Acta Applicandae Mathematicae **78:** 145–154, 2003.

Asymptotically Efficient Nonparametric Estimation of Nonlinear Spectral Functionals

M. S. GINOVIAN
Institute of Mathematics, Armenian National Academy of Sciences, Marshal Bagramian 24-B, Yerevan, Armenia. e-mail: mamgin@instmath.sci.am

(Received: 5 August 2002)

Abstract. The paper considers a problem of construction of asymptotically efficient estimators for functionals defined on a class of spectral densities. We define the concepts of H_0- and IK-efficiency of estimators, based on the variants of Hájek–Ibragimov–Khas'minskii convolution theorem and Hájek–Le Cam local asymptotic minimax theorem, respectively. We prove that $\Phi(\widehat{\theta}_T)$, where $\widehat{\theta}_T$ is a suitable sequence of $T^{1/2}$-consistent estimators of unknown spectral density $\theta(\lambda)$, is H_0- and IK-asymptotically efficient estimator for a nonlinear smooth functional $\Phi(\theta)$.

Mathematics Subject Classifications (2000): 60G10, 60G15, 62G05, 62M15.

Key words: stationary Gaussian process, efficient estimation, spectral density, nonlinear spectral functionals, periodogram, local asymptotic normality.

1. Introduction

Suppose we observe a finite realization $\mathbf{X}_T = \{X(1), \dots, X(T)\}$ of a zero mean real-valued stationary Gaussian process $X(t)$ with an *unknown* spectral density function $\theta(\lambda)$, $\lambda \in [-\pi, \pi]$. We assume that $\theta(\lambda)$ belongs to a given class $\mathbf{\Theta} \subset \mathbf{L}_p = \mathbf{L}_p[-\pi, \pi]$ $(p > 1)$ of spectral densities possessing some smoothness conditions. Let $\Phi(\cdot)$ be some *known* functional, the domain of definition of which contains $\mathbf{\Theta}$. The distribution of the process $X(t)$ is completely determined by the spectral density, and we consider $\theta(\lambda)$ as an infinite-dimensional 'parameter' on which the distribution of $X(t)$ depends.

The problem is to estimate the value $\Phi(\theta)$ of the functional $\Phi(\cdot)$ at an unknown point $\theta \in \mathbf{\Theta}$ on the basis of an observation $\mathbf{X}_T$, and to investigate the asymptotic (as $T \to \infty$) properties of the suggested estimators. The main objective is construction of asymptotically efficient estimators for $\Phi(\theta)$.

The problem of asymptotically efficient nonparametric estimation of different kind of functionals of a spectral density function has been considered by Millar [16], Ibragimov and Khas'minskii [9, 12], Dahlhaus and Wefelmeyer [2], Ginovian [6, 7]. The main restriction imposed on spectral density $\theta(\lambda)$ was existence of constants C_1 and C_2, such that $0 < C_1 \leqslant \theta(\lambda) \leqslant C_2 < \infty$. The objective of the present paper is to extend some results of these papers for broader class $\mathbf{\Theta}$

that contains spectral densities possessing singularities. Asymptotically efficient estimators we construct by a method suggested by Ibragimov and Khas'minskii [9, 12]. Our plan will be as follows:

- We define the concept of local asymptotic normality (LAN) in the spirit of Ibragimov and Khas'minskii [12], and derive conditions under which the underlying family of Gaussian distributions is LAN at a point $\theta_0 \in \mathbf{\Theta}$.
- Using LAN we state variants of Hájek–Le Cam local asymptotic minimax theorem and Hájek–Ibragimov–Khas'minskii convolution theorem.
- We define the concepts of H_0- and IK-asymptotically efficient estimators, and prove that the statistic $\Phi(\widehat{\theta_T})$, where $\widehat{\theta_T}$ is a suitable sequence of $T^{1/2}$-consistent estimators of unknown spectral density $\theta(\lambda)$, is H_0- and IK-asymptotically efficient estimator for a nonlinear smooth functional $\Phi(\theta)$.

2. Basic Definitions and Preliminary Results

2.1. THE MODEL

Statistical analysis of stationary Gaussian processes usually involves two type of conditions imposed on the spectral density $\theta(\lambda)$. Conditions of the first type controls the singularities of $\theta(\lambda)$, and describe the dependence structure of the underlying process $X(t)$. Conditions of the second type refer to smoothness properties of $\theta(\lambda)$. To specify the model we need the following definition [10].

