Identifying jumps variations in high-frequency time series

William Gonzalo Rojas Durán

Tese apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Doutor em Ciências

Programa: Estatística

Orientador: Prof. Dr. Pedro Alberto Morettin

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES-PDSE/CNPq

São Paulo, fevereiro de 2019

Identifying jumps variations in high-frequency time series

Esta é a versão original da tese elaborada pelo candidato William Gonzalo Rojas Durán, tal como submetida à Comissão Julgadora.

Identifying jumps variations in high-frequency time series

Esta versão da tese contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 25/03/2019. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Pedro Alberto Morettin (orientador) - IME-USP
- Prof. Dra. Chang Chiann - IME-USP
- Prof. Dr. Aluisio S. Pinheiro - IMECC-UNICAMP
- Prof. Dr. Ronaldo Dias - IMECC-UNICAMP
- Prof. Dr. Pedro L. Valls Pereira - FGV-SP

Acknowledgements

Firstly, I would like to thank my advisor Pedro A Morettin for his support during these four years of excellent guidance. I would also like to thank Professor Ruey S. Tsay, for the experience gained in Chicago and for the opportunity to have worked as an research assistant at the University of Chicago.

Second, I want to thank my mother, sisters and nieces for the encouragement and support received, without them would not have so much motivation.

Also, I would like to thank my friend Fabio Ortolano and my friends from IME-USP, Ana Paula, Andressa Cerqueira, Andressa Siroky, Joscelanio and Elisangela for the study talks and the good times shared.

Finally, this work was conducted during a scholarship supported by the International Cooperation Program CAPES/PDSE at the University of Chicago. Financed by CAPES-Brazilian Federal Agency for Support and Evaluation of Graduate Education within the Ministry of Education of Brazil.

Abstract

ROJAS, W. G. **Identifying jumps variations in high-frequency time series**. 2018. Tese (Doutorado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2018.

Stochastic models based on diffusions are often used to describe complex dynamical systems in biology, engineering, finance, physics etc. However, these models, when applied in finance, for example, do not take into account possible price jumps during a business session on a stock exchange due to the arrival of market information. In diffusion models, price movements are conditionally Gaussian, so large and sudden movements do not occur. On the other hand, in practice, price jumps can give rise to substantial losses or gains. Therefore, it is important to analyze the functional volatility for high frequency data, taking into account the presence of these jumps. This work consist of two parts. The first part refers to detection of jumps in a time series using wavelets. The second part is devoted to studying a test statistic of the Cramér-Von Mises type test statistic to identify variations in time series jumps with high frequency data. The main result and contribution of this study shows that the distribution function of the proposed test statistic follows approximately a gamma distribution. This is of vital importance because it enables us to determine the critical region for the rejection of the null hypothesis of interest. We observe better results in comparison with the Kolmogorov-Smirnov (KS) test. Specifically, we show that the power and the error rate of the test using Cramér-von Mises (Cv-M) statistic is better than those using the KS test statistic, showing a higher detection power and lower error rate. We applied the proposed test to three real data sets, namely, the stock returns of Google, Apple and Goldman Sachs (GS), and found that the proposed test can capture the dynamics of the series.

Keywords: Itô semimartingale, Activity Jumps, Variation Jumps, Distribution function, Cramér-von Mises.

Resumo

ROJAS, W. G. **Identificação de variações em séries de tempo com saltos em dados de alta frequência**. 2018. Tese (Doutorado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2018.

Modelos estocásticos baseados em difusões são usados frequentemente para descrever sistemas dinâmicos complexos em biologia, engenharia, finanças, física etc. Contudo, esses modelos, quando aplicados em finanças, por exemplo, não levam em conta possíveis saltos nos preços durante uma sessão de negócios em uma bolsa de valores devido à chegada de informações do mercado. Nos modelos de difusão, os movimentos dos preços são condicionalmente gaussianos, portanto movimentos grandes e repentinos não ocorrem. Por outro lado, nos modelos que incorporam saltos, estes podem dar origem a grandes perdas ou ganhos . Torna-se importante, portanto, a análise da volatilidade funcional para dados de alta frequência, levando-se em conta a presença desses saltos. Este trabalho consiste em duas partes. A primeira parte refere-se à detecção de saltos em uma série temporal usando wavelets. A segunda parte é dedicada ao estudo de uma estatística de teste do tipo Cramér-von Mises para identificar variações em séries tempo com saltos em dados de alta freqüência. O principal resultado e contribuição deste estudo mostra que a função de distribuição da estatística de teste proposta segue aproximadamente uma distribuição Gamma. Isto é de vital importância porque permite determinar a região critica para a rejeição da hipotese nula de interesse. Encontramos melhores resultados em comparação com o teste de Kolmogorov-Smirnov. Especificamente, mostramos que a taxa de erro e o poder do teste, usando a estatística de teste Cramér-von Mises (Cv-M) é melhor do que a estatística de teste Kolmogorv-Smirnov (KS), mostrando um alto poder de detecção e baixa taxa de erro. Aplicamos o teste proposto a três conjuntos de dados reais, retornos da Google, Apple e Goldman Sachs (GS) e encontramos que o teste proposto captura a dinâmica das series.

Palavras-chave: Itô semimartingal, Saltos, variação de saltos, função de distribuição, Cramér von Mises.

Contents

List of Abbreviations

cdf	cumulative distribution function
q	quantile
Cv-M	Cramér-von Mises
KS	Kolmogorov-Smirnov
TA1	Test statistic using the truncate variation estimator
TA2	Test statistic using the local estimator
RV	Realized Volatility
RVBP	Realized Volatility bi-power
GS	Goldman Sachs
ARCH	Autoregressive Conditional Heterocedasticity
FDA	Functional Data Analysis
RB	white noise
ROC	Receiver Operating Characteristic Curve

List of Symbols

ω	Angular frequency
ψ	*Wavelet* analysis function
Ψ	Fourier transform ψ
Φ	Cumulative normal distribution function
W	Brownian motion
Δ	Increments
$N(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
$O(\cdot)$	order of magnitude
$G(a, b)$	Gamma distribution with shape parameter a and scale parameter b

List of Figures

List of Tables

Chapter 1

Introduction

Stochastic models based on diffusions are often used to describe the complex dynamical systems in biology, engineering, finance, and physics. However, these models, when applied in finance, do not take into account possible price jumps during a business session on a stock exchange due to the arrival of market information. In diffusion models, price movements are conditionally Gaussian, so large and sudden movements do not occur. On the other hand, in practice, price jumps can give rise to substantial losses or gains.

The volatility of financial time series has been treated in two ways. The first approach is parametric and postulates a latent variable via conditional heteroscedastic models to describe the volatility, such as the ARCH (Autoregressive Conditional Heterocedasticity) family models proposed by Engle (1982) and the stochastic volatility models. There is a huge literature on this approach and for details, see, for example, Tsay (2005) or Morettin (2017). This parametric approach is used mainly for low frequency data (daily, weekly, monthly) of financial asset returns. Recently, with the availability of intraday data (also called high-frequency data), a second approach, which is nonparametric, was considered. The main motivation for using the nonparametric approach is because parametric models often fail to adequately capture the movements of intraday volatility (see Andersen and Labys (2003)).

A nonparametric approach consists of constructing the daily realized volatility (RV) of a financial series using intraday returns, sampled at intervals Δt, of the order of 5 or 15 minutes, for example. RV is obtained from sums of squares of intraday returns (see Andersen and Labys (2003) and Barndorff-Nielsen and Shephard (2002)). Another method is to construct the realized bi-power variance (RVBP), from the sums of cross products of absolute adjacent, properly scaled returns (Barndorff-Nielsen and Shephard (2006)). For details, see Ait-Sahalia and Jacod (2014).

Fan and Wang (2008), Barndorff-Nielsen and Shephard (2006) and Ait-Sahalia and Jacod (2009b) derived test statistics for detecting the existence of jumps. Other authors established tests for the necessity of adding a Brownian force, see Ait-Sahalia and Jacod (2010), Jing and Liu (2012b), Kong and Jing (2015), Todorov and Tauchen (2014) and Todorov (2015). Ait-Sahalia and Jacod (2011) studied whether the jump component is of finite activity when the Brownian force is present. Recently, Kong (2017) studied whether it is necessary to add an infinite variation jump term in addition to a continuous local martigale, using a Kolmogorov-Smirnov type test statistic.

In recent years, marked efforts have been devoted to study the distribution function of

the jump test statistics. Csorgo and Faraway (1996) showed the exact and asymptotic distribution of Cramér-von Mises statistic when the empirical distribution function is a uniform distribution function. Todorov and Tauchen (2014) suggested using other measures of discrepancy between distributions like the Cramér-von Mises test for the presence of diffusive component in X_t.

However, finding the distribution function of a jump test statistic is not easy so that numerical methods are often employed. In this study, we propose a new test statistic of the modified Cramér-von Mises type.
This work consists of two parts. The first part refers to the detection of jumps in a time series using wavelets. The second part is devoted to studying a Cramer-von Mises type test statistic to identify variations in time series jumps with high frequency data.

1.1 Setup 1

Stochastic models based on diffusions are often used to describe complex dynamical systems as stated above. It follows that there is a growing demand for the development of inferential methods for these models, for example (Prakasa Rao (1999)).

The problem to consider is

- detect the jumps in the trajectories described by a process of diffusion with jumps, which can be done by methods that use wavelets (Wang (1995)).

A diffusion process with jumps has the form

$$X_t = X_0 + \int_0^t \tilde{\mu}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \sum_{i=1}^{N_t} Z_i, \tag{1.1}$$

where the second term corresponds to drift, the third to the diffusion, and the fourth to the jumps of X_t. Here, N_t is a counting process (Poisson, for example), that represents the number of jumps between 0 and t, Z_i represents the size of the "i"-th jump and W_s represents a Wiener process. The log-prices of (1.1) has a quadratic variation

$$[X, X]_t = \int_0^t \tilde{\sigma}_s dWs + \sum_{i=1}^{N_t} Z_i^2,$$

with two parts: integrated volatility and variation of jumps,

$$\Theta = \int_0^t \tilde{\sigma}_s dWs, \qquad \Phi = \sum_{i=1}^{N_t} Z_i^2.$$

One purpose is:

- detect jumps and estimate Θ and Φ. For this, we can use techniques via wavelets, presented by Wang (1995) and Raimondo (1998), techniques for estimating jumps by Fan and Wang (2007) and the development by Muller and Stadtmuller (2011) for diffusions.

1.2 Setup 2

Let (Ω, F_t, F, P) be a filtered probability space. The standard jump-diffusion model used for modeling many stochastic processes is an Itô semimartingale given by the following differential equation:

$$dX_t = \alpha_t dt + \sigma_t dW_t + dZ_t, \tag{1.2}$$

where α_t and σ_t are processes with cadlag paths, W_t is a standard Brownian motion, and Z_t is an Itô semimartingale process of pure-jump type. Todorov and Tauchen (2014) provide a formal setup and assumptions. They generalize the setup (1.2) to accommodate the alternative hypothesis in which X_t can be of pure-jump type. Itô semimartingale plays an important role in stochastic calculus and we make the following assumption:

Assumption 1. *Suppose that Y_t follows a non-parametric volatility model:*

$$Y_t = X_t + \epsilon_t, \ t \in [0, 1], \tag{1.3}$$

in which X_t is a continuous Itô semimartingale, that is,

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + Z_t, \tag{1.4}$$

where $\int_0^t b_s ds$ is the drift term and cadlag process, $\int_0^t \sigma_s dW_s$ is a continuous local martingale with σ_s being an adapted process, W_s is a standard Brownian motion, and Z_t is a skewed β-stable Lévy process.

Kong (2017) for the first time provided a theoretical test for the presence of infinite variation jumps in the simultaneous presence of a diffusion term and a jump component of finite variation and established the asymptotic theory of the empirical distribution of the "devolatilized" increments of Itô semimartingales with infinitely active or even infinite variation jumps. To estimate the volatility, he uses the local method and splits the interval into non-overlapping shrinking blocks consisting of $2k_n$ intervals of length Δ_n, where k_n is some integer depending on n. Here $\Delta_i^n X = X_{t_i} - X_{t_{i-1}}$ denote the ith one-step increment for $1 \leq i \leq n$. The local estimate of $\sigma_j^2 = \sigma_{2jk_n\Delta_n}$ is

$$\hat{\sigma}_j^2(u) = c_j(u) - \frac{1}{u^2 k_n}(\sinh(u^2 c_j(u)))^2, \tag{1.5}$$

where

$$\begin{aligned} c_j(u) &= -\frac{1}{u^2}\log\left(L_j(u) \vee \frac{c}{\sqrt{k_n}}\right), \ 0 \leq j \leq \lfloor n/(2k_n)\rfloor - 1, \\ L_j(u) &= \frac{1}{k_n}\sum_{l=1}^{k_n} \cos\left(u\frac{\Delta_{2jk_n+2l}^n X - \Delta_{2jk_n+2l-1}^n X}{\sqrt{\Delta_n}}\right), \end{aligned}$$

where the subtracted term in (1.5) is used to correct the bias due to the jumps, $a \vee b = \max(a, b)$, the lower threshold $\frac{c}{\sqrt{k_n}}$ in $c_j(u)$ with constant c is to assure that the logarithmic function is well-defined, and $\lfloor m \rfloor$ denotes the integer part of m.

There are other methods to estimate the spot volatility. See, for instance, Todorov and Tauchen (2014) and Fan and Wang (2007). The major advantage of this Laplace-transform-based lo-

cal estimator is that it can easily separate the effect of the Brownian force and the stable-like driving force. For properly chosen m_n and u_n the empirical distribution function of the devolatilized increments is defined as

$$\hat{F}_n(u_n,\tau) = \frac{1}{\lfloor n/2k_n \rfloor m_n} \sum_{j=1}^{[n/2k_n]} \sum_{i=2jk_n+1}^{2jk_n+m_n} I\left(\frac{\Delta_i^n X}{\sqrt{\hat{\sigma}_j^2(u_n)\Delta_n}} \leq \tau \right), \tag{1.6}$$

for $\tau \in \mathbb{R}$, for details, see Kong (2017). Furthermore, the author defines the empirical process as

$$\hat{Y}_n(\tau) = \sqrt{\lfloor n/2k_n \rfloor m_n}(\hat{F}_n(u_n,\tau) - \Phi(\tau)), \tag{1.7}$$

where $\Phi(\tau)$ denotes the c.d.f. of a standard normal random variable.

In this thesis, we consider the following hypotheses

$$\begin{aligned} H_0: &\ \Delta_i^n X \ \sim \textbf{standard normal model}; \\ H_1: &\ \Delta_i^n X \ \sim \textbf{jump-diffusion model}. \end{aligned}$$

Along the work, we assume that the available data set $\{X_{t_j}; 0 \leq j \leq n\}$ which are discretely sampled from X, and are equally spaced in the fixed interval $[0,T]$, i.e, $t_j = j\Delta_n$ with $\Delta_n = T/n$ for $0 \leq j \leq n$.

Our aim is to show that the Cramer-Von Mises statistic to be defined in Section (1.3.2) has better properties in terms of size and power compared to the Kolmogorov Smirnov test.

1.3 Basic concepts

1.3.1 Wavelets

Wavelets are functions that satisfy certain properties. They can be smooth or not and can have simple mathematical expressions or not. Before starting the basic ideas about the wavelets, we will make some comments in relation to the analogies and differences between two analyses, fourier and wavelets. An analogy is that given an integrable square function, it can be written as a edge overlay of sines and cosines or wavelets. The difference is that the functions of a base of wavelets are indexed by two parameters, whereas on the basis of Fourier we have a single parameter, λ, which has the physical interpretation of frequency (see Morettin (2014)).

By analogy with Fourier analysis, consider the space $L^2(\mathbb{R})$ of all measurable square functions integrable on $\mathbb{R}$. Here, the functions $f(t)$ must fall to zero, when $|t| \to \infty$. Therefore, the exponentials do not belong to this space.

The ψ and $\psi_{j,k}$ functions satisfy certain properties, as follows:

- $\int_{-\infty}^{\infty} \psi(t)dt = 0$ (Admissibility).
- $\int_{-\infty}^{\infty} |\psi(t)|dt < \infty$.
- $c_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty$, where $\Psi(\omega)$ is the Fourier transform of $\psi(t)$. A required condition for c_ψ assert is that $\Psi(0) = 0$, which is equivalent to the admissibility property.

- $\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1$ or $\int_{-\infty}^{\infty} |\Psi(\omega)|^2 d\omega = 2\pi$. The first $r-1$ moments of ψ cancel out, that is, $\int_{-\infty}^{\infty} t^j \psi(t) dt = 0,\ j = 0, 1, \ldots, r-1$, for some $r \geq 1$ and $\int_{-\infty}^{\infty} |t^r \psi(t)| dt < \infty$.

The value of r is linked to the degree of smoothness (regularity) of ψ: The higher r, the smoother will be ψ. Some wavelets have compact support, as we will see below, which is a desirable property and have to do with the fact that the wavelets are localized in time. Not all wavelets generate orthogonal systems, for example the Mexican hat, also defined below. In the case of wavelets with compact support, the value of r is also related to the wavelet support. We will see that one way to generate wavelets is by the scale function, or father wavelet, ϕ, which is a solution of equation

$$\phi(t) = \sqrt{2} \sum_k l_k \phi(2t - k). \tag{1.8}$$

This function generates an orthonormal family in $L^2(\mathbb{R})$,

$$\phi_{j,k}(t) = 2^{j/2} \phi(2^j t - k),\ j, k \in \mathbb{Z},\ \mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}. \tag{1.9}$$

Under this conditions, ψ can be obtained from ϕ by

$$\psi(t) = \sqrt{2} \sum_k h_k \phi(2t - k), \tag{1.10}$$

where $h_k = (-1)^k l_{1-k}$ is called quadrature mirror relation. As a matter of fact, l_k and h_k are low-pass and high-pass filter coefficients, respectively, used to calculate the discrete wavelet transform. These coefficients are given by

$$l_k = \sqrt{2} \int_{-\infty}^{\infty} \phi(t) \phi(2t - k) dt, \tag{1.11}$$

$$h_k = \sqrt{2} \int_{-\infty}^{\infty} \psi(t) \phi(2t - k) dt. \tag{1.12}$$

Equations (1.8) and (1.10) are called expansion equations. The ϕ wavelets generate certain subspaces of a multiresolution analysis, for details see Morettin (2014).

It can be shown that $\sum_k l_k = \sqrt{2}$, $\sum_k h_k = 0$, $\sum_k l_k^2 = 1$ and $\sum_k l_k h_{k-2m} = 1$ if $m = 0$, and equal to zero, otherwise. It is convenient, then, to consider the orthonormal system

$$\{\phi_{j_0,k}(t), \psi_{j,k}, j, k \in \mathbb{Z}, j \geq j_0\} \tag{1.13}$$

so that we can write, for $f(t) \in L^2(\mathbb{R})$,

$$f(t) = \sum_k c_{j_0,k} \phi_{j_0,k}(t) + \sum_{j \geq j_0} \sum_k d_{j,k} \psi_{j,k}(t), \tag{1.14}$$

in which

$$c_{j_0,k} = \int_{-\infty}^{\infty} f(t) \phi_{j_0,k} dt, \tag{1.15}$$

$$d_{j,k} = \int_{-\infty}^{\infty} f(t)\psi_{j,k}dt. \tag{1.16}$$

In (1.14), j_0 is the lowest resolution scale (coarsest scale).

1.4 Organization

The rest of the work is organized as follows.

- Chapter 2 is dedicated to study the functional volatility. First, we introduce some properties to be able to show the asymptotic value of the jump term in the stochastic differential equation. Second, we present a method of detection of jump, already existing and its respective estimation. We use Holder's inequality to prove that $|TJ(a,t)|$ is bounded.

- Chapter 3 is devoted to presenting the empirical cumulative distribution function and estimator for the volatility. Specifically, we used the local estimator proposed by Todorov and Tauchen (2014) to estimate σ_t^2. We present two functions of cumulative distribution with their respective estimators for the variance and we show the performance of each one. Then, in Section 3.3, the specific case of the quantile analysis, here we show our quantile for different sample sizes and $\alpha = 5\%$; our chosen cumulative distribution function and we show that it is approximately a gamma distribution.

- Chapter 4 is used for our simulation study of both parts of this thesis. First, we use the method proposed by Wang (1995) to detect and locate jumps at different scales. In Section 4.2 we present the table for different probability values and sample sizes. Specifically, we conduct simulations studies to check the performance of the test.

- Finally, Chapter 5 is dedicated to show the real data analysis, for case 1 and 2. In Section 5.1 we present the application for Google Stock the period considered is from November 11th to November 12th in 2014 and in Section 5.2 we collet intraday transaction price of the Google, Apple and Goldman Sachs Index, from November 11th to November 12th in 2014, with a sampling frequency up to every 15 seconds.

