Afrika Statistika

Afrika Statistika
Vol. 15 (1), 2020, pages 2225 - 2246.
DOI: http://dx.doi.org/10.16929/as/2020.2225.155

ISSN 2316-090X

Estimation and asymptotic properties of a stationary univariate GARCH(p, q) process

Aka Roger Kadjo (1,2,∗), **Ouagnina Hili** (2) **and Aubin N'dri** (2,3)

(1) University Félix Houphouët-Boigny, Abidjan-Cocody, Côte d'Ivoire.
(2) Laboratory of Mathematics and New Technologies of Information, National Polytechnic Institute Félix Houphouët-Boigny, BP 1911, Yamoussoukro, Côte d'Ivoire.
[3]University Jean Lorougnon Guédé, Côte d'Ivoire.

Received July 24,2019; Accepted January 28, 2020

Abstract. In this paper, we determine the Minimum Hellinger Distance estimator of a stationary univariate 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.

Résumé. Dans ce papier, nous déterminons l'Estimateur du Minimum de Distance de Hellinger d'un processus GARCH univarié stationnaire. Nous construisons un estimateur basé sur la méthode du Minimum de Distance de Hellinger. Sous les conditions de ϕ-mélange du processus GARCH, nous établissons les propriétés asymptotiques de cet estimateur.

Key words: Hellinger distance estimation; GARCH process; phi-mixing process; consistence; asymptotic normality.
AMS 2010 Mathematics Subject Classification : 62F12, 62G35, 62H12, 60G10.

(∗) Aka Roger Kadjo: roger.kadjo@yahoo.fr
Ouagnina Hili: o_hili@yahoo.fr
Aubin N'dri: aubin_ndri@yahoo.com

The authors.

(1) Aka Roger Kadjo, Ph.D. Student, is preparing his Ph.D. dissertation under the supervision of the second author, at University Félix Houphouët-Boigny, Abidjan-Cocody, Côte d'Ivoire.

(2) Ouagnina Hili, (Ph.D.), full professor, at National Polytechnic Institute Félix Houphouët-Boigny, Yamoussoukro, Côte d'Ivoire.

(3) Aubin N'dri, (Ph.D.), Assistant Professor, at University Jean Lorougnon Guédé, Côte d'Ivoire.

NB. All notation are given in the Appendix, page 2246.

1. Introduction

The GARCH (General Autoregressive Conditionally Heteroscedastic) models were pioneered by Engle (1982) and Bollerslev, 1986, and have ever since been widely used to analyze financial time series. The method of estimating parameters of the most widely used GARCH models in the literature is QMLE (See Berkes (2003) and Francq and Zakoian (2004)). Studies have shown that the resulting estimator is efficient but not robust. Thus, for the contaminated data the results obtained by the QMLE are not reassuring.

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 advantages of this method is that the obtained estimators are efficient and robust against disturbances (See Beran (1977)). Which can be interesting for contaminated data.

To show the performance of the MHD estimator, we compare it (in simulation section) to the QML estimator, to the MD estimator and to the M-estimator.

The minimum Hellinger distance estimators have been used in parameter estimation for independent observations (See Beran (1977)), for nonlinear time series models (See Hili (1995)) and recently for univariate long memory linear processes (See Bitty and Hili (2010)), for nonlinear univariate and multivariate gaussian process (See N'dri and Hili (2011) and N'dri and Hili (2013)), for parameter estimation of one-dimensional diffusion process (See N'drin and Hili (2013)).

The paper is organized as follows. In section 2 we give the definitions and some properties of the univariate GARCH model. Section 3 contains the definition of the estimator and some assumptions. In 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 did the conclusion.

2. Definitions and some properties of GARCH model

Definition 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 $E(\varepsilon_t) = 0$, $\mathrm{E}\left(\varepsilon_t^2\right) = 1$ 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.

Let $\theta = (w, \alpha_1, ..., \alpha_p, \beta_1, ..., \beta_q)^T \in \Theta \subset {}^{p+q+1}$ be the vector of the parameters of interest where Θ is a compact set, θ_0 the vector of the true value and T the transpose.

Proposition 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 (2.1) admits a unique strictly stationary solution.

