isid/ms/2003/24
September 11, 2003
http://www.isid.ac.in/~statmath/eprints

Self-Similar Processes, Fractional Brownian Motion and Statistical Inference

B. L. S. Prakasa Rao

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

Self-Similar Processes, Fractional Brownian Motion and Statistical Inference

B.L.S. PRAKASA RAO
INDIAN STATISTICAL INSTITUTE, NEW DELHI

Abstract

Self-similar stochastic processes are used for stochastic modeling whenever it is expected that long range dependence may be present in the phenomenon under consideration. After discusing some basic concepts of self-similar processes and fractional Brownian motion, we review some recent work on parametric and nonparametric inference for estimation of parameters for linear systems of stochastic differential equations driven by a fractional Brownian motion.

Keywords and phrases: Self-similar process; fractional Brownian motion; fractional Ornstein-Uhlenbeck type process; Girsanov-type theorem; Maximum likelihood estimation; Bayes estimation; Nonparametric inference; Linear stochastic systems.

AMS Subject classification (2000): Primary 62M09, Secondary 60G15.

1 Introduction

"Asymptotic Distributions in Some Nonregular Statistical Problems" was the topic of my Ph.D. Dissertation prepared under the guidance of Prof. Herman Rubin at Michigan State University in 1966. One of the nonregular problems studied in the dissertation was the problem of estimation of the location of cusp of a continuous density. The approach adapted was to study the limiting distribution if any of the log-likelihood ratio process and then obtain the asymptotic properies of the maximum likelihood estimator. It turned out that the limiting process is a nonstationary gaussian process . The name fractional Brownian motion was not in vogue in those years and the limiting process is nothing but a functional shift of a fractional Brownian motion. Details of these results are given in Prakasa Rao (1966) or Prakasa Rao (1968). The other nonregular problems discussed in the dissertation dealt with inference under order restrictions where in it was shown that, for the existence of the limiting distribution if any of the nonparametric maximum likelihood density estimators under order restrictions such as unimodality of the density function or monotonicity of the failure rate function, one needs to scale the estimator by the cube root of n, the sample size rather than the square root of n as in the classical parametric inference (cf. Prakasa Rao (1969, 1970). These type of asymptotics are presently known as cube root asymptotics in the literature. It gives me a great pleasure to contribute this paper to the festschrift in honour of my "guruvu" Prof. Herman Rubin. A short review of some properties of self-similar processes is given in the Section 2. Stochastic differential equations driven by a fractional Brownian motion (fBm) are introduced in the Section 3. Asymptotic properties of the maximum likelihood estimators and the Bayes estimators for parameters involed in linear stochastic differential equations driven by a fBm with a known Hurst index are reviewed in the Section 4. Methods for statistical inference such as the maximum likelihood estimation and the sequential maximum likelihood estimation are discussed for the special case of the fractional Ornstein-Uhlenbeck type process and some new results on the method of minimum L_1-norm estimation are presented in the Section 5. Identification or nonparametric estimation of the "drift" function for linear stochastic systems driven by a fBm are studied in the Section 6.

2 Self-similar processes

Long range dependence phenomenon is said to occur in a stationary time series $\{X_n, n \geq 0\}$ if the $Cov(X_0, X_n)$ of the time series tends to zero as $n \to \infty$ and yet it satisfies the condition

$$\sum_{n=0}^{\infty} |Cov(X_0, X_n)| = \infty. \tag{2. 1}$$

In other words the covariance between X_0 and X_n tends to zero but so slowly that their sums diverge. This phenonmenon was first observed by the hydrologist Hurst (1951) on projects involving the design of reservoirs alnog the Nile river (cf. Montanari (2003)) and by others in hydrological time series. It was recently observed that a similar phenomenon occurs in problems connected with traffic patterns of packet flows in high speed data net works such as the internet (cf. Willinger et al. (2003) and Norros (2003)). Long range dependence is also related to the concept of self-similarity for a stochastic process in that the increments of a self-similar process with stationary increments exhibit long range dependence. Long range dependence pattern is also observed in macroeconomics and finance (cf. Henry and Zafforoni (2003)). A recent monograph by Doukhan et al. (2003) discusses the theory and applications of long range dependence.

A real-valued stochastic process $Z = \{Z(t), -\infty < t < \infty\}$ is said to be *self-similar* with index $H > 0$ if for any $a > 0$,

$$\mathcal{L}(\{Z(at), -\infty < t < \infty\}) = \mathcal{L}(\{a^H Z(t), -\infty < t < \infty\}) \tag{2. 2}$$

where $\mathcal{L}$ denotes the finite dimensional distributions and the equality indicates the equality of the finite dimensional distributions of both the processes. The index H is called the *scaling exponent* or the *fractal index* or the *Hurst parameter* of the process. If H is the scaling exponent of a self-similar process Z, then the process Z is called H- self similar process or H-ss process for short. It can be checked that a nondegenerate H-ss process cannot be a stationary process. In fact if $\{Z(t), t > 0\}$ is a H-ss process, then the process

$$Y(t) = e^{-tH} Z(e^t), -\infty < t < \infty \tag{2. 3}$$

is a stationary process. Conversely if $Y = \{Y(t), -\infty < t < \infty\}$ is a stationary process, then $Z = \{t^H Y(\log t), t > 0\}$ is a H-ss process.

Suppose $Z = \{Z(t), -\infty < t < \infty\}$ is a H-ss process with finite variance and stationary increments, that is

$$\mathcal{L}(Z(t+h) - Z(t)) = \mathcal{L}(Z(t) - Z(0)), -\infty < t, h < \infty. \tag{2. 4}$$

Then the following properties hold:

(i) $Z(0) = 0$ a.s;

(ii) If $H \neq 1$, then $E(Z(t)) = 0, -\infty < t < \infty$;

(iii) $\mathcal{L}(Z(-t)) = \mathcal{L}(-Z(t))$;

(iv)$E(Z^2(t)) = |t|^{2H} E(Z^2(1))$;

(v)The covariance function $\Gamma_H(t,s)$ of the process Z is given by

$$\Gamma_H(t,s) = \frac{1}{2}\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\}. \tag{2. 5}$$

(vi)The self-similarity parameter, also called the scaling exponent or fractal index H, is less than or equal to one.

(vii) If $H = 1$, then $Z(t) = tZ(1)$ a.s. for $-\infty < t < \infty$.

(viii) Let $0 < H \leq 1$. Then the function

$$R_H(s,t) = \{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\} \tag{2. 6}$$

is nonnegative definite. For proofs of the above properties, see Taqqu (2003).

A gaussian process H-ss process $W^H = \{W_H(t), -\infty < t < \infty\}$ with stationary increments and with fractal index $0 < H < 1$ is called a *fractional Brownian motion* (fBm). It is said to be standard if $Var(W^H(1)) = 1$. For any $0 < H < 1$, there exists a version of the fBm for which the sample paths are continuous with probability one but are not differentiable even in the L^2-sense. The continuity of the sample paths follows from the Kolmogorov's continuity condition and the fact that

$$E|W^H(t_2) - W^H(t_1)|^\alpha = E|W^H(1)|^\alpha |t_2 - t_1|^{\alpha H} \tag{2. 7}$$

from the property that the fBm is a H-ss process with stationary increments. We can choose α such that $\alpha H > 1$ to satisfy the Kolmogorov's continuity condition. Further more

$$E|\frac{W^H(t_2) - W^H(T_1)}{t_2 - t_1}|^2 = E[W^H(1)^2]|t_2 - t_1|^{2H-2} \tag{2. 8}$$

and the last term tends to infinity as $t_2 \to t_1$ since $H < 1$. Hence the paths of the fBm are not L^2-differentiable. It is interesting to note that the fractional Brownian motion reduces to the Brownian motion or the Wiener process for the case when $H = \frac{1}{2}$. As was mentioned above, self-similar processes have been used for stochastic modeling in such diverse areas as hydrology, geophysics, medicine, genetics and financial economics and more recently in modeling internet traffic patterns. Recent additional applications are given in Buldyrev et al. (1993), Ossandik et al. (1994), Percival and Guttorp (1994) and Peng et al.(1992, 1995a,b). It is important to estimate the constant H for modeling purposes. This problem has been considered by Azais (1990), Geweke and Porter-Hudak (1983), Taylor and Taylor (1991), Beran and Terrin (1994), Constantine and Hall (1994), Feuverger et al. (1994), Chen et al. (1995), Robinson (1995), Abry and Sellan (1996), Comte (1996), McCoy and Walden (1996), Hall et al. (1997), Kent and Wood (1997), and more recently in Jensen (1998), Poggi and Viano (1998) and Coeurjolly (2001). It was observed that there are some phenomena which exhibit self-similar behaviour locally but the nature of self-similarity changes as the phenomenon evolves. It was suggested that the parameter H must be allowed to vary as function of time for modeling such data. Goncalves and Flandrin (1993) and Flandrin and Goncalves (1994) propose a class of processes which are called *locally self-similar* with dependent scaling exponents and discuss their applications. Wang et al. (2001) develop procedures using wavelets to construct local estimates for time varying scaling exponent $H(t)$ of a locally self-similar process.

