Residual Empirical Processes for Nearly Unstable Long-memory Time Series

LIU, Weiwei

A Thesis Submitted in Partial Fulfilment
of the Requirements for the Degree of
Doctor of Philosophy
in
Statistics

Supervised by

Prof. Ngai Hang CHAN

June, 2009

中文摘要

本文研究了漸進不穩定長相依時間序列的剩余經驗過程。在 Ho 和 Hsing（1997）的文章中提出了基于經驗過程的 K-S 統計量用于檢驗一列具有穩定分布的長相依隨機變量序列的邊際分布。基于這一理論，Chan 和 Ling（2007）研究了自回歸時間序列的剩余經驗過程并提出了基于其的 K-S 統計量。在他們的研究中表明，當不穩定自回歸時間序列的特征多項式具有單位根時，由 Ho 和 Hsing 提出的基于經驗過程的 K-S 統計量的極限分布不同于由 Chan 和 Ling 提出的基于剩余經驗過程的 K-S 統計量的極限分布。本文研究的主要目的是通過研究漸進不穩定長相依時間序列的剩余經驗過程提出一個基于此剩余經驗過程的 K-S 統計量用來檢驗長相依更新序列的邊際分布。本文的研究表明，由 Chan 和 Ling 提出的 K-S 統計量可以應用到漸進不穩定長相依時間序列模型當中，并且此統計量的極限分布不同于 Ho 和 Hsing 提出的 K-S 統計量的極限分布，它的極限分布可以表示為由分數布朗運動生成的 Ornstein-Uhlenbeck 過程的函數。

This is the cover page.

Abstract of thesis entitled:

Residual Empirical Processes for Nearly Unstable Long-memory Time Series

Submitted by LIU, Weiwei

for the degree of Doctor of Philosophy

at The Chinese University of Hong Kong in June, 2009

The first part of this thesis considers the residual empirical process of a nearly unstable long-memory time series. Chan and Ling [8] showed that the usual limit distribution of the Kolmogorov-Smirnov test statistics does not hold when the characteristic polynomial of the unstable autoregressive model has a unit root. A key question of interest is what happens when this model has a near unit root, that is, when it is nearly non-stationary. In this thesis, it is established that the statistics proposed by Chan and Ling can be extended. The limit distribution is expressed as a functional of an Orenstein-Uhlenbeck process that is driven by a fractional Brownian motion. This result extends and generalizes Chan and Ling's results to a nearly non-stationary long-memory time series.

The second part of the thesis investigates the weak convergence of weighted sums of random variables that are functionals of moving average processes. A non-central limit theorem is established in which the Wiener integrals with respect to the Hermite processes appear as the limit. As an application of the non-central limit theorem, we examine the asymptotic theory of least squares estimators (LSE) for a nearly unstable AR(1) model when the innovation sequences are functionals of moving average processes. It is shown that the limit distribution of the LSE appears as functionals of the Ornstein-Uhlenbeck processes driven by Hermite processes.

Acknowledgement

This thesis was written during my time as a Ph.D. student in the Department of Statistics. First of all, I would like to thank my supervisor, Prof. Chan Ngai Hang, for sharing with me his knowledge of long-memory time series and limit theory and for introducing me to the topics investigated herein. He has always had confidence in me and given me his support and continuous encouragement. I am also grateful to having so many nice colleagues in the statistics department. I have a very good time working with them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Empirical Processes

Let $\{\epsilon_t\}_{t\geq 1}$ be a sequence of identically distributed random variables with a common distribution, $F(x) = P(\epsilon_1 \leq x)$. Define the empirical distribution function by

$$F_n(x) = \frac{1}{n}\sum_{t=1}^{n} 1(\epsilon_t \leq x). \tag{1.1}$$

By virtue of the Glivenko-Cantelli theorem, $D_n = \sup_x |F_n(x) - F(x)| \to 0$, a.s. Quantity D_n can be used to construct statistics to test the hypothesis of F. These statistics are usually known as the Kolmogorov-Smirnov statistics. It is important to understand the asymptotic properties of D_n, particularly the rate of D_n decaying to zero. When $\{\epsilon_t\}_{t\geq 1}$ are i.i.d. random variables, the empirical process is defined by

$$Y_n(x) = \sqrt{n}\left(F_n(x) - F(x)\right). \tag{1.2}$$

Komlós, Major and Tusnády [31] proved that $Y_n(x)$ converges weakly to a Brownian bridge.

An interesting problem is what happens to $Y_n(x)$ when $\{\epsilon_t\}_{t\geq 1}$ is strongly dependent. The properties of $Y_n(x)$ can be very different depending on the structure of the underlying process $\{\epsilon_t\}$.

This thesis is concerned with a situation in which $\{\epsilon_t\}_{t\geq 1}$ is a linear moving average process, of which the long-memory process constitutes an important example. A moving average process is defined by

$$\epsilon_t = \sum_{i=0}^{\infty} a_i e_{t-i}, \tag{1.3}$$

where $\{e_i, i \in \mathcal{Z}\}$ are i.i.d. random variables with zero mean and finite variance, and the coefficients $\{a_i, i \geq 0\}$ are square-summable. A lot of work have been conducted on the case in which $\{a_i\}$ takes the specific form $a_i = i^{H-3/2}L(i)$, $H \in (0,1)$, and $L(x)$ is a slowly varying function. When $H \in (1/2, 1)(H < 1/2)$, the process $\{\epsilon_t\}$ exhibits a long-memory (short-memory) phenomenon. This thesis focuses on the results of the long-memory case.

When $\{\epsilon_t\}_{t\geq 1}$ is a long-memory Gaussian process, weak convergence of the empirical process can be derived through the Hermite expansion approach; see Dehling and Taqqu [18], Koul and Surgails [34]. The important paper published by Ho and Hsing [27] derived an asymptotic expansion of the empirical process (1.3) and developed a new approach to the study of the non-linear functionals of non-Gaussian moving averages. Specifically, let

$$K_n(x) = \frac{1}{\sigma_n} \sum_{t=1}^{n} \left[1\left(\epsilon_t \leq x\right) - F(x)\right] \tag{1.4}$$

be the empirical process, where $\sigma_n^2 = \operatorname{Var}\left(\sum_{t=1}^{n} \epsilon_t\right)$. Ho and Hsing proved that

$$\sup_x |K_n(x) + \frac{1}{\sigma_n} F'(x) \sum_{t=1}^{n} \epsilon_t| = o(1) \ \text{ a.s.} \tag{1.5}$$

$$\sigma_n^2 \sim c(H)^2 n^{2H} L^2(n) \ \text{ and } \ \sigma_n^{-1} \sum_{t=1}^{n} \epsilon_t \rightarrow_{\mathcal{L}} N(0,1), \tag{1.6}$$

where

$$c(H) = \{H(2H-1) \times [\int_0^\infty (x+x^2)^{H-3/2}\,dx]^{-1}\}^{1/2}. \quad (1.7)$$

Herein, $a_n \sim b_n$ means that $a_n/b_n \to 1$, as $n \to \infty$, and $\to_{\mathcal{L}}$ denotes convergence in distribution, as $n \to \infty$. By (1.5),

$$|\sup_x F'(x)|^{-1} \sup_x |K_n(x)| \to_{\mathcal{L}} |N(0,1)|. \quad (1.8)$$

If $\sup_x |F'(x)| < \infty$, then this becomes the Kolmogorow-Smirnov test statistic used by Ho and Hsing [27] to test the distribution $F(x)$.

In other words, the asymptotic distribution of the empirical process of long-memory moving averages changes according to the dependence structure of $\{\epsilon_t\}$. If $\{\epsilon_t\}$ is weakly dependent, then $n^{1/2}(F_n(x) - F(x))$ converges in distribution to a Brownian bridge. When $\{\epsilon_t\}$ exhibits the long-memory phenomenon, $F_n(x) - F(x)$ is of order n^θ, $0 < \theta < 1/2$, and the weak limit of $n^\theta(F_n(x) - F(x))$ exists.

1.2 Residual Empirical Processes

Empirical processes have many applications in testing the specific distribution of the underlying random variables encountered, including income and wealth distribution, the distributions of asset returns and the distribution of profit and loss in risk management.

Let the time series $\{y_t\}$ be generated by the model

$$y_t = \rho' X_t + \epsilon_t \text{ and } \epsilon_t = \sum_{i=0}^{\infty} a_i e_{t-i}, \quad (1.9)$$

where $\{X_t\}$ is a sequence of p-dimensional time series that are measurable with respect to $\mathcal{F}_{t-1} = \sigma\{\epsilon_{t-1}, \epsilon_{t-2}, \ldots\}$ and $\{\epsilon_t\}$ is a long-memory moving average process defined by (1.3). By reference to the results of

Ho and Hsing [27], we have

$$\sup_x |K_n(x) + \frac{1}{\sigma_n} F'(x) \sum_{t=1}^{n} \epsilon_t| = o(1) \quad \text{a.s.} \tag{1.10}$$

where K_n is defined in (1.4).

In model (1.9), as ϵ_t is unobservable, the statistics have to be evaluated based on an estimator of ϵ_t. Under such circumstances, a key issue of interest is to determine the validity of (1.10) for the Kolmogorov-Smirnov statistics when $\{\epsilon_t\}$ is replaced by its corresponding estimated residual process. Furthermore, when (1.10) becomes invalid, how can one test for the distribution of $\{\epsilon_t\}$? These two issues have been investigated extensively for when $\{\epsilon_t\}$ is i.i.d. or exhibits a short-memory phenomenon. Ling [38] established the weak convergence of the residual empirical process for nonstationary autoregressive models and showed that this process converges weakly to a Kiefer process when the characteristic polynomial does not include the unit root 1; otherwise, it converges weakly to a Kiefer process plus a functional of the stochastic integrals of a standard Brownian motion. Bai [3] investigated the weak convergence of the residual process for ARMA(p, q) models. For both model (1.9) and for the Kolmogorov-Smirnov statistic considered in Ho and Hsing [27], these two important issues were investigated by Chan and Ling [10]. In their paper, a uniform expansion of the residual empirical process of $\{\epsilon_t\}$ was established under a general framework. They showed that the test statistic (1.4) considered by Ho and Hsing [27] is no longer valid when the characteristic polynomial of the unstable AR model has a unit root. Furthermore, Chan and Ling [8] propose a new statistic to test the distribution of long-memory noises. Inspired by their paper, a key question in this study becomes what happens when the model has an approximate unit root. Does the statistic proposed by Chan and Ling still hold? If the answer is affirmative, then what kind of approximation should be used for it?

The remainder of this thesis is organized as follows. In Chapter 2, it is established that the statistic proposed by Chan and Ling can be extended to test for the distribution of long-memory noises. The limit distribution is expressed as a functional of an Ornstein-Unlenbeck process driven by a fractional Brownian motion. This result extends and generalizes the results of Chan and Ling to a nearly non-stationary long-memory time series. In Chapter 3, we investigate the weak convergence of weighted sums of random variables that are functionals of moving average processes. To this end, a non-central limit theorem is established, in which the Wiener integrals with respect to Hermite processes appear as the limit. In Chapter 4, we investigate the asymptotic theory of the least squares estimator (LSE) for a nearly unstable AR(1) model when the innovation sequences are functionals of moving average processes. It is shown that the limit distribution of the LSE is expressed as a functional of the Ornstein-Uhlenbeck processes driven by Hermite processes. Chapter 5 concludes the thesis.

□ **End of chapter.**

Chapter 2

Residual Empirical Processes

Let the time series $\{y_{t,n}\}$ be generated by the model

$$y_{t,n} = \rho_n y_{t-1,n} + \epsilon_t \quad \text{and} \quad \epsilon_t = \sum_{i=0}^{\infty} a_i e_{t-i}, \tag{2.1}$$

where $\rho_n = 1 - \gamma/n$ and γ is a real number. The coefficients a_i satisfy $\sum_{i=1}^{\infty} a_i^2 < \infty$; $a_0 = 1$ and $a_k = k^{H-3/2}L_0(k)$ for some slowly varying function L_0, with $H < 1$. $\{e_t\}$ is a sequence of i.i.d. mean zero random variables with $\sigma_e^2 = \mathrm{E}e_t^2 < \infty$. The process $\{\epsilon_t\}$ exhibits a long-memory (short-memory) phenomenon when $H \in (1/2, 1)$ $(H < 1/2)$. Here, we focus on the case $H \in (1/2, 1)$.

This chapter is organized as follows. In Section 2.1, we examine the expansion of the empirical residual processes investigated by Chan and Ling [8]. It is shown that this expansion can be extended to model (2.1) under similar assumptions. In Section 2.2, it is shown that model (2.1) satisfies these assumptions. To this end, a new statistic is proposed to test the distribution of the long-memory noises for the nearly unstable model. In section 2.3, a number of simulations are presented to assess the finite sample behavior of the limit distribution established in the previous section.

2.1 Expansion of Empirical Residual Processes

Let $\hat{\rho}_n$ be the least squares estimator of ρ_n in model (2.1). Let $\hat{\epsilon}_{t,n} = y_{t,n} - \hat{\rho}_n y_{t,n}$ be the residual of model (2.1). Further, define the empirical process based on residuals $\{\hat{\epsilon}_{t,n}\}$ by

$$\hat{K}_n(x) = \frac{1}{\sigma_n} \sum_{t=1}^{n} \left[1(\hat{\epsilon}_{t,n} \le x) - F(x)\right], \tag{2.2}$$

where $\sigma_n^2 \sim c(H,k)^2 n^{2H} L^2(n)$ and $k = 1$

$$\begin{aligned} c(H,k) &= \{k!(1-k(1-H))(1-k(2-2H)) \\ &\quad \times [\int_0^\infty (x+x^2)^{H-3/2}\, dx]^{-k}\}^{1/2}. \end{aligned} \tag{2.3}$$

Let G_0 be the common distribution of $\{e_t\}$. Let $\epsilon_t = e_t + \xi_{t-1}$ and $A_t(x) = G_0'(x - \xi_{t-1}) - \mathrm{E}[G_0'(x - \xi_{t-1})]$, where $\xi_{t-1} = \sum_{i=1}^{\infty} a_i e_{t-i}$.

First, let us introduce the following two assumptions.

Assumption 2.1. *G_0 is three times differentiable with bounded, continuous and integrable derivatives such that $\int x^4\, dG_0(x) < \infty$.*

Assumption 2.2. *Assume that the following statements hold:*

(a) $n(\hat{\rho}_n - \rho_n) = O_p(1)$,

(b) $\sigma_n^{-1} \sum_{t=1}^{n} \mathrm{E} \mid \frac{y_{t-1,n}}{n} \mid = O(1)$,

(c) $\sigma_n^{-1} \sum_{t=1}^{n} \mathrm{E} \left(\frac{y_{t-1,n}}{n}\right)^2 = o(1)$,

(d) $\sigma_n^{-1} \sup_x \mid \sum_{t=1}^{n} A_t(x) \frac{y_{t-1,n}}{n} \mid = o_p(1)$.

Theorem 2.1. *Assume that Assumptions 2.1 and 2.2 hold, then*

$$\sup_x \mid \hat{K}_n(x) - K_n(x) - R_n F'(x) \mid = o_p(1), \tag{2.4}$$

where $R_n = \sigma_n^{-1}(\hat{\rho}_n - \rho_n) \sum_{t=1}^{n} y_{t-1,n} = O_p(1)$.