Proof of Proposition 1. See Bollerslev, 1986. ■

Definition 2. Let $(X_t)_{t\in\mathbb{Z}}$ be a strictly stationary process, $\mathcal{F}_{-\infty}^{0}$ the σ-field generated by $\{..., X_{-2}, X_{-1}, X_0\}$ and $\mathcal{F}_k^{\infty}$ the σ-field generated by $\{X_k, X_{k+1}, X_{k+2}, ...\}$. Let's define the following :

$$\phi(k) = \sup_{C\in\mathcal{F}_{-\infty}^{0}, D\in\mathcal{F}_k^{\infty}} |P(D \mid C) - P(D)|.$$

The process (X_t) is said to be ϕ-mixing if $\phi(k) \longrightarrow 0$ when $k \longrightarrow \infty$.

Proposition 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 $0 < \rho < 1$).*

Proof of Proposition 2. See Davis and Mikosch (2008). ■

3. Assumptions and Estimation

We observe random variables, $X_1, \ldots, X_n$, a sequence of univariate GARCH process whose density belongs to the parametric family $\{f_\theta\}_{\theta\in\Theta}$ where Θ is the parameter space, a compact set of $\mathbb{R}^{p+q+1}$. We specify that in our study, the form of the density is not explicit.

We denote by f_n the kernel density estimator of f_θ, a non parametric estimator which is defined as

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

where $\mathrm{K}(.)$ is a kernel function and (b_n) is a sequence of bandwidths.

We construct an estimator $\widehat{\theta}_n$ of θ_0 over the family $\{f_\theta\}_{\theta\in\Theta}$. In order to do this, we choose the value of θ which minimizes the Hellinger distance (denoted by H_2) between f_n and f_θ defined by :

$$\widehat{\theta}_n = \arg\min_{\theta\in\Theta} H_2(f_n; f_\theta),$$

where

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

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

3.1. *Assumption (A3.1)*

The ε_t has a positive Lebesgue density on a neighborhood of 0 and

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

Assumption (**A 3.1**) gives the stationarity and the ϕ-mixing conditions of GARCH(p, q) model.

3.2. *Assumption (A3.2)*

1-For each $\theta \in \Theta$, the function $x \mapsto f_\theta(x)$ is positive and twice continuously differentiable.

2-For each $x \in \mathbb{R}$, the function $\theta \mapsto f_\theta(x)$ is continuously differentiable.

Assumption (**A 3.2**) is a technical assumption on the density of the GARCH(p, q) model.

3.3. *Assumption (A3.3)*

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

Assumption (**A 3.3**) is also a technical assumption on the density of the GARCH(p, q) model.

3.4. *Assumption (A3.4)*

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

Assumption (**A 3.4**) specifies the choice of the bandwidth.

3.5. *Assumption (A3.5)*

The kernel K is bounded with compact support, such that

$$\int_{\mathbb{R}} u\mathrm{K}(u)du = 0, \quad \int_{\mathbb{R}} u^2\mathrm{K}(u)du < \infty$$

and

$$\int_{\mathbb{R}} \mathrm{K}^2(u)du < \infty.$$

Assumption (**A 3.5**) specifies the choice of the kernel.

3.6. *Assumption (A3.6)*

For θ_1, $\theta_2 \in \Theta$, $\theta_1 \neq \theta_2$ implies that $\{x \in \mathbb{R},\ f_{\theta_1}(x) \neq f_{\theta_2}(x)\}$ is a set of positive Lebesgue measure.

Assumption (**A 3.6**) is the identifiability assumption on the parametrization.

The assumptions (A 3.2.2) and (A 3.6) ensure the existence of the estimator $\widehat{\theta}_n$.

4. Consistency of the estimator

Theorem 1. *[Almost sure convergence] Suppose that assumptions (A 3.1)-(A 3.6) 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 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

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

$U(g)$ may have multiple values, so we shall assume that it stands for any one of those values.

We have

$$|f_n(x) - f_{\theta_0}(x)| \leq |f_n(x) - Ef_n(x)| + \sup_{x\in\mathbb{R}} |Ef_n(x) - f_{\theta_0}(x)|\,.$$

By lemmas 1 and 2

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

From the continuity of the Hellinger distance (See Beran (1977), Theorem 1),

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

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

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

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

The following lemmas (Lemma 1 and Lemma 2) give the details of the proof of Theorem 1.