3 Stochastic differential equations driven by fBm

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a stochastic basis satisfying the usual conditions. The natural fitration of a process is understood as the P-completion of the filtration generated by this process. Let $W^H = \{W_t^H, t \geq 0\}$ be a normalized fractional Brownian motion (fBm) with Hurst parameter $H \in (0,1)$, that is, a gaussian process with continuous sample paths such that $W_0^H = 0, E(W_t^H) = 0$ and

(3. 1) $$E(W_s^H W_t^H) = \frac{1}{2}[s^{2H} + t^{2H} - |s-t|^{2H}], t \geq 0, s \geq 0.$$

Let us consider a stochastic process $Y = \{Y_t, t \geq 0\}$ defined by the stochastic integral equation

(3. 2) $$Y_t = \int_0^t C(s)ds + \int_0^t B(s)dW_s^H, t \geq 0$$

where $C = \{C(t), t \geq 0\}$ is an $(\mathcal{F}_t)$-adapted process and $B(t)$ is a nonvanishing nonrandom function. For convenience we write the above integral equation in the form of a stochastic differential equation

(3. 3) $$dY_t = C(t)dt + B(t)dW_t^H, t \geq 0$$

driven by the fractional Brownian motion W^H. The integral

(3. 4) $$\int_0^t B(s)dW_s^H$$

is not a stochastic integral in the Ito sense but one can define the integral of a deterministic function with respect to the fBm in a natural sense (cf. Gripenberg and Norros (1996); Norros et al. (1999)). Even though the process Y is not a semimartingale, one can associate a semimartingale $Z = \{Z_t, t \geq 0\}$ which is called a *fundamental semimartingale* such that the natural filtration $(\mathcal{Z}_t)$ of the process Z coincides with the natural filtration $(\mathcal{Y}_t)$ of the process Y (Kleptsyna et al. (2000)). Define, for $0 < s < t$,

$$k_H = 2H\ \Gamma(\frac{3}{2} - H)\ \Gamma(H + \frac{1}{2}), \tag{3. 5}$$

$$k_H(t,s) = k_H^{-1} s^{\frac{1}{2}-H}(t-s)^{\frac{1}{2}-H}, \tag{3. 6}$$

$$\lambda_H = \frac{2H\ \Gamma(3-2H)\ \Gamma(H+\frac{1}{2})}{\Gamma(\frac{3}{2}-H)}, \tag{3. 7}$$

$$w_t^H = \lambda_H^{-1} t^{2-2H}, \tag{3. 8}$$

and

$$M_t^H = \int_0^t k_H(t,s)dW_s^H, t \geq 0. \tag{3. 9}$$

The process M^H is a Gaussian martingale, called the *fundamental martingale* (cf. Norros et al. (1999)) and its quadratic variation $< M_t^H >= w_t^H$. Further more the natural filtration of the martingale M^H coincides with the natural fitration of the fBm W^H. In fact the stochastic integral

$$\int_0^t B(s)dW_s^H \tag{3. 10}$$

can be represented in terms of the stochastic integral with respect to the martingale M^H. For a measurable function f on $[0,T]$, let

$$K_H^f(t,s) = -2H\frac{d}{ds}\int_s^t f(r)r^{H-\frac{1}{2}}(r-s)^{H-\frac{1}{2}}dr, 0 \leq s \leq t \tag{3. 11}$$

when the derivative exists in the sense of absolute continuity with respect to the Lebesgue measure(see Samko et al. (1993) for sufficient conditions). The following result is due to Kleptsyna et al. (2000).

Therorem 3.1: Let M^H be the fundamental martingale associated with the fBm W^H defined by (3.9). Then

$$\int_0^t f(s)dW_s^H = \int_0^t K_H^f(t,s)dM_s^H, t \in [0,T] \tag{3. 12}$$

a.s $[P]$ whenever both sides are well defined.

Suppose the sample paths of the process $\{\frac{C(t)}{B(t)}, t \geq 0\}$ are smooth enough (see Samko et al. (1993)) so that

$$Q_H(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{C(s)}{B(s)}ds, t \in [0,T] \tag{3. 13}$$

is well-defined where w^H and k_H are as defined in (3.8) and (3.6) respectively and the derivative is understood in the sense of absoulute continuity. The following theorem due to Kleptsyna et al. (2000) associates a *fundamental semimartingale* Z associated with the process Y such that the natural filtration $(\mathcal{Z}_t)$ coincides with the natural filtration $(\mathcal{Y}_t)$ of Y.

Theorem 3.2: Suppose the sample paths of the process Q_H defined by (3.13) belong P-a.s to $L^2([0,T], dw^H)$ where w^H is as defined by (3.8). Let the process $Z = (Z_t, t \in [0,T])$ be defined by

$$Z_t = \int_0^t k_H(t,s) B^{-1}(s) dY_s \tag{3. 14}$$

where the function $k_H(t,s)$ is as defined in (3.6). Then the following results hold:
(i) The process Z is an $(\mathcal{F}_t)$ -semimartingale with the decomposition

$$Z_t = \int_0^t Q_H(s) dw_s^H + M_t^H \tag{3. 15}$$

where M^H is the fundamental martingale defined by (3.9), (ii) the process Y admits the representation

$$Y_t = \int_0^t K_H^B(t,s) dZ_s \tag{3. 16}$$

where the function K_H^B is as defined in (3.11), and (iii) the natural fitrations of $(\mathcal{Z}_t)$ and $(\mathcal{Y}_t)$ coincide.

Kleptsyna et al. (2000) derived the following Girsanov type formula as a consequence of the Theorem 3.2.

Theorem 3.3: Suppose the assumptions of Theorem 3.2 hold. Define

$$\Lambda_H(T) = \exp\{-\int_0^T Q_H(t) dM_t^H - \frac{1}{2}\int_0^t Q_H^2(t) dw_t^H\}. \tag{3. 17}$$

Suppose that $E(\Lambda_H(T)) = 1$. Then the measure $P^* = \Lambda_H(T)P$ is a probability measure and the probability measure of the process Y under P^* is the same as that of the process V defined by

$$V_t = \int_0^t B(s) dW_s^H, 0 \le t \le T. \tag{3. 18}$$

.

4 Statistical inference for linear SDE driven by fBm

Statistical inference for diffusion type processes satisfying stochastic differential equations driven by Wiener processes have been studied earlier and a comprehensive survey of various methods is given in Prakasa Rao (1999a, b). There has been a recent interest to study similar problems for stochastic processes driven by a fractional Brownian motion for modeling stochastic phemonena with possible long range dependence. Le Breton (1998) studied parameter estimation and filtering in a simple linear model driven by a fractional Brownian motion. In a recent paper, Kleptsyna and Le Breton (2002) studied parameter estimation problems for fractional Ornstein-Uhlenbeck type process. This is a fractional analogue of the Ornstein-Uhlenbeck process, that is, a continuous time first order autoregressive process $X = \{X_t, t \geq 0\}$ which is the solution of a one-dimensional homogeneous linear stochastic differential equation driven by a fractional Brownian motion (fBm) $W^H = \{W_t^H, t \geq 0\}$ with Hurst parameter $H \in [1/2, 1)$. Such a process is the unique Gaussian process satisfying the linear integral equation

$$X_t = \theta \int_0^t X_s ds + \sigma W_t^H, t \geq 0. \tag{4. 1}$$

They investigate the problem of estimation of the parameters θ and σ^2 based on the observation $\{X_s, 0 \leq s \leq T\}$ and prove that the maximum likelihood estimator $\hat{\theta}_T$ is strongly consistent as $T \to \infty$. We now discuss more general classes of stochastic processes satisfying linear stochastic differential equations driven a fractional Brownian motion and review some recent work connected with the asymptotic properties of the maximum likelihood and the Bayes estimators for parameters involved in such processes. We will also discuss some aspects of sequential estimation and minimum distance estimation problems for fractional Ornstein-Uhlenbeck type processes in the next section. Let us consider the stochastic differential equation

$$dX(t) = [a(t, X(t)) + \theta\, b(t, X(t))]dt + \sigma(t) dW_t^H, X(0) = 0, t \geq 0 \tag{4. 2}$$

where $\theta \in \Theta \subset R, W = \{W_t^H, t \geq 0\}$ is a fractional Brownian motion with known Hurst parameter H and $\sigma(t)$ is a positive nonvanishing function on $[0, \infty)$. In other words $X = \{X_t, t \geq 0\}$ is a stochastic process satisfying the stochastic integral equation

$$X(t) = \int_0^t [a(s, X(s)) + \theta\, b(s, X(s))]ds + \int_0^t \sigma(s) dW_s^H, t \geq 0. \tag{4. 3}$$

Let

$$C(\theta, t) = a(t, X(t)) + \theta\, b(t, X(t)), t \geq 0 \tag{4. 4}$$

and assume that the sample paths of the process $\{\frac{C(\theta,t)}{\sigma(t)}, t \geq 0\}$ are smooth enough so that the the process

$$Q_{H,\theta}(t) = \frac{d}{dw_t^H} \int_0^t k_H(t, s) \frac{C(\theta, s)}{\sigma(s)} ds, t \geq 0 \tag{4. 5}$$

is well-defined where w_t^H and $k_H(t, s)$ are as defined in (3.8) and (3.6) respectively. Suppose the sample paths of the process $\{Q_{H,\theta}, 0 \leq t \leq T\}$ belong almost surely to $L^2([0, T], dw_t^H)$. Define

$$Z_t = \int_0^t \frac{k_H(t, s)}{\sigma(s)} dX_s, t \geq 0. \tag{4. 6}$$

Then the process $Z = \{Z_t, t \geq 0\}$ is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_{H,\theta}(s) dw_s^H + M_t^H \tag{4. 7}$$

where M^H is the fundamental martingale defined by (3.9) and the process X admits the representation

(4. 8) $$X_t = \int_0^t K_H^\sigma(t,s)dZ_s$$

where the function K_H^σ is as defined by (3.11). Let P_θ^T be the measure induced by the process $\{X_t, 0 \leq t \leq T\}$ when θ is the true parameter. Following Theorem 3.3, we get that the Radon-Nikodym derivative of P_θ^T with respect to P_0^T is given by

(4. 9) $$\frac{dP_\theta^T}{dP_0^T} = \exp[\int_0^T Q_{H,\theta}(s)dZ_s - \frac{1}{2}\int_0^T Q_{H,\theta}^2(s)dw_s^H].$$

Maximum likelihood estimation We now consider the problem of estimation of the parameter θ based on the observation of the process $X = \{X_t, 0 \leq t \leq T\}$ and study its asymptotic properties as $T \to \infty$.