DEFINITION 1. We say that a spectral density function $\theta(\lambda)$ satisfies the *Muckenhoupt condition* (A_2) (or has Muckenhoupt type singularities), if

$$\sup \frac{1}{|J|^2} \int_J \theta(\lambda)\, \mathrm{d}\lambda \int_J \frac{1}{\theta(\lambda)}\, \mathrm{d}\lambda < \infty, \tag{A_2}$$

where the supremum is over all intervals $J \subset [-\pi, \pi]$, and $|J|$ denotes the length of J. The class of spectral densities satisfying (A_2) we denote by $\mathcal{A}_2$.

Remark 1. The spectral densities from the class $\mathcal{A}_2$ may possess singularities. In particular, $\mathcal{A}_2$ contains functions of the form $\theta(\lambda) \sim |\lambda|^\alpha$, $-1/2 < \alpha < 1/2$.

Remark 2. In fact (A_2) is a weak dependence condition, meaning that the maximal coefficient of correlation between the 'past' and the 'future' of the process $X(t)$ is less than 1. The class of processes $X(t)$ with spectral densities $\theta(\lambda)$ satisfying (A_2) is contained in the class of linearly regular processes and contains the class of completely regular processes.

Given numbers $0 < \alpha < 1$, $r \in \mathbb{N}_0$ ($\mathbb{N}_0$ is the set of nonnegative integers). We put $\beta = r + \alpha$, and denote by $\mathbf{H}_p(\beta)$ the Hölder class of functions, i.e. the class of functions $\psi(\lambda) \in \mathbf{L}_p$, which have rth derivatives in $\mathbf{L}_p$ and satisfy

$$\|\psi^{(r)}(\cdot + u) - \psi^{(r)}(\cdot)\|_p \leqslant C|u|^\alpha,$$

where C is a positive constant. Also by $\mathbf{\Sigma}_p(\beta)$ we denote the set of all spectral densities which belong to the class $\mathbf{H}_p(\beta)$.

The basic assumption on the observed process $X(t)$ is the following.

ASSUMPTION 1. $X(t)$ $(t \in \mathbb{Z})$ is a zero mean real-valued stationary Gaussian process with a spectral density $\theta(\lambda)$ satisfying Muckenhoupt condition (A_2) and belonging to a Hölder class $\mathbf{\Sigma}_p(\beta)$.

2.2. LOCAL ASYMPTOTIC NORMALITY

The notion of local asymptotic normality (LAN) of families of distributions plays an important role in asymptotic estimation theory. Le Cam, Hájek, Ibragimov and Khas'minskii and others have shown (see, for instance, Ibragimov and Khas'minskii [11, 12], Le Cam [14]) that many important properties of statistical estimators (characterization of limiting distributions, lower bounds on the accuracy, asymptotic efficiency, etc.) follow from LAN condition. The significance of LAN for nonparametric estimation problems has been emphasized by Levit [15], Millar [16], Ibragimov and Khas'minskii [12] and others. The LAN condition for families of distributions generated by a stationary Gaussian process with spectral density depending on a finite-dimensional parameter has been studied by Davies [3], Dzhaparidze [4] and Ginovian [5]. In [12] Ibragimov and Khas'minskii suggested a new definition of LAN concept of families of distributions in the case where the parametric set is a subset of an infinite-dimensional normed space.

Let $\mathbb{P}_{T,\theta}$ be the probability distribution of the vector $\mathbf{X}_T = \{X(1), \dots, X(T)\}$ with spectral density $\theta(\lambda)$. The next definition follows Ibragimov and Khas'minskii [12] (see also Ginovian [8]).

DEFINITION 2. A family of distributions $\{\mathbb{P}_{T,\theta},\ \theta \in \mathbf{\Theta}\}$ is called *locally asymptotically normal* (*LAN*) at $\theta_0 \in \mathbf{\Theta}$ in the direction $\mathbf{L}_2$ with norming factors $A_T = A_T(\theta_0)$ if there exist a linear manifold $H_0 = H_0(\theta_0) \subset \mathbf{L}_2$ with closure $\overline{H_0} = \mathbf{L}_2$ and a family $\{A_T\}$ of linear operators $A_T\colon \mathbf{L}_2 \to \mathbf{L}_2$ that satisfy:

