International Journal of Statistical Sciences *ISSN 1683–5603*
Vol. 9(Special Issue), 2009, pp 151-162

Comparing Two Nonparametric Regression Curves with Long Memory Errors under Random Design

Fang Li
Department of Mathematical Sciences
Indiana University Purdue University Indianapolis
Indianapolis, IN 46202, USA
E-mail: fli@math.iupui.edu

[Received December 31, 2008; Revised February 17, 2009; Accepted February 21, 2009]

Abstract

This paper discusses the problem of testing the equality of two nonparametric regression functions against two-sided alternatives for random design on $[0, 1]$ with long memory moving average errors. The standard deviations are possibly different for the two errors. The paper applied the marked empirical processes to construct the tests and derives their asymptotic null distributions. The paper also shows that these tests are consistent for general alternatives.

Keywords and Phrases: Marked Empirical Process, Long Memory Process, Nonparametric Regression, Random Design.

AMS Classification: 62M10; 62F03.

1 Introduction

This paper is concerned with testing the equality of two regression functions against the two-sided alternatives when the errors form long-memory moving averages under a random design. More precisely, let μ_1 and μ_2 be real valued continuous functions on $[0, 1]$, σ_1, σ_2 be positive numbers. Suppose $\frac{1}{2} < H < 1$ and let

$$\alpha_j = 0, \quad \text{for } j < 0, \qquad \alpha_0 = 1, \qquad \alpha_j : \;=\; j^{\frac{2H-3}{2}} \quad \text{for } j \geq 1. \tag{1.1}$$

In the problem of interest, one observes two stochastic processes $Y_{1,i}$ and $Y_{2,i},\ i = \{0, 1, \cdots\}$ such that

$$Y_{1,i} = \mu_1(X_i) + \sigma_1 u_{1,i}, \qquad u_{1,i} = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{1,i-j}, \tag{1.2}$$
$$Y_{2,i} = \mu_2(X_i) + \sigma_2 u_{2,i}, \qquad u_{2,i} = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{2,i-j},$$

where, $\{\varepsilon_{1,i},\ i \in \mathbb{Z} = \{0, \pm 1, \cdots\}\}$ are i.i.d. standard r.v.'s and so are $\{\varepsilon_{2,i},\ i \in \mathbb{Z}\}$; $\{X_i,\ i \in \mathbb{Z}\}$ are i.i.d. r.v.'s on $[0,\ 1]$. Moreover, $\{\varepsilon_{1,i},\ i \in \mathbb{Z}\}$, $\{\varepsilon_{2,i},\ i \in \mathbb{Z}\}$ and $\{X_i,\ i \in \mathbb{Z}\}$ are mutually independent.

Note that

$$\sum_{j=0}^{\infty} \alpha_j = \infty, \qquad \sum_{j=0}^{\infty} \alpha_j^2 < \infty.$$

Hence, the error processes $u_{1,i}$ and $u_{2,i}$ have long-memory.

The problem of interest is to test the null hypothesis:

$$H_0 :\ \mu_1(x) = \mu_2(x), \qquad \forall\, x \in [0, 1],$$

against the two-sided alternative hypothesis

$$H_a :\ \mu_1(x) \neq \mu_2(x), \quad \text{for some } x \in [0, 1]. \tag{1.3}$$

based on the data $(X_i,\ Y_{1,i},\ Y_{2,i})$, $i = 1, \cdots, n$, where n is a positive integer.

The motivation of studying long memory processes is from their important applications in hydrology, economics, finance and various other physical sciences. For example, long memory processes describe well with financial data such as exchange rates, stock returns and inflation rates. Please see Beran (1992, 1994) and Baillie (1996) and the references therein.

In both one-sample and two-sample settings with independent errors, related testing problems have been addressed by several authors. For one-sample setting, see Cox, Koh, Wahba and Yandell (1988), Eubank and Spiegelman (1990), Raz (1990), Härdel and Mammen (1993), Koul and Ni (2004) and monograph of Hart (1997). For two sample setting, see Härdle and Marron (1990), Scheike (2000), Hall, Huber and Speckman (1997) and Koul and Schick (1997, 2003).