Strong consistency: Let $L_T(\theta)$ denote the Radon-Nikodym derivative $\frac{dP_\theta^T}{dP_0^T}$. The maximum likelihood estimator (MLE) is defined by the relation

(4. 10) $$L_T(\hat{\theta}_T) = \sup_{\theta \in \Theta} L_T(\theta).$$

We assume that there exists a measurable maximum likelihood estimator. Sufficient conditions can be given for the existence of such an estimator (cf. Lemma 3.1.2, Prakasa Rao (1987)). Note that

(4. 11) $$\begin{aligned} Q_{H,\theta}(t) &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{C(\theta,s)}{\sigma(s)}ds \\ &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{a(s,X(s))}{\sigma(s)}ds + \theta\frac{d}{dw_t^H}\int_0^t k_H(t,s)\frac{b(s,X(s))}{\sigma(s)}ds \\ &= J_1(t) + \theta J_2(t).(say) \end{aligned}$$

Then

(4. 12) $$\log L_T(\theta) = \int_0^T (J_1(t) + \theta J_2(t))dZ_t - \frac{1}{2}\int_0^T (J_1(t) + \theta J_2(t))^2 dw_t^H$$

and the likelihood equation is given by

(4. 13) $$\int_0^T J_2(t)dZ_t - \int_0^T (J_1(t) + \theta J_2(t))J_2(t)dw_t^H = 0.$$

Hence the MLE $\hat{\theta}_T$ of θ is given by

(4. 14) $$\hat{\theta}_T = \frac{\int_0^T J_2(t)dZ_t + \int_0^T J_1(t)J_2(t)dw_t^H}{\int_0^T J_2^2(t)dw_t^H}.$$

Let θ_0 be the true parameter. Using the fact that

(4. 15) $$dZ_t = (J_1(t) + \theta_0 J_2(t))dw_t^H + dM_t^H,$$

it can be shown that

$$\frac{dP_\theta^T}{dP_{\theta_0}^T} = \exp[(\theta-\theta_0)\int_0^T J_2(t)dM_t^H - \frac{1}{2}(\theta-\theta_0)^2\int_0^T J_2^2(t)dw_t^H. \tag{4. 16}$$

Following this representation of the Radon-Nikodym derivative, we obtain that

$$\hat{\theta}_T - \theta_0 = \frac{\int_0^T J_2(t)dM_t^H}{\int_0^T J_2^2(t)dw_t^H}. \tag{4. 17}$$

Note that the quadratic variation $< Z >$ of the process Z is the same as the quadratic variation $< M^H >$ of the martingale M^H which in turn is equal to w^H. This follows from the relations (3.15) and (3.9). Hence we obtain that

$$[w_T^H]^{-1}\lim_n \Sigma[Z_{t_{i+1}^{(n)}} - Z_{t_i^{(n)}}]^2 = 1 \text{ a.s } [P_{\theta_0}]$$

where $(t_i^{(n)}$ is a partition of the interval $[0,T]$ such that $\sup|t_{i+1}^{(n)} - t_i^{(n)}|$ tends to zero as $n\to\infty$. If the function $\sigma(t)$ is an unknown constant σ, the above property can be used to obtain a strongly consistent estimator of σ^2 based on the continuous observation of the process X over the interval $[0,T]$. Here after we assume that the nonrandom function $\sigma(t)$ is known.

We now discuss the problem of maximum likelihood estimation of the parameter θ on the basis of the observation of the process X or equivalently the process Z on the interval $[0,T]$. The following result holds.

Theorem 4.1: The maximum likelihood estimator $\hat{\theta}_T$ is strongly consistent, that is,

$$\hat{\theta}_T \to \theta_0 \text{ a.s } [P_{\theta_0}] \text{ as } T\to\infty \tag{4. 18}$$

provided

$$\int_0^T J_2^2(t)dw_t^H \to \infty \text{ a.s } [P_{\theta_0}] \text{ as } T\to\infty. \tag{4. 19}$$

Remark: For the case fractional Ornstein-Uhlenbeck type process investigated in Kleptsyna and Le Breton (2002), it can be checked that the condition stated in equation (4.18) holds and hence the maximum likelihood estimator $\hat{\theta}_T$ is strongly consistent as $T\to\infty$.

Limiting distribution: We now discuss the limiting distribution of the MLE $\hat{\theta}_T$ as $T\to\infty$.

Theorem 4.2: Assume that the functions $b(t,s)$ and $\sigma(t)$ are such that the process $\{R_t, t\geq 0\}$ is a local continuous martingale and that there exists a norming function $I_t, t\geq 0$ such that

$$I_T^2 < R_T >= I_T^2\int_0^T J_2^2(t)dw_t^H \xrightarrow{p} \eta^2 \text{ as } T\to\infty \tag{4. 20}$$

where $I_T\to 0$ as $T\to\infty$ and η is a random variable such that $P(\eta>0)=1$. Then

$$(I_T R_T, I_T^2 < R_T >) \xrightarrow{\mathcal{L}} (\eta Z, \eta^2) \text{ as } T\to\infty \tag{4. 21}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

For the proofs of Theorems 4.1 and 4.2, see Prakasa Rao (2003a).

Theorem 4.3: Suppose the conditions stated in the Theorem 4.2 hold. Then

$$I_T^{-1}(\hat{\theta}_T - \theta_0) \xrightarrow{\mathcal{L}} \frac{Z}{\eta} \quad \text{as } t \to \infty \tag{4. 22}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

Remarks: If the random variable η is a constant with probability one, then the limiting distribution of the maximum likelihood estimator is normal with mean 0 and variance η^{-2}. Otherwise it is a mixture of the normal distributions with mean zero and variance η^{-2} with the mixing distribution as that of η. The rate of convergence of the distribution of the maximum likelihood estimator is discussed in Prakasa Rao (2003b).

Bayes estimation Suppose that the parameter space Θ is open and Λ is a prior probability measure on the parameter space Θ. Further suppose that Λ has the density $\lambda(.)$ with respect to the Lebesgue measure and the density function is continuous and positive in an open neighbourhood of θ_0, the true parameter. Let

$$\alpha_T \equiv I_T R_T = I_T \int_0^T J_2(t) dM_t^H \tag{4. 23}$$

and

$$\beta_T \equiv I_T^2 < R_T >= I_T^2 \int_0^T J_2^2(t) dw_t^H. \tag{4. 24}$$

We have seen earlier in (4.17) that the maximum likelihood estimator satisfies the relation

$$\alpha_T = (\hat{\theta}_T - \theta_0) I_T^{-1} \beta_T. \tag{4. 25}$$

The posterior density of θ given the observation $X^T \equiv \{X_s, 0 \leq s \leq T\}$ is given by

$$p(\theta|X^T) = \frac{\frac{dP_\theta^T}{dP_{\theta_0}^T} \lambda(\theta)}{\int_\Theta \frac{dP_\theta^T}{dP_{\theta_0}^T} \lambda(\theta) d\theta}. \tag{4. 26}$$

Let us write $t = I_T^{-1}(\theta - \hat{\theta}_T)$ and define

$$p^*(t|X^T) = I_T \;\; p(\hat{\theta}_T + tI_T|X^T). \tag{4. 27}$$

Then the function $p^*(t|X^T)$ is the posterior density of the transformed variable $t = I_T^{-1}(\theta - \hat{\theta}_T)$. Let

$$\begin{aligned} \nu_T(t) &\equiv \frac{dP_{\hat{\theta}_T + tI_T}/dP_{\theta_0}}{dP_{\hat{\theta}_T}/dP_{\theta_0}} \\ &= \frac{dP_{\hat{\theta}_T + tI_T}}{dP_{\hat{\theta}_T}} a.s. \end{aligned} \tag{4. 28}$$

and

(4. 29) $$C_T = \int_{-\infty}^{\infty} \nu_T(t)\lambda(\hat{\theta}_T + tI_T)dt.$$

It can be checked that

(4. 30) $$p^*(t|X^T) = C_T^{-1}\nu_T(t)\lambda(\hat{\theta}_T + tI_T)$$

and

(4. 31)
$$\begin{aligned}\log \nu_T(t) &= I_T^{-1}\alpha_T[(\hat{\theta}_T + tI_T - \theta_0) - (\hat{\theta}_T - \theta_0)] \\ &\quad -\frac{1}{2}I_T^{-2}\beta_T[(\hat{\theta}_T + tI_T - \theta_0)^2 - (\hat{\theta}_T - \theta_0)^2] \\ &= t\alpha_T - \frac{1}{2}t^2\beta_T - t\beta_T I_T^{-1}(\hat{\theta}_T - \theta_0) \\ &= -\frac{1}{2}\beta_T t^2\end{aligned}$$

in view of the equation (4.25).

Suppose that the convergence in the condition in the equation (4.20) holds almost surely under the measure P_{θ_0} and the limit is a constant $\eta^2 > 0$ with probability one. For convenience, we write $\beta = \eta^2$. Then

(4. 32) $$\beta_T \to \beta \text{ a.s } [P_{\theta_0}] \text{ as } T \to \infty.$$

Further suppose that $K(t)$ is a nonnegative measurable function such that, for some $0 < \varepsilon < \beta$,

(4. 33) $$\int_{-\infty}^{\infty} K(t)\exp[-\frac{1}{2}t^2(\beta - \varepsilon)]dt < \infty$$

and the maximum likelihood estimator $\hat{\theta}_T$ is strongly consistent, that is,

(4. 34) $$\hat{\theta}_T \to \theta_0 \text{ a.s } [P_{\theta_0}] \text{ as } T \to \infty.$$

In addition, suppose that the following condition holds for every $\varepsilon > 0$ and $\delta > 0$:

(4. 35) $$\exp[-\varepsilon I_T^{-2}]\int_{|u|>\delta} K(uI_T^{-1})\lambda(\hat{\theta}_T + u)du \to 0 \text{ a.s.}[P_{\theta_0}] \text{ as } T \to \infty.$$

Then we have the follwing theorem which is an analogue of the Bernstein - von Mises theorem proved in Prakasa Rao (1981) for a class of processes satisfying a linear stochastic differential equation driven by the standard Wiener process.

