Republic of Iraq
Ministry of Higher Education
and Scientific Rescarch
University of Al-Qadisiyah
College of Administration and Economics
Statistics Department

Some nonparametric methods to estimate the heavy-tailed index in Stochastic Differential Equations for Iraq Stock Exchange

A thesis submitted to the council of the college of Administration and Economics at University of Al-Qadisiyah as partial fulfillment of the requirements for the degree of master in Statistics

By

Noor Abd Hassan

Supervised by

Assist. Prof. Dr. Muhannad F. Al-Saadony

1443 A.H **2021 A.D.**

بسم الله الرحمن الرحيم

نَرْفَعُ دَرَجَاتٍ مَنْ نَشَاءُ وَفَوْقَ كُلِّ

ذِي عِلْمٍ عَلِيمٌ

صدق الله العلي العظيم

سورة يوسف المباركة (الآية 76)

Dedication

To the one who left me with her body, but her soul still flutters in the sky of my life…… my mother.

To my support and my destiny…… my father.

To my sister and my brothers.

To the one who took my hand on the path to success …….my husband.

To the light of my eyes and the joy of my heart my children Ali, Maryam and Lian.

To all my relatives and friends.

Acknowledge

The Most Compassionate all praise be to Allah the Lord of the worlds.

I would like to express my gratitude to my supervisor Assist. Prof. Dr. Muhannad. F. Al-Saadony for his supervision of this work and showed me the right ways to get my destination, through the useful comments and remarks. I am really thankful to you.

My thanks go also to Dr. Muhammed Habib, Dr. Tahir Reisan, Dr. Rahim Alhamzawi, Dr Ali Alkenani, Dr. Ahmed Naeem, Dr. Hassan Sami, Dr. Taha Alshaybawe and Dr. Fadil Alhusseini for their contribution in my teaching.

My sincere thanks to my professors in primary studies Dr. Abbas Laftah, Basim Muslim, Khader Al Bayati and everyone who taught me a letter since the beginning of my school life.

Finally my parents. Thank you Dad for always believing in me. you made me what I am today. Mom, I am forever grateful. My deepest gratitude to my sister and brothers.

Abstract

Heavy-tailed distributions are very important branch of statistical analysis. The heavy-tailed occurs because there are some extreme values in the distributions. In this thesis, we estimated the tail parameter α-stable with $(0 \leq \alpha \leq 2)$ for a set of independent and identical distributed observations using three non-parametric methods (the direct, the Bootstrap and the Double Bootstrap). The methods were compared in order to choose the best among them, which represents the smallest average mean square error. We mentioned the Geometric Brownian motion and Levy process as two famous examples of Stochastic Differential Equations that have been used to generate data in simulations. The tail parameter estimation methods were applied in the simulation and real data for the daily data set of the Iraq Stock Exchange using the R program.

Contents

Chapter one

Chapter two

Chapter Three

Chapter four

Chapter five

List of symbols and abbreviations

(Ω, F, P)	Probability space
Ω	Sample space
F	σ - field
$(S, \sum)$	Measurable space
X(t)	Stochastic process
W_t	Brownian motion
μ	Drift parameter
σ	Volatility parameter
S_t	Spot price
S_0	Initial value
G.B.M.	Geometric Brownian motion
S.D.E	Stochastic Differential Equation
L	Levy process
L_{Ito}	Levy- Ito
L^x	Levy filtration
θ	Scale parameter
i.i.d.	identically independent distribution

Ψ	exponent
(a, σ^2, v)	a triple Levy process
c.d.f.	cumulative distribution function
E.V.T.	Extreme Value Theory
α	Tail index or Tail parameter
$\hat{\alpha}$	Hill estimator
S	The thresholds
$avH_{k,n}$	Smoothing Hill
r.v.	random variable
$M_X(t)$	Moment generating function
$P_{mL}(x)$	mild - left heavy - tailed
$P_{mR}(x)$	mild – right heavy – tailed
$P_{m2}(x)$	mild – two heavy - tailed
$P_{eL}(x)$	extremely heavy left – tailed
$P_{eR}(x)$	extremely heavy right – tailed
$P_{e2}(x)$	extremely – two heavy – tailed
IQR	Inter – Quartile Range
ρ, β	Second order regular variation parameters
$\hat{\alpha}_{(k)}$, $\hat{\rho}_{(k)}$ and $\hat{\beta}_{\tau}(k)$	Consistent estimators for α, ρ, β respectively

ISX	Iraq Stock Exchange
S.P	Stock Prices
$\bar{F}$	tail distribution function
E.p.w	Exponent power weibull
AMSE	Average Mean Square Error

Chapter one

1.1. Introduction

Estimation of the tail parameter plays a significant role in heavy – tailed distributions. Assuming that $x_{1,}\ x_{2,}\ldots,x_n$ be i.i.d. random variables with a function of a real variable whose behavior at infinity is similar to the behavior of a power law function called a regularly varying tail **(Danielsson et al, 2001)** :

$$1 - F(X) = X^{-\alpha}\, L(X) \qquad X \to \infty\ , \quad \textbf{(1-1)}$$

where,

$F(X)$: the distribution function

α: the tail index

$L(X)$: slowly varying function (it is a function of a real variable whose behavior at infinity is somewhat similar to that of a function converging at infinity).

This is if (F) is in field of attraction of a stable distribution with parameter $(0 < \alpha < 2)$ or of extreme – value distribution with positive index. Many researchers have estimated the tail index in various ways such as (Hill, 1975); (Hall, 1982); (Hall & welsh, 1985) and (Davis & Resnick, 1984). We focus on a commonly used estimator which is the Hill estimator. (which will be covered in more detail in chapter 2). The thesis problem is how the data behave in the case of the presence of the heavy tail and the extent of the impact of the heavy tail on the data, so we estimated the tail index in the Stochastic Differential Equations (SDE), where we used two models, namely the Geometric Brownian motion (G.B.M.) model, which has a log-normal distribution, and the levy process. We used three methods to estimate the tail parameter (the Direct; Bootstrap and Double Bootstrap). These methods mainly depended on the Hill estimator. In order to show which method is best than the others, we used the mean square error of the difference between the tail parameter that was estimated by one of the methods and Hill estimator as follows:

$$MSE = E[\hat{\alpha}(K) - \alpha(K)]^2, \qquad \textbf{(1-2)}$$

where:

$\hat{\alpha}(K)$: the tail index estimated by one of the methods.

$\alpha(K)$: the tail index estimated by Hill estimator.

1.2. The Aim of thesis

Estimation of the tail parameter in two models of Stochastic Differential Equations, namely the Geometric Brownian Motion G.B.M and Levy model using three non-parametric methods (the Direct, Bootstrap and Double Bootstrap) and comparison among methods by the Average Mean Square Error (MSE).

1.3. Literature Review

In this section, we present the studies conducted by researchers that are related to the topic of our research.

Hall & Welsh (1985) explained that the problem of estimating the shape and scale parameters of a regularly varying tails distribution is related to the problem of non-parametric estimation of density at a fixed point. They explained how to overcome this problem using adaptive method. By assuming a Hill estimator, they give the optimum number with the largest order stats as function of some of the parameters for unknown distribution of the function. In addition, a consistent estimator K_n^{opt} is suggested if 2^{nd} order parameter ρ of F is known.

Hall (1990) describes bootstrap method for estimating the smoothing parameter and the mean square error of nonparametric problems. This method involved the use of substitution for a volume smaller than the original sample. Several applications have been used, including nonparametric regression, nonparametric density estimation and tail parameter estimation.

Dress & Kaufmann (1998) presented a sequential procedure, which has yielded a consistent estimator of K_n^{opt} in complete model with no need for prior information on the second order parameter ρ. The result of an adaptive Hill estimator has been shown to be asymptotically effective according to optimum number of the order statistics.

Danielsson et. al. (2001) showed that the estimate of the tail parameter depends on its accuracy in selecting the sample fraction. They provided a full solution for selecting the fraction of the sample by two-step sub-sample smoothing , and explained that this method determines the sample fraction that minimizes the asymptotic mean square error without the need for prior knowledge of ρ and also no necessity for a prior estimation of the tail index.

Gomes et.al. (2002) presented a class of semi-parametric estimators for ρ with a regularly varying tail and showed that 2^{nd} order parameter has a very significant impact when dealing with the problems of the optimization in the statistics of the extreme values. The asymptotic normality and consistency are proven under suitable conditions.

Beirlant & Goegebeur (2004) estimated the tail index when data of independent sets were available. The proposed methods depended on regression models that link statistics related to the tail of basic distribution function to the index of the extreme value and parameters describing the behavior of the tail. The optimal number of extremes for using in the estimate was derived from the asymptotic mean squared error matrix. The simulation results showed that combining data from several groups greatly improves estimation of extreme value index.

Gardes & Girard (2008) provided nonparametric estimators of the tail index for the Pareto distribution when the covariate information was available. The estimations were based upon the weighted sum of the log-spacing between the chosen observations, as this selection is accomplished by a random threshold on variable of interest and the moving window approach on covariate domain. The researchers used real data for a limited sample, and the Asymptotic normality was demonstrated under mild regularity conditions.

Ciuperca & Mercadier (2010) have generalized many studies on the extreme value theory to estimate 2nd order parameter and extreme value index. By performing some numerical calculations and asymptotic normality and consistency are proven under classical assumptions.

Baek & Pipiras (2010) estimated the parameters in the heavy-tailed distributions within the 2nd order regular variation framework when ρ of tail was known. They clarified that 2nd order tail parameter is known as a large class of common random difference equations (such as ARCH models). The focus was on the least square estimates which generalized QQ estimates and rank-based estimates. The results of the Monte Carlo simulation showed that the least square estimators are better in performance and easier to use for finite samples.

Gomes et. al. (2012) discussed an algorithm for adaptive estimation of the positive extreme value index (α) which is the main parameter of extreme stats. They proposed to consider associated 2nd order corrected bias estimators as well as utilization of the re-sampling based computer intensive methods to choose a consistent asymptote to the thresholds for utilization in the adaptive estimation of α.

Jia (2014) explained that the tail index can be inferred using new methods based on the whole sample, not just the tail. The researcher proposed two methods: the first is a regression technique that depends on the characteristic function, as for the second method, the researcher used the scaling function (graphical method). The researcher applied these two methods to discover the laws of force in experimental data sets and distinguish them from the log-normal distributions. The regression method was the best in most cases.

Hashemifard et.al. (2016) focused on heavy-tailed stochastic signals generated through continuous time auto-regressive models evoked through the infinite-variance α-stable processes with $(0 < \alpha < 2)$. Their aim was estimating the continuous time model parameters. The consistency of the estimator of desired values is illustrated in the case where the sample size and sampling frequency approach infinity. The suggested method was applied to two real data types, and the experimental results showed good agreement between a model and this data.

Danielsson et.al. (2019) mentioned how difficult it is to choose upper order statistics (u.o.s.) in tail estimation and showed that most of the methods depend on minimizing the asymptotic mean square error, which does not work well in finite samples. Therefore, they introduce a data driven method that reduces the maximum distance between the fitted pareto distribution tail and the observed quantity. A comparison has been made between the methods based on the finite sample and other methods, and the first has proven its efficiency in estimation.

Kallgren (2019) used the numerical methods to simulate a stochastic differential equation showing stability due to variance. The researcher stated that this property of the processes means that the stable behavior is imposed through the volatility of the process and that this property takes a long time in simulation and in statistical methods. The researcher used simulation of time change and Euler scheme as statistical methods for estimating the parameters in the models.

Nemeth (2020) showed that improving the performance of the Hill estimator using the Bootstrap or the Kolmogrov-Smirnov method can fail the estimates if the tail index $\alpha > 0.5$. The researcher presented a new experimental methods that combine the advantages of Kolmogrov-Smirnov and Bootstrap and showed that the estimators have the ability to estimate well the parameters of the large tail index and also the small sample size.

Ahmad et.al. (2020) proposed 9 new methods for defining new distributions suitable for modeling heavy right-tailed data. A special three-parameter model that is referred to as the Exponent Power Weibull (E.P.W) distribution were studied. A simulation work was conducted to illustrate the importance of the suggested method. For proving the importance of (E.P.W) distribution, the researchers took two insurance datasets and compared their suitability for other distributions. The suggested model outperformed competitive models.

Our work in this thesis is to estimate the tail parameter in Stochastic Differential Equations using three non-parametric methods.