(1) for any $h \in H_0$, $\|A_T h\|_2 \to 0$ as $T \to \infty$, where $\|\cdot\|_2$ denotes the L_2-norm;
(2) for any $h \in H_0$ there is $T(h)$ such that $\theta_0 + A_T h \in \mathbf{\Theta}$ for all $T > T(h)$;
(3) for any $h \in H_0$ and $T > T(h)$, the representation

$$\ln \frac{\mathrm{d}\mathbb{P}_{T,\theta_0+A_T h}}{\mathrm{d}\mathbb{P}_{T,\theta_0}}(\mathbf{X}_T) = \Delta_T(h,\theta_0) - \tfrac{1}{2}\|h\|_2^2 + \phi(T,h,\theta_0)$$

is valid, where $\Delta_T(h) = \Delta_T(h,\theta_0)$ is a random linear function on H_0 asymptotically (as $T \to \infty$) $N(0, \|h\|_2^2)$ – normally distributed for any $h \in H_0$ and $\phi(T,h,\theta_0) \to 0$ as $T \to \infty$ in $\mathbb{P}_{T,\theta_0}$ – probability.

Note that the presence of LAN property depends on the point θ_0, the space $\mathbf{L}_2$ and the family $\{A_T\}$. As regards $H_0 = H_0(\theta_0)$, we need only that $\overline{H_0} = \mathbf{L}_2$.

DEFINITION 3. We say that a pair of functions $(f(\lambda), g(\lambda))$ satisfies *condition* $(\mathcal{H}_1)$, if $f(\lambda) \in \mathbf{\Sigma}_p(\beta)$ for $1 < p \leqslant 2$ and $\beta > 1/p$, and $g(\lambda) \in \mathbf{L}_q$, where q is the conjugate of p: $1/p + 1/q = 1$.

We will always assume that $\mathbf{\Theta}$ is a subset of the space $\mathbf{L}_p$ $(p \geqslant 1)$ consisting of spectral densities satisfying Muckenhoupt condition (A_2) and belonging to the Hölder class $\mathbf{\Sigma}_p(\beta)$. Define $H_0 = H_0(\theta)$ to be the linear manifold consisting of bounded on $[-\pi, \pi]$ functions $h(\lambda)$ such that the pair $(\theta, h\theta^{-1})$ satisfies the condition $(\mathcal{H}_1)$. We also define A_T: $\mathbf{L}_2 \to \mathbf{L}_2$ by $A_T h = [T^{-1/2}\theta] \cdot h$, that is, A_T is the operator of multiplication by function $T^{-1/2}\theta(\lambda)$. As an immediate consequence of Theorem 1 in [8] we have

THEOREM 1. *Let* $\mathbf{\Theta}$, H_0 *and* A_T *be defined as above. Then the family of distributions* $\{\mathbb{P}_{T,\theta},\ \theta \in \mathbf{\Theta}\}$ *satisfies LAN condition at any point* $\theta \in \mathbf{\Theta}$ *in the direction* $\mathbf{L}_2$ *with norming factors* A_T *and*

$$\Delta_T(h) = \frac{T^{1/2}}{4\pi} \int_{-\pi}^{\pi} \frac{I_T(\lambda) - \theta(\lambda)}{\theta(\lambda)} h(\lambda)\, \mathrm{d}\lambda, \tag{1}$$

where

$$I_T(\lambda) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} X(t) \mathrm{e}^{-i\lambda t} \right|^2 \tag{2}$$

is the periodogram of the process $X(t)$.

2.3. CHARACTERIZATION OF LIMITING DISTRIBUTION. H_0-EFFICIENCY

We now consider the problem of estimating the value $\Phi(\theta)$ of a known functional $\Phi(\cdot)$ at an unknown point $\theta \in \mathbf{\Theta}$ on the basis of an observation $\mathbf{X}_T$, which has distribution $\mathbb{P}_{T,\theta}$. We assume that the family $\{\mathbb{P}_{T,\theta},\ \theta \in \mathbf{\Theta}\}$ satisfies the LAN condition at a point $\theta_0 = f \in \mathbf{\Theta}$ in the direction $\mathbf{L}_2$ with norming factors A_T. We will also assume the functional $\Phi(\theta)$ to be Fréchet differentiable with derivative $\Phi'(\theta)$ satisfying the condition: for $f \in \mathbf{\Theta}$

$$0 < \|\Phi'(f) f\|_2 < \infty. \tag{3}$$

Let $\widehat{\Phi}_T$ be a statistical estimator of $\Phi(\theta)$, i.e. a measurable mapping $\widehat{\Phi}_T = \widehat{\Phi}_T(\mathbf{X}_T)$: $\mathbb{R}^T \to \mathbb{R}^1$. We need a version of Hájek–Ibragimov–Khas'minskii convolution theorem.

