Journal of Mathematics Research; Vol. 9, No. 3; June 2017
ISSN 1916-9795 E-ISSN 1916-9809
Published by Canadian Center of Science and Education

Minimum Hellinger Distance Estimation of a Univariate GARCH Process

Roger Kadjo[1,2], Ouagnina Hili[2] & Aubin N'dri[2,3]

[1] University Félix HOUPHOUET-BOIGNY, Abidjan, Côte d'Ivoire

[2] Laboratory of Mathematics and New Technologies of Information, National Polytechnic Institute Félix HOUPHOUET BOIGNY, BP 1911, Yamoussoukro, Côte d'Ivoire

[3] University Jean Lorougnon GUEDE, Côte d'Ivoire

Correspondence: Roger Kadjo, BP 96 Alépé, Côte d'Ivoire. E-mail: roger.kadjo@yahoo.fr

Received: March 20, 2017 Accepted: May 22, 2017 Online Published: May 28, 2017

doi:10.5539/jmr.v9n3p80 URL: https://doi.org/10.5539/jmr.v9n3p80

Abstract

In this paper, we determine the Minimum Hellinger Distance estimator of a stationary GARCH process. We construct an estimator of the parameters based on the minimum Hellinger distance method. Under conditions which ensure the ϕ-mixing of the GARCH process, we establish the almost sure convergence and the asymptotic normality of the estimator.

Keywords. Hellinger distance estimation, GARCH process, ϕ-mixing process, consistence, asymptotic normality.

1. Introduction

General Autoregressive conditionally heteroscedastic (GARCH) models were pioneered by Engle(1982) and Bollerslev(1986), and have ever since been widely used to analyze financial time series. Parameters of GARCH models are usually estimated by the quasi-maximum likelihood estimator (QMLE) (Berkes, Horváth, & Kokoszka, 2003) and (Francq & Zakoïan, 2004). The QML estimator is well-known for its efficiency asymptotic properties under regular conditions, however it has very bad robustess properties.

In this paper we estimate the parameters of GARCH process using the minimum Hellinger distance (MHD) method, under uniform mixing (or ϕ-mixing) condition.

The interest for this method of parametric estimation is that the minimum Hellinger distance estimation method gives efficient and robust estimators (Beran, 1977). The minimum Hellinger distance estimators have been used in parameter estimation for independent observations (Beran, 1977), for nonlinear time series models (Hili, 1995) and recently for univariate long memory linear processes (Bitty & Hili, 2010), for nonlinear univariate and multivariate gaussian process (N'dri & Hili, 2011, 2013), for parameter estimation of one-dimensional diffusion process (Apala & Hili, 2013).

The paper is organized as follows. In section 2 we give the definition and some properties of the GARCH model. Section 3 contains the definition of the estimator and some assumptions. Sections 4 and 5 are the main results of the paper. They respectively establish the consistency and the asymptotic normality of the estimator $\widehat{\theta}_n$. In section 6 we did some numerical simulations. In section 7 we apply MHD method to a financial time series. In section 8 we open problem.

2. Definition and Some Properties of GARCH Model

Definition 2.1. *The process $(X_t)_{t\in\mathbb{Z}}$ is called a GARCH(p, q) if*

$$X_t = \varepsilon_t \sqrt{h_t}, \tag{2.1}$$

where ε_t are i.i.d random variables, with $E(\varepsilon_t) = 0$, $E\left(\varepsilon_t^2\right) = 1$ and

$$h_t = \sigma_t^2 = w + \sum_{i=1}^{p}\alpha_i X_{t-i}^2 + \sum_{i=1}^{q}\beta_i h_{t-i}.$$

The α_i and β_i are nonnegative constants and ω is a (strictly) positive constant.

$\theta = \left(w, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q\right)^T \in \Theta \subset \mathbb{R}^{p+q+1}$ is the vector of the parameters of interest and $\theta_0 = (\omega_0, \alpha_{01}, ..., \alpha_{0q}, \beta_{01}, ..., \beta_{0p})^T$ the vector of the true values, where T denotes the transpose.

Proposition 2.1 (Bollerslev, 1986, Theorem 1). *If*

$$\sum_{i=1}^{p}\alpha_i + \sum_{i=1}^{q}\beta_i < 1,$$

then, The GARCH(p, q) process $(X_t)_{t\in\mathbb{Z}}$ defined in (1.2) admits a unique strictly stationary solution.

For the following properties see Davis & Mikosch (2008).

Proposition 2.2. *If ε_t has a positive Lebesgue density on a neighborhood of 0, the strictly stationary GARCH process (X_t) defined in (2.1) is ϕ-mixing, moreover, the mixing rate ϕ_k decays to 0 geometrically ($\phi_k \le C\rho^k$ with $C > 0$ and $0 < \rho < 1$).*

3. Definition of the Estimator and Some Assumptions

Let $X_1, \ldots, X_n$ be an observed sequence of GARCH processes with the density belonging to a specified parametric family $\{f_\theta\}_{\theta\in\Theta}$ where Θ is the parameter espace, a compact set of $\mathbb{R}^{p+q+1}$. Note that in our study the form of the density is not explicit.

Let f_n be a nonparametric estimator of the density f_θ defined as

$$f_n(x) = \frac{1}{nb_n}\sum_{t=1}^{n}\mathrm{K}(\frac{x - X_t}{b_n}),\ x \in \mathbb{R}$$

where K (.) is a kernel function and (b_n) is a sequence of bandwidths.

The Minimum Hellinger Distance estimator $\widehat{\theta_n}$ of θ_0 is the value in the parameter espace Θ which minimizes the Hellinger distance (denoted H_2) between f_n and f_θ defined by:

$$\widehat{\theta_n} = \arg\min_{\theta\in\Theta} H_2(f_n; f_\theta),\ \text{where} \tag{3.1}$$

$$H_2(f_n; f_\theta) = \left\{\int_{\mathbb{R}}\left|f_n^{\frac{1}{2}}(x) - f_\theta^{\frac{1}{2}}(x)\right|^2 dx\right\}^{\frac{1}{2}}.$$

To establish the asymptotic properties of the estimator $\widehat{\theta_n}$, we need the following assumptions.

Assumption A1

The GARCH(p,q) process is geometrically ϕ-mixing.

Assumption A2

For each $\theta \in \Theta$, the density f_θ of X_t is positive over all $\mathbb{R}$ and twice continuously differentiable.

Assumption A3

For each $\theta \in \Theta$, $\left\|f_\theta^{(i)}\right\|_\infty = \sup_x \left|f_\theta^{(i)}(x)\right| < \infty \ \ i = 0, 1, 2.$

Assumption A4

We chose b_n such that $\lim_{n\longrightarrow\infty} b_n = 0$, $\lim_{n\longrightarrow\infty} nb_n = +\infty$, $\lim_{n\longrightarrow\infty} \sqrt{n}b_n^2 = +\infty$ and $\lim_{n\longrightarrow\infty} n^{\frac{1}{4}}b_n^2 = 0$.