Chapter two

The theoretical part

2.1.Introduction

In this chapter, we will explain some important features such as Brownian motion, stochastic integration and the Ito formula. Then, we study two models of stochastic differential equations : the Geometric Brownian motion and the Levy process, where some important definitions and properties are mentioned for each of them, and then we have touched on the topic of extreme values and their importance in our thesis as being the main reason for the existence of heavy tails, which is the focus of our thesis, where we studied in detail the heavy tails with the mention of the important definitions, and then we have estimated the tail parameter in three ways : the direct, the Hall Bootstrap and the Double Bootstrap methods.

2.2.Stochastic processes

A stochastic process can be defined as a set of random variables that can be represented by time series. It is used as a mathematical model for phenomena that seem to vary randomly. It is applied in many fields such as Physics, Chemistry, Biology, Communications, Computer, Finance,…, so on.

The term random function refers to the stochastic processes as seen in **(Shorokhod, 2005)** and **(Gusak et.al., 2010)** because the stochastic processes can be interpreted as a random element in the function space.

Definition:2.2.1

If we have a probability space$(\Omega\,, F, p)$ where Ω represent the sample space, F represent the σ- algebra, p represents the measure of probability and a measurable space $(S, \sum)$, the stochastic process is a set of S of r.vs written in the following form **(Florescu, 2014)**

$$\{\boldsymbol{X}(\boldsymbol{t}), \boldsymbol{t} \in \boldsymbol{T}\},$$

where:

X(t) is a random variable representing a value observed at time t **(Borovkov, 2013)**. The stochastic process can also be written as follows **(Lindgren et.al., 2013)**:-

$$\{\boldsymbol{X}(\boldsymbol{t}, \boldsymbol{w}), \boldsymbol{where}\ \boldsymbol{t} \in \boldsymbol{T}\ \boldsymbol{and}\ \boldsymbol{w} \in \Omega\}$$

A common example of stochastic processes is the **Brownian motion** or **Wiener process**.

2.3.Brownian motion

This process is discovered in 1827 by **Robert Brown** and then developed by Albert Einstein in 1905. It is sometimes called the Wiener process, due to the scientist Wiener who studied Brownian motion more deeply **(Iacus, 2009 and 2011)**.

Definition:2.3.1

Brownian motion (W_t) is a stochastic process in continuous time $\{W_t, t \geq 0\}$ where t is a positive real number. There are some properties of the Brownian motion, which are as follows **(Franke et.al., 2004)**:-

1- $W_t = 0$ when t=0

2- W_t is continuous in the time t , $t \geq 0$ and W_t is distributed normally with mean zero and variance t.

$$W_t \sim N(0, t)$$

3- W_t has independent increments, this means for $0 \leq s < t$, $W_t - W_s$ is independent of W_s.

4- for $0 \leq s < t$ the increment $(W_t - W_s)$ is distributed normally with mean zero and variance (t-s).

$$W_t - W_s \sim N(0, t - s)$$

This means the distribution of $(W_t - W_s)$ depends on the length (t-s) of the time interval.

2.4.Stochastic integral

Stochastic integration is a branch of mathematics used in modeling stochastic systems. Stochastic integrals, in particular the Ito integrals, are used as a solution to stochastic differential equations. Among the most important stochastic processes to which stochastic calculus is applied is the Wiener process, which used to model Brownian motion **(Iacus, 2009)**:-

$$\boldsymbol{dS_t = \mu S_t dt + \sigma s(t) dw_t} \qquad \textbf{(2-1)}$$

Where:

S_t : is a spot price of underlying asset in time t.

μ: is a drift parameter.

σ: is a volatility parameter.

w_t: Brownian motion.

Equation (2-1) represents the model of the Geometric Brownian Motion (G.B.M.), which can be written in an integrated manner as follows:

$$\boldsymbol{s(t) = s(0) + \mu \int_0^t s(u)du + \sigma \int_0^t s(u)dw(u)} \qquad \textbf{(2-2)}$$

Equation (2-2) introduces the stochastic integral **(Iacus, 2008)**:-

$$\boldsymbol{I(X) = \int_0^t X(u)dw(u)}$$

2.5. $\boldsymbol{It\hat{o}}$ formula

It is one of the very significant tools in the stochastic calculus. This formula is the stochastic version of the Taylor expansion $f(Y)$ which stops at 2nd order where (Y) represents a process of diffusion. The $It\hat{o}$ process means that if $f(t, y)$represents a twice differentiable function for t and y. So, the $It\hat{o}$ process is a special case of Taylor expansion **(Iacus, 2009)** **,** thus:-

$$\boldsymbol{f(t, Y_t) =}$$

$$\boldsymbol{f(0, Y_0) + \int_0^t f_t\,(u, Y_u)du + \int_0^t f_y\,(u, Y_u)dY_u + \frac{1}{2}\int_0^t f_{yy}\,(u, Y_u)(dY_u)^2} \quad , \quad \textbf{(2-3)}$$

where:-

$$f_y(t, y) = \frac{\partial f(t, y)}{\partial y}$$

$$f_t(t, y) = \frac{\partial f(t, y)}{\partial t}$$

$$f_{yy}(t,y) = \frac{\partial^2 f(t,y)}{\partial y^2},$$

or

$$df(t,Y_t) = f_t(t,Y_t)dt + f_y(t,Y_t)dY_t + \frac{1}{2}f_{yy}(t,Y_t)(dY_t)^2$$

If (Y_t) is a Brownian motion, this simplifies the following

$$f(t,W_t) = f(0,0) + \int_0^t \left(f_t(u,W_u) + \frac{1}{2}f_{yy}(u,W_u)\right)du + \int_0^t f_y\,(u,W_u)dW_u$$

Or

$$df(t,W_t) = \left(f_t(t,W_t) + \frac{1}{2}f_{yy}(t,W_t)\right)dt + f_y(t,W_t)dW_t$$

There are some features for the stochastic integral and $It\hat{o}$ process which are as follows:-

1-If (Y) is $It\hat{o}$ integrable, so

$$E\left[\int_o^T Y(s)dW(s)\right] = 0$$

And

$Var\left[\int_0^T Y(s)dW(s)\right] = \int_0^T EY^2\,(t)dt \;\ldots.$ (($It\hat{o}$ isometry))

2-If y and x are two $It\hat{o}$ integrable operations and d and c are two constants, then linear function is :-

$$\int_0^T \big(cX(t) + dy(t)\big)dW(t) = d\int_0^T Y(t)dW(t) + c\int_0^T X(t)dW(t)$$

3-It follows from the linearity property (mentioned in the second point) :

$\int_0^T cdW(t) = c\int_0^T dW(t) = cW(T).$

4- $\int_0^T W(t)\,dW(t) = \frac{1}{2}W^2(T) - \frac{1}{2}T$

5- $Z(t) = Z(0) + \int_0^t X(s)\,dW(s)$ is martingale process with Z(0) a constant.

$It\hat{o}$ process $[Y_t, 0 \leq t \leq T]$ represents stochastic process which can be expressed as:-

$$\boldsymbol{Y_t = Y_0 + \int_0^t g(s)ds + \int_0^t h(s)\, dWs} \quad , \qquad \textbf{(2-4)}$$

where:-

$g(t,w)$ and $h(t,w)$ are two adaptive, progressively measurable, random functions:-

$$P\left[\int_0^T |g(t,w)|\, dt < \infty\right] = 1\,,$$

and

$$P\left[\int_0^T h(t,w)^2 dt < \infty\right] = 1$$

2.6. Stochastic Differential Equations (SDEs)

SDEs are differential equations with a stochastic term and they have the stochastic solution. SDEs are utilized for modeling many different phenomena like the physical systems and unstable stock prices. SDEs contains a variable which is a random white noise that is calculated as one of the derivatives of the Wiener process or Brownian motion.

It shall be noted that the direct application of $It\hat{o}$ lemma can be helped to find the solution of SDEs **(Iacus, 2009) and (Imkeller & Schmalfuss, 2001).**

The general formula of (SDE) is **(Franke et.al., 2004):-**

$$\boldsymbol{dx_t = \mu dt + \sigma dw_t} \quad , \qquad \textbf{(2-5)}$$

where:-

dx_t: change of x_t in a continuous time t.

μ: drift parameter.

σ: volatility parameter.

dw_t: change of standard Brownian motion.

A standard Brownian motion is a continuous space and continuous time stochastic process that describes the process of evolution of the value of any random variable. It is sometimes called the Wiener process **(Iacus, 2011).**

We will present the Geometric Brownian Motion and Levy process as a popular examples of Stochastic Differential Equations .

2.6.1.Geometric Brownian Motion

The **(G.B.M.)**, also called the Exponential Brownian Motion (E.B.M.) is a continuous- space and continuous-time stochastic process in which the logarithm of a randomly varying quantity which follows a Brownian motion with drift **(Ross, 2014).**

As a simple model of market prices, many economists prefer the Geometric Brownian motion because it is positive everywhere (with probability 1) **(Dunbar, 2016).**

The **G.B.M.** is an important example of stochastic differential equations as it is used to model stock prices in mathematical finance which is called **Black and Scholes (Mikosch, 2004) and (Iacus, 2009)** model by **Fisher Black and Myron Scholes (Fisher and Scholes, 1973)**. It estimates the price (s) over time (t). It is a contract to sell or buy an underlying asset at a specified price at time t **(Mikosch, 1998).**

2.6.1.1.The use of Geometric Brownian Motion in finance

The main reasons for using Geometric Brownian Motion to model stock prices are as follows **(Hull, 2009)**:-

1-The returns that are expected from (GBM) are not dependent on stock price and agrees with what we actually expect.

2-This process assumes only positive values which are the same as real s. p..

3-The GBM process exhibits the same type of coarseness in its paths as in the real stock prices.

4-Ease of computations using this process.

Yet, the geometric Brownian motion is not an entirely realistic model as it falls short of the reality in what follows:-

1-Volatility in the changes of the real stock prices throughout time while volatility is assumed to be constant in Geometric Brownian motion.

2-In fact, stock prices often show jumps due to unpredictable events while the path is continuous in GBM.

Definition: 2.6.1.1

A stochastic process S_t is said to follow a Geometric Brownian Motion if it satisfies the following stochastic differential equation **(Franke et.al., 2004) :-**

$$dS_t = \mu S_t d_t + \sigma S_t dw_t \qquad \textbf{(2-6)}$$

It is easier way to work with returns : $y_t = \log \frac{s_t}{s_0}$,

where:

S_0 is an initial value.

Using the $It\hat{o}$ lemma, we can transform equation (2-6) into an equation for y_t .

$$\partial y_t = \left[\frac{\partial y_t}{\partial s_t}\,\mu\, S_t + \frac{1}{2}\,\sigma^2\, S_t^2\,\frac{\partial^2 y_t}{\partial s_t^2}\right] dt + \frac{\partial y_t}{\partial s_t}\,\sigma\, S_t\, dw_t \quad , \qquad (2\text{-}7)$$

where:

$$\frac{\partial y_t}{\partial S_t} = \frac{1}{S_t}$$

$$\frac{\partial^2 y_t}{\partial S_t^2} = -\frac{1}{S_t^2}$$

and

$$\frac{\partial y_t}{\partial t} = 0$$

this leads to :-

$$\partial y_t = \left[\frac{1}{S_t}\,\mu\, S_t + \frac{1}{2}\,\sigma^2\, S_t^2\,(-\frac{1}{S_t^2})\right] dt + \frac{1}{S_t}\,\sigma\, S_t dw_t$$

$$= \left(\mu - \frac{1}{2}\,\sigma^2\right) dt + \sigma\, dw_t \qquad (2\text{-}8)$$

Now, to convert G.B.M. from continuous to discrete time we use Euler scheme **(Iacus, 2009)**.

The goal of applying the Euler scheme is to make it easier to deal with discrete time rather than continuous time.

The Euler scheme of y_t is **(AL-Saadony[b], 2016)** :-

$$y_{t+\Delta} = y_t + \int_t^{t+\Delta}\left(\mu - \frac{1}{2}\,\sigma^2\right) ds + \int_t^{t+\Delta} \sigma\, dw_t \qquad (2\text{-}9)$$

By integrated equation (2-9), we now let $\Delta \to 0$.

When converting from continuous to discrete time, the positivity of the volatility parameter becomes unsecured, so we replace σ with $|\sigma|$**(AL-Saadony[b], 2016)**.

$$\boldsymbol{y_{t+\Delta} = y_t + \left(\mu - \frac{1}{2}\,\sigma^2\right)\Delta + |\sigma|\,\sqrt{\Delta}\,z_t\,,\ \ z_t \sim N(0,1)} \qquad \textbf{(2-10)}$$

2.6.1.2.Derivation of the (SDE) for the G.B.M.