There have been some works on fitting a regression function in the presence of long memory, including Cs*ö*rg*ö* and Mielniczuk (1995, 1999, 2000), Robinson (1997), Koul and Stute (1998), Koul, Baillie and Surgailis (2004), Hurvich, Lang and Soulier (2005) and Guo and Koul (2007). Koul and Stute (1998) studied a class of such tests based on partial sum processes of certain residuals when the design is either fixed or random. Koul, Baillie and Surgailis (2004) studied those tests further when the covariate is of dimension one and also forms a long memory moving average process.

The papers that address the above two-sided nonparametric testing problem with independent errors include Hall and Hart (1990), Kulasekera (1995), Delgado (1993) and Ferreira and Stute (2004). Delgado (1993) used the absolute difference of the cumulative regression functions assuming the same covariates in the two samples while the two samples are possibly dependent on each other. Li (2006) extends Delgado's test to the above long-memory set up, but under a fixed design, i.e. $X_i = i/n$. Ferreira and Stute (2004) introduced a marked empirical process to serve as a basic test process for the above two sided testing problem with design variable being stationary and with independent errors. For our model (1.2) under random design and with long memory error, this current paper will adapt the idea of marked empirical process in Ferreira and Stute (2004) to construct our test.

As in Ferreira and Stute (2004), let

$$D_j := Y_{1,j} - Y_{2,j}, \quad j = 1, \cdots, n$$

and define

$$U_n(x) := \sum_{j=1}^{n} D_j I(X_j \le x), \quad x \in [0,1].$$

Hence, U_n is an empirical process marked by D_j.

It can be shown that $\frac{1}{n} U_n(x)$ provides a uniformly consistent estimator of

$$\Delta(t) := \int_0^x \Big(\mu_1(x) - \mu_2(x)\Big)\, dQ(x), \quad \forall\, 0 \le t \le 1, \tag{1.4}$$

where Q is the distribution function (d.f.) of the design variable, assumed to be continuous. This suggests to base tests of H_0 on some suitable functions of this process, for example, the Kolmogorov-Smirnov and Cramér-von Mises type tests. In this paper we shall focus on the Kolmogorov-Smirnov type tests based on $\sup_{0\le t\le 1} |U_n(t)|$.

To determine the large sample distribution of the process $U_n(t)$, one needs to normalize this process suitably. As in Li (2006), the normalizing sequence depends on the parameters H and σ_i^2.

To make this more precise, let

$$c^2(H) = \frac{1}{H(2H-1)} \int_0^\infty (y + y^2)^{-\frac{1+\theta}{2}}\, dy, \quad \theta := 2 - 2H, \tag{1.5}$$

and

$$\tau_{n,i}^2 = \sigma_i^2 c^2(H) n^{2H}, \quad \tau_n = (\tau_{n,1}^2 + \tau_{n,2}^2)^{\frac{1}{2}}. \tag{1.6}$$

Now define

$$T: \;=\; \sup_{0\le x\le 1} \Big|\frac{1}{\tau_n} U_n(x)\Big|. \tag{1.7}$$

In the case H and σ_i's are known, the tests of H_0 could be based on T, being significant for its large value. But, usually these parameters are unknown which renders T of little use. This suggests to replace the parameters in T by their estimates. Therefore, the proposed tests will be based on the adaptive versions of T, namely

$$\hat{T} := \sup_{0\le x\le 1}\Big|\frac{1}{\hat{\tau}_n}U_n(x)\Big|,\ \hat{\tau}_n = (\hat{\tau}_{n,1}^2 + \hat{\tau}_{n,2}^2)^{\frac{1}{2}},\ \hat{\tau}_{n,i}^2 = c^2(\hat{H})\hat{\sigma}_i^2 n^{2\hat{H}},\ i = 1, 2, \tag{1.8}$$

where, $\hat{\sigma}_i,\ (i = 1, 2)$ and $\hat{H}$ are the estimates of σ_i and H.