Remark 2.1. *Assumptions 2.1 and 2.2 are identical to those given in Chan and Ling [10]. Assumption 2.2 (a) is satisfied by the weak convergence of least squares estimator, see Buchmann and Chan [8].*

Proof. Let $\hat{u}_n = n(\hat{\rho}_n - \rho_n)$, multiply $y_{t-1,n}$ on both sides of this equation, we have

$$\begin{aligned}\hat{u}_n y_{t-1,n} &= n(y_{t-1,n}\hat{\rho}_n - y_{t-1,n}\rho_n) \\ &= n(y_{t,n} - y_{t-1,n}\rho_n - (y_{t,n} - y_{t-1,n}\hat{\rho}_n)) \\ &= n(\epsilon_t - \hat{\epsilon}_t)\end{aligned}$$

so that $\hat{\epsilon}_t = \epsilon_t - \frac{1}{n}\hat{u}_n y_{t-1,n}$. Therefore,

$$\hat{K}_n(x) - K(x) - R_n F'(x) =$$
$$\frac{1}{\sigma_n}\sum_{t=1}^{n}[1(\epsilon_t \le x + \frac{1}{n}\hat{u}_n y_{t-1,n}) - 1(\epsilon_t \le x) - \frac{1}{n}\hat{u}_n y_{t-1,n}F'(x)].$$

To study the preceding process, consider the following process

$$\begin{aligned}A_n(x,u) = \quad & \frac{1}{\sigma_n}\sum_{t=1}^{n}[1(\epsilon_t \le x + \frac{1}{n}u y_{t-1,n}) \\ & -1(\epsilon_t \le x) - \frac{1}{n}u y_{t-1,n}F'(x)]\end{aligned}$$

for all $u \in R$ and $x \in R$. By Assumption 2.2 (a), if we can show that

$$\sup_{u\in[-\triangle,\triangle]}\ \sup_x \mid A_n(x,u) \mid = o_p(1) \tag{2.5}$$

for every $\triangle \in (0,\infty)$, we then add and subtract $F(x)$, $F(x+\frac{1}{n}u\sum_{t=1}^{n} y_{t-1,n})$ to $A_u(x,u)$ and split it into two parts

$$A_n(x,u) = Z_n(x,u) + H_n(x,u). \tag{2.6}$$

By the triangular inequality, we have

$$\mid A_n(x,u) \mid \le \mid Z_n(x,u) \mid + \mid H_n(x,u) \mid, \tag{2.7}$$

where

$$\begin{aligned} Z_n(x,u) = & \frac{1}{\sigma_n}\sum_{t=1}^{n}[1(\epsilon_t \le x + \frac{1}{n}u\sum_{t=1}^{n} y_{t-1,n}) \\ & -F(x+\frac{1}{n}u\sum_{t=1}^{n} y_{t-1,n}) - 1(\epsilon_t \le x) + F(x)] \end{aligned}$$

and

$$\begin{aligned} H_n(x,u) = & \frac{1}{\sigma_n}\sum_{t=1}^{n}[F(x+\frac{1}{n}u\sum_{t=1}^{n} y_{t-1,n}) \\ & -F(x) - \frac{1}{n}F'(x)u\sum_{t=1}^{n} y_{t-1,n}]. \end{aligned}$$

Since $\sup_x \mid G_0''(x) \mid < \infty$, we have $\sup_x \mid F''(x) \mid < \infty$. By Assumption 2.1 and Taylor expansion, it suffices to show that

$$\sup_x \mid H_n(x,u) \mid = o_p(1) \qquad \text{for all } u \in R. \tag{2.8}$$

Now we need to proof that the following equation holds

$$\sup_{u\in[-\triangle,\triangle]} \sup_x \mid Z_n(x,u) \mid = o_p(1). \tag{2.9}$$

Let $g_t(u,\lambda) = \frac{y_{t-1,n}}{n}u + \lambda|\frac{y_{t-1,n}}{n}|$. Partition $[-\triangle,\triangle]$ into m parts $\{l_1,\ldots,l_m\}$ each interval with length less than δ. Take one point in each l_r and denote it by u_r. Then for any $u \in l_r$, we have

$$\mid g_t(u,\lambda) - g_t(u_r,\lambda) \mid \le \delta|\frac{y_{t-1,n}}{n}|. \tag{2.10}$$

Thus, $g_t(u_r,\lambda-\delta) \le g_t(u,\lambda) \le g_t(u_r,\lambda+\delta)$.

$$
\begin{aligned}
Z_n(x,u) &= \frac{1}{\sigma_n}\sum_{t=1}^{n}[1(\epsilon_t \le x + u\frac{y_{t-1,n}}{n}) - F(x + u\frac{y_{t-1,n}}{n}) \\
&\quad -1(\epsilon_t \le x) + F(x)] \\
&= \frac{1}{\sigma_n}\sum_{t=1}^{n}[1(\epsilon_t \le x + g_t(u,0)) - F(x + g_t(u,0)) \\
&\quad -1(\epsilon_t \le x) + F(x)] \\
&\le \frac{1}{\sigma_n}\sum_{t=1}^{n}[1(\epsilon_t \le x + g_t(u_r,\delta)) - F(x + g_t(u_r,\delta)) \\
&\quad -1(\epsilon_t \le x) + F(x)] \\
&\quad +[\frac{1}{\sigma_n}\sum_{t=1}^{n} F(x + g_t(u_r,\delta)) - \frac{1}{\sigma_n}\sum_{t=1}^{n} F(x + g_t(u,0))] \\
&= \tilde{Z}_n(x,u_r,\delta) \\
&\quad +\frac{1}{\sigma_n}\sum_{t=1}^{n}[F(x + g_t(u_r,\delta)) - F(x + g_t(u,0))],
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{Z}_n(x,u_r,\delta) = \quad & \frac{1}{\sigma_n}\sum_{t=1}^{n}[1(\epsilon_t \le x + g_t(u_r,\delta)) - F(x + g_t(u_r,\delta)) \\
& -1(\epsilon_t \le x) + F(x)].
\end{aligned}
$$

Using the same argument, we have

$$
\begin{aligned}
Z_n(x,u) \ \ge \ & \tilde{Z}_n(x,u_r,\delta) \\
& +\frac{1}{\sigma_n}\sum_{t=1}^{n}[F(x + g_t(u_r,-\delta)) - F(x + g_t(u,0))].
\end{aligned}
$$

Thus, for any $\epsilon > 0$ and $\eta > 0$, if we can prove that

$$
\begin{aligned}
P \ \{ \ & \frac{1}{\sigma_n}\max_r \max_{u\in l_r} \sup_x |\sum_{t=1}^{n} F(x + g_t(u_r,\pm\delta)) \\
& -F(x + g_t(u,0))| \ge \frac{\epsilon}{3}\} \le \frac{\eta}{6} \qquad (2.11)
\end{aligned}
$$

and

$$
P\{\max_r \sup_x \mid \tilde{Z}_n(x,u_r,\pm\delta) \mid \ge \frac{\epsilon}{3}\} \le \frac{\eta}{3}, \qquad (2.12)
$$

then

$$\begin{aligned} P \quad & \{ \sup_{u\in[-\triangle,\triangle]} \sup_x \mid Z_n(x,u) \mid \geq \epsilon \} \\ \leq \quad & P\{\max_r \sup_x \mid \tilde{Z}_n(x,u_r,\delta) \mid \geq \frac{\epsilon}{3}\} \\ & +P\{\max_r \sup_x \mid \tilde{Z}_n(x,u_r,\delta) \mid \geq \frac{\epsilon}{3}\} \\ & +P\{\frac{1}{\sigma_n} \max_r \max_{u\in l_r} \sup_x \mid \sum_{t=1}^n [F(x+g_t(u_r,\pm\delta)) \\ & -F(x+g_t(u,0))] \mid \geq \frac{\epsilon}{3}\} \\ \leq \quad & \eta, \end{aligned}$$

when $n \to \infty$. We omit the proofs of (2.11) and (2.12) because they routinely follow from the arguments in Chan and Ling [8]. □

2.2 Kolmogorov-Smirnov Statistic

2.2.1 The Hermite Process

To prepare our analysis in this section, we first introduce the notion of a Hermite process, $Z_H^k = \left(Z_H^k(t)\right)_{t\in R}$, of order $k \geq 1, k \in Z$ with Hurst parameter $H \in (1/2, 1)$. This stochastic process is defined as a multiple Wiener-Itô integral of order k with respect to the standard Brownian motion, $(B(t))_{t\in R}$, as follows.

$$Z_H^k(t) = c(H,k) \int_{R^k} \int_0^t (\prod_{j=1}^k (s-y_i)^{(\frac{1}{2}+\frac{1-H}{k})}) \, ds \, dB(y_1) \cdots dB(y_k), \tag{2.13}$$

where $x_+ = \max(x, 0)$ and

$$\begin{aligned} c(H,k) \quad = \quad & \{k!(1-k(1-H))(1-k(2-2H)) \\ & \times [\int_0^\infty (x+x^2)^{H-3/2} \, dx]^{-k}\}^{1/2}, \end{aligned} \tag{2.14}$$

making $\mathrm{E}[(Z_H^k(t))^2] = 1$. When $k = 1$, this process corresponds to the fractional Brownian motion with Hurst parameter $H \in (0, 1)$. Let $R(t, s)$ denote the covariance function of $Z_H^k(t)$, that is, $R(t, s) = \mathrm{E}[Z_H^k(t)Z_H^k(s)]$. Maejima and Tudor [39] showed that

$$R(t, s) = \frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}). \tag{2.15}$$

Consider the Langevin type stochastic differential equation,

$$X_t = -\lambda \int_0^t X_s \, ds + \sigma Z_H^k(t), \tag{2.16}$$

where $\sigma, \lambda > 0$, the Hurst index $H \in (1/2, 1)$ with $k \geq 1$ is an integer, and Hermite process Z_H^k is represented as a driving noise. For $k = 1$, that is, $Z_H^k = B_H(t)$, B_H is a fractional Brownian motion. The process $\{X_t\}$ was considered in Cheridito, Kawaguchi and Maejima [14], and the unique solution of the stochastic equation (2.16) is

$$Z_{H,\gamma}^1 = \sigma \int_\infty^t e^{-\lambda(t-u)} \, dB_H(u).$$

The process, $Z_{H,\gamma}^1$, is called a fractional Ornstein-Uhlenbeck process. Maejima and Tudor [39] extended this result to the case of $k \geq 1$, that is, the solution of equation (2.16) is given by

$$Z_{H,\gamma}^k = \sigma \int_{-\infty}^t e^{-\lambda(t-u)} \, dZ_H^k(t), \tag{2.17}$$

where $Z_{H,\gamma}^k$ is the Ornstein-Uhlenbeck process driven by the Hermite process. It is shown by Maejima and Tudor [39], for all $t > a$, the integral $\int_a^t e^{\lambda u} \, dZ_H^k(u, \omega)$ exists in the Riemann-Stieltjes sense and it is equal to

$$e^{\lambda t} Z_H^k(t, \omega) - e^{\lambda a} Z_H^k(a, \omega) - \lambda \int_a^t Z_H^k(u, \omega) e^{iu} \, du. \tag{2.18}$$

Therefore, equation (2.17) can be re-written as

$$Z_{H,\gamma}^k(t) = \sigma Z_H^k(t) - \lambda\sigma \int_\infty^t e^{-\lambda(t-u)} Z_H^k(t) \, du. \tag{2.19}$$

Let $C[0,1]$ be the collection of all continuous functions defined on $[0,1]$ and define the metric $d(f,g) = \sup_{0\le t\le 1}|f(t)-g(t)|$ for $f,g\in C[0,1]$. Let $\Rightarrow$ denote the weak convergence of ξ_n to ξ in the space $C[0,1]$ under the metric d, if for any bounded function $f\in C[0,1]$, that is, $\lim_{n\to\infty}\mathrm{E}f(\xi_n)=\mathrm{E}f(\xi)$.

Let $y_{t,n}$ be generated by the model (2.1), then we have the following theorem.

Theorem 2.2. *For $H\in(1/2,1)$,*

$$\{\frac{y_{[nu]}}{\sigma_n}, 0\le u\le 1\}\Rightarrow\{Z^1_{H,\gamma}(u),\quad 0\le u\le 1\},$$

where Z^1_H denotes the fractional Brownian motion and $Z^1_{H,\gamma}$ is the Ornstein-Uhlenbeck process.

Remark 2.2. *A general theorem for $k\ge 1$ is given in Theorem 4.1.*

Proof. Let $f_u(s)=e^{-\gamma(u-s)}$ and $f_{n,u}(s)=(1-\frac{\gamma}{n})^{[nu]-[ns]}$. Obviously, $f_{n,u}(s)\to f_u(s)$.

$$\begin{aligned}\frac{y_{[nu]}}{\sigma_n} &= \sum_{j=1}^{[nu]} f_{n,u}(\frac{j}{n})(\epsilon_j/\sigma_n)-\sum_{j=1}^{[nu]} f_u(\frac{j}{n})(\epsilon_j/\sigma_n)\\ &\quad+\sum_{j=1}^{[nu]} f_u(\frac{j}{n})(\epsilon_j/\sigma_n)\\ &= M_{1,n}+M_{2,n},\end{aligned} \tag{2.20}$$

where

$$M_{1,n}=\sum_{j=1}^{[nu]} f_{n,u}(\frac{j}{n})(\epsilon_j/\sigma_n)-\sum_{j=1}^{[nu]} f_u(\frac{j}{n})(\epsilon_j/\sigma_n) \tag{2.21}$$

and

$$M_{2,n}=\sum_{j=1}^{[nu]} f_u(\frac{j}{n})(\epsilon_j/\sigma_n). \tag{2.22}$$

We now show that $|M_{1n}| = o_p(1)$. Let $d_n = (n/\gamma)\log(1-\gamma/n)$. For fixed γ, $d_n \to -1$ as $n \to \infty$. For any s, u $(0 \le s \le u \le 1)$, we have

$$\begin{aligned} |f_{n,u}(s) - f_u(s)| &= |(1-\frac{\gamma}{n})^{[nu]-[ns]} - e^{-\gamma(u-s)}| \\ &= |e^{(\frac{[nu]}{n} - \frac{[ns]}{n})\gamma d_n} - e^{-\gamma(u-s)}| \\ &= |e^{-\gamma(u-s)} e^{O(\frac{1}{n})} - e^{-\gamma(u-s)}| \\ &= O(1/n). \end{aligned} \tag{2.23}$$

Let $T_n = \sup_{s,u} |f_{n,u}(s) - f_u(s)|$, we have $T_n = O(\frac{1}{n})$.

$$\begin{aligned} \mathrm{E}|M_{1n}| &= \mathrm{E}|\sum_{j=1}^{[nu]} f_{n,u}(\frac{j}{n})(\epsilon_j/\sigma_n) - \sum_{j=1}^{[nu]} f_u(\frac{j}{n})(\epsilon_j/\sigma_n)| \\ &\le CT_n[\sum_j \mathrm{E}|Z_H^1(\frac{j}{n}) - Z_H^1(\frac{j-1}{n})|]. \end{aligned}$$

By the Hölder continuity of the sample path, for $p \ge 1$ and $H > 1$, we have

$$\mathrm{E}|Z_H^1(\frac{j}{n}) - Z_H^1(\frac{j-1}{n})|^p < Cn^{-H}, \tag{2.24}$$

where C is a constant. Combining (2.23) and (2.24), we have $M_{1n} = o_p(1)$. For M_{2n}, one sees that $f_u(s) = e^{-\gamma(u-s)}1_{[0,u]}$ belongs to $L^1(\mathcal{R}) \cap L^2(\mathcal{R})$ and satisfies conditions in Theorem 3.1, thus we have

$$\sum_{j=1}^{[nu]} f_u(\frac{j}{n})(\epsilon_j/\sigma_n) \to_{\mathcal{L}} \int_0^u e^{-\gamma(u-s)}\, dZ_H^1(s). \tag{2.25}$$

Thus the finite-dimensional convergence of $y_{[nu]}/\sigma_n$ to $Z_{H,\gamma}^1(u)$ holds. It remains to prove the tightness. By Theorem 12.3 of Billingsley [6], it suffices to show that for all m there exist $C < \infty$ and $\alpha > 1$ such that

$$\frac{\mathrm{E}(y_{m,n})^2}{\sigma_n^2} \le C\frac{m^\alpha}{n^\alpha}. \tag{2.26}$$

For any $H \in (1/2, 1)$, there exists a positive δ such that $\alpha = 2H - \delta > 1$. Because $\mathrm{E}(y_{m,n})^2 = O(m^{2H}L^{2k}(m))$ and $\sigma_n^2 = O(n^{2H}L^{2k}(u))$. Thus by elementary properties of slowly varying functions, we have

$$\lim_{n\to\infty}\max_{m\le n}\frac{n^\alpha \mathrm{E}(y_{m,n})^2}{m^\alpha\sigma_n^2}=\lim_{n\to\infty}\max_{m\le n}C\frac{n^{\alpha-2H}m^{2H-\alpha}L^{2k}(m)}{L^{2k}(n)}=1, \quad (2.27)$$

which conclude the result. □

Some useful limit distributions are given in the following theorem.

Theorem 2.3. *If $H\in(1/2,1)$, we have as $n\to\infty$,*

(i) $n^{-1}\sum_{t=1}^n y_{t-1,n}^2 \to_{\mathcal{L}} \int_0^1 (Z_{H,\gamma}^1(u))^2\,du$;

(ii) $n^{-1}\sum_{t=1}^n y_{t-1,n} \to_{\mathcal{L}} \int_0^1 Z_{H,\gamma}^1(u)\,du$;

(iii) $\sum_{t=1}^n y_{t-1,n}\epsilon_t/\sigma_n^2 \to_{\mathcal{L}} \frac{1}{2}(Z_{H,\gamma}^1(u))^2+\gamma\int_0^1 (Z_{H,\gamma}^1(u))^2\,du$;

(iv) $\sum_{i=1}^n(\sum_{t=1}^i \frac{\epsilon_t}{\sigma_n})\frac{\epsilon_i\rho_n^{n-i}}{\sigma_n} \to_{\mathcal{L}}$
$B_H(1)Z_{H,\gamma}^1(1)-\frac{1}{2}(Z_{H,\gamma}^1(1))^2-\gamma\int_0^1 (Z_{H,\gamma}^1(u))^2\,du$;

(v) $n(\hat\rho_n-\rho_n)\to_{\mathcal{L}}$
$[\frac{1}{2}(Z_{H,\gamma}^1(1))^2+\gamma\int_0^1 (Z_{H,\gamma}^1(u))^2\,du][\int_0^1 (Z_{H,\gamma}^1(u))^2\,du]^{-1}$.