We can drive the stochastic differential equation (SDE) for the geometric Brownian motion **(Iacus, 2009),**

$$\boldsymbol{S_t = S_0\, exp\left[\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma w_t\right], t} > 0 \qquad \textbf{(2-11)}$$

By choosing $f(t,x) = S_0\ exp\ \left[\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma_x\right]$

Thus,

$$f(t, w_t) = S_t$$

and

$$f_t(t,x) = \left(\mu - \frac{\sigma^2}{2}\right) f(t,x)$$

$$f_x(t,x) = \sigma f(t,x)$$

$$f_{xx}(t,x) = \sigma^2 f(t,x).$$

Hence:-

$$dS_t = df(t, w_t)$$

$$= \left(f_t(t,w_t) + \frac{1}{2}\, f_{xx}(t,w_t)\right) d_t + f_x(t,w_t) dw(t) \qquad \textbf{(2-12)}$$

$$= \left(\left(\mu - \frac{\sigma^2}{2}\right) S_t + \frac{1}{2}\,\sigma^2 S_t\right)\, dt + \sigma S_t dw_t$$

$$= \mu\, S_t\, dt + \sigma\, S_t dw_t$$

It should be noted that the G.B.M has a log-normal distribution with drift (μ)and volatility(σ).

2.6.2. Lévy process

Lévy process (L) can be defined as a stochastic process with stationary and independent increments. It was introduced by French mathematician **Paul Lévy** in 1950 **(Applebaum, 2009), (Kessler et.al., 2012) and (Klebaner, 2012).**

The main idea of the levy process is to work with a small or big jump in continuous time stochastic process. The jump of levy process $\Delta L_t = L_t - L_{t-}$ is very important to understand the behavior of these process. There are many example of the Levy process such as the Brownian motion process, the poisson process, stable process, Inverse Gaussian process ….. so on **(Kyprianou, 2014).** Al levy processes are additive processes .

Definition: 2.6.2.1

The additive process is generalization of the obtained Lévy process to mitigate the hypothesis of similarly distributed increases. A stochastic process is an additive process if it fulfills the following conditions **(Carr et.al., 2007):-**

1-Independent increments //

A stochastic process $X_t, t \geq 0$ has independent increments if:

For any $0 \leq p < r < s < t$ the r.v. $X_t - X_s$ is independent from the r.v. $X_r - X_p$

2-Continuity in probability//

A stochastic process $X_t, t \geq 0$ is continuous in probability if:

$\lim_{s \to t} pr(|X_s - X_t| \geq \varepsilon) = 0$, for any $0 \leq s < t$

The probability density function of Lévy process is **(Applebaum, 2009)**:-

$$\boldsymbol{f_X(x)} = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(X-\theta)^{\frac{3}{2}}} \, \boldsymbol{exp}\left\{-\frac{\sigma}{2(X-\theta)}\right\} \quad \boldsymbol{for\ X} > \theta \quad , \qquad \textbf{(2-13)}$$

where:

X: location parameter

$\boldsymbol{\theta}$**:** scale parameter

There are some important definitions and features of L which are as follows **(Applebaum, 2009), (korn et.al., 2010) and(onalan, 2009)** :-

Definition:2.6.2.2

The process $L = [L_t; t \geq 0]$, defined on a space of probability (Ω, F, p), which is called the Lévy Process in the case where it satisfies the following conditions:-

1 - $L_0 = 0$

2 -$P(L_0 = 0) = 1$

3 -The increments are independent: for any $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n < \infty$, $L_{t2} - L_{t1}$, $L_{t3} - L_{t2}$,..., $L_{tn} - L_{tn-1}$.

4 -Stationary increments, i.e., for any $t > 0$, $(L_{t+s} - L_t)$ has the same distribution as (L_s), where $s < t$.

5 – Cadlag paths ,i.e., the path of L is continuous from the right and limited from the left.

6 - L is a continuous stochastic process, i.e., for all $\varepsilon > 0$ and $s \geq 0$ we have

$$\lim_{t \to s} p(|x_t - x_s| > \varepsilon) = 0$$

Definition:2.6.2.3

If a L$\acute{e}$vy process (L) assigns a set (V) on (R) through the setting of each borel set (S) that doesn't include zero **(Onalan, 2009)**

$$\boldsymbol{V}(\boldsymbol{S}) = \boldsymbol{E}[\{\boldsymbol{t} \in [\boldsymbol{0}, \boldsymbol{1}]: \Delta \boldsymbol{L}_{\boldsymbol{t}} \neq \boldsymbol{0}, \Delta \boldsymbol{L}_{\boldsymbol{t}} \in \boldsymbol{S}\}] \qquad \textbf{(2-14)}$$

And

V{0} = 0.

Therefore (V) represents a positive measure called the L$\acute{e}$vy measure and it has the following features **(Barndorff & Shepherd, 2001), (Ken-Iti, 1999)** :-

1- $V(S) < \infty$ represents a Borel set and it is bounded away from 0.

2- $\int_{-\infty}^{\infty} (|x|^2)\, v(dx) < \infty$,

where:

V(S): the expected number of the jumps for each unit time whose sizes belong to a set S**.**

Definition:2.6.2.4

Lévy filtration (L^X) be the Lévy process, (L^X) achieve the used conditions of right continuity and completeness **(Al-Saadony[a], 2016).**

2.6.2.1.properties of Lévy process

There are some important properties of the Levy process, and they can be summarized as follows **(Feller, 2008), (Barker, 2019)** :-

1-Infinitely divisible distribution

Poisson and normal distributions are infinitely divisible. It is known that these distributions can be written as the sum of Poisson / normal distributions, respectively, which can be easily verified as being infinitely divisible.

Examples of the infinitely divisible distribution cases correspond with simplest Lévy processes: Poisson counting processes and Brownian motion **(Barker, 2019) .**

Definition:1.1

The random variable X has an infinite distribution that is divisible if for each $n \in N$, there are (i.i.d.)r.v's $X_1, X_2, \ldots, X_n$ so that X has the same distribution as $X_1 + X_2 + \cdots + X_n$

The stationary independent increments definition indicate that Lévy processes have infinite divisible distributions. Surely, for each $n \in N$

$$\boldsymbol{X_t =^d X_{\frac{t}{n}} + \left(X_{\frac{2t}{n}} - X_{\frac{t}{n}}\right) + \cdots + \left(X_{\frac{nt}{n}} - X_{\frac{(n-1)t}{n}}\right)}, \qquad \textbf{(2-15)}$$

where each of these increases are independent with the same distribution. On the contrary, every infinitely divisible distribution leads to the unique Lévy process, so there's a one-to-one match between infinitely divisible distributions & Lévy processes. For further details see **(Feller, 2008).**

2-The Lévy – khintchine formula

It is a very important formula that laid the foundations for the modern study of Lévy process, as this formula gives an analytical expression to the characteristic function associated with the Lévy process in general through three main keys. The characteristic function defines the process distribution which gives a simple and elegant way to work with the Lévy process.

However, often work is done with the characteristic exponents Ψ determined by the equation below(2-16) and the characteristic function itself is not studied **(Barker, 2019)**.

$$E\left[e^{i\lambda X_t}\right] = e^{-t\Psi(\lambda)} \quad \text{, for all } \lambda \in R \quad , \qquad \text{(2-16)}$$

where:

$\boldsymbol{\Psi(\lambda)}$**:** Lévy – khintchine.

Since this relationship applies to every $t \geq 0$, this confirms that Lévy process have infinite divisible distribution **[see the subsection above 2-6-2-1(1)]**, as well as the stationary and independent increments property.

Lévy – khintchine formula was firstly proven in 1934 by **Lévy** but a simpler guide is presented in 1937 **(khintchine, 1937).**

Theorem 2.1. (Lévy – khintchine formula)

Let **L** be a Lévy process real – value random variable is infinite divisible with characteristic exponent Ψ **(AL-Saadony[a], 2016) and (Onalan, 2009):-**

$$\int_R e^{i\lambda x}\mu(dx) = e^{-\Psi(\lambda)}, \text{ for } \lambda \in R$$

Then, with the triple (a, σ^2, v), where $a \in R$, $\sigma \in R$ and V is a Lévy measure on $R\backslash\{0\}$ **(Kyprianou, 2014) and (Barker, 2019).** Then

$$\Psi(\lambda) = ia\lambda + \frac{1}{2}\sigma^2\lambda^2 + \int_R \left(1 - e^{i\lambda X} + i\lambda X\, 1_{(|L|<1)}\right) V dL \qquad \textbf{(2-17)}$$

For every $\lambda \in R$, a and σ^2 are constant, Moreover, the triple (a, σ^2, v) is unique, where:

(a, σ^2, v) represent the triple Lévy process

$(ia\lambda)$ represent the drift term parameter

$(\sigma^2\lambda^2)$ represent the Brownian motion

3- Lévy - $\hat{I}to$ Formula

This formula is used to describe the path of the sample and it is also contains basic information about simulating the Lévy process. It also works by relating the distribution to the process. Using to distinguish (L) depending on the triple (a, σ^2, v).

The form of Lévy - $I\hat{t}o$ can be written as follows **(Mousa, 2016)** :-

$$L_{Ito} = a_t + \sigma\, w_t + \int_0^t \int_{|L|\geq 1} Ln^L(ds, dL) + [\int_0^t \int_{|L|<1} Ln^L(ds,dL)] - t \int_{|L|<1} Lv(dL) \;, \qquad \textbf{(2-18)}$$

where

a_t: represents a drift

$\sigma\ w_t$: is the Brownian motion

$\int_0^t \int_{|L|\geq 1} Ln^L(ds, dL)$: is the independent compound poisson process

$[\int_0^t \int_{|L|<1} Ln^L(ds,dL)] - t \int_{|L|<1} L\nu(dL)$: is the pure jump martingale

n^L: is the stochastic measure of the number of jumps.

Given the presence of extreme values in heavy-tailed distributions, we will dwell on them in some detail..

2.7.Extreme value theory

Assume $X_1, X_2, \dots \dots X_n$ are identically independent distribution (i.i.d.) r.vs with cumulative distribution function (c.d.f.). Suppose we are careful about the probability that maximal value doesn't override a certain threshold x. This probability is:-

$$\boldsymbol{P[max\ (X_1, X_2, \dots \dots X_n) \leq x] = P[X_1 \leq x_1, X_2 \leq x_2, \dots \dots X_n \leq x_n]}$$

$$\boldsymbol{\overset{iid}{\approx} [F(x)]^n} \qquad \textbf{(2-19)}$$

Extreme value theory (EVT) gives the conditions under which sequences of constants values (a_n) and (b_n) such that:-

$$\boldsymbol{\lim_{n\to\infty}[F(a_n x + b_n)]^n \to G(x)} \ , \qquad \textbf{(2-20)}$$

where:

$G(x)$: represents a well-defined non-degenerate c.d.f..

There are three possible $G(x)$ that depend on the tail shape of F(x). Here we will focus on distributions that have a regularly varying tail **(Danielsson et.al., 2019)** **:-**

$$\frac{1-F(x)}{x^{-\alpha}\,L(x)} = 1 \qquad as \quad x > 0,\ x \to \infty \quad , \qquad (2\text{-}21)$$

where:-

$\alpha = \frac{1}{y}$: is the tail index or the index of regular variation, $y > 0$

L is slowly varying function i.e,

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

α determines how heavy the tail is.

There are many different estimates for estimating tail index α **(Hill, 1975), (Pickands, 1975), (Mason[b], 1982), (Hall, 1982), (Davis & Resnick, 1984), (Hall & Welsh, 1985)and (Csorgo et.al., 1985)** but the most common tool is the Hill estimator which will be explained later.

Now how to find the absolute extreme value

1- Find all the critical numbers for f in an interval [a, b].
2- Enter every one of the critical numbers from step 1 in f(x)
3- Insert end-points a & b in f(x)
4- The largest value represents absolute maximum and the smallest value represents absolute minimum.

Given the importance of the tail index in extreme values, we will estimate :

2.8.Tail index estimation

There are many estimators of the tail index (α) in equation **(2-21)**, but here we will use the well-known Hill estimator **(Hill, 1975)** to address some important issues and solutions for estimating the tail index such as optimal choice of the sample fraction (k) and goodness-of-fit test .

