ESTIMATION OF ANALYTIC FUNCTIONS

I. IBRAGIMOV[1]

St.Petersburg Branch of Steklov Mathematical Institute Russian Ac.Sci.

In this paper we present a review of some results on nonparametric estimation of analytic functions and in particular derive minimax bounds under different conditions on these functions.

AMS subject classifications: 62G05, 62G20.
Keywords and phrases: Nonparametric estimation, Minimax bounds.

1 Introduction

The aim of this paper is to present a review of some results about nonparametric estimation of analytic functions. A part of it is written in exposition style and summarizes some recent work of the author on the subject (detailed versions have already been published). The rest of the paper contains new results in the area. Sometimes the proofs are only outlined and will be published elsewhere.

Generally the problem looks as follows. We are given a class $\mathbf{F}$ of functions defined on a region $D \subset R^d$ and analytic in a vicinity of D. It means that all $f \in \mathbf{F}$ admit analytic continuation into a domain $G \supset D$ of the complex space C^d. To esimate an unknown function $f \in \mathbf{F}$ one makes observations X_ε. Consider as risk functions of estimators $\hat{f}$ for f the averaged $L_p(D)$-norms

$$\mathbf{E}_f\|\hat{f} - f\|_p = \mathbf{E}_f \left(\int_D |\hat{f}(t) - f(t)| dt \right)^{1/p}, \qquad 1 \leq p < \infty,$$
$$\mathbf{E}_f\|\hat{f} - f\|_\infty = \mathbf{E}\{\sup_{x \in D} |\hat{f}(x) - f(x)|\},$$

where in the case of noncontinuous $\hat{f}$ the supremum is understood as an essential supremum. Put

$$\Delta_p(\varepsilon, \mathbf{F}) = \Delta(\mathbf{F}) = \inf_{\hat{f}} \sup_{f \in \mathbf{F}} \mathbf{E}_f\|\hat{f} - f\|_p.$$

[1] This work is partly supported by the Russian Foundation for Basic Research, grants 99-01-00111, 96-15-96199, 00-15-96019 and INTAS, grant 95-0099, 99-1317.

The last expression gives some understanding of our estimation problem and below we study the asymptotic ($\varepsilon \to 0$, say) behaviour of $\Delta_p(\varepsilon, \mathbf{F})$. The rate of convergence of Δ depends on $\mathbf{F}$ and the understanding of this dependence is the main interest of the paper. Of course the problem is formulated too general and to have reasonable results we are to consider concrete variants of it. Below we are dealing with the following problems.

Problem I An observed signal X_ε on the interval $[a, b]$ is of the form

$$dX_\varepsilon(t) = f(t)dt + \varepsilon dw(t).$$

Here $w(t)$ is the Wiener process, ε is a known small parameter and the unknown function f belongs to a known class $\mathbf{F}$ of functions analytic in a vicinity of $[a, b]$.

Problem II Assume that one observes iid random variables $X_1, X_2, \dots X_n$ taking values in R^1 and having a density function $f \in \mathbf{F}$ where again $\mathbf{F}$ consists of functions analytic in a vicinity of an interval $D = [a, b]$. The problem is to estimate f on the interval $[a, b]$.

Problem III Let $\{X_t\}$ be a real-valued stationary Gaussian process with discrete or continuous time t, mean zero and spectral density $f(\lambda)$. We assume that the spectral density f is unknown and should be estimated on the base of observations $X_t, 0 \leq t \leq T$. Assume further that $f \in \mathbf{F}$ where $\mathbf{F}$ is a given (known) class of spectral densities. Let $[a, b]$ be an interval and assume that we would like to estimate the restriction of f to $[a, b], \{f(\lambda), a \leq \lambda \leq b\}$, when all $f \in \mathbf{F}$ are analytic in a vicinity of $[a, b]$.

Problem IV Let f be an unknown function belonging to a known class $\mathbf{F}$ of functions analytic in a vicinity of an interval $[a, b]$. To estimate f one makes n observations of the function f at the points $X_1, X_2, \dots X_n$ and observes

$$Y_j = f(X_j) + G_j(X_j, \omega),$$

where $\mathbf{E}(G_j(X_j, \omega)|X_1, \dots X_{j-1}) = 0$ and the noise variables G_j are conditionally independent under a given observation plan (see in more detail [14]).

One may ask why such attention to the estimation of analytic functions, isn't the class of such functions too special and too narrow? I believe that apart of their mathematical beauty these problems are rather natural. I appeal to two arguments of authority.

"When I asked Hilbert what were the general considerations which brought him to the hypothesis (proved later by myself) that all solutions to regular problems of the calculus of variations are analytic, the celebrated geometer answered that by his opinion solutions to all naturally posed problems should be analytic." (S. Bernstein, Sur la deformation de surfaces, Ann.Math., 1905).

"It may seem that analytic functions are a refined, a recherché object and that such a refined property as analyticity is not necessary in numerical analysis. In fact it is not true. Analytical functions constitute a very disseminated object and many natural scientific problems lead to analytic functions. Ignoring the analytic nature of solutions we narrow drastically our possibilities and operate squanderingly." (K. Babenko, Foundations on numerical analysis).

Notice also that in the case of stationary processes a fast decay of the correlation function implies the analyticity of the spectral density. If the correlation radius is finite, the spectral density is even an integer function.

Below we denote through c, C with or without indices various constants. In different places the same notation may stay for different constants.

