Journal of Mathematical Sciences: Advances and Applications
Volume 49, 2018, Pages 59-79
Available at http://scientificadvances.co.in
DOI: http://dx.doi.org/10.18642/jmsaa_7100121926

MINIMUM HELLINGER DISTANCE ESTIMATES OF LONG-RANGE DEPENDENCE GAUSSIAN PROCESSES

A. Y. N'DRI[1,2] and O. HILI[2]

[1]University Jean Lorougnon Guédé
Daloa
Côte d'Ivoire
e-mail: aubin_ndri@yahoo.com

[2]Laboratory of Mathematics and New
Technologies of Information
National Polytechnic Institute Félix
Houphouët-Boigny of Yamous-soukro
Côte d'Ivoire

Abstract

In the present paper, we propose a new method to determine the minimum Hellinger distance estimator of stationary Gaussian univariate processes with long-range dependence. Under some assumptions which ensure some probabilistic properties, we establish strong consistency and asymptotic distribution properties of this estimator.

1. Introduction

Many empirical observations of both natural and industrial data exhibit a correlation decreasing at a rate slower. The dependence among

2010 Mathematics Subject Classification: 34D05.
Keywords and phrases: asymptotic properties, long-range dependence, Hellinger distance estimation, Hermite polynomials.
Received January 30, 2018; Revised February 21, 2018

observations separated by a large time span may therefore be not negligible. Time series with long-range dependence appear in many contexts. Well known examples of phenomena affected by long-range dependence are the levels of the Nile river, the temperature of the earth (Beran [1], p. 20-29), economic time series (Granger [5]). Some research that suggests the use of models incorporating long dependency have been done for processes encountered in telecommunication networks (De Giovanni and Naldi [4]). In particular such models have been validated for internet networks (Willinger et al. [14]). Two of the best known classes stochastic processes are the increments of Fractional Brownian Motion (known as Fractional Gaussian Noise (FGN) Beran ([1], p. 55), Samorodnitsky et al. ([13], p. 332)) and Fractional ARIMA (FARIMA) (Granger et al. [6], Hosking [8]).

This paper is developed under the framework of parametric estimation using the Minimum Hellinger Distance (MHD) method. We opt for this method of estimation because it gives an estimator which is strongly consistent and especially robust. Our study is essentially based on the long range dependence process as in N'dri and Hili ([10], [11]). In their different studies, they used a power of the sample size with a decay hypothesis on the kernel density estimator to get the strong consistency and asymptotic distribution properties of the parametric estimator. We resume their paper (univariate case) to improve it. We work directly on the sample size of the process to get the same results. Without loss of generality, we use a kernel density function with compact support.

We consider the sequence $(X_n)_{n\geq 1}$ that is a $\mathbb{R}$-valued stationary Gaussian process with density $f(x, \theta_0)$, where θ_0 is assumed to belong to a compact subset Θ of $\mathbb{R}^q$. We suppose that $E(X_n)^2 = \sigma^2(\theta_0)$ and defined $\mathbb{E}(X_n X_{n+t}) := \gamma(t) = t^{-\alpha(\theta_0)}\mathcal{L}(t)$, where $t, n \in \mathbb{N}^*$ and $\mathcal{L}$, a slowly varying function at infinity:

$$\lim_{s\to+\infty} \frac{\mathcal{L}(st)}{\mathcal{L}(s)} = 1,\ t > 0 \text{ and } 0 < \alpha(\theta_0) < 1.$$

The results in this paper are organized as follows. In Section 2, we define the notations, the useful hypotheses and the main results. That is to say, almost sure convergence and asymptotic distribution properties of the estimator $\hat{\theta}_n$. Proofs of these properties are detailed in Section 3. The results of some simulations are described in Section 4. They are designed to evaluate the performance in finite samples of the proposed estimator.

The following notations are adopted throughout the paper: $\sim$, $[x]$, $\to^P$, $\to^{\mathcal{D}}$ denote, respectively, asymptotically equivalent, nearest integer to x, convergence in probability, and convergence in distribution.

2. Statement of Assumptions and Results

2.1. Notations

Let $\mathcal{F} = \{f(., \theta)\}$ be a family of functions indexed over a compact parameter set $\Theta \subset \mathbb{R}^q$, such that for each $\theta \in \Theta$, $f(x, \theta) = \frac{1}{\sqrt{2\pi}\sigma(\theta)} \exp\left(-\frac{x^2}{2\sigma^2(\theta)}\right)$.

The Hellinger distance H_2 and the functional T in question are defined as

$$H_2 : \mathcal{F} \times \mathcal{B} \to \mathbb{R}$$

$$(f(., \theta), g) \mapsto H_2(f(., \theta), g) = \left(\int_{\mathbb{R}} |f^{\frac{1}{2}}(x, \theta) - g^{\frac{1}{2}}(x)|^2 dx\right)^{\frac{1}{2}},$$

and

$$T : \mathcal{B} \to \Theta$$

$$g \mapsto T(g) = \arg\min_{\theta \in \Theta} H_2(f(., \theta), g),$$

where $\mathcal{B} = L^1(\mathbb{R})$.