Assumption A5

The continuous function K is symmetric positive, bounded function with compact support such that:

$$\int_{\mathbb{R}}\mathrm{K}(u)du = 1,\ \int_{\mathbb{R}} u\,\mathrm{K}(u)du = 0 \text{ and } \int_{\mathbb{R}}|u|^2\,\mathrm{K}(u)du < \infty.$$

Assumption A6

For $\theta_1, \theta_2 \in \Theta$, $\theta_1 \ne \theta_2$ implies that $\{x \in \mathbb{R}/\ f_{\theta_1}(x) \ne f_{\theta_2}(x)\}$ is a set of positive Lebesgue measure.

4. Consistency of the Estimator $\widehat{\theta_n}$

Theorem 4.1 (Almost sure convergence). *Suppose that assumptions (A1)-(A6) are satisfied. If θ_0 is in the interior of Θ, then, $\widehat{\theta_n} \longrightarrow \theta_0$ a.s when $n \longrightarrow +\infty$.*

Proof of theorem 4.1.

Let F denote the set of all densities with respect to the Lebesgue measure on $\mathbb{R}$.

Define the functional $U\colon F \longrightarrow \Theta$ as

$$U(g) = \arg\min_{\theta\in\Theta} H_2(g, f_\theta),$$

provided such minimum exists. In cas $U(g)$ is multiple-value, the notation $U(g)$ will represent one of possible values chosen arbitrarily.

We have

$$\left|f_n(x) - f_{\theta_0}(x)\right| \leq |f_n(x) - Ef_n(x)| + \left|Ef_n(x) - f_{\theta_0}(x)\right|.$$

By lemmas 4.2 and 4.3

$$\left|f_n(x) - f_{\theta_0}(x)\right| \longrightarrow 0 \text{ almost surely when } n \longrightarrow \infty,$$

consequently,

$$\Pr\text{ob}\left\{\lim_{n\longrightarrow\infty} f_n^{\frac{1}{2}}(x) = f_{\theta_0}^{\frac{1}{2}}(x) \text{ for all } x\right\} = 1.$$

Since

$$\int_{\mathbb{R}} f_n(x)dx = \int_{\mathbb{R}} f_{\theta_0}(x)dx = 1,$$

then,

$$H_2(f_n, f_{\theta_0}) = \left\{\int_{\mathbb{R}} \left|f_n^{\frac{1}{2}}(x) - f_{\theta_0}^{\frac{1}{2}}(x)\right|^2 dx\right\}^{\frac{1}{2}} \longrightarrow 0 \; a.s \text{ when } n \longrightarrow \infty.$$

Thus $f_n(x) \longrightarrow f_{\theta_0}(x)$ $a.s$ when $n \longrightarrow \infty$ in the Hellinger topology.

From the continuity of the functional U (Beran, 1977, Theorem 1), we obtain

$$\widehat{\theta_n} = U(f_n(x)) \longrightarrow U(f_{\theta_0}(x)) = \theta_0 \; a.s \text{ when } n \longrightarrow \infty.$$

Lemma 4.2. *Suppose that assumptions (A1)-(A5) are satisfied. Then,*

$$|f_n(x) - Ef_n(x)| \longrightarrow 0 \; a.s \text{ when } n \longrightarrow \infty.$$

Proof of lemma 4.2.

We have

$$|f_n(x) - Ef_n(x)| = \frac{1}{nb_n}\left|\sum_{t=1}^{n} \Delta_t\right|$$

where

$$\Delta_t = \mathrm{K}(\frac{x - X_t}{b_n}) - E\,\mathrm{K}(\frac{x - X_t}{b_n}).$$

Using assumption (A5) and Jensen's inequality, we get

$$\begin{aligned}
|\Delta_1| &\leq \left|\mathrm{K}(\frac{x - X_1}{b_n})\right| + \left|E\left(\mathrm{K}(\frac{x - X_1}{b_n})\right)\right| \\
&\leq \left|\mathrm{K}(\frac{x - X_1}{b_n})\right| + E\left(\left|\mathrm{K}(\frac{x - X_1}{b_n})\right|\right) \\
&\leq \sup_x \left|\mathrm{K}(\frac{x - X_1}{b_n})\right| + E\left(\sup_x \left|\mathrm{K}(\frac{x - X_1}{b_n})\right|\right) \\
&\leq 2K_0
\end{aligned}$$

where K_0 is a constant.

We have also

$$
\begin{aligned}
E\left|\Delta_1\right|^2 &= E\left|\mathrm{K}(\frac{x-X_1}{b_n}) - E\,\mathrm{K}\left(\frac{x-X_1}{b_n}\right)\right|^2 \\
&= E\,\mathrm{K}^2(\frac{x-X_1}{b_n}) - \left(E\,\mathrm{K}\left(\frac{x-X_1}{b_n}\right)\right)^2 \\
&\leq E\,\mathrm{K}^2(\frac{x-X_1}{b_n}) \\
&= \int_{\mathbb{R}} \mathrm{K}^2(\frac{x-s}{b_n}) f_{\theta_0}(s)ds \\
&\leq \int_{\mathbb{R}} \mathrm{K}^2(\frac{x-s}{b_n}) \sup_{s\in\mathbb{R}} f_{\theta_0}(s)ds \\
&= \sup_{s\in\mathbb{R}} f_{\theta_0}(s) \int_{\mathbb{R}} \mathrm{K}^2(\frac{x-s}{b_n})ds \\
&= \sup_{s\in\mathbb{R}} f_{\theta_0}(s) b_n \int_{\mathbb{R}} \mathrm{K}^2(u)du \\
&= Cb_n \leq K_1,
\end{aligned}
$$

where $C = \sup_{s\in\mathbb{R}} f_{\theta_0}(s) \int_{\mathbb{R}} \mathrm{K}^2(u)du$ and K_1 is a constant.

Then, by the relation (20) in Hang et al (2015) and for all $\epsilon > 0$, we obtain

$$
\begin{aligned}
P\left(\frac{n^{\frac{1}{4}}}{nb_n}\left|\sum_{t=1}^{n}\Delta_t\right| > \epsilon\right) &= P\left(\frac{1}{n}\left|\sum_{t=1}^{n}\Delta_t\right| > n^{-\frac{1}{4}}\epsilon b_n\right) \\
&\leq 2\exp\left\{\left(-\frac{\left(n^{-\frac{1}{4}}\epsilon b_n\right)^2 n}{8C_\phi\left(4K_1 + 2K_0 n^{-\frac{1}{4}}\epsilon b_n\right)}\right)\right\} \\
&= 2\exp\left\{\left(-\frac{\epsilon^2\sqrt{n}b_n^2}{8C_\phi\left(4K_1 + 2K_0\epsilon n^{-\frac{1}{4}} b_n\right)}\right)\right\},
\end{aligned}
$$

where $C_\phi = \sum_{k=1}^{\infty}\phi_k$.