Lemma 1. *Suppose that assumptions (A 3.1)-(A 3.5) are satisfied. Then,*

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

Proof of Lemma 1. We have

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

where

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

Using Assumption (A 3.5) and Jensen's inequality, we get

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

Finally $\sup|\Delta_t| \leq 2K_0$, where K_0 is a constant.

We also have the following

$$
\begin{aligned}
E\left|\Delta_t\right|^2 &= E\left|\mathrm{K}(\frac{x-X_t}{b_n}) - E\mathrm{K}\left(\frac{x-X_t}{b_n}\right)\right|^2 \\
&= E(\mathrm{K}^2(\frac{x-X_t}{b_n})) - \left(E\mathrm{K}\left(\frac{x-X_t}{b_n}\right)\right)^2 \\
&\leq E(\mathrm{K}^2(\frac{x-X_t}{b_n})) \\
&= \int_{\mathbb{R}} \mathrm{K}^2(\frac{x-s}{b_n}) f_{\theta_0}(s) ds \\
&\leq \int_{\mathbb{R}} \mathrm{K}^2(\frac{x-s}{b_n}) \sup_{s\in\mathbb{R}} f_{\theta_0}(s) ds \\
&= \sup_{s\in\mathbb{R}} f_{\theta_0}(s) \int_{\mathbb{R}} \mathrm{K}^2(\frac{x-s}{b_n}) ds \\
&= \sup_{s\in\mathbb{R}} f_{\theta_0}(s) b_n \int_{\mathbb{R}} \mathrm{K}^2(u) du \\
&= C_1 b_n,
\end{aligned}
$$

where $C_1 = \sup_{s\in\mathbb{R}} f_{\theta_0}(s) \int_{\mathbb{R}} \mathrm{K}^2(u)du$.

Then, from the Bernstein-type inequality for ϕ-mixing processes established by Hanyuan *et al.*(2016) and for all $\epsilon > 0$, we obtain

$$
\begin{aligned}
P\left(\frac{n^{\frac{1}{4}}}{nb_n}\left|\sum_{t=1}^{n}\Delta_t\right| > \epsilon\right) &= P\left(\frac{1}{n}\left|\sum_{t=1}^{n}\Delta_t\right| > n^{-\frac{1}{4}}\epsilon b_n\right) \\
&\leq 2\exp\left\{\left(-\frac{\left(n^{-\frac{1}{4}}\epsilon b_n\right)^2 n}{8C_\phi\left(4C_1 b_n + 2K_0 n^{-\frac{1}{4}}\epsilon b_n\right)}\right)\right\} \\
&= 2\exp\left\{\left(-\frac{\epsilon^2\sqrt{n}b_n}{8C_\phi\left(4C_1 + 2K_0\epsilon n^{-\frac{1}{4}}\right)}\right)\right\},
\end{aligned}
$$

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

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

Then, using Assumption (A 3.4) 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,

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

Lemma 2. *Suppose that assumptions (A 3.1)-(A 3.5) are satisfied. Then,*

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

Proof of Lemma 2. We have

$$\begin{aligned} Ef_n(x) &= \frac{1}{b_n} E\left(\mathrm{K}(\frac{x - X_t}{b_n})\right) \\ &= \frac{1}{b_n}\int_{\mathbb{R}} \mathrm{K}(\frac{x-s}{b_n}) f_{\theta_0}(s)ds \\ &= \int_{\mathbb{R}} \mathrm{K}(u) f_{\theta_0}(x - b_n u)du. \end{aligned}$$

Using assumptions (A 3.1.1), (A 3.5) and Taylor's formula in one variable gives for x such that $|\delta - x| < |b_n u|$, we obtain

$$\begin{aligned} Ef_n(x) - f_{\theta_0}(x) &= \int_{\mathbb{R}} \mathrm{K}(u)\left[f_{\theta_0}(x - b_n u) - f_{\theta_0}(x)\right] du \\ &= \int_{\mathbb{R}} \mathrm{K}(u)\left[\frac{1}{2} b_n^2 u^2 f_{\theta_0}^{''}(\delta)\right] du. \end{aligned}$$