2.9.Hil estimator

It is one of the most important estimators used to detect the presence of heavy tails for the marginal distribution of stable sequences of random variables. To define the Hill estimator, assume that the observations $(X_1, X_2, \dots \dots, X_n)$ are nonnegative. For $1 \leq i \leq n$, write $X_{(i)}$ for the (i th) largest value of $(X_1, X_2, \dots \dots, X_n)$, so that

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$$

Then Hill's estimator is defined as **(Hill, 1975)**:-

$$\hat{\boldsymbol{y}} = \frac{\mathbf{1}}{\hat{\boldsymbol{\alpha}}} = \left[\frac{\mathbf{1}}{\boldsymbol{k}} \sum_{i=1}^{k} \mathbf{log} \frac{\boldsymbol{x}_{n,n-i+1}}{\boldsymbol{x}_{n,n-k}}\right]$$

Hence :

$$\hat{\boldsymbol{\alpha}} = \left[\frac{\mathbf{1}}{\boldsymbol{k}} \sum_{i=1}^{k} \mathbf{log} \frac{\boldsymbol{x}_{n,n-i+1}}{\boldsymbol{n,n-k}}\right]^{-1} \quad , \qquad \textbf{(2-22)}$$

where:-

K: number of upper-order statistics used in estimating α.

$\hat{\alpha}(k)$ is a consistent estimator for the tail index if the following are achieved **(Masson[a], 1982)**:-

$$k = k(n) \to \infty \quad and \; \frac{k}{n} \to 0 \;\; as\, n \to \infty$$

Using a small (k) leads to large variance; and it is possible that the estimator will be biased if (k) is too large **(Nemeth & Zempleni, 2020).** Therefore the Hill estimator is strongly dependent on optimal k selection, see, e.g.**(Danielsson et.al., 2019), (Gomes et.al., 2009)** and others.

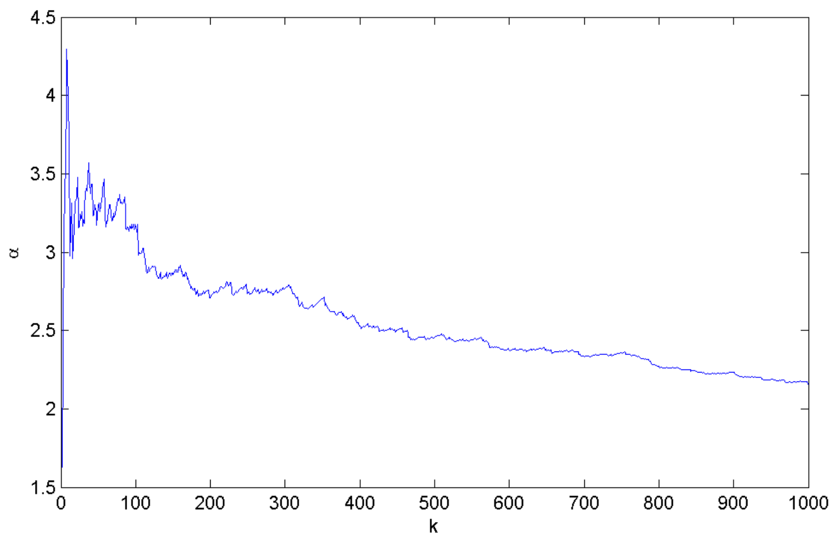

Figure 2-1: Hill plot for distribution of student-t(4) (Danielsson et.al., 2019),

Figure 2-1 illustrates the reciprocity of Hill's estimates of a sample that has been taken from a student-t (4) distribution that has been plotted against increasing number of the order statistics **k**. Selecting **k** is very important to obtain the correct estimate as the estimate of $\boldsymbol{\alpha}$ varies greatly according to **k**.

Figure 2-1, we notice that when **k** is small, the variance in the Hill estimator is high. With the increases in the value of **k**, the variability decreases and bias increases. Variation and bias can be found in the estimator of parameter distributions for sub-class of distributions in equation (2-21) satisfying what is referred to as Hall expansion.

$$\mathbf{1 - F(x) = Ax^{-\alpha}\left[1 + Bx^{-\beta} + 0\left(x^{-\beta}\right)\right]}\ , \qquad \textbf{(2-23)}$$

where

A,α,B: represent the parameter values.

β: constant.

Utilizing Hall expansion to show asymptotic bias as follows:

$$\boldsymbol{E}\left\{\frac{\mathbf{1}}{\widehat{\boldsymbol{\alpha}}} - \frac{\mathbf{1}}{\boldsymbol{\alpha}}\middle|\boldsymbol{x}_{\boldsymbol{n},\boldsymbol{n}-\boldsymbol{i}} > S\right\} = \frac{-\boldsymbol{\beta}\boldsymbol{B}\boldsymbol{S}^{-\boldsymbol{\beta}}}{\boldsymbol{\alpha}(\boldsymbol{\alpha}+\boldsymbol{\beta})} + \mathbf{0}\left(\boldsymbol{S}^{-\boldsymbol{\beta}}\right) \quad , \qquad \textbf{(2-24)}$$

where:

S: is the threshold

Equation (2-24) shows the relationship between the bias of Hill estimator and the threshold S, the bias increases as S gets smaller (in other words, threshold moves towards the distribution center).

This chart shows estimation of (α) for another level of (k). The sample is taken from the student-t distribution with (4) freedom degrees such that $(\alpha = 4)$. The size of the sample is 10,000, this diagram is refer to as Hill plot.

The asymptotic variance of the Hill estimator is as follows:

$$\boldsymbol{Var}\left(\frac{\mathbf{1}}{\widehat{\boldsymbol{\alpha}}}\right) = \frac{\boldsymbol{S}^{\boldsymbol{\alpha}}}{\boldsymbol{nA}}\frac{\mathbf{1}}{\boldsymbol{\alpha}^{\mathbf{2}}} + \mathbf{0}\left(\frac{\boldsymbol{S}^{\boldsymbol{\alpha}}}{\boldsymbol{n}}\right) \quad , \qquad \textbf{(2-25)}$$

where:

S: represent the threshold.

A,α: the parameters values.

The variance is a function of (S). Variance becomes smaller as (S) decreases, and variance predominates if (S) is large and bias is small.

When using the Hill estimator there are many difficulties including the following **(Resnick, 2007):-**

1 - Determine what value of K we should use.

2 - The chart may show a large volatility .

3 - Hill estimator is not fixed on location. This means that the Hill estimator is very sensitive to changes in location.

Now we can overcome these difficulties by using some techniques to smoothen and re-measure the Hill plot through the smoothing Hill method.

2.9.1.Smoothing Hill

Resnick and **Starica** in 1995 proposed this technique to reduce the Hill plot estimator variance as mean values of Hill estimator corresponding to different numbers of the order statistics and as follows **(Resnick & Starica, 1997):**

$$\boldsymbol{av\ H_{k,n} := \frac{1}{(r-1)k} \sum_{p=k+1}^{rk} H_{p,n}} \qquad \textbf{(2-26)}$$

It is a consistent estimate of $\frac{1}{\alpha}$, **r** integers greater than one **(often 2 or 3).**

2.10.Order statistics

The order statistics are sample values arranged in ascending order. For a sample of independent observations $X_1, X_2, X_3, \dots \dots X_n$ on a distribution (F), then :

$$X_1 = \min\,(X_1, X_2, X_3, \dots \dots X_n)$$

$$X_n = \max(X_1, X_2, X_3, \dots \dots X_n)$$

The ordered sample values are:

$$X_{(1)} \le X_{(2)} \le X_{(3)} \le \cdots \le X_{(n)}$$

That are called the order statistics.

Now we will present the most important part of our thesis, which is the heavy tails, and some important definitions and details will be clarified. we will also present the types, classes and properties of the heavy- tailed distributions as follows:

2.11.Heavy-tailed distribution

Heavy tail can be defined as the fact that there 's a higher probability to get very high values. The heavy – tailed distribution goes to zero slower than the exponential distribution. This type of distribution tends to have many extreme values with very high values. There will be more density under the p.d.f. curve. The greater the weight of the tail, the more likely it is to obtain one or more disproportionate values in the sample.

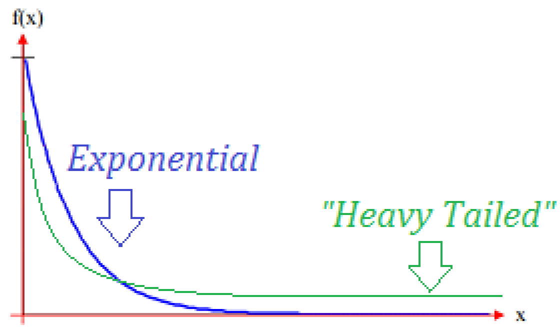

Figure2-2: the heavy-tailed distribution (Bryson, 1974)

In probability theory, a heavy-tailed distribution is a probability distribution whose tail is not exponentially constrained **(Asmussen, 2003).** In numerous applications the distribution's right tail is important, while the left tail could be heavy or both tails may be heavy.

There is some discrepancy about the extent to which the term heavy tail is used. Some authors use the term for distributions whose moments are not limited, while others use it to refer to distributions whose variance is infinite.

Definition:2.11.1

the distribution function F(X) is considered to be having a (right) heavy tailed with tail index if ($\alpha > 0$) satisfies **(Peng & Qi, 2017) :-**

$$\lim_{t\to\infty}\frac{1-F(tx)}{F(t)} = x^{-\alpha} \quad \text{for all } x > 0 \qquad \text{(2-27)}$$

Similarly, the random variable X that has such the distribution function F in equation (2-27) is called the heavy tailed random variable.

Definition:2.11.2

A distribution of the (r.v.) x with the distribution function is considered to be having a heavy (right) tail in the case where the moment generating function $M_X(t)$ of x is infinite for each $t > 0$ **(Rolski et.al., 2009), (Foss et.al., 2011), i.e.,**

$$\int_{-\infty}^{\infty} e^{tx} dF(x) = \infty \quad \text{for all } t > 0$$

The implications of that:

$$\lim_{x\to\infty} e^{tx}\, Pr[X > x] = \infty \quad \text{for all } t > 0$$

the tail distribution function is written as follows:

$$\overline{F}(x) \equiv Pr[X > x]$$

Such as

$$\lim_{x\to\infty} e^{tx}\,\overline{F}(x) = \infty \quad \text{for all } t > 0$$

There are three types of heavy-tailed distributions as follows:-

2.11.1.Long-tailed distribution

A distribution of the r.v. x with F is said to be having a long right- tailed in the case where **(Asmussen, 2003)** :-

$$\lim_{x\to\infty} Pr\{X > x + t | X > x\} = 1 \quad \text{for all } t > 0$$

Or

$$\overline{F}(x + t) \sim \overline{F}(x) \quad \text{as } x \to \infty$$

This means that in the case where the long-tailed right-tailed quantity exceeds a high level, the probability approaching 1 will exceed all other higher levels.

2.11.2. Fat-tailed distribution

Fat –tailed distribution represents the heavy-tailed distribution with infinite variation. It is a probability distribution that shows a large skewness or kurtosis relative to the exponential or normal distribution. Often the term heavy-tailed and fat-tailed are synonymous.

The log-normal distribution is one example of a fat-tailed distribution **(Bahat et.al. , 2005)**.

several authors state that the fat-tailed distribution is a probability distribution with a tail that appears fatter than usual.

2.11.3.Sub - exponential distribution

It is a distribution in which the largest value in the sample makes a large contribution to the overall total **(Mikosch, 1999)**.

And also sub - exponential is characterized in terms of convolution of the probability distribution. For two i.i.d. random variable (X_1, X_2) with common distribution function (F) convolution of (F) with itself, (F^{*2}) is convolution square, utilizing the Lebesgue- stieltjes integration through:

$$p_r[x_1 + x_2 \leq x] = F^{*2}(x) = \int_0^x F(x-y)dF(y)$$

The n-fold convolution (F^{*n}) is defined inductively through the base:

$$F^{*n}(x) = \int_0^x F(x-y)dF^{*\,n-1}(y)$$

The tail distribution function $(\bar{F})$ is defined by the following forula:

$$\bar{F}(x) = 1 - F(x)$$

The distribution (F) on positive half-line is a subexponential **(Asmussen, 2003) (Chistyakov, 1964) and (Teugels & joze, 1975)** if:-

$$\overline{F^{*2}}(x) \sim 2\bar{F}(x) \quad \text{as } x \to \infty$$

This implies that **(Embrechts & et al., 2013)**

$$\overline{F^{*n}}(x) \sim n\bar{F}(x) \quad \text{as } x \to \infty \text{ for any } n \geq 1$$

The probabilistic interpretation of this is: for a summation (n) of independent r.v. $(x_1, \ldots, x_n)$ with common distribution (F)**(Embrechts et al., 2013):**

$$p_r[X_1 + \cdots + X_n > x] \sim p_r[max(X_1, \ldots, X_n) > x] \quad \text{as } x \to \infty$$

This is called the single big jump **(Foss et al., 2007)** or catastrophe principle **(Wierman & Adam , 2014).**The random variable (x) supported on real line represents a sub exponential if and only if [$X^{+} = max(0, X)$] is a subexponential.