2 Functions analytic on bounded regions

In this section we suppose that the class $\mathbf{F}$ consists of functions analytic on the *bounded* region $G \supset [a, b]$ of the complex plane and bounded there by a constant M. The region G and the constant M are supposed to be known. Thus

$$\Delta_p(\mathbf{F}) = \inf_{\hat{f}} \sup_{f \in \mathbf{F}} \Big(\int_a^b |\hat{f}(t) - f(t)|^p dt \Big)^{1/p}.$$

Theorem 2.1 *Under the conditions of Problem I when $\varepsilon \to 0$ the minimax risk $\Delta_p(\mathbf{F}, \varepsilon)$ satisfies the following asymptotic relations*

$$\begin{aligned}
&\Delta_p(\mathbf{F})) \asymp \varepsilon\sqrt{\ln(1/\varepsilon)}, \qquad 1 \le p < 4,\\
&\Delta_4(\mathbf{F}) \asymp \varepsilon\sqrt{\ln(1/\varepsilon)}(\ln\ln(1/\varepsilon))^{1/4},\\
&\Delta_p(\mathbf{F}) \asymp \varepsilon(\ln(1/\varepsilon))^{1-2/p}, \qquad 4 < p \le \infty.
\end{aligned} \tag{2.1}$$

The constants under the $\asymp$ sign depend on G, M, and p only.

Recall that the sign "$a_n \asymp b_n$" means that $0 < c = \liminf(a_n/b_n) \le \limsup(a_n/b_n) = C < \infty$.

Theorem 2.2 *Under the conditions of Problem II when $n \to \infty$ the* *minimax risk* $\Delta_p(\mathbf{F}, n)$ *satisfies the following asymptotic relations*

$$\begin{aligned} \Delta_p(\mathbf{F}) &\asymp \sqrt{\frac{\ln n}{n}}, \qquad 1 \le p < 4, \\ \Delta_4(\mathbf{F}) &\asymp \sqrt{\frac{\ln n}{n}}(\ln\ln n)^{1/4}, \\ \Delta_p(\mathbf{F}) &\asymp \sqrt{\frac{1}{n}}(\ln n)^{1-2/p}, \qquad 4 < p \le \infty. \end{aligned} \tag{2.2}$$

The constants depend on G, M, *and* p *only.*

Theorem 2.3 *Let* X_t *be a stationary gaussian process with discrete time. Under the conditions of Problem III when* $T \to \infty$ *the minimax risk* $\Delta_p(\mathbf{F}, T)$ *satisfies the following asymptotic relations*

$$\begin{aligned} \Delta_p(T) &\asymp \sqrt{\frac{\ln T}{T}}, \qquad 1 \le p < 4, \\ \Delta_4(T) &\asymp \sqrt{\frac{\ln T}{T}}(\ln\ln T)^{1/4}, \\ \Delta_p(T) &\asymp \sqrt{\frac{1}{T}}(\ln T)^{1-2/p}, \qquad 4 < p \le \infty. \end{aligned} \tag{2.3}$$

The constants depend on G, M, *and* p *only.*

In the Theorems 2.4 and 2.5 below we suppose that X_t is a stationary Gaussian process with continuous time (may be a generalized one).

Theorem 2.4 *Let the set* $\mathbf{F}$ *consist of all spectral densities of generalized processes analytic in some bounded region* $G, [a, b] \subset G$ *and bounded there by a common constant* M. *Then there exist estimators* $\hat{f}_T$ *such that*

$$\begin{aligned} \sup_{f\in F} \mathbf{E}_f|\hat{f}_T - f|_p &\le C\sqrt{\frac{\ln T}{T}}, \qquad 1 \le p < 4, \\ \sup_{f\in F} \mathbf{E}_f|\hat{f}_T - f|_4 &\le C\sqrt{\frac{\ln T}{T}}(\ln\ln T)^{1/4}, \\ \sup_{f\in F} \mathbf{E}_f|\hat{f}_T - f|_p &\le CT^{-1/2}(\ln T)^{1-2/p}, \qquad 4 < p \le \infty. \end{aligned} \tag{2.4}$$

The constants C *depend on* G, M, *and* p *only.*

Theorem 2.5 *Let the set* $\mathbf{F}$ *consist of all spectral densities of real-valued (not generalized) processes analytic in some bounded region* $G, [a, b] \subset G$

and bounded there by a common constant M. Then there exist positive constants $c > 0$ such that

(2.5)
$$\Delta_p(T) \geq c\sqrt{\frac{\ln T}{T}}, \qquad 1 \leq p < 4$$
$$\Delta_4(T) \geq c\sqrt{\frac{\ln T}{T}}(\ln\ln T)^{1/4},$$
$$\Delta_p(T) \geq cT^{-1/2}(\ln T)^{1-2/p}, \qquad 4 < p \leq \infty.$$

The constants c depend on G, M, and p only.

Theorem 2.6 *Let under the conditions of Problem IV the variables G_j have moments of all orders. Then when $n \to \infty$ the minimax risk $\Delta_p(\mathbf{F}, n)$ satisfies the following asymptotic relations*

(2.6)
$$\Delta_p(\mathbf{F}) \asymp \sqrt{\frac{\ln n}{n}}, \qquad 1 \leq p < \infty,$$
$$\Delta_\infty(\mathbf{F}) \asymp \sqrt{\frac{\ln n}{n}}(\ln\ln n)^{1/2}.$$

The constants depend on G, M, and p only.

Theorems 2.1, 2.2 and 2.6 and their multidimensional analogues are proved in [1]. Theorem 2.3 is proved in [2]. The proofs of Theorems 2.4 and 2.5, which also have been mentioned in [2], will be published elsewhere.