Chapter 2

Jump Detection

The classical diffusion model, used by Black and Scholes (1973) to model prices, is given by

$$\frac{dX(t)}{X(t)} = \mu dt + \sigma dW(t), \quad t \geq 0, \tag{2.1}$$

where $W(t)$ is the standard Wiener process (standard Brownian motion), $\sigma > 0$ is the volatility and μ is the term for "drift", both supposed to be independent of time. This model is simplified and does not reflect facts observed in the data, for example volatility varying in time. In addition, prices are not obtained continuously, but in a regular grid of values $t_j = j\Delta$, $j = 1, 2, \ldots [T/\Delta]$, being $[0, T]$ the interval at which the process is observed, for example $\Delta = 5$ min. Barndorff-Nielsen and Shephard (2002) suggest the model

$$d \log X(t, \omega) = \mu dt + \beta\sigma^2(t, \omega)dt + \sigma(t, \omega)dW(t, \omega), \tag{2.2}$$

where β is the "risk premium", σ is a stationary process ("spot volatility"), that can be modeled, *e.g*, by a process of Ornestein-Uhlenbech. Muller and Stadtmuller (2011) propose a variant of equation (2.2) within the focus of a general diffusion model with random drift, and the observations constitute a sample of n process realizations

$$d \log X(t, \omega) = \tilde{\mu}_i(t, \omega)dt + \tilde{\sigma}_i(t, \omega)dW_i(t, \omega), \quad 0 \leq t \leq T, \ i = 1, \ldots, n. \tag{2.3}$$

Here, $\tilde{\mu}_i$ and $\tilde{\sigma}_i$ are copies i.i.d of stochastic processes $\tilde{\mu}, \tilde{\sigma}$, supposed smooth but not stationary and W_i are independent Wiener processes. The availability of multiple copies is crucial for the application of Functional Data Analysis (FDA) techniques. Let $Z_\Delta(t)$ and $W_\Delta(t)$ the returns obtained from the prices $X(t)$ and the discretized diffusion terms, respectively. The model (2.3) can be written as

$$\begin{aligned} Z_\Delta(t) &= \frac{1}{\sqrt{\Delta}} \int_t^{t+\Delta} \tilde{\mu}(v)dv + \frac{1}{\sqrt{\Delta}} \int_t^{t+\Delta} \tilde{\sigma}(v)dW(v) \\ &= \tilde{\mu}(t)\sqrt{\Delta} + \tilde{\sigma}(t)W_\Delta(t) + R_1(t, \Delta) + R_2(t, \Delta). \end{aligned} \tag{2.4}$$

The authors prove that

$$Z_\Delta(t) \approx \tilde{\sigma}(t)W_\Delta(t). \tag{2.5}$$

The purpose is to estimate the process of functional volatility

$$V(t) = \log[\tilde{\sigma}(t)]^2,$$

from the data $Y_{i,j\Delta}$, where

$$\log([Z_\Delta(t_j)]^2) - q_0 \approx Y_\Delta(t_j) = V(t_j) + U_\Delta(t_j), \tag{2.6}$$

with $q_0 \approx -1.27$, $U_\Delta(t) = \log([W_\Delta]^2) - q_0$.
The procedure is based on the analysis of principal components applied to the data $Y_{i,j\Delta}$ using weighted linear local smoothing. The prediction of future volatilities is made using functional linear regression (see Muller and Stadtmuller (2011)).

2.1 Asymptotic volatility model

Consider (2.4) plus jumps, that is, let $M_\Delta(t)$ and $W_\Delta(t)$ the returns obtained from prices $X(t)$ and the discretized diffusion terms, respectively.

$$\begin{aligned}
M_\Delta(t) &= \frac{1}{\sqrt{\Delta}}\int_t^{t+\Delta}\tilde{\mu}(v)dv + \frac{1}{\sqrt{\Delta}}\int_t^{t+\Delta}\tilde{\sigma}(v)dW(v) + \frac{1}{\sqrt{\Delta}}\int_t^{t+\Delta}\sum_{i=1}^{N_t} Z_i ds \\
&= \tilde{\mu}(t)\sqrt{\Delta} + \frac{1}{\sqrt{\Delta}}\int_t^{t+\Delta}\tilde{\mu}(v)dv - \tilde{\mu}(t)\sqrt{\Delta} + \tilde{\sigma}(t)W_\Delta(t) \\
&+ \frac{1}{\sqrt{\Delta}}\int_t^{t+\Delta}\tilde{\sigma}(v)dW(v) - \tilde{\sigma}(t)W_\Delta(t) + Z_{N_{t+1}}\sqrt{\Delta} + \frac{1}{\sqrt{\Delta}}\int_t^{t+\Delta}\sum_{i=1}^{N_t} Z_i ds \\
&- Z_{N_{t+1}}\sqrt{\Delta} \\
&= \tilde{\mu}(t)\sqrt{\Delta} + \tilde{\sigma}(t)W_\Delta(t) + Z_{N_t+1}\sqrt{\Delta} + R_1(t,\Delta) + R_2(t,\Delta) + R_3(t,\Delta).
\end{aligned}$$

The quantity $Z_{N_{t+1}}\sqrt{\Delta}$ in $M_\Delta(t)$ was obtained as shown in Figure 2.1.

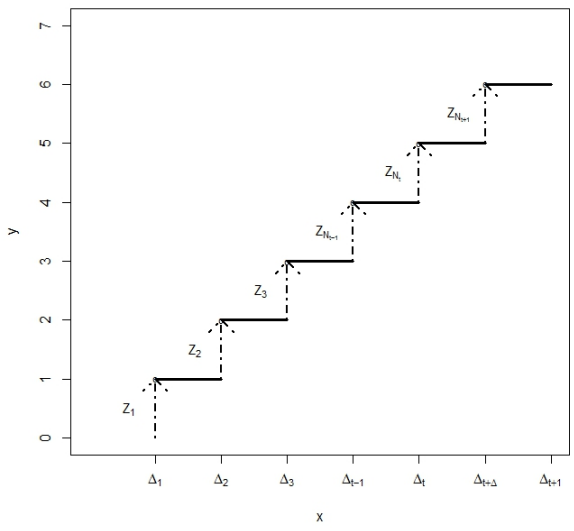


Figure 2.1: *Simulation of the jumps in a fixed time grid.*

We calculate the area until t,

$$Z_1 \times \Delta + (Z_1 + Z_2) \times \Delta + \ldots + \left(Z_1 + \ldots + Z_{N_{t+1}}\right) \times \Delta$$

and the area until $t+\Delta$,

$$Z_1 \times \Delta + (Z_1 + Z_2) \times \Delta + \ldots + \left(Z_1 + \ldots + Z_{N_{t+1}}\right) \times \Delta + Z_{N_{t+1}}\Delta.$$

Thus,

$$\int_t^{t+\Delta} \sum_{i=1}^{N_t} Z_i ds \approx Z_{N_{t+1}}.\Delta \tag{2.7}$$

Multiplying by $\frac{1}{\sqrt{\Delta}}$ on both sides, we have

$$\begin{aligned} \frac{1}{\sqrt{\Delta}} \int_t^{t+\Delta} \sum_{i=1}^{N_t} Z_i ds &\approx \frac{1}{\sqrt{\Delta}} Z_{N_{t+1}} \Delta \\ &\approx Z_{N_{t+1}} \sqrt{\Delta}. \end{aligned}$$

The terms $R_1(t,\Delta), R_2(t,\Delta)$ and $R_3(t,\Delta)$ are uniformly small and therefore can be disregarded asymptotically.

Assumptions:

(M1) Process $(\tilde{\mu}(t))_{t\in[0,T]}$ and $(\tilde{\sigma}(t))_{t\in[0,T]}$ in (2.3) have trajectories that are uniformly continuous Lipschitz of order 1, *i.e.*, there are constant $L_1 > 0$ and $L_2 > 0$ such that

$$\begin{aligned} |\tilde{\mu}(t) - \tilde{\mu}(s)| &< L_1 |t-s| \ \ and \\ |\tilde{\sigma}(t) - \tilde{\sigma}(s)| &< L_2 |t-s| . \end{aligned}$$

(M2) Process $\tilde{\mu}$ satisfy $E\left[sup_{t\in[0,T]} |\tilde{\mu}(t)|\right] < \infty$

(M3) Process $\tilde{\sigma}$ satisfy $sup_{t\in[0,T]}\ E\left[|\tilde{\sigma}(t)^2|\right] < \infty$
(M4) Trajectories of the process $\tilde{\sigma}$ are smooth and the derivative $\frac{d}{dt}\tilde{\sigma}(t)$ satisfy

$$E\left[sup_{t\in[0,T]} \left|\frac{d}{dt}\tilde{\sigma}(t)\right|^2\right] = O(1).$$

(M5) The process Z_t satisfy $E\left[sup_{t\in[0,T]} |Z_t|\right] < \infty$.

Now, let us see that the terms $R_1(t,\Delta), R_2(t,\Delta)$ and $R_3(t,\Delta)$ are uniformly small and therefore can be disregarded asymptotically. For $R_1(t,\Delta)$ Muller and Stadtmuller (2011) proved that under (M1) and (M4):

1) $E\left[sup_{t\in[0,T]} |R_1(t,\Delta)|\right] = O\left(\Delta^{3/2}\right)$.

2) $R_2(t,\Delta) = O\left(\Delta^{1/2}\right)$.

Proposition 1: For the term $R_3(t,\Delta)$, we have

$$E\left[sup_{t\in[0,T]} |R_3(t,\Delta)|\right] = O\left(\Delta^{1/2}\right).$$

Proof:

$$E\left[sup_{t\in[0,T]}\left|R_3(t,\Delta)\right|\right] = E\left[sup_{t\in[0,T]}\left|\frac{1}{\sqrt{\Delta}}\int_t^{t+\Delta}\sum_{i=1}^{N_t}Z_i ds - Z_{N_{t+1}}\sqrt{\Delta}\right|\right].$$

Let

$$Z'_{N_{t+1}}\sqrt{\Delta} = \frac{1}{\sqrt{\Delta}}\int_t^{t+\Delta}\sum_{i=1}^{N_t}Z_i ds.$$

Then;

$$\begin{aligned}
& E && \left[sup_{t\in[0,T]}\left|Z'_{N_{t+1}}\sqrt{\Delta} - Z_{N_{t+1}}\sqrt{\Delta}\right|\right] \\
\overset{Triangular.ineq}{\overbrace{\leq}} & E && \left[sup_{t\in[0,T]}\left\{\left|Z'_{N_{t+1}}\sqrt{\Delta}\right| + \left|Z_{N_{t+1}}\sqrt{\Delta}\right|\right\}\right] \\
\overset{Sup.property}{\overbrace{\leq}} & E && \left[sup_{t\in[0,T]}\left|Z'_{N_{t+1}}\sqrt{\Delta}\right| + sup_{t\in[0,T]}\left|Z_{N_{t+1}}\sqrt{\Delta}\right|\right] \\
\overset{L.mean}{\overbrace{=}} & E && \left[sup_{t\in[0,T]}\left|Z'_{N_{t+1}}\sqrt{\Delta}\right|\right] + E\left[sup_{t\in[0,T]}\left|Z_{N_{t+1}}\sqrt{\Delta}\right|\right] \\
= & \sqrt{\Delta}E && \left[sup_{t\in[0,T]}\left|Z'_{N_{t+1}}\right|\right] + \sqrt{\Delta}E\left[sup_{t\in[0,T]}\left|Z_{N_{t+1}}\right|\right].
\end{aligned}$$

Let $E\left[sup_{t\in[0,T]}\left|Z'_{N_{t+1}}\right|\right] = E\left[sup_{t\in[0,T]}\left|Z_{N_{t+1}}\right|\right] + \epsilon(t,\Delta)$, thus,

$$\begin{aligned}
E\left[sup_{t\in[0,T]}\left|R_3(t,\Delta)\right|\right] &= \sqrt{\Delta}\left(E\left[sup_{t\in[0,T]}\left|Z_{N_{t+1}}\right|\right] + \epsilon(t,\Delta)\right) \\
&+ \sqrt{\Delta}E\left[sup_{t\in[0,T]}\left|Z_{N_{t+1}}\right|\right] \\
&= 2*\sqrt{\Delta}E\left[sup_{t\in[0,T]}\left|Z_{N_{t+1}}\right|\right] + \sqrt{\Delta}*\epsilon(t,\Delta).
\end{aligned}$$

From (M5) we have to $E\left[sup_{t\in[0,T]}\left|Z_{N_{t+1}}\right|\right] < \infty$ and $\exists t^* : \epsilon(t,\Delta) \leq \epsilon(t^*,\Delta) = f(\Delta)\forall t \in [0,T]$. Then,

$$\begin{aligned}
E\left[sup_{t\in[0,T]}\left|R_3(t,\Delta)\right|\right] &= O(\Delta^{1/2}) + O(\Delta^{1/2})*O(f(\Delta)) \\
&= O(\Delta^{1/2}) + O(\Delta^{1/2}f(\Delta)) \\
&= O(max\left\{\Delta^{1/2}, \Delta^{1/2}f(\Delta)\right\}) \\
&= O(\Delta^{1/2}max\left\{1, f(\Delta)\right\}).
\end{aligned}$$

Finally,

$$E\left[sup_{t\in[0,T]}\left|R_3(t,\Delta)\right|\right] = \begin{cases} O(\Delta^{1/2}), & \text{if } f(\Delta)\leq 1,\ \forall\Delta \\ O(\Delta^{1/2}f(\Delta)), & \text{if } f(\Delta) > 1,\ \forall\Delta \end{cases}$$

Note: 1) assuming that $Z \sim N(\mu,\sigma^2)$, then $|Z|$ has Half-normal distribution with $E\left[|Z'|\right] < \infty$ and $E\left[|Z|\right] < \infty$.
2) we can choose Δ = 5min for example; then $\Delta \leq K$ where K is a constant. Thus, $f(\Delta) = O(1)$.
As we had shown that $R_1(t,\Delta)$, $R_2(t,\Delta)$ and $R_3(t,\Delta)$ are uniformly small, the approximation

is

$$M_\Delta(t) \approx \tilde{\mu}(t)\sqrt{\Delta} + \tilde{\sigma}(t)W_\Delta(t) + Z_{N_{t+1}}\sqrt{\Delta} \tag{2.8}$$

Assumptions:

- $sup_{t\in[0,T]} |\tilde{\mu}(t)| = O_p(1)$
- $sup_{t\in[0,T]} \left|Z_{N_{t+1}}\right| = O_p(1)$.

Thus, we arrive at the same approximation of Muller and Stadtmuller (2011) given in (2.5):

$$\begin{aligned} M_\Delta(t) &\approx O_p(1) + \tilde{\sigma}(t)W_\Delta(t) + O_p(1) \\ &\approx O_p(max\{1,1\}) + \tilde{\sigma}(t)W_\Delta(t) \\ &\approx \tilde{\sigma}(t)W_\Delta(t) + O_p(1) \end{aligned}$$

Therefore, jumps do impact asymptotically in the estimation of volatility.

2.2 Jump analysis

As previously mentioned, one purpose of the thesis is to detect the jumps and for this we can use techniques via wavelets, presented by Wang (1995) and Raimondo (1998). Let ψ be a mother wavelet and define

$$\psi_{ab}(x) = a^{1/2}\psi\left(\frac{x-b}{a}\right),$$

the continuous wavelet transform of $f(x)$ is

$$Tf(a,b)(x) = \int \psi_{a,b}\left(\frac{x-b}{a}\right) f(x)dx.$$

$Tf(a,b)(x)$ is a function of scale or frequency, here a, b are real parameters. On small scales $Tf(a,b)(x)$ provides localized information about local regularity of $f(x)$; this local regularity is often measured with Lipschitz exponents.

Definition 3: The function $f(x)$ is said to be Lipschitz α in x_0, if there is $k > 0$, such that for $h \to 0$,

$$|f(x_0 + h) - f(x_0)| \leq k\,|h|^\alpha .$$

If $f(x)$ is Lipschitz at all points of $[0,1]$, then $f(x)$ is said to be uniformly Lipschitz α on $[0,1]$. The local and global regularity Lipschitz can be characterized by the asymptotic decay of the wavelet transform in small scales.

2.2.1 Jump detection

A diffusion process with jumps has the form given in $M_\Delta(t)$, and applying the wavelet transform to the model (2.1) plus the jump component, that is

$$\frac{dX(t)}{X(t)} = \mu dt + \sigma dW(t) + J(t), \quad t \geq 0, \tag{2.9}$$

we have

$$TX(a,t) = Tu(a,t) + \sigma TW(a,t) + TJ(a,t),$$

with $J(a,t) = \sum_{i=1}^{N_t} Z_i$. The wavelet transform of the white noise $W(ds)$ is given by

$$TW(a,t) = \int \psi_a(t-u)W(du),$$

and the wavelet transform of the jump component is

$$TJ(a,t) = \int \psi_a(t-u)J(u)du.$$

Steps to follow:

- At a certain scale a, we will find the orders of convergence of the transforms and verify the convergence speed, this will be verified making the transforms in fine and coarse scales;
- then the jumps will be detected by verifying the values of $TX(a,t)$.

Wang (1995), shows the following lemma to find the limiting factor for $|TW(a,x)|$.

Lemma 1: If ψ is differentiable with probability 1, there is a positive constant $K < \infty$ such that for all x and a small,

$$|TW(a,x)| \leq K \left|loga\right|^{1/2}.$$

See the Proof in Raimondo (1998).
For the case of the wavelet transformation corresponding to the jumps we have

$$TJ(a,t) = \int \psi_a(t-u)J(u)du.$$

We should find a bound for $|TJ(a,t)|$, for this we use the following Proposition.

Proposition 2: (Holder inequality) Let $p > 1$ and $q < \infty$ be, that is, $\frac{1}{p} + \frac{1}{q} = 1$. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ be functions $f \in L^p$, $g \in L^q$ and $V \subseteq D$, then

$$\left| \int_V f(x)g(x)dx \right| \leq \left(\int_V |f(x)|^p \, dx \right)^{1/p} \left(\int_V |g(x)|^q \, dx \right)^{1/q}.$$

We will prove the following proposition.

Proposition 3: Let $\psi : D \to \mathbb{R}$ and $J : D \to \mathbb{R}$ be functions $\psi \in L^2$, $J \in L^2$ and $V \subseteq D$, applying the Proposition 2 for $p = q = 2$ the conditions are satisfied and then $|TJ(a,t)|$ is bounded.

Proof: We have, applying the previous proposition,

$$\begin{aligned} |TJ(a,t)| &= \left| \int \psi_a(x-u)J(u)du \right| \\ &\leq \left(\int |\psi_a(x-u)|^2 \, du \right)^{1/2} \left(\int |J(u)|^2 \, du \right)^{1/2}, \end{aligned}$$

and we know that

$$\int_{\infty}^{\infty} |\psi(t)|^2 \, dt = 1. \tag{2.10}$$

We have that:

$$\int |\psi_a(x-u)|^2 \, du = \int \left|a^{1/2}\psi(x-u/a)\right|^2 du = a^{1/2} \int |\psi(x-u/a)|^2 \, du.$$

Therefore, from (2.10) we have $\left(\int |\psi(x-u/a)|^2 \, du\right) \leq K_1$, where K_1 is a constant. Thus,

$$|TJ(a,t)| \leq a^{1/2} K_1 \left(\int |J(u)|^2 \, du \right)^{1/2},$$

but

$$\int |J(u)|^2 \, du = \int \left| \sum_{i=1}^{N_t} Z_i \right|^2 du \leq \Delta^2 Z_{N_t+1}$$

or

$$\begin{aligned} \int |J(u)|^2 \, du &= \int \left| \sum_{i=1}^{N_t} Z_i \right|^2 du \\ &= \sum_i Z_i(u) \overline{\sum_j Z_j(u)} = \sum_i \sum_j Z_i(u) \overline{Z_j(u)} \end{aligned}$$

Finally,

$$|TJ(a,t)| \leq K_1 K_2 a^{1/2} \leq K a^{1/2},$$

where $K_2 = \Delta^2 Z_{N_t+1}$ and $K = K_1 K_2$. As previously stated, the jumps will be detected by checking the values of

$$TX(a,t) = Tu(a,t) + \sigma TW(a,t) + TJ(a,t).$$

We have:

- $|TW(a,t)| = O(|loga|^{1/2})$,
- $|TJ(a,t)| = O(a^{1/2})$.

That is, $TW(a,t)$ and $TJ(a,t)$ are of the orders $O(|loga|^{1/2})$ and $O(a^{1/2})$, respectively. Figure 2.2 shows the convergence speed for $a = 9$.