2.12.Classification of the distributions based on their heavy tailedness

The existence of extreme in the sample with independent observations is dependent on the type of distribution and the sample size. The distributions are classified according to their probabilities of having either moderate or extreme values.

Heavy-tailed distributions are categrized into three parts as follows:

2.12.1.Classification of the distributions w.r.t. left-tails heaviness (Pavlina & Monika,2017)

Definition:2.12.1.1

we call a random variable (r.v.) and its cumulative distribution function (c.d.f.) F, $P_{mL}(x)$ mild-heavy left-tailed if

$$\boldsymbol{P_{mL}(X) = P\big(Q_1(F) - 3\ IQR(F) < X \leq Q_1(F) - 1.5\ IQR(F)\big)} \qquad \textbf{(2-28)}$$

Where:

Q_1: is the first quartile

IQR: inter – quartile range

Definition:2.12.1.2

The r.v.(x) and the r.v.(y) belong to the same P_{mL} class if:

$$P_{mL}(X) = P_{mL}(Y)\ .$$

(see figure 3,(b)). if $P_{mL}(X) < P_{mL}(Y)$ this means the r.v.(x) is of a lighter P_{mL} in comparison with an a r.v.(y).

If $X = Y$ in the distribution then $P_{mL}(X) = P_{mL}(Y)$. But if $P_{mL}(X) = P_{mL}(Y)$ neither this doesn't mean the fact that X & Y belong to the same distribution type nor they have the same variance or mean.

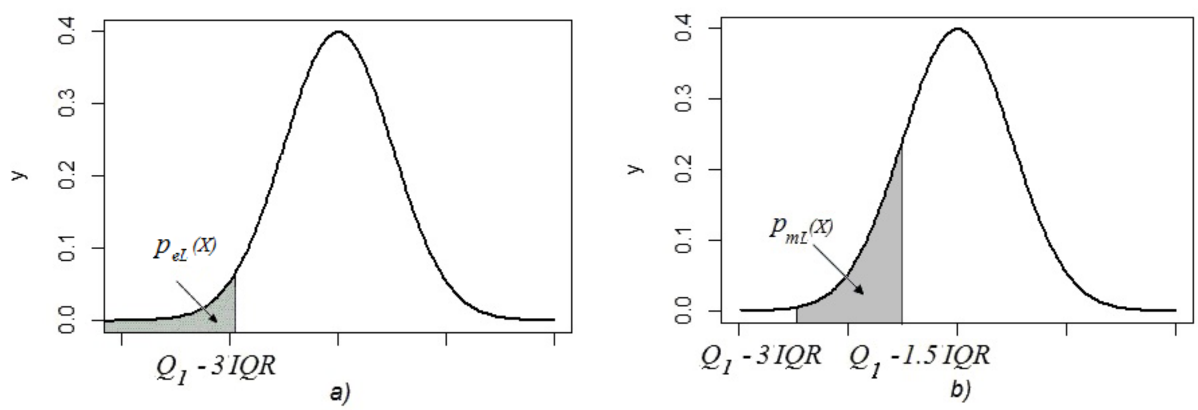

Figure 2-3: Relation between plot of p.d.f. of random variable X with the c.d.f., peL (X) and pmL (X) (Wikipedia)

Definition:2.12.1.3

We call the random variable X and the cumulative distribution function of x, $P_{eL}(X)$-extremely heavy left-tailed if:

$$\boldsymbol{P_{eL}(X) = P\big(X < Q_1(F) - 3\ IQR(F)\big)} \qquad \text{(2-29)}$$

Definition:2.12.1.4

It is said that the r.v. (x) and the r.v. (y) are part of the same P_{eL} class if :-

$$\boldsymbol{P_{eL}(X) = P_{eL}(Y)}$$

Likewise, the r.v. (x) has lighter P_{eL} in comparison with the r.v. (y) in the case where:-

$$P_{eL}(X) < P_{eL}(Y)$$

2.12.2.Classification of the distributions w.r.t right-tails heaviness

There is a possibility of working with right-tails in a similar way to the previous subsection. (see figure 2-4,(a) and (b)).

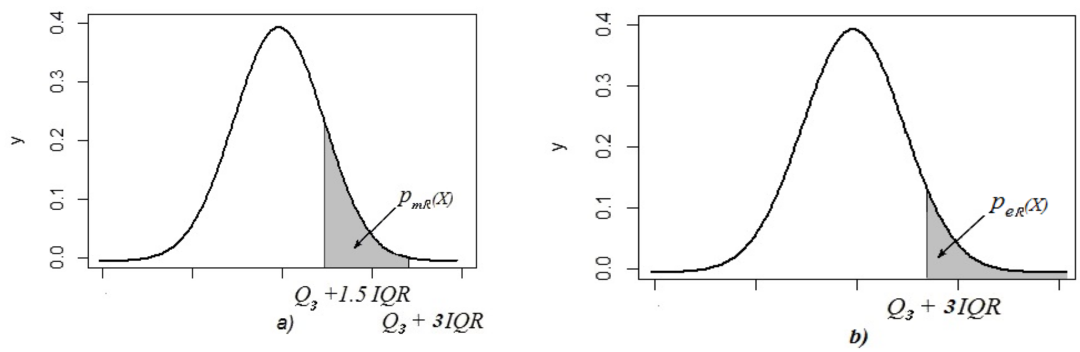

Figure 2-4: Relation between the plot of the p.d.f. of the r.v. (X) with c.d.f. (F), pmR(X) and peR(X) (Wikipedia)

Definition:2.12.2.1

A random variable (r.v.) X and its c.d.f. (F) are called $P_{mR}(X)$ mild-right heavy-tailed if **(Pavlina & Monika,2017):-**:-

$$P_{mR}(X) = P\big(Q_3(F) + 1.5\, IQR(F) < X \le Q_3(F) + 3\, IQR(F)\big) \; , \qquad \textbf{(2-30)}$$

where:

Q_3 is the third quartile.

Definition:2.12.2.2

The r.v. (X) and the r.v. (Y) belong to the same P_{mR} (mild-right heavy- tailed) class If:

$$\boldsymbol{P_{mR}(X) = P_{mR}(Y)}$$

If $P_{mR}(X) < P_{mR}(Y)$ this means a r.v. (X) has lighter P_{mR} than a r.v. (Y).

Definition:2.12.2.3

We call the random variable (X) and its c.d.f. (F), $P_{eR}(X)$-extremely heavy right – tailed in the case where:

$$\boldsymbol{P_{eR}(X) = P\big(X > Q_3(F) + 3\ IQR(F)\big)} \qquad \textbf{(2-31)}$$

Definition:2.12.2.4

A random variable X and a random variable Y belong to the same P_{eR} class if:

$$\boldsymbol{P_{eR}(X) = P_{eR}(Y)}$$

Likewise, a random variable X has lighter P_{eR} than a random variable Y if:

$$\boldsymbol{P_{eR}(X) < P_{eR}(Y)}$$

2.12.3.Classification of the distributions w.r.t. two sided-tails heaviness

Definition:2.12.3.1

We call the random variable X and its c.d.f., $P_{m2}(X)$ mild - two heavy - tailed if **(Pavlina & Monika,2017):-**

$$P_{m2}(X) = P\big(Q_1(F) - 3\ IQR(F) < X \le Q_1(F) - \ 1.5\ IQR(F) \cup \ Q_3(F) + 1.5\ IQR(F) < X \le Q_3(F) + 3\ IQR\ (F)\big) \quad (2\text{-}32)$$

Definition:2.12.3.2

A r.v. (X) and a r.v. (Y) are part of the same P_{m2} class if:

$$P_{m2}(X) = P_{m2}(Y)$$

If $P_{m2}(X) < P_{m2}(Y)$ this means the random variable X with cumulative distribution function F has lighter P_{m2} than the random variable y.

Definition:2.12.3.3

We call the random variable (X) and its c.d.f., $P_{e2}(X)$ extremely- two heavy- tailed if **(Pavlina & Monika,2017):-**

$$P_{e2}(X) = P\big(X < Q_1(F) - 3\ IQR\ (F)\ \cup X > Q_3(F) + 3\ IQR\ (F)\big) \qquad (2\text{-}33)$$

Definition:2.12.3.4

The random variable X and the random variable Y are part of the same P_{e2} (extremely heavy two – tailed) class in the case where :

$$P_{e2}(X) = P_{e2}(Y)$$

In a similar way, a random variable (X) has lighter P_{e2}than a random variable (Y) if:

$$P_{e2}(X) < P_{e2}(Y)$$

properties of the Heavy Tailed Distribution are follows:

1- the central limit theory works misleading.
2- order statistics are used because some moments do not exist.

2.13.Light-tailed distributions

They are the probability distributions whose tails are thinner than the exponential distribution. These distributions go to zero faster than the exponential distribution, thus in the tail their mass is less. There are many light-tailed distributions, the most important of which are Gamble distribution, t-distribution and the normal distribution.

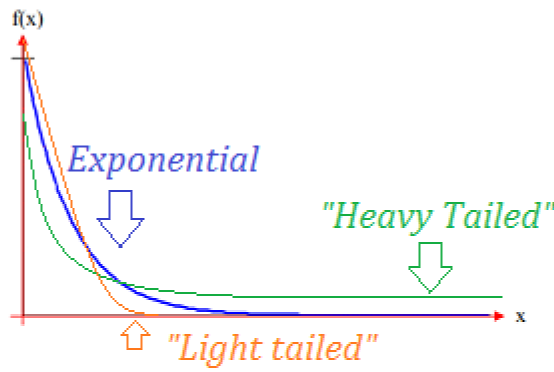

Figure 2-5// light tailed distribution (Bryson, 1974)

2.14.Choose the optimal k methods

We will use three methods to choose the optimal k that balances bias and variance and thus will reduce the mean square error. Relying on k selection, we will estimate the tail index in heavy-tailed distributions as follows:

2.15.Direct estimation method

A simple method for selecting optimal k in the **equation (2.34)** directly is given by **(Peng & Qi, 2017)** :

$$K_{opt} = argmin_k \left[\frac{\alpha^4}{(1+\rho)^2}\,\beta^2\left(\frac{k}{n}\right)^{2\rho} + \frac{\alpha^2}{k}\right], \quad \textbf{(2-34)}$$

where:

α**:** is the Hill estimator mentioned in (2-22)

k: represent the **equation (2-35)**

n**:** sample size

ρ and β : are the second order regular variation parameters, they can be calculated using the following steps:-

a) steps to calculate ρ **(Gomes & Pestana, 2007):**

1- $$k = \min\left[n-1, \left(\frac{2n}{\ln\ln n}\right)\right] \quad \textbf{(2-35)}$$

2- $$H_n^{(\varepsilon)}(k) = \left[\frac{1}{k}\sum_{i=1}^{k}\ln\frac{X_{n,n-i+1}}{X_{n,n-k}}\right]^{\varepsilon} \quad \textbf{(2-36)} \quad \varepsilon=1,2,3,4$$

3- if $\tau = 0$

$$T_{(n)}^{(\tau)}(k) = \left[\frac{\ln\left(H_n^{(1)}(k)\right)-\frac{1}{2}\ln\left(H_n^{(2)}(k)/2\right)}{\frac{1}{2}\ln\left(H_n^{(2)}(k)/2\right)-\frac{1}{3}\ln\left(H_n^{(3)}(k)/6\right)}\right] \quad \textbf{(2-37)}$$

4- if $\rho < 0$

$$\hat{\rho}_{(k)} = -\left|\frac{3\left(T_n^{(\tau)}(k)-1\right)}{T_n^{(\tau)}(k)-3}\right| \quad \textbf{(2-38)}$$

b) we can calculate β as follows **(Caeiro & Gomes, 2006):**

$$\hat{\beta}_{\tau}(k) = -\frac{2(2+\hat{\rho}_k)}{\tau\,\hat{\rho}_k\hat{\alpha}_k}\left(\frac{n}{k}\right)^{\hat{\rho}_k}\frac{\left[\left(H_n^{(1)}(k)\right)^{\tau}-\left(H_n^{(2)}(k)/2\right)^{\frac{\tau}{2}}\right]^2}{\left(H_n^{(2)}(k)/2\right)^{\tau}-\left(H_n^{(4)}(k)/24\right)^{\frac{\tau}{2}}} \quad , \quad \textbf{(2-39)}$$

where $\hat{\alpha}_{(k)}$, $\hat{\rho}_{(k)}$ and $\hat{\beta}_{\tau}(k)$ are consistent estimators for α, ρ and β respectively.