3 Functions analytic in strips

In this section we suppose that the interval $[a, b] = (-\infty, \infty)$ and the set $\mathbf{F}$ consists of functions $f(z), z = t + is$, analytic inside the strip $|s| < \Lambda$ and bounded in the closure of this strip by a common constant $M, |f(z)| \leq M$. The strip and the constant are supposed to be known.

Theorem 3.1 *Under the conditions of Problem II when $n \to \infty$ the minimax risk $\Delta_p(\mathbf{F}, n)$ satisfies the following asymptotic relations*

(3.1)
$$\Delta_p(\mathbf{F}) \asymp \sqrt{\frac{\ln n}{n}}, \; 2 \leq p < \infty,$$
$$\Delta_\infty(\mathbf{F}) \asymp \sqrt{\frac{\ln n}{n}}\sqrt{\ln\ln n}.$$

The constants depend on Λ, M, and p only.

Note that the restriction $p \geq 2$ here is essential. For $p < 1$ the behaviour of $\Delta_p(\mathbf{F})$ depends on p. For $p = 1$ we can not even have consistent estimators for f (see [3]).

Theorem 3.2 *Let under the conditions of Problem III X_t be a real valued Gaussian process with continuous time. Then when $T \to \infty$ the minimax risk $\Delta_p(\mathbf{F}, T)$ satisfies the following asymptotic relations*

$$\begin{aligned} \Delta_p(\mathbf{F}) &\asymp \sqrt{\frac{\ln T}{T}}, \ 1 \leq p < \infty, \\ \Delta_\infty(\mathbf{F}) &\asymp \sqrt{\frac{\ln T}{T}} \sqrt{\ln \ln T}. \end{aligned} \tag{3.2}$$

The constants depend on Λ, M, and p only.

Theorem 3.3 *Let under the conditions of Problem III X_t be a real valued Gaussian process with discrete time. If $\mathbf{F}$ consists of all 2π-periodic functions analytic and uniformly bounded inside a strip $|s| < \Lambda$, then when $T \to \infty$ the minimax risk $\Delta_p(\mathbf{F}, T)$ satisfies the following asymptotic relations*

$$\begin{aligned} \Delta_p(\mathbf{F}) &\asymp \sqrt{\frac{\ln T}{T}}, \ 1 \leq p < \infty, \\ \Delta_\infty(\mathbf{F}) &\asymp \sqrt{\frac{\ln T}{T}} \sqrt{\ln \ln T}. \end{aligned} \tag{3.3}$$

The constants depend on Λ, M, and p only

Evidently in the case of Problems I or IV we can not expect reasonable results in the $L_p(R^1)$ norms. So we suppose that the risk functions are defined with respect to a bounded interval $[a, b]$ and that the estimated functions f are $(b - a)$- periodic functions analytic and bounded inside the strip $|s| < \Lambda$. Under such conditions relations (3.1) are true for both these problems (in the case of Problem I we have to substitute ε instead of $1/\sqrt{n}$).

All the above mentioned theorems of this Section have been found by Ibragimov and Khasminskii beginning from the middle of the 70's (see [4], [12], [14]). In 1980 M. Pinsker [5] worked out a new method which in the case of L_2 norms allowed to compute the precise asymptotics of the type $\sim$ not $\asymp$(the relation $a_n \sim b_n, n \to \infty$, means that $\lim \frac{a_n}{b_n} = 1$). His method heavily used the fact that in the case of L_2 spaces the sets $\mathbf{F}$ are ellipsoids in Hilbert space L_2. Later Korostelev, Donoho, Golubev, Levit, Tsybakov have spread Pinsker's results on non Hilbert norms case. I quote below some last results of F. Guerre and A. Tsybakov [6].

Denote by $\mathbf{F} = \mathbf{F}(L, \gamma)$ the class of all real valued functions $f \in L_2$ such that

$$\int_{-\infty}^{\infty} \cosh^2(\gamma t)|g(t)|^2 dt \le 2\pi L$$

where g is the Fourier transform of f and L, γ are positive constants. Evidently the functions $f \in \mathbf{F}$ are analytic inside the strip $|s| < \gamma$.

Consider now the Problem I when the signal $f(t)$ is observed on the *whole* real line. Denote by $|\cdot|_p$ the norm in the space $L_p[0,1]$.

Theorem 3.4 **(F. Guerre, A. Tsybakov)** *Let* $1 \le p \le \infty$. *Then*

$$\lim_{\varepsilon \to 0} \inf_{\hat{f}} \sup_{f \in \mathbf{F}} \varepsilon^{-1} (\ln(1/\varepsilon))^{-1/2} \mathbf{E}_f |\hat{f} - f|_p = M_p$$

where

$$M_p = \sqrt{\frac{2}{\gamma\pi}} \left(\frac{1}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right)\right)^{1/p}.$$

The case $p = \infty$ has been treated earlier by Golubev, Levit, Tsybakov [7].

Theorem 3.5 **(Golubev, Levit, Tsybakov)** *When* $\varepsilon \to 0$

$$\lim \inf \sup \varepsilon^2 \ln(1/\varepsilon) \ln\ln(1/\varepsilon)^{-1/2} \mathbf{E}_f |\hat{f} - f|_\infty = \left(\frac{2}{\gamma\pi}\right)^{1/2}.$$

Notice that one may expect the same results for the other problems mentioned in Section 1. Notice also that in the frames of Theorems 3.4 and 3.5 it would be natural to make observations on the interval $[0, 1]$ only. The final formulas would change of course but how?