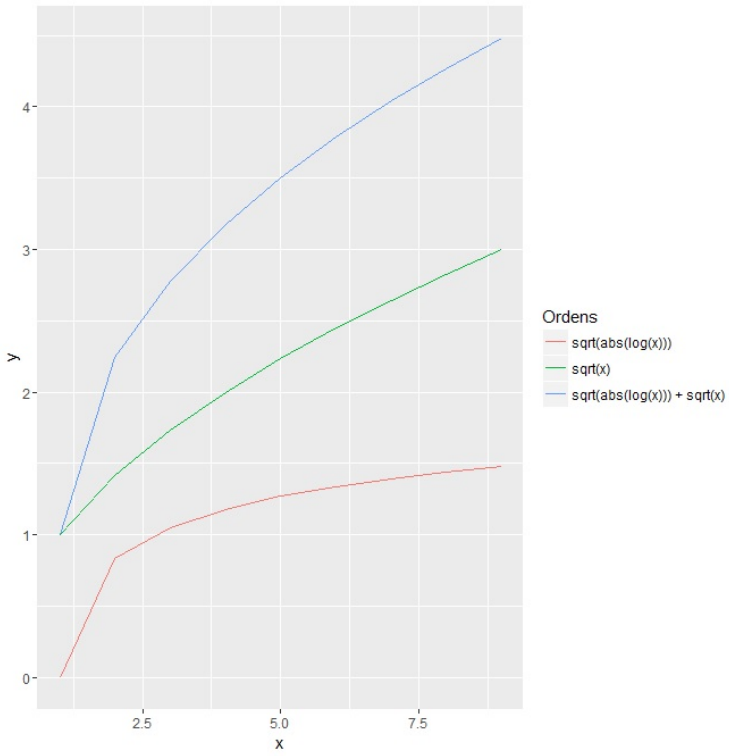


Figure 2.2: *Convergence speed for a=9.*

2.2.2 Jump estimation

To estimate the jump variation Φ, we apply the wavelet method to the observed data and locate all the jumps in trajectory X_t, then use the estimated jump localizations to estimate the jump size. Here the variation of the jump is estimated by the sum of the squares of all sizes of the estimated jumps. Consider the following non-parametric volatility model

$$Y_t = X_t + \epsilon_t, \ t \in [0,1]. \tag{2.11}$$

Here our latent variable is Y_t, that is, the logarithm of the transaction value and is observed in times $t_i = i/n$, with $i = 0, ..., n$; $\epsilon_t \sim RB(0, \eta^2)$ are i.i.d and with the fourth finite moment. Let $X_{a,k}$, $Y_{a,k}$ and $\epsilon_{a,k}$ be the wavelet coefficients of X_t, Y_t and ϵ_t respectively. Then, from the model (2.11) we obtain that

$$Y_{a,k} = X_{a,k} + \epsilon_{a,k}, \ k = 1, ..., 2^a, \ a = 1, ..., log_2(n). \tag{2.12}$$

We used a threshold T_n to adjust $|Y_{a_n,k}|$ and estimate the jump localizations of the trajectory X_t by the locations of $|Y_{a_n,k}|$ that exceed T_n. That is, if $|Y_{a_n,k}| > T_n$ for some $k = 1, ..., 2^a$, the corresponding localization of the jump is estimated by $\widehat{\beta} = k2^{-a_n}$. A threshold choice is the universal threshold given by Dohono and Johnstone (1995), $T_n = d\sqrt{2logn}$, where $d = \frac{|Y_{a_n,k}|}{0{,}6745}$.

2.2.3 Estimation of jump variation

To estimate how the jump of trajectory X_t varies, we need the size of the jump. For each localization $\widehat{\beta}_l$ of X_t we choose a small neighborhood $\widehat{\beta}_l \pm \delta_n$ for some $\delta_n > 0$. We denote by $\overline{Y}_{\widehat{\beta}_{l+}}$ and $\overline{Y}_{\widehat{\beta}_{l-}}$ the averages of Y_{t_i} on $\left[\widehat{\beta}_l, \widehat{\beta}_l + \delta_n\right]$ and $\left[\widehat{\beta}_l - \delta_n, \widehat{\beta}_l\right)$ respectively. We use

$\widehat{Z}_l = \overline{Y}_{\widehat{\beta}_{l+}} - \overline{Y}_{\widehat{\beta}_{l-}}$ for estimate the true size of the jump $Z_l = X_{\beta_l} - X_{\beta_{l^-}}$. The variation of the jump $\Phi = \sum_{l=1}^{N_1} Z_l^2$ as stated above is estimated by the sum of the squares of all sizes of the estimated jumps, that is

$$\widehat{\Phi} = \sum_{l=1}^{\widehat{q}} \left(\overline{Y}_{\widehat{\beta}_{l+}} - \overline{Y}_{\widehat{\beta}_{l-}} \right)^2 .$$

For an example of simulation and real data, see Chapter 4 and 5, respectively.

Chapter 3

Empirical c.d.f. and Estimators for Volatility

We used the local estimator proposed by Todorov and Tauchen (2014) to estimate σ_t^2. On each of the blocks the local estimator of σ_t^2 is given by

$$\hat{V}_j^n = \frac{\pi}{2}\frac{n}{k_n - 1}\sum_{i=(j-1)k_n+2}^{jk_n} |\Delta_{i-1}^n X||\Delta_i^n X|,\ j = 1, \ldots, \lfloor n/2k_n \rfloor, \tag{3.1}$$

which is the bipower variation for measuring the quadratic variation of the diffusion component of X. Todorov and Tauchen (2014) remove the high-frequency increments that contain "big" jumps. The total number of increments used in their statistic is thus given by

$$N^n(\alpha, \bar{w}) = \sum_{j=1}^{\lfloor n/2k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} I\left(|\Delta_i^n X| \leq \alpha\sqrt{\hat{V}_j^n} n^{\bar{w}}\right),$$

where $\alpha > 0$ and $\bar{w} \in (0, 1/2)$. They use a time-varying threshold in the truncation to account for the time varying σ_t. The scaling of every high-frequency increment is done after adjusting $\hat{V}_j^n$ to exclude the contribution of that increment in its formation:

$$\hat{V}_j^n(i) = \begin{cases} \frac{k_n-1}{k_n-3}\hat{V}_j^n - \frac{\pi}{2}\frac{n}{k_n-3}|\Delta_i^n X||\Delta_{i+1}^n X|, & \text{for } i = (j-1)k_n + 1, \\ \frac{k_n-1}{k_n-3}\hat{V}_j^n - \frac{\pi}{2}\frac{n}{k_n-3}(|\Delta_{i-1}^n X||\Delta_i^n X| + |\Delta_i^n X||\Delta_{i+1}^n X|), & \text{for } i = (j-1)k_n + 2, \ldots, jk_n - 1 \\ \frac{k_n-1}{k_n-3}\hat{V}_j^n - \frac{\pi}{2}\frac{n}{k_n-3}|\Delta_{i-1}^n X||\Delta_i^n X|, & \text{for } i = jk_n. \end{cases} \tag{3.2}$$

They then define

$$\hat{F}_n(\tau) = \frac{1}{N^n(\alpha, \bar{w})}\sum_{j=1}^{\lfloor n/2k_n \rfloor} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} I\left(\frac{\sqrt{n}\Delta_i^n X}{\sqrt{\hat{V}_j^n(i)}} \leq \tau\right) I_{\left(|\Delta_i^n X| \leq \alpha\sqrt{\hat{V}_j^n} n^{\bar{w}}\right)}, \tag{3.3}$$

which is simply the empirical c.d.f. of the devolatilized increments that do not contain any big jumps. In the jump-diffusion case of (1.2), $\hat{F}_n(\tau)$ should be approximately the c.d.f. of a standard normal random variable.

Todorov and Tauchen (2014) use an alternative estimator of the volatility that is the truncated variation defined as

$$\hat{C}_j^n = \frac{n}{k_n} \sum_{i=(j-1)k_n+1}^{jk_n} |\Delta_i^n X|^2 I\left(|\Delta_i^n X| \leq \alpha n^{\bar{w}}\right), \ j = 1, \ldots, \lfloor n/2k_n \rfloor,$$

where $\alpha > 0$ and $\bar{w} \in (0, 1/2)$ and the corresponding one excluding the contribution of the ith increment for $i = (j-1)k_n + 1, \ldots, jk_n$, is

$$\hat{C}_j^n(i) = \frac{k_n}{k_n - 1}\hat{C}_j^n - \frac{n}{k_n - 1}|\Delta_i^n X|^2 I\left(|\Delta_i^n X| \leq \alpha n^{\bar{w}}\right), \ j = 1, \ldots, \lfloor n/2k_n \rfloor. \tag{3.4}$$

They also define the corresponding empirical c.d.f. of the devolatilized (and truncated) high-frequency increments as

$$\hat{F'}_n(\tau) = \frac{1}{N'^n(\alpha, \bar{w})} \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} I\left(\frac{\sqrt{n}\Delta_i^n X}{\sqrt{\hat{C}_j^n(i)}} \leq \tau\right) I_{\left(|\Delta_i^n X| \leq \alpha n^{\bar{w}}\right)}, \tag{3.5}$$

where $\alpha > 0$, $\bar{w} \in (0, 1/2)$ and

$$N'^n(\alpha, \bar{w}) = \sum_{j=1}^{[n/k_n]} \sum_{i=(j-1)k_n+1}^{(j-1)k_n+m_n} I\left(|\Delta_i^n X| \leq \alpha n^{\bar{w}}\right).$$

3.1 Test Statistic: Cramér-von Mises

We define the empirical process as:

$$\hat{Y}_n(\tau) = \sqrt{\lfloor n/2k_n \rfloor \, m_n}(\hat{F}_n(u_n, \tau) - \Phi(\tau)), \tag{3.6}$$

where $\hat{F}_n(u_n, \tau)$ is an empirical distribution function of the devolatized increments. In this work we will consider two possible empirical distribution functions in (3.6). Theorem 4 of Todorov and Tauchen (2014) and Theorem 1 of Kong (2017) motivate us to propose a measure of discrepancy between distributions using the "Cramér-Von Mises" type statistic as

$$T_A^n = d^2 \sqrt{\lfloor n/2k_n \rfloor \, m_n}, \tag{3.7}$$

where A is a compact set in $\mathbb{R}$ and

$$d^2 = \int_{-\infty}^{\infty} [F_n(x) - F(x)]^2 dF(x),$$

where $F(x)$ is a cumulative distribution function and $F_n(x)$ is an empirical distribution function. By notation we will replace $F_n(x)$ by $\hat{F'}_n(x)$ given by (3.5) and $F(x)$ by $\Phi(x)$, so we will have,

$$d^2 = \int_{-\infty}^{\infty} [\hat{F'}_n(x) - \Phi(x)]^2 d\Phi(x).$$

The choice for this test statistic comes from the fact that it has better power than the KS statistic that measures the maximum distance. The critial region of the proposed test is

given by

$$C_n = \left\{ d^2 \sqrt{\lfloor n/2k_n \rfloor \, m_n} > q_n(\alpha, A) \right\},$$

where $q_n(\alpha, A)$ is the $(1-\alpha)$-quantile of

$$\sqrt{[n/(2k_n)]m_n} \int_{-\infty}^{\infty} [\hat{F}'_n(x) - \Phi(x)]^2 d\Phi(x). \tag{3.8}$$

We evaluate $q_n(\alpha, A)$ via simulation. The test rejects H_0 if $T_A^n > q_n(\alpha, A)$.

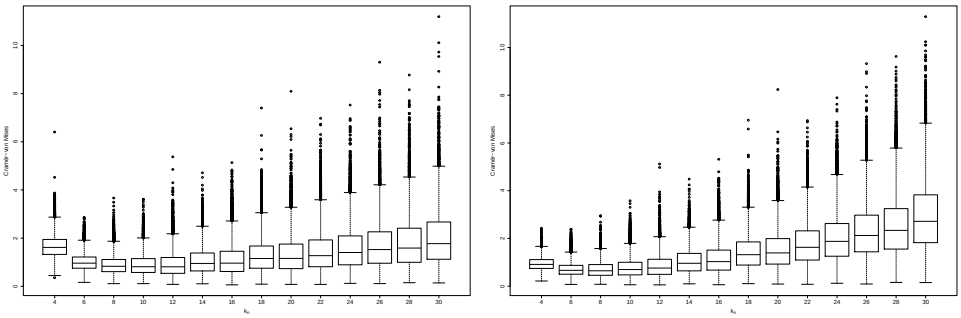
Our simulation study showed that the proposed test statistic follows approximately a known distribution function. Comparing the respective quantiles at the significance levels of 5% and 10%, the values of the proposed distribution and the sample distribution are very close.

3.2 Performance of the empirical c.d.f.

We consider the cumulative distribution functions with their respective volatility estimators as described in (3.3) that contains the bipower variation estimator and in (3.5) that contains the truncated variation estimator. A simulation study was performed with 10,000 replications and different values of n, k_n and m_n. We set the pair of (k_n, m_n) for $n = 1000$, 2000 and 5000 to be:

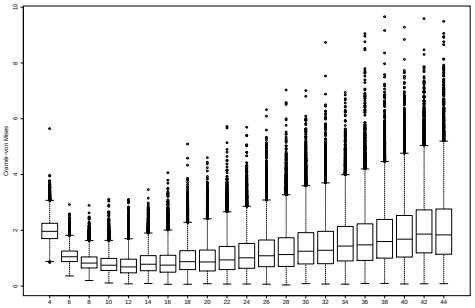
n	(k_n, m_n)
1000	(4,2), (6,3), (8,4), (10,5), (12,5), (14,8), (16,7) (18,11), (20,9), (22,11), (24,12), (26,13), (28,13), (30,16).
2000	(4,2), (6,3), (8,4), (10,5), (12,5), (14,8), (16,7) (18,11), (20,9), (22,11), (24,12), (26,13), (28,13), (30,16), (32,15), (34,18), (36,17), (38,20), (40,20), (42,26), (44,21).
5000	(4,2), (6,3), (8,4), (10,5), (12,5), (14,8), (16,7) (18,11), (20,9), (22,11), (24,12), (26,13), (28,13), (30,16), (32,15), (34,18), (36,17), (38,20), (40,20), (42,26), (44,21), (46,25), (48,23), (50,28), (52,27), (54,34), (56,31), (58,30), (60,34), (62,33), (64,30), (66,40), (68,35), (70,33).

Figure 3.1 shows the behavior of the test statistic Cv-M against k_n for the jump-diffusion model. From the plots, we observe some outliers indicating the existence of some discrepant values. Also, the value of the test statistic increases as k_n increases. In cases where $k_n = 4$ and $k_n = 6$ the values of the test statistic present some inconsistencies. In these cases we simply call F the empirical c.d.f $\hat{F}_n(\tau)$ that contains the local estimator given in (3.2) and F' as the empirical c.d.f. $\hat{F}'_n(\tau)$ that contains the truncated variation estimator given in (3.4). Further, we call the test statistic using $\hat{F}_n(\tau)$ as $TA1$ and the test statistic using $\hat{F}'_n(\tau)$ as $TA2$. Figure 3.2 shows the behavior of the test statistic Cv-M against k_n for the standard normal model.

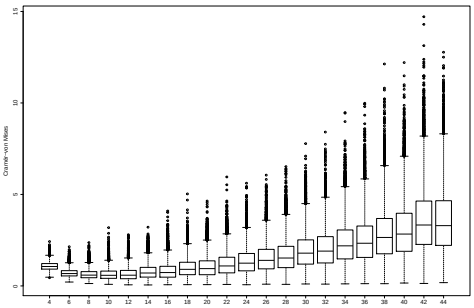


(a) *Cv-M using* $\hat{F}$, $n = 1000$.

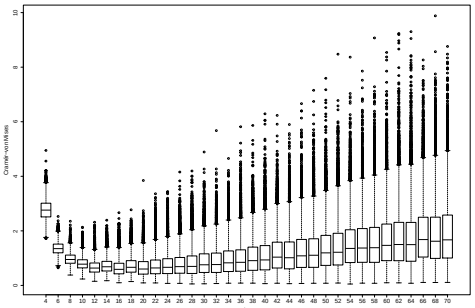
(b) *Cv-M using* $\hat{F}'$, $n = 1000$.



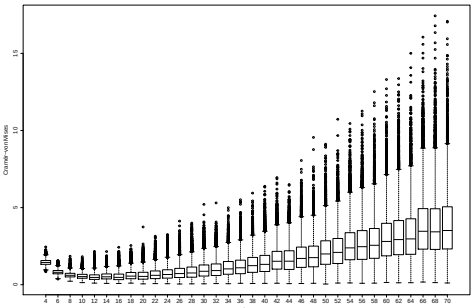
(c) *Cv-M using* $\hat{F}$, $n = 2000$.



(d) *Cv-M using* $\hat{F}'$, $n = 2000$.

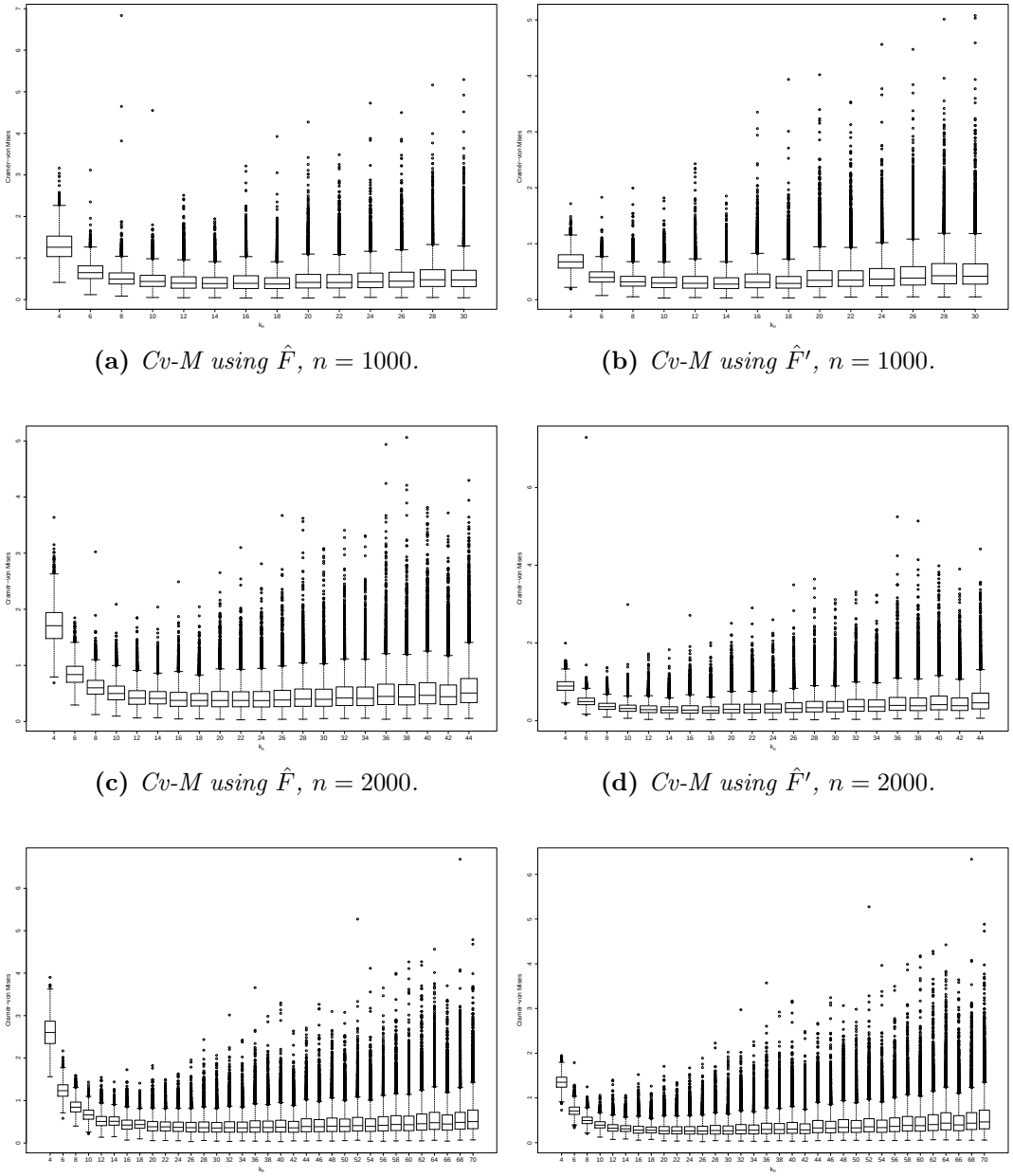


(e) *Cv-M using* $\hat{F}$, $n = 5000$.



(f) *Cv-M using* $\hat{F}'$, $n = 5000$.

Figure 3.1: *Left panel: the test statistic Cv-M against* k_n *using* F*; Right panel: the test statistic Cv-M against* k_n *using* F' *for the jump difussion model.*



(a) *Cv-M using* $\hat{F}$, $n = 1000$.

(b) *Cv-M using* $\hat{F}'$, $n = 1000$.