Remark 1. $T(g)$ may have multiple values, so we shall assume that it stands for any one of those values.

2.2. Assumptions

In the rest of our study, (b_n) will design a sequence of bandwidth.

Assumptions A

(A1) There exists a non-decreasing sequence of positive integers (d_n) such that $d_n < n$ and

$$(\gamma^*(d_n)n / d_n) + \left((n / d_n)b^2_{n/d_n}\right)^{-1} \ln(d_n) = o(1),$$

where $\gamma^*(n) = \sup_{|n|\le t}(\gamma(t))$.

(A2) The kernel K is bounded with compact support, such that $\int_{\mathbb{R}} uK(u)du = 0$, $\int_{\mathbb{R}} u^2K(u)du \neq 0$, and $\int_{\mathbb{R}} |u|^2K(u)du < +\infty$.

(A3) Suppose that $\lim_{n\to+\infty} b_n = 0$, $\lim_{n\to+\infty} \sqrt{n}b_n^2 = +\infty$, and $\lim_{n\to+\infty} n^{m\alpha(\theta_0)/2}b_n^2 = 0$.

Assume $\mathcal{F}$ satisfies the following:

Assumptions B

(B1) For each $\theta \in \Theta$, the function $x \mapsto f(x, \theta)$ is twice continuously differentiable.

(B2) For almost every $x \in \mathbb{R}$, the function $\theta \mapsto f(x, \theta)$ is continuous.

(B3) For almost every $x \in \mathbb{R}$, the function $\theta \mapsto \frac{\partial}{\partial\theta_j} f^{1/2}(x, \theta)$, for $1 \le j \le q$ is continuous and for every j, the function $\theta \mapsto \frac{\partial}{\partial\theta_j} f^{1/2}(x, \theta)$ is in $L^2(\mathbb{R}^q)$.

(B4) For almost every $x \in \mathbb{R}$, the function $\theta \mapsto \frac{\partial^2}{\partial\theta_j \partial\theta_k} f^{1/2}(x, \theta)$, $1 \le j, k \le q$ is continuous and for every j, k, the function $\theta \mapsto \frac{\partial^2}{\partial\theta_j \partial\theta_k} f^{1/2}(x, \theta)$ is in $L^2(\mathbb{R}^q)$.

(B5) For $\theta_1, \theta_2 \in \Theta$, $\theta_1 \neq \theta_2$ implies $f(x, \theta_1) \neq f(x, \theta_2)$ on a set of positive Lebesgue measure.

Remark 2. The assumptions (B2) and (B5) ensure existence of $T(g)$. On the other hand, (B5) is the identifiability assumption on the parametrization.

2.3. Main results

The main result in this subsection discusses the asymptotic behaviour of the MHD estimator. We study in Theorem 1, the efficiency property of this estimator. The study of asymptotic distribution property in Theorem 2 is very important because it is useful in the selection criteria of the estimators. Knowing the asymptotic distribution of the estimator can allow to solve the estimation problem to the interval confidence construction or the hypotheses tests.

We first observe random variables $X_1, \ldots, X_n$ with true density $f(., \theta_0)$, where $\theta_0 \in \Theta$. We denote by f_n the kernel density estimator of $f(., \theta_0)$, a nonparametric estimator which is defined as

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

where K a kernel function.

Then, we construct an estimator $\widehat{\theta_n}$ of the parameter θ over the class $\mathcal{F}$. In order to do this we choose the value of θ which minimizes the functional $H_2(f(., \theta), f_n(.))$. As a matter of fact, $\widehat{\theta_n} = T(f_n)$ and $\theta_0 = T(f(., \theta_0))$.

Theorem 1 (Almost sure convergence)**.** *Let assumptions* (A1)-(A3), (B2) *and* (B5) *be fulfilled. If* θ_0 *is in the interior of* Θ, *then* $\hat{\theta}_n$ *almost surely converges to* θ_0 *when* $n \to +\infty$.

For the following theorem, we consider I_d the identity function, $I\{.\}$ the indicator function, $F(.)$ the continuous marginal distribution function of the sequence $(X_n)_{n\geq 1}$ and $F_n(x) = n^{-1}\sum_{i=1}^{n} I\{X_i \leq x\}$, $x \in \mathbb{R}$, the sample distribution function. The integer m is the Hermite rank of the family $\{I\{I_d(.) \leq x\} - F(x) : x \in \mathbb{R}\}$ and $F(x) = \text{Prob}(X_i \leq x)$.

Otherwise, let's pose $F_n(x) - F(x) = n^{-1}\sum_{i=1}^{n} B_x(X_i)$, with $B_x(.) = I\{I_d(.) \leq x\} - F(x)$, then in $L^2\phi = \left\{g : \int_{-\infty}^{+\infty} g^2(z)\phi(z)dz < +\infty\right\}$ with $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$, let's consider the Fourier-Hermite expansion $B_x(.) = \sum_{k=m}^{+\infty} J_k(x)\frac{H_k(.)}{k!}$, where $H_k(z) = (-1)^k \exp(z^2/2)\frac{d^k}{dz^k}\left(\exp(-z^2/2)\right)$, $z \in \mathbb{R}$ is the k-th Hermite polynomial and $J_k(x) = E(H_k(.)B_x(.))$, $k = 0, 1, \ldots$.