The estimates of these parameters might be necessarily based on the residuals $Y_{i,j} - \hat{\mu}_i(X_j),\ i = 1, 2; j = 1, \cdots, n$, where $\hat{\mu}_i,\ i = 1, 2$, are some estimators of the regression functions $\mu_i,\ i = 1, 2,$. The latter estimation problem in the presence of long memory has been addressed by some authors. See Csörgő and Mielniczuk (1995), Csörgő and Mielniczuk (2000) and Robinson (1997). They all studied the kernel regression function estimators. We shall also use the kernel method to estimate the two nonparametric regression functions in our model when needed.

The estimation of the long memory parameter H are also of interests here. Fox and Taqqu (1986) and Dahlhaus (1989) studied the MLE and Whittle estimators. Robinson (1995a) discussed a form of log-periodogram regression estimator under the condition of Gaussianity. Robinson (1995b) considered another estimator that maximizes an approximate form of frequency domain Gaussian likelihood in a semiparametric setting. Its consistency and asymptotic normality is obtained with a rate less than $n^{1/2}$. In Robinson (1997), this consistency rate is shown to be $\log n$ under some mild conditions. Li (2006) adapted Robinson (1997)'s estimator and here we shall again use this estimation method to estimate H under our model.

We shall study the asymptotic behaviors of $\hat{T}$ as the sample size n tend to infinity. Theorem 2.1 of section 2 shows that under H_0, T weakly converge to the absolute value of standard normal. Then in Corollary 2.1, under a general set of assumptions on the estimates σ and H, we derived the same asymptotic distribution of $\hat{T}$ under H_0. Remark 2.1 proves that the test based on $\hat{T}$ is consistent, at the fixed alternative (1.3). In section 3, under some additional conditions, appropriate estimates $\hat{\sigma}_1,\ \hat{\sigma}_2$ and $\hat{H}$ are constructed.

2 Asymptotic behavior of T and $\hat{T}$

This section investigates the asymptotic behaviors of T given in (1.7) and the adaptive statistic $\hat{T}$ given in (1.8) under the null hypothesis and the alternatives (1.3). We write P for the underline probability measures and E for the corresponding expectations.

First give the following assumption:

(A.1) For $i = 1, 2$, $E\varepsilon_{i,1}^4 < \infty$, where $\varepsilon_{i,1}$ is as in model (1.2).

For $x \in [0, 1]$, define

$$U_{n,i}(x) := \sum_{j=1}^{n} \sigma_i u_{i,j} I(X_j \le x),$$
$$U^1_{n,i}(x) := \sum_{j=1}^{n} [\sigma_i u_{i,j}(I(X_j \le x) - Q(x))], \quad U^2_{n,i} := \sum_{j=1}^{n} \sigma_i u_{i,j}.$$

Recall that Q is the distribution function of the design variable X_j. It is easy to see that $U_{n,i}(x) = U^1_{n,i}(x) + Q(x)U^2_{n,i}$. We are now ready to state the following lemmas:

Lemma 2.1. *Under model (1.2) and Assumption (A.1), we have*

$$n^{-1/2} \sup_{x \in [0,1]} |U^1_{n,i}(x)| = O_p(1). \tag{2.1}$$

Proof: The proof appears in Koul and Stute (1998). Hence it is omitted here. □

The proof of the next lemma is implied by the invariance principle of long memory linear processes, which is proved in Davydov (1970), Taqqu (1975), Avram and Taqqu (1987), Sowel (1990), among others.

Lemma 2.2. *The long memory linear processes $u_{i,j}$, $i = 1, 2$ of model (1.2) satisfy:*

$$\frac{1}{\tau_{n,i}} \sum_{j=1}^{n} \sigma_i u_{i,j} = \frac{1}{\tau_{n,i}} U^2_{n,i} \Longrightarrow Z, \tag{2.2}$$

where Z represents standard normal.

Next, let

$$T_1(x) = \frac{1}{\tau_n} (U_{n,1}(x) - U_{n,2}(x)), \quad 0 \le x \le 1. \tag{2.3}$$

It is easy to see that $T = \sup_{x \in [0,1]} |T_1(x)|$ under the null hypothesis. Now, we are ready to give the first main result.