(c) *Cv-M using* $\hat{F}$, $n = 2000$.

(d) *Cv-M using* $\hat{F}'$, $n = 2000$.

(e) *Cv-M using* $\hat{F}$, $n = 5000$.

(f) *Cv-M using* $\hat{F}'$, $n = 5000$.

Figure 3.2: *Left panel: the test statistic Cv-M against* k_n *using* F*; Right panel: the test statistic Cv-M against* k_n *using* F' *for the standard normal model.*

The first simulations were based on $10,000$ replicas with $n = 1000$ and the same pairs of k_n and m_n as mentioned above. The idea is to observe which test statistic performs better using the two variance estimators mentioned in the previous section. It can be observed in Figure 3.3 that for the jump-diffusion model, in average, the quadratic distance of the test statistic increases as k_n increases. On the other hand, for the normal model, in average, the quadratic distance of the test statistic stabilizes, that is, the values of the test statistics are close as k_n increases.

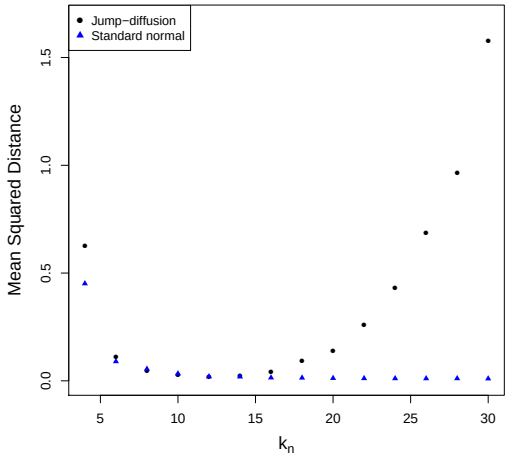


Figure 3.3: *Mean Squared Distance of the two test statistics for different values of k_n, in data generated from the two models with $n = 1000$.*

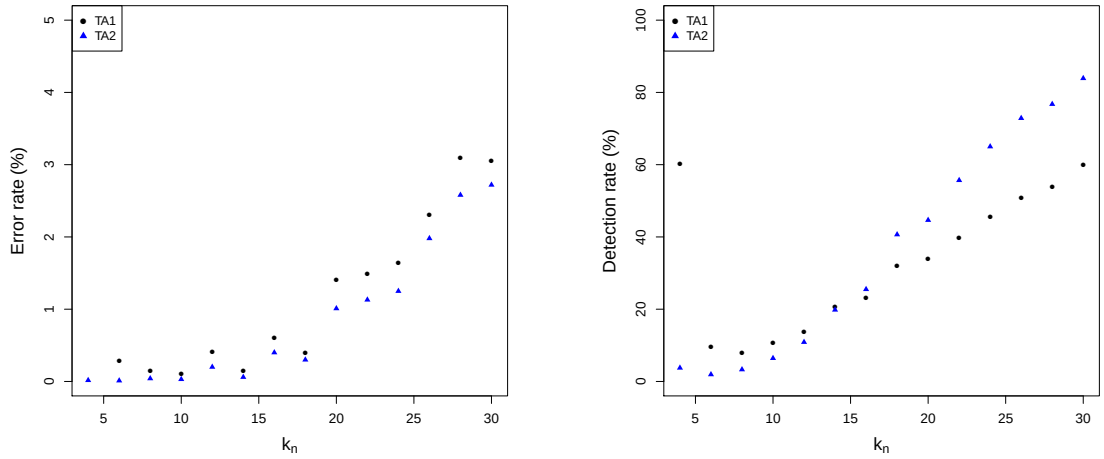
We present a method for choosing the "best" critical value. The ROC curve (Receiver Operating Characteristic Curve) is widely used to determine the cutoff point. We test the critical value "q" between 1 and 3 and we use the ROC curve to show the cutoff point. We compare detection and error rates for this value and a fixed value.

		$q = 1.5$				$q = 1.2322$			
		Error rate		Dectection rate		Error rate		Dectection rate	
k_n	m_n	$TA1$	$TA2$	$TA1$	$TA2$	$TA1$	$TA2$	$TA1$	$TA2$
4	2	0.2728	0.0002	0.6026	0.0376	0.5286	0.0052	0.8217	0.1453
6	3	0.0029	0.0001	0.0952	0.0193	0.0204	0.0004	0.2378	0.0604
8	4	0.0015	0.0004	0.0805	0.0330	0.0056	0.0010	0.1793	0.0863
10	5	0.0011	0.0003	0.1061	0.0643	0.0057	0.0008	0.2074	0.1402
12	5	0.0042	0.0020	0.1372	0.1088	0.0100	0.0052	0.2324	0.1993
14	8	0.0014	0.0006	0.2071	0.1982	0.0054	0.0028	0.3248	0.3242
16	7	0.0060	0.0040	0.2323	0.2551	0.0164	0.0113	0.3490	0.3833
18	11	0.0039	0.0030	0.3204	0.4066	0.0110	0.0082	0.4578	0.5502
20	9	0.0140	0.0101	0.3381	0.4465	0.0287	0.0213	0.4616	0.5796
22	11	0.0149	0.0113	0.3984	0.5568	0.0306	0.0247	0.5208	0.6814
24	12	0.0164	0.0125	0.4569	0.6501	0.0354	0.0258	0.5795	0.7586
26	13	0.0231	0.0198	0.5094	0.7286	0.0451	0.0378	0.6255	0.8170
28	13	0.0309	0.0258	0.5378	0.7675	0.0585	0.0482	0.6472	0.8477
30	16	0.0305	0.0272	0.5989	0.8390	0.0599	0.0500	0.7057	0.8937

Table 3.1: *Detection and error rates for k_n values between 4 and 30 for fixed critical value $q = 1.5$ and simulated the critical value $q = 1.2322$, in 10000 replicas of data generated with $n = 1000$.*

Table 3.1 shows that for a fixed critical value 1.5, the error rate for $k_n = 30$ with the

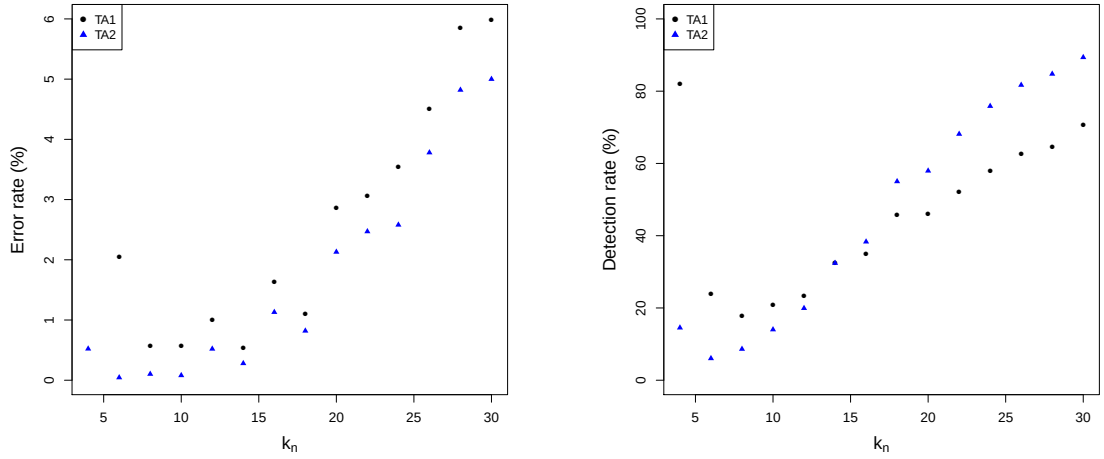
test statistic $TA2$ is around 2.8% and the detection rate is 84%. In contrast, for the test statistic $TA1$, the error rate is around 3% and the detection rate is 60%. It can also be seen that, for low values of k_n in $TA1$, there are low error rates but with low detection power. For example for $q = 1.5$, in the case of $k_n = 20$, there is an error rate around 1.4% and a detection rate of 34%. In contrast for $TA2$ we observe 44% of detection. Besides for the optimal critical value $q = 1.2322$, with $k_n = 30$ we have an error rate of 5% and a detection power of 90% for the test statistic $TA2$ and for the test statistic $TA1$, we observe an error rate of 6% and a detection rate of 70%. Figure 3.4 shows the detection and error rates for a fixed critical value of 1.5. Figure 3.5 shows the detection and error rates of a simulated critical value $q = 1.2322$, that was obtained by ROC curve (Figure 3.6).



(a) *Error rate* **(b)** *Detection rate*

Figure 3.4: *Detection and error rates with critical value* 1.5 *and different values of* k_n

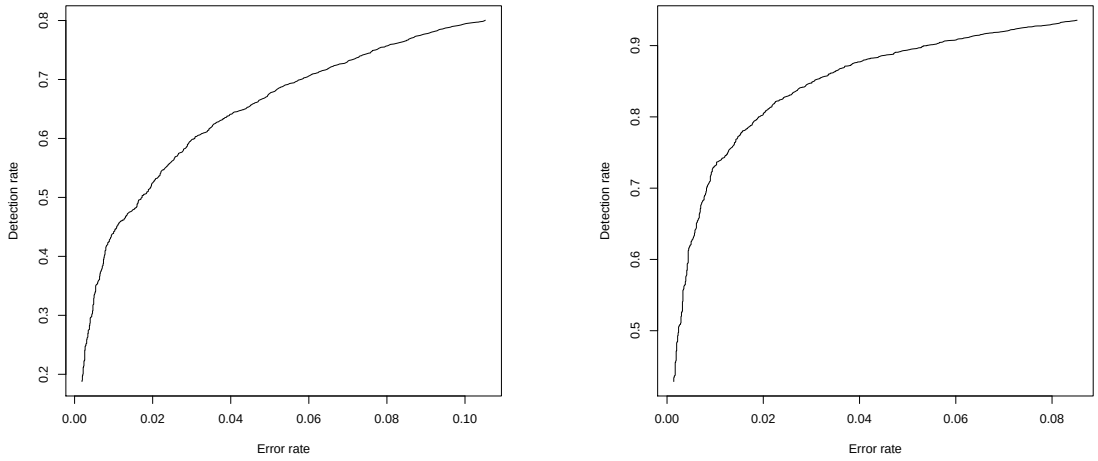
.



(a) *Error rate* **(b)** *Detection rate*

Figure 3.5: *Detection and error rates with critical value* 1.2322

We repeat our simulation for $n = 2000$ and the same pairs of k_n and m_n already mentioned above. The conclusion is similar to the case $n = 1000$ and with respect to the average quadratic distance of the statistic, as we can see in Figure 3.7.



(a) *ROC curve TA1* **(b)** *ROC curve TA2*

Figure 3.6: *ROC curve for* $TA1$ *and* $TA2$ *statistics.*

Figure 3.8 shows the detection and error rates for a fixed critical value of 1.5. The Table 3.2 shows the detection and error rates for two different critical values, which lead us to conclude that $TA2$ is better. For example, for $q = 1.3663$ we have an error rate of 5% and a detection power of 91% for the test statistic $TA2$. With the test statistic $TA1$, we observe an error rate of 5.8% and a detection rate of 66%.

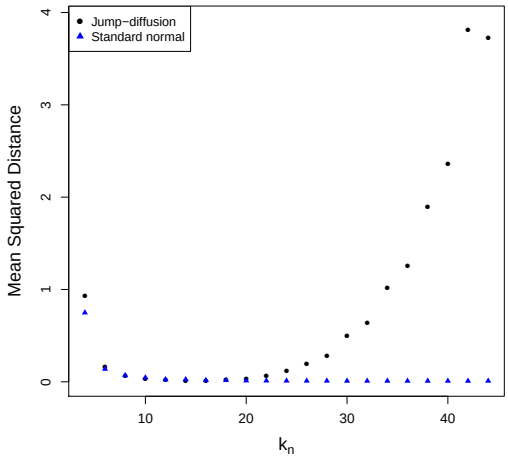


Figure 3.7: *Mean Squared Distance of the two test statistics for different values of* k_n*, in data generated from the two models with* $n = 2000$*.*

Figure 3.9 shows the detection and error rates for a fixed $q = 1.3663$, obtained by simulation, with the ROC curve (Figure 3.10). It can be seen that as k_n increases the error rate for $TA1$ is higher than for $TA2$ and for the detection rate as k_n increases and $TA2$ has a higher detection rate than $TA1$. Again $TA2$ seems to be better than $TA1$.

We can observe in Figure 3.11 that the quadratic distance of the statistics has a similiar behavior to the previous analysis. We are interested in the largest quadratic distance, so for the pairs of k_n and m_n already mentioned above and $n = 5000$ we will do our study for the aforementioned reasons. Figure 3.12 shows the detection and error rates for a fixed critical value 1.5. Figure 3.13 shows the detection and error rate considering a simulated critical value obatined by ROC curve (Figure 3.14).

Table 3.3 shows that for a fixed critical value 1.5, we see that, the error rate for $k_n = 70$

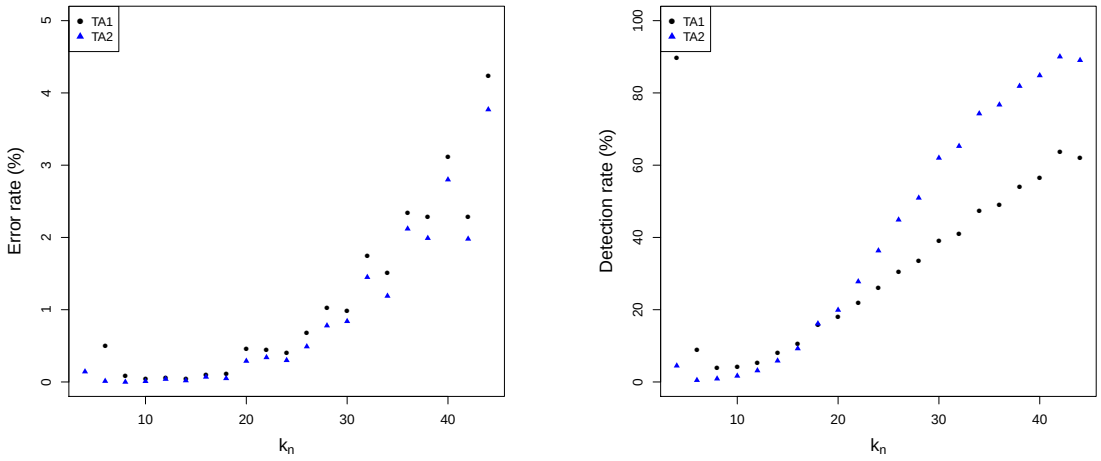
		$q = 1.5$				$q = 1.3663$			
		Error rate		Detection rate		Error rate		Detection rate	
k_n	m_n	$TA1$	$TA2$	$TA1$	$TA2$	$TA1$	$TA2$	$TA1$	$TA2$
4	2	0.7289	0.0014	0.8966	0.0449	0.8380	0.0090	0.9519	0.1101
6	3	0.0050	0.0001	0.0877	0.0049	0.0164	0.0002	0.1565	0.0110
8	4	0.0008	0.0000	0.0387	0.0090	0.0018	0.0000	0.0688	0.0162
10	5	0.0004	0.0001	0.0424	0.0168	0.0011	0.0003	0.0681	0.0288
12	5	0.0006	0.0004	0.0530	0.0316	0.0015	0.0008	0.0793	0.0512
14	8	0.0004	0.0002	0.0806	0.0587	0.0006	0.0003	0.1223	0.0892
16	7	0.0010	0.0007	0.1066	0.0926	0.0018	0.0010	0.1471	0.1277
18	11	0.0011	0.0005	0.1572	0.1615	0.0021	0.0008	0.2067	0.2145
20	9	0.0046	0.0029	0.1799	0.1992	0.0063	0.0041	0.2227	0.2533
22	11	0.0044	0.0034	0.2201	0.2779	0.0069	0.0051	0.2713	0.3444
24	12	0.0041	0.0030	0.2596	0.3635	0.0075	0.0056	0.3168	0.4325
26	13	0.0068	0.0049	0.3052	0.4489	0.0110	0.0076	0.3589	0.5166
28	13	0.0103	0.0078	0.3350	0.5095	0.0152	0.0116	0.3876	0.5730
30	16	0.0099	0.0084	0.3908	0.6201	0.0158	0.0125	0.4465	0.6761
32	15	0.0174	0.0145	0.4094	0.6527	0.0249	0.0208	0.4670	0.7058
34	18	0.0151	0.0119	0.4726	0.7428	0.0223	0.0183	0.5326	0.7878
36	17	0.0234	0.0212	0.4911	0.7673	0.0341	0.0295	0.5451	0.8075
38	20	0.0228	0.0199	0.5409	0.8189	0.0310	0.0263	0.5943	0.8515
40	20	0.0311	0.0280	0.5648	0.8483	0.0430	0.0381	0.6188	0.8776
42	26	0.0228	0.0198	0.6370	0.9004	0.0327	0.0274	0.6876	0.9205
44	21	0.0423	0.0377	0.6201	0.8905	0.0576	0.0500	0.6659	0.9118

Table 3.2: *Detection and error rates for k_n values between 4 and 44 for fixed critical value $q = 1.5$ and simulated critical value $q = 1.3663$, in 10000 replicas of data generated with $n = 2000$.*

with the test statistic $TA2$ is around 3.9% and the detection rate of 90%. In contrast, for the test statistic $TA1$, the error rate is around 4.3% and detection rate of 56%. It can also be seen that, for low values of k_n in $TA1$, there are low error rates but with low detection power. For example, in the case of $k_n = 48$ with the test statistic $TA1$, there is an error rate around 1.2% and a detection rate of 32%.

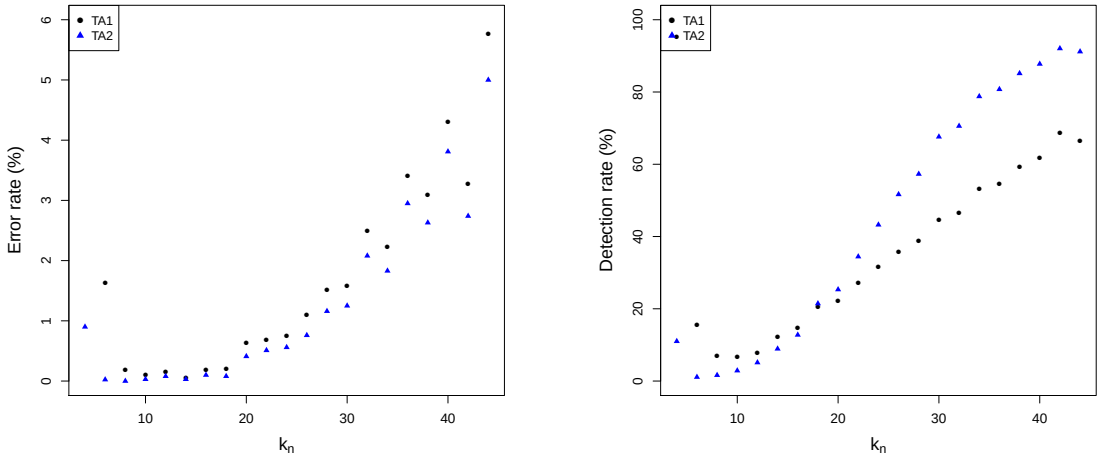
The critical value obtained by ROC curve is $q = 1.3663$. With this, we have an error rate of 4.9% and a detection power of 91% for the test statistic $TA2$. With the test statistic using the $TA1$, we observe an error rate of 5.5% and a detection rate of 59%. We will use the $TA2$ test statistic, which uses the truncated variance estimator defined in (3.4), since it presented better results in relation to low error rate and high detection power. The ideal is to choose a k_n that shows a low error rate and high detection power. In contrast, for the test statistic $TA1$, there is high error rate and low detection power for most k_n values. Table 3.4 shows a summary of the results for the 3 possible sample sizes.

Note that in Table 3.1, 3.2 and 3.3 for values of $k_n = 4$ and $k_n = 6$ the error and detection rate shows inconsistent values, this is because it has a high false alarm rate and not so much detection power. Therefore, we do not recommend using low values of k_n.