Then,

$$n^{\frac{1}{4}} \sup_{x\in\mathbb{R}} |Ef_n(x) - f_{\theta_0}(x)| \leq \frac{1}{2} n^{\frac{1}{4}} b_n^2 \sup_{x\in\mathbb{R}} \left|f_{\theta_0}^{''}(\delta)\right| \int_{\mathbb{R}} u^2 \mathrm{K}(u)du.$$

By assumptions (A 3.3), (A 3.4) and (A 3.5)

$$n^{\frac{1}{4}} \sup_{x\in\mathbb{R}} |Ef_n(x) - f_{\theta_0}(x)| \longrightarrow 0 \text{ when } n \longrightarrow \infty. \tag{4.2}$$

Hence,

$$\sup |Ef_n(x) - f_{\theta_0}(x)| = o\left(n^{-\frac{1}{4}}\right) \text{ when } n \longrightarrow \infty \ \blacksquare$$

5. Asymptotic normality of the estimator

For the following theorem, let's note,

$$g_\theta = f_\theta^{\frac{1}{2}}, \ \dot{g}_\theta = \frac{\partial g_\theta}{\partial \theta}, \ \ddot{g}_\theta = \frac{\partial^2 g_\theta}{\partial\theta\partial\theta^T}$$

and

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

Theorem 2. *[Asymptotic normality of the estimator]*

Suppose that assumptions (A 3.1)-(A 3.6) are satisfied. Furthermore, assume that

(i) the components of $\dot{g}_\theta$ and $\ddot{g}_\theta$ 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 $\int_{\mathbb{R}} \ddot{g_{\theta_0}}(x) g_{\theta_0}(x)dx$ 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

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

Proof of Theorem 2. From Theorem 2 in Beran (1977), we deduce that

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

where A_n is a $((p+q+1)\times(p+q+1))$-matrix whose components tends to zero as $n \longrightarrow +\infty$.

We have

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

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

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

We have

$$n^{\frac{1}{4}}\left|f_n(x)-f_{\theta_0}(x)\right| \leq n^{\frac{1}{4}}\left|f_n(x)-Ef_n(x)\right| + n^{\frac{1}{4}}\sup_{x\in\mathbb{R}}\left|Ef_n(x)-f_{\theta_0}(x)\right|.$$

By (4.1) and (4.2), we get

$$n^{\frac{1}{4}}\left|f_n(x)-f_{\theta_0}(x)\right| \longrightarrow 0 \ a.s \text{ when } n \longrightarrow \infty.$$

Hence,

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

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

On the other hand,

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

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

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

Therefore, by the lemmas 3 and 4, we get

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

where

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

We conclude that,

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

where

$$\Sigma^2 = \frac{1}{4}\left(\int_{\mathbb{R}^d} \dot{g}_{\theta_0}(x)\dot{g_{\theta_0}}(x)^T dx\right)^{-1} \blacksquare$$

The next lemmas (Lemma 3 and Lemma 4) allow us to prove the Theorem 2.

Lemma 3. *Let $h_\theta(.)$ the continuous function defined in (5.1). Suppose that assumptions (A 3.1)-(A 3.5) are 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 3. We have

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

Thus,

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

Step 1 : we prove that

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

Using Cauchy-Schwarz's inequality, we obtain

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} E\left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(X_i + ub_n\right) - h_{\theta_0}\left(X_i\right)\right) \mathrm{K}(u)du\right)^2 \\
&= \frac{1}{n}\left\{n \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right) \mathrm{K}(u)du\right)^2 f_{\theta_0}(x)dx\right\} \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right) \mathrm{K}(u)du\right)^2 f_{\theta_0}(x)dx \\
&\leq \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 du \times \int_{\mathbb{R}} \mathrm{K}^2(u)du\right] f_{\theta_0}(x)dx \\
&= C_3 \int_{\mathbb{R}}\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x)dudx \\
&= C_3 \int_{\mathbb{R}}\int_{\mathbb{R}} \left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x)dxdu.
\end{aligned}
$$

where C_3 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 2 imply that $\dot{g}_{\theta_0}$ is bounded (for θ_0 fixed). Then, h_{θ_0} is also bounded (for θ_0 fixed).