4 Integer functions

In this section we will suppose that the class $\mathbf{F}$ consists of functions analytic on the whole complex plane. One may expect that now the rate of $\Delta_p(\mathbf{F})$ will depend on the growth of functions $f \in \mathbf{F}$ at ∞. We consider the following well known standard classes of integer functions (see [20]).

1. The class $\mathbf{F}_1(M, c, \rho)$ consists of all integer functions $f(z), z = x + iy$, such that

$$\sup_x |f(x + iy)| \le M \exp\{c|y|^\rho\}.$$

2. The class $\mathbf{F}_2(M, c, \rho)$ consists of all integer functions $f(z)$ such that

$$\sup_{|z| \le r} |f(z)| \le M \exp\{cr^\rho\}.$$

Theorem 4.1 *Let under the conditions of Problem II the unknown density function belong to a class $\mathbf{F}_1$ with $\rho > 1$. Let the minimax risks Δ_p be defined through $L_p(R^1)$ norms. Then when $n \to \infty$*

$$\begin{aligned} \Delta_p(\mathbf{F}) &\asymp \frac{1}{\sqrt{n}}(\ln n)^{(\rho-1)/(2\rho)},\ 2 \le p < \infty, \\ \Delta_\infty(\mathbf{F}) &\asymp \frac{1}{\sqrt{n}}(\ln n)^{(\rho-1)(2\rho)}\sqrt{\ln\ln n}. \end{aligned} \tag{4.1}$$

The constants depend on M, c, ρ, and p only.

Proof of Theorem 4.1 **Upper bounds.** Consider estimators $f_T(x)$ defined by the formula

$$f_T(x) = \frac{1}{\pi n}\sum_1^n \frac{\sin T(X_j - x)}{X_j - x} \tag{4.2}$$

and study the behaviour of its bias and stochastic term.

1. Bias. Let

$$\phi(t) = \mathbf{E}_f e^{itX_1} = \int_{-\infty}^{\infty} e^{itx} f(x)dx$$

be the characteristic function of X_j under the density function f. Then the bias of the estimator $f_T(x)$ at the point x is

$$\begin{aligned} f(x) - \mathbf{E}_f f_T(x) &= f(x) - \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin T(y-x)}{y-x} f(y)dy \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)dt - \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)dy\int_{-T}^{T} e^{it(y-x)}dt \\ &= \frac{1}{2\pi}\int_{|t|>T} e^{-itx}\phi(t)dt. \end{aligned} \tag{4.3}$$

A function $g \in L_p, p \ge 2$, and its Fourier transform $\tilde{g}$ satisfy the following inequalities (see [8], ch. IV):

$$\|g\|_p \le C\|\tilde{g}\|_q, \qquad 1/p + 1/q = 1. \tag{4.4}$$

Hence for $2 \le p \le \infty$ the L_p-norm of the bias satisfies

$$\|f - \mathbf{E}_f f_T\|_p \le C\int_{|t|>T} |\phi(t)|^q dt \le C\int_{|t|>T} |\phi(t)|dt. \tag{4.5}$$

To estimate the last integral consider an analytic function $\psi \in L_2$ whose Fourier transform $\tilde{\psi}$ is zero outside an interval $[-A, A]$. Then for any y

$$\psi(x+iy) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\lambda x}e^{\lambda y}\tilde{\psi}(\lambda)d\lambda$$

and Parseval's identity implies that

$$\int_{-\infty}^{\infty} f(x)\psi(x+iy)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(\lambda)e^{-\lambda y}\tilde{\psi}(-\lambda)d\lambda.$$

Changing on the left the contour of integration we find that for any y

$$\int_{-\infty}^{\infty} f(x+iy)\psi(x)dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(\lambda)e^{\lambda y}\tilde{\psi}(-\lambda)d\lambda. \tag{4.6}$$

Take at the last relation

$$\tilde{\psi}(\lambda) = \tilde{\psi}_A(\lambda) = \begin{cases} \phi(\lambda)(1-|\lambda|/A), & \text{if } |\lambda| \le A, \\ 0, & \text{if } |\lambda| > A. \end{cases}$$

where A is a positive number. Then

$$\begin{aligned}\psi_A(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\lambda x}\tilde{\psi}_A(\lambda) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi(\lambda)(1-|\lambda|/A)e^{-i\lambda x}d\lambda = \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)\frac{\sin^2\frac{x-y}{2}A}{A(\frac{x-y}{2})^2}dy \ge 0\end{aligned} \tag{4.7}$$

and

$$\int_{-\infty}^{\infty}\psi_A(x)dx = 1. \tag{4.8}$$

It follows from (4.6) and (4.7), (4.8) that

$$\begin{aligned}\frac{1}{2\pi}\int_{-A}^{A}|\phi(\lambda)|^2 e^{\lambda y}(1-|\lambda|/A)d\lambda &= \int_{-\infty}^{\infty} f(x+iy)\psi_A(x)dx \\ &\le \sup_x |f(x+iy)| \le M\exp\{c|y|^\rho\}.\end{aligned}$$

The last inequality implies that for any y

$$\int_{-\infty}^{\infty}|\phi(\lambda)|^2\exp\{|\lambda||y|\}d\lambda \le 2\pi M e^{c|y|^\rho}. \tag{4.9}$$