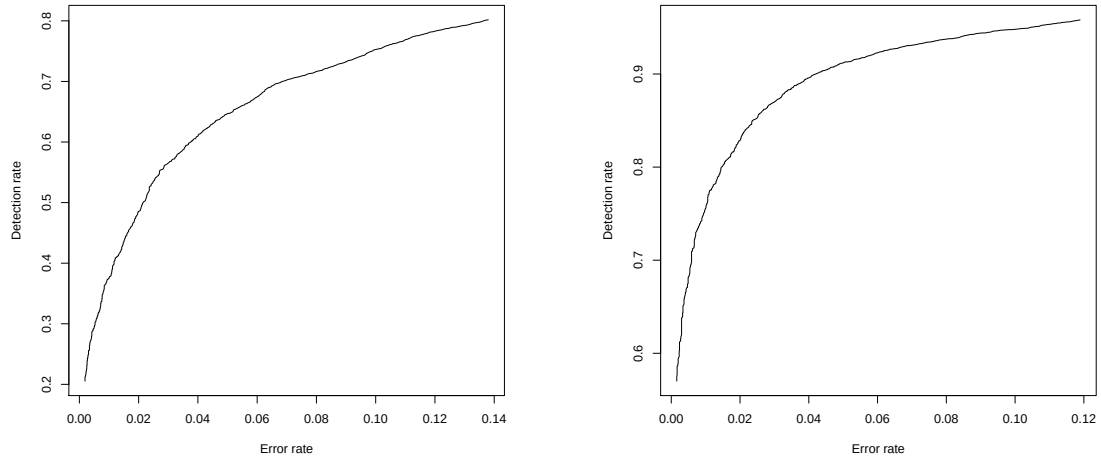
(a) *Error rate* **(b)** *Detection rate*

Figure 3.8: *Detection and error rates with critical value* 1.5



(a) *Error rate* **(b)** *Detection rate*

Figure 3.9: *Detection and error rates with critical value* 1.3663



(a) *ROC curve TA1* **(b)** *ROC curve TA2*

Figure 3.10: *ROC curve for* $TA1$ *and* $TA2$ *statistic.*

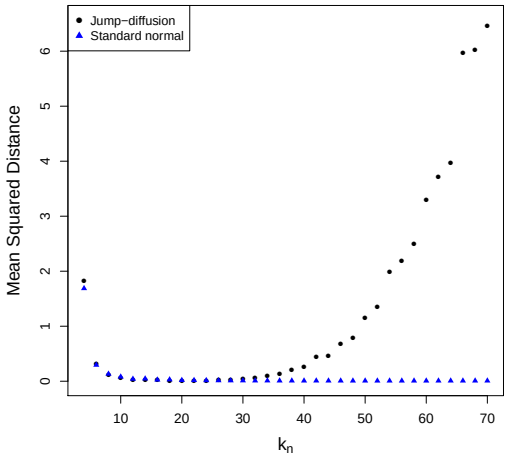
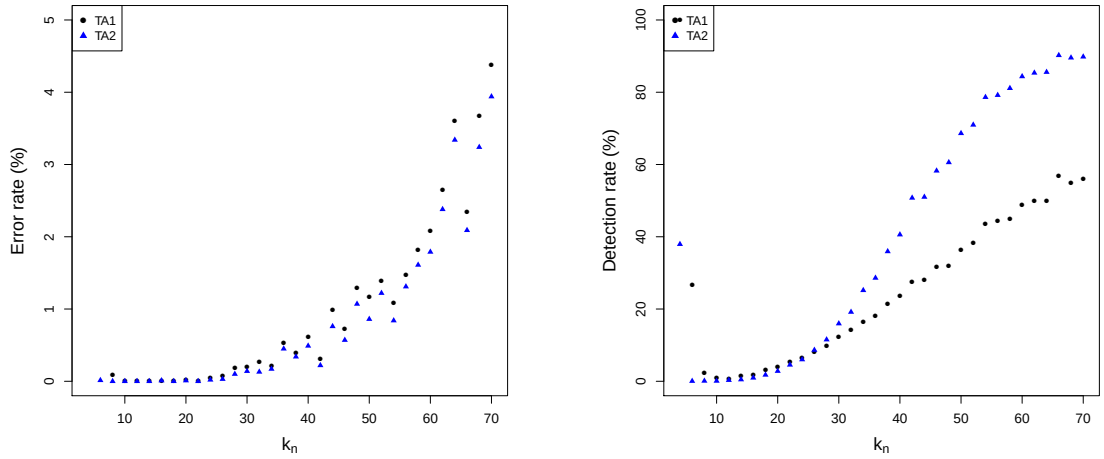


Figure 3.11: *Mean Squared Distance of the two test statistics for different values of k_n, in data generated from the two models with $n = 5000$.*



(a) *Error rate* **(b)** *Detection rate*

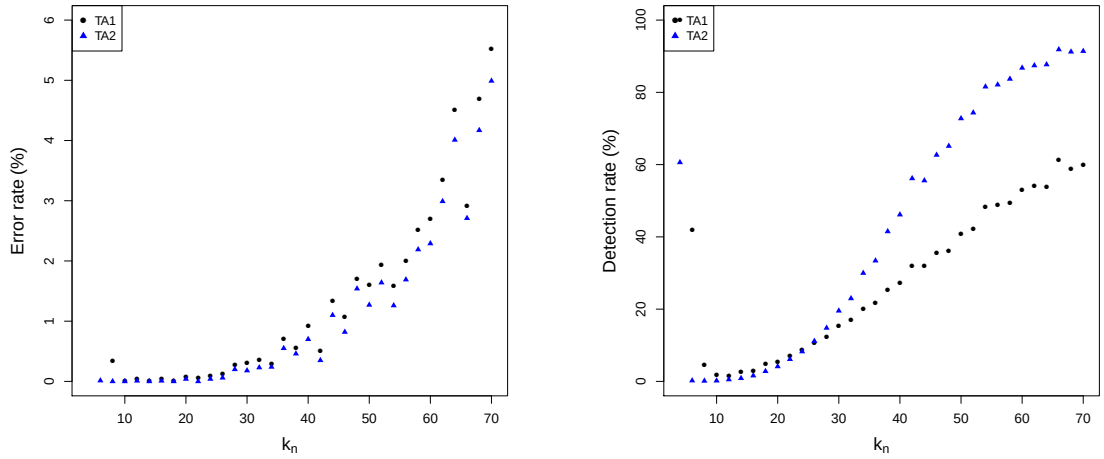
Figure 3.12: *Detection and error rates with critical value* 1.5

		$q = 1.5$				$q = 1.3943$			
		Error rate		Detection rate		Error rate		Detection rate	
k_n	m_n	$TA1$	$TA2$	$TA1$	$TA2$	$TA1$	$TA2$	$TA1$	$TA2$
4	2	1.0000	0.1851	1.0000	0.3794	1.0000	0.4053	1.0000	0.6061
6	3	0.1101	0.0001	0.2681	0.0005	0.2176	0.0001	0.4208	0.0022
8	4	0.0009	0.0000	0.0226	0.0010	0.0034	0.0000	0.0452	0.0013
10	5	0.0000	0.0000	0.0094	0.0009	0.0001	0.0000	0.0179	0.0019
12	5	0.0001	0.0000	0.0082	0.0031	0.0004	0.0001	0.0145	0.0052
14	8	0.0000	0.0000	0.0150	0.0048	0.0001	0.0000	0.0261	0.0087
16	7	0.0001	0.0001	0.0171	0.0096	0.0004	0.0001	0.0279	0.0159
18	11	0.0000	0.0000	0.0314	0.0177	0.0001	0.0000	0.0481	0.0281
20	9	0.0002	0.0001	0.0387	0.0282	0.0008	0.0004	0.0548	0.0413
22	11	0.0000	0.0000	0.0536	0.0455	0.0006	0.0000	0.0699	0.0613
24	12	0.0005	0.0002	0.0638	0.0602	0.0009	0.0004	0.0875	0.0826
26	13	0.0008	0.0003	0.0815	0.0864	0.0012	0.0006	0.1080	0.1115
28	13	0.0018	0.0010	0.0978	0.1153	0.0028	0.0020	0.1246	0.1479
30	16	0.0020	0.0014	0.1233	0.1597	0.0031	0.0018	0.1538	0.1954
32	15	0.0027	0.0013	0.1432	0.1917	0.0036	0.0023	0.1705	0.2296
34	18	0.0021	0.0017	0.1649	0.2518	0.0030	0.0024	0.2011	0.2998
36	17	0.0053	0.0045	0.1800	0.2860	0.0070	0.0055	0.2174	0.3341
38	20	0.0039	0.0034	0.2135	0.3594	0.0055	0.0046	0.2530	0.4150
40	20	0.0061	0.0049	0.2368	0.4059	0.0093	0.0070	0.2742	0.4615
42	26	0.0031	0.0022	0.2755	0.5076	0.0050	0.0035	0.3185	0.5618
44	21	0.0099	0.0076	0.2806	0.5101	0.0133	0.0110	0.3189	0.5559
46	25	0.0072	0.0057	0.3156	0.5826	0.0108	0.0082	0.3556	0.6267
48	23	0.0129	0.0107	0.3206	0.6058	0.0171	0.0154	0.3626	0.6514
50	28	0.0117	0.0086	0.3629	0.6863	0.0160	0.0127	0.4086	0.7277
52	27	0.0139	0.0122	0.3823	0.7097	0.0193	0.0164	0.4226	0.7437
54	34	0.0109	0.0084	0.4359	0.7863	0.0158	0.0126	0.4834	0.8156
56	31	0.0147	0.0131	0.4448	0.7919	0.0200	0.0169	0.4888	0.8212
58	30	0.0182	0.0161	0.4508	0.8111	0.0252	0.0219	0.4944	0.8371
60	34	0.0208	0.0179	0.4876	0.8437	0.0270	0.0229	0.5301	0.8677
62	33	0.0265	0.0238	0.4999	0.8536	0.0334	0.0299	0.5411	0.8743
64	30	0.0360	0.0334	0.4983	0.8557	0.0451	0.0401	0.5380	0.8770
66	40	0.0235	0.0209	0.5697	0.9022	0.0291	0.0271	0.6127	0.9186
68	35	0.0367	0.0324	0.5480	0.8955	0.0470	0.0417	0.5882	0.9123
70	33	0.0438	0.0394	0.5613	0.8982	0.0552	0.0499	0.5994	0.9140

Table 3.3: *Detection and error rates for k_n values between 4 and 70 for fixed critical value $q = 1.5$ and simulated the critical value $q = 1.3943$, in 10000 replicas of data generated with $n = 5000$.*

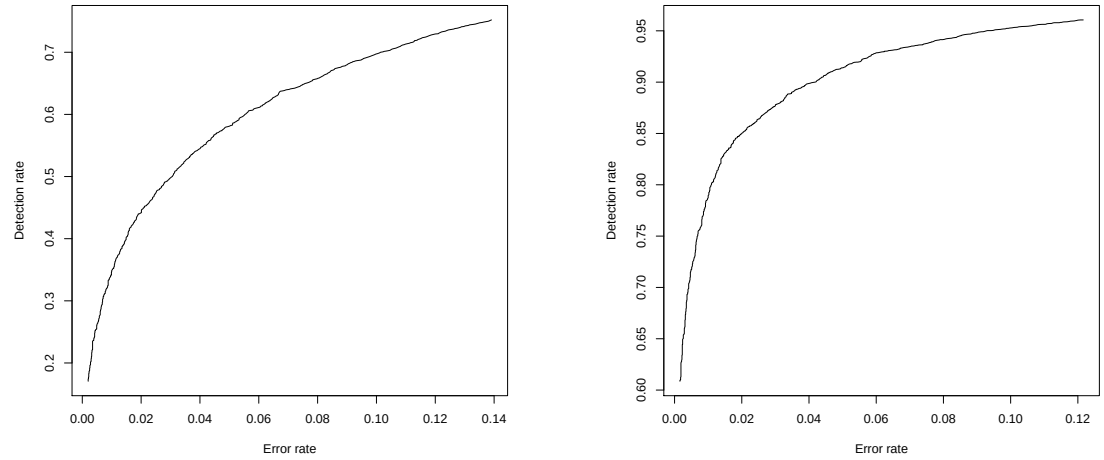
			$q = 1.5$		Optimal q	
n	k_n	m_n	Error rate	Detection rate	Error rate	Detection rate
1000	30	16	0.0272	0.8390	0.05 ($q = 1.2322$)	0.8937
2000	44	21	0.0377	0.8905	0.05 ($q = 1.3663$)	0.9118
5000	70	33	0.0394	0.8982	0.049 ($q = 1.3943$)	0.9140

Table 3.4: *Error and detection rates of $TA2$ statistic for a critical value fixed $q = 1.5$ and optimal critical values for some scenarios.*



(a) *Error rate* **(b)** *Detection rate*

Figure 3.13: *Detection and error rates with critical value* 1.3943



(a) *ROC curve TA1* **(b)** *ROC curve TA2*

Figure 3.14: *ROC curve for* $TA1$ *and* $TA2$ *statistic.*

3.3 Quantile Analysis

3.3.1 Practical Considerations

The conditions in Theorem 1 and as discussed in Kong (2017), for a finite sample n, k_n should be smaller than $\sqrt{n}$ and m_n should be smaller than k_n, with k_n/m_n ranging from 1.5 to 2.5 and having an increasing trend. We will use the $TA2$ test statistic, which uses the truncated variation estimator defined in (3.4), since it presented better results in relation to low error rate and high detection power. The ideal is to choose a k_n that shows a low error rate and high detection power. In contrast, for the test statistic $TA1$, there is high error rate and low detection power for most k_n values.

3.3.2 Quantile

The purpose is to find the theoretical quantile of our test statistic as stated above; for this we must find the cumulative distribution function of our statistic Cramér-von Mises given by:

$$\begin{aligned} P\left(T_A^n \leq x\right) &= P\left(d^2\sqrt{\lfloor n/2k_n \rfloor\, m_n} \leq x\right) \\ &= P\left(\sqrt{\lfloor n/2k_n \rfloor\, m_n}\int_{-\infty}^{\infty}[\hat{F'}_n(\tau) - \Phi(\tau)]^2 d\Phi(\tau) \leq x\right), \tau \in \mathbb{R}, x \geq 0. \end{aligned} \tag{3.9}$$

where $\hat{F'}_n(\tau)$ is the empirical c.d.f. given in (3.5) and $\Phi(\tau)$ is the c.d.f of standard normal. Here, $\Phi'(\tau) = d\Phi(\tau) = \phi(\tau)d\tau$. The expression (3.9) does not have a closed form (for details see the Appendix). So we have to resort to numerical methods to calculate it.

Our simulation study shows that the c.d.f. of our $TA2$ test statistic is approximately a gamma distribution with parameters of shape and scale, a and b, respectively. The gamma distribution has as density

$$f(x) = \frac{1}{b^a\Gamma(a)}x^{a-1}\exp^{-x/b},\ x \geq 0,\ a > 0, \text{and } b > 0.$$

The mean and variance are $E[x] = ab$ and $Var[x] = ab^2$. Since $E[x] = ab$ then $a = \frac{E[x]}{b}$. On the other hand

$$Var[x] = ab^2 = \frac{E[x]}{b}b^2 = E[x]b \Rightarrow b = \frac{Var[x]}{E[x]}.$$

Thus, the gamma c.d.f. is given by

$$\begin{aligned} F(x) &= \int_0^x f(t)dt = \int_0^x \frac{1}{b^a\Gamma(a)}t^{a-1}e^{-t/b}dt = \frac{1}{\Gamma(a)}\int_0^x \frac{t^{a-1}}{b^a}e^{-t/b}dt \\ &= \frac{1}{\Gamma(a)}\int_0^x \left(\frac{t}{b}\right)^{a-1} e^{-t/b}dt. \end{aligned}$$

Making the transformation $y = t/b \Rightarrow dy = dt/b$, we have

$$F(x) = \frac{1}{\Gamma(a)}\int_0^x y^{a-1}e^{-y}dy.$$

The gamma c.d.f. involves the incomplete gamma function given by

$$\int_0^x y^{a-1}e^{-y}dy = \gamma(a,x).$$

So finally,

$$F(x) = \frac{\gamma(a,x)}{\Gamma(a)},\ a > 0.$$

Then, $TA2 \sim^a Gamma(a,b)$. For large sample sizes, we observe that $a \approx 2$ and $b \approx 0.25$.

n	k_n	m_n	a	b	Quantile of $qgamma(a,b)$ 0.95	Sample Quantile 0.95
500	18	11	2.254	0.190	0.982	0.970
	20	11	2.109	0.238	1.174	1.165
	22	12	1.938	0.274	1.273	1.243
1000	26	13	2.083	0.231	1.128	1.116
	28	13	1.993	0.267	1.266	1.270
	30	19	2.064	0.228	1.106	1.084
2000	40	18	1.922	0.283	1.308	1.296
	42	22	1.972	0.267	1.259	1.246
	44	22	1.953	0.282	1.320	1.304
5000	66	36	1.993	0.261	1.239	1.251
	68	38	1.974	0.264	1.241	1.247
	70	34	1.885	0.308	1.403	1.380
10000	96	62	1.947	0.253	1.181	1.179
	98	48	1.894	0.301	1.379	1.357
	100	54	1.972	0.284	1.334	1.322

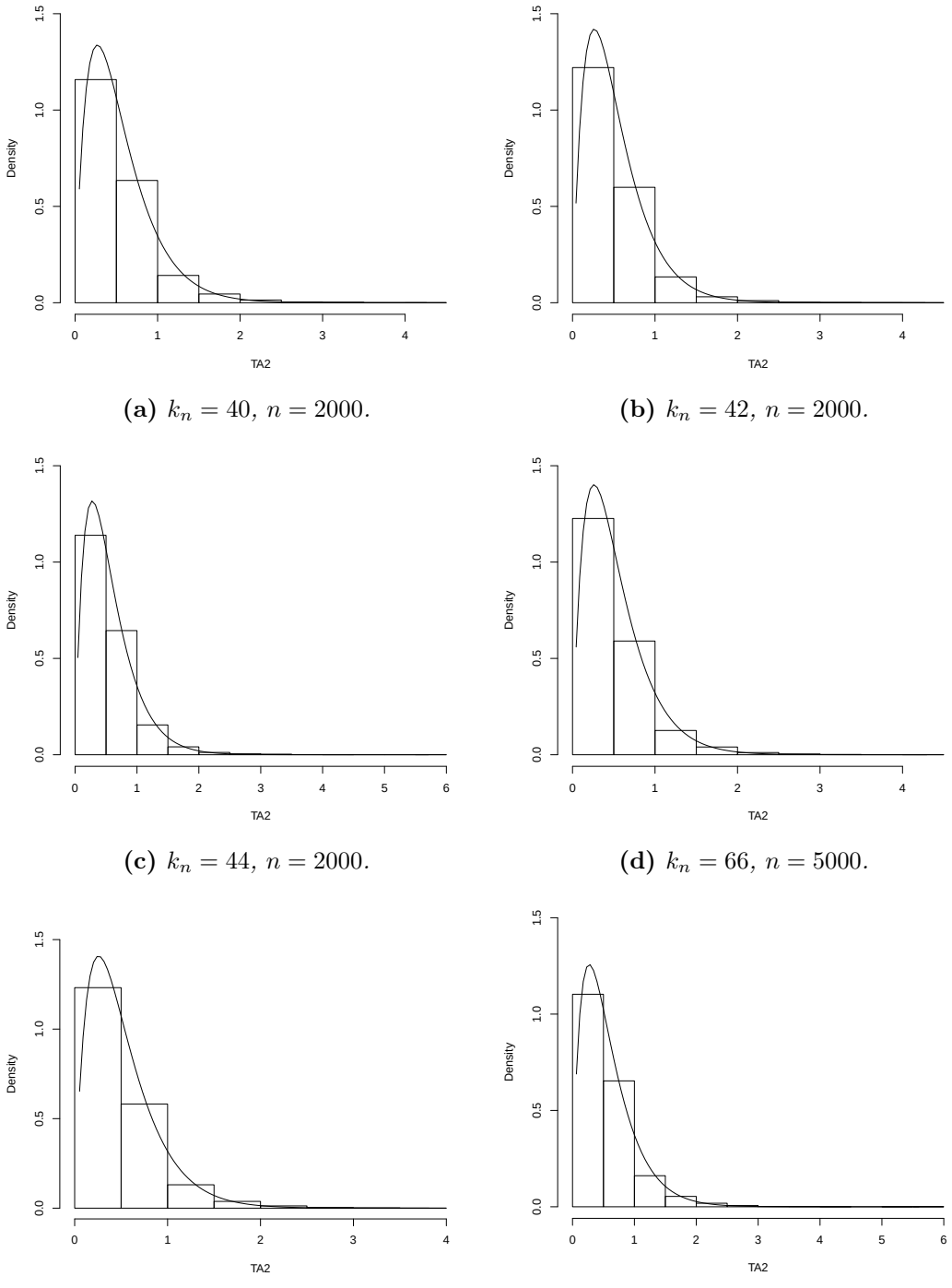
Table 3.5: *Comparison between sample quantile and gamma quantile.*

Hence, approximately

$$TA2 \sim^a Gamma(2, 0.25),$$

with $E[TA2] = 0.5$ and $Var[TA2] = 0.125$.

In Figure A.1 (See Appendix) we observe that the sampling c.d.f. of $TA2$ and Gamma c.d.f. are very close. In Figure 3.15 was plotted the sample density against the gamma density, we can observe that they are similar.



(a) $k_n = 40$, $n = 2000$.