Finally, denote by $d_{m,n}^2 = \mathrm{Var}\left(\sum_{i=1}^{n} H_m(X_i)\right)$, $J'_m(x) = \dfrac{dJ_m(x)}{dx}$, $Y_{m,n} = \dfrac{1}{d_{m,n}} \sum_{i=1}^{n} \dfrac{H_m(X_i)}{m!}$ and the long-range dependence condition for the sequence $(X_n)_{n\geq 1}$ becomes $0 < \alpha(\theta_0) < 1/m$. In addition, $nd_{m,n}^{-1} \sim \sqrt{\dfrac{(1 - m\alpha(\theta_0))(2 - m\alpha(\theta_0))}{2m!}}\, \dfrac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)}$.

Theorem 2 (Asymptotic distribution)**.** *Let assumptions* (A1)-(A3) *and* (B1)-(B5) *be fulfilled. If* θ_0 *lies in the interior of* Θ *and* $\int_{-\infty}^{+\infty} \dot{S}(x, \theta_0)\, \dot{S}(x, \theta_0)^t dx$ *is a non-singular* $(q \times q)$*-matrix, then the limiting distribution of*

$$nd_{m,n}^{-1}(\hat{\theta}_n - \theta_0) \text{ is } (Z_m(1)/m!) \int_{-\infty}^{+\infty} \left(\rho(x, \theta_0)/2f^{1/2}(x, \theta_0)\right) J'_m(x) dx,$$

where $(Z_m(t))_{0\leq t\leq 1}$ *is the Hermite process,*

$$S(., \theta) = f^{1/2}(., \theta),\ \dot{S}(., \theta) = \left(\frac{\partial}{\partial\theta_1} f^{1/2}(., \theta), \cdots, \frac{\partial}{\partial\theta_q} f^{1/2}(., \theta)\right)^t,$$

$\dot{S}(., \theta)^t$ *is the transpose of* $\dot{S}(., \theta)$ *and*

$$\rho(x, \theta) = \left[\int_{-\infty}^{+\infty} \dot{S}(x, \theta)\dot{S}(x, \theta)^t dx\right]^{-1} \dot{S}(x, \theta).$$

3. Proofs

Proof of Theorem 1. First we state the following lemma required in proof.

Lemma. *If the assumptions* (A1)-(A3), (B2) *and* (B5) *are satisfied, then* f_n *almost surely converges to* $f(., \theta_0)$ *in the Hellinger topology.*

Proof. In the proof of this Lemma, we use a similar technical as in Ho [7].

We have $f_n(x) - f(x, \theta_0) = f_n(x) - Ef_n(x) + (Ef_n(x) - f(x, \theta_0))$ and

$$f_n(x) - Ef_n(x) = \frac{1}{nb_n}\sum_{j=1}^{n}\left(K\left(\frac{x - X_j}{b_n}\right) - EK\left(\frac{x - X_j}{b_n}\right)\right)$$

$$= \frac{1}{nb_n}\sum_{j=1}^{n}\left(K\left(\frac{y - Y_j}{h_n}\right) - EK\left(\frac{y - Y_j}{h_n}\right)\right),$$

where $y = x / \sigma(\theta_0)$, $h_n = b_n / \sigma(\theta_0)$ and $Y_j = X_j / \sigma(\theta_0)$ is standard Gaussian. We subdivide the centered sums $K\left(\frac{y - Y_j}{h_n}\right) - EK\left(\frac{y - Y_j}{h_n}\right)$ into disjoint blocks of length d_n.

$\left[\frac{n}{d_n}\right]$ is the number of blocks from the original series $\left(Y_j\right)_{1\le j\le n}$ with replacement by

$$\mathcal{B}(j, d_n) = \left(Y_{i+(j-1)d_n}\right)_{1\le j\le [n/d_n]}, \tag{1}$$

for $1 \le i \le d_n$.

As the sample size $n \to +\infty$, the block size $[n / d_n]$ grows more slowly than the sample size: $[n / d_n] = o(n)$.

Then,

$$\sum_{j=1}^{n}\left(K\left(\frac{y - Y_j}{h_n}\right) - E\left(K\left(\frac{y - Y_j}{h_n}\right)\right)\right)$$

$$= \sum_{i=1}^{d_n}\sum_{j=1}^{[n/d_n]}\left(K\left(\frac{y - Y_{i+jd_n}}{h_{[n/d_n]}}\right) - E\left(K\left(\frac{y - Y_{i+jd_n}}{h_{[n/d_n]}}\right)\right)\right),$$

and for $\varepsilon > 0$,

$$\mathbb{P}(|f_n(x) - Ef_n(x)| \geq \varepsilon)$$

$$\leq \sum_{i=1}^{d_n} \mathbb{P}\left(\frac{1}{[n/d_n]b_{[n/d_n]}} \left| Z_{[n/d_n]}(y, Y) \right| \geq \varepsilon \right)$$

