T=1
mu=0.2
sig=1
xzero=1
S=2^13
ds=1/S
M=10
dw=matrix(0,S,M)
w=matrix(0,S,M)
for(m in 1:M) {
dw[,m]=sqrt(ds)*rnorm(S) }
```

```
for(m in 1:M) {
w[,m]=cumsum(dw[,m]) }
k=matrix(seq(ds,T,ds),S,1)
xtrue=matrix(0,S,M)
for(s in 1:S){
for(m in 1:M){
xtrue[s,m]=xzero*exp((mu-0.5*sig^2)*(k[s,1])+sig*w[s,m])}}
U=1
Da=U*ds
F=S/U
winc=matrix(0,F,M)
for(m in 1:M){
for(l in 1:L){
winc[l,m]=sum(dw[(U*(l-1)+1):(U*l),m])}}
xem=matrix(0,F,M)
for(m in 1:M){
xem[1,m]=xzero }
for(m in 1:M){
for(l in 2:L){
xem[l,m]=xem[l-1,m]+Da*mu*xem[m-1,l]+sig*xem[m-1,l]*winc[m-1,l]}}
b=numeric(10)
for(m in 1:10){
b[m]=xtrue[6,m]-xem[6,m]}
abs(mean(b))
```

% The following codes generate Figures and Table in Chapter 5. Also, x contains original data values in the following code.

```
z1=matrix(0,46,8)
z2=matrix(0,46,8)
z3=matrix(0,46,8)
```

```
z4=matrix(0,46,8)
for(m in 1:46){
for(n in 1:8){
z1[m,n]=x[m]+0.0361*x[m]+sqrt(0.529*x[m])*rnorm(1)}}
for(m in 1:46){
for(n in 1:8){
z2[m,n]=z1[m,n]+0.0361*z1[m,n]*1/2+sqrt(0.529*z1[m,n])*sqrt(1/2)*rnorm(1)}}
for(m in 1:46){ for(n in 1:8){
z3[m,n]=z2[m,n]+0.0361*z2[m,n]*1/3+sqrt(0.529*z2[m,n])*sqrt(1/3)*rnorm(1)}}
for(m in 1:46){
for(n in 1:8){
z4[m,n]=z3[m,n]+0.0361*z3[m,n]*1/4+sqrt(0.529*z3[m,n])*sqrt(1/4)*rnorm(1)}}
plot(veriYear, x, lty = 3, pch = 10, col = "red", xlab = "Year", ylab = "Population")
lines(veriYear,x,lty=1,col="red")
points(seq(1940,1985,1), z4[,1], col="black", pch="*")
points(seq(1940,1985,1), z4[,2], col="black", pch="*")
points(seq(1940,1985,1), z4[,3], col="black", pch="*")
points(seq(1940,1985,1), z4[,4], col="black", pch="*")
points(seq(1940,1985,1), z4[,5], col="black", pch="*")
points(seq(1940,1985,1), z4[,6], col="black", pch="*")
points(seq(1940,1985,1), z4[,7], col="black", pch="*")
points(seq(1940,1985,1), z4[,8], col="black", pch="*")
d=matrix(0,46,8)
for(n in 1:46){
for(m in 1:8){
if(x[n+1]>=z4[n,m]) {d[n,m]=1} else {d[n,m]=0}}}
M=numeric(46)
for(m in 1:46){
M[m]=sum(d[m,])}
sort(M)
table(M)
```

```
z1=matrix(0,46,8)
z2=matrix(0,46,8)
z3=matrix(0,46,8)
z4=matrix(0,46,8)
for(m in 1:46){
for(n in 1:8){
z1[m,n]=x[m]+(3510)/(x[m]^2)+13500/(x[m]^2)*rnorm(1)}}
for(m in 1:46){
for(n in 1:8){
z2[m,n]=z1[m,n]+(3510/(z1[m,n]^2))*1/2+(13500/(z1[m,n]^2))*sqrt(1/2)*rnorm(1)}}
for(m in 1:46){
for(n in 1:8){
z3[m,n]=z2[m,n]+(3510/(z2[m,n]^2))*1/3+(13500/(z2[m,n]^2))*sqrt(1/3)*rnorm(1) }}
for(m in 1:46){
for(n in 1:8){
z4[m,n]=z3[m,n]+(3510/(z3[m,n]^2))*1/4+(13500/(z3[m,n]^2))*sqrt(1/4)*rnorm(1)}}
plot(veriYear, x, lty = 3, pch = 10, col = "red", xlab = "Year", ylab = "Population")
lines(veriYear,x,lty=1,col="red")
points(seq(1940,1985,1), z4[,1], col="black", pch="*")
points(seq(1940,1985,1), z4[,2], col="black", pch="*")
points(seq(1940,1985,1), z4[,3], col="black", pch="*")
points(seq(1940,1985,1), z4[,4], col="black", pch="*")
points(seq(1940,1985,1), z4[,5], col="black", pch="*")
points(seq(1940,1985,1), z4[,6], col="black", pch="*")
points(seq(1940,1985,1), z4[,7], col="black", pch="*")
points(seq(1940,1985,1), z4[,8], col="black", pch="*")
```

% The following codes are used to check assumption, parameter estimation, GOF procedure, MAPE value calculation in Chapter 5. Before using the codes, we install necessary packages in R.