Recall that a statistical estimator $\widehat{\Phi}_T$ of $\Phi(\theta)$ is called H_0-*regular* at $\theta_0 \in \mathbf{\Theta}$, if for any $h \in H_0$ there exists a proper limit distribution F of $T^{1/2}(\widehat{\Phi}_T - \Phi(\theta_h))$, where $\theta_h = \theta_0 + A_T h$, and this limit distribution does not depend on h.

The next theorem follows from Theorem 1 and Theorem 3.1 in [12].

THEOREM 2. *Let $\widehat{\Phi}_T$ be a H_0-regular estimator of $\Phi(\theta)$ at $f \in \mathbf{\Theta}$. Assume that the pair $(f, \Phi'(f))$ satisfies* (3). *Then under the assumptions of Theorem* 1 *the limit distribution F of $T^{1/2}(\widehat{\Phi}_T - \Phi(f))$ is a convolution of a probability distribution G and a centered normal distribution with variance $\|\Phi'(f)f\|_2^2$:*

$$F = N(0, \|\Phi'(f)f\|_2^2) * G. \tag{4}$$

By a lemma of Anderson (see, e.g., [11], Section 2.10), the distribution F in (4) is less concentrated in symmetric intervals than $N(0, \|\Phi'(f)f\|_2^2)$. This justifies the following definition of H_0-efficiency (cf. [2, 16]).

DEFINITION 4. Let the family $\{\mathbb{P}_{T,\theta},\ \theta \in \mathbf{\Theta}\}$ be LAN at a point $f \in \mathbf{\Theta}$. An estimator $\widehat{\Phi}_T$ of $\Phi(\theta)$ is called *H_0-asymptotically efficient* at f (in the class of H_0-regular estimators) with asymptotic variance $\sigma^2 = \|\Phi'(f)f\|_2^2$, if

$$\mathcal{L}\{T^{1/2}(\widehat{\Phi}_T - \Phi(\theta_h)) \mid \mathbb{P}_{T,\theta_h}\} \Longrightarrow N(0, \sigma^2) \quad \text{as } T \to \infty,$$

that is the distribution G in (4) is degenerate.

2.4. A LOWER BOUND FOR THE ASYMPTOTIC MINIMAX RISK. IK-EFFICIENCY

Denote by $\mathbf{\Phi}_T$ the set of all estimators of $\Phi(\theta)$ constructed on the basis of an observation $\mathbf{X}_T$. Let $\mathbf{W}$ denote the set of all loss functions w: $\mathbb{R}^1 \to \mathbb{R}^1$, which are symmetric and nondecreasing on $(0, \infty)$, and satisfy $w(x) \geqslant 0$, $w(0) = 0$. Also by $\mathbf{W}_e$ we denote the subset of loss functions $w \in \mathbf{W}$ which for some constants $C_1, C_2 > 0$ satisfy the condition $w(x) \leqslant C_1 \exp\{C_2|x|\}$. The next theorem, which is an immediate consequence of Theorem 1 and Theorem 4.1 in [12], contains a minimax lower bound for the risk of all possible estimators $\widehat{\Phi}_T$ of $\Phi(\cdot)$ in the neighborhood of $f \in \mathbf{\Theta}$ (cf. [5, 9]).

THEOREM 3. *Assume that the pair $(f, \Phi'(f))$ satisfies* (3). *Then under the assumptions of Theorem* 1, *for all $w \in \mathbf{W}$*

$$\liminf_{\delta\to 0} \lim_{T\to\infty} \inf_{\widehat{\Phi}_T \in \mathbf{\Phi}_T} \sup_{\|\theta - f\|_2 < \delta} \mathbb{E}_\theta\{w(T^{1/2}(\widehat{\Phi}_T - \Phi(f)))\} \geqslant \mathbb{E}w(\xi),$$

where ξ is a centered normal random variable with variance $\|\Phi'(f)f\|_2^2$.