Theorem 4.4: Let the assumptions (4.32) to (4.35) hold where $\lambda(.)$ is a prior density which is continuous and positive in an open neighbourhood of θ_0, the true parameter. Then

(4. 36) $$\lim_{T\to\infty}\int_{-\infty}^{\infty} K(t)|p^*(t|X^T) - (\frac{\beta}{2\pi})^{1/2}\exp(-\frac{1}{2}\beta t^2)|dt = 0 \text{ a.s } [P_{\theta_0}].$$

As a consequence of the above theorem, we obtain the following result by choosing $K(t) = |t|^m$, for some integer $m \geq 0$.

Theorem 4.5: Assume that the following conditions hold:

$$(4.\ 37) \qquad (C1)\ \hat{\theta}_T \to \theta_0 \ \text{ a.s } [P_{\theta_0}] \ \text{ as } T \to \infty,$$

$$(4.\ 38) \qquad (C2)\ \beta_T \to \beta > 0 \ \text{ a.s } [P_{\theta_0}] \ \text{ as } T \to \infty.$$

Further suppose that (C3)$\lambda(.)$ is a prior probability density on Θ which is continuous and positive in an open neighbourhood of θ_0, the true parameter and

$$(4.\ 39) \qquad (C4)\ \int_{-\infty}^{\infty} |\theta|^m \lambda(\theta) d\theta < \infty$$

for some integer $m \geq 0$. Then

$$(4.\ 40) \qquad \lim_{T\to\infty} \int_{-\infty}^{\infty} |t|^m |p^*(t|X^T) - (\frac{\beta}{2\pi})^{1/2} \exp(-\frac{1}{2}\beta t^2)| dt = 0 \ \text{ a.s } [P_{\theta_0}].$$

In particular, choosing $m = 0$, we obtain that

$$(4.\ 41) \qquad \lim_{T\to\infty} \int_{-\infty}^{\infty} |p^*(t|X^T) - (\frac{\beta}{2\pi})^{1/2} \exp(-\frac{1}{2}\beta t^2)| dt = 0 \ \text{ a.s } [P_{\theta_0}]$$

whenver the conditions (C1), (C2) and (C3) hold. This is the analogue of the Bernstein-von Mises theorem for a class of diffusion processes proved in Prakasa Rao (1981) and it shows the asymptotic convergence in the L_1-mean of the posterior density to the normal distribution.

For proofs of above results, see Prakasa Rao (2003a).

As a Corollory to the Theorem 4.5, we also obtain that the conditional expectation, under P_{θ_0}, of $[I_T^{-1}(\hat{\theta}_T - \theta)]^m$ converges to the corresponding m-th abosolute moment of the normal distribution with mean zero and variance β^{-1}.

We define a *regular Bayes estimator* of θ, corresponding to a prior probability density $\lambda(\theta)$ and the loss function $L(\theta, \phi)$, based on the observation X^T, as an estimator which minimizes the posterior risk

$$(4.\ 42) \qquad B_T(\phi) \equiv \int_{-\infty}^{\infty} L(\theta, \phi) p(\theta|X^T) d\theta.$$

over all the estimators ϕ of θ. Here $L(\theta, \phi)$ is a loss function defined on $\Theta \times \Theta$.

Suppose there exists a measurable regular Bayes estimator $\tilde{\theta}_T$ for the parameter θ (cf. Theorem 3.1.3, Prakasa Rao (1987).) Suppose that the loss function $L(\theta, \phi)$ satisfies the following conditions:

$$(4.\ 43) \qquad L(\theta, \phi) = \ell(|\theta - \phi|) \geq 0$$

and the function $\ell(t)$ is nondecreasing for $t \geq 0$. An example of such a loss function is $L(\theta, \phi) = |\theta - \phi|$. Suppose there exist nonnegative functions $R(t)$, $J(t)$ and $G(t)$ such that

$$(4.\ 44) \qquad (D1)\ R(t)\ell(tI_T) \leq G(t) \ \text{ for all } T \geq 0,$$

$$(4.45) \qquad (D2) \;\; R(t)\ell(tI_T) \to J(t) \;\text{ as }\; T \to \infty$$

uniformly on bounded intervals of t. Further suppose that the function

$$(4.46) \qquad (D3) \;\; \int_{-\infty}^{\infty} J(t+h)\exp[-\frac{1}{2}\beta t^2]dt$$

has a strict minimum at $h=0$, and
(D4)the function $G(t)$ satisfies the conditions similar to (4.33) and (4.35).

We have the following result giving the asymptotic properties of the Bayes risk of the estimator $\tilde{\theta}_T$.

Thoerem 4.6: Suppose the conditions (C1) to (C3) in the Theorem 4.5 and the conditions (D1) to (D4) stated above hold. Then

$$(4.47) \qquad I_T^{-1}(\tilde{\theta}_T - \hat{\theta}_T) \to 0 \;\text{ a.s }\; [P_{\theta_0}] \;\text{ as }\; T \to \infty$$

and

$$(4.48) \qquad \begin{aligned} \lim_{T\to\infty} R(T)B_T(\tilde{\theta}_T) &= \lim_{T\to\infty} R(T)B_T(\hat{\theta}_T) \\ &= (\frac{\beta}{2\pi})^{1/2}\int_{-\infty}^{\infty} K(t)\exp[-\frac{1}{2}\beta t^2]dt \;\text{ a.s }\; [P_{\theta_0}] \end{aligned}$$

This theorem can be proved by arguments similar to those given in the proof of Theorem 4.1 in Borwanker et al. (1971).

We have observed earlier that

$$(4.49) \qquad I_T^{-1}(\hat{\theta}_T - \theta_0) \xrightarrow{\mathcal{L}} N(0,\beta^{-1}) \;\text{ as }\; T \to \infty.$$

As a consequence of the Theorem 4.6, we obtain that

$$(4.50) \qquad \tilde{\theta}_T \to \theta_0 \;\text{ a.s }\; [P_{\theta_0}] \;\text{ as }\; T \to \infty$$

and

$$(4.51) \qquad I_T^{-1}(\tilde{\theta}_T - \theta_0) \xrightarrow{\mathcal{L}} N(0,\beta^{-1}) \;\text{ as }\; T \to \infty.$$

In other words, the Bayes estimator is asymptotically normal and has asymptotically the same distribution as the maxiumum likelihood estimator. The asymptotic Bayes risk of the estimator is given by the Theorem 4.6.

5 Statistical inference for fractional Ornstein-Uhlenbeck type process

In a recent paper, Kleptsyna and Le Breton (2002) studied parameter estimation problems for fractional Ornstein-Uhlenbeck type process. This is a fractional analogue of the Ornstein-Uhlenbeck process, that is, a continuous time first order autoregressive process $X = \{X_t, t \geq 0\}$

which is the solution of a one-dimensional homogeneous linear stochastic differential equation driven by a fractional Brownian motion (fBm) $W^H = \{W_t^H, t \geq 0\}$ with Hurst parameter $H \in (1/2, 1)$. Such a process is the unique Gaussian process satisfying the linear integral equation

$$X_t = \theta \int_0^t X_s ds + \sigma W_t^H, t \geq 0. \tag{5. 1}$$

They investigate the problem of estimation of the parameters θ and σ^2 based on the observation $\{X_s, 0 \leq s \leq T\}$ and prove that the maximum likelihood estimator $\hat{\theta}_T$ is strongly consistent as $T \to \infty$. It is well known that the sequential estimation methods might lead to equally efficient estimators, as compared to the maximum likelihood estimators, from the process observed possibly over a shorter expected period of observation time. Novikov (1972) investigated the asymptotic properties of a sequential maximum likelihood estimator for the drift parameter in the Ornstein-Uhlenbeck process. Maximum likelihood estimators are not robust. Kutoyants and Pilibossian (1994) developed a minimum L_1-norm estimator for the drift parameter. We now discuss the asymptotic properties of a sequential maximum likelihood estimators and minimum L_1-norm estimators for the drift parameter for a fractional Ornstein-Uhlenbeck type process.

Maximum likelihood estimation Let

$$K_H(t,s) = H(2H-1)\frac{d}{ds}\int_s^t r^{H-\frac{1}{2}}(r-s)^{H-\frac{3}{2}}dr, 0 \leq s \leq t. \tag{5. 2}$$

The sample paths of the process $\{X_t, t \geq 0\}$ are smooth enough so that the process Q defined by

$$Q(t) = \frac{d}{dw_t^H}\int_0^t k_H(t,s)X_s ds, t \in [0,T] \tag{5. 3}$$

is well-defined where w^H and k_H are as defined in (3.7) and (3.5) respectively and the derivative is understood in the sense of absolute continuity with respect to the measure generated by w^H. More over the sample paths of the process Q belong to $L^2([0,T], dw^H)$ a.s. [P]. Define the process Z as in (4.6).

As an application of the Girsanov type formula given in Theorem 3.3 for the fractional Brownian motions derived by Kleptsyna et al. (2000) , it follows that the Radon-Nikodym derivative of the measure P_θ^T, generated by the stochastic process X when θ is the true parameter, with respect to the measure generated by the process X when $\theta = 0$, is given by

$$\frac{dP_\theta^T}{dP_0^T} = \exp[\theta\int_0^T Q(s)dZ_s - \frac{1}{2}\theta^2\int_0^T Q^2(s)dw_s^H]. \tag{5. 4}$$

Further more the quadratic variation $< Z >_T$ of the process Z on $[0,T]$ is equal to $\sigma^2 w_T^H$ a.s. and hence the parameter σ^2 can be estimated by the relation

$$\lim_n \Sigma[Z_{t_{i+1}^{(n)}} - Z_{t_i^{(n)}}]^2 = \sigma^2 w_T^H \text{ a.s.} \tag{5. 5}$$

where $(t_i^{(n)})$ is an appropriate partition of $[0,T]$ such that

$$\sup_i |t_{i+1}^{(n)} - t_i^{(n)}| \to 0$$

as $n \to \infty$. Hence we can estimate σ^2 almost surely from any small interval as long as we have a continuous observation of the process. For further discussion, we assume that $\sigma^2 = 1$.

We consider the problem of estimation of the parameter θ based on the observation of the process $X = \{X_t, 0 \leq t \leq T\}$ for a fixed time T and study its asymptotic properties as $T \to \infty$. The following results are due to Kleptsyna and Le Breton (2002) and Prakasa Rao (2003a).