Theorem 2.1. *Under model (1.2) and assumption (A.1), we have*

$$T_1(x) \overset{D[0,1]}{\Longrightarrow} Q(x)Z, \qquad 0 \le x \le 1, \tag{2.4}$$

in $D[0, 1]$. $\overset{D[0,1]}{\Longrightarrow}$ stands for the weak convergence of random elements with values in the Skorohod space $D[0, 1]$, with respect to the uniform metric. Consequently, under the null hypothesis, T of (1.7) satisfy

$$T \Longrightarrow |Z|$$

Proof: It suffices to prove (2.4) since $T = \sup_{x\in[0,1]} |T_1(x)|$ under the null hypothesis. Now, by decomposition,

$$\frac{1}{\tau_{n,i}} U_{n,i}(x) = \frac{n^{1/2}}{\tau_{n,i}} n^{-1/2} U^1_{n,i}(x) + Q(x)\tau^{-1}_{n,i} U^2_{n,i}, \quad i = 1, 2.$$

This, Lemma 2.1 and 2.2, and the fact that $\frac{n^{1/2}}{\tau_{n,i}} \to 0$ give

$$\frac{1}{\tau_{n,i}} U_{n,i}(x) \overset{D[0,1]}{\Longrightarrow} Q(x)Z, \qquad \text{in } D[0,1], \quad i = 1, 2$$

This, $\frac{\tau_{n,i}}{\tau_n} \to \sigma_i/\sqrt{\sigma_1^2 + \sigma_2^2}$ for $i = 1, 2$ and the fact that $\{u_{1,j}\}$'s are independent of $\{u_{2,j}\}$'s implied (2.4) and hence completes the proof of the theorem. □

Next, we need the following additional assumptions to obtain the asymptotic distribution of $\hat{T}$ given in (1.8).

Assumption 2.1 Let the estimators $\hat{H}$, $\hat{\sigma}^2$ of H and $\sigma^2 = \sigma_1^2 + \sigma_2^2$ be such that under null hypotheses,

$$(\log n)(\hat{H} - H) \to_P 0, \qquad \hat{\sigma}^2 - \sigma^2 \to_P 0, \quad n \to \infty. \tag{2.5}$$

Corollary 2.1. Suppose that the Assumption 2.1 holds. Then under the model (1.2) and the null hypothesis,

$$\hat{T} \Longrightarrow |Z|.$$

Proof: By Theorem 2.1, it suffices to prove

$$\frac{\hat{\tau}_n^2}{\tau_n^2} \to_P 1. \tag{2.6}$$

By the definition of $c(H)$ and simple calculation, (2.5) implies that

$$c(\hat{H})^2 - c(H)^2 \to_P 0, \quad \text{and} \quad \frac{c(\hat{H})^2\hat{\sigma}^2}{c(H)^2\sigma^2} \to_P 1,$$

This together with the first part of (2.5) gives

$$\frac{\hat{\tau}_n^2}{\tau_n^2} = \frac{c(\hat{H})^2\hat{\sigma}^2 n^{2\hat{H}}}{c(H)^2\sigma^2 n^{2H}} = \frac{c(\hat{H})^2\hat{\sigma}^2}{c(H)^2\sigma^2} \exp\{2\log(n)(\hat{H} - H)\} \to_P 1,$$

which implies (2.6). This corollary is proved. □

Remark 2.1. **Testing property of $\hat{T}$.** Under the model (1.2), consider the following alternative that is the same as in (1.3):

$$H_a: \quad \mu_1(x) - \mu_2(x) = \delta(x) \neq 0 \quad \text{for some } x \in [0, 1],$$

where δ is continuous on [0,1] since $\mu_1,\ \mu_2$ are continuous. Theorem 2.1 and its corollary suggest to reject the null hypothesis for large values of $\hat{T}$ given in (1.8).