We have

$$C_\phi = \sum_{k=1}^{\infty}\phi_k < \infty \text{ and } \lim_{n\longrightarrow\infty} n^{-\frac{1}{4}} b_n = 0.$$

Then, using assumption (A4) and Borel Cantelli's lemma, we get

$$n^{\frac{1}{4}}\left|f_n(x) - Ef_n(x)\right| \longrightarrow 0 \; a.s \text{ when } n \longrightarrow \infty\,. \tag{4.1}$$

Hence

$$\left|f_n(x) - Ef_n(x)\right| = o\left(n^{-\frac{1}{4}}\right) \; a.s \text{ when } n \longrightarrow \infty\,.$$

Lemma 4.3. *Suppose that assumptions (A1)-(A5) are satisfied. Then,*

$$\sup_{x\in\mathbb{R}}\left|Ef_n(x) - f_{\theta_0}(x)\right| \longrightarrow 0 \textit{ when } n \longrightarrow \infty.$$

Proof of lemma 4.3.

Using assumption (A5) and the Taylor's expansion in a neighbourhood of x, we have

$$\begin{aligned} & Ef_n(x) - f_{\theta_0}(x) \\ & = E\left(\frac{1}{nb_n}\sum_{t=1}^{n} \mathrm{K}(\frac{x-X_t}{b_n})\right) - f_{\theta_0}(x) \end{aligned}$$

$$\begin{aligned} &= \frac{1}{b_n}E\left(\mathrm{K}(\frac{x-X_1}{b_n})\right) - f_{\theta_0}(x) \\ &= \frac{1}{b_n}\int_{\mathbb{R}} \mathrm{K}(\frac{x-s}{b_n}) f_{\theta_0}(s)ds - f_{\theta_0}(x) \\ &= \int_{\mathbb{R}} \mathrm{K}(u) f_{\theta_0}(x-b_n u)du - \int_{\mathbb{R}} \mathrm{K}(u)du f_{\theta_0}(x) + b_n f'_{\theta_0}(x)\int_{\mathbb{R}} u\,\mathrm{K}(u)du \\ &= \int_{\mathbb{R}} \mathrm{K}(u)\left[f_{\theta_0}(x) - b_n u f'_{\theta_0}(x) + \frac{1}{2}b_n^2u^2 f''_{\theta_0}(\delta) - f_{\theta_0}(x) + b_n u f'_{\theta_0}(x)\right]du \\ &= \int_{\mathbb{R}} \mathrm{K}(u)\frac{1}{2}b_n^2u^2 f''_{\theta_0}(\delta)du. \end{aligned}$$

where $\delta = x + \varphi(-b_n u)$ with $0 < \varphi < 1$.

Then, by assumption (A3) and (A5)

$$\sup_{x\in\mathbb{R}}\left|Ef_n(x) - f_{\theta_0}(x)\right| \le \frac{1}{2}b_n^2 \sup_{x\in\mathbb{R}}\left|f''_{\theta_0}(\delta)\right| \int_{\mathbb{R}} |u|^2\,\mathrm{K}(u)du \le C_1 b_n^2. \tag{4.2}$$

where C_1 is a constant.

Using assumption (A4), we conclued that $\sup_{x\in\mathbb{R}}\left|Ef_n(x) - f_{\theta_0}(x)\right| \longrightarrow 0$ when $n \longrightarrow \infty$.

5. Asymptotic Normality of the Estimator $\widehat{\theta}_n$

For the following theorem, denote by

$$S_\theta = f_\theta^{\frac{1}{2}}, \quad \dot{S}_\theta = \frac{\partial S_\theta}{\partial\theta}, \qquad \ddot{S}_\theta = \frac{\partial^2 S_\theta}{\partial\theta\partial\theta^T}$$

when these quantities exist. Furthermore, let

$$V_\theta(x) = \left[\int_{\mathbb{R}} \dot{S}_\theta(x)\dot{S}_\theta^T(x)dx\right]^{-1}\dot{S}_\theta(x) \text{ and } h_\theta(x) = \frac{\dot{S}_\theta(x)}{2f_\theta^{\frac{1}{2}}(x)}. \tag{5.1}$$

Theorem 5.1 (asymptotic normality of the estimator)**.** *Suppose that assumptions (A1)-(A6) are satisfied. Furthermore, assume that*

(i) if the components of $\dot{S}_\theta$ and $\ddot{S}_\theta$ are in L_2 and the norms of these components are continuous functions at θ and

(ii) if θ_0 lies in the interior of Θ and if $\int_{\mathbb{R}} \ddot{S}_{\theta_0}(x)S_{\theta_0}(x)dx$ is a non singular $(p+q+1)\times(p+q+1)-$matrix, then, the limiting distribution of $\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right)$ is $\mathcal{N}\left(0,\Sigma^2\right)$ where

$$\Sigma^2 = \frac{1}{4}\left[\int_{\mathbb{R}} \dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}^T(x)dx\right]^{-1}.$$

Proof of theorem 5.1.

From theorem 2 of Beran (1977), we can write :

$$\sqrt{n}\left(\widehat{\theta_n}-\theta_0\right)= \sqrt{n}\int_{\mathbb{R}} V_{\theta_0}(x)\left[f_n^{\frac{1}{2}}(x)-f_{\theta_0}^{\frac{1}{2}}(x)\right]dx + \sqrt{n}A_n\int_{\mathbb{R}} \dot{S}_{\theta_0}(x)\left[f_n^{\frac{1}{2}}(x)-f_{\theta_0}^{\frac{1}{2}}(x)\right]dx$$

where the components of $A_n \longrightarrow 0$ when $n \longrightarrow \infty$ and $V_{\theta_0}(x)$ defined in (5.1).