```
library(Sim.DiffProc)
library(tseries)
% x contains mandarin data from year 1980 to 2014 in the following code.

theta1=1
theta2=1
xs=ts(x,1980,2014)
fx <- expression((theta[1])*(xs))
gx <- expression(sqrt(theta[2]*(xs)))
fitmod <- fitsde(data = xs, drift = fx, diffusion = gx, start = list(theta1=1, theta2=1),pmle="euler")
summary(fitmod)
z1=matrix(0,34,5)
z2=matrix(0,34,5)
z3=matrix(0,34,5)
z4=matrix(0,34,5)
for(j in 1:34){
for(i in 1:5){
z1[j,i]=x[j]+0.049*(x[j])+sqrt(x[j]*3752)*rnorm(1)}}
for(j in 1:34){
for(i in 1:5){
z2[j,i]=z1[j,i]+0.049*(z1[j,i])*1/2+sqrt(z1[j,i]*3752)*sqrt(1/2)*rnorm(1)}}
for(j in 1:34){
for(i in 1:5){
z3[j,i]=z2[j,i]+0.049*(z2[j,i])*1/3+sqrt(z2[j,i]*3752)*sqrt(1/3)*rnorm(1)}}
for(j in 1:34){
for(i in 1:5){
z4[j,i]=z3[j,i]+0.049*(z3[j,i])*1/4+sqrt(z3[j,i]*3752)*sqrt(1/4)*rnorm(1)}}
d=matrix(0,34,5)
for(i in 1:34){
for(j in 1:5){
if(x[i+1]>=z4[i,j]) {d[i,j]=1} else {d[i,j]=0}}}
```

```
M=numeric(34)
for(j in 1:34){
M[j]=sum(d[j,])}
sort(M)
table(M)
r=c(rep(5.67,6))
u=matrix(c(table(M),r),6,2)
k=numeric(6)
for(i in 1:6){k[i]=(((u[i,1]-u[i,2])^2)/5.67)}
1-pchisq(sum(k),3)
sum(k)
plot(seq(1980,2014,1),x,lty=3,pch=10,col="red",xlab="Year",ylab="Mandarin Produc-
tion")
lines(seq(1980,2014,1),x,lty=1,col="red")
points(seq(1981,2014,1), z4[,1], col="black", pch="*")
points(seq(1981,2014,1), z4[,2], col="black", pch="*")
points(seq(1981,2014,1), z4[,3], col="black", pch="*")
points(seq(1981,2014,1), z4[,4], col="black", pch="*")
points(seq(1981,2014,1), z4[,5], col="black", pch="*")
```

% The following codes are utilised for alternative model for mandarin quantity.

```
fx <- expression((theta[1])*(xs))
gx <- expression(((xs^theta[2])))
fitmod <- fitsde(data = xs, drift = fx, diffusion = gx, start = list(theta1=1, theta2=1),pmle="euler")
summary(fitmod)
z1=matrix(0,34,5)
z2=matrix(0,34,5)
z3=matrix(0,34,5)
z4=matrix(0,34,5)
for(j in 1:34){
```

```
for(i in 1:5){
z1[j,i]=x[j]+0.055*(x[j])+(x[j] ^0.82)*rnorm(1)}}
for(j in 1:34){
for(i in 1:5){
z2[j,i]=z1[j,i]+0.055*(z1[j,i])*1/2+(z1[j,i]^0.82)*sqrt(1/2)*rnorm(1)}}
for(j in 1:34){
for(i in 1:5){ z3[j,i]=z2[j,i]+0.055*(z2[j,i])*1/3+(z2[j,i]^0.82)*sqrt(1/3)*rnorm(1)}}
for(j in 1:34){
for(i in 1:5){
z4[j,i]=z3[j,i]+0.055*(z3[j,i])*1/4+(z3[j,i]^0.82)*sqrt(1/4)*rnorm(1)}}
d=matrix(0,34,5)
for(i in 1:34){
for(j in 1:5){
if(x[i+1]>=z4[i,j]) {d[i,j]=1} else {d[i,j]=0}}}
M=numeric(34)
for(j in 1:34){
M[j]=sum(d[j,])}
sort(M)
table(M)
r=c(rep(5.67,6))
u=matrix(c(table(M),r),6,2)
k=numeric(6)
for(i in 1:6){k[i]=(((u[i,1]-u[i,2])^2)/5.67)}
1-pchisq(sum(k),3)
sum(k)
plot(seq(1980,2014,1),x,lty=3,pch=10,col="red",xlab="Year",ylab="Mandarin Produc-
tion")
lines(seq(1980,2014,1),x,lty=1,col="red")
points(seq(1981,2014,1), z4[,1], col="black", pch="*")
points(seq(1981,2014,1), z4[,2], col="black", pch="*")
points(seq(1981,2014,1), z4[,3], col="black", pch="*")
```

```
points(seq(1981,2014,1), z4[,4], col="black", pch="*")
points(seq(1981,2014,1), z4[,5], col="black", pch="*")
x1=c(1156365,1337037,1550469,1650000,1750000)
z101=matrix(0,6,100000)
for (i in 1:100000){z101[1,i]=x[35]}
for(j in 1:5){
for(i in 1:100000){
z101[j+1,i]=z101[j,i]+0.049*(z101[j,i])+sqrt(z101[j,i]*3752)*rnorm(1)}}
f=apply(z101,1,mean)[-1]
100*(sum(abs(x1-f)/x1)/5)
apply(z101,1,bconfint,level=0.95)
```