Theorem 5.1: The maximum likelihood estimator θ from the observation $X = \{X_t, 0 \leq t \leq T\}$ is given by

$$\hat{\theta}_T = \{\int_0^T Q^2(s) dw_s^H\}^{-1} \int_0^T Q(s) dZ_s. \tag{5. 6}$$

Then the estimator $\hat{\theta}_T$ is strongly consistent as $T \to \infty$, that is,

$$\lim_{T\to\infty} \hat{\theta}_T = \theta \text{ a.s. } [P_\theta] \tag{5. 7}$$

for every $\theta \in R$.

We now discuss the limiting distribution of the MLE $\hat{\theta}_T$ as $T \to \infty$.

Theorem 5.2: Let

$$R_T = \int_0^T Q(s) dZ_s. \tag{5. 8}$$

Assume that there exists a norming function $I_t, t \geq 0$ such that

$$I_T^2 \int_0^T Q^2(t) dw_t^H \xrightarrow{p} \eta^2 \text{ as } T \to \infty \tag{5. 9}$$

where $I_T \to 0$ as $T \to \infty$ and η is a random variable such that $P(\eta > 0) = 1$. Then

$$(I_T R_T, I_T^2 < R_T >) \xrightarrow{\mathcal{L}} (\eta Z, \eta^2) \text{as } T \to \infty \tag{5. 10}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

Observe that

$$I_T^{-1}(\hat{\theta}_T - \theta_0) = \frac{I_T R_T}{I_T^2 < R_T >} \tag{5. 11}$$

Applying the Theorem 5.2, we obtain the following result.

Theorem 5.3: Suppose the conditions stated in the Theorem 5.2 hold. Then

$$I_T^{-1}(\hat{\theta}_T - \theta_0) \xrightarrow{\mathcal{L}} \frac{Z}{\eta} \text{ as } T \to \infty \tag{5. 12}$$

where the random variable Z has the standard normal distribution and the random variables Z and η are independent.

Remarks: If the random variable η is a constant with probability one, then the limiting distribution of the maximum likelihood estimator is normal with mean 0 and variance η^{-2} Otherwise it is a mixture of the normal distributions with mean zero and variance η^{-2} with the mixing distribution as that of η. Berry-Esseen type bound for the MLE is discussed in Prakasa Rao (2003b) when the limiting distribution of the MLE is normal.

Sequential maximum likelihood estimation We now consider the problem of sequential maximum likelihood estimation of the parameter θ. Let h be a nonnegative number. Define the stopping rule $\tau(h)$ by the rule

$$\tau(h) = \inf\{t : \int_0^t Q^2(s)dw_s^H \geq h\}. \tag{5. 13}$$

Kletptsyna and Le Breton (2002) have shown that

$$\lim_{t\to\infty} \int_0^t Q^2(s)dw_s^H = +\infty \text{ a.s. } [P_\theta] \tag{5. 14}$$

for every $\theta \in R$. Then it can be shown that $P_\theta(\tau(h) < \infty) = 1$. If the process is observed up to a previuosly determined time T , we have observed that the maximum likelihood estimator is given by

$$\hat{\theta}_T = \{\int_0^T Q^2(s)dw_s^H\}^{-1} \int_0^T Q(s)dZ_s. \tag{5. 15}$$

The estimator

$$\begin{aligned} \hat{\theta}(h) &\equiv \hat{\theta}_{\tau(h)} \\ &= \{\int_0^{\tau(h)} Q^2(s)dw_s^H\}^{-1} \int_0^{\tau(h)} Q(s)dZ_s \\ &= h^{-1} \int_0^{\tau(h)} Q(s)dZ_s \end{aligned} \tag{5. 16}$$

is called the *sequential maximum likelihood estimator* of θ. We now study the asymptotic properties of the estimator $\hat{\theta}(h)$.

The following lemma is an analogue of the Cramer-Rao inequality for sequential plans $(\tau(X), \hat{\theta}_\tau(X))$ for estimating the parameter θ satisfying the property

$$E_\theta\{\hat{\theta}_\tau(X)\} = \theta \tag{5. 17}$$

for all θ.

Lemma 5.4: Suppose that differentiation under the integral sign with respect to θ on the left side of the equation (5.17) is permissible. Further suppose that

$$E_\theta\{\int_0^{\tau(X)} Q^2(s)dw_s^H\} < \infty \tag{5. 18}$$

for all θ. Then

(5. 19) $$Var_\theta\{\hat{\theta}_\tau(X)\} \geq \{E_\theta\{\int_0^{\tau(X)} Q^2(s)dw_s^H\}^{-1}$$

for all θ.

A sequential plan $(\tau(X), \hat{\theta}_\tau(X))$ is said to be *efficient* if there is equality in (5.19) for all θ. We now have the following result.

Theorem 5.5: Consider the fractional Ornstein-Uhlenbeck type process governed by the stochastic differential equation (5.1) with $\sigma = 1$ driven by the fractional Brownian motion W^H with $H \in (\frac{1}{2}, 1)$. Then the sequential plan $(\tau(h), \hat{\theta}(h))$ defined by the equations (5.13) and (5.16) has the following properties for all θ.

(i) $\hat{\theta}(h) \equiv \hat{\theta}_{\tau(h)}$ is normally distributed with $E_\theta(\hat{\theta}(h)) = \theta$ and $Var_\theta(\hat{\theta}(h)) = h^{-1}$;

(ii) the plan is efficient; and

(iii) the plan is closed, that is, $P_\theta(\tau(h) < \infty) = 1$.

For proof, see Prakasa Rao (2003c).

Minimum L_1-norm estimation In spite of the fact that maximum likelihood estimators (MLE) are consistent and asymptotically normal and also asymptotically efficient in general, they have some short comings at the same time. Their calculation is often cumbersome as the expression for the MLE involve stochastic integrals which need good approximations for computational purposes. Further more the MLE are not robust in the sense that a slight perturbation in the noise component will change the properties of the MLE substantially. In order to circumvent such problems, the minimum distance approach is proposed. Properties of the minimum distance estimators (MDE) were discussed in Millar (1984) in a general frame work. We now obtain the minimum L_1-norm estimates of the drift parameter of a fractional Ornstein-Uhlenbeck type process and investigate the asymptotic properties of such estimators following the work of Kutoyants and Pilibossian (1994). We now consider the problem of estimation of the parameter θ based on the observation of fractional Ornstein-Uhlenbeck type process $X = \{X_t, 0 \leq t \leq T\}$ satisfying the stochastic differential equation

(5. 20) $$dX_t = \theta X(t)dt + \varepsilon dW_t^H, X_0 = x_0, 0 \leq t \leq T$$

for a fixed time T where $\theta \in \Theta \subset R$ and study its asymptotic properties as $\varepsilon \to 0$.

Let $x_t(\theta)$ be the solution of the above differential equation with $\varepsilon = 0$. It is obvious that

(5. 21) $$x_t(\theta) = x_0 e^{\theta t}, 0 \leq t \leq T.$$

Let

(5. 22) $$S_T(\theta) = \int_0^T |X_t - x_t(\theta)|dt.$$

We define θ^*_ε to be a *minimum L_1-norm estimator* if there exists a measurable selection θ^*_ε such that

(5. 23) $$S_T(\theta^*_\varepsilon) = \inf_{\theta\in\Theta} S_T(\theta).$$

Conditions for the existence of a measurable selection are given in Lemma 3.1.2 in Prakasa Rao (1987). We assume that there exists a measurable selection θ^*_ε satisfying the above equation. An alternate way of defining the estimator θ^*_ε is by the relation

(5. 24) $$\theta^*_\varepsilon = \arg\inf_{\theta\in\Theta}\int_0^T |X_t - x_t(\theta)|dt.$$

Consistency:

Let $W_T^{H*} = \sup_{0\leq t\leq T}|W_t^H|$. The self-similarity of the fractional Brownian motion W_t^H implies that the random variables W_{at}^H and $a^H W_t$ have the same probability distribution for any $a > 0$. Further more it follows from the self-similarity that the supremum process W^{H*} has the property that the random variables W_{at}^{H*} and $a^H W_t^{H*}$ have the same probability distribution for any $a > 0$. Hence we have the following observation due to Novikov and Valkeila (1999).

Lemma 5.6: Let $T > 0$ and $\{W_t^H, 0 \leq t \leq T\}$ be a fBm with Hurst index H. Let $W_T^{H*} = \sup_{0\leq t\leq T} W_t^H$. Then

(5. 25) $$E(W_T^{H*})^p = K(p,H)T^{pH}$$

for every $p > 0$, where $K(p,H) = E(W_1^{H*})^p$.

Let θ_0 denote the true parameter, For any $\delta > 0$, define

(5. 26) $$g(\delta) = \inf_{|\theta-\theta_0|>\delta}\int_0^T |x_t(\theta) - x_t(\theta_0)|dt.$$

Note that $g(\delta) > 0$ for any $\delta > 0$.