Next, we define the notion of asymptotically efficient estimators in the spirit of Ibragimov and Khas'minskii (IK-efficiency) [9, 12].

DEFINITION 5. Let the family $\{\mathbb{P}_{T,\theta},\ \theta \in \mathbf{\Theta}\}$ be LAN at $f \in \mathbf{\Theta}$. An estimator $\widehat{\Phi}_T$ of $\Phi(\theta)$ is called *IK-asymptotically efficient* at f for the loss function $w(x) \in \mathbf{W}$, with asymptotic variance $\sigma^2 = \|\Phi'(f)f\|_2^2$, if

$$\liminf_{\delta\to 0} \lim_{T\to\infty} \sup_{\|\theta - f\|_2 < \delta} \mathbb{E}_\theta\{w(T^{1/2}(\widehat{\Phi}_T - \Phi(f)))\} = \mathbb{E}w(\xi),$$

where ξ is a centered normal random variable with variance $\|\Phi'(f)f\|_2^2$.

3. Asymptotically Efficient Estimators

First assume that the functional $\Phi(f)$ is linear and continuous in $\mathbf{L}_p[-\pi,\pi]$, $p>1$. It is well-known (see, e.g., [13]) that $\Phi(f)$ admits the representation

$$\Phi(f)=\int_{-\pi}^{\pi} f(\lambda)g(\lambda)\,\mathrm{d}\lambda, \tag{5}$$

where $g(\lambda)\in L_q;\ 1/p+1/q=1$. As an estimator for the functional $\Phi(f)$ we consider the averaged periodogram statistics:

$$\widehat{\Phi}_T=\Phi(I_T)=\int_{-\pi}^{\pi} I_T(\lambda)g(\lambda)\,\mathrm{d}\lambda, \tag{6}$$

where $I_T(\lambda)$ is the periodogram of $X(t)$ defined by (2).

THEOREM 4. *Let* $\Phi(f)$ *and* $\widehat{\Phi}_T$ *be defined by* (5) *and* (6). *Assume that the pair of functions* (f,g) *satisfies the conditions* $(\mathcal{H}_1)$ *and* $0<\|fg\|_2<\infty$. *Then under the assumptions of Theorem* 1, *the statistics* $\widehat{\Phi}_T$ *is*

(a) *H_0-regular and H_0-asymptotically efficient estimator of* $\Phi(f)$ *with asymptotic variance* $\|fg\|_2^2$;
(b) *IK-asymptotically efficient estimator of* $\Phi(f)$ *for* $w(x)\in\mathbf{W}_e$ *with asymptotic variance* $\|fg\|_2^2$.

The problem of asymptotically efficient estimation becomes somewhat more complicated for nonlinear functionals. In this case the statistics $\Phi(I_T)$ is not necessary a consistent estimator for the functional $\Phi(f)$, and hence instead of the periodogram $I_T(\lambda)$, we need to use a suitable sequence of consistent estimators $\widehat{f_T}$ of f (see [2, 9]). On the other hand, if $\widehat{f_T}$ is a sequence of consistent estimators for f, the estimators $\Phi(\widehat{f_T})$, in general, will converge to $\Phi(f)$ too slowly to be asymptotically efficient (cf. [9]).

We consider a sequence $\{\widehat{f_T}\}$ of estimators for f which are consistent of order $T^{2\alpha}$ $(\alpha\leqslant 1/2)$, and derive conditions under which the statistics $\widehat{\Phi}_T=\Phi(\widehat{f_T})$ is asymptotically efficient estimator for $\Phi(f)$. (An estimator $\widehat{f_T}$ of f is called $T^{2\alpha}$-consistent with asymptotic variance σ^2, if $\lim_{T\to\infty}T^{2\alpha}\mathbb{E}(\widehat{f_T}-f)^2=\sigma^2$.)

We assume that $f\in\mathbf{\Sigma}_p(\beta)$, and as an estimator for unknown f we take

$$\widehat{f_T}(\lambda)=\int_{-\pi}^{\pi} W_T(\lambda-\mu)I_T(\mu)\,\mathrm{d}\mu. \tag{7}$$

For the kernel $W_T(\lambda)$ we set down the following assumptions ([2, 9, 17]).

ASSUMPTION 2. $W_T(\lambda)=M_TW(M_T\lambda)$, where $M_T=\mathrm{O}(T^\alpha)$. The choice of α $(0<\alpha<1)$ will depend on the appriori knowledge about f and Φ.