First, suppose Assumption 2.1 is satisfied. Let

$$\hat{T}(x) = \frac{1}{\hat{\tau}_n} U_n(x), \qquad \hat{T}_1(x) = \frac{1}{\hat{\tau}_n} (U_{n,1}(x) - U_{n,2}(x)).$$

Then,

$$\hat{T}(x) = \hat{T}_1(x) + h(x), \qquad h(x) = \frac{1}{\hat{\tau}_n} \sum_{j=1}^{[nx]} \delta(\frac{j}{n})$$

By (1.8) and the fact that $\frac{1}{n} \sum_{j=1}^{[nx]} \delta(\frac{j}{n}) \to \int_0^x \delta(t)\, dt$ uniformly,

$$\hat{T}(x) = \frac{\tau_n}{\hat{\tau}_n} T(x), \quad h(x) = \frac{n}{\hat{\tau}_n} \frac{1}{n} \sum_{j=1}^{[nx]} \delta(\frac{j}{n}) \sim \frac{n}{\hat{\tau}_n} \int_0^x \delta(t)\, dt = O_P(n^{1-\hat{H}}). \tag{2.7}$$

By (1.6), (1.8), (2.3) and Theorem 2.1,

$$\sup_{x\in[0,1]} |\hat{T}_1(x)| = \frac{\tau_n}{\hat{\tau}_n} \sup_{x\in[0,1]} |T_1(x)| = O_P(n^{H-\hat{H}}) = o_P(n^{1-\hat{H}})$$

Hence,

$$\hat{T} = \sup_{0\le x\le 1} |\hat{T}(x)| = \sup_{0\le x\le 1} |\hat{T}_1(x) + h(x)| \to_P \infty. \tag{2.8}$$

Let z_α be the $(1-\alpha) * 100$ percentile of standard normal. By Corollary (2.1),

$$\lim_{n\to\infty} P(\hat{T} > z_{\alpha/2}) = \alpha \text{ under } H_0 \quad \text{and} \quad \lim_{n\to\infty} P(\hat{T} > z_{\alpha/2}) = 1 \text{ under } H_a.$$

Therefore, the test based on $\hat{T}$ is consistent for H_a.

3 Construction of $\hat{H}$ and $\hat{\sigma}^2$

In this section, under some additional conditions, we shall now construct estimates of $H,\ \sigma^2 = \sigma_1^2 + \sigma_2^2$ that satisfy Assumption 2.1. Now consider the following conditions:

(A.2) The regression functions $\mu_1,\ \mu_2$ are Lipschitz-continuous on $[0,1]$.

(A.3) For some $1 > \Delta_2 \ge H \ge \Delta_1 > 1/2$ and for some $\delta_0 > 0$, as $n \to \infty$,

$$(\log n)^2 (\frac{m}{n})^\beta + \frac{n^{2-2H}}{m^{1-2\max(\delta_0,\, H-\Delta_1)}} \to 0, \qquad \beta = H - \frac{1}{2} > 0.$$

First, we will construct the estimates of H that satisfy Assumption 2.1. The following estimator is analogues of the estimator defined at (4.8)-(4.10) in Robinson (1997).

For $k = 1, \cdots, m \in [1, \frac{n}{2})$, Let $\lambda_k = 2\pi k/n$ and

$$\tilde{G}(h) = \frac{1}{m}\sum_{k=1}^{m}\lambda_k^{2h-1}I_{DD}(\lambda_k),\ \tilde{R}(h) = \log\tilde{G}(h) - (2h-1)\frac{1}{m}\sum_{k=1}^{m}\log\lambda_k, \tag{3.1}$$

$$I_{DD}(\lambda) = w_D(\lambda)w_D(-\lambda), \qquad w_D(\lambda) = \frac{1}{(2\pi n)^{\frac{1}{2}}}\sum_{t=1}^{n}D_t e^{\mathbf{i}t\lambda}. \tag{3.2}$$

Recall, $e^{\mathbf{i}x}$ always represent the complex value $\cos x + \mathbf{i}\sin x$, $\mathbf{i} = \sqrt{-1}$. Define

$$\hat{H} := \arg\min_{h\in[\Delta_1,\Delta_2]}\tilde{R}(h). \tag{3.3}$$

The following lemma shows the consistency of the estimator.