2.16.Hall's Bootstrap method

It is an important method used to estimate statistics on a population by sampling a dataset with replacement. Bootstrap is helpful in the case where there isn't any analytical form or normal theory for helping the estimation of statistics distribution, due to the fact that the Bootstrap approaches may be applied to most random quantities.**(Hall , 1990)** suggested the bootstrap method for the estimation of the Mean Squared Error (MSE) and selection of the smoothing parameters in the non-parametric methods. It should be noted that the subsample bootstrap method is required to capture the term bias of the tail parameters estimator. Now suppose $X_1, X_2, \dots \dots, X_n$ denote observations from the distribution with distribution function (F) and assume :

$$\mathbf{1 - F(X) \sim CX^{-\alpha}} \quad , \qquad \textbf{(2-40)}$$

where, C and $\alpha > 0$. We would like to estimate α, given the sample

$x_n = \{X_1, X_2, \dots \dots X_n\}$. **(Hill, 1975)** proposed the estimator $\hat{\alpha}$ [mentioned in **equation (2-40)**]. $X_{n1} \geq X_{n2} \geq \cdots \geq X_{nn}$ denote the order statistics of x_n and k is a smoothing parameter. We will choose k to minimize the mean square error (MSE) of $\hat{\alpha}$ **(Hall, 1990).** Put

$$\textbf{MSE(n, k)} = \mathbf{E}[\hat{\boldsymbol{\alpha}}(\boldsymbol{K}) - \boldsymbol{\alpha}]^{\mathbf{2}} \quad , \qquad \textbf{(2-41)}$$

where:

$\hat{\alpha}(K)$ is the Hill estimator based on Bootstrap.

α: Hill estimator

To select K the Bootstrap method includes the following:

Draw a resample $x_{n1}^* = \{x_1^*, x_2^*, \dots \dots x_{n1}^*\}$ from x_n; $n_1 \leq n$.

Let

$x^*_{n1,1} \geq x^*_{n1,2} \geq \cdots \geq x^*_{n1,n1}$ denote the order statistics of $x^*_1, x^*_2, \ldots \ldots x^*_{n1}$, and put

$$\widehat{\boldsymbol{\alpha}}^*(\boldsymbol{n1}, \boldsymbol{k1}) = \left[\frac{\mathbf{1}}{\boldsymbol{k}} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{k1}} \log \boldsymbol{x}^*_{\boldsymbol{n1},\boldsymbol{n1}-\boldsymbol{i}+\mathbf{1}} - \log \boldsymbol{x}^*_{\boldsymbol{n1},\boldsymbol{n1}-\boldsymbol{k1}}\right]^{-\mathbf{1}}$$

The Bootstrap estimate of $MSE(n_1, k_1)$ is **(Peng & Qi, 2017)**:-

$$\widehat{\boldsymbol{MSE}}\,(\boldsymbol{n_1}, \boldsymbol{k_1}) = \boldsymbol{E}[\{\widehat{\boldsymbol{\alpha}}^*(\boldsymbol{n_1}, \boldsymbol{k_1}) - \widehat{\boldsymbol{\alpha}}(\boldsymbol{n}, \boldsymbol{k})\}^2 | \boldsymbol{x_1}, \boldsymbol{x_2}, \ldots, \boldsymbol{x_n}] \qquad \textbf{(2-42)}$$

Then choose $\hat{k}_1$ to minimize $\widehat{MSE}\,(n_1, k_1)$. It shall be noted that **equation(2-42)** can fail if α is unsmooth, and in this case the bias of $\hat{\alpha}(K)$ is often a major contributor to the MSE, and the Bootstrap method does not accurately estimate bias. **Equation (2-42)** is calculated as an average of $(\widehat{\boldsymbol{\alpha}}^*(\boldsymbol{n_1}, \boldsymbol{k_1}) - \widehat{\boldsymbol{\alpha}}(\boldsymbol{n}, \boldsymbol{k}))^2$ through a large number of resampling**.** When the optimal $\boldsymbol{k_{opt}} = \boldsymbol{cn^y}$ for a known y ∈ (*0,1*)(in many cases the value of y is $\frac{2}{3}$), but an unknown $c > 0$, **(Hall, 1990)** proposed to estimate k_{opt} by:

$$\widehat{\boldsymbol{k}}_{\boldsymbol{H}} = \widehat{\boldsymbol{k}}_{\mathbf{1}}\left(\frac{\boldsymbol{n}}{\boldsymbol{n1}}\right)^{\boldsymbol{y}} \qquad \textbf{(2-43)}$$

2.17.Double Bootstrap method

This method states that the estimation of the tail index depends on its accuracy in selecting the sample fraction. This method offers a solution for selecting the sample fraction by utilizing two-step sub-sample bootstrap approach. in this approach, the sample fraction reduces the MSE (n,k) and it is important to note that this method enables us to dispense with a necessity for prior estimation of the tail index **(Danielsson et al. 2001),** where:

$$\widehat{\boldsymbol{M}}(\boldsymbol{k}) = \left\{\frac{1}{k}\sum_{i=1}^{k}\left(\log\frac{X_{n,n-i+1}}{X_{n,n-k}}\right)^{2}\right\}^{-1} \qquad \textbf{(2-44)}$$

Now, draw a resample $[X_1^*, \ldots\ldots, X_{n1}^*]$ from $[X_1, \ldots\ldots, X_n]$ with a smaller sample size $n_1 = O\left(n^{1-\delta}\right)$ for some $\delta \in \left(0, \frac{1}{2}\right)$. Determine the corresponding estimators of $\hat{\alpha}(k)$ and $\widehat{M}(k)$ based on bootstrap sample as $\hat{\alpha}^*(k)$ and $\widehat{M}^*(k)$, then choose:

$$\widehat{\boldsymbol{k}}_{\boldsymbol{1}} = \boldsymbol{arg\ min}_{\boldsymbol{k1}}\boldsymbol{E}\left\{\left(\frac{1}{2}\left(\widehat{\boldsymbol{\alpha}}^*(\boldsymbol{k_1})\right)^2 - \widehat{\boldsymbol{M}}^*(\boldsymbol{k_1})\right)^2 \middle| \boldsymbol{x_1}, \ldots\ldots, \boldsymbol{x_n}\right\}, \qquad \textbf{(2-45)}$$

where:

$\widehat{\boldsymbol{\alpha}}^*(\boldsymbol{k_1})$**:** represent the Hill estimator

Equation(2-45) is calculated through a large number of resampling. We repeat the **equation(2-44)** with $n_2 = \frac{n_1^2}{n}$ and we get $\hat{k}_2$.

Then the optimal k_{opt} can be estimated by**(Peng & Qi, 2017)**:

$$\widehat{\boldsymbol{k}}_{\boldsymbol{DHPV}} = \frac{\widehat{K}_1^2}{\widehat{K}_2}\left\{\frac{(\boldsymbol{log}\,\widehat{\boldsymbol{k}}_{\boldsymbol{1}})^2}{(\boldsymbol{2logn_1 - log\widehat{k}_1})^2}\right\}^{\frac{\boldsymbol{logn_1 - log\widehat{k}_1}}{\boldsymbol{logn_1}}} \qquad \textbf{(2-46)}$$

Hence

$$\textbf{MSE(n, k)} = \textbf{E}[\widehat{\boldsymbol{\alpha}}(\boldsymbol{K}) - \boldsymbol{\alpha}]^{\boldsymbol{2}} \qquad \textbf{(2-47)}$$

$\hat{\alpha}(K)$ is the Hill estimator based on Double Bootstrap.

α: Hill estimator

Chapter Three

Simulation part

3.1: Simulation study

In order to compare the Hill estimator with other nonparametric estimators represented by the Direct method , Bootstrap and Double Bootstrap method, The simulation technique was adopted. We generate the data using two models of SDE (Geometric Brownian motion model and Levy model) with $\mu = 0.5$ and $\sigma = 0.05$. The following sizes of the samples (N= 50, 100, 150, 200, 250, 500, 800 and 1000) and AMSE criterion are used to compare these methods. The method with less value of AMSE is the best. Then we have get tail parameter (α); K_{opt} and AMSE for each sample based on 100 replications . The results of the simulation were obtained based on a program written by R in appendix. The simulation studies have two part A and B.

Part A

The data is simulated using equation **(2-11)**

3.2: The results

We get the value of the tail parameter , K_{opt} and AMSE depending on the G.B.M. as follows:

Table 3.1: α, AMSE and K_{opt} for five samples driven by G.B.M. for simulation data.

N=50	Direct	Bootstrap	Double Bootstrap
α	0.07107605	0.0102245247	0.033000000
AMSE	0.01004159	0.0001292511	0.001716803
K_{opt}	7.11581631	3.0000000000	2.000000000
N=100			
α	0.041829245	0.0105181566	0.0223157895
AMSE	0.002247387	0.0001373524	0.0004991821

K_{opt}	0.490268148	3.0000000000	2.0000000000
N=150			
α	0.04216307	0.0162245871	0.045960000
AMSE	0.00261230	0.0003937763	0.002791125
K_{opt}	0.25461379	3.4100000000	2.610000000
N=200			
α	0.050240318	0.0164169738	0.054640000
AMSE	0.002988345	0.0003870336	0.003323239
K_{opt}	0.168603382	3.9500000000	2.970000000
N=250			
α	0.056774807	0.0224340013	0.066000000
AMSE	0.003709766	0.0008237568	0.005643042
K_{opt}	0.041839528	4.1200000000	3.350000000
N=500			
α	0.051385442	0.0229479856	0.13840000
AMSE	0.002843486	0.0007349886	0.02520428
K_{opt}	0.028233823	6.0200000000	5.24000000
N=800			
α	0.045265601	0.0234060623	0.24053333
AMSE	0.002127585	0.0007046003	0.07481925
K_{opt}	0.021606282	8.2000000000	7.44000000
N=1000			
α	0.06808131	0.0210652667	0.16240000
AMSE	0.08998870	0.0006364354	0.03122137
K_{opt}	0.02087413	9.1500000000	8.52000000

Table (3.1) represent the value of α, AMSE and K_{opt} for our model using Direct, Bootstrap and Double Bootstrap methods. It is obvious that the Bootstrap method is much better than the others. We also note that the Direct method is better than the Double Bootstrap method when the sample size is between 50-800 while the Double Bootstrap excels when N=1000.

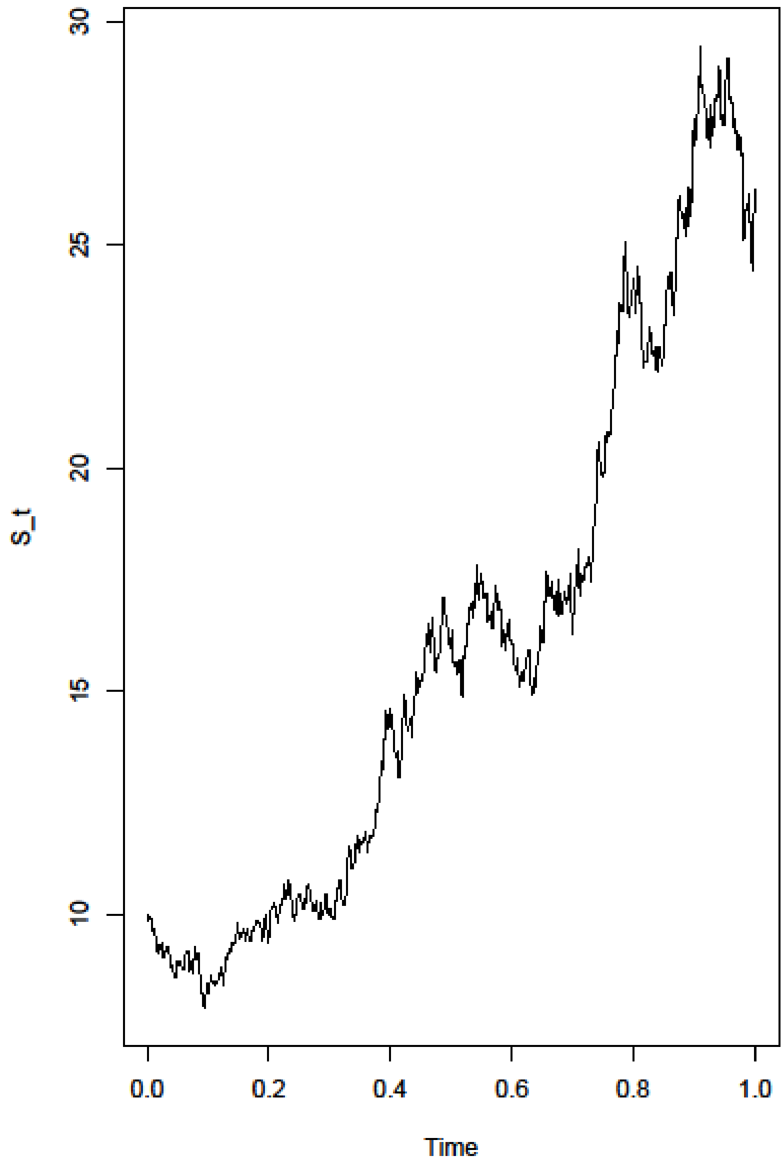

Figure 3.1: the Geometric Brownian Motion model through time.