But then for $|y| \ge 1$

$$\begin{aligned}\int_{-\infty}^{\infty} e^{|\lambda||y|}|\phi(\lambda)|d\lambda &\le \Big(\int_{-\infty}^{\infty} e^{-2|\lambda||y|}d\lambda\Big)^{1/2}\Big(\int_{-\infty}^{\infty} e^{4|\lambda||y|}|\phi(\lambda)|^2 d\lambda\Big)^{1/2} \\ &\le (2\pi M)^{1/2}e^{c4^\rho|y|^\rho} = M_1 e^{c_1|y|^\rho}.\end{aligned}$$

Hence for any $y > 0$

$$\int_T^{\infty}|\phi(t)|dt \le e^{-Ty}\int_{-\infty}^{\infty} e^{|t|y}|\phi(t)|dt \le M_1\exp\{-Ty + c_1|y|^\rho\}.$$

Take here $y = T^{\frac{1}{\rho-1}}$. We find that

$$\int_T^\infty |\phi(t)|dt \leq M_1 \exp\{-c_2 T^{\rho/(\rho-1)}\}. \tag{4.10}$$

The last inequality together with (4.5) gives us the following bound for the bias. For any $p, 2 \leq p \leq \infty$,

$$\|f_T - \mathbf{E}_f f_T\|_p \leq B \exp\{-bT^{\rho/(\rho-1)}\}, \tag{4.11}$$

where the constants $B, b > 0$ depend on M, c, ρ, and p only.

2. The stochastic term is the $L_p(R^1)$-norm of the random function

$$n^{-1} \sum_{j=1}^{n} \left[\frac{\sin T(X_j - x)}{X_j - x} - \mathbf{E}_f \frac{\sin T(X_j - x)}{X_j - x} \right].$$

To estimate it we use the following inequality of Rosenthal ([9], see also [10]):

Lemma 4.2 *Let $\xi_1, \ldots, \xi_n$ be independent random variables with $\mathbf{E}\xi_j = 0$. Then for $2 \leq p < \infty$*

$$\mathbf{E}|\sum_1^n \xi_j|^p \leq c_p\Big(\sum_1^n \mathbf{E}|\xi_j|^p + \Big(\sum_1^n \mathbf{E}\xi_j^2\Big)^{p/2}\Big). \tag{4.12}$$

It follows from this inequality that for $2 \leq p < \infty$

$$\mathbf{E}\|f_T - \mathbf{E}f_T\|_p^p \leq \frac{2^p c_p}{\pi^p n^{p-1}} \int_{-\infty}^{\infty} \mathbf{E} \left| \frac{\sin T(X_1 - x)}{X_1 - x} \right|^p dx + \frac{c_p}{\pi^p n^{p/2}} \int_{-\infty}^{\infty} \left(\mathbf{E} \frac{\sin^2 T(X_1 - x)}{(X_1 - x)^2} \right)^{p/2} dx.$$

Now for $p \geq 2$

$$\begin{aligned} &\int_{-\infty}^{\infty} \left(\mathbf{E} \frac{\sin^2 T(X_1 - x)}{(X_1 - x)^2} \right)^{p/2} dx \\ &\leq M^{p/2-1} \left(\int_{-\infty}^{\infty} \frac{\sin^2 Tx}{x^2} dx \right)^{p/2-1} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} \frac{\sin^2 T(x-y)}{(x-y)^2} f(y)dy \\ &\leq M^{p/2-1} \pi^{p/2} T^{p/2}. \end{aligned}$$

It follows from these two inequalities that

$$\mathbf{E}\|f_T - \mathbf{E}f_T\|_p^p \leq c_p^{(1)} \left\{ \frac{T^{p-1}}{n^{p-1}} \int_{-\infty}^{\infty} \left| \frac{\sin x}{x} \right|^p dx + M^{p/2} \left(\frac{T}{n} \right)^{p/2} \right\}. \tag{4.13}$$

We always suppose that $T = o(n)$. The last inequality implies then that for $2 \le p < \infty$

$$\mathbf{E}\|f_T - \mathbf{E}f_T\|_p \le C_p\sqrt{\frac{T}{n}}. \tag{4.14}$$

The case $p = \infty$ needs a special consideration. At first we need to know the behaviour of C_p in (4.14) for large p. Letting

$$\xi_j(x) = \xi_j = \frac{\sin T(X_j - x)}{X_j - x} - \mathbf{E}\left(\frac{\sin T(X_j - x)}{X_j - x}\right)$$

we find for even $p = 2k$ and $p = o(n)$ that

$$\begin{aligned}\mathbf{E}|\sum_1^n \xi_j|^{2k} &= \sum_{\substack{l_1 + \ldots + l_n = 2k \\ l_j \ne 1}} \mathbf{E}\xi_1^{l_1} \ldots \mathbf{E}\xi_n^{l_n} \\ &= n\mathbf{E}\xi_1^{2k} + \binom{n}{2}\frac{2k!}{(2k-2)!2!}\mathbf{E}\xi_1^{2k-2}\mathbf{E}\xi_1^2 + \ldots + \binom{n}{k}\frac{2k!}{(2!)^k}\left(\mathbf{E}\xi_1^2\right)^k.\end{aligned}$$

It follows from this inequality that if $p < n^{1-\gamma}, \gamma > 0$, which will be the case, that

$$\int_{-\infty}^{\infty} |\sum_1^n \xi_j(x)|^{2k} dx \le (Ck)^k \left(\frac{T}{n}\right)^k \tag{4.15}$$

where C is a constant. It means that (4.14) can be rewritten as

$$\mathbf{E}_f\|f_T - \mathbf{E}_f f_T\|_p \le c\sqrt{p}\sqrt{T/n} \tag{4.16}$$

where the constant c does not depend on p. We apply now the following Nikolskii inequality(see [11]):

Lemma 4.3 *Let $g(x)$ be an integer function such that*

$$|g(z)| \le C\exp\{T|z|\}. \tag{4.17}$$

Then for $1 \le p < q \le \infty$

$$\|g\|_q \le cT^{1/p-1/q}\|g\|_p \tag{4.18}$$

where c is an absolute constant.