Lemma 3.1. *Suppose the model (1.2), and the Assumptions (A.1) - (A.3) hold. Then, the estimates $\hat{H}$ given in (3.3) satisfy Assumption 2.1, i.e.*

$$(\log n)(\hat{H} - H) \to_P 0, \quad \textit{under } H_0. \tag{3.4}$$

Proof: First, under H_0, $D_i = \sigma_1 u_{1,i} - \sigma_2 u_{2,i} = \sum_{j=0}^{\infty}\sigma_1\alpha_j\varepsilon_{1,i-j} + \sigma_2\alpha_j\varepsilon_{2,i-j}$. Let $\varepsilon_{i-j} = \frac{\sigma_1\varepsilon_{1,i-j}+\sigma_2\varepsilon_{2,i-j}}{\sqrt{\sigma_1^2+\sigma_2^2}}$. Then, $\{\varepsilon_j\}$ are i.i.d. standard r.v.'s and the process $D_i = \sigma\sum_{j=0}^{\infty}\alpha_j\varepsilon_{i-j}$ is another long memory moving average process. By Theorem 4 in Robinson (1997), to prove (3.4), it suffices to verify the assumptions of his theorem for our model. By a careful comparison of our assumptions to that of Robinson (1997) and by a close inspection of the proof in Robinson (1997), it suffices to prove the following two assumptions. The labels here are correspondent to that in Robinson's paper.

Assumption 3.1. In a neighborhood $(0, \delta)$ of the origin, $\alpha(\lambda) = \sum_{j=-\infty}^{\infty}\alpha_j e^{\mathbf{i}j\lambda}$ is differentiable and $(d/d\lambda)\alpha(\lambda) = O(|\alpha(\lambda)|/\lambda)$, as $\lambda \to 0^+$.

Assumption 3.2. For $H \in [\Delta_1, \Delta_2]$, there exist some $\beta \in (0, 2]$ and $G > 0$,

$$f(\lambda) \sim G\lambda^{1-2H}\left(1 + O(\lambda^\beta)\right) \quad \text{as } \lambda \to 0^+, \qquad f(\lambda) = |\alpha(\lambda)|^2/2\pi.$$

In view of (1.2) and (2.3.11) of Zygmund (1968, page 70), Assumption 3.2 is satisfied with $\beta = 2H - 1$, while Li (2006) has shown that Assumption 3.1 is also satisfied here. Hence we proved the lemma. □

Next, we are to give the estimator of $\sigma^2 = \sigma_1^2 + \sigma_2^2$ as below:

$$\hat{\sigma}^2 = \frac{1}{2n\hat{F}}\sum_{j=1}^{n-1}(D_{j+1} - D_j)^2, \qquad F = \sum_{j=0}^{\infty}\alpha_j(\alpha_j - \alpha_{j+1}), \tag{3.5}$$

where, $\hat{F}$ is F given above with H in the expression of α_j's replaced by its estimator $\hat{H}$ given in (3.3). This estimating method is adapted from Delgado (1993). The following lemma shows its consistency under H_0.

Lemma 3.2. *Suppose the model (1.2) and the assumptions (A.1)-(A.3) hold. Then the sequence of estimators $\hat{\sigma}^2$ given in (3.5) satisfies the Assumption 2.1, i.e.*

$$\hat{\sigma}^2 \to_P \sigma^2 := \sigma_1^2 + \sigma_2^2, \quad \textit{under } H_0. \tag{3.6}$$

To prove Lemma 3.2, we need the following two lemmas.

Lemma 3.3. *The coefficients $\{\alpha_j\}$ of (1.1) satisfy*

$$\sum_{j=0}^{\infty} \alpha_j \alpha_{j+|t-s|} \sim H(2H-1)c^2(H)|t-s|^{2H-2}, \quad |t-s| \to \infty.$$

Proof: By the Karamata Theorem

$$\begin{aligned} & \sum_{j=0}^{\infty} \alpha_j \alpha_{j+|t-s|} \\ = \;& |t-s|^{2H-2} \sum_{j=1}^{\infty} \left(\frac{j}{|t-s|} \left(\frac{j}{|t-s|} + 1 \right) \right)^{-\frac{3-2H}{2}} \frac{1}{|t-s|} + \alpha_{i,|t-s|} \\ \sim \;& H(2H-1)c^2(H)\,|t-s|^{-\theta_i}, \qquad |t-s| \to \infty, \qquad i = 1,2. \end{aligned} \tag{3.7}$$