Figure (3.1) shows the movement of Geometric Brownian motion through time. It is clear that the process affects by Brownian motion and always is positive.

Part B

The data is simulated using the following equation:

$$S_t = S_0 \, exp\left[\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma L_t\right], t > 0 \qquad (3\text{-}1)$$

where

L_t: represent Levy process.

3.3: The results

We get the value of the tail parameter, K_{opt} and AMSE depending on the Levy process as follows:

Table 3.2: α, AMSE and K_{opt} for five samples driven by Levy process for simulation data.

N=50	Direct	Bootstrap	Double Bootstrap
α	0.035319927	0.0111697464	0.0210000000
AMSE	0.002396321	0.0001245574	0.0004164086
K_{opt}	5.465085974	3.0000000000	2.0000000000
N=100			
α	0.031221234	0.0172252821	0.045560000
AMSE	0.001191061	0.0004293483	0.002385551
K_{opt}	0.215970520	3.4400000000	3.000000000
N=150			
α	0.038914153	0.0205659604	0.053600000
AMSE	0.001836429	0.0007201316	0.003359408
K_{opt}	0.146922287	4.0000000000	3.000000000
N=200			
α	0.038106131	0.022992670	0.077800000
AMSE	0.001644134	0.000778499	0.007320747
K_{opt}	0.100875938	4.870000000	4.000000000
N=250			

α	0.034531429	0.030054351	0.081680000
AMSE	0.001272489	0.001684765	0.007830792
K_{opt}	0.087406402	5.000000000	4.110000000
N=500			
α	0.031824993	0.026969259	0.19573333
AMSE	0.000177517	0.000893379	0.04707668
K_{opt}	0.059326323	7.980000000	7.00000000
N=800			
α	0.0293744686	0.028605673	0.348000
AMSE	0.0009357992	0.001203129	0.147316
K_{opt}	0.0492388616	10.900000000	10.860000
N=1000			
α	0.0277368102	0.026801045	0.25140000
AMSE	0.0008210364	0.001114412	0.07700598
K_{opt}	0.0425225095	13.220000000	13.00000000

Table (3.2) represent the value of α, AMSE and K_{opt} for our model using Direct, Bootstrap and Double Bootstrap methods. it is obvious that the Bootstrap method is better than other methods when the sample size is 50,100,150 and 200 while the direct method is better when the sample size is greater than or equal to 250. This is because the Bootstrap method fails as the sample size increases.

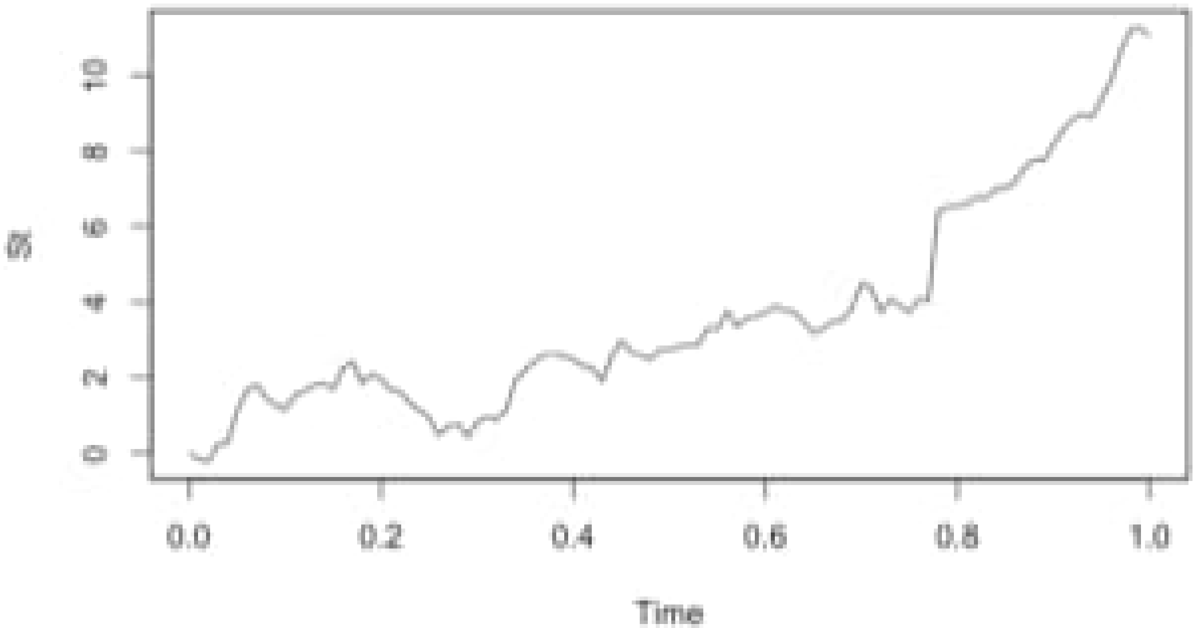

Figure 3.2: the Levy process through time.

Figure (3.2) represents the Black-Scholes model measured by the Levy process. The Levy process here represents the Inverse Gaussian distribution that depends on $\mu = 0.5$ and $\sigma = 0.05$, which is always a positive value.

Chapter Four

Applied part

4.1. Introduction

The stock market is a very important concept in the economic and financial field in any country. In 2004 AD, the Iraq Stock Exchange (ISX) was established in Baghdad, Iraq, this market operates under a supervision from Iraqi Securities Commission, which is independent body established along lines of American Securities and Exchange Commission. In the period before 2003 AD, the current market has been called the Baghdad stock Exchange that has been managed by Ministry of Finance of Iraq, however, it is now a self – regulatory body like stock Exchange of New York, and as of 2005, ISX became the only Iraqi stock Exchange.

In this section, we apply the methods mentioned in the theoretical side to the real data (ISX) for the dinar for the period 1/1/2017 - 1/1/2020.The data was obtained from (Homepage www.isx.iq.net). We used the daily returns for the mentioned period as follows:

$$r_t = \log \frac{S_t}{S_{t-1}} \quad , \qquad \textbf{(4-1)}$$

where

r_t: represent daily returns at time t.

S_t: exchange rate at time t.

4.2. Kolmgorov-Smirnov test

We use this test to see if the data follows a normal distribution or not. The null hypothesis of the test states that the data have a normal distribution. The p-value of the test is ($2.2e^{-16}$) at the significant level of (5%). Therefore, the null hypothesis was rejected, which means that the data do not follow a normal distribution.

4.3. Barndorff-Nielsen and Shephard jump test

In order to check for jumps in the data, we use the Barndorff-Nielsen and Shephard jump test. The null hypothesis of this test states that there are no jumps. At the significant level (5%), the test value is (1.3239) and the p-value is (0.09276)for the data. As for the returns, it was the test value is (-0.34325) and the p-value is (0.6343). Therefore, we will accept the null hypothesis, which means that there are no jumps, whether the test is for data or returns.

The comparison between the studied methods was done by calculating the MSE of the data, where the tail index of the studied methods was compared with the tail index of the Hill estimator. We used N=897, $\mu = 0.5$ and $\sigma = 0.05$. After analyzing the data using R- program, we obtained some results presented in the following table.

Table 4.1: shows the summary of the real data

N=898	Direct	Bootstrap	Double Bootstrap
α	0.006878175	0.0017532978	0.02000000
AMSE	0.000111708	0.0002463041	0.000006515888
K_{opt}	5.832650928	4.0000000000	2.00000000

Table (4.1) represent the value of α, MSE and K_{opt} for the real data using Direct, Bootstrap and Double Bootstrap methods. It is obvious that the Double Bootstrap method is better than other methods. We also note that the Direct method is better than the Bootstrap method because of the large sample size.

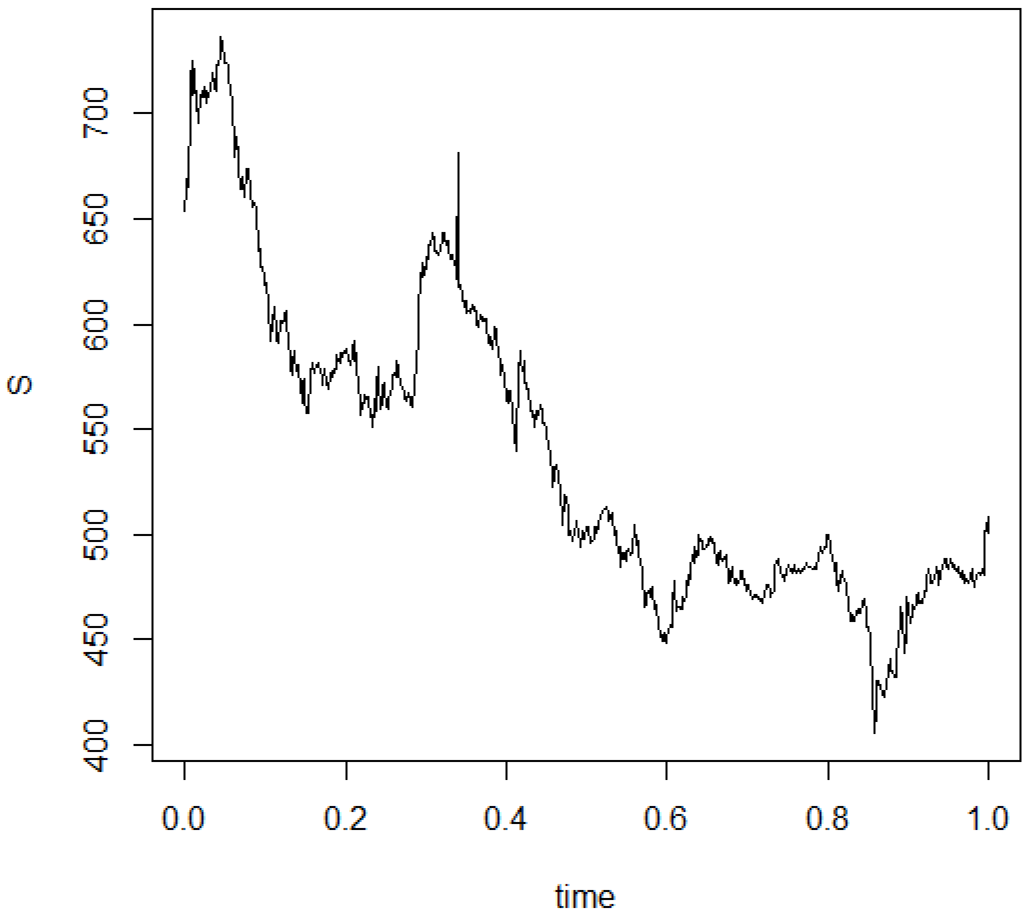

Figure 4.1: the real data through time.

Figure (4.1) represent the real data (ISX) during 2017-2020. It is clear that the behavior of our index follows the Stochastic Differential Equation.

Chapter five

Conclusions

Recommendations

5.1. Conclusions

We conclude from our study the following:

1- In the simulation in part A, the Bootstrap method was the best for all sample sizes. For part B, the Bootstrap method was also best when the sample size was less than 250. while for large values, the Direct method outperformed the others.

2- In the real data, the Double Bootstrap method was the best, and there is a very clear convergence in the results of the other methods.

The explanation for this is that the Bootstrap method fails in large sample sizes in addition to the nature of the data.

5.2. Recommendations

The methods can be used to estimate the tail index with $\alpha > 2$. Also, the tail index for independent data can be estimated based on other estimators instead of Hill estimators such as kernel estimators, linear combinations of intermediate order statistics and least- squares estimators. Other models of stochastic differential equations can be used to estimate the tail index. Finally, we recommend the Iraq Stock Exchange to use:

1- Geometric Brownian Motion in the absence of jumps.

2- Levy model when there are jumps.

The jumps can be determined by the Barndorff-Nielsen and Shephard test (that were discussed in the practical part).