(b) $k_n = 42$, $n = 2000$.

(c) $k_n = 44$, $n = 2000$.

(d) $k_n = 66$, $n = 5000$.

(e) $k_n = 68$, $n = 5000$.

(f) $k_n = 70$, $n = 5000$.

Figure 3.15: *Histograms of the sample density function for different sizes of k_n and n, compared with the gamma density.*

Chapter 4

Simulation study

4.1 Simulation Case 1.

4.1.1 Example

To test the method proposed by Fan and Wang (2007), let's take a simulated example in Wang (1995), that is illustrated in Figure 4.1. For our example $n = 2^{10}$ with $a = 0, ..., 9$. The function clearly has a jump and a peak. We will perform the detection and localization of the jumps in different scales.

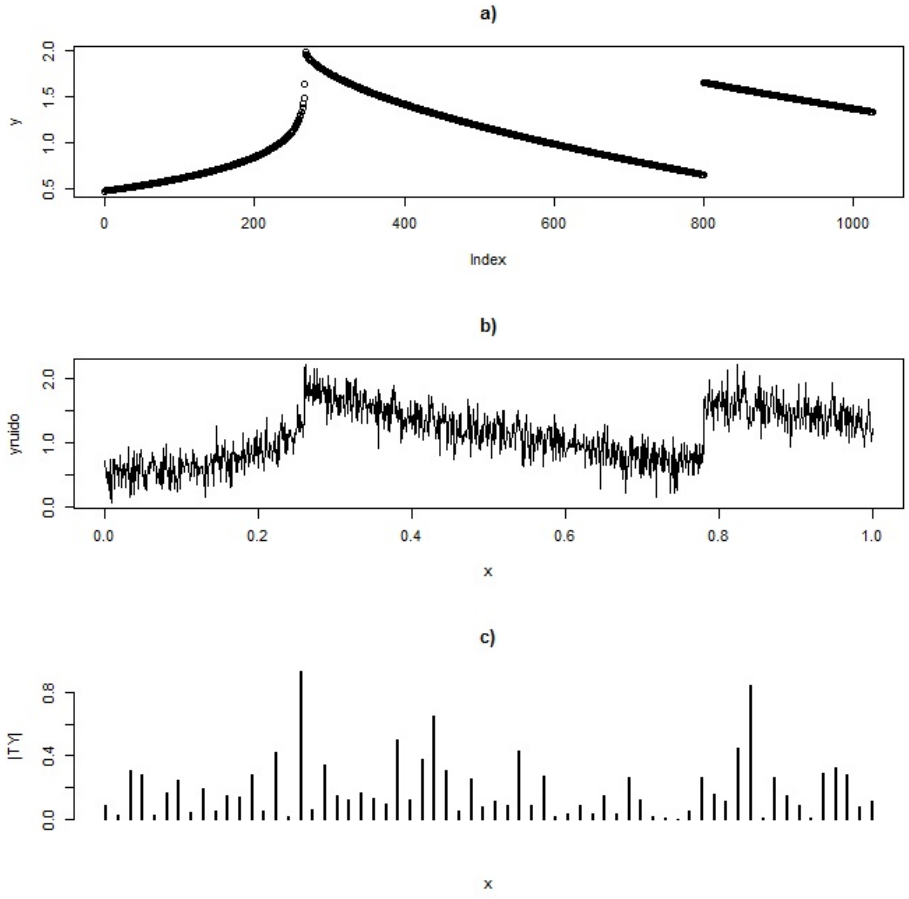


Figure 4.1: *Simulated data from the model* $y_i = f(i/n) + \epsilon_i$, $f(x) = 2 - 2\,|x - 0.26|^{1/5}\, I(x \leq 0.26) - 2\,|x - 0.26|^{3/5}\, I(x > 0.26) + I(x \geq 0.78)$, $\epsilon_i \sim N(0, \sigma^2)$, $\sigma = 0.2$ *e* $n = 1024$*; a) real curve b) curve with noise and c) Absolute value of the wavelet coefficients, scale* $a = 7$.

When checking the wavelet coefficients on the 10 scales, we can find dyadic intervals in some scales, whose corresponding absolute value of the wavelet coefficient exceeds the threshold, and are significantly higher than the others. In Figure 4.1 (c), the wavelet coefficients are significantly large and exceed the threshold line, only where the functions have jump and peak. With $n = 1024$ in the example, we increased σ from 0.2 to 1 in steps of 0.2. Detection works well for values of σ up to 0.4. After that, the jump and the peak become difficult to detect. The jump and the peak tend to be detected by wavelet coefficients at lower levels of resolution and thus the detection is increasingly accurate. In particular, the peak is often located at low resolution levels or it may not be detected. For $\sigma \geq 1$ the method fails and jumps and spikes are difficult to detect. It can be seen in Table 4.1 that in the scale $a = 4$

Table 4.1: *Estimated values with* $\sigma = 0.2$ *of the number of jumps , jump localization (t), jump size (Z) and jump variation* (Φ)

a (Scale)	# Jumps	t (Jump Loc)	Z (Jump size)	Φ
4	1	832	0,02	0,0004
5	1	288	-0,005	0,000025
6	1	272	-0,16	0,0256

the method can detect a jump in the observation 832, note that we have to consider an error and the value of the observation can be around that value, so for the scales $a = 5, 6$, one jump was detected in each scale in the observations 288 and 272 respectively. With $\sigma = 0.2$ it was observed that no scale could detect the two jumps together.

Table 4.2: *Estimated values with* $\sigma = 0.1$ *of the number of jumps, jump localization (t), jump size (Z) and jump variation* (Φ)

a (Scale)	# Jumps	t (Jump Loc)	Z (Jump size)	Φ
4	1	832	0,03	0,0009
5	1	**288**	-0,38	0,1444
6	1	272	-0,2	0,04
8	2	**288 800**	-0,30 -0,0066	0,09
9	1	**800**	0,064	0,004096

For $\sigma = 0.1$ in Table 4.2 the method was able to detect the two jumps together in the scale $a = 8$. We can say that the function jumped in the observations 288 and 800 as shown in Figure 4.2 c) of the wavelet coefficients in the $a = 8$ scale.

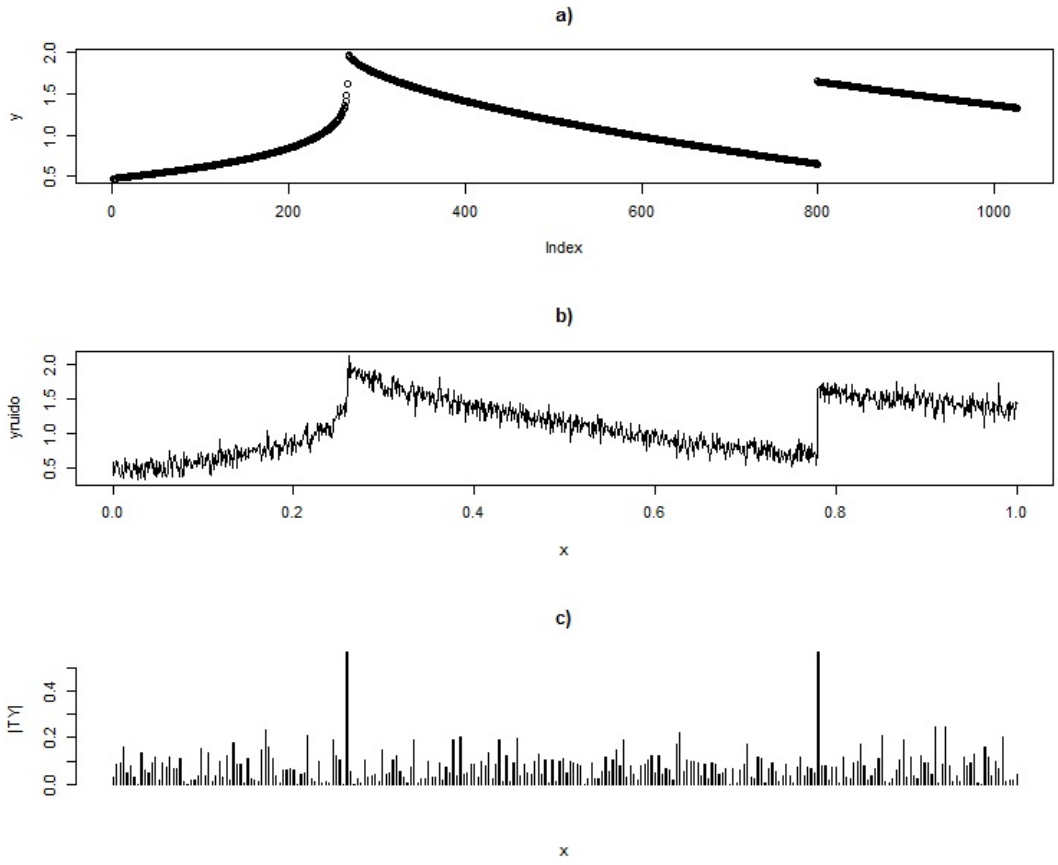


Figure 4.2: *Simulated data from the model* $y_i = f(i/n) + \epsilon_i$, $f(x) = 2 - 2\,|x - 0.26|^{1/5} I(x \leq 0.26) - 2\,|x - 0.26|^{3/5} I(x > 0.26) + I(x \geq 0.78)$, $\epsilon_i \sim N(0, \sigma^2)$, $\sigma = 0.1$ *e* $n = 1024$*; a) real curve b) curve with noise and c) Absolute value of the wavelet coefficients, scale* $a = 8$

4.1.2 Simulating a jump process

When the Lévy process has a Gaussian component and a jump component of a composite Poisson process, the two independent components can be simulated separately. Figure 4.3 shows a path of discrete data for such a process. Here, in this example, the jump size has normal distribution with zero mean and standard deviation 0.5, the jump intensity is 10, the diffusion volatility is 1 and there is not drift, that is, $\mu = 0$. The discretized trajectory has the form:

$$X(t_i) = \mu t_i + \sum_{k=1}^{i} Z_k + \sum_{j=1}^{N} I_{U_j<t_i} Y_j.$$

Here Z_i is a normal random variable with variance $Var(Z_i) = (t_i - t_{i-1})\sigma^2$, $t_0 = 0$. The third term is simulated as a compound Poisson process below.

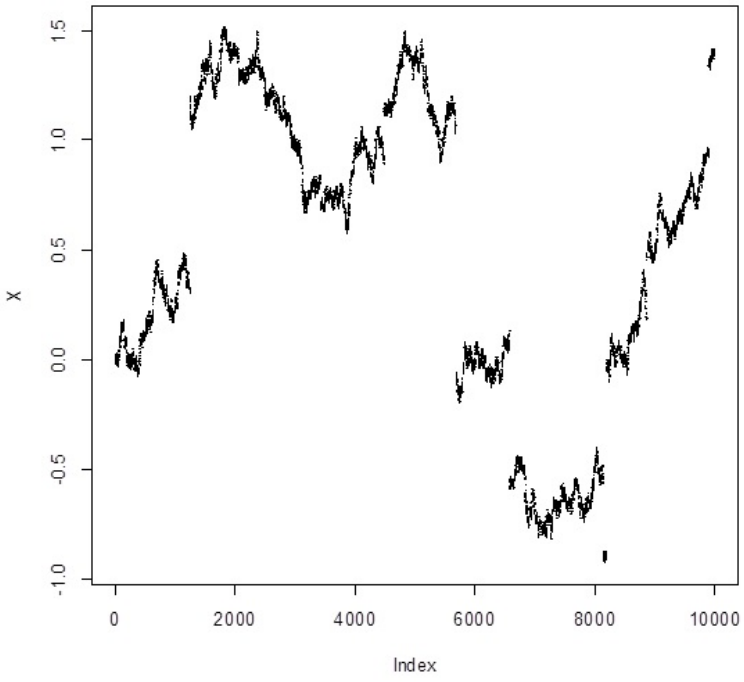


Figure 4.3: *Simulating jump of a diffusion process in a fixed time grid*

On the other hand we have that the trajectory

$$X(t) = \mu t + \sum_{i=1}^{N(t)} Y_i.$$

where $N(t) = sup\left\{k : \sum_{i=1}^{k} T_i \leq t\right\}$. This is a typical trajectory of a compound Poisson process. Here the jump size has standard normal distribution, the jump intensity is 10 and the drift is 3 (See Figure 4.4).
We will see that many infinite activity Lévy processes can be well approximated by a process of such type: the small jumps are truncated and replaced with a properly renormalized Brownian motion (see Figure 4.5).

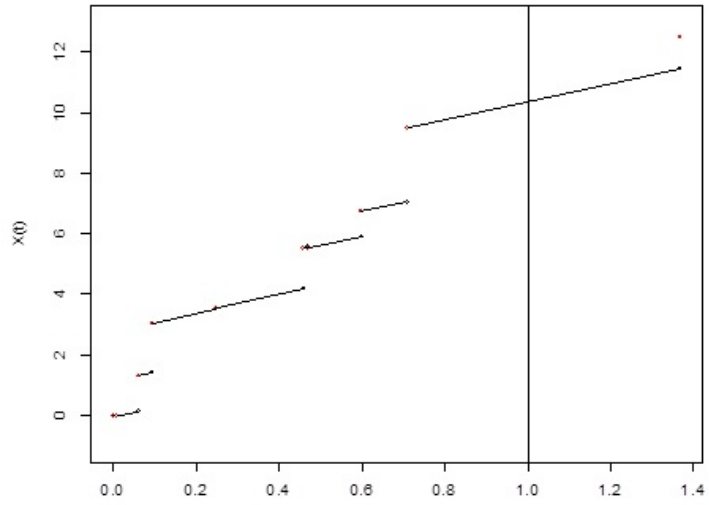


Figure 4.4: *Simulating a compound Poisson process with Brownian Motion.*

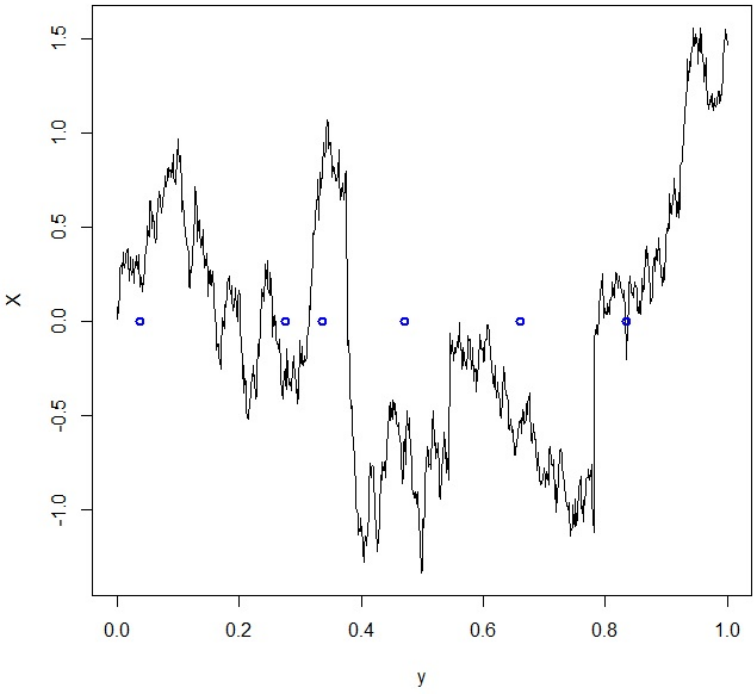


Figure 4.5: *Simulating a compound Poisson process*

4.2 Simulation Case 2.

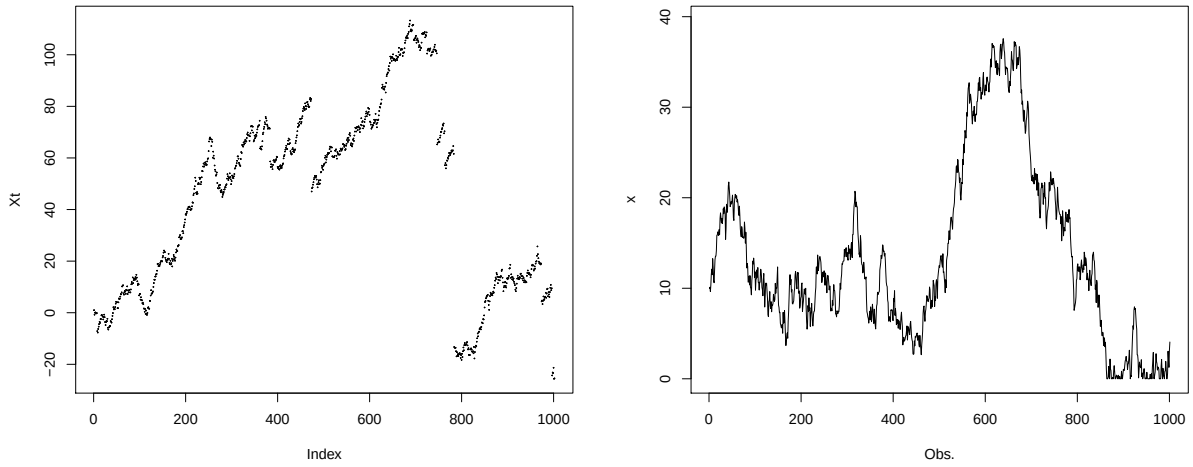
In this section, we conduct simulations studies to check the performance of the Cv-M test. We consider the following two models. In the first scenario, we have the following stochastic volatility model:

$$\begin{aligned} X_t &= X_0 + \int_0^t \sqrt{c_s} dW_s + 0.5 Z_t,\ 0 \le t \le T, \\ c_t &= c_0 + \int_0^t 0.03(1.0 - c_s) ds + 0.15 \int_0^t \sqrt{c_s} dW'_s, \end{aligned}$$

where W_s, W'_s are independent Wiener processes and Z_t is a skewed β stable Lévy process. The volatility c_t is a square root diffusion process which is widely used in financial applications. The parameter in c_t is specified as in Jacob and Todorov (2014). The second model is

$$\tilde{X}_t = X_{t-1} + a_t,\ a_t \sim N(0,1), \tag{4.1}$$

where X_0 is the initial value that should defined.



(a) X_t, *jump diffusion model.* **(b)** $\tilde{X}_t$, *standard normal model.*

Figure 4.6: *Example of time series generate with the two models.*

We considered $10,000$ replicas for sample sizes $n = 500, 1,000, 2,000, 5,000$ and $10,000$. For the truncation of the increments, as is typical in the literature, we set $\alpha = 3.0$ and $\bar{w} = 0.49$. Hence for the sampling frequencies mentioned above, we set the pair of (k_n, m_n) to be $(22, 12)$, $(30, 19)$, $(44, 22)$, $(70, 34)$ and $(100, 54)$ with k_n/m_n ranging from 1.5 to 2.15 and having an increasing trend.

In Table 4.3 we can observe that the test power increases as the sample size increases, and also for a fixed n the test power is bigger for largest k_n values. The test power was calculated with the gamma quantile values, and we note that these values are very close to the sample quantile.

To evaluate the performance of our test, we compare with another test that also measures driscrepancy between distributions, the Kolmorogov-Smirnov (KS) test. In Table 4.4 we note that our test shows better results. Besides, the KS test power decreases and the error rate

n	k_n	m_n	a	b	Quantile of $gamma(a,b)$ 0.95	Sample Quantile 0.95	Test Power
500	4	2	6.323	0.087	0.962	0.905	0.564
	6	3	2.302	0.165	0.865	0.689	0.615
	8	4	3.592	0.096	0.693	0.684	0.818
	18	11	2.254	0.190	0.982	0.968	0.956
	20	11	2.109	0.238	1.174	1.165	0.928
	22	12	1.938	0.274	1.273	1.243	0.930
1000	4	2	13.991	0.049	1.017	1.004	0.875
	6	3	6.924	0.060	0.709	0.680	0.958
	8	4	5.350	0.064	0.623	0.621	0.981
	26	13	2.083	0.231	1.128	1.116	0.999
	28	13	1.993	0.267	1.266	1.270	0.997
	30	19	2.064	0.228	1.106	1.084	0.999
2000	4	2	27.702	0.032	1.202	1.196	0.990
	6	3	14.850	0.033	0.736	0.731	0.999
	8	4	9.147	0.041	0.612	0.607	0.999
	40	18	1.922	0.283	1.308	1.296	1
	42	22	1.972	0.267	1.259	1.246	1
	44	22	1.953	0.282	1.320	1.304	1
5000	4	2	69.609	0.019	1.633	1.631	1
	6	3	35.490	0.020	0.921	0.920	1
	8	4	21.383	0.023	0.694	0.694	1
	66	36	1.993	0.261	1.239	1.251	1
	68	38	1.974	0.264	1.241	1.247	1
	70	34	1.885	0.308	1.403	1.380	1
10000	4	2	132.233	0.014	2.155	2.154	1
	6	3	70.155	0.013	1.171	1.170	1
	8	4	41.810	0.015	0.839	0.838	1
	96	62	1.947	0.253	1.181	1.179	1
	98	48	1.894	0.301	1.379	1.357	1
	100	54	1.972	0.284	1.334	1.322	1

Table 4.3: *Comparison between the sampling and theoretical quantiles for $TA2$.*

increases as n increases. For $n = 500$ the KS test power is bigger than $TA2$ power, however $TA2$ has error rate 5% and the KS test has an error rate bigger than 76%. For larger sample sizes, $TA2$ shows better power with low error rate.