Theorem 5.7: For every $p > 0$, there exists a positive constant $K(p,H)$ such that, for every $\delta > 0$,

(5. 27) $$\begin{aligned} P_{\theta_0}^{(\varepsilon)}\{|\theta^*_\varepsilon - \theta_0| > \delta\} &\leq 2^p T^{pH+p} K(p,H) e^{|\theta_0|Tp}(g(\delta))^{-p}\varepsilon^p \\ &= O((g(\delta))^{-p}\varepsilon^p). \end{aligned}$$

Proof: Let $||.||$ denote the L_1-norm. Then

(5. 28) $$\begin{aligned} P_{\theta_0}^{(\varepsilon)}\{|\theta^*_\varepsilon - \theta_0| > \delta\} &= P_{\theta_0}^{(\varepsilon)}\{\inf_{|\theta-\theta_0|\leq\delta}||X - x(\theta)|| > \inf_{|\theta-\theta_0|>\delta}||X - x(\theta)||\} \\ &\leq P_{\theta_0}^{(\varepsilon)}\{\inf_{|\theta-\theta_0|\leq\delta}(||X - x(\theta_0)|| + ||x(\theta) - x(\theta_0)||) \\ &\qquad > \inf_{|\theta-\theta_0|>\delta}(||x(\theta) - x(\theta_0)|| - ||X - x(\theta_0)||)\} \end{aligned}$$

$$
\begin{aligned}
&= P_{\theta_0}^{(\varepsilon)}\{2||X - x(\theta_0)|| > \inf_{|\theta-\theta_0|>\delta} ||x(\theta) - x(\theta_0)||\} \\
&= P_{\theta_0}^{(\varepsilon)}\{||X - x(\theta_0)|| > \frac{1}{2} g(\delta)\}.
\end{aligned}
$$

Since the process X_t satisfies the stochastic differential equation (5.20), it follows that

$$
\begin{aligned}
X_t - x_t(\theta_0) &= x_0 + \theta_0 \int_0^t X_s ds + \varepsilon W_t^H - x_t(\theta_0) \\
&= \theta_0 \int_0^t (X_s - x_s(\theta_0)) ds + \varepsilon W_t^H
\end{aligned}
\tag{5. 29}
$$

since $x_t(\theta) = x_0 e^{\theta t}$. Let $U_t = X_t - x_t(\theta_0)$. Then it follows from the above equation that

$$
U_t = \theta_0 \int_0^t U_s \ \ ds + \varepsilon W_t^H. \tag{5. 30}
$$

Let $V_t = |U_t| = |X_t - x_t(\theta_0)|$. The above relation implies that

$$
V_t = |X_t - x_t(\theta_0)| \leq |\theta_0| \int_0^t V_s ds + \varepsilon |W_t^H|. \tag{5. 31}
$$

Applying Gronwall-Bellman Lemma, we obtain that

$$
\sup_{0 \leq t \leq T} |V_t| \leq \varepsilon e^{|\theta_0 T|} \sup_{0 \leq t \leq T} |W_t^H|. \tag{5. 32}
$$

Hence

$$
\begin{aligned}
P_{\theta_0}^{(\varepsilon)}\{||X - x(\theta_0)|| > \frac{1}{2} g(\delta)\} &\leq P\{\sup_{0 \leq t \leq T} |W_t^H| > \frac{e^{-|\theta_0 T|} g(\delta)}{2\varepsilon T}\} \\
&= P\{W_T^{H*} > \frac{e^{-|\theta_0 T|} g(\delta)}{2\varepsilon T}\}.
\end{aligned}
\tag{5. 33}
$$

Applying the Lemma 5.6 to the estimate obtained above, we get that

$$
\begin{aligned}
P_{\theta_0}^{(\varepsilon)}\{|\theta_\varepsilon^* - \theta_0| > \delta\} &\leq 2^p T^{pH+p} K(p,H) e^{|\theta_0 T|p} (g(\delta))^{-p} \varepsilon^p \\
&= O((g(\delta))^{-p} \varepsilon^p).
\end{aligned}
\tag{5. 34}
$$

Remarks: As a consequence of the above theorem, we obtain that θ_ε^* converges in probability to θ_0 under $P_{\theta_0}^{(\varepsilon)}$-measure as $\varepsilon \to 0$. Further more the rate of convergence is of the order $(O(\varepsilon^p))$ for every $p > 0$.

Asymptotic distribution We will now study the asymptotic distribution if any of the estimator θ_ε^* after suitable scaling. It can be checked that

$$
X_t = e^{\theta_0 t}\{x_0 + \int_0^t e^{-\theta_0 s} \varepsilon dW_s^H\} \tag{5. 35}
$$

or equivalently

$$
X_t - x_t(\theta_0) = \varepsilon e^{\theta_0 t} \int_0^t e^{-\theta_0 s} dW_s^H. \tag{5. 36}
$$

Let

(5. 37) $$Y_t = e^{\theta_0 t}\int_0^t e^{-\theta_0 s} dW_s^H.$$

Note that $\{Y_t, 0 \le t \le T\}$ is a gaussian process and can be interpreted as the "derivative" of the process $\{X_t, 0 \le t \le T\}$ with respect to ε. Applying Theorem 3.1, we obtain that, P-a.s.,

(5. 38) $$Y_t e^{-\theta_0 t} = \int_0^t e^{-\theta_0 s} dW_s^H = \int_0^t K_H^f(t,s) dM_s^H, t \in [0,T]$$

where $f(s) = e^{-\theta_0 s}, s \in [0,T]$ and M^H is the fundamental gaussian martingale associated with the fBm W^H. In particular it follows that the random variable $Y_t e^{-\theta_0 t}$ and hence Y_t has normal distribution with mean zero and further more, for any $h \ge 0$,

(5. 39)
$$\begin{aligned} Cov(Y_t, Y_{t+h}) &= e^{2\theta_0 t+\theta_0 h} E[\int_0^t e^{-\theta_0 u} dW_u^H \int_0^{t+h} e^{-\theta_0 v} dW_v^H] \\ &= e^{2\theta_0 t+\theta_0 h} H(2H-1)\int_0^t\int_0^t e^{-\theta_0(u+v)}|u-v|^{2H-2} du dv \\ &= e^{2\theta_0 t+\theta_0 h}\gamma_H(t) \ \text{(say)}. \end{aligned}$$

In particular

(5. 40) $$Var(Y_t) = e^{2\theta_0 t}\gamma_H(t).$$

Hence $\{Y_t, 0 \le t \le T\}$ is a zero mean gaussian process with $Cov(Y_t, Y_s) = e^{\theta_0(t+s)}\gamma_H(t)$ for $s \ge t$.

Let

(5. 41) $$\zeta = \arg \inf_{-\infty<u<\infty} \int_0^T |Y_t - utx_0 e^{\theta_0 t}| dt.$$

Theorem 5.8: The random variable converges in probability to a random variable whose probability distribution is the same as that of ζ under P_{θ_0}.

Proof: Let $x_t'(\theta) = x_0 t e^{\theta t}$ and let

(5. 42) $$Z_\varepsilon(u) = ||Y - \varepsilon^{-1}(x(\theta_0+\varepsilon u) - x(\theta_0))||$$

and

(5. 43) $$Z_0(u) = ||Y - ux'(\theta_0)||.$$

Further more, let

(5. 44) $$A_\varepsilon = \{\omega : |\theta_\varepsilon^* - \theta_0| < \delta_\varepsilon\}, \delta_\varepsilon = \varepsilon^\tau, \tau \in (\frac{1}{2}, 1), L_\varepsilon = \varepsilon^{\tau-1}.$$

Observe that the random variable $u_\varepsilon^* = \varepsilon^{-1}(\theta_\varepsilon^* - \theta_0)$ satisfies the equation

(5. 45) $$Z_\varepsilon(u_\varepsilon^*) = \inf_{|u|<L_\varepsilon} Z_\varepsilon(u), \omega \in A_\varepsilon.$$

Define

$$\zeta_\varepsilon = \arg \inf_{|u|<L_\varepsilon} Z_0(u). \tag{5. 46}$$

Observe that, with probability one,

$$\begin{aligned} \sup_{|u|<L_\varepsilon} |Z_\varepsilon(u) - Z_0(u)| &= |||Y - ux'(\theta_0) - \frac{1}{2}\varepsilon u^2 x''(\tilde{\theta})|| - ||Y - ux'(\theta_0)||| \\ &\leq \frac{\varepsilon}{2} L_\varepsilon^2 \sup_{|\theta-\theta_0<\delta_\varepsilon} \int_0^T |x''(\theta)| dt \\ &\leq C\varepsilon^{2\tau-1}. \end{aligned} \tag{5. 47}$$

Here $\tilde{\theta} = \theta_0 + \alpha(\theta - \theta_0)$ for some $\alpha \in (0,1)$. Note that the last term in the above inequality tends to zero as $\varepsilon \to 0$. Further more the process $\{Z_0(u), -\infty < u < \infty\}$ has a unique minimum u^* with probability one. This follows from the arguments given in Theorem 2 of Kutoyants and Pilibossian (1994). In addition, we can choose the interval $[-L, L]$ such that

$$P_{\theta_0}^{(\varepsilon)}\{u_\varepsilon^* \in (-L, L)\} \geq 1 - \beta g(L)^{-p} \tag{5. 48}$$

and

$$P\{u^* \in (-L, L)\} \geq 1 - \beta g(L)^{-p} \tag{5. 49}$$

where $\beta > 0$. Note that $g(L)$ increases as L increases. The processes $Z_\varepsilon(u), u \in [-L, L]$ and $Z_0(u), u \in [-L, L]$ satisfy the Lipschitz conditions and $Z_\varepsilon(u)$ converges uniformly to $Z_0(u)$ over $u \in [-L, L]$. Hence the minimizer of $Z_\varepsilon(.)$ converges to the minimizer of $Z_0(u)$. This completes the proof.

Remarks : We have seen earlier that the process $\{Y_t, 0 \leq t \leq T\}$ is a zero mean gaussian process with the covariance function

$$Cov(Y_t, Y_s) = e^{\theta_0(t+s)}\gamma_H(t)$$

for $s \geq t$. Recall that

$$\zeta = \arg \inf_{-\infty<u<\infty} \int_0^T |Y_t - utx_0 e^{\theta_0 t}| dt. \tag{5. 50}$$

It is not clear what the distribution of ζ is. Observe that for every u, the integrand in the above integral is the absolute value of a gaussian process $\{J_t, 0 \leq t \leq T\}$ with the mean function $E(J_t) = -utx_0 e^{\theta_0 t}$ and the covariance function

$$Cov(J_t, J_s) = e^{\theta_0(t+s)}\gamma_H(t)$$

for $s \geq t$.

6 Identification for linear stochasic systems driven by fBm

We now discuss the problem of nonparametric estimation or identification of the "drift" function $\theta(t)$ for a class of stochastic processes satisfying a stochastic differential equation

$$dX_t = \theta(t)X_t dt + dW_t^H, X_0 = \tau, t \geq 0 \tag{6. 1}$$

where τ is a gaussian random variable independent of the process $\{W_t^H\}$ which is a fBm with known Hurst parameter. We use the method of sieves and study the asymptotic properties of the estimator. Identification of nonstationary diffusion models by the method of sieves is studied in Nguyen and Pham (1982).