References

Reference

Ahmad, Z., Mahmoudi, E., Hamedani, G. G. & Kharazmi, O. (2020). New methods to define heavy-tailed distributions with applications to insurance data. *Journal of Taibah University for Science*,14(1), 359-382.

Al-Saadony[a], M. F. (2016). A Simulation Study on Stochastic Differential Equation Driven by Lévy processes. *Journal of Al kut for Economic and administrative sciencess*, 23, 1-13.

Al-Saadony[b], M. F. (2016). Utilizing Parametric and Non- parameter Methods in Black – Scholes Process with Application to Finance. *Journal of Al- Qadisiyah for Computer Science and mathematics*, 8(1).

Applebaum, D. (2009). Lévy processes and stochastic calculus. Cambrige University press.

Asmussen, S. (2003). Steady-State Properties of gi/g/1. *Applied probability and Queues*, 266-301.

Baek, C. & Pipiras, V. (2010). Estimation of parameters in heavy-tailed distribution when its second order tail parameter is known. *Journal of Statistical Planning and Inference*, 140(7), 1957-1967.

Bahat, D. , Rabinovich, A. & Frid, V. (2005). Tensile Fracturing in Rocks (p.570). Springer-verlag Berlin Heidelberg.

Balkema, A. A. & De Haan, L. (1974). Residual life time at great age. *Annals of Probability*, 2(5), 792-804.

Barker, A. (2019). Path Properties of Lévy processes (Doctoral dissertation, University of Reading).

Barndorff – Nielsen, O. E. & Shephard ,N. (2001). Non – Gaussian Ornstien – Uhlenbeck Based Models and some of their uses in financial economics. *Journal of the Royal statistical society: series B (statistical Methodology)*, 63(2), 167-241.

Beirlant, J. and Goegebeur, Y. (2004). Simultaneous tail index estimation. *Journal of statistical*, 2(1).

Borovkov, A. A. (2013). Probability theory. Springer science & Business Media, ISBN 978-1-4471-5201-9.

Caeiro, F., & Gomes, M. I. (2006). A new class of estimators of a " scale" second order parameter. Extremes, 9(3-4), 193-211.

Carr, P., Geman, H., Madan, D.B. & Yor, M. (2007). Self-Decomosability and option pricing. *Mathematical Finance*, 17(1), 31-57.

Chistyakov, V.P. (1964). A Theorem on Sums of Independent Positive Random Variables and Its Applications to Branching Random Processes. *Theory of probability & Its Applications* 9(4), 640-648.

Ciuperca, G. & Mercadier, C. (2010). Semi-parametric estimation for heavy tailed distributions. *Extremes*, 13(1), 55-87.

Csorgo, S., Deheuvels, P. & Mason, D. (1985). Kernel estimates of the tail index of a distribution. *Annals of statistics*, 1050-1077.

Danielsson, J., de Haan, L., Peng, L. & de Vries, C. G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*, 76(2), 226-248.

Danielsson, J., Ergun, L. M., de Haan, L. & de Vries, C. G. (2019). Tail index estimation: Quantile driven threshold selection. *Available at SSRN* 2717478.

Davis, R. & Resnick, S. (1984). Tail estimates motivated by extreme value theory. *Annals of statistics*, 12(4), 1467-1487.

Drees, H., & Kaufmann, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Processes and their applications*, 75(2), 149-172.

Dunbar, S. R. (2016). Stochastic Processes and Advanced Mathematical Finance. *The Definition of Brownian Motion and the Wiener Process, Department of Mathematics*, 203, 68588-0130.

Embrechts, P. ; klueppelberg, C. & Mikosch,T. (2013). Modelling extremal events: for insurance and finance. Springer Science & Business Media.

Feller, W. (2008). An Introduction to Probability Theory and its Applications, vol 2. John Wiley & Sons.

Fisher, B. & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 6, 637- 654.

Florescu, I. (2014). Probability and stochastic processes. John Wiley & Sons.

Foss, S., Korshunov, D. & Zachary, S. (2011). An Introduction to Heavy-Tailed and Sub exponential Distributions. (vol. 6, pp. 0090-6778). New York: springer.

Foss, S.; Konstantopoulos, T. & Zachary, S. (2007). Discrete and Continuous Time Modulated Random Walks with Heavy-Tailed Increments. *Journal of Theoretical Probability*, 20(3), 581-612.

Franke, J. , Hardle, W. K. & Hafner, C. M. (2004). *Statistics of Financial Markets* (vol. 2). Berlin: Springer.

Gardes, L. & Girard, S. (2008). A moving window approach for nonparametric estimation of the conditional tail index. *Journal of Multivariate Analysis*, 99(10), 2368- 2388.

Gomes, M. I., De Haan, L. & Peng, L. (2002). Semi-Parametric Estimation of the Second Order Parameter in Statistics of Extremes. *Extremes*, 59(4), 387-414.

Gomes, M. I., Figueiredo, F. & Neves, M. M. (2012). Adative estimation of heavy right tails: Resampling – based method in action. *Extremes*,15(4), 463-489.

Gomes, M. I., Pestana, D. & Caeiro, F. (2009). A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Statistics & Probability Letters*, 79(3), 295-303.

Gomes, M.I., & Pestana, D. (2007). A simple second-order reduced bias tail index estimator. *Journal of statistical computation and simulation, 77(6),* 487-502.

Gusak, D., Kukush, A., Kulik, A., Mishura, Y. & Pilipenko, A. (2010). Theory of stochastic processes: with Applications to Financial Mathematics and Risk theory. Springer Science & Business Media, ISBN 978-0-387-87862-1.

Hall, P. (1982). Great planning disasters: with a new Introduction (vol. 1). Univ. of California press.

Hall, P. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis*, 32(2), 177-203.

Hall, P. and Welsh, A.H. (1985). Adaptive Estimates of Parameters of Regular Variation. *The Annals of Statistics*, 331-341.

Hashemifard, Z., Amindavar, H. & Amini, A. (2016). Parameters Estimation for Continuous-Time Heavy-Tailed Signals Modeled By α- Stable Autoregressive Processes. *Digital signal processing*, 57, 79-92.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 1163-1174.

Hull, J. (2009). Options, futures, and other derivatives / John C. Hull. Upper Saddle River, NJ: prentice Hall,.

Iacus, S. M. (2009). Simulation and Inference for stochastic Differential Equation with R Examples. Springer Science & Business Media.

Iacus, S. M. (2011). Option Pricing and Estimation of Financial Model with R. New York: Wiley.

Ross, S. M. (2014). Introduction to probability Models. Academic press.Steven, R. (2016).

Stochastic Processes and Advanced Mathematical Finance. *The Definition of Brownian Motion and the Wiener process, Department of Mathematics*, 203, 68588-0130.

Imkeller, P. & Schmalfuss, B. (2001). The conjugacy of stochastic and random differential equations and the existence of global attractors. *Journal of Dynamics and Differential Equations*, 13(2), 215-249.

Jia, M. (2014). Heavy-tailed phenomena and tail index inference (Doctoral dissertation, University of Trento).

Jordanova, P. K. & Petkova, M. P. (2017). Measuring heavy-tailedness of distributions. In A/P conference proceedings (vol. 1910, No. 1, p.060002).

Kallgren, C. (2019). Simulation and statistical method for stochastic differential equations. University of Gothenburg, Sweden.

Ken-Iti, S. (1999). Levy processes and infinitely divisible distributions. Cambridge University press.

Ken-Iti, S. (1999). Lévy processes Infinitely Divisible distributions. Cambridge University Press.

Kessler, M.; Lindner, A. and Sorensen, M. (2012). Statistical Methods for Stochastic Differential Equations (Monographs on Statistical & Applied Probability).

Khintchine, A. (1937). A new dervation of one formula by Lévy P. *Bulletin of Moscow State University*, 1(1), 1.

Klebaner, F. C. (2012). Introduction to Stochastic Calculus with Applications.World Scientific publishing company.

Korn, R., Kom, E. & Kroisandt , G. (2010). Monte Carlo Methods and Modeling in Finance and Insurance. CRC press.

Kyprianou, A. E. (2014). Fluctuattions of Lévy processes with Applications. Springer Science & Business Media.

Lindgren, G., Rootzen, H. & Sandsten, M. (2013). Stationary stochastic processes for scientists and engineers. CRC press.

Mandelbrot, B. B. (1997). The Variation of Certain Speculative Prices. In Fractals and scaling in finance (pp. 371-418). Springer, New York, NY.

Mason[a], D. M. (1982). Law of Large Number for Sums of Extreme Values. *Annals of Probability*, 754-764.

Mason[b], I. (1982). A Model for Assessment of Weather Forecasts. Aust. Meteor. Mag, 30(4), 291-303.

Mikosch, T. (1998). Elementary Stochastic Calculus With Finance in View. World Scientific.

Mikosch, T. (1999). Regular Variation, Sub exponentiality and Their Applications in Probability Theory (vol.99). Eindhoven, the Netherlands: Eindhoven University of Technology.

Mikosch, T. (2004). Elementary Stochastic Calculus (5th ed.), World Scientific.

Mousa, M. (2016). Methods of estimating Model Parameters Lévy processes and emoloyment in returns stock prices North and United Bank**.** Baghdad University.

Nemeth, L. & Zempleni, A. (2020). Regression estimation for the tail index. *Journal of Statistical Theory and Practice*, 14(3), 1-23.

Onalan, O. (2009**)**. Financial Modelling with Ornstein-Uhlenbeck Processes Driven by Lévy Process. *In proceedings of the world congress on Engineering* (vol. 2, pp. 1-3).

Peng, L. & Qi, Y. (2017). Inference for Heavy-Tailed Data: Applications in Insurance and Finance. *Academic press*.

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

Qi, Y. (2008). Bootstrap and empirical likelihood methods in extremes**.** *Extremes*, 11(1), 81-97.

Qi, Y. (2008). Bootstrap and empirical likelihood methods in extremes. *Extremes*, 11(1), 81-97.

Resnick, S. & Starica, C. (1997). Smoothing The Hill Estimator. Advance in Applied Probability, 271-293.

Resnick, S. I. (2007). Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Science & Business Media.

Rolski, T., Schmidli, H., Scmidt, V. & Teugels, J. L. (2009). Stochastic Processes for Insurance and Finance. (Vol. 505). *John wiley & Sons*.

Ross, S. M. (2014). Introduction to stochastic dynamic programming. Academic press.

Shorokhod, V. (2005). Basic principles and Aplications of robability theory. Springer Science & Business Media.

Tankov, p. (2003). Financial modeling with jump processes. Chapman and Hall/CRC.

Teugels, J. L. (1975). The Class of Sub exponential Distributions. *The Annals of probability*, 1000-1011.

Wireman, A. (2014). Catastrophes, conspiracies, and Sub exponential Distributions (Part III), (http:// rigorandrelevance.Wordpress.com/catastrophes-conspiracies-and-Sub exponential-distributions-part-iii/).Rigor + Relevance blog. RSRG, Caltech.

المستخلص

التوزيعات ذات الذيل الثقيل هي فرع مهم جدا من التحليل الإحصائي. يحدث الذيل الثقيل بسبب وجود بعض القيم المتطرفة في التوزيعات. في هذه الرسالة، قدرنا معلمة الذيل المستقرة مع $(2 \leq \alpha \leq \mathbf{0})$ لمجموعة من المشاهدات المستقلة والمتطابقة التوزيع باستخدام ثلاث طرق لا معلميه (المباشرة والبوتستراب والدبل بوتستراب). تمت المقارنة بين الطرق من أجل اختيار الأفضل من بينها والذي يمثل اصغر متوسط مربع للخطأ. ذكرنا الحركة البروانية الهندسية وعملية ليفي كمثالين مشهورين للمعادلات التفاضلية العشوائية التي تم استخدامها لتوليد البيانات في المحاكاة. تم تطبيق طرق تقدير معلمة الذيل في المحاكاة والبيانات الحقيقية لمجموعة البيانات اليومية لسوق الأوراق المالية العراقية باستخدام برنامج R.

جمهورية العراق

وزارة التعليم العالي والبحث العلمي

جامعة القادسية

كلية الإدارة والاقتصاد

قسم الإحصاء

بعض الطرائق اللامعلميه لتقدير مؤشر الذيل الثقيل في المعادلات التفاضلية العشوائية لسوق الأوراق المالية العراقية

رسالة ماجستير مقدمة إلى مجلس كلية الإدارة والاقتصاد/ جامعة القادسية وهي جزء من متطلبات نيل درجة الماجستير في علوم الإحصاء

من قبل الطالبة

نور عبد حسن

إشراف

أ.م.د مهند فائز السعدون

1442 هـ **2021** م