Remark 2.3. *In (v), asymptotic distributions of the least squares estimator of a nearly unstable AR(1) process with long memory errors are derived.*

Proof. (i) and (ii) follow from continuous mapping theorem together with Theorem 2.2.

To prove (iii), squaring and summing (2.1), we decompose $\sum_{t=1}^n y_{t-1,n}\epsilon_t$ into three terms which we analyze separately, that is, for all n

$$\sum_{t=1}^n y_{t-1,n}\epsilon_t=\frac{n}{2(n-\gamma)}y_{n,n}^2+\frac{\gamma(2n-\gamma)}{2n(n-\gamma)}\sum_{t=1}^n y_{t-1,n}^2-\frac{n}{2(n-\gamma)}\sum_{t=1}^n\epsilon_t^2.$$

Define the auxiliary random variables

$$\begin{aligned}T_{1,n} &= \frac{1}{\sigma_n^2}\sum_{t=1}^n y_{t-1,n}\epsilon_t\\ &= \frac{n}{2(n-\gamma)}T_{2n}+\frac{\gamma(2n-\gamma)}{n-\gamma}(T_{3,n})^2-\frac{n}{2(n-\gamma)}\sum_{t=1}^n\epsilon_t^2,\end{aligned}$$

where

$$T_{2,n} = \frac{y_{n,n}^2}{\sigma_n^2}, \qquad T_{3,n} = \frac{1}{n}\sum_{t=1}^{n}\left(y_{t-1,n}^2/\sigma_n^2\right), \qquad T_{4,n} = \frac{1}{\sigma_n^2}\sum_{t-=1}^{n}\epsilon_t^2.$$

As the function $f \to (f(1), \int_0^t f(s)^2\,ds)$ is a continuous mapping from $D[0,1]$ into R^2, we have the following results

$$\begin{pmatrix} T_{2,n} \\ T_{3,n} \end{pmatrix} \to_{\mathcal{L}} \begin{pmatrix} (Z_{H,\gamma}^1(1))^2 \\ \int_0^1 (Z_{H,\gamma}^1(u))^2\,du \end{pmatrix}.$$

By definition, we have

$$T_{1,n} \to_{\mathcal{L}} \frac{1}{2}(Z_{H,\gamma}^1(1))^2 + \gamma\int_0^1 (Z_{H,\gamma}^1(u))^2\,du.$$

For part (iv), since

$$\begin{aligned}
\sum_{i=1}^{n}\Big(\sum_{t=1}^{i}\frac{\epsilon_t}{\sigma_n}\Big)\frac{\epsilon_i\rho_n^{n-i}}{\sigma_n} &= \sum_{t=1}^{n}\Big(\sum_{i=t+1}^{n}\frac{\epsilon_i\rho_n^{n-i}}{\sigma_n}\Big)\frac{\epsilon_t}{\sigma_n} \\
&= \sum_{t=1}^{n}\Big(\sum_{i=1}^{n}\frac{\epsilon_i\rho_n^{n-i}}{\sigma_n}\Big)\frac{\epsilon_t}{\sigma_n} - \sum_{t=1}^{n}\Big(\sum_{i=1}^{t}\frac{\epsilon_i\rho_n^{n-i}}{\sigma_n}\Big)\frac{\epsilon_t}{\sigma_n} \\
&= \Big(\sum_{t=1}^{n}\frac{\epsilon_t}{\sigma_n}\Big)\Big(\sum_{i=1}^{n}\frac{\epsilon_i\rho_n^{n-i}}{\sigma_n}\Big) - \sum_{t=1}^{n}\frac{y_t\epsilon_t}{\sigma_n^2},
\end{aligned}$$

by Theorem 1 of Wu [62] (see also Corollary 3.3 of Ho and Hsing [28]), we have

$$\sum_{t=1}^{n}\frac{\epsilon_t}{\sigma_n} \to_{\mathcal{L}} B_H(1). \tag{2.28}$$

Also it follows from Theorem 2.1 in Buchmann and Chan [10] that

$$\sum_{i=1}^{n}\frac{\epsilon_i\rho_n^{n-i}}{\sigma_n} \to_{\mathcal{L}} Z_{H,\gamma}^1(1). \tag{2.29}$$

Thus by the continuous mapping theorem together with (2.28), (2.29) and (iii), we have

$$\sum_{i=1}^{n}\Big(\sum_{t=1}^{i}\frac{\epsilon_t}{\sigma_n}\Big)\frac{\epsilon_i\rho_n^{n-i}}{\sigma_n} \to_{\mathcal{L}}$$
$$B_H(1)Z_{H,\gamma}^1(1) - \frac{1}{2}(Z_{H,\gamma}^1(1))^2 - \gamma\int_0^1 (Z_{H,\gamma}^1(u))^2\,du.$$

For (v), by a similar argument in (iii) and (iv) and the fact that

$$n(\hat{\rho}_n - \rho_n) = \frac{\sum_{t=1}^{n} y_{t-1,n}\epsilon_t}{\frac{1}{n}\sum_{t=1}^{n} y_{t-1,n}^2},$$

(v) follows from (i), (ii) and (iii). □

Theorem 2.4. *For* $H \in (1/2, 1)$,

$$[\sup_x F'(x)]^{-1} \sup_x \mid \hat{K}_n(x) \mid$$
$$\rightarrow_{\mathcal{L}} K = |B_H(1) - [\frac{1}{2}(Z_{H,\gamma}^1(1))^2 + \gamma \int_0^1 (Z_{H,\gamma}^1(u))^2\, du]$$
$$\times [\int_0^1 Z_{H,\gamma}^1(u)\, du][\int_0^1 (Z_{H,\gamma}^1(u))^2 du]^{-1}|.$$

Proof. By (i) (ii) and (iv) of Theorem 2.3, Assumption 2.2 (a), (b) and (c) hold. Now we consider Assumption 2.2 (d). The proof follows from the argument of Chan and Ling [8]. First, since G_0' is bounded and by Theorem 2.1 of Buchmann and Chan [8], we have $\mathrm{E}\sup_{|x|>M} A_t^2(x) \to 0$ as $M \to \infty$ and $\max_{1\le t\le n} \sigma_n^{-2}\mathrm{E}y_{t-1,n}^2 = O(1)$. Thus for any given $\epsilon > 0$ and $\eta > 0$, there exists a constant $M > 0$ such that

$$P\left(\sup_{|x|>M} \left|\frac{1}{n\sigma_n}\sum_{t=1}^{n} A_t(x) y_{t-1,n}\right| > \eta\right)$$
$$\le \frac{\sqrt{\mathrm{E}\sup_{|x|>M}|A_t(x)|^2}}{\eta n \sigma_n}\sum_{t=1}^{n}\sqrt{\mathrm{E}|y_{t-1,n}|^2} < \epsilon, \qquad (2.30)$$

uniformly in n. Partition $[-M, M]$ into $m = [4M\delta^{-1}]$ sub-intervals such that $-M = x_0 \le x_1 \le \cdots \le x_m = M$ with $x_{r+1} - x_r < \delta$, for any given $\delta > 0$. Thus

$$\begin{aligned}
&\sup_{|x|\le M} |\frac{1}{n\sigma_n}\sum_{t=1}^{n} A_t(x) y_{t-1,n}| \\
&\le \max_r \sup_{x_{r-1}\le x\le x_r} |\frac{1}{n\sigma_n}\sum_{t=1}^{n} A_t(x) y_{t-1,n}| \\
&\le \max_r \sup_{x_{r-1}\le x\le x_r} |\frac{1}{n\sigma_n}\sum_{t=1}^{n} [A_t(x) - A_t(x_r)] y_{t-1,n}| \\
&+ \max_r |\frac{1}{n\sigma_n}\sum_{t=1}^{n} A_t(x_r) y_{t-1,n}| = J_{1n} + J_{2n}. \qquad (2.31)
\end{aligned}$$

Since $\sup_x |A_t'(x)| < \infty$, by Taylor expansion, we have

$$J_{1n} \le O(\delta)[\frac{1}{n\sigma_n}\sum_{t=1}^{n} |y_{t-1,n}|] = O_p(\delta). \qquad (2.32)$$

For J_{2n}, we have the following decomposition

$$\begin{aligned}
\frac{1}{n\sigma_n}\sum_{t=1}^{n} A_t(x) y_{t-1,n} &= \frac{1}{n\sigma_n}\sum_{t=1}^{n} A_t(x)[\sum_{i=1}^{t} \rho_n^{t-i}\epsilon_i] \\
&= \frac{1}{n\sigma_n}\sum_{i=1}^{n}[\sum_{t=i+1}^{n} A_t(x)]\rho_n^{t-i}\epsilon_i \\
&= \frac{1}{n\sigma_n}[\sum_{i=1}^{n}\epsilon_i\rho_n^{n-i}][\sum_{t=1}^{n} A_t(x)] \\
&\quad -\frac{1}{n\sigma_n}\sum_{i=1}^{n}[\sum_{t=1}^{i} A_t(x)]\epsilon_i\rho_n^{n-i} \\
&= U_{1n}(x) - U_{2n}(x). \qquad (2.33)
\end{aligned}$$

Since $\sum_{t=1}^{n} A_t(x)/n = o_p(1)$ for each x and $\frac{1}{\sigma_n}\sum_{i=1}^{n}\epsilon_i\rho_n^{n-i} = O_p(1)$, we have $\max_r |U_{1n}(x_r)| = o_p(1)$ for a given $\delta > 0$. Let $R_t(x) = A_t(x) -$

$G_0''\xi_{t-1}$, we have

$$\begin{aligned} U_{2n}(x) &= \frac{1}{n\sigma_n}\sum_{i=1}^{n}[\sum_{t=1}^{i} A_t(x)]\epsilon_i\rho_n^{n-i} \\ &= \frac{1}{n\sigma_n}\sum_{i=1}^{n}[\sum_{t=1}^{i}[A_t(x) - G_0''(x)\xi_{t-1}]]\epsilon_i\rho_n^{n-i} \\ &\quad +\frac{G_0''(x)}{n\sigma_n}\sum_{i=1}^{n}(\sum_{t=1}^{i}\xi_{t-1})\epsilon_i\rho_n^{n-i} \\ &= U_{3n}(x) + U_{4n}(x). \end{aligned} \tag{2.34}$$

For each x and ζ, by Theorem 3.1 in Ho and Hsing [28], we have

$$\mathrm{E}[\sum_{t=1}^{i} R_t(x)]^2 = O(i^{\max\{1,4(H-1/2)+2\zeta\}}). \tag{2.35}$$

For any $\eta > 0$ and $\sigma > 0$, we have

$$\begin{aligned} P(\max_r |U_{3n}(x_r)| > \eta) &\le \frac{1}{\eta}\sum_{r=1}^{m}\mathrm{E}|U_{3n}(x_r)| \\ &\le \frac{1}{\eta n\sigma_n}\sum_{r=1}^{m}\sum_{i=1}^{n}\{\mathrm{E}[\sum_{t=1}^{i} R_t(x)]^2\mathrm{E}\epsilon_i^2\rho_n^{2n-2i}\}^{1/2} \\ &\le \frac{1}{\eta n\sigma_n}\sum_{m}^{r=1}\sum_{i=1}^{n}\{\mathrm{E}[\sum_{t=1}^{i} R_t(x)]^2\mathrm{E}\epsilon_i^2\}^{1/2} \\ &= O(n^{-\alpha}L_0^{-1}(n)) \to 0, \end{aligned} \tag{2.36}$$

when $n \to \infty$, where $\alpha = \min\{H - 1/2, 1 - H - \zeta\} > 0$ and the last inequality follows from (4.8) in Chan and Ling [10]. Note that

$$\begin{aligned} U_{4n}(x) &= \frac{G_0'''(x)}{n\sigma_n}\sum_{i=1}^{n}(\sum_{t=1}^{i}\xi_{t-1})\epsilon_i\rho_n^{n-i} \\ &= \frac{G_0'''(x)}{n\sigma_n}\sum_{i=1}^{n}(\sum_{t=1}^{i}\epsilon_t)\epsilon_i\rho_n^{n-i} - \frac{G_0'''(x)}{n\sigma_n}\sum_{i=1}^{n}(\sum_{t=1}^{i}\epsilon_t)\epsilon_i\rho_n^{n-i}. \end{aligned}$$

By (iv) of Theorem 2.4 and $\sigma_n = O(n^H L_0(n))$, we have that the first term in $U_{4n}(x)$ is $o_p(1)$ uniformly in $x \in R$. Note that $\sum_{i=1}^{n}|\epsilon_i\rho_n^{n-i}|/n \le$

$\sum_{i=1}^n |\epsilon_i|/n = O_p(1)$ by the ergodic theorem and $\max_{1\le i\le n} \frac{|\sum_{t=1}^i e_t|}{\sqrt{n}} = O_p(1)$, the second term in U_{4n} is $o_p(1)$ uniformly in $x \in R$. Thus, we have $\max_x |U_{4n}(x)| = o_p(1)$ uniformly in $x \in R$. Furthermore, by (2.36) and (2.34), $\max_r |U_{2n}(x_r)| = o_p(1)$ for any given δ when $H \in (1/2, 1)$. Thus, Assumption 2.2 (d) holds when $H \in (1/2, 1)$.

Now we are ready to analyze the limit distribution of R_n. First, note the following identity holds,

$$\begin{aligned} R_n &= \sigma_n^{-1}(\hat{\rho}_n - \rho_n)\sum_{t=1}^n y_{t-1,n} \\ &= \sigma_n^{-1}\frac{\sum_{t=1}^n y_{t-1,n}\epsilon_t}{\sum_{t=1}^n y_{t-1,n}^2}\sum_{t=1}^n y_{t-1,n}. \end{aligned} \tag{2.37}$$

By the continuous mapping theorem and (i), (ii), (iii) of Theorem 2.3, the limit distribution of R_n becomes

$$\begin{aligned} R_n &= \sigma_n^{-1}(\hat{\rho}_n - \rho_n)\sum_{t=1}^n y_{t-1,n} \\ &\to_{\mathcal{L}} [\frac{1}{2}(Z_{H,\gamma}^1(1))^2 + \gamma\int_0^1 (Z_{H,\gamma}^1(u))^2\,du] \\ &\quad \times[\int_0^1 Z_{H,\gamma}^1(u)\,du][\int_0^1 (Z_{H,\gamma}^1(u))^2\,du]^{-1}. \end{aligned} \tag{2.38}$$

By (3.2) in Ho and Hsing [27], we have

$$|\sup_x F'(x)|^{-1}\sup_x |K_n(x)| \to_{\mathcal{L}} |B_H(1)|\,. \tag{2.39}$$

Combining (2.4), (2.38) and (2.39), we can conclude the result. □

Theorem 2.4 gives the limit distribution of the Kolmogorov-Smirnov statistic. It can be used to test for the distribution of the long-memory noise in model (2.1). The percentiles of the limit distribution are tabulated in section 2.3.

2.3 Simulation Study

To assess the finite sample behavior of the limit distribution established in the preceding section, a number of simulation studies are presented here.

2.3.1 Simulating the Sample Paths of the OU Processes

First recall the AR(1) model defined in equation (2.1). Let the time series $\{y_{t,n}\}$ be generated by model

$$y_{t,n} = \rho_n y_{t-1,n} + \epsilon_t \quad \text{and} \quad \epsilon_t = \sum_{i=0}^{\infty} a_i e_{t-i}, \tag{2.40}$$

where $\rho_n = 1 - \gamma/n$ and the coefficients a_i satisfy $\sum_{i=1}^{\infty} a_i^2 < \infty$; $a_0 = 1$ and $a_k = k^{H-3/2}L_0(k)$ for some slowly varying function L_0 with $H \in (1/2, 1)$; and $\{e_t\}$ is a sequence of i.i.d.mean zero random variables with $\sigma_e^2 = \mathrm{E}e_t^2 < \infty$.