Estimation by the method of sieves We assume that $\theta(t) \in L^2([0,T], dt)$. In other words $X = \{X_t, t \geq 0\}$ is a stochastic process satisfying the stochastic integral equation

$$X(t) = \tau + \int_0^t \theta(s)X(s)ds + W_t^H, 0 \leq t \leq T. \tag{6. 2}$$

where $\theta(t) \in L^2([0,T], dt)$. Let

$$C_\theta(t) = \theta(t)\ X(t), 0 \leq t \leq T \tag{6. 3}$$

and assume that the sample paths of the process $\{C_\theta(t), 0 \leq t \leq T\}$ are smooth enough so that the process

$$Q_{H,\theta}(t) = \frac{d}{dw_t^H} \int_0^t k_H(t,s)C_\theta(s)ds, 0 \leq t \leq T \tag{6. 4}$$

is well-defined where w_t^H and $k_H(t,s)$ are as defined in (3.8) and (3.6) respectively. Suppose the sample paths of the process $\{Q_H(t), 0 \leq t \leq T\}$ belong almost surely to $L^2([0,T], dw_t^H)$. Define

$$Z_t = \int_0^t k_H(t,s)dX_s, 0 \leq t \leq T. \tag{6. 5}$$

Then the process $Z = \{Z_t, 0 \leq t \leq T\}$ is an $(\mathcal{F}_t)$-semimartingale with the decomposition

$$Z_t = \int_0^t Q_{H,\theta}(s)dw_s^H + M_t^H \tag{6. 6}$$

where M^H is the fundamental martingale defined by (3.9) and the process X admits the representation

$$X_t = X_0 + \int_0^t K_H(t,s)dZ_s \tag{6. 7}$$

where the function K_H is as defined by (3.11) with $f \equiv 1$. Let P_θ^T be the measure induced by the process $\{X_t, 0 \leq t \leq T\}$ when $\theta(.)$ is the true "drift" function. Following Theorem 3.3, we get that the Radon-Nikodym derivative of P_θ^T with respect to P_0^T is given by

$$\frac{dP_\theta^T}{dP_0^T} = \exp[\int_0^T Q_{H,\theta}(s)dZ_s - \frac{1}{2}\int_0^T Q_{H,\theta}^2(s)dw_s^H]. \tag{6. 8}$$

Suppose the process X is observable on $[0,T]$ and $X_i, 1 \leq i \leq n$ is a random sample of n independent observations of the process X on $[0,T]$. Following the representation of the Radon-Nikodym derivative of P_θ^T with respect to P_0^T given above, it follows that the log-likelihood function corresponding to the observations $\{X_i, 1 \leq i \leq n\}$ is given by

$$
\text{(6. 9)} \qquad \begin{aligned} L_n(X_1,\ldots,X_n;\theta) &\equiv L_n(\theta) \\ &= \sum_{i=1}^{n}\Big(\int_0^T Q_{H,\theta}^{(i)}(s)dZ_i(s) - \frac{1}{2}\int_0^T [Q_{H,\theta}^{(i)}]^2(s)dw_s^H\Big). \end{aligned}
$$

where the process $Q_{H,\theta}^{(i)}$ is as defined by the relation (6.4) for the process X_i. For convenience in notation, we write $Q_{i,\theta}(s)$ hereafter for $Q_{H,\theta}^{(i)}(s)$. Let $\{V_n, n \geq 1\}$ be an increasing sequence of subspaces of finite dimensions $\{d_n\}$ such that $\cup_{n\geq 1}V_n$ is dense in $L^2([0,T],dt)$. The method of sieves consists in maximizing $L_n(\theta)$ on the subspace V_n. Let $\{e_i\}$ be a set of linearly independent vectors in $L^2([0,T],dt)$ such that the set of vectors $\{e_1,\ldots,e_{d_n}\}$ is a basis for the subspace V_n for every $n \geq 1$. For $\theta \in V_n$, $\theta(.) = \sum_{j=1}^{d_n}\theta_j e_j(.)$, we have

$$
\text{(6. 10)} \qquad \begin{aligned} Q_{i,\theta}(t) &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)\theta(s)X_i(s)ds \\ &= \frac{d}{dw_t^H}\int_0^t k_H(t,s)[\sum_{j=1}^{d_n}\theta_j e_j(s)]X_i(s)ds \\ &= \sum_{j=1}^{d_n}\theta_j\frac{d}{dw_t^H}\int_0^t k_H(t,s)e_j(s)X_i(s)ds \\ &= \sum_{j=1}^{d_n}\theta_j\Gamma_{i,j}(t) \ \ \text{(say)}. \end{aligned}
$$

Further more

$$
\text{(6. 11)} \qquad \begin{aligned} \int_0^T Q_{i,\theta}(t)dZ_i(t) &= \int_0^T [\sum_{j=1}^{d_n}\theta_j\Gamma_{i,j}(t)]dZ_i(t) \\ &= \sum_{j=1}^{d_n}\theta_j\int_0^T \Gamma_{i,j}(t)dZ_i(t) \\ &= \sum_{j=1}^{d_n}\theta_j R_{i,j} \ \ \text{(say)} \end{aligned}
$$

and

$$
\text{(6. 12)} \qquad \begin{aligned} \int_0^T Q_{i,\theta}^2(t)dw_t^H &= \int_0^T [\sum_{j=1}^{d_n}\theta_j\Gamma_{i,j}(t)]^2 dw_t^H \\ &= \sum_{j=1}^{d_n}\sum_{k=1}^{d_n}\theta_j\theta_k\int_0^T \Gamma_{i,j}(t)\Gamma_{i,k}(t)dw_t^H \\ &= \sum_{j=1}^{d_n}\sum_{k=1}^{d_n}\theta_j\theta_k < R_{i,j}, R_{i,k} > \end{aligned}
$$

where $< .,. >$ denotes the quadratic covariation. Therefore the log-likelihood function corresponding to the observations $\{X_i, 1 \leq i \leq n\}$ is given by

$$
\begin{aligned}
(6.\ 13) \qquad L_n(\theta) &= \sum_{i=1}^{n}\left(\int_0^T Q_{i,\theta}(t)dZ_i(t) - \frac{1}{2}\int_0^T Q_{i,\theta}^2(t)dw_t^H\right) \\
&= \sum_{i=1}^{n}\left[\sum_{j=1}^{d_n}\theta_j R_{i,j} - \frac{1}{2}\sum_{j=1}^{d_n}\sum_{k=1}^{d_n}\theta_j\theta_k < R_{i,j}, R_{i,k} >\right] \\
&= n\left[\sum_{j=1}^{d_n}\theta_j B_j^{(n)} - \frac{1}{2}\sum_{j=1}^{d_n}\sum_{k=1}^{d_n}\theta_j\theta_k A_{j,k}^{(n)}\right]
\end{aligned}
$$

where

$$(6.\ 14) \qquad B_j^{(n)} = n^{-1}\sum_{i=1}^{n} R_{i,j}, \quad 1 \leq j \leq d_n$$

and

$$(6.\ 15) \qquad A_{j,k}^{(n)} = n^{-1}\sum_{i=1}^{n} < R_{i,j}, R_{i,k} >, \quad 1 \leq j,k \leq d_n.$$

Let $\theta^{(n)}, B^{(n)}$ and $A^{(n)}$ be the vectors and the matrix with elements $\theta_j, j = 1, \ldots, d_n$, $B_j^{(n)}, j = 1, \ldots, d_n$ and $A_{j,k}^{(n)}, j,k = 1, \ldots, d_n$ as defined above. Then the log-likelihood function can be written in the form

$$(6.\ 16) \qquad L_n(\theta) = n[B^{(n)}\theta^{(n)} - \frac{1}{2}\theta^{(n)\prime}A^{(n)}\theta^{(n)}].$$

Here α' denotes the transpose of the vector α. The restricted maximum likelihood estimator $\hat{\theta}^{(n)}(.)$ of $\theta(.)$ is given by

$$(6.\ 17) \qquad \hat{\theta}^{(n)}(.) = \sum_{j=1}^{d_n}\hat{\theta}_j^{(n)} e_j(.)$$

where

$$(6.\ 18) \qquad \hat{\theta}^{(n)} = (\hat{\theta}_1^{(n)}, \ldots, \hat{\theta}_{d_n}^{(n)})$$

is the solution of the equation

$$(6.\ 19) \qquad A^{(n)}\hat{\theta}^{(n)} = B^{(n)}.$$

Assuming that $A^{(n)}$ is invertible, we get that

$$(6.\ 20) \qquad \hat{\theta}^{(n)} = (A^{(n)})^{-1}B^{(n)}.$$

Asymptotic properties of the estimator $\hat{\theta}^{(n)}(.)$ are studied in Prakasa Rao (2003e). We do not go into the details here.

References

Abry, P. and Sellan, F. (1996) The wavelet-based synthesis for fractional Brownian motion proposed by F.Sellan and Y.Meyer: Remarks and fast implementation. *Applied and Computational Harmonic Analysis*, **3**, 377-383.

Azais, J.M. (1990) Condition for convergence of number of crossings to the local time. Application to stable processes with independent increments and to gaussian processes. *Probab. Math. Stat.*, **11**, 19-36.

Beran, J. and Terrin, N. (1994) Estimation of the long-memory parameter based on multivariate central limit theorem, *J. Time Ser. Anal.* ,**15**, 269-278.

Borwankaer, J.D., Kallianpur, G. and Prakasa Rao, B.L.S. (1971) The Bernstein- von Mises theorem for Markov processes, *Ann. Math. Statist.*, **42**, 1241-1253.

Buldyrev, S.V., Goldberger, A.L., Havlin, S., Peng, C.-K., Stanley, H.E., Stanley, M.H.R. and Simons, M. (1993) Fractal landscapes and molecular evolution: modeling the Moyosin heavy chain gene family. *Biophys. J.*, **65**, 2673-2679.