Thus, there exist a constant $C_4 > 0$ such that for all $n \in \mathbb{N}$

$$\left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 \leq C_4.$$

Since by Assumption (A 3.3) $0 < f_{\theta_0} < \infty$, we have

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

For all $n \in \mathbb{N}$ and for all $x \in \mathbb{R}$,

$$\left(h_{\theta_0}\left(x + ub_n\right) - h_{\theta_0}\left(x\right)\right)^2 f_{\theta_0}(x) \leq C_4 f_{\theta_0}(x).$$

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 get

$$\begin{aligned}
&\left|\int_{\mathbb{R}} \left(h_{\theta_0}(x+ub_n) - h_{\theta_0}(x)\right)^2 f_{\theta_0}(x)dx\right| \\
&\leq C_4 \int_{\mathbb{R}} f_{\theta_0}(x)dx \\
&= C_4.
\end{aligned}$$

Therefore, by the dominated convergence theorem

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

Step 2 : we prove that

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

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

$$\begin{aligned}
|\psi(x)| &= \left|\int_{\mathbb{R}} \left(h_{\theta_0}(x+ub_n) - h_{\theta_0}(x)\right)\mathrm{K}(u)du\right| \\
&\leq \int_{\mathbb{R}} \left|\left(h_{\theta_0}(x+ub_n) - h_{\theta_0}(x)\right)\right|\mathrm{K}(u)du \\
&\leq C_5 \int_{\mathbb{R}} \mathrm{K}(u)du \\
&= C_5.
\end{aligned}$$

Thus, there exist a constant C_6 such that

$$E\left|\psi(X_i)\right|^2 < C_6. \tag{5.2}$$

On the other hand, we know that

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

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

Therefore, $\phi^{\frac{1}{2}}(k) \leq C\text{exp}\left(\frac{1}{2}\text{k log}\rho\right) = C\text{exp}\left(-\frac{1}{2}\nu k\right)$ with $\nu = -\log\rho$.

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

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

We have

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

Using lemma 20.1 in Billingsley (1968) and (5.2), we obtain

$$\begin{aligned}
\frac{2}{n}\sum_{1\le i<j\le n} \left|E\left(\psi(X_i)\,\psi(X_j)\right)\right| &= \frac{2}{n}\sum_{j=1}^{n-1} j\left|E\left(\psi(X_1)\,\psi(X_{j+1})\right)\right| \\
&\le \frac{2}{n}\sum_{j=1}^{n-1} 2j\phi^{\frac{1}{2}}(j)\left(E\left|\psi(X_1)\right|^2\right)^{\frac{1}{2}}\left(E\left|\psi(X_{j+1})\right|^2\right)^{\frac{1}{2}} \\
&\le \frac{4C_6}{n}\sum_{j=1}^{n-1} j\phi^{\frac{1}{2}}(j) \\
&= \frac{4C_6}{n}\sum_{k=1}^{n-1}\sum_{l=k}^{n-1} \phi^{\frac{1}{2}}(l) \\
&\le \frac{4C_6}{n}\sum_{k=1}^{n-1}\sum_{l=k}^{n-1} \chi(l) \\
&= \frac{4C_6}{n}\sum_{k=1}^{n-1}\sum_{l=k}^{n-1} \left(\chi(l)\right)^{\frac{1}{2}}\left(\chi(l)\right)^{\frac{1}{2}}.
\end{aligned}$$

The fact that χ is a decreasing function, we get

$$\begin{aligned}
\frac{2}{n}\sum_{1\le i<j\le n} \left|E\left(\psi(X_i)\,\psi(X_j)\right)\right| &\le \frac{4C_6}{n}\sum_{k=1}^{n-1}\sum_{l=k}^{n-1} \left(\chi(l)\right)^{\frac{1}{2}}\left(\chi(k)\right)^{\frac{1}{2}} \\
&\le \frac{4C_6}{n}\sum_{k=1}^{\infty}\left(\chi(k)\right)^{\frac{1}{2}}\sum_{l=1}^{\infty}\left(\chi(l)\right)^{\frac{1}{2}} \\
&\le \frac{4C_6}{n}\left(\sum_{i=1}^{\infty}\left(\chi(i)\right)^{\frac{1}{2}}\right)^2 \longrightarrow 0.
\end{aligned}$$

Combining **step1** and **step2**, 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 conclude that,