We have

$$f_n^{\frac{1}{2}}(x)-f_{\theta_0}^{\frac{1}{2}}(x) = \frac{(f_n(x)-f_{\theta_0}(x))}{2f_{\theta_0}^{\frac{1}{2}}(x)} - \frac{\left(f_n^{\frac{1}{2}}(x)-f_{\theta_0}^{\frac{1}{2}}(x)\right)^2}{2f_{\theta_0}^{\frac{1}{2}}(x)} = \frac{(f_n(x)-f_{\theta_0}(x))}{2f_{\theta_0}^{\frac{1}{2}}(x)} - \frac{(f_n(x)-f_{\theta_0}(x))^2}{2f_{\theta_0}^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x)+f_{\theta_0}^{\frac{1}{2}}(x)\right)^2}.$$

Thus,

$$\begin{aligned}&\sqrt{n}\int_{\mathbb{R}} V_{\theta_0}(x)\left[f_n^{\frac{1}{2}}(x)-f_{\theta_0}^{\frac{1}{2}}(x)\right]dx\\ &= \sqrt{n}\int_{\mathbb{R}} V_{\theta_0}(x)\frac{(f_n(x)-f_{\theta_0}(x))}{2f_{\theta_0}^{\frac{1}{2}}(x)}dx - \sqrt{n}\int_{\mathbb{R}} V_{\theta_0}(x)\frac{(f_n(x)-f_{\theta_0}(x))^2}{2f_{\theta_0}^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x)+f_{\theta_0}^{\frac{1}{2}}(x)\right)^2}dx\\ &= \sqrt{n}\int_{\mathbb{R}} V_{\theta_0}(x)\frac{(f_n(x)-f_{\theta_0}(x))}{2f_{\theta_0}^{\frac{1}{2}}(x)}dx + B_n\end{aligned}$$

where

$$B_n = -\sqrt{n}\int_{\mathbb{R}} V_{\theta_0}(x)\frac{(f_n(x)-f_{\theta_0}(x))^2}{2f_{\theta_0}^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x)+f_{\theta_0}^{\frac{1}{2}}(x)\right)^2}dx.$$

Since

$$2f_{\theta_0}^{\frac{1}{2}}(x)\left(f_n^{\frac{1}{2}}(x)+f_{\theta_0}^{\frac{1}{2}}(x)\right)^2 > 2f_{\theta_0}^{\frac{3}{2}}(x),$$

thus,

$$|B_n| \le \int_{\mathbb{R}} \frac{\left|V_{\theta_0}(x)\right| \sqrt{n}\,(f_n(x)-f_{\theta_0}(x))^2}{2f_{\theta_0}^{\frac{3}{2}}(x)}dx.$$

Using (4.1), (4.2) and assumption (A4), we get

$$\begin{aligned} n^{\frac{1}{4}}\left|f_n(x)-f_{\theta_0}(x)\right| &\le n^{\frac{1}{4}}\left|f_n(x)-Ef_n(x)\right| + n^{\frac{1}{4}}\sup_{x\in\mathbb{R}}\left|Ef_n(x)-f_{\theta_0}(x)\right|\\ &\le n^{\frac{1}{4}}\left|f_n(x)-Ef_n(x)\right| + C_1 n^{\frac{1}{4}} b_n^2 \longrightarrow 0 \ a.s \text{ when } n\longrightarrow\infty.\end{aligned}$$

In conclusion,

$$\sqrt{n}\,(f_n(x)-f_{\theta_0}(x))^2 \longrightarrow 0 \ a.s \text{ when } n\longrightarrow\infty.$$

Conditions (i) and (ii) of theorem 5.1 imply that V_{θ_0} is continuous and bounded (for θ_0 fixed), futhermore, applying Vitali's theorem on the sequence $\left|V_{\theta_0}(x)\right| \sqrt{n}\,(f_n(x)-f_{\theta_0}(x))^2$, we obtain $|B_n| \longrightarrow 0$ in probability when $n \longrightarrow \infty$.

On the other hand,

$$\begin{aligned}
&\sqrt{n}\int_{\mathbb{R}} V_{\theta_0}(x)\frac{(f_n(x)-f_{\theta_0}(x))}{2f_{\theta_0}^{\frac{1}{2}}(x)}dx\\
&=\left(\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\right)^{-1}\int_{\mathbb{R}}\sqrt{n}\dot{S}_{\theta_0}(x)\frac{(f_n(x)-f_{\theta_0}(x))}{2f_{\theta_0}^{\frac{1}{2}}(x)}dx\\
&=\left(\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\right)^{-1}\left\{\int_{\mathbb{R}}\sqrt{n}\frac{\dot{S}_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x)}f_n(x)dx-\frac{1}{2}\sqrt{n}\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)f_{\theta_0}^{\frac{1}{2}}(x)dx\right\}\\
&=\left(\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\right)^{-1}\int_{\mathbb{R}}\sqrt{n}\frac{\dot{S}_{\theta_0}(x)}{2f_{\theta_0}^{\frac{1}{2}}(x)}f_n(x)dx-0\text{ (see (5.4))}\\
&=\left(\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\right)^{-1}\int_{\mathbb{R}}\sqrt{n}h_{\theta_0}(x)\,f_n(x)dx.
\end{aligned}$$

Since $A_n \longrightarrow 0$ when $n \longrightarrow \infty$, then, the limit distribution of $\sqrt{n}\left(\widehat{\theta_n}-\theta_0\right)$ is reduced to the limit distribution of

$$\left(\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\right)^{-1}\int_{\mathbb{R}}\sqrt{n}h_{\theta_0}(x)\,f_n(x)dx.$$

Therefore, by the lemmas 5.3 and 5.4

$$\int_{\mathbb{R}}\sqrt{n}h_{\theta_0}(x)\,f_n(x)dx \xrightarrow{Đ} \mathcal{N}\left(0\,,\,\Gamma^2\right)\text{ when } n\longrightarrow\infty,$$

where

$$\Gamma^2=\frac{1}{4}\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\,.$$

We conclued that,

$$\left(\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\right)^{-1}\int_{\mathbb{R}}\sqrt{n}h_{\theta_0}(x)\,f_n(x)dx \xrightarrow{Đ} \mathcal{N}\left(0\,,\,\Sigma^2\right)$$

where

$$\begin{aligned}
\Sigma^2&=\left(\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\right)^{-2}\frac{1}{4}\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\\
&=\frac{1}{4}\left(\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx\right)^{-1}.
\end{aligned}$$

Lemma 5.2. *Let $\left\{Y_j : j\in\mathbb{N}\right\}$ be a stationary sequence which is ϕ-mixing with $\sum_{j=1}^{\infty}\phi_j^{\frac{1}{2}}<\infty$. Assume $E\left|Y_1\right|^2<C<\infty$ and $E\left(Y_1\right)=0$. Then,*

$$\lim_{n\longrightarrow\infty}\frac{1}{n}E\left(\sum_{j=1}^{n}Y_j\right)^2=E\left(Y_1^2\right)+2\sum_{j=1}^{\infty}E\left(Y_1Y_{j+1}\right).$$

Proof of lemma 5.2.

Using lemma 20.1 in Billingsley (1968), we have

$$\begin{aligned}
\sum_{j=1}^{\infty}\left|E\left(Y_1Y_{j+1}\right)\right| &\le \sum_{j=1}^{\infty}2\phi^{\frac{1}{2}}(j)\left(E\left|Y_1\right|^2\right)^{\frac{1}{2}}\left(E\left|\left(Y_{j+1}\right)\right|^2\right)^{\frac{1}{2}}\\
&\le 2C\sum_{j=1}^{\infty}\phi^{\frac{1}{2}}(j)<\infty.
\end{aligned}$$

Hence, the lemma follows from Ibragimov and Linnik (1971, theorem 18.5.2).