Chen, G., Hall, P. and Poskitt, D.S. (1995) Periodogram-based estimators of fractal properties. *Ann. Statist.*, **23**, 1684-1711.

Coeurjolly, Jean-Francois. (2001) Estimating the parameters of a fractional Brownian motion by discrete variations of its sample paths, *Statist. Infer. Stoch. Proc.*, **4**, 199-207..

Comte, F. (1996) Simulation and estimation of long memory continuous time models, *J. Time Ser. Anal.*, **17**, 19-36.

Constantine, A.G., Hall, P. (1994) Characterizing surface smoothness via estimation of effective fractal dimension, *J.Roy. Statist. Soc.*, **56**, 97-113.

Doukhan, P., Oppenheim, G. and Taqqu, M.S. (2003) *Theory of Long-Range Dependence*, Birkhauser, Boston.

Feuverger, A., Hall. P. and Wood, A.T.A. (1994) Estimation of the fractal index and fractal dimension of a gaussian process by counting the number of level crossings, *J. Time Ser. Anal.* , **15**, 587-606.

Flandrin, P., and Goncalves, P. (1994) From wavelets to time-scale energy distributions, In *Recent Advances in wavelet Analysis*, Ed. L.L.Schumaker and G.Webb, Academic Press, Boston, pp. 309-334.

Geweke, J. and Porter-Hudak, S. (1983) The estimation and application of long-memory time series models, *J. Time Ser. Anal.*, **4**, 221-238.

Goncalves, P and Flandrin, P. (1993) Bilinear time-scale analysis applied to a local scaling exponents estimation, In *Progress in Wavelet Analysis and Applications*, Ed. Y.Meyer and S.Roques, Frontieres, Paris, pp. 271-276.

Gripenberg, G. and Norros, I. (1996) On the prediction of fractional Brownian motion, *J. Appl. Prob.*, **33**, 400-410., Hall, P., Loul, H.L. and Turlach, B.A. (1997) Note on convergence rates of semiparametric estimators of dependence index, *Ann. Statist.*, **25**, 1725-1739.

Henry, M. and Zafforoni, P. (2003) The long-range dependence pardigm for macroeconomics and finance, In *Theory of Long-Range Dependence*, Ed. Doukhan, P., Oppenheim, G. and Taqnu, M.S., Birkhauser, Boston, pp. 417-438..

Hurst, H.E. (1951) Long term storage capacity of reservoirs (with discussion), *Trans. Amer. Soc. Civ. Eng.*, **116**, 770-808.

Jensen, M. (1998) An approximate wavelet MLE of short and long memory parameters, *Studies in Nonlinear Dynamics and Econometrics*, **3**, 239-253.

Kent, J.T. and Wood, A,T.A. (1997) Estimating the fractal dimension of a locally self-similar gaussian process by using increments, *J. Roy. Statist. Soc. B* , **59**, 679-700.

Kleptsyna, M.L. and Le Breton, A. (2002) Statistical analysis of the fractinal Ornstein-Uhlenbeck type process, *Statist. Infer. Stoch. Proc.*, **5**, 229-248.

Kleptsyna, M.L. and Le Breton, A. and Roubaud, M.-C. (2000) Parameter estimation and optimal filtering for fractional type stochastic systems, *Statist. Infer. Stoch. Proc.*, **3**, 173-182.

Kutoyants, Yu. and Pilibossian, P. (1994) On minimum L_1-norm estimate of the parameter of the Ornstein-Uhlenbeck process, *Statist. Probab. Lett.*, **20**, 117-123.

Le Breton, A. (1998) Filtering and parameter estimation in a simple linear model driven by a fractional Brownian motion, *Statist. Probab. Lett.*, **38**, 263-274.

Mc Coy, E.J. and Walden, A.T. (1996) Wavelet analysis and synthesis of stationary long memory processes, *J. Comput. Graphical Statist.*,**5**, 26-56.

Millar, P.W. (1984) A general approach to the optimality of the minimum distance estimators, *Trans. Amer. Math. Soc.*, **286**, 249-272.

Montanari, A. (2003) Long-range dependence in hydrology, In *Theory of Long-Range Dependence*, Ed. Doukhan, P., Oppenheim, G. and Taqqu, M.S., Birkhauser, Boston, pp. 461-472.

Nguyen, Hung T. and Pham, Tuan D. (1982) Identification of nonstationary diffusion model by the method of sieves, *SIAM J. Control and Optimization*, **20**, 603- 611.

Norros, I. (2003) Large deviations of queues with long-range dependent input, In *Theory of Long-Range Dependence*, Ed. Doukhan, P., Oppenheim, G. and Taqqu, M.S., Birkhauser, Boston, pp. 409-415.

Norros, I., Valkeila, E., and Virtamo, J. (1999) An elementary approach to a Girsanov type formula and other analytical results on fractional Brownian motion, *Bernoulli*, **5**, 571-587.

Novikov, A.A. (1972) Sequential estimation of the parameters of diffusion proocesses, *Mathematical Notes*, **12**, 812-818.

Novikov, A.A. and Valkeila, E. (1999) On some maximal inequalities for fractional Brownian motion, *Statist. Probab. Lett.*, **44**, 47-54.

Ossandik, S.M. Buldyrev, S.V., Goldberger, A.L., Havlin, S., Mantegna, R.N., Peng, C.-K., Simons, M. and Stanley, H.E. (1993) Correlation approach to identify coding regions in DNA sequences *Biophys. J.*, **67**, 64-70.

Peng, C.-K., Buldyrev, S.V., Goldberger, A.L., Havlin, S., Sciortino, F., Simons, M and Stanley, H.E. (1992) Long-range correlation in nucleotide sequences, *Nature*, **356**, 168-170.

Peng, C.-K., Hausdorff, J.M., Mietus, J.E., Havlin,S., Stanley, H.E. and Goldberger, A.L. (1995a) Fractals in physiological control from heartbeat to gait, In *Levy Flights and Related Phenomena in Physics*, Ed. M.F. Shlesinger,, G.M. Zaslavsky, and U.Frisch, Proceedings of the 1994 International Conference on Levy flights, Springer, Berlin, pp. 315-330.

Peng, C.-K., Havlin, S., Stanley, H.E., and Goldberger, A.L. (1995b) Quantification of scaling exponents and crossover phenomena in noinstationary heartbeat time series, *Chaos*, **5**, 82-87.

Percival, D.P. and Guttorp, P. (1994) Long-memory processes, the Allan variance and wavelets, In *Wavelets in Geophysics*, Ed. Foufoula-Georgiou, E. and kumar, P., Academic Press, New York, pp. 325-357.

Poggi, Jean-Michel.,Viano, Marie-Claude. (1998) An estimate of the fractal index using multiscale aggregates, *J. Time Ser. Anal.*, **19**, 221-233.

Prakasa Rao, B.L.S. (1966) *Asymptotic Distributions in Some Nonregular Statistical Problems*, Ph.D. Dissertation, Michigan State University.

Prakasa Rao, B.L.S. (1968) Estimation of the location of the cusp of a continuous density, *Ann. Math. statist.*, **39**, 76-87.

Prakasa Rao, B.L.S. (1969) Estimation of a unimodal density, *Sankhya Ser.A*, **31**, 23-36.

Prakasa Rao, B.L.S. (1970) Estimation for distributions with monotone failure rate, *Ann. Math. Statist.*, **41**, 507-519.

Prakasa Rao, B.L.S. (1981) The Bernstein- von Mises theorem for a class of diffusion processes, *Teor. Sluch. Proc.* ,**9**, 95-101 (In Russian).

Prakasa Rao, B.L.S. (1987) *Asymptotic Theory of Statistical Inference*, Wiley, New York.

Prakasa Rao, B.L.S. (1999a) *Statistical Inference for Diffusion Type Processes*, Arnold, London and Oxford University Press, New York.

Prakasa Rao, B.L.S. (1999b) *Semimartingales and Their Statistical Inference*, CRC Press, Boca Raton and Chapman and Hall, London.

Prakasa Rao, B.L.S. (2003a) Parametric estimation for linear stochastic differential equations driven by fractional Brownian motion, *Random Operators and Stochastic Equations* (To appear).

Prakasa Rao, B.L.S. (2003b) Berry-Esseen bound for MLE for linear stochastic differential equations driven by fractional Brownian motion, Preprint, Indian Statistical Institute, New Delhi.

Prakasa Rao, B.L.S. (2003c) Sequential estimation for fractional Ornstein-Uhlenbeck type process, *Sequential Analysis* (To appear).

Prakasa Rao, B.L.S. (2003d) Minimum L_1-norm estimation for fractional Ornstein-Uhlenbeck type process, Preprint, Indian Statistical Institute, New Delhi.

Prakasa Rao, B.L.S. (2003e) Identification for linear stochastic systems driven by fractional Brownian motion, Preprint, Indian Statistical Institute, New Delhi.

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

Samko, S.G., Kilbas, A.A., and Marichev, O.I. (1993) *Fractional Integrals and derivatives*, Gordon and Breach Science.

Taylor, C.C. and Taylor, S.J. (1991) Estimating the dimension of a fractal, *J. Roy. Statist. Soc. B*, **53**, 353-364.

Taqqu, M. (2003) Fractional Brownian motion and long-range dependence, In *Theory of Long-Range Dependence*, Ed. Doukhan, P., Oppenheim, G. and Taqqu, M.S., Birkhauser, Boston, pp.5-38.

Wang, Y., Cavanaugh. J.E., and Song, C. (2001) Self-similarity index estimation via wavelets for locally self-similar processes, *J. Stat. Plan. Infer.*, **99**, 91-110.

Willinger, W., Paxson, V., Riedi, R.H. and Taqqu, M. (2003) Long-range dependence and data network traffic, In *Theory of Long-Range Dependence*, Ed. Doukhan, P., Oppenheim, G. and Taqqu, M.S., Birkhauser, Boston, pp. 373-407.