The function $g_T(x) = f_T(x) - \mathbf{E}f_T(x)$ is a random integer function which satisfies (4.17). The inequalities (4.16), (4.18) imply then that

$$\mathbf{E}\|f_T - \mathbf{E}f_T\|_\infty \le cT^{1/p}\mathbf{E}\|f_T - \mathbf{E}f_T\|_p \le C\sqrt{p}T^{1/p}\sqrt{T/n}.$$

Take here $p = \ln T$. We find that

$$\mathbf{E}\|f_T - \mathbf{E}f_T\|_\infty \le C\sqrt{\ln T}\sqrt{T/n}. \tag{4.19}$$

Now we are ready to establish the upper bounds of the theorem. If $p < \infty$, we apply (4.11) and (4.14) (or (4.16)) and get that

$$\begin{aligned}\mathbf{E}_f\|f_T - f\|_p \le \|\mathbf{E}_f f_T - f\|_p + \mathbf{E}_f\|f_T - \mathbf{E}_f f_T\|_p \le \\ \le C(\exp\{-T^{\rho/(\rho-1)}\} + \sqrt{T/n}).\end{aligned}$$

Take here $T \sim (\ln n)^{(\rho-1)/\rho}$. We find that for such f_T

$$\mathbf{E}_f\|f_T - f\|_p \le C\frac{1}{\sqrt{n}}(\ln n)^{(\rho-1)/2\rho}. \tag{4.20}$$

In the same way we find applying (4.11) and (4.19) that for the same T

$$\mathbf{E}\|f_T - f\|_\infty \le c\frac{1}{\sqrt{n}}(\ln n)^{(\rho-1)/2\rho}\sqrt{\ln\ln n}. \tag{4.21}$$

Lower bounds. To prove the lower bounds of the theorem we use methods from Hasminskii and Ibragimov [12]. The following lemma is proved in [12].

Lemma 4.4 *Assume that there are $N(\delta)$ densities $f_{i\delta} \in \mathbf{F}, i = 1, \cdots, N$, such that $\|f_{i\delta} - f_{j\delta}\|_p \ge \delta$ for $i \ne j$. Let $\{f_{0\delta}\}$ be a family of densities . Let*

$$\delta(n, \mathbf{F}) = \sup\Big\{\delta : \frac{1}{\ln N(\delta)} \max_{1\le i,j\le N}\Big\|\frac{f_{i\delta} - f_{j\delta}}{\sqrt{f_{0\delta}}}\Big\|_2^2 \le 1/2n\Big\}. \tag{4.22}$$

Then for any estimator $\hat{f}_n$ of f and all $p \ge 1$

$$\sup_{f\in\mathbf{F}} \mathbf{E}_f\|\hat{f}_n - f\|_p \ge \frac{\delta(n,\mathbf{F})}{4}. \tag{4.23}$$

For the sake of simplicity we consider the technically simpler case $p = \infty$. We begin with constructing a family $\{f_{i\delta}\}$ which will satisfy the conditions of Lemma 4.4. Set

$$f_0(x) = a\frac{\sin(\gamma x + ib)}{\gamma x + ib}\frac{\sin(\gamma x - ib)}{\gamma x - ib}$$

where the real positive constants a, γ, b are chosen to make f_0 a density function from the class $\mathbf{F}_1(1/2M, c)$. Define now

$$f_{k\delta}(x) = f_k(x) = f_0(x) + \varepsilon\left(\frac{\sin N(x - 2\pi(k-1)/N)}{x - 2\pi(k-1)/N}\right)^3, \qquad k = 1, \ldots, N, \tag{4.24}$$

where ε will be fixed later. Evidently $\int_{-\infty}^{\infty} f_k(x)dx = 1$ and if ε is sufficiently small, $\varepsilon < cN^{-3}$, c is a small constant, $f_k(x) > 0$. Hence all f_k are probability densities for ε sufficiently small. The differences

$$\|f_k - f_j\|_\infty \geq \varepsilon N^3.$$

Evidently all f_k are integer functions. Notice that

$$\begin{aligned}\sup_x |f_k(x+iy)| &\leq \frac{1}{2} M \exp\{c|y|^\rho\} + \varepsilon \sup_x \left|\frac{\sin N(x+iy)}{x+iy}\right|^3 \\ &\leq \frac{1}{2} M \exp\{c|y|^\rho\} + \varepsilon N e^{N|y|}.\end{aligned}$$

To ensure that $f_k \in \mathbf{F}$ we have to choose ε in a such way that for all y

$$\varepsilon N e^{N|y|} \leq \frac{1}{2} M e^{c|y|^\rho}.$$

The expression

$$c|y|^\rho + \ln M/2 - N|y| - \ln N - \ln \varepsilon$$

takes its maximal value at the points $|y| = (N/c\rho)^{1/(\rho-1)}$ and it is enough to take as ε any number

$$\lambda \leq \frac{1}{2} MN \exp\{-c^{1/(\rho-1)}(\rho^{-(1/(\rho-1))} - \rho^{\rho/(\rho-1)}) N^{\rho/(\rho-1)}\}.$$

We take

$$\varepsilon = \exp\{-\alpha N^{\rho/(\rho-1)}\}$$

where α is a sufficiently large number.