Lemma 5.3. *Let $h_\theta(.)$ the continuous function defined in (5.1). Suppose that assumptions (A1)-(A5) hold. Then,*

$$\sqrt{n}\left\{\int_{\mathbb{R}} h_{\theta_0}(x)\, f_n(x)dx - \frac{1}{n}\sum_{i=1}^{n} h_{\theta_0}(X_i)\right\} \longrightarrow 0 \text{ in probability.}$$

Proof of lemma 5.3.

We have

$$\begin{aligned}
&\sqrt{n}\left\{\int_{\mathbb{R}} h_{\theta_0}(x)\, f_n(x)dx - \frac{1}{n}\sum_{i=1}^{n} h_{\theta_0}(X_i)\right\}\\
&= \sqrt{n}\left\{\int_{\mathbb{R}} \frac{1}{nb_n}\sum_{i=1}^{n} \mathrm{K}(\frac{x-X_i}{b_n}) h_{\theta_0}(x)\, dx - \frac{1}{n}\sum_{i=1}^{n} h_{\theta_0}(X_i)\right\}\\
&= \sqrt{n}\left\{\frac{1}{n}\sum_{i=1}^{n}\left(\int_{\mathbb{R}} \frac{1}{b_n}\mathrm{K}(\frac{x-X_i}{b_n}) h_{\theta_0}(x)\, dx - h_{\theta_0}(X_i)\right)\right\}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int_{\mathbb{R}} \frac{1}{b_n}\mathrm{K}(\frac{x-X_i}{b_n}) h_{\theta_0}(x)\, dx - h_{\theta_0}(X_i)\right)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int_{\mathbb{R}} \mathrm{K}(u) h_{\theta_0}(X_i+ub_n)\, du - h_{\theta_0}(X_i)\int_{\mathbb{R}} \mathrm{K}(u)du\right)\\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int_{\mathbb{R}} \left(h_{\theta_0}(X_i+ub_n) - h_{\theta_0}(X_i)\right)\mathrm{K}(u)du\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
&E\left(\sqrt{n}\left\{\int_{\mathbb{R}} h_{\theta_0}(x)\, f_n(x)dx - \frac{1}{n}\sum_{i=1}^{n} h_{\theta_0}(X_i)\right\}\right)^2\\
&= E\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(\int_{\mathbb{R}} \left(h_{\theta_0}(X_i+ub_n) - h_{\theta_0}(X_i)\right)\mathrm{K}(u)du\right)\right)^2\\
&= \frac{1}{n}E\left(\sum_{i=1}^{n}\left(\int_{\mathbb{R}} \left(h_{\theta_0}(X_i+ub_n) - h_{\theta_0}(X_i)\right)\mathrm{K}(u)du\right)\right)^2\\
&= \frac{1}{n}\sum_{i=1}^{n}E\left(\int_{\mathbb{R}} \left(h_{\theta_0}(X_i+ub_n) - h_{\theta_0}(X_i)\right)\mathrm{K}(u)du\right)^2\\
&+ \frac{2}{n}\sum_{i<j}E\left(\left(\int_{\mathbb{R}} \left(h_{\theta_0}(X_i+ub_n) - h_{\theta_0}(X_i)\right)\mathrm{K}(u)du\right)\right.\\
&\left.\times\left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(X_j+ub_n\right) - h_{\theta_0}\left(X_j\right)\right)\mathrm{K}(u)du\right)\right)
\end{aligned}$$

Step 1: we prove that

$$\frac{1}{n}\sum_{i=1}^{n}E\left(\int_{\mathbb{R}} \left(h_{\theta_0}(X_i+ub_n) - h_{\theta_0}(X_i)\right)\mathrm{K}(u)du\right)^2 \longrightarrow 0 \text{ when } n \longrightarrow \infty.$$

Using Cauchy-Schwarz's inequality and Fubini's theorem, we obtain

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} E\left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(X_i + ub_n\right) - h_{\theta_0}\left(X_i\right)\right) \mathrm{K}(u)du\right)^2 \\
&= \frac{1}{n}\left\{n\int_{\mathbb{R}}\left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right) \mathrm{K}(u)du\right)^2 f_{\theta_0}(x)dx\right\} \\
&= \int_{\mathbb{R}}\left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right) \mathrm{K}(u)du\right)^2 f_{\theta_0}(x)dx \\
&\leq \int_{\mathbb{R}}\left[\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 du \times \int_{\mathbb{R}} \mathrm{K}^2(u)du\right] f_{\theta_0}(x)dx \\
&= C_2\int_{\mathbb{R}}\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x)dudx \\
&= C_2\int_{\mathbb{R}}\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x)dxdu
\end{aligned}$$

where C_2 is a constant.

By the continuity of h_{θ_0}, we get

$$\left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 \longrightarrow 0 \text{ when } n \longrightarrow \infty.$$

Conditions (i) of theorem 5.1 imply that $\dot{S}_{\theta_0}$is bounded (for θ_0 fixed). Then, h_{θ_0} is also bounded (for θ_0 fixed). Thus, there existe a constant $C_3 > 0$ such that for all $n \in \mathbb{N}$

$$\left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 \leq C_3.$$

Since by assumption (A3) $0 < f_{\theta_0} < \infty$, we have

$$\left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x) \longrightarrow 0 \text{ when } n \longrightarrow \infty$$

and for all $n \in \mathbb{N}$

$$\left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x) \leq C_3 f_{\theta_0}(x).$$

Then, by the dominated convergence theorem,

$$\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x)dx \longrightarrow 0 \text{ when } n \longrightarrow \infty$$

where $u \in$supp(K) the support of kernel density K (.) a compact set.

On other hand, for all $u \in$supp(K) and for all $n \in \mathbb{N}$ we obtain

$$\begin{aligned}
&\left|\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x)dx\right| \\
&\leq C_3\int_{\mathbb{R}} f_{\theta_0}(x)dx \\
&= C_3.
\end{aligned}$$

Therefore, by the dominated convergence theorem

$$\int_{\mathbb{R}}\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x)dxdu \longrightarrow 0 \text{ when } n \longrightarrow \infty.$$

Step 2: we prove that

$$\begin{aligned}
&\frac{2}{n}\sum_{i<j} E\left(\left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(X_i + ub_n\right) - h_{\theta_0}\left(X_i\right)\right) \mathrm{K}(u)du\right)\right. \\
&\left.\times\left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(X_j + ub_n\right) - h_{\theta_0}\left(X_j\right)\right) \mathrm{K}(u)du\right)\right) \longrightarrow 0 \text{ when } n \longrightarrow \infty.
\end{aligned}$$

Let ψ be the function defined by:

for all $u \in$supp(K) a compact set and for all $x \in \mathbb{R}$

$$\psi(x) = \int_{\mathbb{R}} \left(h_{\theta_0}(x + ub_n) - h_{\theta_0}(x)\right) \mathrm{K}(u)du.$$

ψ is a continuous function. Therefore $\psi(X_i)$ is ϕ-mixing.