$$= \sum_{i=1}^{d_n} \mathbb{P}\left(\left| Z_{[n/d_n]}(y, Y) \right| \geq [n/d_n]b_{[n/d_n]}\varepsilon \right),$$

where $Z_{[n/d_n]}(y, Y) = \sum_{j=1}^{[n/d_n]} \left(K\left(\frac{y - Y_{i+jd_n}}{h_{[n/d_n]}} \right) - E\left(K\left(\frac{y - Y_{i+jd_n}}{h_{[n/d_n]}} \right) \right) \right)$. We have

$$\mathbb{P}\left(\left| Z_{[n/d_n]}(y, Y) \right| \geq [n/d_n]b_{[n/d_n]}\varepsilon \right)$$

$$= \int_{B(n,i,\varepsilon)} \tilde{f}_n(y_1, y_2, \ldots, y_{[n/d_n]}) dy_1 dy_2 \ldots dy_{[n/d_n]},$$

where

$$B(n, i, \varepsilon) = \left\{ (y_1, y_2, \ldots, y_{[n/d_n]}), \left| \sum_{j=1}^{[n/d_n]} G(y_j) \right| \geq [n/d_n]b_{[n/d_n]}\varepsilon \right\},$$

and $\tilde{f}_n$ the joint probability of

$$\left\{ Y_{i+jd_n},\ j = 1, \ldots, [n/d_n],\ i = 1, \ldots, d_n \right\}.$$

Consider the set of $[n/d_n]$ Gaussian random variables $\left\{ Y_{i+jd_n},\ j = 1, \ldots, [n/d_n],\ i = 1, \ldots, d_n \right\}$; the covariance of any two of the random variables $Y_{i+j_1 d_n}$ and $Y_{i+j_2 d_n}$, $j_1 \neq j_2$ satisfies

$$[n/d_n]|E(Y_{i+j_1 d_n} Y_{i+j_2 d_n})| = [n/d_n]|\gamma(|j_1 - j_2|d_n)|$$

$$\leq \gamma^*(d_n)[n/d_n]$$

$$< 1 \text{ for large } n \text{ (by(A1))}.$$

This enables us to use Lemma 3.3 of Taqqu [12] to express $\tilde{f}_n$ as a uniformly convergent series over $\mathbb{R}^{[n/d_n]}$.

For $n \geq n_0$,

$$\tilde{f}_n(y_1, y_2, \ldots, y_{[n/d_n]})$$

$$= \sum_{q=0}^{\infty} \sum_{\substack{k_1+\ldots+k_{[n/d_n]}=2q \\ 0<k_1,\ldots,k_{[n/d_n]}<q}} \left\{ E \prod_{j=1}^{[n/d_n]} H_{k_j}(Y_{i+jd_n}) \right\} \prod_{j=1}^{[n/d_n]} \frac{H_{k_j}(y_j)}{k_j!} \phi(y_j), \qquad (2)$$

where $H_q(y)$ denotes the q-th Hermite polynomial of $K\left(\frac{x-X}{b_n}\right) - EK\left(\frac{x-X}{b_n}\right)$ in $L^2(\phi)$.

By Holder's inequality and using $\int H_q^2(y)\phi(y)dy = q!$ we get,

$$\int_{B(n,i,\varepsilon)} \tilde{f}_n(y_1, y_2, \ldots, y_{[n/d_n]}) dy_1 dy_2 \ldots dy_{[n/d_n]}$$

$$\leq \left\{ \sum_{q=0}^{\infty} \sum_{\substack{k_1+\ldots+k_{[n/d_n]}=2q \\ 0<k_1,\ldots,k_{[n/d_n]}<q}} \left| E \prod_{j=1}^{[n/d_n]} \frac{H_{k_j}(y_{i+jd_n})}{\sqrt{k_j!}} \right| \right\}$$

$$\left\{ \int_{B(n,i,\varepsilon)} \prod_{j=1}^{[n/d_n]} \phi(y_j) dy_j \right\}^{1/2}.$$

Apply Lemma 3.1 of Taqqu [12] (cf. the proof of the Lemma 3.1 of the paper) to obtain

$$\sum_{q=0}^{\infty} \sum_{\substack{k_1+\ldots+k_{[n/d_n]}=2q \\ 0<k_1,\ldots,k_{[n/d_n]}<q}} \left| E \prod_{j=1}^{[n/d_n]} \frac{H_{k_j}(y_{i+jd_n})}{\sqrt{k_j!}} \right|$$

$$\leq \sum_{q=0}^{\infty} \sum_{\substack{k_1+\ldots+k_{[n/d_n]}=2q \\ 0<k_1,\ldots,k_{[n/d_n]}<q}} \gamma^*(d_n)^{(k_1+\ldots+k_{[n/d_n]})/2} \prod_{j=1}^{[n/d_n]} ([n/d_n]-1)^{k_j/2}$$

$$\leq \left(\sum_{k=0}^{\infty} (\gamma^*(d_n)([n/d_n]-1))^{k/2} \right)^{[n/d_n]}$$

$$= \exp(-[n/d_n]\ln(1-V(n))), \tag{3}$$

where $V(n) = (\gamma^*(d_n)([n/d_n]-1))^{1/2}$.