Let us find now $\delta(n, \mathbf{F})$ from (4.22). We have

$$\begin{aligned}\left\|\frac{f_k - f_j}{\sqrt{f_0}}\right\|_2^2 &\leq 4\varepsilon^2 a \sup_k \int_{-\infty}^{\infty} \left|\frac{\sin Nx}{x - 2\pi(k-1)/N}\right|^6 \cdot \frac{\gamma^2 x^2 + b^2}{|\sin(\gamma x + ib)|^2} dx \\ &\leq c\varepsilon^2 N^5 \end{aligned} \tag{4.25}$$

where c is a positive constant. Hence we may take any N which satisfies the inequality

$$c(\ln N)^{-1} \varepsilon^2 N^5 = c(\ln n)^{-1} \exp\{-2\alpha N^{\rho/(\rho-1)}\} N^5 \geq 1/2n. \tag{4.26}$$

In particular we may take

$$N \sim c(\ln n)^{(\rho-1)/\rho} \tag{4.27}$$

where c is a sufficiently small positive number. For such a choice of N the inequalities (4.25), (4.26) imply that

$$\delta(N, F) \geq \varepsilon N^3 \geq \frac{c_0}{\sqrt{n}} \sqrt{N \ln N} \geq \frac{c_1}{\sqrt{n}} (\ln n)^{(\rho-1)/2\rho} \sqrt{\ln \ln n} \tag{4.28}$$

which gives the needed lower bound for the case $p = \infty$.

If $p < \infty$, then (following again [12]) we construct the family $f_{j\delta}$ as

$$f_{\bar{a}}(x) = f_0(x) + \varepsilon \sum_1^N a_k \left(\frac{\sin N(x - 2\pi(k-1)/N)}{x - 2\pi(k-1)/N} \right)^3$$

where the vectors $\bar{a} = (a_1, \dots a_N)$ run a set of vectors $\bar{a}$ with $a_j = \pm 1$. Combining arguments from [12] with those we have just used we get the lower bound for $p < \infty$. The theorem is proved.

Remark. The restriction $p \geq 2$ is essential. For $p = 1$ we cannot even expect the existence of consistent estimators, see [12], [13].

The last result can be strengthened if $\rho = 1$. Namely in 1982 Ibragimov and Khasminskii proved the following result [13].

Theorem 4.5 *Let under the conditions of Problem II the observations X_j take their values in R^d and have density function f belonging to a class* **F***of functions whose Fourier transform is zero outside a compact convex symmetric set K. Then when $n \to \infty$*

$$\Delta_2^2(\mathbf{F}) \sim \frac{1}{n} \frac{m(K)}{(2\pi)^{2d}} \tag{4.29}$$

where $m(K)$ denotes the Lebesgue measure of K.

Analogous results can be proved also for stationary processes.

Theorem 4.6 *Let under the conditions of Problem III X_t be a real valued Gaussian process with continuous time. Let Δ_p be defined through $L_p(R^1)$-norms. Then when $T \to \infty$ the minimax risk $\Delta_p(\mathbf{F}_1(M, c, \rho)), \rho > 1$, satisfies the following asymptotic relations*

$$\begin{aligned} \Delta_p(\mathbf{F}_1) &\asymp \frac{1}{\sqrt{T}} (\ln T)^{(\rho-1)/2\rho}, \qquad 2 \leq p < \infty, \\ \Delta_\infty(\mathbf{F}_1) &\asymp \frac{1}{\sqrt{T}} (\ln T)^{(\rho-1)/2\rho} \sqrt{\ln \ln T}. \end{aligned} \tag{4.30}$$

The constants depend on M, c, ρ, and p only.

The theorem will be proved elsewhere. We show below only that in the case $\rho = 1$ the rate of convergence of Δ_p is $T^{-1/2}$.

Theorem 4.7 *Let under the conditions of Problem III $X(t)$ be a stationary Gaussian process with continuous time and spectral density function f belonging to a class $\mathbf{F} = \mathbf{F}_1(M, a, 1) \cap \{f : \|f\| \leq \sigma^2\}$. Then when $T \to \infty$*

$$\limsup T \Delta_2^2(\mathbf{F}) \leq 4a\sigma^2.$$

Proof By the Paley-Wiener theorem the correlation function $R(t)$ of the process is zero outside the interval $[-a,a]$. Thus the spectral density

$$f(\lambda) = \frac{1}{2\pi}\int_{-a}^{a} R(t)dt.$$

We estimate R(t) by

$$\hat{R}(t) = \begin{cases} \frac{1}{T-|t|}\int_0^{T-|t|} X(t+u)X(u)\,du, & \text{if } |t| \le a, \\ 0, & \text{if } |t| > a. \end{cases}$$

and $f(\lambda)$ by

$$\hat{f}(\lambda) = \frac{1}{2\pi}\int \hat{R}(t)dt.$$

Not difficult computations show then that

$$\begin{aligned} \mathbf{E}\|f-\hat{f}\|_2^2 &= \mathbf{E}\{\frac{1}{2\pi}\|R-\hat{R}\|_2^2\} \\ &= \frac{1}{2\pi}\int_{-a}^{a}(T-|t|)^{-2}dt\int_0^{T-|t|}\int_0^{T-|t|}(R^2(u-v) \\ &\qquad + R(t+u-v)R(t-u+v))\,du\,dv \\ &= \frac{1}{2\pi T}(1+o(1))\left(2a\|R\|_2^2 + \left(\int_{-a}^{a} R(u)du\right)^2\right). \end{aligned}$$

Further $\frac{1}{2\pi}\|R\|_2^2 = \|f\|_2^2 \le \sigma^2$ and

$$\left(\int_a^a R(u)du\right)^2 \le 2a\|R\|_2^2 \le 4\pi a\sigma^2.$$

The theorem is proved. ∎

Consider now problems when $f \in \mathbf{F}_2(M,c,\rho)$. (Recall that the last condition means that $\sup_{|z|\le r}|f(z)| \le M\exp\{c|y|^{\rho}\}$, see p.6).