Lemma is proved. □

Lemma 3.4. *The long memory linear processes $u_{i,j}$, $i = 1, 2$ of model (1.2) satisfy*

$$\frac{1}{n} \sum_{j=1}^{n} u_{i,j}^2 \to_P A \quad A = \sum_{j=0}^{\infty} \alpha_j^2, \qquad i = 1,2. \tag{3.8}$$

Proof: The proof appears in Li (2006), page 631-632. Hence it is omitted here. □

We are now ready to present the

Proof of Lemma 3.2: Since $H, \hat{H} < \infty$, by the consistency of $\hat{H}$, guaranteed by Lemma 3.1, and the continuity of $\hat{F}$ in $\hat{H}$, we obtain that $\hat{F} \to_P F$. Thus, by decomposition, to prove (3.6), it suffices to prove the following results:

$$\frac{1}{n} \sum_{j=1}^{n-1} (u_{i,j+1} - u_{i,j})^2 \to_P F, \qquad i = 1,2, \tag{3.9}$$

$$\frac{1}{n} \sum_{j=1}^{n-1} (u_{1,j+1} - u_{1,j})(u_{2,j+1} - u_{2,j}) \to_P 0. \tag{3.10}$$

First, to prove of (3.10), it suffices to prove

$$\frac{1}{n} \sum_{j=1}^{n} u_{1,j} u_{2,j} \to_P 0. \tag{3.11}$$

By Lemma 3.3,

$$\begin{aligned} E(u_{1,j}u_{2,j}u_{1,k}u_{2,k}) &= E(u_{1,j}u_{1,k})E(u_{2,j}u_{2,k}) = \left(\sum_{i=0}^{\infty}\alpha_i\alpha_{i+|j-k|}\right)^2 \\ &\sim H^2(2H-1)^2c(H)^4|j-k|^{4H-4}. \end{aligned}$$

Hence,

$$\begin{aligned} &E\left(\frac{1}{n}\sum_{j=1}^{n}u_{1,j}u_{2,j}\right)^2 \\ \sim\ & \frac{1}{n^2}\sum_{j=1}^{n}\sum_{j\neq k=1}^{n}H^2(2H-1)^2c^4(H)|j-k|^{4H-4}+\frac{1}{n^2}\sum_{j=1}^{n}V(u_{1,j})V(u_{2,j}) \\ \leq\ & H^2(2H-1)^2c^4(H)\frac{n^{2H-2}}{n^2}\sum_{j=1}^{n}\sum_{j\neq k=1}^{n}\left|\frac{j}{n}-\frac{k}{n}\right|^{2H-2}+\frac{A^2}{n} \\ =\ & O\left(n^{2H-2}\int_0^1\int_0^1|x-y|^{2H-2}\,dxdy\right) \\ \to\ & 0, \end{aligned}$$

which implies (3.11) by the Chebyshev inequality, and hence proves (3.10).

Finally, consider (3.9) for $i=1$. Let

$$\tilde{u}_{1,j+1}=\sum_{k=0}^{\infty}\tilde{\alpha}_k\varepsilon_{1,j+1-k}, \qquad \tilde{\alpha}_k=\frac{1}{\sqrt{2}}\left(\alpha_{1,k}-\alpha_{1,k-1}\right)$$

Then the left hand side of (3.9) can be rewritten as

$$\frac{1}{n}\sum_{j=1}^{n-1}\tilde{u}_{1,j+1}^2\to_P\sum_{k=0}^{\infty}\tilde{\alpha}_k^2=F,$$

by Lemma 3.4. This proves (3.9), and hence completes the proof of the lemma. □

Acknowledgments

I am thankful to the Editor for his time on processing this paper.

References

[1] Baillie, R. T. (1996). Long memory processes and fractional integration in economics. *Journal of Econometrics*, **73**, 5-59.

[2] Beran, J. (1992). Statistical methods for data with long-range dependence (with discussion). *Statistical Science*, **7**, 404-427.

[3] Beran, J. (1994). *Statistics for Long-Memory Processes*. Monographs on Statistics and Applied probability, **61**, Chapman and Hall.