ASSUMPTION 3. $W(\lambda)$ is nonnegative bounded even function with $W(\lambda) \equiv 0$ for $|\lambda| > 1$ and with $r = [\beta]$,

$$\int_{-1}^{1} W(\lambda)\,\mathrm{d}\lambda = 1, \qquad \int_{-1}^{1} \lambda^k W(\lambda)\,\mathrm{d}\lambda = 0, \quad k = 1, 2, \ldots, r.$$

We assume the functional $\Phi(\cdot)$ to be Fréchet differentiable with derivative $\Phi'(\cdot)$ satisfying (3) and a Hölder condition: there exist constants $C > 0$ and $\delta > 0$ such that for any $f_1, f_2 \in \mathbf{L}_2$,

$$\|\Phi'(f_1) - \Phi'(f_2)\| \leqslant C\|f_1 - f_2\|_2^{\delta}. \tag{8}$$

THEOREM 5. *Let the spectral density $f(\cdot)$ and the functional $\Phi(\cdot)$ be such that:*

(i) *the pair $(f, \Phi'(f))$ satisfies conditions $(\mathcal{H}_1)$ and* (3);
(ii) *$\Phi(\cdot)$ satisfies condition* (8) *with $\delta \geqslant (2\beta - 1)^{-1}$.*

Let the estimator $\widehat{f_T}$ for f be defined by (7) *with the kernel $W_T(\lambda)$ satisfying Assumptions* 2 *and* 3 *with $\frac{1}{2\beta} < \alpha < \frac{\delta}{\delta+1}$. Then under the conditions given in Theorem* 1 *the statistics $\Phi(\widehat{f_T})$ is:*

(a) *H_0-regular and H_0-asymptotically efficient estimator for $\Phi(f)$ with asymptotic variance $\|\Phi'(f)f\|_2^2$;*
(b) *IK-asymptotically efficient estimator of $\Phi(f)$ for $w(x) \in \mathbf{W}_e$ with asymptotic variance $\|\Phi'(f)f\|_2^2$.*

4. Proofs

In this section we outline the proof of Theorems 5. The following lemma is essential for the proof of assertion (a) of the theorem (see [2]).

LEMMA 1. *Assume that the family $\{\mathbb{P}_{T,\theta},\ \theta \in \boldsymbol{\Theta}\}$ is LAN at $f \in \boldsymbol{\Theta}$. Then an estimator $\widehat{\Phi}_T$ of $\Phi(f)$ is H_0-regular and H_0-asymptotically efficient at f with asymptotic variance $\|f\Phi'(f)\|_2^2$ if and only if*

$$T^{1/2}[\widehat{\Phi}_T - \Phi(f)] - \Delta_T(f\Phi'(f)) = \mathrm{o}_P(1) \quad \textit{as } T \to \infty,$$

where $\Delta_T(\cdot)$ is defined by (1).

The proof of the next lemma is similar to the proof of Theorem 3 in [6].

LEMMA 2. *Let $\Phi(f)$ and $\widehat{\Phi}(I_T)$ be defined by* (5) *and* (6) *and let the pair (f, g) satisfy the conditions $(\mathcal{H}_1)$ and $0 < \|fg\|_2 < \infty$. Then for $w \in \mathbf{W}_e$*

$$\lim_{T\to\infty} \mathbb{E}_f\{w(T^{1/2}(\widehat{\Phi}(I_T) - \Phi(f)))\} = \mathbb{E}w(\xi),$$

where ξ is a centered normal random variable with variance $\|fg\|_2^2$.

LEMMA 3. *Assume that $f \in \mathbf{\Sigma}_p(\beta)$, and let $\widehat{f}_T$ be defined by* (7) *with the kernel $W_T(\lambda)$ satisfying Assumptions* 2 *and* 3 *with $1/2\beta < \alpha < \delta/(\delta+1)$, then*

$$T^{1/2}\|\widehat{f}_T - f\|_2^{1+\delta} = \mathrm{o}_P(1) \quad \text{as } T \to \infty.$$

Proof. Using the arguments of the proof of Theorem 3.1 in [1] we get

$$\|\widehat{f}_T - f\|_2 = \mathrm{O}_P\big(M_T^{1/2}T^{-1/2}\big) + \mathrm{O}_P\big(M_T^{-\beta}\big).$$