On the basis of Theorem 2.2, the sample path of the Ornstein-Uhlenbeck process can be simulated approximately by the linear process on the left side of equation (2.20). However, this linear process has an infinite number of terms, which makes it difficult to implement. To deal with this problem, the truncation scheme proposed in Dietrich and Newsam [19] is adopted; see also Wu, Michailidis and Zhang [63]. In this scheme, the first m terms are used to simulate the linear process, that is, ϵ_t is approximated by $\epsilon_{t,m} = \sum_{i=0}^{m} a_i e_{t-i}$. Here, m is the truncation number, which is chosen to ensure the difference between the truncated version and the infinite-sum version is sufficiently small. Moreover, to speed up calculations, we embed the coefficients a_i in a circulant matrix. In this case, $\{\epsilon_{t,m}\}$ is generated by the following relation.

$$\begin{pmatrix} \epsilon_{1,m} \\ \vdots \\ \epsilon_{m,m} \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \dots & a_{m-1} & a_m \\ a_2 & a_3 & \dots & a_m & a_1 \\ \dots & \dots & \dots & \dots & \dots \\ a_m & a_1 & \dots & a_{m-2} & a_{m-1} \end{pmatrix} \mathbf{e} = A\mathbf{e},$$

where A is the circulant matrix and $\mathbf{e}^T = (e_1, \dots, e_m)$. Then, the Ornstein-Uhlenbeck processes are simulated approximately by $\sum_{t=1}^{[nu]} \left(1 - \frac{\gamma}{n}\right)^{[nu]-[nt]} \epsilon_{t,m}/\sigma_n (0 \le u \le 1)$. Here, we investigate the case in which $\mathbf{e}$ is standard normal. In this simulation study, m is set to 4,000. Sample paths of Ornstein-Uhlenbeck processes with different parameters, H and γ, are shown in the following figures.

Figure 2.1: Sample paths of Ornstein-Uhlenbeck process with Hurst index and $H = 0.65$.

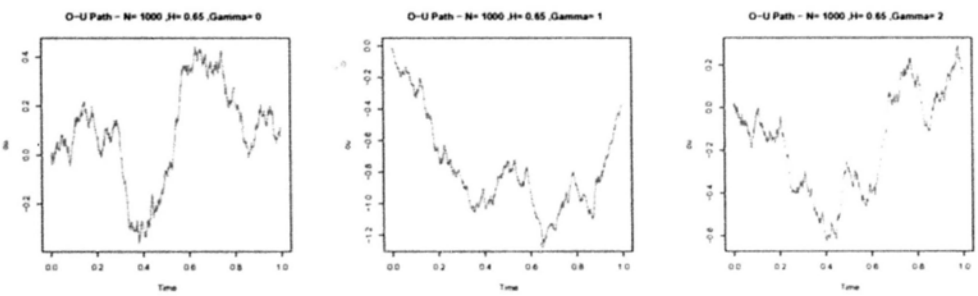


Figure 2.2: Sample paths of Ornstein-Uhlenbeck process with Hurst index $H = 0.75$.

Figure 2.3: Sample paths of Ornstein-Uhlenbeck process with Hurst index $H = 0.85$.

2.3.2 Percentiles of K

The critical values of K for different parameters are tabulated in the following tables. These tables are constructed as follows. First, the interval $[0, 1]$ is partitioned evenly with $t_{j+1} - t_j = 1/n$, and the integral $\int_0^1 e^{-\gamma(t-s)}\, dB_H(s)$ is approximated by $\sum \exp(-\gamma(t-s_j)) \cdot \triangle B_j^H$, where $\triangle B_j^H = B_H(t_{j+1}) - B_H(t_j)$. K is simulated independently N times to obtain the percentiles of K. By increasing n and N, the percentiles of K exhibit ignorable differences from those in Tables 2.1, 2.2 and 2.3. Here, let $n = 500$ and $N = 10,000$. Histograms of K are plotted in Figures 2.4-2.6.

Table 2.1: Percentiles of K for $H = 0.65$ (sample size=10,000).

	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
$\gamma = 0.0$	0.0097	0.0238	0.0490	0.0916	0.6197	0.7226	0.8186	0.9151
$\gamma = 1.0$	0.0067	0.0171	0.0362	0.0698	0.4922	0.5891	0.6848	0.8113
$\gamma = 2.0$	0.0069	0.0151	0.0288	0.0536	0.4177	0.5057	0.6038	0.7341
$\gamma = 3.0$	0.0045	0.0112	0.0228	0.0446	0.3585	0.4432	0.5445	0.6694
$\gamma = 4.0$	0.0039	0.0099	0.0194	0.0398	0.3287	0.4099	0.4911	0.5967

Figure 2.4: Histogram of statistic with Hurst index $H = 0.65$.

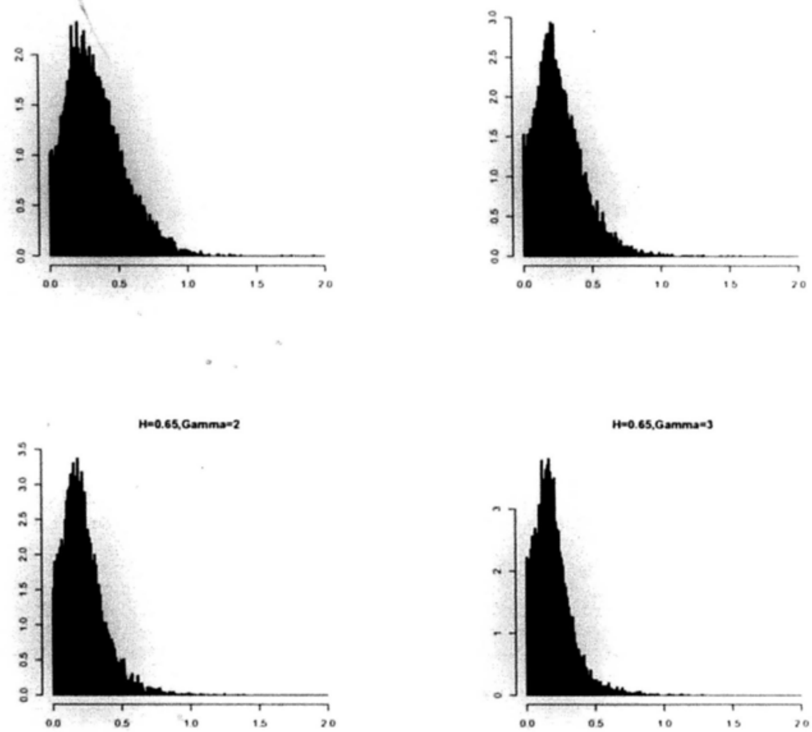


Table 2.2: Percentiles of K for $H = 0.75$ (sample size=10,000).

	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
$\gamma = 0.0$	0.0091	0.0245	0.0473	0.0826	0.0553	0.6491	0.7287	0.8382
$\gamma = 1.0$	0.0066	0.0165	0.0331	0.0616	0.4463	0.5218	0.5952	0.6988
$\gamma = 2.0$	0.0048	0.0140	0.0263	0.0493	0.3600	0.4306	0.5004	0.5974
$\gamma = 3.0$	0.0039	0.0100	0.0201	0.0398	0.3079	0.3706	0.4284	0.5301
$\gamma = 4.0$	0.0034	0.0084	0.0173	0.0331	0.2683	0.3266	0.3910	0.4884

Table 2.3: Percentiles of K for $H = 0.85$ (sample size=10,000).

	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
$\gamma = 0.0$	0.0079	0.0213	0.0399	0.0710	0.4693	0.5499	0.6263	0.7147
$\gamma = 1.0$	0.0050	0.0135	0.0288	0.0528	0.3775	0.4415	0.4978	0.5708
$\gamma = 2.0$	0.0049	0.0110	0.0218	0.0419	0.3013	0.3533	0.4052	0.4704
$\gamma = 3.0$	0.0036	0.0083	0.0174	0.0346	0.2518	0.2953	0.3427	0.3966
$\gamma = 4.0$	0.0030	0.0073	0.0151	0.0306	0.2170	0.2541	0.2973	0.3571

Figure 2.5: Histogram of statistic with Hurst index $H = 0.75$ and $\gamma = 0$ (left), $\gamma = 1$ (middle) and $\gamma = 2$ (right).

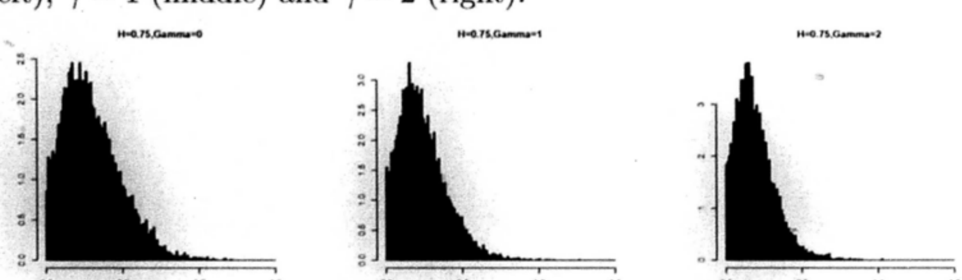
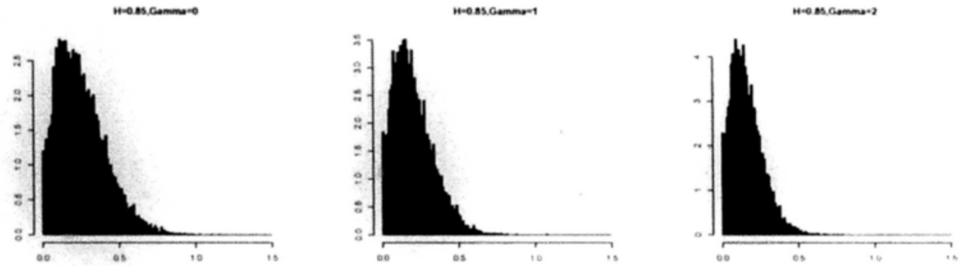


Figure 2.6: Histogram of statistic with Hurst index $H = 0.85$ and $\gamma = 0$ (left), $\gamma = 1$ (middle) and $\gamma = 2$ (right).



2.3.3 Examples

We now consider some simple examples to examine the performance of the test statistic.

Example 2.5. *Consider the nearly unstable model*

$$y_{t,n} = \rho_n y_{t-1,n} + \epsilon_t, \tag{2.41}$$

where $\rho_n = 1-\gamma/n$ *and* $\epsilon_t = \sum_{i=0}^{\infty} a_i e_{t-i}$. *The coefficients* $a_i = i^{H-3/2}L(i)$ *are square summable. Here, we consider the case that* $\{e_i\}$ *is a sequence of centered stationary Gaussian random variables with variance* σ^2. ϵ_t *is approximated by its truncated version,* $\epsilon_{t,m} = \sum_{i=0}^{m} a_i e_{t-i}$. $L(i)$ *is chosen, such that* $\sum_{i=0}^{m} a_i^2 = 1$. *Obviously,* $\epsilon_{t,m}$ *follows a normal distribution with variance* σ^2.

Consider the following Hypothesis.

$$H_0 : \sigma^2 = 1 \quad versus \quad H_1 : \sigma^2 \neq 1. \tag{2.42}$$

We take sample sizes $m = 1,000$ *and* $2,000$, *and* $1,000$ *replications are used.*

The power and size are given in Tables 2.4 and 2.5 at the selected level and parameters. From these tables, we can see that when n

is increased, the size is very close to the nominal significance level. As σ^2 is increased from 1, the rejection rate also increases. This is reasonable, as the test depends on the difference between the shape of the empirical distribution and the true distribution. When σ^2 is far from 1, the difference between the shape of $N(0,1)$ and $N(0,\sigma^2)$ is large, which means the rejection rate is high. These simulation studies indicate that the proposed test has satisfactory size and power behavior in the finite samples. They should be useful in practice.

Table 2.4: Size and power of test statistic $\hat{K}_n$ with $H = 0.65$ and $\gamma = 0$.

	$n = 1,000$			$n = 2,000$		
α	0.025	0.05	0.1	0.025	0.05	0.1
			size			
$\sigma^2 = 1$	0.015	0.031	0.087	0.019	0.042	0.097
			power			
$\sigma^2 = 1.2$	0.017	0.051	0.110	0.033	0.070	0.164
$\sigma^2 = 1.3$	0.043	0.093	0.208	0.066	0.159	0.307
$\sigma^2 = 1.4$	0.078	0.135	0.271	0.131	0.286	0.496
$\sigma^2 = 1.5$	0.127	0.256	0.453	0.283	0.479	0.716
$\sigma^2 = 1.6$	0.204	0.346	0.549	0.404	0.643	0.849
$\sigma^2 = 1.7$	0.289	0.466	0.711	0.613	0.821	0.956
$\sigma^2 = 1.8$	0.411	0.606	0.806	0.772	0.917	0.986
$\sigma^2 = 1.9$	0.495	0.709	0.905	0.889	0.977	0.998
$\sigma^2 = 2.0$	0.624	0.826	0.956	0.963	0.993	0.998
$\sigma^2 = 2.1$	0.713	0.886	0.918	0.985	0.999	1.000
$\sigma^2 = 2.2$	0.813	0.938	0.988	0.997	1.000	1.000
$\sigma^2 = 2.3$	0.887	0.978	0.998	1.000	1.000	1.000
$\sigma^2 = 2.4$	0.948	0.994	0.998	1.000	1.000	1.000
$\sigma^2 = 2.5$	0.959	0.994	1.000	1.000	1.000	1.000

Figure 2.7: Power Function.

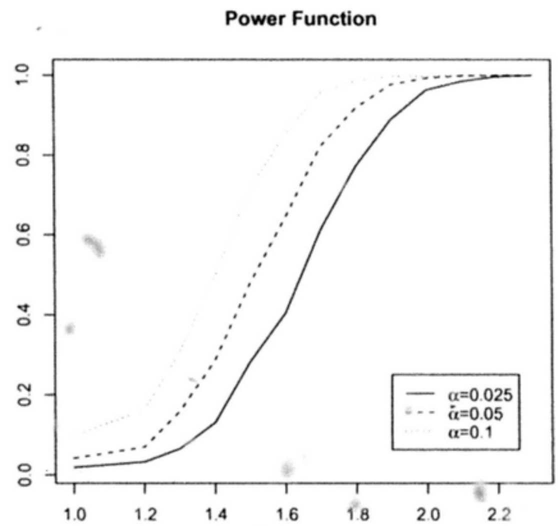
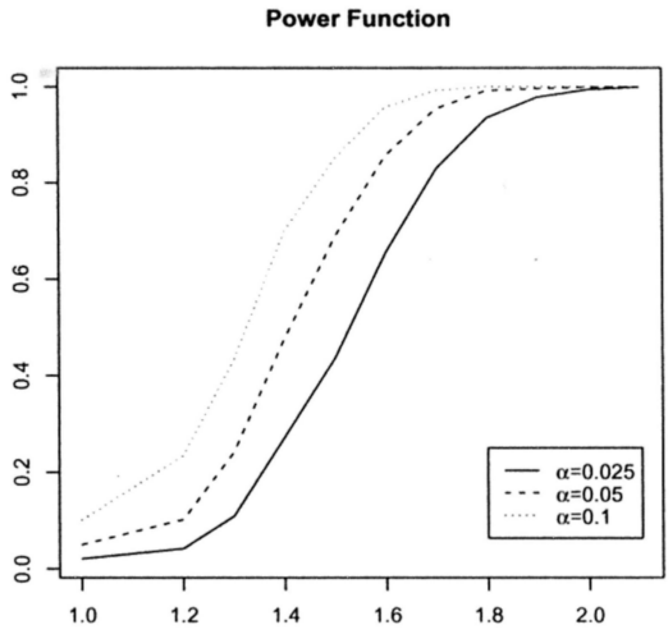


Table 2.5: Size and power of test statistic $\hat{K}_n$ with $H = 0.65$ and $\gamma = 1$.

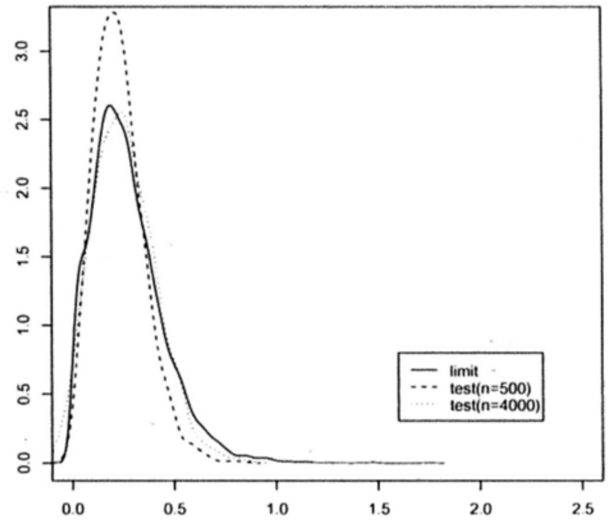
	$n = 1,000$			$n = 2,000$		
α	0.025	0.05	0.1	0.025	0.05	0.1
			size			
$\sigma^2 = 1$	0.013	0.046	0.129	0.021	0.050	0.101
			power			
$\sigma^2 = 1.2$	0.023	0.074	0.176	0.042	0.102	0.234
$\sigma^2 = 1.3$	0.067	0.158	0.303	0.109	0.241	0.432
$\sigma^2 = 1.4$	0.122	0.240	0.440	0.272	0.476	0.699
$\sigma^2 = 1.5$	0.176	0.356	0.577	0.436	0.688	0.850
$\sigma^2 = 1.6$	0.312	0.558	0.764	0.655	0.856	0.957
$\sigma^2 = 1.7$	0.464	0.729	0.889	0.829	0.954	0.992
$\sigma^2 = 1.8$	0.609	0.826	0.939	0.935	0.992	1.000
$\sigma^2 = 1.9$	0.750	0.911	0.980	0.978	0.996	1.000
$\sigma^2 = 2.0$	0.857	0.961	0.996	0.994	0.999	1.000
$\sigma^2 = 2.1$	0.914	0.989	0.999	0.999	0.999	1.000
$\sigma^2 = 2.2$	0.956	0.993	0.999	1.000	1.000	1.000
$\sigma^2 = 2.3$	0.986	0.998	1.000	1.000	1.000	1.000
$\sigma^2 = 2.4$	0.989	1.000	1.000	1.000	1.000	1.000

Figure 2.8: Power Function.