We have

$$\begin{aligned}
|\psi(x)| &= \left| \int_{\mathbb{R}} \left(h_{\theta_0}(x + ub_n) - h_{\theta_0}(x)\right) \mathrm{K}(u)du \right| \\
&= \int_{\mathbb{R}} \left| \left(h_{\theta_0}(x + ub_n) - h_{\theta_0}(x)\right) \right| \mathrm{K}(u)du \\
&\leq C_4 \int_{\mathbb{R}} \mathrm{K}(u)du \\
&= C_4.
\end{aligned}$$

Thus, there exists a constant C_5 such that

$$E\left|\psi(X_i)\right|^2 < C_5. \tag{5.2}$$

On the other hand, we know that

$$\phi(k) \leq C\rho^k$$

where $C > 0$ and $0 < \rho < 1$.

Therefore,

$$\phi^{\frac{1}{2}}(k) \leq C \exp\left(\frac{1}{2} k Log\rho\right) = C \exp\left(-\frac{1}{2}\nu k\right)$$

with $\nu = -Log\rho$.

Let χ be the decreasing function defined by: for all $k \in \mathbb{N}$

$$\chi(k) = C \exp\left(-\frac{1}{2}\nu k\right). \tag{5.3}$$

We have

$$\begin{aligned}
&\frac{2}{n}\sum_{i<j} E\left(\left(\int_{\mathbb{R}} \left(h_{\theta_0}(X_i + ub_n) - h_{\theta_0}(X_i)\right) \mathrm{K}(u)du\right)\right. \\
&\left.\times\left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(X_j + ub_n\right) - h_{\theta_0}\left(X_j\right)\right) \mathrm{K}(u)du\right)\right) \\
&= \frac{2}{n}\sum_{i<j} E\left(\psi(X_i)\,\psi\left(X_j\right)\right).
\end{aligned}$$

Using lemma 20.1 in Billingsley (1968), (5.2) and the fact that χ is a decreasing function, we obtain

$$
\begin{aligned}
\frac{2}{n}\sum_{1\le i<j\le n}\left|E\left(\psi(X_i)\psi\left(X_j\right)\right)\right| &= \frac{2}{n}\sum_{j=1}^{n-1} j\left|E\left(\psi(X_1)\psi\left(X_{j+1}\right)\right)\right| \\
&\le \frac{2}{n}\sum_{j=1}^{n-1} 2j\phi^{\frac{1}{2}}(j)\left(E|\psi(X_1)|^2\right)^{\frac{1}{2}}\left(E\left|\psi\left(X_{j+1}\right)\right|^2\right)^{\frac{1}{2}} \\
&\le \frac{4C_5}{n}\sum_{j=1}^{n-1} j\phi^{\frac{1}{2}}(j) \\
&= \frac{4C_5}{n}\sum_{k=1}^{n-1}\sum_{l=k}^{n-1}\phi^{\frac{1}{2}}(l) \\
&\le \frac{4C_5}{n}\sum_{k=1}^{n-1}\sum_{l=k}^{n-1}\chi(l) \\
&= \frac{4C_5}{n}\sum_{k=1}^{n-1}\sum_{l=k}^{n-1}(\chi(l))^{\frac{1}{2}}(\chi(l))^{\frac{1}{2}} \\
&\le \frac{4C_5}{n}\sum_{k=1}^{n-1}\sum_{l=k}^{n-1}(\chi(l))^{\frac{1}{2}}(\chi(k))^{\frac{1}{2}} \\
&\le \frac{4C_5}{n}\sum_{k=1}^{\infty}(\chi(k))^{\frac{1}{2}}\sum_{l=1}^{\infty}(\chi(l))^{\frac{1}{2}} \\
&\le \frac{4C_5}{n}\left(\sum_{i=1}^{\infty}(\chi(i))^{\frac{1}{2}}\right)^2 \longrightarrow 0.
\end{aligned}
$$

By **Step 1** and **Step 2** we obtain

$$
E\left(\sqrt{n}\left\{\int_{\mathbb{R}} h_{\theta_0}(x)\, f_n(x)dx - \frac{1}{n}\sum_{i=1}^{n} h_{\theta_0}(X_i)\right\}\right)^2 \longrightarrow 0 \text{ when } n \longrightarrow \infty.
$$

We conclued that,

$$
\sqrt{n}\left\{\int_{\mathbb{R}} h_{\theta_0}(x)\, f_n(x)dx - \frac{1}{n}\sum_{i=1}^{n} h_{\theta_0}(X_i)\right\} \longrightarrow 0 \text{ in probability.}
$$

Lemma 5.4. *Suppose that assumptions (A1)-(A5) are hold. Then,*

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^{n} h_{\theta_0}(X_i) \xrightarrow{\text{Đ}} \mathcal{N}\left(0\,,\,\Gamma^2\right),\ \textit{where } \Gamma^2 = \frac{1}{4}\int_{\mathbb{R}} \dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx \textit{ and}
$$

"Đ" denote the convergence in distribution.

Proof of lemma 5.4.

We have

$$
\begin{aligned}
E\left(h_{\theta_0}(X_i)\right) = \int_{\mathbb{R}} h_{\theta_0}(x)\, f_{\theta_0}(x)dx &= \frac{1}{2}\int_{\mathbb{R}} \dot{S}_{\theta_0}(x) f_{\theta_0}^{\frac{1}{2}}(x)dx \\
&= \frac{1}{4}\int_{\mathbb{R}} 2\dot{S}_{\theta_0}(x) S_{\theta_0}(x)dx = 0
\end{aligned}
\tag{5.4}
$$

and

$$
\begin{aligned}
E\left(h^2_{\theta_0}(X_i)\right) = \int_{\mathbb{R}} h^2_{\theta_0}(x)\, f_{\theta_0}(x)dx &= \frac{1}{4}\int_{\mathbb{R}} \dot{S}^2_{\theta_0}(x)dx \\
&= \frac{1}{4}\int_{\mathbb{R}} \dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx < \infty.
\end{aligned}
$$