In Table 4.5 we can find the quantile values of gamma and sample distribuiton for different probabilities. Therefore, note that the theoretical quantiles compared with the sample quantiles are closer for probabilities higher than 0.90. Often in hypotheses test it is used a significance level less than 0.10, so we can use the gamma distribution to choose a critical value for the test.

n	k_n	m_n	a	b	Quantile of $gamma(a,b)$ 0.95	Sample Quantile 0.95	Test Power	Test Power KS	Error Rate KS
500	4	2	6.323	0.087	0.962	0.905	0.564	0.958	0.763
	6	3	2.302	0.165	0.865	0.689	0.615	0.958	0.963
	8	4	3.592	0.096	0.693	0.684	0.818	0.959	0.987
	18	11	2.254	0.190	0.982	0.968	0.956	0.960	0.999
	20	11	2.109	0.238	1.174	1.165	0.928	0.958	1
	22	12	1.938	0.274	1.273	1.243	0.930	0.960	1
1000	4	2	13.991	0.049	1.017	1.004	0.875	0.911	0.691
	6	3	6.924	0.060	0.709	0.680	0.958	0.912	0.913
	8	4	5.350	0.064	0.623	0.621	0.981	0.911	0.974
	26	13	2.083	0.231	1.128	1.116	0.999	0.913	1
	28	13	1.993	0.267	1.266	1.270	0.997	0.914	1
	30	19	2.064	0.228	1.106	1.084	0.999	0.913	1
2000	4	2	27.702	0.032	1.202	1.196	0.990	0.836	0.528
	6	3	14.850	0.033	0.736	0.731	0.999	0.836	0.897
	8	4	9.147	0.041	0.612	0.607	0.999	0.836	0.977
	40	18	1.922	0.283	1.308	1.296	1	0.836	1
	42	22	1.972	0.267	1.259	1.246	1	0.838	1
	44	22	1.953	0.282	1.320	1.304	1	0.837	1
5000	4	2	69.609	0.019	1.633	1.631	1	0.635	0.462
	6	3	35.490	0.020	0.921	0.920	1	0.635	0.818
	8	4	21.383	0.023	0.694	0.694	1	0.635	0.959
	66	36	1.993	0.261	1.239	1.251	1	0.638	1
	68	38	1.974	0.264	1.241	1.247	1	0.637	1
	70	34	1.885	0.308	1.403	1.380	1	0.637	1

Table 4.4: *Comparison between our test statistic $TA2$ and KS test.*

Table 4.5: *Percentage points for $TA2$. Entries in the table are x such that $P\{d^2\sqrt{\lfloor n/2k_n \rfloor\, m_n} \leq x\} = P\{TA2 \leq x\} = p$. The upper number in a double entry is the critical value calculated by using the gamma distribution; the lower is based on the sample distribution*

n	Percentage points for the following values of p:																n
	0.01	0.025	0.05	0.1	0.15	0.20	0.25	0.50	0.75	0.80	0.85	0.90	0.95	0.975	0.99	0.999	
500	0.037	0.061	0.090	0.137	0.177	0.215	0.251	0.443	0.717	0.798	0.901	1.041	1.273	1.498	1.788	2.495	500
	0.121	0.145	0.167	0.201	0.231	0.259	0.285	0.430	0.650	0.722	0.825	0.964	1.243	1.589	2.054	3.194	
1000	0.037	0.059	0.087	0.129	0.164	0.198	0.230	0.397	0.633	0.702	0.790	0.909	1.106	1.297	1.542	2.138	1000
	0.107	0.129	0.152	0.186	0.212	0.235	0.259	0.382	0.569	0.641	0.727	0.859	1.084	1.361	1.754	2.700	
2000	0.040	0.064	0.095	0.143	0.185	0.224	0.262	0.461	0.745	0.829	0.935	1.080	1.320	1.553	1.853	2.583	2000
	0.117	0.142	0.169	0.206	0.238	0.267	0.296	0.443	0.679	0.759	0.863	1.023	1.304	1.625	2.072	3.278	
5000	0.037	0.0624	0.096	0.145	0.189	0.230	0.270	0.482	0.785	0.875	0.989	1.145	1.403	1.654	1.978	2.769	5000
	0.119	0.148	0.176	0.217	0.250	0.279	0.306	0.463	0.712	0.794	0.907	1.090	1.380	1.693	2.235	3.403	
10000	0.040	0.066	0.097	0.147	0.189	0.228	0.267	0.468	0.754	0.840	0.947	1.093	1.334	1.569	1.871	2.606	10000
	0.118	0.141	0.169	0.209	0.242	0.270	0.298	0.449	0.692	0.770	0.880	1.045	1.322	1.610	2.048	3.252	

Chapter 5

Real Data Analysis

5.1 Real Data Analysis, Case 1

The application was made for the Google Stock. The period considered is November 11th to November 12th in 2014, for a total of 27.511 observations. For this analysis we will choose the first 1024 observations. The price studied was the price of daily closing of the index. The following is the graph of the wavelet coefficients (see Figure 5.2). Here the number of

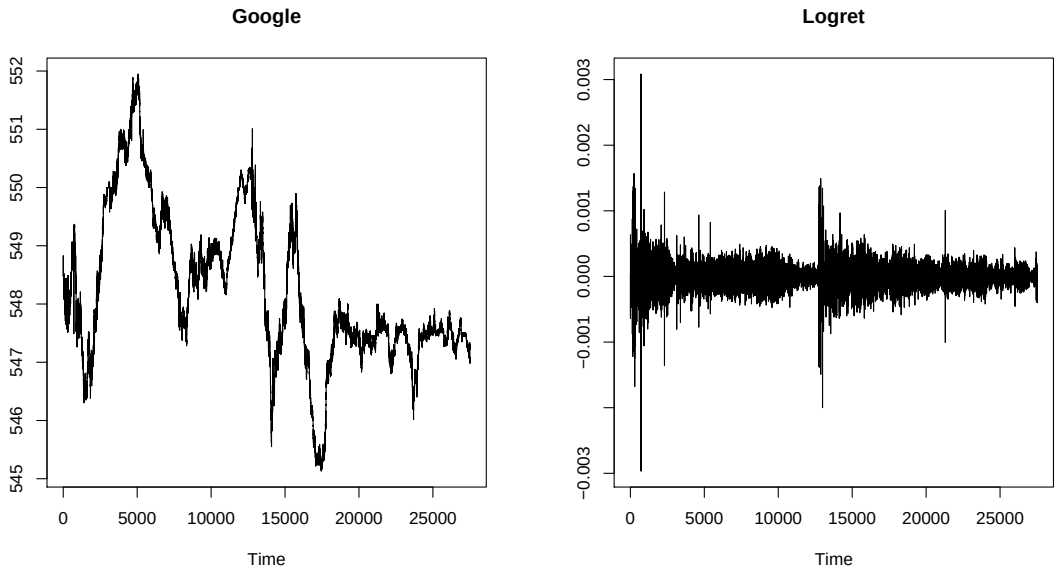


Figure 5.1: *Series of the Google and Log-return in the period November 11th to November 12th in 2014.*

coefficients is given by $2^a = 2^{10} = 1024$, so we would have 10 resolution levels, $a = 0, ..., 9$. The coefficients will not be shown, since the vector d has 1024 components and the matrix W also will not be shown, since it has dimension 1024×1024. The following table shows the different jumps that were detected in the series by means of the wavelet method. In Figure 5.3 the wavelet coefficients for the scale $a = 3$ are significantly large and exceed the threshold line, it is visible where the function detected a jump in observation 896. The jumps tend to be detected by wavelet coefficients at lower levels of resolution and thus the detection is increasingly accurate. In particular, the estimation of the jump 896 was detected in two scales, for $a = 3, 4$ and for the estimation of the jump 832 was also detected in two scales $a = 4, 5$ (see Figure 5.5).

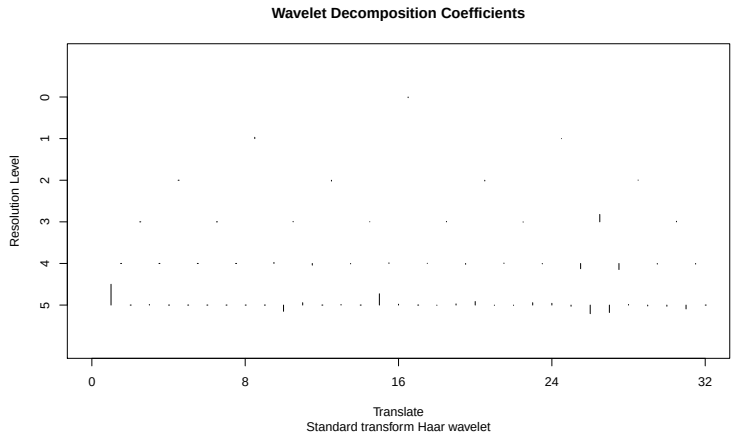


Figure 5.2: *Wavelet coefficients for Google stock*

Table 5.1: *Estimated values of number jumps, jump localization (t), jump size (Z) and jump variation* (Φ) *at different scales.*

a (Scale)	# Jumps	t (Jump Loc)	Φ
3	1	**896**	3.742584e-10
4	2	**832 896**	2.520881e-08
5	7	32 320 480 640 **832** 864 992	2.664882e-07

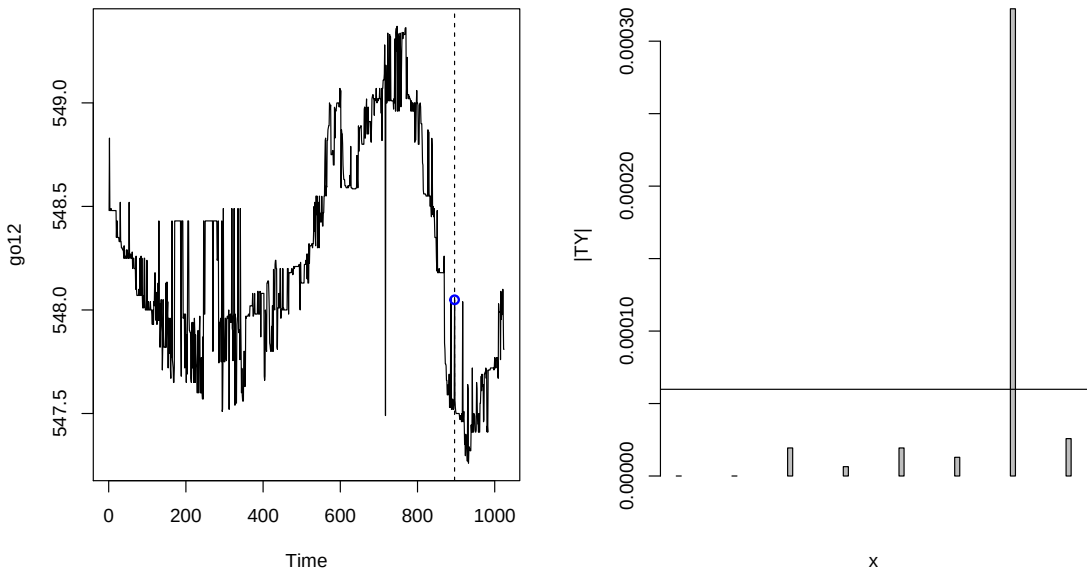


Figure 5.3: *Google Stock and wavelet coefficients,* $a = 3$.

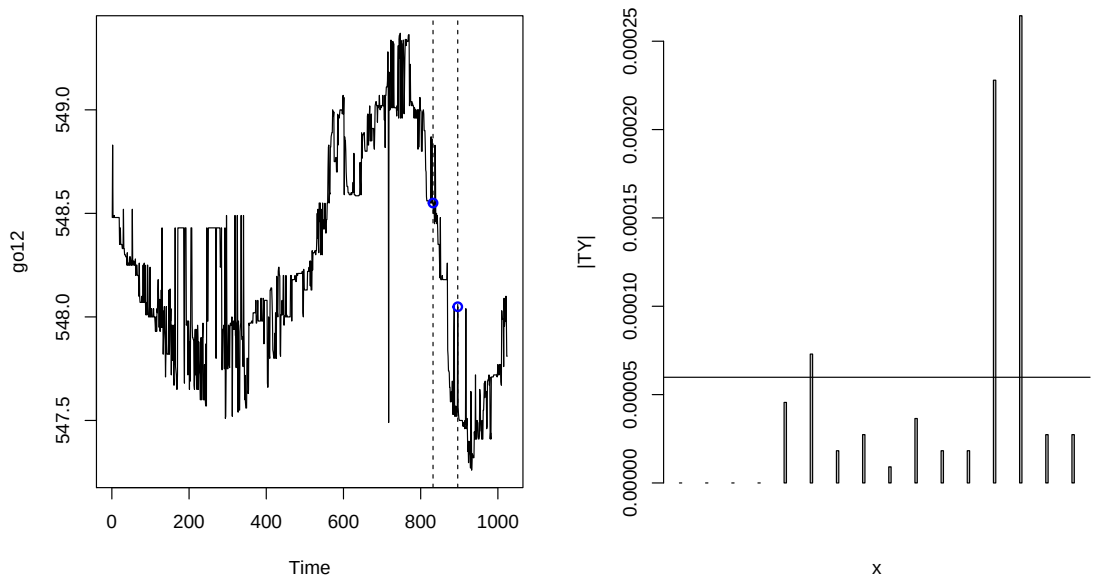


Figure 5.4: *Google Stock and wavelet coefficients,* $a = 4$.

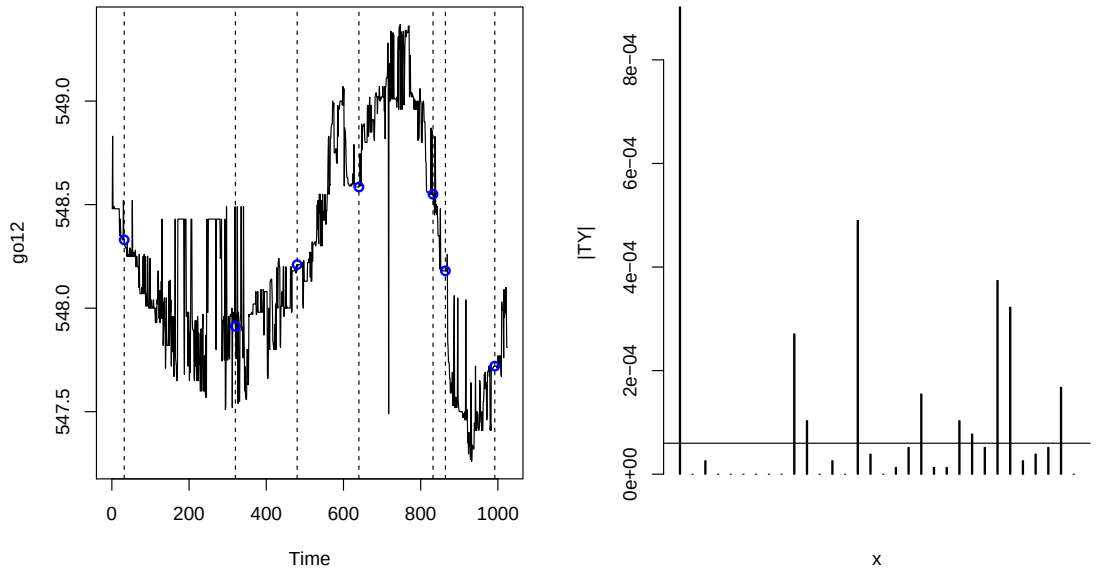
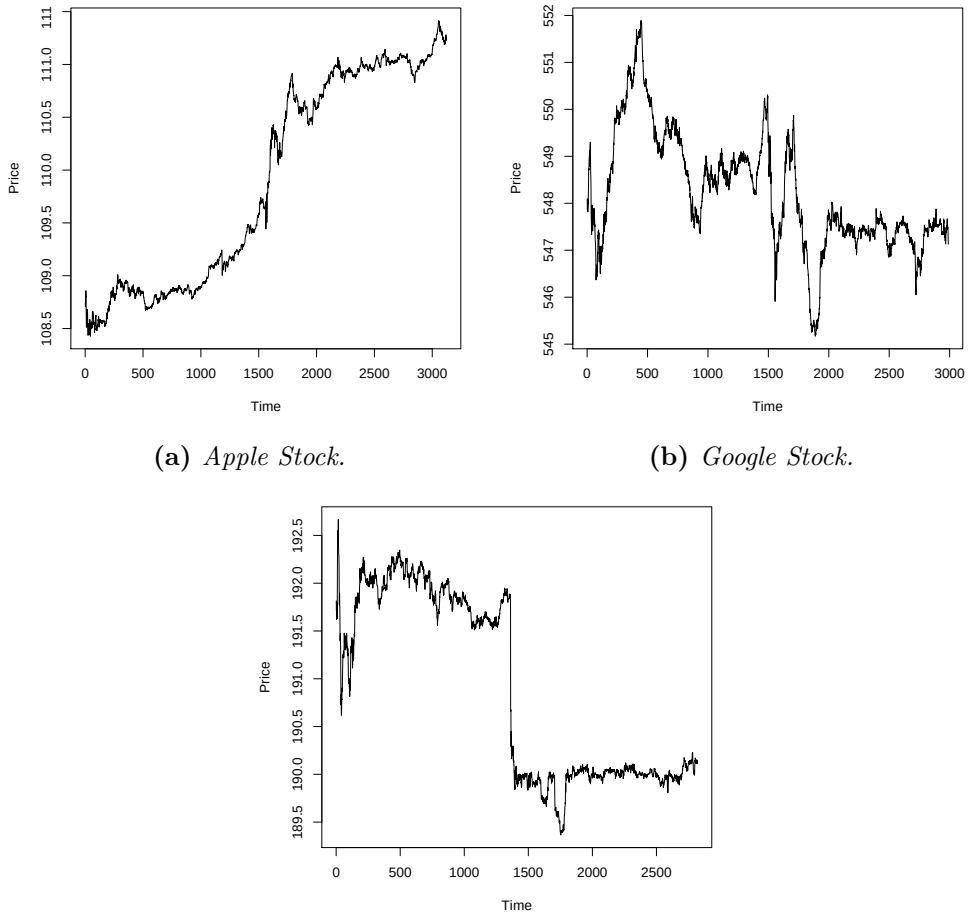


Figure 5.5: *Google Stock and wavelet coefficients,* $a = 5$.

5.2 Real Data Analysis, Case 2

In this section, we collect intraday transaction prices of the Google, Apple and Goldman Sachs (GS) stocks, respectively, from November 11th to November 12th in 2014, with a sampling frequency to every 15 seconds. The transaction records are excluded if they are outside the ranges of quotes, in our case the range is 9:30 a.m to 4:00 p.m. There are in total 3122, 2898 and 2819 stock prices, respectively. We aim to plot the observed test statistics against different values of k_n and m_n.