The estimated density functions of the statistics with parameters $H = 0.65$ and $\gamma = 1$ for $n = 500$ (dashed line) and $n = 4000$ (dotted line) are plotted in Figure 2.9. The density function of the limit distributions of K (solid line) is also given on the same figure. We can see that when the sample size is increased, the estimated density function of the statistic approaches to that of its limit distribution, indicating that convergence behavior of the statistic is fast.

Figure 2.9: Density Functions



Example 2.6. *In this example, we restrict our attention to test the normality. Consider the following model.*

$$y_{t,n} = \rho_n y_{t-1,n} + \epsilon_t, \tag{2.43}$$

where $\rho_n = 1 - \gamma/n$ *and* $\epsilon_t = \sum_{i=0}^{\infty} a_i e_{t-i}$. *The coefficients* $a_i = i^{H-3/2}L(i)$ *are square summable. When* $\{e_i\}_{i\geq 1}$ *is an i.i.d. Gaussian sequence, then* ϵ_t *is also Gaussian. Let* F *denote the d.f. of* ϵ_t, *and consider the following hypothesis.*

$$H_0 \quad F = \Phi \quad versus \quad F \neq \Phi. \tag{2.44}$$

In this simulation study, the selected alternatives of e_i *are* $A_1 : \sqrt{3/5}t_5$, $A_2 : \sqrt{1/2}t_4$ *and* $A_3 : \sqrt{1/3}t_3$, *respectively.* ϵ_t *is approximated by its truncated version,* $\epsilon_{t,m} = \sum_{i=0}^{m} a_i e_{t-i}$, $\{e_i\}_{i=0}^{m}$ *is simulated independently, and* $L(i)$ *is chosen such that* $\sum_{i=0}^{m} a_i^2 = 1$. *We take the sample sizes* $m = 1,000$ *and* $2,000$, *and* $1,000$ *replications are used.*

Table 2.6: Size and power of test statistic K_n with $H = 0.65$ and $\gamma = 1$.

	$n = 1000$				$n = 2000$		
α	0.025	0.05	0.1		0.025	0.05	0.1
				size			
$N(0,1)$	0.013	0.046	0.129		0.021	0.050	0.101
				power			
A_1	0.019	0.068	0.213		0.028	0.078	0.221
A_2	0.027	0.106	0.296		0.079	0.204	0.395
A_3	0.285	0.514	0.752		0.477	0.713	0.874

From Table 2.6, we can see that when sample size n is increased, the size becomes very close to the nominal significant level for all selected levels. The power of this test depends on the difference between the shape of $N(0,1)$ and that of $\{A_i\}_{i=1,2,3}$. This is reasonable, as the difference between the shapes of $N(0,1)$ and A_3 is more than that between the shapes of $N(0,1)$ and A_1. This example shows that even if the sample comes from the distribution family that deviates from that hypothesized, the reject rate remains satisfactory.

□ **End of chapter.**

Chapter 3

Auxiliary Results I

In this chapter, we investigate the limit distribution of the weighted sum of the functionals of moving average processes. It is shown that this distribution can be expressed as Wiener integrals with respect to Hermite processes. Now let us begin with the definition of the stochastic integral with respect to the Hermite process when the integrands are some deterministic functions.

3.1 Background

Self-similar processes have been widely applied to models of various phenomena, including hydrology, network traffic analysis and finance.

An interesting class of self-similar processes is the Hermite process, given by

$$Z_H^k = c(H,k)\int_{\mathcal{R}^k}\int_0^t \Big(\prod_{j=1}^{k}(s-y_i)_+^{-(\frac{1}{2}+\frac{1-H}{k})}\Big)\, ds\, dB(y_1)\cdots dB(y_k), \quad (3.1)$$

where the foregoing integral is a Wiener-Itô multiple integral of order k with respect to the standard Brownian motion $(B(y))_{y\in\mathcal{R}}$ and $c(H,k)$ is a positive normalization constant such that $\mathrm{E}[Z_H^k]^2 = 1$. When $k = 1$,

$Z_H^k(t)$ corresponds to the fractional Brownian motion (fBm), $B_H(t)$. Due to the various applications of the fractional Brownian motion, a rich theory of stochastic integration with respect to the fBm has been established. For example, Decreusefond and Üstünel [17] defined a stochastic integral with respect to the fBm by using the stochastic calculus of variations (also known as the Malliavin calculus) and the fact that the fBm is a Gaussian process. A different approach, that of pathwise integration, was taken by Dudley and Norvaiš [20] and Zähle [66]. They used the specific path properties of the fBm, namely, p-variation in Dudley and Norvaiš [20] and Hölder continuity in Zähle [66]. A particular case in which the integrands are not random was examined by Pipiras and Taqqu [43], who investigate a class of functions can be used to construct the integral with respect to the fBm. For the general case, Maejima and Tudor [39] introduce Wiener integrals with respect to the Hermite process of order k when the integrands are deterministic functions of specific classes.

Of central interest here is to investigate the limit of the weighted sums of a sequence of long-range dependent random variables, which appears as the Wiener integral with respect to the Hermite process, Z_H^k. Pipiras and Taqqu [42] proved that if f is a deterministic function, then the sequence

$$\frac{1}{n^H}\sum_j f(\frac{j}{n})X_j,$$

where $(X_j)_{j\in\mathcal{Z}}$ is in the domain of attraction of the fBm, converges weakly, as $n\to\infty$, to the Wiener integral $\int_{\mathcal{R}} f(u)\,dB_H(u)$, where B_H denotes the fractional Brownian motion. A natural extension of this result is to consider the convergence of sequences:

$$\frac{1}{n^H}\sum_j f(\frac{j}{n})g(X_j). \tag{3.2}$$

When $\{X_j\}_{j\in Z}$ are stationary Gaussian sequences and g has the Hermite expansion $g(x) = \sum_{l=0}^{\infty} c_l H_l(x)$ with Hermite rank k, where H_l is the Hermite polynomial of degree l and $c_l = \frac{1}{n!}\mathrm{E}\left[g(X_0)H_l(X_0)\right]$, as shown in Maejima and Tudor [39],

$$\frac{1}{n^H}\sum_j f(\frac{j}{n})g(X_j) \to_{\mathcal{L}} \int f\, dZ_H^k, \quad \text{as } n \to \infty, \tag{3.3}$$

where $\to_{\mathcal{L}}$ denotes weak convergence and Z_H^k is defined as in (3.1).

In this chapter, we relax the Gaussian assumption and consider $\{X_j\}$ to be moving average processes and g to have power rank k. It is shown that, under certain conditions,

$$\frac{1}{n^H L^k(n)}\sum_j f(\frac{j}{n})g(X_j) \to_{\mathcal{L}} \int f\, dZ_H^k. \tag{3.4}$$

When dealing with the non-Gaussian case, no appropriate expansion of $g(X_j)$ can be used to prove the convergence of partial sums in the L^2 sense. To fix this problem, we restrict our attention to the case in which X_i is a moving average process, that is, $g(X_i)$ is a functional of the moving average process. Ho and Hsing [27] developed a new technique to decompose $g(X_i)$ into two asymptotically uncorrelated terms to obtain the weak convergence of the partial sum of $g(X_i)$. They showed that the weighted sum of the functional of the moving average processes weakly converges to a limit of the Wiener integral with respect to the Hermite process.

3.2 The Hermite Process and Wiener Integrals with respect to it

3.2.1 The Hermite Process

First, some of the basic properties of the Hermite process $Z_H^k = \left(Z_H^k(t)\right)_{t\in R}$ of order $k \geq 1, k \in Z$, with Hurst parameter $H \in (1/2, 1)$, are presented. This stochastic process is defined as a multiple Wiener-Itô integral of order k with respect to the standard Brownian motion, $(B(t))_{t\in R}$, as follows.

$$Z_H^k(t) = c(H,k)\int_{R^k}\int_0^t \Big(\prod_{j=1}^{k}(s-y_i)^{(\frac{1}{2}+\frac{1-H}{k})}\Big)\,ds\,dB(y_1)\cdots dB(y_k), \tag{3.5}$$

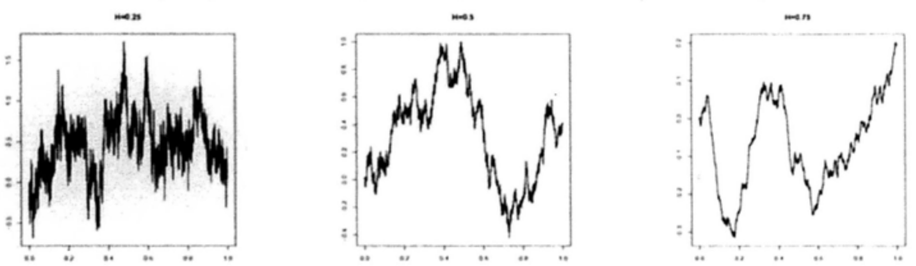
where $x_+ = \max(x, 0)$ and

$$\begin{aligned} c(H,k) &= \{k!(1-k(1-H))(1-k(2-2H)) \\ &\quad \times \Big[\int_0^\infty (x+x^2)^{H-3/2}\,dx\Big]^{-k}\}^{1/2} \end{aligned}$$

making $\mathrm{E}[\big(Z_H^k(t)\big)^2] = 1$. When $k = 1$, the process corresponds to the fractional Brownian motion with Hurst parameter $H \in (0, 1)$. Let $R(t, s)$ denote the covariance function of $Z_H^k(t)$, that is $R(t, s) = \mathrm{E}\left[Z_H^k(t)Z_H^k(s)\right]$. Maejima and Tudor [39] showed that

$$R(t,s) = \frac{1}{2}(t^{2H} + s^{2H} - |t-s|^{2H}). \tag{3.6}$$

Figure 3.1: Sample paths of the fractional Brownian motion with Hurst index $H = 0.25$(left), $H = 0.5$(middle) and $H = 0.75$ (right).



The following are some of the properties of the Hermite process used in this thesis.

- Process Z_H^k is H-selfsimilar with stationary increments.

- From the stationarity of the increments and the foregoing self-similarity, it follows that, for any $p \geq 1$,

$$\mathrm{E}\left[\left|Z_H^k(t) - Z_H^k(s)\right|^p\right] = c(p, H, k)\,|t - s|^{pH}. \tag{3.7}$$

As a consequence, the Hermite process has Hölder-continuous paths of order $\delta < H$.

3.2.2 Wiener Integrals with Respect to the Hermite Process

Let Z_H^k denote the Hermite process of order k for a deterministic function f; $\int f\, dZ_H^k$ is called the integral on the real line of f with respect to this Hermite process. The following question now arises. What is the class of functions from which the integral on the real line can be well-defined? This interesting isssue has been extensively investigated; see Pipiras and Taqqu [42] and Maejima and Tudor [39]. Now, let us review their results beginning with the definition of the integral on the real line with respect to Z_H^k.

Let

$$f(u) = \sum_{i=1}^{n} c_i 1_{(t_i, t_{i+1}]}(u), \quad t_i < t_{i+1}, c_i \in R, \quad i = 1, \dots, n, \tag{3.8}$$

be an element from the set, $\mathcal{E}$, of elementary functions. For such an $f(u)$, it is natural to define its Wiener integral with respect to Hermite process Z_H^k by

$$\int_R f(u)\, dZ_H^k(u) = \sum_{i=1}^{n} c_i(Z_H^k(t_{i+1}) - Z_H^k(t_i)). \tag{3.9}$$

Let $\overline{sp}(Z_H^k)$ be the closure of the span of the increments of Z_H^k. It is obvious that the right side in (3.9) belongs to $\overline{sp}(Z_H^k)$. It is also well-known that when $Z_H^k = Z_{1/2}^1$ is the usual Brownian motion, element $X \in \overline{sp}(B)$ can be characterized by the unique function $f_X \in L^2(\mathcal{R})$, in which case one writes X in the integral form $X = \int_R f_X(u)\, dB(u)$ and $B(u)$ is a standard Brownian motion. From a different perspective, the space $L^2(R)$ forms a class of integrands for the integral on the real line with respect to $B(u)$. As $L^2(R)$ is a complete space and $\mathcal{E}$ is dense in $L^2(R)$, we can associate X with a sequence of elementary functions, f_n, $f_n \to f$, such that $\int f_n(u)\, dB(u) \to_{\mathcal{L}} \int f(u)\, dB(u)$ in the L^2 sense. In this case, the integral $\int f(u)\, dB(u)$ can be approximated by a sequence of the integral of f_n with respect to $B(u)$, that is,

$$\int f(u)\, dB(u) = \lim_{n\to\infty} \int f_n(u)\, dB(u)$$

and

$$\mathrm{Var}(\int f(u)\, dB(u)) = \lim_{n\to\infty} \mathrm{Var}(\int f_n(u)\, dB(u)) = \int f(u)^2\, du.$$

Because the mapping $f \mapsto \int f\, dB$ is a linear and onto map between $L^2(\mathcal{R})$ and $\overline{sp}(B)$, which preserves inner products, it is an isometry, and the Hilbert space, $\overline{sp}(B)$, and $L^2(R)$ are isometric.

Based on this idea, Pipiras and Taqqu [42] investigated a case in which $Z_H^k (k = 1)$ is the fractional Brownian motion. In that work, they explored whether a similar characterization of the elements of $\overline{sp}(B_H)$ could be obtained when $H \in (1/2, 1)$. These characterizations form a class of integrands for the integral on the real line with respect to $B_H(u)$. These classes of integrands are inner product spaces. If the space of the integrands is not complete, then that space characterizes only a strict subset of $\overline{sp}(B_H)$. From the left side of (3.14), it is obvious that $\mathcal{E}$ is contained in this space. Now, we introduce the space constructed by Pipiras and Taqqu [42].

Let

$$\mathcal{H} = \{f : \mathcal{R} \to \mathcal{R} \mid \int_{\mathcal{R}} \int_{\mathcal{R}} f(u) f(v)\, |u - v|^{2H-2}\, du\, dv < \infty\}$$

be equipped with the inner product

$$< f, g >_{\mathcal{H}} = H(2H-1)\int_{\mathcal{R}}\int_{\mathcal{R}} f(u)g(v)\,|u-v|^{2H-2}\,du\,dv.$$

The following are some of the properties proved by Pipiras and Taqqu [42].

- Set $\mathcal{E}$ is dense in $\mathcal{H}$.
- For $H \in (0, 1/2)$, $\mathcal{H}$ is a complete space. For $H \in (1/2, 1)$, $\mathcal{H}$ is not a complete space.
- The space $|\mathcal{H}|$ is not complete with respect to norm $\|\cdot\|_{\mathcal{H}}$, but is a Banach space with respect to norm

$$\| f \|^2_{|\mathcal{H}|} = \int_{\mathcal{R}}\int_{\mathcal{R}} |f(u)|\,|f(v)|\,|u-v|^{2H-2}\,dv\,dv. \qquad (3.10)$$

- $|\mathcal{H}|$ is a strict subspace of $\mathcal{H}$ and is, in fact (see Pipiras and Taqqu [42]),

$$L^1(\mathcal{R}) \cap L^2(\mathcal{R}) \subset |\mathcal{H}| \subset \mathcal{H}, \qquad (3.11)$$

where $L^1(\mathcal{R})$ and $L^2(\mathcal{R})$ are defined, respectively,

$$L^1(\mathcal{R}) = \{f : \int_{\mathcal{R}} |f(u)|\,du < \infty\} \qquad (3.12)$$

and

$$L^2(\mathcal{R}) = \{f : \int_{\mathcal{R}} |f(u)|^2\,du < \infty\}. \qquad (3.13)$$

As $\mathcal{E}$ is dense in $|\mathcal{H}|$, for any arbitrary function $f \in |\mathcal{H}|$, there exists a sequence of elementary functions $\{f_n\}_{n\geq 1}$ such that $|f_n - f|_{|\mathcal{H}|} \to 0$. Therefore, we have

$$\int f(u)\,dB_H(u) = \lim_{n\to\infty}\int f_n(u)\,dB_H(u) \qquad (3.14)$$

and

$$\begin{aligned}
&\mathrm{Var}\Big(\int f(u)\,dB_H(u)\Big) \\
&= \lim_{n\to\infty} \mathrm{Var}\Big(\int f_n(u)\,dB_H(u)\Big) \\
&= H(2H-1)\int\int f(u)f(v)|u-v|^{2H-2}\,du\,dv. \qquad (3.15)
\end{aligned}$$

The integral on the real line with respect to B_H with $H \in (1/2, 1)$ is well-defined in (3.14).