Hence, taking into account that $M_T = \mathrm{O}(T^{\alpha})$ we can write

$$T^{1/2}\|\widehat{f}_T - f\|_2^{1+\delta} = \mathrm{O}_P\big(T^{\frac{1}{2}+(\frac{\alpha}{2}-\frac{1}{2})(1+\delta)}\big) + \mathrm{O}_P\big(T^{\frac{1}{2}-\alpha\beta(1+\delta)}\big). \tag{9}$$

The assumptions imply

$$\frac{1}{2} + \left(\frac{\alpha}{2} - \frac{1}{2}\right)(1+\delta) < 0 \quad \text{and} \quad \frac{1}{2} - \alpha\beta(1+\delta) < 0.$$

Hence both terms in (9) are $\mathrm{o}_P(1)$ as $T \to \infty$, and the result follows. □

LEMMA 4. *Assume that $f \in \mathbf{\Sigma}_p(\beta)$, and Assumptions* 2 *and* 3 *are satisfied. Let $\psi(\lambda)$ be a continuous even function on $[-\pi, \pi]$ such that the pair (f, ψ) satisfies the conditions $(\mathcal{H}_1)$ and $0 < \|f\psi\|_2 < \infty$. Then the distribution of the random variable*

$$\eta_T = T^{1/2}\int_{-\pi}^{\pi} \psi(\lambda)[\widehat{f}_T(\lambda) - f(\lambda)]\,\mathrm{d}\lambda$$

as $T \to \infty$ tends to the normal distribution $N(0, \sigma^2)$, where

$$\sigma^2 = 4\pi\int_{-\pi}^{\pi} \psi^2(\lambda) f^2(\lambda)\,\mathrm{d}\lambda.$$

Proof. By Lemma 2 the distribution of the random variable

$$\xi_T = T^{1/2}\int_{-\pi}^{\pi} \psi(\lambda)[I_T(\lambda) - f(\lambda)]\,\mathrm{d}\lambda$$

tends to $N(0, \sigma^2)$. To complete the proof it is enough to show that

$$|\xi_T - \eta_T| = \mathrm{o}_P(1) \quad \text{as } T \to \infty. \tag{10}$$

Putting $M_T(\lambda - \mu) = t$ with some easy calculations we have

$$\begin{aligned}\eta_T &= T^{1/2}\int_{-\pi}^{\pi} \Psi_T(\lambda)(I_T(\lambda) - f(\lambda))\,\mathrm{d}\lambda + \\ &\quad + T^{1/2}\int_{-\pi}^{\pi} \psi(\lambda)\left[\int_{-\pi}^{\pi} f(\mu)W_T(\lambda-\mu)\,\mathrm{d}\mu - f(\mu)\right]\mathrm{d}\lambda \\ &= \eta_T^{(1)} + S_T \quad \text{(say)},\end{aligned} \tag{11}$$

where

$$\Psi_T(\lambda) = \int_{M_T(-\pi-\lambda)}^{M_T(\pi-\lambda)} \psi\left(\lambda + \frac{t}{M_T}\right) W_T(t)\,\mathrm{d}t - \psi(\lambda).$$

Along the lines of the proof of Theorem 3 in [17] and Lemma 2 we obtain

$$|\xi_T - \eta_T^{(1)}| = \mathrm{o}_P(1) \quad \text{as } T \to \infty. \tag{12}$$

Applying Hölder and Minkowski generalized inequalities we can show that

$$|S_T| \leqslant T^{1/2}\|\psi\|_q(1 + \|W_T\|_1)E_{M_T,p}(f), \tag{13}$$

where $E_{m,p}(f)$ is the best approximation of f by polynomials of degree m in the metric of L_p. The condition $f \in \mathbf{\Sigma}_p(\beta)$ implies $E_{m,p}(f) \leqslant Cm^{-\beta}$. Since by assumption $\alpha > \frac{1}{2\beta}$, in view of (13) we get

$$S_T = \mathrm{O}\big(T^{1/2}M_T^{-\beta}\big) = \mathrm{O}\big(T^{1/2-\alpha\beta}\big) \to 0 \quad \text{as } T \to \infty. \tag{14}$$