Using Bernstein inequality, we obtain for some $C > 0$ and for some $\kappa > 0$

$$\left\{ \int_{B(n,i,\varepsilon)} \prod_{j=1}^{[n/d_n]} \phi(y_j)\, dy_j \right\}^{1/2}$$

$$\leq 2\exp\left(-\frac{[n/d_n]^2 b^2_{[n/d_n]}\varepsilon^2}{2[n/d_n]C + 2\kappa[n/d_n]b_{[n/d_n]}\varepsilon} \right) \tag{4}$$

$$= 2\exp\left(-\frac{[n/d_n]b^2_{[n/d_n]}\varepsilon^2}{2C + 2b_{[n/d_n]}\kappa\varepsilon} \right). \tag{5}$$

Combining (3) and (5) gives for $\varepsilon > 0$,

$$\mathbb{P}(|f_n(x) - Ef_n(x)| \geq \varepsilon)$$

$$\leq 2\exp\left(-\frac{[n/d_n]b^2_{[n/d_n]}\varepsilon^2}{2C + 2b_{[n/d_n]}\kappa\varepsilon} - [n/d_n]\ln(1 - V(n)) + \ln(d_n)\right)$$

$$= \exp\left(-[n/d_n]b^2_{[n/d_n]}\left(\frac{\varepsilon^2}{2C + 2b_{[n/d_n]}\kappa\varepsilon} + U(n)\right)\right),$$

where

$$U(n) = \frac{\sqrt{[n/d_n]}\ln(1 - V(n))}{\sqrt{[n/d_n]}b^2_{[n/d_n]}} - \frac{\ln(d_n)}{[n/d_n]b^2_{[n/d_n]}}.$$

Using assumption (A1) and the fact that $\sqrt{[n/d_n]}\ln(1 - V(n)) \sim -[n/d_n]\gamma^*(d_n)$, we obtain

$$\frac{\varepsilon^2}{2C + 2b_{[n/d_n]}\kappa\varepsilon} + \frac{\sqrt{[n/d_n]}\ln(1 - V(n))}{\sqrt{[n/d_n]}b^2_{[n/d_n]}} - \frac{\ln(d_n)}{[n/d_n]b^2_{[n/d_n]}} \to \frac{\varepsilon^2}{2C}$$

as $n \to +\infty$. Hence there is $n_1(\varepsilon)$ such that, for all $n \geq n_1(\varepsilon)$

$$\mathbb{P}(|f_n(x) - Ef_n(x)| \geq \varepsilon) \leq \exp\left(-\left([n/d_n]b^2_{[n/d_n]}\left(\frac{\varepsilon^2}{2C}\right)\right)\right).$$

Then

$$\sum_{n=1}^{\infty}\mathbb{P}(|f_n(x) - Ef_n(x)| > \varepsilon)$$

$$\leq \sum_{n=1}^{\infty}\left(\exp\left(-\left([n/d_n]b^2_{[n/d_n]}\left(\frac{\varepsilon^2}{2C}\right)\right)\right)\right) < \infty. \tag{6}$$

On the other hand,

$$Ef_n(x) - f(x, \theta_0) = \int_{\mathbb{R}} K(u)(f(x - b_n u, \theta_0) - f(x, \theta_0))du,$$

and for each $x \in \mathbb{R}$, $K(u)|f(x - b_n u, \theta_0) - f(x, \theta_0)| \leq K(u)$. By the continuity of the density and by the dominated convergence theorem, we conclude that $Ef_n(x) - f(x, \theta_0) \to 0$ as $n \to +\infty$ for each $x \in \mathbb{R}$.

Then, using (6), we conclude that for all $x \in \mathbb{R}$, $f_n(x)$ almost surely (a.s.) converges to $f(x, \theta_0)$ and

$$\mathbb{P}\left(\lim_{n \to +\infty} f_n^{1/2}(x) = f^{1/2}(x, \theta_0), \forall x \in \mathbb{R}\right) = 1.$$

Furthermore, since $\int_{\mathbb{R}} f_n(x)dx = \int_{\mathbb{R}} f(x, \theta_0)dx = 1$, hence

$$\lim_{n \to +\infty}\left(\int_{\mathbb{R}} |f_n^{1/2}(x) - f^{1/2}(x, \theta_0)|^2 dx\right)^{1/2} = 0 \text{ a.s.}$$

Therefore $f_n \to f$ a.s. when $n \to +\infty$ in the Hellinger topology and

$$\hat{\theta}_n = T(f_n) \to T(f(., \theta_0)) = \theta_0 \quad \text{a.s. as } n \to +\infty.$$

This completes the proof of Theorem 1 □

Proof of Theorem 2. From Theorem 2 in Beran [2], we deduce that

$$\frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)}(\hat{\theta}_n - \theta_0)$$

$$= \frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \rho(x, \theta_0)(f_n^{1/2}(x) - f^{1/2}(x, \theta_0))dx$$

$$+ V_n \int_{\mathbb{R}} \frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \dot{S}(x, \theta_0)(f_n^{1/2}(x) - f^{1/2}(x, \theta_0))dx,$$

where V_n is a $(q \times q)$-matrix whose components tends to zero in probability as $n \to +\infty$.