We have

$$E\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}h_{\theta_0}(X_k)\right)^2 = E\left(h_{\theta_0}(X_1)\right)^2 + \frac{2}{n}\sum_{i<j}E\left(h_{\theta_0}(X_i)\,h_{\theta_0}\left(X_j\right)\right)$$

and

$$\begin{aligned}\frac{2}{n}\sum_{1\le i<j\le n}\left|E\left(h_{\theta_0}(X_i)\,h_{\theta_0}\left(X_j\right)\right)\right| &= \frac{2}{n}\sum_{j=1}^{n-1}j\left|E\left(h_{\theta_0}(X_1)\,h_{\theta_0}\left(X_{j+1}\right)\right)\right| \\ &\le \frac{2}{n}\sum_{j=1}^{n-1}2j\phi^{\frac{1}{2}}(j)\left(E\left|h_{\theta_0}(X_1)\right|^2\right)^{\frac{1}{2}}\left(E\left|h_{\theta_0}\left(X_{j+1}\right)\right|^2\right)^{\frac{1}{2}} \\ &\le \frac{4C_0}{n}\sum_{j=1}^{n-1}j\phi^{\frac{1}{2}}(j)\,.\end{aligned}$$

Using the same arguments as in the lemma 5.3, we can prove that

$$\frac{2}{n}\sum_{i<j}\left|E\left(h_{\theta_0}(X_i)\,h_{\theta_0}\left(X_j\right)\right)\right| \le \frac{4C_0}{n}\left(\sum_{i=1}^{\infty}(\chi(i))^{\frac{1}{2}}\right)^2 \longrightarrow 0,$$

where χ is defined in (5.3).

We conclued that,

$$E\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}h_{\theta_0}(X_k)\right)^2 \longrightarrow \frac{1}{4}\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx \text{ when } n\longrightarrow\infty.$$

Using the convergence limit theorem in Ibragimov (1975)

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}h_{\theta_0}(X_k) \xrightarrow{\text{Đ}} \mathcal{N}\left(0\,,\,\Gamma^2\right),$$

where

$$\Gamma^2 = \frac{1}{4}\int_{\mathbb{R}}\dot{S}_{\theta_0}(x)\dot{S}_{\theta_0}(x)^T dx \text{ when } n\longrightarrow\infty.$$

6. Simulations

In this section we give some numerical simulations for minimum Hellinger distance estimator (MDHE) to show its performance. Note that the model density f_θ is intractable. However, the GARCH process generating the X_t can be simulated. Thus, on the basis of X_t simulated, we can obtain the nonparametric density estimate denoted $\widetilde{f}_{n,\theta}$, which becomes an alternative to intractable f_θ in expression (3.1) (see Gouriéroux & Monfort, 1996). Using the method of Takada (2007), we define $\widetilde{f}_{n,\theta}$ as follows:

let $\left(\widetilde{X}_1^s(\theta),...,\widetilde{X}_n^s(\theta)\right)$ be the s-th replication of the simulated sequence from the model GARCH, $s = 1,2,...,S$. That simulated sequence has the length $S\times n$.

$$\widetilde{f}_{n,\theta}(x) = \frac{1}{S}\sum_{s=1}^{S}\left[\frac{1}{nb_n}\sum_{t=1}^{n}\mathrm{K}\left(\frac{x-\widetilde{X}_t^s(\theta)}{b_n}\right)\right],\ x\in\mathbb{R}.$$

We generate a $\phi-$mixing GARCH(1,1) stationary process with true parameter $\theta_0 = (w = 0.09, \alpha = 0.15, \beta = 0.4)$. We choose a sample length: $n = 500$. The process ε_t is gaussian with mean 0 and standard deviation 1. To calculate the MDH estimator, we choose the bandwidth $b_n = n^{-0.23}$. The nonparametric estimator $\widetilde{f}_{n,\theta}$ is calculated by choosing $S = 100$; b_n is the same for $\widetilde{f}_{n,\theta}$ and f_n. The kernel is gaussian with mean 0 and standard deviation 1. For the simulations, we use "fgarch and fbasics" in R packages.

To confirm the performance of the estimator, we use the sample bias and the root means quare error (RMSE) defined as follows:

$$\mathrm{BIAS}(\theta) = \frac{1}{S}\sum_{s=1}^{S}\left(\theta_0 - \widehat{\theta}_s\right)$$

and

$$\mathrm{RMSE}(\theta) = \sqrt{\frac{1}{S}\sum_{s=1}^{S}\left(\theta_0 - \widehat{\theta}_s\right)^2}.$$

Table 1 shows the consistency of the MHD estimator.

Table 1. Consistency of the MHD estimator and the comparison between MHDE and QMLE

	$\widehat{w}$	$\widehat{\alpha}$	$\widehat{\beta}$
MHDE	0.09107473	0.15162636	0.40128264
BIAS	-0.00107473	-0.00162636	-0.00128264
RMSE	0.385817373	0.05417584	0.03038960
QMLE	0.08569830	0.17748747	0.30008846
BIAS	0.0043015	-0.00274874	0.09991154
RMSE	0.38228581	0.73444438	0.11449742

The above results show the good performance of the MHD method because all estimations biases are close to 0. Also, the table 1 shows that the estimations biases and the RMSE of the MHD method are small or almost equal to the estimation biases and the RMSE of QMLE. The MHD estimator seems better performed than the QML estimator.

To illustrate the robustness of the MHD estimator, we proceed in this manner: Let

$$f_{n,\lambda} = (1-\lambda)f_n + \lambda\delta[0,1], \text{ where } \lambda \in [0,1],$$

and $\delta[0,1]$ the uniform density on the interval $[0,1]$. We vary λ between 0 and 1 and consider the MHD associated estimator. In each case, we replace f_n by $f_{n,\lambda}$. This gives the following table with nine values of λ.

Table 2. Robustness of the MHD estimator

λ	$\widehat{w}$	$\widehat{\alpha}$	$\widehat{\beta}$
0.1	0.08971371	0.14954614	0.39994447
0.2	0.08988248	0.14988466	0.40021855
0.3	0.09113607	0.15321818	0.39774967
0.4	0.09000711	0.14994950	0.40003438
0.5	0.09107550	0.15284870	0.40281130
0.6	0.08961155	0.14484359	0.39813678
0.7	0.08963666	0.14753569	0.39873747
0.8	0.08997696	0.15081873	0.40006430
0.9	0.08943709	0.14916746	0.39941254

The results of table 2 seem to indicate a certain robustness of the MHD estimator.

7. Application in Finance

The data used for our empirical study are daily returns of S&P500 index of 1272 observations. The study period is from 21-01-2009 to 02-01-2004 (cf figure 1 page 22).

We define the daily returns r_t of S&P500 index as follows:

$$r_t = \log\frac{p_t}{p_{t-1}},$$

where p_t is the price at the end of trading day t.