A combination of (11), (12) and (14) yields (10). Lemma 4 is proved. □

Proof of Theorem 5. It follows from (8) that (cf. [13], p. 454)

$$\left|\Phi(\widehat{f_T}) - \Phi(f) - \int_{-\pi}^{\pi} \Phi'(f)(\lambda)(\widehat{f_T}(\lambda) - f(\lambda))\,\mathrm{d}\lambda\right|$$
$$\leqslant \|\widehat{f_T} - f\| \sup_{0\leqslant\theta\leqslant 1} \|\Phi'(f + \theta(\widehat{f_T} - f)) - \Phi'(f)\| \leqslant C\|\widehat{f_T} - f\|^{1+\delta}.$$

Therefore

$$T^{1/2}[\Phi(\widehat{f_T}) - \Phi(f)] = T^{1/2}\int_{-\pi}^{\pi} \Phi'(f)(\lambda)\big[\widehat{f_T}(\lambda) - f(\lambda)\big]\,\mathrm{d}\lambda + {} + \mathrm{O}\big(T^{1/2}\|\widehat{f_T} - f\|^{1+\delta}\big) \quad \text{as } T \to \infty.$$

Hence, by Lemmas 3 and 4 we can write

$$T^{1/2}[\Phi(\widehat{f_T}) - \Phi(f)] = T^{1/2}\int_{-\pi}^{\pi} \Phi'(f)(\lambda)[I_T(\lambda) - f(\lambda)]\,\mathrm{d}\lambda + \mathrm{o}_P(1). \tag{15}$$

Now the assertion (a) of the theorem follows from Lemma 1 and (15), while the assertion (b) is an immediate consequence of Lemma 2 and (15). Theorem 5 is proved. □

References

1. Bentkus, R. J. and Rudzkis, R. A.: On the distributions of some statistical estimates of spectral density, *Theory Probab. Appl.* **27** (1982), 739–756.
2. Dahlhaus, R. and Wefelmeyer, W.: Asymptotically optimal estimation in misspecified time series models, *Ann. Statist.* **24** (1996), 952–974.

3. Davies, B. R.: Asymptotic inference in stationary time-series, *Adv. Appl. Probab.* **5** (1973), 469–497.
4. Dzhaparidze, K. O.: *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*, Springer-Verlag, New York, 1986.
5. Ginovian, M. S.: Local asymptotic normality of families of Gaussian distributions, *Zapiski Nauchn. Semin. LOMI* **136** (1984), 13–27.
6. Ginovian, M. S.: Asymptotically efficient nonparametric estimation of functionals of spectral density with zeros, *Theory Probab. Appl.* **33** (1988), 315–322.
7. Ginovian, M. S.: On Toeplitz type quadratic functionals in Gaussian stationary process, *Probab. Theory Related Fields* **100** (1994), 395–406.
8. Ginovian, M. S.: Locally asymptotically normal families of Gaussian distributions, *J. Contemp. Math. Anal.* **34** (1999), 18–29.
9. Has'minskii, R. Z. and Ibragimov, I. A.: Asymptotically efficient nonparametric estimation of functionals of a spectral density function, *Probab. Theory Related Fields* **73** (1986), 447–461.
10. Hunt, R. A., Muckenhoupt, B. and Wheeden, R. L.: Weighted norm inequalities for the conjugate function and Hilbert transform, *Trans. Amer. Math. Soc.* **76** (1973), 227–251.
11. Igragimov, I. A. and Khasminskii, R. Z.: *Statistical Estimation: Asymptotic Theory*, Springer-Verlag, New York, 1981.
12. Ibragimov, I. A. and Khas'minskii, R. Z.: Asymptotically normal families of distributions and efficient estimation, *Ann. Statist.* **19** (1991), 1681–1724.
13. Kolmogorov, A. N. and Fomin, S.: *Elements of Functional Analysis*, Nauka, Moscow, 1972.
14. Le Cam, L.: *Asymptotic Methods in Statistical Decision Theory*, Springer-Verlag, New York, 1986.
15. Levit, B. Ya.: On optimality of some statistical estimates, In: *Proc. Prague Sympos. Asymptotic. Statistics II*, 1974, pp. 215–238.
16. Millar, P. W.: Non-parametric applications of an infinite dimensional convolution theorem, *Z. Wahrshch. Verw. Geb.* **68** (1985), 545–556.
17. Taniguchi, M.: Minimum contrast estimation for spectral densities of stationary processes, *J. Roy. Statist. Soc. Ser. B* **49** (1987), 315–325.