For $b \geq 0,\ a > 0,$ we have the algebraic identity

$$b^{1/2} - a^{1/2} = (b-a)/(2a^{1/2}) - (b-a)^2 / \left[2a^{1/2}(b^{1/2} + a^{1/2})^2\right].$$

Thus

$$\int_{\mathbb{R}} \rho(x,\ \theta_0)(f_n^{1/2}(x) - f^{1/2}(x,\ \theta_0))dx$$

$$= \int_{\mathbb{R}} \Psi(x,\ \theta_0)(f_n(x) - f(x,\ \theta_0))dx - \int_{\mathbb{R}} \frac{\Psi(x,\ \theta_0)(f_n(x) - f(x,\ \theta_0))^2}{\left(f_n^{1/2}(x) + f^{1/2}(x,\ \theta_0)\right)^2}\, dx,$$

where

$$\Psi(x,\ \theta) = \rho(x,\ \theta) / 2f^{1/2}(x,\ \theta).$$

Let

$$H(n,\ \theta_0) = \left| \int_{\mathbb{R}} \frac{\Psi(x,\ \theta_0)(f_n(x) - f(x,\ \theta_0))^2}{(f_n^{1/2}(x) + f^{1/2}(x,\ \theta_0))^2}\, dx \right|,$$

and

$$A(n,\ \theta_0) = \int_{\mathbb{R}} \left| \frac{\rho(x,\ \theta_0)}{f^{3/2}(x,\ \theta_0)} \right| (Ef_n(x) - f(x,\ \theta_0))^2 dx,$$

we have

$$\frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} H(n,\ \theta_0)$$

$$\leq \frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \left| \frac{\rho(x,\ \theta_0)}{f^{3/2}(x,\ \theta_0)} \right| (f_n(x) - Ef_n(x))^2 dx + \frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} A(n,\ \theta_0).$$

We use the same approach as in the proof of Theorem 1, it follows that for $\varepsilon > 0,$

$$\mathbb{P}(|f_n(x) - Ef_n(x)|^2 \geq \varepsilon)$$

$$\leq \sum_{i=1}^{d_n} \mathbb{P}\left(\frac{1}{[n/d_n]^2 b_{[n/d_n]}^2} |Z_{[n/d_n]}(y, Y)|^2 \geq \varepsilon \right)$$

$$= \sum_{i=1}^{d_n} \mathbb{P}\left(|Z_{[n/d_n](y,Y)}| \geq [n/d_n] b_{[n/d_n]} \varepsilon^{1/2} \right).$$

Hence there is $n_2(\varepsilon)$, such that for all $n \geq n_2(\varepsilon)$

$$\mathbb{P}(|f_n(x) - Ef_n(x)|^2 \geq \varepsilon) \leq \exp\left(- [n/d_n] b_{[n/d_n]}^2 \left(\frac{\varepsilon}{2C} \right) \right).$$

Then

$$\sum_{n=1}^{\infty} n^{m\alpha(\theta_0)/2} \mathbb{P}\left(|f_n(x) - Ef_n(x)|^2 > \varepsilon \right)$$

$$\leq \sum_{n=1}^{\infty} n^{m\alpha(\theta_0)/2} \left(\exp\left(- [n/d_n] b_{[n/d_n]}^2 \left(\frac{\varepsilon}{2C} \right) \right) \right) < \infty. \tag{7}$$

Using the dominated convergence theorem, we conclude that

$$\frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \left| \frac{\rho(x, \theta_0)}{f^{3/2}(x, \theta_0)} \right| (f_n(x) - Ef_n(x))^2 dx \to 0 \text{ a.s. as } n \to +\infty.$$

By assumptions (A2), (A3), and (B1), Taylor's formula in one variable gives for x such as $|u^* - x| < |b_n u|$,

$$Ef_n(x) - f(x, \theta_0) = \int_{-\infty}^{+\infty} K(u)(f(x - b_n u, \theta_0) - f(x, \theta_0)) du$$

$$= b_n f(x, \theta_0) \int_{-\infty}^{+\infty} uK(u) du$$

$$+ \frac{b_n^2}{2} f^{(2)}(u^*, \theta_0) \int_{-\infty}^{+\infty} u^2 K(u) du.$$

We get,

$$|Ef_n(x) - f(x,\ \theta_0)| \ \leq \frac{b_n^2}{2} \sup_x |f^{(2)}(u^*,\ \theta_0)| \int_{-\infty}^{+\infty} |u|^2 K(u) du < +\infty,$$

and we deduce

$$\lim_{n\to+\infty} n^{m\alpha(\theta_0)/2} (Ef_n(x) - f(x,\ \theta_0))^2 = \lim_{n\to+\infty} n^{m\alpha(\theta_0)/2} b_n^4 \tau^2(\theta_0) = 0, \quad (8)$$

where

$$\tau^2(\theta_0) = \sup_x \left(\frac{|f^{(2)}(u^*,\ \theta_0)|}{2} \right) \int_{-\infty}^{+\infty} |u|^2 K(u) du < +\infty.$$

Consequently, by (7) and (8) and by the dominated convergence theorem,

$$\frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \left| \frac{\rho(x,\ \theta_0)}{f^{3/2}(x,\ \theta_0)} \right| (f_n(x) - f(x,\ \theta_0))^2 dx \to 0 \text{ as } n \to +\infty.$$

Thus,

$$\frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} (\hat{\theta}_n - \theta_0),$$

and

$$\frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \Psi(x,\ \theta_0)(f_n(x) - f(x,\ \theta_0)) dx$$

have the same asymptotic distribution.