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

Lemma 4. *Suppose that assumptions (A 3.1)-(A 3.5) are hold. Then,*

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

Proof of Lemma 4.
We have $\int_{\mathbb{R}} f_{\theta_0}(x)dx = 1$, and by Assumption (A 3.2.2), we get

$$\begin{aligned}\int_{\mathbb{R}} h_\theta(x)\, f_\theta(x)dx &= \frac{1}{2}\int_{\mathbb{R}} \dot{g}_\theta(x) f_\theta^{\frac{1}{2}}(x)dx \\ &= \frac{1}{2}\int_{\mathbb{R}} \frac{\partial}{\partial\theta} f_\theta^{\frac{1}{2}}(x) f_\theta^{\frac{1}{2}}(x)dx \\ &= \frac{1}{2}\int_{\mathbb{R}} \frac{\frac{\partial f_\theta(x)}{\partial\theta}}{2 f_\theta^{\frac{1}{2}}(x)} f_\theta^{\frac{1}{2}}(x)dx \\ &= \frac{1}{4}\int_{\mathbb{R}} \frac{\partial f_\theta(x)}{\partial\theta}dx \\ &= 0.\end{aligned}$$

Then,

$$E\left(h_{\theta_0}(X_i)\right) = \int_{\mathbb{R}} h_{\theta_0}(x)\, f_{\theta_0}(x)dx = 0.$$

Moreover

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

We have

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

Since h_{θ_0} is bounded (for θ_0 fixed), thus there exist a constant C_0 such that

$$E\left|h_{\theta_0}\left(X_i\right)\right|^2 \leq C_0.$$

Then, we obtain

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

We use the same approach as in the proof of Lemma 3, it follow that

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

where χ is defined in (5.3).

We conclude that,

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

By the convergence limit theorem in Dürr (1986), we get

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

where

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

6. Simulations

In this section we give some simulations for the minimum Hellinger distance estimator to show its performance. For this purpose, we consider the univariate GARCH model $\left(\widetilde{X}_1^s\left(\theta\right),...,\widetilde{X}_n^s\left(\theta\right)\right)$ for $s = 1, 2, ..., S$, obtained from s replications of the basic model. We define as in Takada, 2007 the function

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

which is an alternative to the intractable density function f_θ. For details one can consult Gouriéroux and Monfort, 1996. The following lemma justifies this method.

Lemma 5. *Let assumptions (A 3.1)-(A 3.6) be fulfilled, then*

$$\left| f_\theta(x) - \widetilde{f}_{n,\theta}(x) \right| \longrightarrow 0 \; a.s. \text{ when } n \longrightarrow \infty.$$

Proof of Lemma 5.
We have

$$\begin{aligned}
\left| f_\theta(x) - \widetilde{f}_{n,\theta}(x) \right| &\leq \sup_{x\in\mathbb{R}} \left| f_\theta(x) - E\left(\widetilde{f}_{n,\theta}(x)\right) \right| \\
&\quad + \left| E\left(\widetilde{f}_{n,\theta}(x)\right) - \widetilde{f}_{n,\theta}(x) \right| \\
&\leq \frac{1}{s}\sum_{s=1}^{S} \sup_{x\in\mathbb{R}} \left| f_\theta(x) - E\left(\widetilde{f}^{(s)}_{n,\theta}(x)\right) \right| \\
&\quad + \frac{1}{s}\sum_{s=1}^{S} \left| E\left(\widetilde{f}^{(s)}_{n,\theta}(x)\right) - \widetilde{f}^{(s)}_{n,\theta}(x) \right|
\end{aligned}$$

where for $s = 1, ..., S$,

$$\widetilde{f}^{(s)}_{n,\theta}(x) = \frac{1}{nb_n}\sum_{t=1}^{n} \mathrm{K}\left(\frac{x - \widetilde{X}^s_t(\theta)}{b_n}\right).$$

Using Lemma 2, we show that

$$\sup_{x\in\mathbb{R}} \left| f_\theta(x) - E\left(\widetilde{f}^{(s)}_{n,\theta}(x)\right) \right| \longrightarrow 0 \text{ when n} \longrightarrow \infty,$$

for each s.