(a) *Apple Stock.*

(b) *Google Stock.*

(c) *Goldman Sachs Stock.*

Figure 5.6: *Time series of the data from* $9:30$ *a.m to* $4:00$ *p.m.*

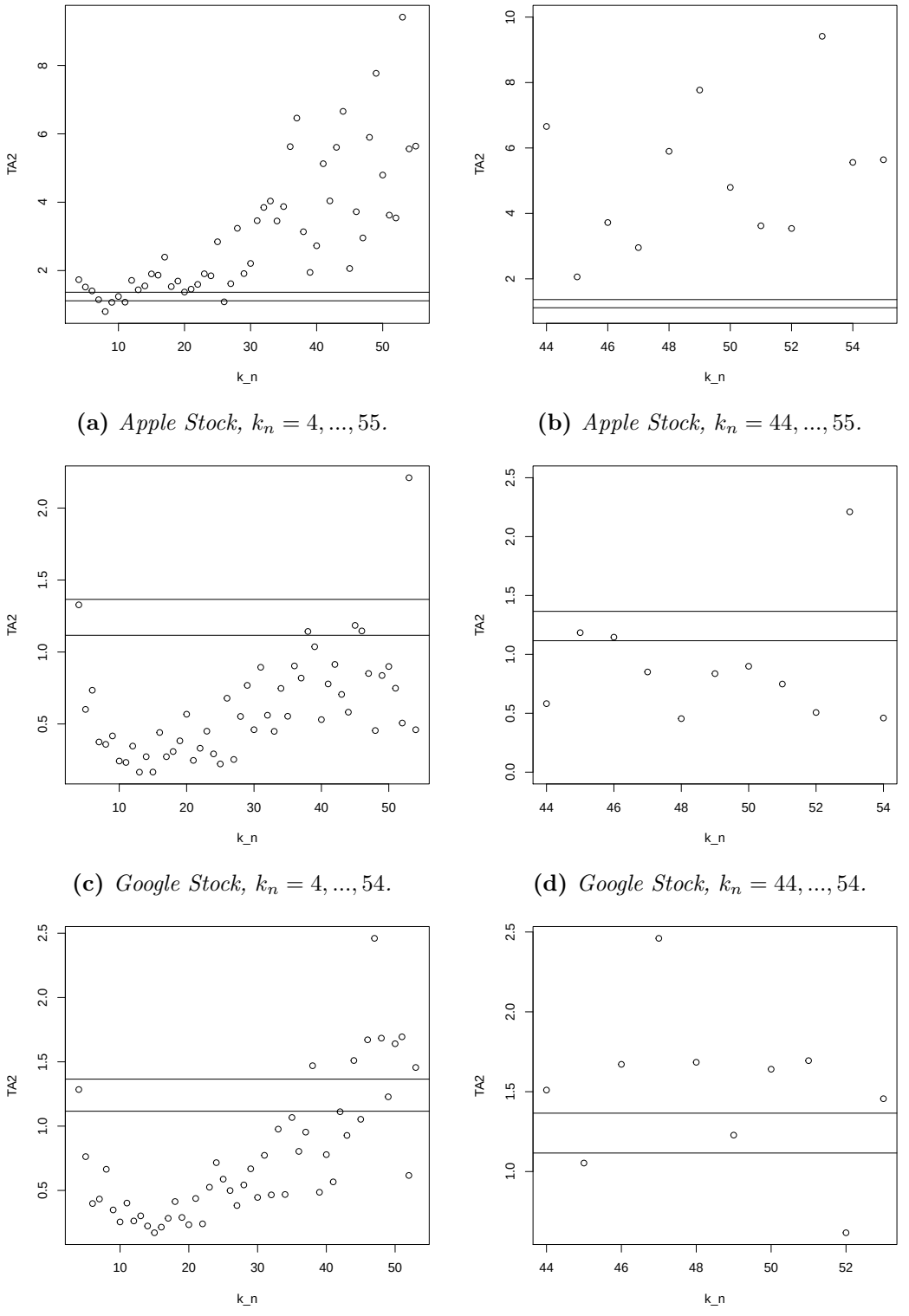
Figure 5.6 shows for three differents stock market data. Figure 5.7 shows the values of our test statistic, compared to the values of k_n. Additionally, we illustrate the quantile values at the significance level of 5% and 10%. The graphics on the right of Figure 5.7 consider the values of k_n between $[4, 55]$, $[4, 54]$ and $[4, 53]$ respectively. The graphics on the left show k_n values starting from 44. It can be observed that at a level of significance of 5%, Apple stock rejects null hypotheses 88%, Google and GS stocks also reject with 98% and 84% respectively. On the other hand, when we increase the level of significance to 10%, the percentage increases for all stocks, with Apple still rejecting with 92%, Google reject with 90% and GS also reject with 80%. This using all k_n values.

Stock	n	k_n	m_n	$\alpha = 5\%(q = 1.361)$	$\alpha = 10\%(q = 1.112)$
Apple	3122	[4,55]	[2,27]	88%	92%
Google	2989	[4,54]	[2,27]	98%	90%
GS	2819	[4,53]	[2,26]	84%	80%

Table 5.2: *Rejection rate of the test, considering all possible* k_n *values.*

As might be expected, when we use more subdivisions of our interval, that is; the value of k_n, and the greater the value of k_n, our test statistic tends to detect more precisely the dynamics of the series. This can be seen in Table 5.3.

We construct a test statistic that capture the dynamics of a series using high frequency



(a) *Apple Stock,* $k_n = 4, ..., 55$.

(b) *Apple Stock,* $k_n = 44, ..., 55$.

(c) *Google Stock,* $k_n = 4, ..., 54$.

(d) *Google Stock,* $k_n = 44, ..., 54$.

(e) *Goldman Sachs Stock,* $k_n = 4, ..., 53$.

(f) *Goldman Sachs Stock,* $k_n = 44, ..., 53$

Figure 5.7: $TA2$ *statistic for different* k_n.

transaction prices. We compare the test power of our statistic with the KS test statistic and the performance of $TA2$ show better results. The simulation show that we can approximate our c.d.f. of the test statistic $TA2$ with the gamma distribution, so we can use it in practice to apply the test.

The empirical data analysis shows that our statistic $TA2$ is useful to identify if the

Stock	n	k_n	m_n	$\alpha = 5\%(q = 1.361)$	$\alpha = 10\%(q = 1.112)$
Apple	3122	[44,55]	[22,27]	100%	100%
Google	2989	[44,54]	[22,27]	91%	73%
GS	2819	[44,53]	[22,26]	30%	20%

Table 5.3: *Rejection rate of the test, considering k_n values that are at least* 44.

variation in the series is different from a standard normal variation in high frequency data of prices. For Apple and GS stock, our test indicates that there is not a standard normal variation for the largest k_n values.

Chapter 6

Conclusion

In this work, we first discussed a method proposed by Fan and Wang (2007) for detection and localization of the jumps in different scales. We also show that the jumps tend to be detected by wavelet coeffients at lower levels of resolution and thus detection is increasingly accurate. When checking the wavelet coefficients on the differents scales, we can find dyadic intervals in some scales, whose corresponding absolute value of the wavelet coefficient exceeds the threshold, and are significantly higher than the others. We applied the procedure for the Google Stock, and we detected 3 scales for the estimation of the jumps.

In Chapter 2 a volatility model with jumps is entertained and a wavelet analysis is proposed.

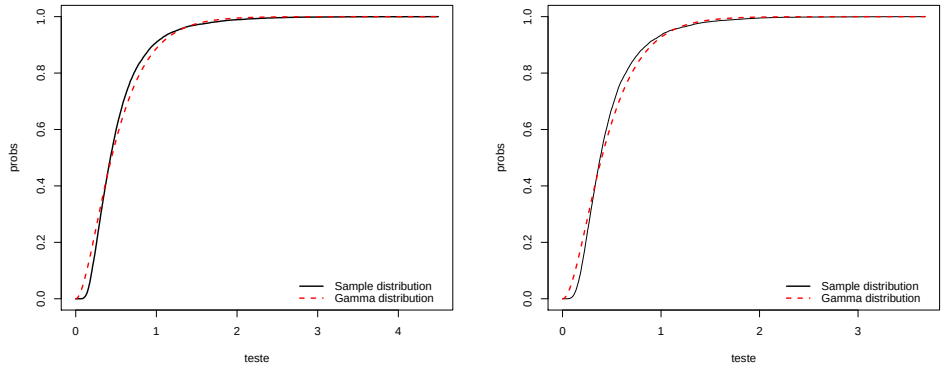
We also construct a test statistic that capture the dynamics of a series using high frequency transaction prices. We compare the test power of our statistic with the Kolmogorov-Smirvov test statistic and $TA2$ shows better results.
The simulation show that we can approximate the c.d.f. of the test statistic $TA2$ with the gamma distribution, so we can use it in practice to apply the test. The empirical data analysis shows that the statistic $TA2$ is useful to identify if the variation in the series is different from a standard normal variation in high frequency data of prices. For Apple and GS stock, the test indicates that there is not a standard normal variation for the largest k_n values.

Also, it can be observed that at a level of significance of 5%, Apple stock rejects null hypotheses 88%, Google and GS stocks also reject with 98% and 84% respectively. On the other hand, when we increase the level of significance to 10%, the percentage increases for all stocks, with Apple still rejecting with 92%, Google reject with 90% and GS also does not reject with 80%. This using all k_n values.

Future research topics include the following:

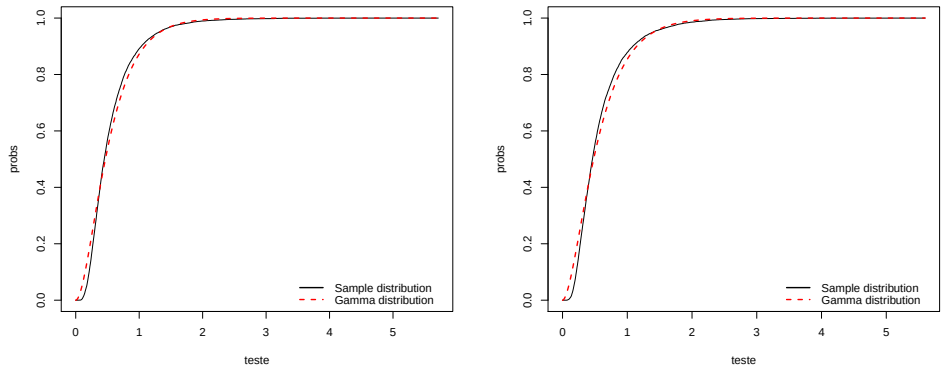
- Consider other estimators for variance and also other cumulative distribution function.
- In the hypothesis test consider other types of models for H_0 and H_1.
- In this work our test was compared with the Kolmogorov-Smirnov test, so we can compare the performance of our test statistic with other tests.
- Study the sensitivity of the variables k_n and m_n.
- As a challenge we might consider finding the closed form of our test statistic.

Appendix A



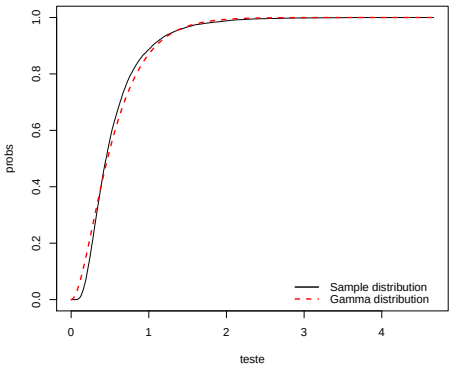
(a) $k_n = 22$, $n = 500$.

(b) $k_n = 30$, $n = 1000$.



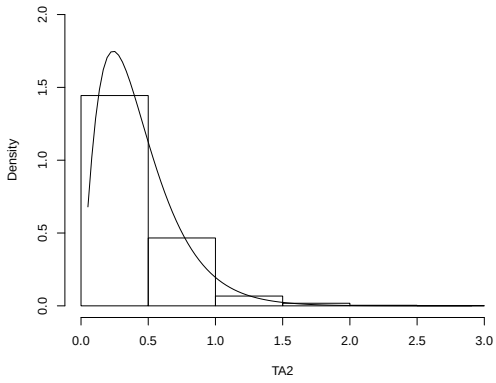
(c) $k_n = 44$, $n = 2000$.

(d) $k_n = 70$, $n = 5000$.

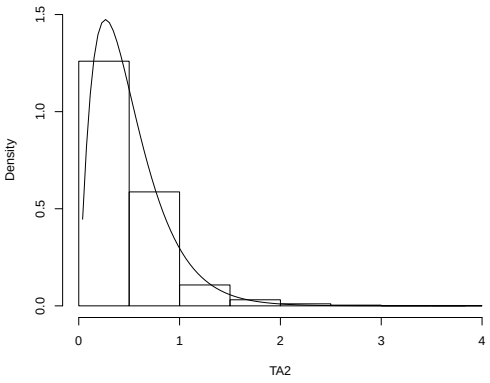


(e) $k_n = 100$, $n = 10000$.

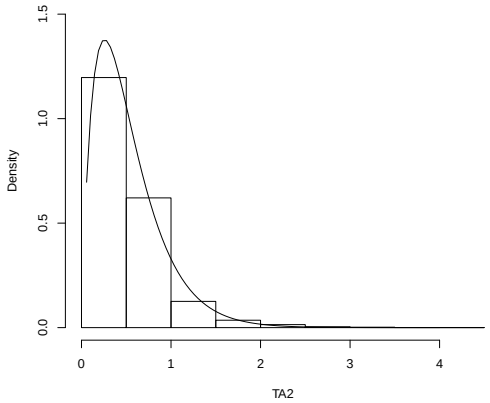
Figure A.1: *Comparison between Sampling cumulative distribution function and Gamma cumulative distribution function.*



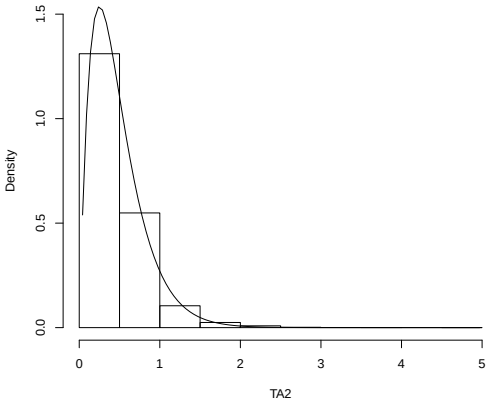
(a) $k_n = 18$, $n = 500$.



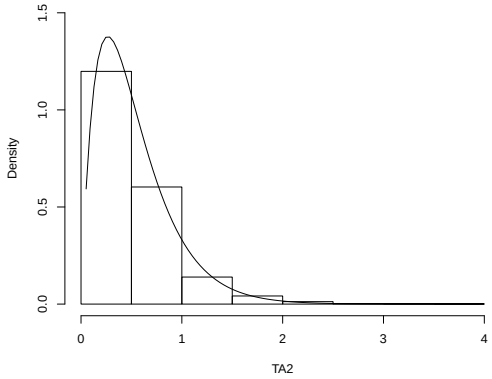
(b) $k_n = 20$, $n = 500$.



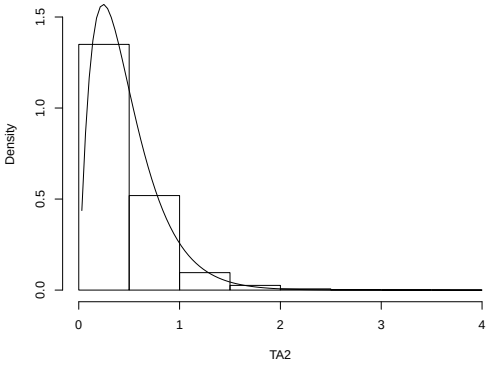
(c) $k_n = 22$, $n = 500$.



(d) $k_n = 26$, $n = 1000$.

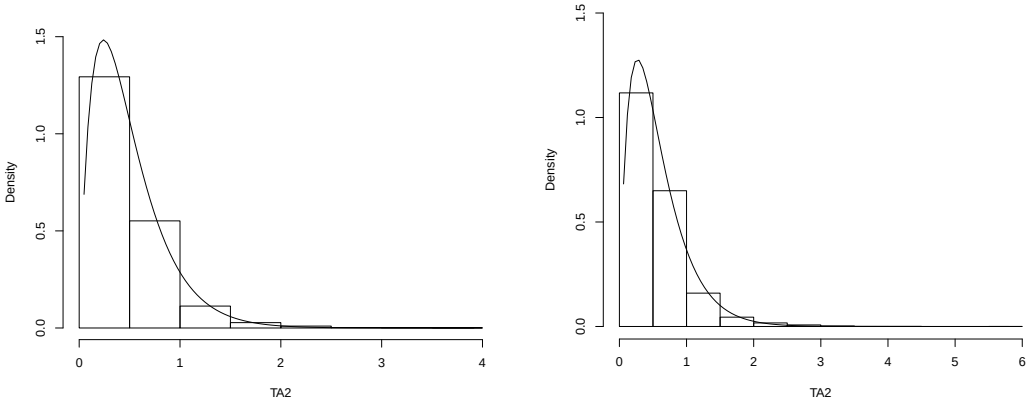


(e) $k_n = 28$, $n = 1000$.



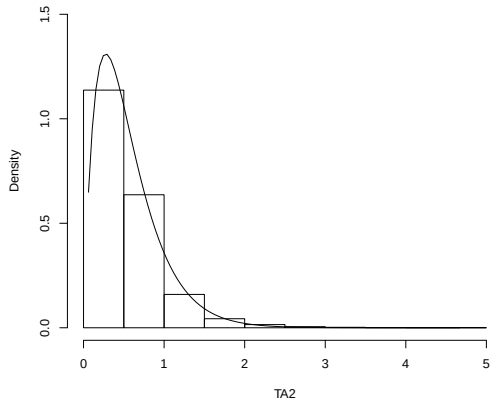
(f) $k_n = 30$, $n = 1000$.

Figure A.2: *Histograms of the sample density function for different sizes of k_n and n, compared with the gamma density.*



(a) $k_n = 96$, $n = 10000$.

(b) $k_n = 98$, $n = 10000$.



(c) $k_n = 100$, $n = 10000$.

Figure A.3: *Histograms of the sample density function for different sizes of k_n and n, compared with the gamma density.*

Bibliography

Ait-Sahalia and Jacod(2009b) Y Ait-Sahalia and J. Jacod. Testing for jumps in discretly observed process. *Ann. Stat*, 37:184–222. 1

Ait-Sahalia and Jacod(2011) Y. Ait-Sahalia and J. Jacod. Testing whether jumps have finite or infinite activity. *Ann. Stat*, 39:1689–1719. 1

Ait-Sahalia and Jacod(2014) Y. Ait-Sahalia and J. Jacod. High-frequency financial econometric. *Princeton University Press*. 1

Ait-Sahalia and Jacod(2010) Y Ait-Sahalia and J. Jacod. Is brownian motion necessary to model high-frequency data?. *Ann. Stat*, 38:3093–3128. 1

Andersen and Labys(2003) Bollerslev T. Diebold F. X Andersen, T. G. and P. Labys. Modeling and forecasting realized volatility. *Econometrica*, 71:579–625. 1

Barndorff-Nielsen and Shephard(2006) O. E. Barndorff-Nielsen and N. Shephard. Econometrics of testing for jumps in financial economics using bipower variation. *Journal Financial Econometrics*, 2:1–48. 1

Barndorff-Nielsen and Shephard(2002) O. E. Barndorff-Nielsen and N. Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society Series B*, 64:253–280. 1, 7

Black and Scholes(1973) F Black and M. Scholes. The pricing of options and corporate liabilities. journal of political economy. *Journal of Political Economy*, 81:635–654. 7

Csorgo and Faraway(1996) S. Csorgo and J. Faraway. The exact and asymptotic distributions of cramér-von mises statistics. *J. R. Statist. Soc. Series B*, 58:221–234. 2

Dohono and Johnstone(1995) D.L Dohono and I.M. Johnstone. Adapting to unknown smoothness via wavelet shrinking. *Journal of the American Statistical Association*, 90: 1200–1224. 14

Engle(1982) R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50:987–1008. 1

Fan and Wang(2007) J. Fan and Y. Wang. Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association*, 102: 1349–1362. 2, 3, 33, 49

Fan and Wang(2008) J. Fan and Y. Wang. Spot volatility estimation for high-frequency data. *Stat. Interf*, 1:279–288. 1

Jing and Liu(2012b) Kong X. B. Jing, B. Y. and Z. Liu. Modeling high-frequency financial data by pure jump models. *Ann. Stat*, 40:759–784. 1

Kong and Jing(2015) Liu Z. Kong, X. B. and B. Y. Jing. Testing for pure-jump processes for high-frequency data. *Ann. Stat*, 43:847–877. 1

Kong(2017) X. B. Kong. Lack of fit test for infinite variation jumps at high frequencies. *Statistica Sinica*. 1, 3, 4, 18, 30

Morettin(2014) P. A. Morettin. *Ondas e Ondaletas.* Edusp, segunda edição. 4, 5

Morettin(2017) P. A. Morettin. *Econometria Financeira.* Blucher, terceira edição. 1

Muller and Stadtmuller(2011) R Muller, H.G.; Sen and U. Stadtmuller. Functional data analysis for volatility. *Journal of Econometrics*, 165:233–245. 2, 7, 8, 9, 11

Prakasa Rao(1999) B.L.S. Prakasa Rao. *Statistical Inference for Diffusion Type Processes*. London: Arnold. 2

Raimondo(1998) M. Raimondo. Minimax estimation of sharp change points. *The Annals of Statistics*, 26:1379–1397. 2, 11, 12

Todorov(2015) V. Todorov. Jump activity estimation for pure-jump semimartingales via self-normalized statsitics. *Ann. Stat*, 43:1831–1864. 1

Todorov and Tauchen(2014) V. Todorov and G. Tauchen. Limit theorems for the empirical distribution function of scaled increments of ito semimartingales at high frequencies. *Ann. Appl. Prob*, 24:1850– 1888. 1, 2, 3, 6, 17, 18

Tsay(2005) R. S. Tsay. *Analysis of Financial Time Series.* Wiley, third edition edição. 1

Wang(1995) Y. Wang. *Jump and sharp cusp detection via wavelets*. Biometrika, 82 edição. 2, 6, 11, 12, 33