Based on the observation that the covariance structure of the Hermite process is similar to that of the fractional Brownian motion, Maejima and Tudor [39] extended this result to the Hermite process using $|\mathcal{H}|$ as the classes of deterministic integrands. Let $S_{|\mathcal{H}|}$ be the corresponding subspace of $\overline{sp}(B_H)$; then, the mapping $f \mapsto \int f\,dB_H$, $f \in |\mathcal{H}|$ and $H \in (1/2, 1)$ is an isometry from $|\mathcal{H}|$ to $S_{|\mathcal{H}|}$. As $|\mathcal{H}|$ is a complete space, we can make the following assertions.

- Every element of linear subspace $S_{|\mathcal{H}|}$ can be expressed as an integral of a deterministic function with respect to B_H.

- The integral on the real line with respect to B_H with $H \in (1/2, 1)$ is well-defined for the functions from $|\mathcal{H}|$.

3.3 Non-central Limit Theorem

Consider a moving average process

$$X_j = \sum_{i=0}^{\infty} a_i \epsilon_{j-i}, \qquad (3.16)$$

where ϵ_j are mean zero i.i.d. random variables having at least finite second moments, and the moving average coefficients a_i satisfy $\sum_{i=1}^{\infty} a_i^2 < \infty$. Our goal is to investigate the asymptotic behavior of

$\sum_j f\left(\frac{j}{n}\right) K(X_j)$, as $n \to \infty$, where f is a deterministic function on $\mathcal{R}$ and K comes from a general class of measurable functions.

We focus on the primary case where the a_i's are regularly varying with exponent $-\beta$, denoted by $a_i \in RV_{-\beta}$, for some $\beta \in (1/2, 1)$, that is, $a_i = i^{-\beta} L(x)$ and $L(x)$ is slowly varying at ∞.

Let $\mathcal{F}_t = \{\ldots, \epsilon_{t-1}, \epsilon_t\}$ and define the truncated process

$$\tilde{X}_{j,i} = \mathrm{E}[X_j \mid \mathcal{F}_{j-i}] = \sum_{m=i}^{\infty} a_m \epsilon_{j-m} \tag{3.17}$$

and let

$$X_{j,i} = X_j - \tilde{X}_{j,i} = \sum_{0 \le m \le i-1} a_m \epsilon_{j-m}. \tag{3.18}$$

Let F_i, $\tilde{F}_i$, G_i and G be the distribution functions of $X_{j,i}$, $\tilde{X}_{j,i}$, $a_i\epsilon_1$ and ϵ_1 respectively. For $i > 0$, define

$$K_i(x) = \int K(x+y)\, dF_i(y) \quad , \quad K_\infty(x) = \int K(x+y)\, dF(y)$$

and

$$K_\infty^{(r)}(x) = \frac{d^r}{dx^r} \int K(x+y)\, dF(y).$$

Definition 3.1. *We say that K has power rank k for some positive integer k, if $K_\infty^{(k)}(0)$ exists and is nonzero and $K_\infty^{(m)}(0) = 0$ for $1 \le m < k$.*

If the t-th derivative $K_i^{(t)}$ of K_i exists, define

$$K_{i,\lambda}^{(t)}(x) = \sup_{|y| \le \lambda} |K_i^{(t)}(x+y)|, \quad \lambda \ge 0.$$

Condition 1. *Let $E(|\epsilon_t|^q) < \infty$ for some $2 < q \le 4$ and $K_i \in \mathcal{C}^{p+1}(\mathcal{R})$ for all large i. Assume that for some $\lambda > 0$,*

$$\sum_{t=0}^{k+1} \| \quad K_{i-1,\lambda}^{(t)}(X_{i,0}) \| + \sum_{t=0}^{k-1} \| \, |\epsilon_1|^{q/2} K_{i-1,\lambda}^{(t)}(X_{i,1}) \| \tag{3.19}$$

$$+ \| \epsilon_1 K_{i-1}^{(k)}(X_{i,1}) \| = O(1). \tag{3.20}$$

Condition 2. *If $\beta \in (1/2, 1)$ and $k(2\beta - 1) < 1$, k is the power rank of $K(\cdot)$ and condition 1 holds with $q = 4$.*

Let $H = 1 - k(\beta - \frac{1}{2})$ and $\sigma_n = c(H,k)n^H L^k(n)$, introduce the sequence of stochastic processes $Z_H^{k,n}$ defined by

$$Z_H^{k,n}(u) = \frac{1}{\sigma_n}\sum_{j=1}^{[nu]} K(X_j) \ (u \geq 0) \tag{3.21}$$

and

$$Z_H^{k,n}(u) = \frac{1}{\sigma_n}\sum_{j=-[nu]-1}^{0} K(X_j) \ (u \leq 0), \tag{3.22}$$

where K is a function of power rank k. By Theorem 1 of Wu [62], under condition (2), it follows that

$$Z_H^{k,n}(u) \to_{\mathcal{L}} K_\infty^{(k)}(0) Z_H^k(u) \ \text{ as } n \to \infty,$$

where $\to_{\mathcal{L}}$ denote weak convergence and Z_H^k is a Hermite process defined in (3.5).

We also use the notation that, if f is a function on R,

$$f_n(x) = \sum_{j=-\infty}^{\infty} f(\frac{j}{n}) 1_{(\frac{j}{n}, \frac{j+1}{n}]},$$

and

$$f_{n,T}^+(x) = \sum_{j=0}^{T} f(\frac{j}{n}) 1_{(\frac{j}{n}, \frac{j+1}{n}]}, \quad f_{n,T}^-(x) = \sum_{j=-T}^{-1} f(\frac{j}{n}) 1_{(\frac{j}{n}, \frac{j+1}{n}]}.$$

Also, let $f_n^+ = f_{n,\infty}^+$ and $f_n^- = f_{n,\infty}^-$.

Theorem 3.1. *Under condition 2, for $k(2\beta - 1) < \beta$, let $f \in |\mathcal{H}|$ such that $f_n^\pm \in |\mathcal{H}|$ for every $n \geq 1$. If $|f_n - f|_{|\mathcal{H}|} \to 0$, as $n \to \infty$, and for every n, $|f_{n,T}^\pm - f_n^\pm| \to 0$, as $T \to \infty$, then, as $n \to \infty$,*

$$\frac{1}{\sigma_n}\sum_j f(\frac{j}{n}) K(X_j) \to_{\mathcal{L}} K_\infty^{(k)}(0) \int_R f(u)\, dZ_H^k(u).$$

Remark 3.1. *$k(2\beta-1)<\beta$ is a strict condition, however for $K(X)=X$, that is, $k=1$, Theorem 3.1 coincides with the results in Pipiras and Taqqu [42].*

Proof. Let us prove first that the sum

$$S(n)=\frac{1}{n^H L^k(n)}\sum_j f(\frac{j}{n})K(X_j), \tag{3.23}$$

where $H=1-k(\beta-\frac{1}{2})$, is convergent. Note that

$$\mathrm{E}[S(n)^2]=\frac{1}{n^{2H}L^{2k}(n)}\sum_{j_1,j_2} f(\frac{j_1}{n})f(\frac{j_2}{n})\mathrm{E}K(X_{j_1})K(X_{j_2}).$$

Let $K(X_j)=\sum_{i=1}^{\infty}[K_{i-1}(\tilde{X}_{j,i-1})-K_i(\tilde{X}_{j,i})]$ and suppose $j_1-i_1\neq j_2-i_2$, then we have

$$\mathrm{E}[K_{i_1-1}(\tilde{X}_{j_1,i_1-1})-K_{i_1}(\tilde{X}_{j_1,i_1})][K_{i_2-1}(\tilde{X}_{j_2,i_2-1})-K_{i_2}(\tilde{X}_{j_2,i_2})]=0. \tag{3.24}$$

According to the crucial observation in Ho and Hsing [28]

$$K_i(\tilde{X}_{j,i})=\mathrm{E}\left(K(X_j)\mid \mathcal{F}_{j-i}\right),$$

where $\mathcal{F}_s$ is the σ-field generated by ϵ_k, $k\le s$. Suppose $j_1-i_1\neq j_2-i_2$ and without loss of generality assume that $j_1-i_1>j_2-i_2$, then we have

$$\begin{aligned}
& \mathrm{E}[K_{i_1-1}(\tilde{X}_{j_1,i_1-1}) - K_{i_1}(\tilde{X}_{j_1,i_1})] \\
& \times[K_{i_2-1}(\tilde{X}_{j_2,i_2-1}) - K_{i_2}(\tilde{X}_{j_2,i_2})] \\
= \; & \mathrm{E}[\mathrm{E}(K(x) \mid \mathcal{F}_{j_1-i_1+1}) - \mathrm{E}(K(x) \mid \mathcal{F}_{j_1-i_1})] \\
& \times[\mathrm{E}(K(x) \mid \mathcal{F}_{j_2-i_2+1}) - \mathrm{E}(K(x) \mid \mathcal{F}_{j_2-i_2})] \\
= \; & \mathrm{E}[\mathrm{E}(K(x) \mid \mathcal{F}_{j_1-i_1+1}) - \mathrm{E}(K(x) \mid \mathcal{F}_{j_1-i_1})] \\
& \times\mathrm{E}[[\mathrm{E}(K(x) \mid \mathcal{F}_{j_2-i_2+1}) - \mathrm{E}(K(x) \mid \mathcal{F}_{j_2-i_2})] \mid \mathcal{F}_{j_1-i_1+1}] \\
= \; & \mathrm{E}[\mathrm{E}(K(x) \mid \mathcal{F}_{j_1-i_1+1}) - \mathrm{E}(K(x) \mid \mathcal{F}_{j_1-i_1})] \\
& \times\mathrm{E}[\mathrm{E}(K(x) \mid \mathcal{F}_{j_1-i_1+1}) - \mathrm{E}(K(x) \mid \mathcal{F}_{j_1-i_1+1})] \\
= \; & 0.
\end{aligned}$$

Observe that

$$\begin{aligned}
\mathrm{E}[S(n)^2] &= \frac{1}{n^{2H}L^{2k}(n)} \sum_{j_1,j_2} f(\frac{j_1}{n})f(\frac{j_2}{n})\mathrm{E}K(X_{j_1})K(X_{j_2}) \\
&= \frac{1}{n^{2H}L^{2k}(n)} \sum_{j_1,j_2} f(\frac{j_1}{n})f(\frac{j_2}{n}) \sum_{j_1-i_1=j_2-i_2} \\
& \quad \mathrm{E}\{[K_{i_1-1}(\tilde{X}_{j_1,i_1-1}) - K_{i_1}(\tilde{X}_{j_1,i_1})] \\
& \quad \times[K_{i_2-1}(\tilde{X}_{j_2,i_2-1}) - K_{i_2}(\tilde{X}_{j_2,i_2})]\}.
\end{aligned}$$

Assume $j_2 \geq j_1$, since $[K_{i-1}(\tilde{X}_{j,i-1}) - K_i(\tilde{X}_{j,i})]^2 \leq C(a_{i-1}^2 + a_i^2)$, by the Cauchy-Schwarz inequality

$$\begin{aligned}
& \sum_{j_1-i_1=j_2-i_2} \mathrm{E}[K_{i_1-1}(\tilde{X}_{j_1,i_1-1}) - K_{i_1}(\tilde{X}_{j_1,i_1})] \\
& \times[K_{i_2-1}(\tilde{X}_{j_2,i_2-1}) - K_{i_2}(\tilde{X}_{j_2,i_2})] \\
= \; & \sum_{i_1=1,i_2=i_1+j_2-j_1} \mathrm{E}[K_{i_1-1}(\tilde{X}_{j_1,i_1-1}) - K_{i_1}(\tilde{X}_{j_1,i_1})] \times \\
& [K_{i_1+j_2-j_1-1}(\tilde{X}_{j_2,i_1+j_2-j_1-1}) - K_{i_1+j_2-j_1}(\tilde{X}_{j_2,i_1+j_2-j_1})] \\
\leq \; & C\sum_{i_1=1}^{\infty}[i_1(i_1+j_2-j_1)]^{-\beta+\xi} \\
\leq \; & C|j_2-j_1|^{-\beta+\xi},
\end{aligned}$$

where ξ is a positive constant. Since for any positive constant ξ, we have $n^{2H}L^{2k}(n) \leq Cn^{2H+\xi}$, thus

$$\begin{aligned} \mathrm{E}[S(n)^2] &\leq \frac{C}{n^2}|f(\frac{j_1}{n})||f(\frac{j_2}{n})||\frac{j_1}{n} - \frac{j_2}{n}|^{2H-2} \\ &= C\int_R\int_R |f(u)||f(v)||u-v|^{2H-2}\,du\,dv \\ &= C \parallel f \parallel^2_{|\mathcal{H}|}. \end{aligned}$$

Since the $|\mathcal{H}|$ space is complete, we can choose a sequence f^r, $r \leq 1$, of elementary functions such that $f^r \to f$ as $r \to \infty$ with respect to the norm $|\mathcal{H}|$. Let

$$f^r_n(x) = \sum_{j\in\mathcal{Z}} f^r(\frac{j}{n})1_{(\frac{j}{n},\frac{j+1}{n}]}(x) \qquad r, n \leq 1, \tag{3.25}$$

and

$$S^r(n) \equiv \sum_{j\in\mathcal{Z}} f^r(\frac{j}{n})(K(X_j)/\sigma_n). \tag{3.26}$$

By Theorem 4.2 in Billingsley [6], there are three steps to complete the proof.

Step 1:

$$S^r(n) \to_{\mathcal{L}} \int f^r(u)\,dZ^k_H(u) \quad as\ \ n \to \infty,\ for\ every\ r \geq 1. \tag{3.27}$$

Step 2:

$$\int f^r(u)\,dZ^k_H(u) \to_{\mathcal{L}} \int f(u)\,dZ^k_H(u). \tag{3.28}$$

Step 3:

$$\overline{\lim_{r\to\infty}\lim_{n\to\infty}}\,\mathrm{E}\left[|S^r(n) - S(n)|^2\right] = 0. \tag{3.29}$$

Step 1: The convergence (3.27) follows from the convergence of the sequence $Z^{k,n}_H$ given by (3.21) and (3.22) to the Hermite process Z^k_H because $S^r(n)$ constitutes a finite linear combination of instants of the process $Z_H^{k,n}$ and $\int_{\mathcal{R}} f^r(u)\, dZ_H^k(u)$ represents a finite combination of instants of Z_H^k.

Step 2: Since

$$\begin{aligned}&\mathrm{E}[|\int f^r(s)\, dZ_H^k(s) - \int f(s)\, dZ_H^k(s)|^2] \\ &= \| f^r - f \|_{\mathcal{H}}^2 \leq \| f^r - f \|_{|\mathcal{H}|}^2 \rightarrow 0.\end{aligned}$$

as $m \rightarrow \infty$.

Step 3: By the Dominated Convergence Theorem, we have

$$\begin{aligned}\overline{\lim_{r\to\infty}\lim_{n\to\infty}} \mathrm{E}\left[\mid S^r(n) - S(n) \mid^2\right] &\leq C\overline{\lim_{r\to\infty}\lim_{n\to\infty}} |f_n^r - f_n|_{|\mathcal{H}|}^2 \\ &= C \lim_{r\to\infty} |f^r - f|_{|\mathcal{H}|}^2 = 0\end{aligned}$$

This concludes the proof. □

The following example illustrates an application of Theorem 3.1.

Example 3.2. *Let $f(s) = e^{\gamma(t-s)} 1_{[0,t]}(s)$. It can be verified that $f(s)$ belongs to $L^1(\mathcal{R}) \cap L^2(\mathcal{R})$, thus by (3.11), $f(s) \in |\mathcal{H}|$. Furthermore we have $\{f_n(x)\}_{n\geq 1}$ satisfies conditions in Theorem 3.1, thus we have*

$$\frac{1}{n^H L^k(n)} \sum_j f(\frac{j}{n}) K(X_j) \rightarrow_{\mathcal{L}} \int_0^t e^{-\gamma(t-s)}\, dZ_H^k(s). \tag{3.30}$$

where $\int_0^t e^{-\gamma(t-s)}\, dZ_H^k(s)$ is a Hermite Ornstein-Uhlenbeck process. As a consequence, at each t, the Hermite Ornstein-Uhlenbeck process can be approximated in law by the partial sum in the left side of (3.30).