[4] Cox, D, Koh, E., Wahba, G. and Yandell, B. S. (1988). Testing the (parametric) null Model Hypothesis in (semiparametric) partial and generalized spline models. *Ann. Statist.*, **16**, 113-119.

[5] Csörgő, S. and Mielniczuk J. (1995). Nonparametric regression under long-range dependent normal errors. *Ann. Statist.*, **23**, 1000-1014.

[6] Csörgő, S. and Mielniczuk, J. (1999). Random-design regression under long-range dependent errors. *Bernoulli*, 5, 209-224.

[7] Csörgő, S. and Mielniczuk J. (2000). The smoothing dichotomy random-design regression with long-memory errors based on moving averages. *Statistica Sinica*, **10**, 771-787.

[8] Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.*, **17**, 1749-1766.

[9] Delgado, M. A. (1993). Testing the equality of nonparametric regression curves. *Statist. & Probab. Letters*, **17**, 199-204.

[10] Eubank, R. L. and Spigelman, C. H. (1990). Testing the goodness-of-fit of a linear model via nonparametric regression techniques. *J. Amer. Statist. Assoc.*, **85**, 387-392.

[11] Ferreira, E. and Stute, W. (2004). Testing for difference between conditional means in a time series context, *Journal of the American Statistical Association*, 99, 169-174.

[12] Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary gaussian time series. *Ann. Statist.*, **14**, 517-532.

[13] Guo, H. and Koul, H. L. (2007). Nonparametric regression with heteroscedastic long memory errors, *Journal of Statistical Planning & Inference*, 137/2, 379-404.

[14] Hall, P. and Hart, J. D. (1990). Bootstrap test for difference between means in nonparametric regression. *J. Amer. Statist. Assoc.*, **85**, 1039-1049.

[15] Hall, P., Huber, C. and Speckman, P. L. (1997). Covariate-matched one-sided tests for the difference between functional means. *J. Amer. Statist. Assoc.*, **92**, 1074-1083.

[16] Härdle, W. and Marron, J. S. (1990). Semiparametric comparisons of regression curves. *Ann. Statist.*, **18**, 63-89.

[17] Hart, Jeffrey D. (1997). *Nonparametric smoothing and lack-of-fit tests.* Springer Series in Statistics. Springer-Verlag, New York.

[18] Hurvich, C., Lang, G. and Soulier, P. (2005). Estimation of long memory in the presence of a smooth nonparametric trend, *Journal of the American Statistical Association*,100(471), 853-871.

[19] Koul, H. L., Baillie, R. T. and Surgailis D. (2004). Regression Model fitting with a long memory covariate process. *Econometric Theory*, **20**, 485-512.

[20] Koul, H. L. and Ni, P. (2004). Minimum distance regression model checking. *J. Statist. Plann. & Inference*, **119(1)** 109-142.

[21] Koul, H. L. and Stute, W. (1998). Regression model fitting with long memory errors. *J. Statist. Plann. & Inference*, **71**, 35-56.

[22] Kulasekera, K. B. (1995). Comparison of regression curves using quasi-residuals. *J. Amer. Statist. Assoc.*, **90**, 1085-1093.

[23] Li, F. (2006). Testing for the equality of two nonparametric regression curves with long memory errors, *Communications in Statistics: Simulation & Computation*, 35, #3, 621-643.

[24] Raz, J. (1990). Testing for no effect when estimating a smooth function by nonparametric regression: a randomization approach. *J. Amer. Statist. Assoc.*, **85**, 132-138.

[25] Robinson P. M. (1995a). Log-periodogram regression of time series with long range dependence. *Ann. Statist.*, **23**, 1048-1072.

[26] Robinson P. M. (1995b). Gaussian semiparametric estimation of long range dependence. *Ann. Statist.*, **23**, 1630-1661.

[27] Robinson P. M. (1997). Large-sample inference for nonparametric regression with dependent errors. *Ann. Statist.*, **25**, 2054-2083.

[28] Scheike, T. H. (2000). Comparison of non-parametric regression functions through their cumulatives. *Statist. & Probab. Letters*, **46**, 21-32.

[29] Zygmund, A. (1968). *Trigonometric Series*. Cambridge Univ. Press.