Denote by

$$\frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \Psi(x,\ \theta_0)(f_n(x) - f(x,\ \theta_0)) dx = B(n,\ \theta_0) + C(n,\ \theta_0),$$

where

$$B(n,\ \theta_0) = \frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \Psi(x,\ \theta_0)(f_n(x) - Ef_n(x)) dx,$$

and

$$C(n, \theta_0) = \frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \Psi(x, \theta_0)(Ef_n(x) - f(x, \theta_0))dx.$$

Assumptions (A2), (A3), (B1) and the dominated convergence theorem implies,

$$\frac{n^{m\alpha(\theta_0)/2}}{\mathcal{L}^{m/2}(n)} \int_{\mathbb{R}} \Psi(x, \theta_0)(Ef_n(x) - f(x, \theta_0))dx \to 0 \text{ as } n \to +\infty.$$

In the other hand,

$$B(n, \theta_0) \sim nd_{m,n}^{-1} \int_{\mathbb{R}} \Psi(x, \theta_0)(f_n(x) - Ef_n(x))dx$$

$$= \int_{\mathbb{R}} \Psi(x, \theta_0)\left(nd_{m,n}^{-1}(f_n(x) - Ef_n(x)) - J'_m(x)Y_{mn}\right)dx$$

$$+ Y_{mn} \int_{\mathbb{R}} \Psi(x, \theta_0)J'_m(x)dx.$$

By Csörgő and Mielniczuk [3],

$$\int_{\mathbb{R}} \Psi(x, \theta_0)\left(nd_{m,n}^{-1}(f_n(x) - Ef_n(x)) - J'_m(x)Y_{mn}\right)dx \to^P 0,$$

and

$$Y_{mn} \int_{\mathbb{R}} \Psi(x, \theta_0)J'_m(x)dx$$

$$\to^{\mathcal{D}} (Z_m(1)/m!) \int_{-\infty}^{+\infty} \Psi(x, \theta_0)J'_m(x)dx,$$

where $(Z_m(t))_{0 \le t \le 1}$ is the Hermite process. Hence,

$$nd_{m,n}^{-1}(\hat{\theta}_n - \theta_0) \to^{\mathcal{D}} (Z_m(1)/m!) \int_{-\infty}^{+\infty} \Psi(x, \theta_0)J'_m(x)dx.$$

This completes the proof of Theorem 2. □

4. Simulations and Performances

In this section, we investigate the finite sample properties of the MHD estimator. For this purpose, we consider the univariate Fractional Gaussian Noise (FGN) process. For $0 < H < 1$, FGN is a mean stationary Gaussian process with auto-covariance sequence

$$\gamma(k) = E(X_i X_{i+k}) = \frac{1}{2}\sigma^2(H)((k+1)^{2H} - 2(k)^{2H} + (k-1)^{2H}).$$

Here H is known as the Hurst parameter, and $\sigma^2 H = \Gamma(2-2H)/(\Gamma(3/2-H))^2$ the innovation variance. For $1/2 < H < 1$, the long-memory behaviour of the process is demonstrated by seeing that

$$\gamma(k) \sim \sigma^2(H)H(2H-1)k^{2H-2},\ k \to +\infty.$$

We simulate the scale $\sigma(H)$ of the process. The kernel density estimator f_n is constructed by using Epanechnikov kernel and the bandwidth $b_n = n^{-2/9}$. Simulations are based on 1000 observations of the (FGN) process with 400 replications. We use in R, "longmemo" package and "MDEstimator", "MLEstimator" functions in "distrMod" package.

Table 1 shows the consistency property of the MHD estimator. Table 2 gives the Maximum Likelihood (ML) estimate. It shows that the ML estimator is less consistent than the MHD estimator.