□ **End of chapter.**

Chapter 4

Auxiliary Results II

In this chapter, we investigate the asymptotic theory of the least squares estimator for nearly unstable processes with a functional of long-memory noises.

4.1 Least Squares Estimator (LSE)

Consider the nearly unstable $AR(1)$ model proposed by Chan and Wei [12],

$$y_{t,n} = \rho_n y_{t-1,n} + \epsilon_t, \tag{4.1}$$

where $\rho_n = 1 - \gamma/n$ and γ is a fixed constant. Let

$$\hat{\rho}_n = \frac{\sum_{t=1}^{n} y_{t-1,n} y_{t,n}}{\sum_{t=1}^{n} y_{t-1,n}^2} \tag{4.2}$$

denote the LSE of ρ_n. When $\{\epsilon_t\}$ is a sequence of independent standard normal random variables, Chan and Wei [12] showed that

$$n(\hat{\rho}_n - \rho_n) \to_{\mathcal{L}} \frac{\int_0^1 Z_{1/2,\gamma}^1(t)\, dZ_{1/2,\gamma}^1}{\int_0^1 (Z_{1/2,\gamma}^1(t))^2\, dt}, \tag{4.3}$$

where $Z_{1/2,\gamma}^1(t)$ is the Ornstein-Uhlenbeck process driven by a standard Brownian motion. When $\gamma = 0$, this model reduces to the $AR(1)$

model, that is,

$$y_t = y_{t-1} + \epsilon_t, \quad \text{for } t \in N, \ y_0 = 0. \tag{4.4}$$

White [60] showed that

$$n(\hat{\rho}_n - 1) \to_{\mathcal{L}} \frac{\int_0^1 B(s)\, dB(s)}{\int_0^1 (B(s))^2\, ds}. \tag{4.5}$$

This particular result is encompassed in (4.3). These results for the nearly unstable AR(1) model are somewhat surprisingly different from those for the strictly stationary or explosive $AR(1)$ model given by Mann and Wald [40] and Anderson [1]. In those papers, it was shown that the limit distribution of the normalized LSE of the autoregressive coefficient is standard normal.

Now, consider the following model. Let

$$y_{t,n} = \rho_n y_{t-1,n} + K(\epsilon_t), \quad \text{for } t = 1, \ldots, n, \tag{4.6}$$

where $\rho_n = 1 - \gamma/n$ and γ is a fixed constant, and K is a function that satisfies certain conditions. Obviously, model (4.6) is a generalization of model (4.1). When $K(x) = x$, (4.6) reduces to (4.1).

In this chapter, we focus on the case in which $\{\epsilon_t\}$ is a moving average process, that is,

$$\epsilon_t = \sum_{i=0}^{\infty} a_i e_{t-i}, \tag{4.7}$$

where $\sum_{i=0}^{\infty} a_i^2 < \infty$ and $\{e_i\}$ is a sequence of i.i.d. mean zero random variables with finite second moments. For the case in which $K(x) = x$ and $\{\epsilon_t\}$ exhibits strong dependence, Buchmann and Chan [8] showed that

$$n(\hat{\rho}_n - \rho_n) \to_{\mathcal{L}} \frac{\int_0^1 Z_{H,\gamma}^1(t)\, dZ_{H,\gamma}^1}{\int_0^1 (Z_{H,\gamma}^1(t))^2\, dt}, \tag{4.8}$$

where $Z_{H,\gamma}^1(t)$ is the Ornstein-Uhlenbeck process driven by a fBm.

For a case in which $K(x)$ takes the more general form, Wu [62] investigated the unit root testing problem (when $\gamma = 0$, model (4.6) is reduced to a unit root AR(1) model). When $\{\epsilon_t\}$ is generated by moving average processes, Wu [62] shows that

$$n(\hat{\rho}_n - 1) \rightarrow_{\mathcal{L}} \frac{\frac{1}{2}(Z^1_{H,0}(1))^2}{\int_0^1 (Z^1_{H,0}(u))^2\, du}, \tag{4.9}$$

where $Z^1_{H,0}$ denotes the fBm. A natural extension would be to extend Wu's result to model (4.10), which is the main theme of this chapter. It is shown that the limit distribution of the LSE is expressed as the functionals of the Ornstein-Uhlenbeck processes driven by the Hermite processes.

4.2 Limit Distributions

Definition 4.1. *Let $\rho_n = 1 - \gamma/n$, for $t = 1, \ldots, n$. Suppose that $y_{t,n}$ satisfies the reparameterized AR(1) model,*

$$y_{t,n} = \rho_n y_{t-1,n} + x_t \quad y_{0,n} = 0, \ \textit{for all } n, \tag{4.10}$$

and

$$x_t = K(\epsilon_t), \qquad \epsilon_t = \sum_{i=0}^{\infty} a_i e_{t-i}. \tag{4.11}$$

The coefficients a_i satisfy $\sum_{i=1}^{\infty} a_i^2 < \infty$, $a_0 = 1$ and $a_k = k^{-\beta} L_0(k)$ for some slowly varying function L_0 with $H < 1$, and $\{e_t\}$ is a sequence of *i.i.d.* mean zero Gaussian random variables with variance σ^2. The process $\{\epsilon_t\}$ exhibits a long-memory phenomenon when $\beta \in (\frac{1}{2}, 1)$. K is the function satisfying $\mathrm{E}K^2(X_1) < \infty$ and $\mathrm{E}K(X_1) = 0$.

Let $H = 1 - k(\beta - \frac{1}{2})$ and $\sigma_n = c(H,k) n^H L^k(n)$. We have the following theorem.

Theorem 4.1. *For the model (4.10), (i) if $k(2\beta - 1) < 1$, then under condition 2 we have*

$$\{y_{[nu]}/\sigma_n, 0 \le u \le 1\} \Rightarrow \{K_\infty^{(k)}(0) Z_{H,\gamma}^k(u), 0 \le u \le 1\} \tag{4.12}$$

where $Z_{H,\gamma}^k$ is a Hermite Ornstein-Uhlenbeck process.

(ii) If $k(2\beta - 1) > 1$ or $k(2\beta - 1) = 1$ and $\sum_{n=1}^{\infty} |L^k(n)|/n < \infty$, then

$$y_{[nu]}/\sqrt{n} \to_{\mathcal{L}} X(u) = \sigma \int_0^u e^{-\gamma(u-s)}\, dB(s), \tag{4.13}$$

where $B(s)$ denotes the standard Brownian motion and $\sigma^2 < \infty$ is a constant given by

$$\sigma^2 = \lim_{n\to\infty} n^{-1} \mathrm{Var}(\sum_{t=1}^{n} x_t). \tag{4.14}$$

.

Proof. (i) Let $f_u(s) = e^{-\gamma(u-s)}$ and $f_{n,u}(s) = (1 - \frac{\gamma}{n})^{[nu]-[ns]}$. Obviously, $f_{n,u}(s) \to f_u(s)$.

$$\begin{aligned} \frac{y_{[nu]}}{\sigma_n} &= \sum_{j=1}^{[nu]} f_{n,u}(\frac{j}{n})(x_j/\sigma_n) - \sum_{j=1}^{[nu]} f_u(\frac{j}{n})(x_j/\sigma_n) \\ &\quad + \sum_{j=1}^{[nu]} f_u(\frac{j}{n})(x_j/\sigma_n) \\ &= M_{1,n} + M_{2,n}, \end{aligned} \tag{4.15}$$

where

$$M_{1,n} = \sum_{j=1}^{[nu]} f_{n,u}(\frac{j}{n})(x_i/\sigma_n) - \sum_{j=1}^{[nu]} f_u(\frac{j}{n})(x_i/\sigma_n) \tag{4.16}$$

and

$$M_{2,n} = \sum_{j=1}^{[nu]} f_u(\frac{j}{n})(x_i/\sigma_n). \tag{4.17}$$

Now let us prove $|M_{1n}| = o_p(1)$. Let $d_n = (n/\gamma)\log(1-\gamma/n)$. For fixed γ, $d_n \to -1$ as $n \to \infty$. For any $s, u (0 \le s \le u \le 1)$, we have

$$\begin{aligned} |f_{n,u}(s) - f_u(s)| &= |(1-\frac{\gamma}{n})^{[nu]-[ns]} - e^{-\gamma(u-s)}| \\ &= |e^{(\frac{[nu]}{n}-\frac{[ns]}{n})\gamma d_n} - e^{-\gamma(u-s)}| \\ &= |e^{-\gamma(u-s)}e^{O(\frac{1}{n})} - e^{-\gamma(u-s)}| \\ &= O(\frac{1}{n}). \end{aligned} \tag{4.18}$$

Let $T_n = \sup_{s,u} |f_{n,u}(s) - f_u(s)|$, we have $T_n = O(\frac{1}{n})$.

$$\begin{aligned} \mathrm{E}|M_{1n}| &= \mathrm{E}|\sum_{j=1}^{[nu]} f_{n,u}(\frac{j}{n})(x_j/\sigma_n) - \sum_{j=1}^{[nu]} f_u(\frac{j}{n})(x_j/\sigma_n)| \\ &\le CT_n[\sum_j \mathrm{E}|Z_H^k(\frac{j}{n}) - Z_H^k(\frac{j-1}{n})|]. \end{aligned} \tag{4.19}$$

By the Hölder continuous path property, for $p \ge 1$ and $H > 1$, we have

$$\mathrm{E}|Z_H^k(\frac{j}{n}) - Z_H^k(\frac{j-1}{n})|^p < Cn^{-H}. \tag{4.20}$$

where C is a constant. Combining (4.18), (4.19) and (4.20), we have $|M_{1n}| = o_p(1)$. For M_{2n}, we see that $f_u(s) = e^{-\gamma(u-s)}1_{[0,u]}$ belongs to $L^1(\mathcal{R}) \cap L^2(\mathcal{R})$ and satisfies conditions of Theorem 3.1. Thus,

$$\sum_{j=1}^{[nu]} f_u(\frac{j}{n})(x_i/\sigma_n) \to_{\mathcal{L}} K_\infty^{(k)}(0) \int_0^u e^{-\gamma(u-s)}\, dZ_H^k(s).$$

The finite-dimensional convergence of $y_{[nu]}/\sigma_n$ to $Z_{H,\gamma}^k(u)$ hold. It remains to prove the tightness. By Theorem 12.3 in Billingsley [6], it suffices to show that for all m there exist $C < \infty$ and $\alpha > 1$ such that

$$\frac{\mathrm{E}(y_{m,n})^2}{\sigma_n^2} \le C\frac{m^\alpha}{n^\alpha}. \tag{4.21}$$

For any $H \in (1/2, 1)$, there is a positive δ such that $\alpha = 2H - \delta > 1$. Because $\mathrm{E}(y_{m,n})^2 = O(m^{2H}L^{2k}(m))$ and $\sigma_n^2 = O(n^{2H}L^{2k}(u))$. Thus by the elementary properties of slowly varying functions, we have

$$\lim_{n\to\infty}\max_{m\leq n}\frac{\mathrm{E}(y_{m,n})^2}{\sigma_n^2}=\lim_{n\to\infty}\max_{m\leq n}C\frac{n^{\alpha-2H}m^{2H-\alpha}L^{2k}(m)}{L^{2k}(n)}=1, \quad (4.22)$$

which completes the proof of part (i).

(ii) We follow the similar argument in part (i) and note that $f_u(s)=e^{-\gamma(u-s)}1_{[0,u]}$ belongs to $L^2(\mathcal{R})$, by theorem 3.1, we can conclude the result. □

Some useful limit distributions are given in the following theorem.

Theorem 4.2. *For the model (4.10), (i) if $k(2\beta-1)<\beta$, then under condition 2 we have*

(i) $n^{-1}\sum_{t=1}^n y_{t-1,n}^2 \to_{\mathcal{L}} \int_0^1 (K_\infty^{(k)}(0)Z_{H,\gamma}^k(u))^2\,du$;

(ii) $n^{-1}\sum_{t=1}^n y_{t-1,n} \to_{\mathcal{L}} \int_0^1 K_\infty^{(k)}(0)Z_{H,\gamma}^k(u)\,du$;

(iii) $\sum_{t=1}^n y_{t,n}x_t/\sigma_n^2 \to_{\mathcal{L}} \frac{1}{2}(K_\infty^{(k)}(0)Z_{H,\gamma}^k(u))^2+\gamma\int_0^1(K_\infty^{(k)}(0)Z_{H,\gamma}^k(u))^2\,du$.

Proof. (i) and (ii) follow from the continuous mapping theorem together with Theorem 4.1.

To prove (iii), squaring and summing (4.10), we decompose the numerators into three terms which we analyze separately in the sequel, that is, for all $n\in N$,

$$\sum_{t=1}^n y_{t-1,n}x_t=\frac{1-\rho_n^2}{2\rho_n}\sum_{t=1}^n y_{t-1,n}^2+\frac{1}{2\rho_n}y_{n,n}^2-\frac{1}{2\rho_n}\sum_{t=1}^n x_t^2. \quad (4.23)$$

We define the auxiliary random variables

$$T_{1,n}=\frac{1}{\sigma_n^2}\sum_{t=1}^n y_{t-1,n}\epsilon_t=\frac{n(1-\rho_n^2)}{2\rho_n}T_{2,n}+\frac{1}{2\rho_n}(T_{3,n}^2)-\frac{1}{2\rho_n}T_{4,n}, \quad (4.24)$$

where

$$T_{2,n}=\frac{1}{n\sigma_n^2}\sum_{t=1}^n y_{t-1,n}^2,\quad T_{3,n}=\frac{y_{n,n}}{\sigma_n},\quad T_{4,n}=\frac{1}{\sigma_n^2}\sum_{t=1}^n x_t^2. \quad (4.25)$$

As the function $f \to (f(1), \int_0^1 f^2\,ds)$ is a continuous function, as $n \to \infty$,

$$\begin{pmatrix} T_{2,n} \\ T_{3,n} \end{pmatrix} = \begin{pmatrix} \int_0^1 (y_{[ns]}/\sigma_n)^2\,ds \\ y_n/\sigma_n \end{pmatrix} \to_{\mathcal{L}} \begin{pmatrix} \int_0^1 K_\infty^{(k)}(0) Z_{H,\gamma}^k(s)\,ds \\ K_\infty^{(k)}(0) Z_{H,\gamma}^k(1) \end{pmatrix} \tag{4.26}$$

For the part $T_{4,n} = \frac{1}{\sigma_n^2}\sum_{t=1}^n x_t^2$, we have $\mathrm{E}[\sum_{t=1}^n x_t^2] = O(n)$ and $\sigma_n = n^H L^k(n)(H > 1/2)$, hence we have $T_{4,n} = \frac{1}{\sigma_n^2}\sum_{t=1}^n x_t^2 = o_p(1)$. Thus, (iii) is proved. □

Now we are in the position to study the limit distribution of the least squares estimators. Let $\hat\rho_n$ denote the least square estimator of ρ_n, given by

$$\hat\rho_n = \sum_{t=1}^n y_{t-1,n} y_{t,n} \Big/ \sum_{t=1}^n y_{t-1,n}^2. \tag{4.27}$$

Let

$$\hat\rho_{1n} = n(\hat\rho_n - \rho_n) \tag{4.28}$$

and

$$\hat\rho_{2n} = \sqrt{n}(\sum_{t=1}^n y_{t-1,n}^2)^{1/2}(\hat\rho_n - \rho_n). \tag{4.29}$$

Let

$$\Theta(Z_{H,\gamma}^k) = \begin{pmatrix} (\int_0^1 (K_\infty^{(k)}(0) Z_{H,\gamma}^k(u))^2\,du)^{-1} \\ (\int_0^1 (K_\infty^{(k)}(0) Z_{H,\gamma}^k(u))^2\,du)^{-1/2} \end{pmatrix}$$

and

$$\Theta(X) = \begin{pmatrix} (\int_0^1 X(u)^2\,du)^{-1} \\ (\int_0^1 X(u)^2\,du)^{-1/2} \end{pmatrix}.$$

We have the following theorem.