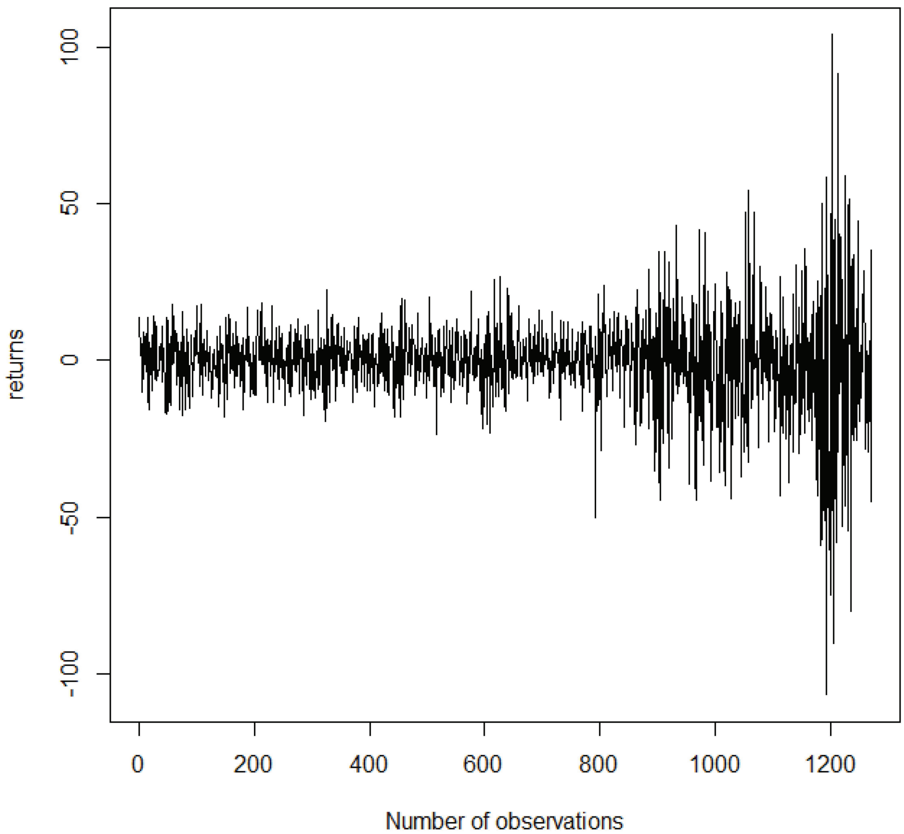

Figure 1. Returns of sp500 index from 02/01/2004 to 21/09/2009

One can adjust a GARCH(1,1) model to the series (r_t). Supposing that we are looking for the return at the date t. Two informations are particulaly interesting: the average and the variance. Particulary, we hope modeling the average and the conditional variance of r_t. Then, for the GARCH modeling, it amounts to write:

$$r_t = \mu + X_t, \tag{7.1}$$

where

$$\begin{aligned} X_t &= \sqrt{h_t}\eta_t \qquad n_t \leadsto i.i.d\mathcal{N}(0,1) \\ h_t &= w + \alpha X_{t-1}^2 + \beta h_{t-1}, \end{aligned}$$

with μ and h_t are respectily the conditional average and conditional variance of r_t.

To estimate the parameters (w, α, β) , we will use the MHD method. For that, we use the method described in the paragraph 6 with 50 replications of the model (7.1) and the bandwidth $b_n = n^{-0.23}$. We obtain

$$\widehat{w} = 1.30022275,\ \widehat{\alpha} = 0.06422973,\ \widehat{\beta} = 0.92449759.$$

and

$$\widehat{\mu} = -0.2311027$$

Finaly we can write,

$$r_t = -0.2311027 + X_t$$

where

$$X_t = \sqrt{h_t}\eta_t \qquad n_t \rightsquigarrow i.i.d\mathcal{N}(0,1)$$
$$h_t = 1.30022275 + 0.06422973X_{t-1}^2 + 0.92449759h_{t-1}.$$

8. Open problem

The multivariate GARCH process study case can be examined using the same method. Also, this study can be examined in EGARCH (Exponential GARCH) model case, IGARCH (integrated GARCH) model case and FIGARCH (Fractionary integrated GARCH) model case.

Acknowledgments

The authors are grateful to the anonymous referees whose suggestions helped them to improve the quality of the paper.

References

Beran, R. (1977). Minimum Hellinger Distance Estimation for parametic models. *The Annals of Statistic, 5*(2) 445-463.

Berkes, L. Horv ath, & P. Kokoszka(2003). GARCH processes: structure and estimation. *Bernoulli, 9*(2), 201-227.

Bitty, A., & Hili, O. (2010). Hellinger distance estimation of long memory linear processes*, C. R. Acad. Sci. Paris Sér. I 348 (2010), 7–8,* 445–448. https://doi:10.1016/j.crma.2010.02.020

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics 31, 307-327.*

Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.

Davis, R. A., & Mikosch, T. (2008). Extreme value theory for GARCH processes*. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.): Handbook of Financial Time Series, 186–200. Springer, New York.*

Engle, R. F. (1982). Autoregressive conditional heterosedasticity with estimates of the variane of United Kingdom infation, *Eonometria, 50,* 987 1008.

Francq, C., & Zakoian, J. M. (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH process. *Bernoulli, 10*(4), 605-637.

Gouriéroux, C., & Monfort, A. (1996). *Simulation-Based Econometric Methods*, Oxford University Press, New York.

Hang Hanyuan, Yunlong Feng, Ingo Steinwart, & Johan, A. K. Suykens (2015). Learning theory estimates with observations from general stationary stochastic processes. *Internal Report, ESAT, KU Leuven, Leuven, Belgium.*

Hili, O. (1995). On the estimation of nonlinear time series models*, Stochastics and Stochastics Reports, 52*, 207-226.

Ibragimov, I. A., & Linnik, Yu. V. (1971). Independent and Stationary Sequences of Random Variables. *Wolters-Noordhoff, Groningen.*

Ibragimov, l. A. (1975). A Note on the CLT for dependent random variables. *Theor. Prob. Appl., 20,* 135-140.

N'dri, A., & Hili, O. (2011). Estimation par la distance de Hellinger des processus gaussiens stationnaires fortement dépendants.*C. R. Acad. Sci. Paris, Ser. I 349,* 991-994. https://doi.org/10.1016/j.crma.2011.07.026

N'dri, A., & Hili, O. (2013). Hellinger Distance Estimation of Strongly Dependent Multi-Dimensional Gaussian Processes. *International Journal of Statistics and Probability, 2*(3).

N'drin, J. A., & Hili, O. (2013). Parameter Estimation of One-Dimensional Diffusion Process by Minimum Hellinger Distance Method. *Random Operators and Stochastic Equations, 21,* 403-424. https://doi.org/ 10.1515/rose-2013-0019

Takada, T. (2007). Robust estimtion of lent variable models with application to stochastic volativity mdels*, faculty of business Oska-city University, sugimoto, Sumiyoshi-ku, Osaka 558-8585, Japan.* https://doi.org/10.1016/j.csda.2008.06.017