Table 1. Minimum Hellinger distance estimate

H	$\sigma(H)$	$\widehat{\sigma}(H)$	Bias
0.55	1.002222	1.146019	– 0.143797
0.6	1.009700	1.114895	– 0.105195
0.65	1.024125	1.174136	– 0.150011
0.7	1.048182	1.158587	– 0.110405
0.75	1.086435	1.109490	– 0.023055
0.8	1.147369	1.176692	– 0.029323
0.85	1.249004	1.113428	0.135576
0.9	1.438784	1.141883	0.296901

Table 2. Maximum likelihood estimate

H	$\sigma(H)$	$\widehat{\sigma}(H)$	Bias
0.55	1.002222	0.484194	0.518028
0.6	1.009700	0.558653	0.451047
0.65	1.024125	0.557807	0.466318
0.7	1.048182	0.567699	0.480483
0.75	1.086435	0.509382	0.5892615
0.8	1.147369	0.514231	0.633138
0.85	1.249004	0.362574	0.88643
0.9	1.438784	0.311558	1.127226

In the following, we examine the property of the quantitative robustness of the MDH estimator. We proceed as follows:

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

and $\delta_{[0,1]}(.)$ the uniform density on $[0; 1]$. We make α vary between 0 and 1 and we consider the estimators of the associated MDH. In each case, $f_n(.)$ is replaced by $f_{n,\alpha}(.)$. The results of Table 3 seem to indicate a robustness of the MDH estimator.

Table 3. Illustration by simulations of the robustness of the MDHE

α	$\widehat{\sigma}\,(H=0.55)$	$\widehat{\sigma}\,(H=0.6)$	$\widehat{\sigma}\,(H=0.7)$	$\widehat{\sigma}\,(H=0.8)$
0.1	1.0352495	1.06008865	1.060344	1.064473
0.2	1.009288	0.9025901	1.006003	0.9899518
0.3	1.085068	0.9943619	0.998759	0.9947557
0.4	1.092368	0.9915725	0.997589	0.893098
0.5	0.92085	0.89118905	0.896811	0.903899
0.6	0.9587001	0.9039753	0.902914	0.862098
0.7	0.916077	0.8767624	0.876066	0.842138
0.8	0.837741	0.835155	0.831019	0.810639
0.9	0.803023	0.808303	0.795467	0.793061

5. Conclusion

At the end of this study, we obtained the almost sure convergence and asymptotic distribution properties of the minimum Hellinger distance estimator of stationary Gaussian univariate processes with long-range dependence. We worked directly on the sample size of the process. We used disjoint blocks technical. Indeed, we have divided the function $K(X) - EK(X)$ into disjoint blocks. The resulting joint density function has been decomposed into Hermite based orthogonal polynomials and then, we used the Bernstein inequality to obtain the convergence results. This paper improve N'dri and Hili [10] to the extent that they worked with a power of the sample size. The kernel function they used is a Parzen kernel of order $l \geq 4$ in reality. Such a kernel is obtained by construction for simulations.

References

[1] J. Beran, Statistics for Long-Memory Process, Chapman & Hall, 1994.

[2] R. Beran, Minimum Hellinger distance estimates for parametric models, Ann. Statist. 5(3) (1977), 445-463.

DOI: http://dx.doi.org/10.1214/aos/1176343842

[3] S. Csörgő and J. Mielniczuk, Density estimation under long-range dependence, Ann. Statist. 23(3) (1995), 990-999.

DOI: http://dx.doi.org/10.1214/aos/1176324632

[4] L. De Giovanni and M. Naldi, Identification of long-range dependence in telecommunication networks, Statistica Applicata-Italian Journal of Applied Statistics 12(1) (2000), 31-58.

[5] C. W. J. Granger, The typical spectral shape of an economic variable, Econometrica 34(1) (1966), 150-161.

DOI: http://dx.doi.org/10.2307/1909859

[6] C. W. J. Granger and R. Joyeux, An introduction to long-memory time series models and fractional differencing, Journal of Time Series Analysis 1(1) (1980), 15-29.

DOI: http://dx.doi.org/10.1111/j.1467-9892.1980.tb00297.x

[7] H. C. Ho, A note on the exponential bounds for sequences of long-range dependence, Soochow Journal of Mathematics 20 (1994), 595-602.

[8] J. R. M. Hosking, Fractional differencing, Biometrika 68(1) (1981), 165-176.

DOI: https://doi.org/10.1093/biomet/68.1.165

[9] W. Leland, M. Taqqu, W. Willinger and D. Wilson, On the self-similar nature of Ethernet traffic, IEEE/ACM Transactions on Networking 2(1) (1994), 1-15.

DOI: https://doi.org/10.1109/90.282603

[10] A. N'dri and O. Hili, Estimation par la distance de Hellinger des processus gaussiens stationnaires fortement dependants, Comptes Rendus Mathématique 349(17-18) (2011), 991-994.

DOI: https://doi.org/10.1016/j.crma.2011.07.026

[11] A. N'dri and O. Hili, Hellinger distance estimation of strongly dependent multi-dimensional Gaussian processes, International Journal of Statistics and Probability 2(3) (2013), 70-84.

DOI: http://dx.doi.org/10.5539/ijsp.v2n3p70

[12] M. S. Taqqu, Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long-range dependence, Z. Wahrscheinlichkeitstheorie verw Gebiete 40(3) (1977), 203-238.

DOI: https://doi.org/10.1007/BF00736047

[13] G. Samorodnitsky and M. Taqqu, Stable Non-Gaussian Processes, Chapman & Hall, 1994.

[14] W. Willinger, M. S. Taqqu, W. E. Leland and D. V. Wilson, Self-similarity in high-speed packet traffic: Analysis and modeling of Ethernet traffic measurements, Statistical Science 10(1) (1995), 67-85.

■