Similarly, the Lemma 1 allows us to show that

$$\left| E\left(\widetilde{f}^{(s)}_{n,\theta}(x)\right) - \widetilde{f}^{(s)}_{n,\theta}(x) \right| \longrightarrow 0 \; a.s. \text{ when } n \longrightarrow \infty,$$

for each s ■

Simulations are based on 500, 1000 or 2000 observations of GARCH (1,1) process. The kernel density f_n and the function $\widetilde{f}_{n,\theta}$ are constructed by using Biweight's kernel or Epanechnikov's kernel and the bandwidth $b_n = n^{-0.27}$. We use in R "fgarch and fbasics" packages. To compare the performance of the estimators, we use the root mean square error (RMSE) defined as follows :

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

Tables 1 to 3 compare, on the one hand the MHDE and QMLE and on the other hand the MHDE and MDE, the MHDE and LADE and QMLE. Table 4 shows the robustness of MHDE.

	$\widehat{w}$	$\widehat{\alpha}$	$\widehat{\beta}$
mean(MHDE)	0.091075	0.153218	0.402811
RMSE	0.386817	0.054185	0.030398
mean(QMLE)	0.085698	0.177487	0.300088
RMSE	0.382285	0.734444	0.114497

Table 1. Comparison of MHDE and QMLE : Results for $n = 500$, $S = 100$, $w = 0.09$, $\alpha = 0.15$, $\beta = 0.4$, Epanechnikov's kernel and $\varepsilon_t \rightsquigarrow \mathcal{N}(0,1)$.

Table 1 shows that the RMSE's of the MHDE are smaller or almost equal to those of QMLE. The MHDE seems to perform better than the QMLE.

	g	$\widehat{\alpha}$ mean	$\widehat{\alpha}$ RMSE	$\widehat{\beta}$ mean	$\widehat{\beta}$ RMSE
MHDE		0.1038	0.0225	0.6004	0.0151
^{a}MDE$_{NW}$	30	0.0959	0.0567	0.5420	0.2444
^{a}MDE$_{NW}$	40	0.0992	0.0575	0.5463	0.2438

Table 2. Comparison of MHDE and MDE : Results for $n = 1000$, $S = 1000$, $w = 0.006$, $\alpha = 0.1$, $\beta = 0.6$, Biweight's kernel and $\varepsilon_t \rightsquigarrow t_5$.

g is the number of autocorrelations

From Table 2 the MHDE's are more efficient than MDE$_{NW}$.

	$\widehat{w}$	$\widehat{\alpha}$	$\widehat{\beta}$
mean(MHDE)	0.1043	0.1072	0.7544
RMSE	0.0207	0.0120	0.0235
^{b}mean(QMLE)	0.1136	0.1063	0.7231
^{b}RMSE	0.0540	0.0379	0.1032
^{b}mean(LADE)	0.1163	0.1085	0.7175
^{b}RMSE	0.0519	0.0334	0.0961

Table 3. Comparison of MHDE, LADE and QMLE : Results for $n = 2000$, $S = 1000$, $w = 0.1$, $\alpha = 0.1$, $\beta = 0.75$, Biweight's kernel and $\varepsilon_t \rightsquigarrow t_5$.

We can see in Table 3 that the MHDE is more efficient than LADE and the QMLE.

a : Baillie and Chung (2001)
b : Tinkl (2013)

To illustrate the robustness of the MHD estimator, we proceed as follows : in the MHD estimation, we replace $\widetilde{f}_{n,\theta}$ by $\widetilde{f}_{n,\theta,\lambda}$ which is defined as follows

$$\widetilde{f}_{n,\theta,\lambda} = (1-\lambda)\widetilde{f}_{n,\theta} + \lambda\delta[0,1], \text{ where } \lambda \in [0,1],$$

and $\delta[0,1]$ the uniform density on the interval $[0,1]$. We make λ vary between 0 and 1 and we consider the estimators of the associated MHD.

λ	$\widehat{w}$	$\widehat{\alpha}$	$\widehat{\beta}$
0.1	0.08971672	0.14978625	0.39864413
0.2	0.08908243	0.1489463	0.39921852
0.3	0.09013601	0.15121177	0.39804661
0.4	0.09010718	0.14984752	0.40005337
0.5	0.08910752	0.15284870	0.40281130
0.6	0.08961155	0.14484359	0.39813678
0.7	0.09063663	0.14853763	0.39963746
0.8	0.08996624	0.15061865	0.40016230
0.9	0.08933807	0.14917768	0.39971352

Table 4. Robustness of the MHDE : Results for $n = 500$, $S = 100$, $w = 0.09$, $\alpha = 0.15$, $\beta = 0.4$, Epanechnikov's kernel and $\varepsilon_t \rightsquigarrow$ Log-$\mathcal{N}(0,1)$.