Theorem 4.3. *(i) If $k(2\beta - 1) < \beta$ and condition (2) holds, then for $n \to \infty$*

$$\begin{pmatrix} \hat{\rho}_{1n} \\ \\ \hat{\rho}_{2n} \end{pmatrix} \to_{\mathcal{L}} (\frac{1}{2}(K_{\infty}^{(k)}(0) Z_{H,\gamma}^{k}(1))^2 + \gamma \int_0^1 (K_{\infty}^{(k)}(0) Z_{H,\gamma}^{k}(u))^2 \, du) \Theta(Z_{H,\gamma}^{k})$$

(ii) If $k(2\beta - 1) > 1$ or $k(2\beta - 1) = 1$ and $\sum_{n=1}^{\infty} |L^k(n)|/n < \infty$, then for $n \to \infty$

$$\begin{pmatrix} \hat{\rho}_{1n} \\ \\ \hat{\rho}_{2n} \end{pmatrix} \to_{\mathcal{L}} \left(\frac{1}{2} X(1)^2 + \gamma \int_0^1 X(u)^2 \, du \right) \Theta(X)$$

Remark 4.1. *The term $(\frac{1}{2}(Z_{H,\gamma}^{k}(1))^2 + \gamma \int_0^1 (Z_{H,\gamma}^{k}(u))^2 \, du)$ can be re-written as (4.30) in terms of stochastic integral*

$$\int_0^1 Z_{H,\gamma}^{k}(s) \, dZ_{H}^{k}(s). \tag{4.30}$$

This type integral is well-defined in the Riemann-Stieltjes sense.

Remark 4.2. *For $K(x) = x$, one can verify that $k = 1$. Thus we have*

$$n(\hat{\rho}_n - \rho_n) \to_{\mathcal{L}} \frac{\frac{1}{2}(Z_{H,\gamma}^{1}(1))^2 + \gamma \int_0^1 (Z_{H,\gamma}^{1}(u))^2 \, du}{\int_0^1 (Z_{H,\gamma}^{1}(u))^2 \, du}, \tag{4.31}$$

where $Z_{H,\gamma}^{1}$ is the fractional Brownian motion. This result agrees with Theorem 2.1(i) of Buchmann and Chan [8].

Remark 4.3. *For $\gamma = 0$, model (4.10) reduces to the unit root $AR(1)$ model proposed by Wu [62]. By Theorem 4.3, we have*

$$n(\hat{\rho}_n - \rho_n) \to_{\mathcal{L}} \frac{\frac{1}{2}(Z_{H}^{k}(1))^2}{\int_0^1 (Z_{H}^{k}(u))^2 \, du}. \tag{4.32}$$

This particular result encompasses the unit root case of Wu [62].

Proof. Note that

$$\hat{\rho}_{1n} = n\left(\sum_{t=1}^{n} y_{t-1,n}\epsilon_t \Big/ \sum_{t=1}^{n} y_{t-1,n}^2\right) \tag{4.33}$$

and

$$\hat{\rho}_{2n} = \frac{\sum_{t=1}^{n} y_{t-1,n}\epsilon_t}{\frac{1}{n}\sum_{t=1}^{n} y_{t-1,n}^2}. \tag{4.34}$$

Thus by theorem 4.2 together with the continuous mapping theorem, we have

$$\begin{aligned}
\begin{pmatrix} \hat{\rho}_{1n} \\ \hat{\rho}_{2n} \end{pmatrix} &= T_{1,n} \begin{pmatrix} (T_{2,n})^{-1} \\ (T_{1,n})^{-1/2} \end{pmatrix} \\
&\to_{\mathcal{L}} \left(\frac{1}{2}(K_{\infty}^{(k)}(0)Z_{H,\gamma}^{k}(1))^2 + \gamma \int_0^1 (K_{\infty}^{(k)}(0)Z_{H,\gamma}^{k}(u))^2\, du\right) \\
&\quad \times \begin{pmatrix} \left(\int_0^1 (K_{\infty}^{(k)}(0)Z_{H,\gamma}^{k}(u))^2\, du\right)^{-1} \\ (\int_0^1 (K_{\infty}^{(k)}(0)Z_{H,\gamma}^{k}(u))^2\, du)^{-1/2} \end{pmatrix} \\
&= \left(\frac{1}{2}(K_{\infty}^{(k)}(0)Z_{H,\gamma}^{k}(1))^2 \right. \\
&\quad \left. +\gamma \int_0^1 (K_{\infty}^{(k)}(0)Z_{H,\gamma}^{k}(u))^2\, du\right) \Theta(Z_{H,\gamma}^{k}),
\end{aligned}$$

where $\Theta(Z_{H,\gamma}^{k})$ is the vector defined in (4.30). This concludes the result.

□

□ **End of chapter.**

Chapter 5

Conclusions and Directions for Further Research

Based on the analytical and experimental investigations presented in this thesis, the following conclusions are reached.

(1) The expansion of the residual empirical processes in Chan and Ling [10] can be extended to a nearly unstable long-memory time series. Moreover, the statistics proposed by Chan and Ling can be used to test for the distribution of long-memory noises. The limit distribution is expressed as a functional of an Ornstein-Uhlenbeck process driven by a fractional Brownian motion.

(2) The results of a number of simulation studies have been presented in this thesis along with the critical values of the proposed test. The performance of this test has also been examined. As shown in Section 2.3, the proposed test has satisfactory size and power behavior in finite samples and should be useful in practice.

(3) The weak convergence of the weighted sums of random variables that are the functionals of moving average processes have also

been considered. To this end, a non-central limit theorem is established, in which the Wiener integrals with respect to the Hermite processes appear as the limit.

(4) The asymptotic theory of the LSE for nearly unstable AR(1) model when the innovation sequences are functionals of moving average processes has been investigated. It is shown that the limit distribution of the LSE is expressed as the functionals of Ornstein-Uhlenbeck processes driven by Hermite processes.

A number of issues remain to be addressed in further research.

(a) In Section 2.3.1., on the basis of Theorem 2.2, the sample paths of the fractional Ornstein-Uhlenbeck processes were simulated by the linear process on the left side of equation (2.20). In these simulations, the truncation scheme proposed by Wu, Michailidis and Zhang [63] was adopted to deal with the infinite number of terms in equation (2.20). This scheme could also be extended to simulate the sample paths of the Hermite Ornstein-Uhlenbeck processes on the basis of Theorem 4.1. Here, a proper function K could be chosen. The sample paths of the Hermite Ornstein-Uhlenbeck processes could be simulated by the linear process on the left side of equation (4.12). The accuracy of these simulations could also be estimated. However, these issues remain to be pursued in further detail.

(b) As the proposed test is applied in practice, the parameter H needs to be estimated. However, the asymptotic validity of such a procedure remains to be examined. To deal with this problem, Durbin [21] suggested a half-sample device in which the innovation sequence $\{\epsilon_t\}$ is i.i.d.. In this device, the unknown parameter vector is estimated from a randomly-chosen half-sample of data and treated as the true parameter. Durbin's study showed that the limit distribution of the Kolmogorov-Smirnov tests so obtained was the same as if the values of the parameters were known. Unfortunately, this device is not valid when $\{\epsilon_t\}$ is dependent. A device for a nearly unstable long-memory time series remains to be proposed.

(c) However, the limit distribution of the Kolmogorov-Smirnov test based on the residuals generally depends on the underlying model parameters, which means that different critical values are needed for different parameter values. To overcome this dependence on the underlying model parameters, Koul and Ling [33] proposed the weighted residual empirical processes. They investigated the tests based on a vector of weighted residual empirical processes for some heteroscedastic time series and showed that the limit distribution of these tests depends only on the fitted distribution, not on the model. A natural extension would be to apply the weighted residual empirical processes in the nearly unstable long-memory time series to derive the Kolmogorov-Smirnov statistics that do not depend on the model.

Bibliography

[1] Anderson, T.W. (1959). On asymptotic distribution of estimates of parameters of stochastic difference equations. *Ann. Math. Statist.* **30**, 676–687.

[2] Avram, F. and Taqqu, M.S. (1987). Noncentral limit theorems and Appell polynomials. *Annals of Probability* **15**, 767–775.

[3] Bai, J. (1994). Weak convergence of the sequential empirical processes of residuals in ARMA models. *Ann. Statist.* **22**, 2051–2061.

[4] Bai, J. (1996). Testing for parameter constancy in linear regressions: an empirical distribution function approach. *Econometrica* **64**, 597–622.

[5] Bai, J. (2003). Testing parametric conditional distribution of dynamic models. *Review of Economics and Statistics* **85**, 531–549.

[6] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[7] Brockwell, P. and Davis, R. A. (1991). *Time Series: Theory and Methods, 2nd ed.* Springer, New York.

[8] Buchmann, B. and Chan, N.H. (2007). Asymptotic theory of least squares estimators for nearly unstable processes under strong dependence. *Ann. Statist.* **35**: 2001–2017.

[9] Beran, J. (1994). *Statistics for Long-memory Processes*. Chapman and Hall, New York.

[10] Chan, N. H. and Ling, S. (2008). Residual empirical processes for long and short memory. *Ann. Statist.* **36**. 2453–2470.

[11] Chan, N. H. and Terrin, N. (1995). Inference for unstable long-memory process with applications to fractional unit root autoregressions. *Ann. Statist.* **23**, 1662–1683.

[12] Chan, N. H. and Wei, C.Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.* **15**, 1050–1063.

[13] Chan, N. H. and Wei, C.Z. (1988). Limiting distribution of least squares estimates of unstable autoregressive processes. *Ann. Statist.* **16**, 367–401.

[14] Cheridito, P., Kawaguchi, H. and Maejima, M. (2003). Fractional Ornstein-Uhlenbeck processes. *Elect. J. Prob.* **8**, 1–14.

[15] Davidson, J. (2002). Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. *J. Econometrics* **106**, 243–269.

[16] Davydov, Yu. A. (1970). The invariance principle for stationary processes. *Theory Probab. Appl.* **15**, 487–498.

[17] Decreusefond, L. and Üstünel, A.S. (1999). Stochastic analysis for the fractional Brownian motion. *Potential Analysis* **10**, 177–214.

[18] Dehling, H. and Taqqu, M.S. (1989). The empirical processes of some long-range dependent sequences with an application to U-statistics. *Ann. Statist.* **17**, 1767–1783.

[19] Dietrich, C. R. and Newsam, G.N. (1997). Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. *SIAM J. Scient. Comput.* **18**, 1088–1107.

[20] Dudley, R. M. and Norvaiša, R. (1999). Concentrated advance course, *An Introduction to p-variation and Young Integrals.* Maphysto, Centre for Mathematical Physics and Stochastics, University of Aarhus, Denmark.

[21] Durbin, J. (1973). *Distribution theory for tests based on the sample distribution function.* Philadephia: Society of Industrial and Applied Mathematics.

[22] Durbin, J. (1973). Weak convergence of the sample distribution function when parameters are estimated. *Ann. Statist.* **1**, 279–290.

[23] Durbin, J. (1976). Kolmogorov-Smirnov tests when parameters are estimated. In Gaenssler, P. and Révész, P. (eds.), *Empirical Distributions and Processes, Lecture Notes in Mathematics* **566**. Springer-Verlag, Berlin.

[24] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* **2**, 2nd ed. Wiley, New York.

[25] Giraitis, L., Koul, H. L. and Surgailis, D. (1996). Asymptotic normality of regression estimators with long-memory errors. *Statistics and Probability Letters* **29**, 317–335.

[26] Hall, P. and Heyde, C. C. (1980). *Maringale Limit Theory and Its Application.* Academic Press, New York.

[27] Ho, H. C. and Hsing, T. (1996). On the asymptotic expansion of the empirical process of long-memory moving averages. *Ann. Statist.* **24**, 992–1024.

[28] Ho, H. C. and Hsing, T. (1997). Limit theorems for functionals of moving averages. *Ann. Statist.* **25**, 1636–1669.

[29] Jeganathan, P. (1991). On the asymptotic behavior of least-squares estimators in AR time series with roots near the unit circle. *Econometric Theory* **7**, 269–306.

[30] Jeganathan, P. (1999). On asymptotic inference in cointegrated time series with fractionally integrated errors. *Econometric Theory* **15**, 583–621.

[31] Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent R.V.'s and the sample DF. *Z. Wahrsch. Verw. Gebiete* **32**, 111–131.

[32] Koul, H. L. (2002). *Weighted Empirical Processes in Dynamic Nonlinear Models. 2nd ed.*. Lecture Notes in Statistics, **166**. Springer-Verlag, New York.

[33] Koul, H. L. and Ling, S. (2006). Fitting an error distribution in some heteroscedastic time series models. *Ann. Statist.* **34**, 994–1012.

[34] Koul, H. L. and Surgailis, D. (1997). Asymptotic expansion of M-estimators with long-memory errors. *Ann. Statist.* **25**, 818–850.

[35] Koul, H. L. and Surgailis, D. (2001). Asymptotic of the empirical processes of long-memory moving averages with infinite variance. *Stochastic Processes and Their Applications* **91**, 309–336.

[36] Lee, S. and Wei, C. Z. (1999). On residual empirical processes of stochastic regression models with applications to time series. *Ann. Statist.* **27**, 237–261.

[37] Ling, S. (1998). Weak convergence of the sequential empirical processes of residuals in nonstationary autoregressive models. *Ann. Statist.* **26**, 741–754.

[38] Ling, S. and Li, W. K. (1998). Limiting distributions of maximum likelihood estimators for unstable autoregressive moving-average time series with general autoregressive heteroscedastic errors. *Ann. Statist.* **26**, 84–125.

[39] Maejima, Makoto and Tudor, Ciprian A. (2007). Wiener integrals with respect to the Hermite process and a non-central limit theorem. *Stochastic Analysis and Applications* **25**, 1043–1056.

[40] Mann, H.B. and Wald, A. (1943). On the statistical treatment of linear stochastic difference equations. *Econometrica* **11**, 173–220.

[41] Nualart, D. (1995) *Malliavin Calculus and Related Topics.* Springer.

[42] Pipiras, V. and Taqqu, M.S. (2000). Convergence of weighted sums of random variables with long range dependence. *Stochastic Processes and their Applications* **90**, 157–174.

[43] Pipiras, V. and Taqqu, M.S. (2000). Integration questions related to fractional Brownian motion. *Probability Theory and Related Fields* **118**, 251–291.

[44] Rao, M. M. (1978). Asymptotic distribution of an estimator of the boundary parameter of an unstable process. *Ann. Statist.* **6**, 185–190.

[45] Robinson, P.M. (1994). Semiparametric analysis of long-memory time series. *Ann. Statist.* **22**, 515–539.

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

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

[48] Robinson, P.M. (2003). *Time Series with Long Memory.* Oxford Univ. Press., Oxford.

[49] Robinson, P.M. (2005). Efficiency improvements in inference on stationary and nonstationary fractional time series. *Ann. Statist.* **33**, 1800–1842.

[50] Robinson, P.M. (2005). The distance between rival nonstationary fractional processes. *J. Econometrics* **128**, 283–300.

[51] Robinson, P.M. and Hidalgo, F.J. (1997). Time series regression with long-range dependence. *Ann. Statist.* **25**, 77–104.

[52] Samorodnitsky, G. and Taqqu, M.S. (1994). *Stable Non-Gaussian Random processes: Stochastic Models with Infinite Variance.* Chapman and Hall, New York.

[53] Sowell, F. (1990). Fractional unit root distribution. *Econometrica* **50**, 495–505.

[54] Surgailis, D. (1982). Zone of attraction of self-similar multiple integrals. *Lithuanian Mathematical Journal* **22**, 327–340.

[55] Surgailis, D. (2000). Long-range dependence and Appell rank. *Annals of Probability* **28**, 478–497.

[56] Taqqu, M.S. (1975). Weak convergence to fractional Brownian motion and to the Rsenblatt process. *Z. Wahrsch. Verw. Gebiete* **31**, 287–302.

[57] Taqqu, M.S. (1979). Convergence of integrated processes of arbitrary Hermite rank. *Z. Wahrsch. Verw. Gebiete* **50**, 53–83.

[58] Truong-Van, B. and Larramendy, I. (1996). Asymptotic distribution of least squares estimators for purely unstable ARMA(m, ∞). *Statistics* **28**, 307–346.

[59] Wang, Q., Lin, Y.X. and Gulati, C.M. (2003). Asymptotics for general fractionally integrated processes with applications to unit root tests. *Econometric Theory* **19**, 143–164.

[60] White, J.S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Ann. Math. Statist.* **29**, 1188–1197.

[61] Wu, W. (2003). Empirical processes of long-memory sequences. *Bernoulli* **9**, 809–831.

[62] Wu, W. (2006). Unit root testing for functional of linear processes. *Econometric Theory* **22**, 1–14.

[63] Wu, W., Michailidis, G. and Zhang, D. (2004). Simulating sample paths of linear fractional stable motion. *IEEE Transactions on information theory* **50**, 1086–1096.

[64] Wu, W.B. and Min, W. (2005). On linear processes with dependent innovations. *Stochastic Processes and Their Applications* **115**, 939–958.

[65] Young, L.C. (1936). An inequality of the Hölder type, connected with Stieltjes integration. *Acta Math.* **67**, 251–282.

[66] Zähle, M. (1998). Integration with respect to fractal functions and stochastic calculus. *Probability Theory and Related Fields* **111**, 333–374.