For each value of λ, the difference between the resulting estimators of the simulation is not meaningful. Therefore, although $\widetilde{f}_{n,\theta}$ are contaminated, the estimators obtained are closed to the true values. Thus, the results of table 4 show the robustness of the MHDE.

7. Conclusion

In this paper; we focused on estimating of the parameter θ_0 of a univariate GARCH process. We construct an estimator $\widehat{\theta}_n$ of θ_0 using the Minimum Hellinger Distance method. The obtained estimator, minimizes the Hellinger distance between the probability density of the univariate GARCH process f_θ and the estimator f_n of this density.

Under regular assumptions, we showed that f_n convergence almost surely to f_{θ_0}. Therefore, using the continuity of the Hellinger distance, we have shown the almost sure convergence of $\widehat{\theta}_n$. We have also studied the asymptotic distribution of $\widehat{\theta}_n$. Through simulations, we have shown that the resulting estimator is efficient and robust.

Acknowledgments. The authors cordially thank the editorial team and the presenter of *Afrika Statistika Journal*. They also thank the referee for his helpful comments which have led to improvements of the manuscript.

References

Baillie T. Richard and Chung Huimin (2001). Estimation of GARCH models from the autocorrelations of the squares of a process. *Journal of time series analysis vol. 22, No. 6*

Beran, R. (1977). Minimum Hellinger distance estimates for parametric models. *Ann. Statist.*, 5(3), 445-463. Doi : http://dx.doi.org/10.1214/aos/1176343842.

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

Bitty, A. L. et Hili, O. (2010). Estimateurs du minimum de distance de hellinger des processus linéaires à longue mémoire, *Comptes Rendus Mathématique*, 348(7-8) : 445-448. Doi : https://doi.org/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. and 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.*

Dürr D., Goldstein S. (1986). Remarks on the central limit theorem for weakly dependent random variables. In: Albeverio S.A., Blanchard P., Streit L. (eds) Stochastic Processes — Mathematics and Physics. Lecture Notes in Mathematics, vol 1158. Springer, Berlin, Heidelberg. DOI https://doi.org/10.1007/BFb0080211.

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

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

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

Hanyuan H., Feng Y., Steinwart I. and Johan A.K. S. (2015) Learning theory estimates with observations from general stationary stochastic processes. *arXiv:1605.02887v1 [stat.ML] 10 May 2016.*

Hili, O. (1995). On the estimation of nonlinear time series models, Stochastics and Stochastics. *Reports*, 52(3-4), 207-226. Doi :http://dx.doi.org /10.1080/17442509508833972.

N'dri, A. and 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.* doi:10.1016/j.crma.2011.07.026

N'dri A. and Hili O. (2013), Hellinger distance estimation of strongly dependent Multidimensional Gaussian processes. *International Journal of Statistics and Probability*, 2(3),70-84. Doi https://doi.org/10.5539ijsp.v2n3p70.

N'drin, J.A and 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 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:10.1016/j.csda.2008.06.017

Tinkl F. (2013). *Asymptotic Theory for M-Estimators in General Autoregressive Conditional Heteroscedastic Time Series Models*. Thesis, Friedrich Alexander Erlangen-Nuremberg University.

Appendix : Notation

$\xrightarrow{D}$	:	Converge in distribution.
$\longrightarrow$ a.s	:	Almost sure convergence.
$C_0, C_1, ..., C_6$	:	Generic constants taking different values from time to time.
$\mathcal{N}(0,1)$	:	Normal distribution with mean 0 and standard deviation 1.
$Log - \mathcal{N}(.,.)$	:	Lognormal distribution.
t_5	:	Student-t distribution with 5 degrees of freedom.
LADE	:	Least Absolute Deviation-type Estimator.
QMLE	:	Quasi-Maximum Likelihood Estimator.
MDE_NW	:	Minimum Distance Estimator use the Newey-West method.
MHDE	:	Minimum Hellinger Distance Estimator.