Theorem 4.8 *Let under the conditions of Problem I the observations be $X_\varepsilon(t), -\infty < a < t < b < \infty$. If the unknown signal $f \in \mathbf{F}_2(M,c,\rho)$, then when $\varepsilon \to 0$ the minimax risk $\Delta_p(\mathbf{F},\varepsilon)$ defined through $L_p(a,b)$-norm satisfies the following asymptotic relations*

$$\begin{aligned} &\Delta_p(\mathbf{F})) \asymp \varepsilon\sqrt{\frac{\ln(1/\varepsilon)}{\ln\ln(1/\varepsilon)}}, \qquad 1 \le p < 4, \\ (4.31)\qquad &\Delta_4(\mathbf{F}) \asymp \varepsilon\sqrt{\frac{\ln(1/\varepsilon)}{\ln\ln(1/\varepsilon)}}(\ln\ln(1/\varepsilon))^{1/4}, \\ &\Delta_p(\mathbf{F}) \asymp \varepsilon\left(\frac{\ln(1/\varepsilon)}{\ln\ln(1/\varepsilon)}\right)^{1-2/p}, \qquad 4 < p \le \infty. \end{aligned}$$

The constants depend on M, c, ρ, and p only.

The same results take place for all other problems if the risk is measured in $L_p(a,b)$-norms, $-\infty < a < b < \infty$ except for Problem IV.

Theorem 4.9 *Let under the conditions of Problem IV the unknown function $f \in \mathbf{F}_2(M,c,\rho)$ and variables G_j have moments of all orders. Then when $n \to \infty$ the minimax risk $\Delta_p(\mathbf{F}, n)$ satisfies the following asymptotic relations*

$$\Delta_p(\mathbf{F}) \asymp \sqrt{\frac{\ln n}{n \ln\ln n}}, \qquad p < \infty, \tag{4.32}$$
$$\Delta_\infty(\mathbf{F}) \asymp \sqrt{\frac{\ln n}{n \ln\ln n}} (\ln\ln n)^{1/2}.$$

The constants depend on M, c, ρ, and p only.

We give the proof of Theorem 4.8 only. The proof of upper bounds in (4.32) is based on the inequality (4.33) below and arguments from [1]; the proof of lower bounds is based on arguments from [14], [15]. A detailed version of the proof will be published elsewhere.

Proof of Theorem 4.8 The proof repeats the main arguments of the proof of Theorem 1 from [1] and we omit the details. Evidently we may and will suppose that $[a,b] = [-1,1]$.

Upper bounds. Consider the Fourier expansion of f with respect to the Legendre polynomials $P_k(x)$:

$$f(x) = \sum_0^\infty a_k P_k(x)$$

and estimate the coefficients $a_k = \int_{-1}^1 P_k(x) f(x) dx$ by the statistics

$$\hat{a}_k = \int_{-1}^1 P_k(x) dX_\varepsilon(x).$$

Introduce now the statistics

$$f_N(x) = \sum_0^N \hat{a}_k P_k(x)$$

and study separately their bias

$$b_N(x) = f(x) - \mathbf{E} f_N(x) = \sum_{N+1}^\infty a_k P_k(x)$$

and the random term

$$z_N(x) = f_N(x) - \mathbf{E} f_N(x).$$

1. The bias. Introduce into the consideration the Chebyshev polynomials $T_k(x) = \frac{2}{\pi}\cos(k \arccos x)$. They are orthonormal on $[-1,1]$ with respect to the weight $(1-x^2)^{-1/2}$. Let

$$f(x) = \sum_0^\infty b_k T_k(x).$$

The value of the best approximation of the function f in the L_2-norm by polynomials Q of degree N is equal to

$$\begin{aligned}\sqrt{\sum_{N+1} a_k^2} &= \left(\inf_Q \int_{-1}^1 |f(x) - Q(x)|^2 dx\right)^{1/2} \\ &\le \left(\inf_Q \int_{-1}^1 |f(x) - Q(x)|^2 (1-x^2)^{-1/2} dx\right)^{1/2} = \sqrt{\sum_{N+1} b_k^2}.\end{aligned}$$

The coefficients b_k have the following representation (see [11], [16])

$$\begin{aligned}b_k &= \int_{-1}^1 \frac{f(x)T_k(x)}{\sqrt{1-x^2}} dx = \frac{2}{\pi}\int_0^\pi f(\cos\theta)\cos k\theta \, d\theta \\ &= \frac{1}{\pi}\int_{-\pi}^\pi f(\cos\theta)\cos k\theta\, d\theta \\ &= \frac{1}{2\pi i}\int_{|z|=1} f(\frac{1}{2}(z+z^{-1}))(z^k + z^{-k})\frac{dz}{z}.\end{aligned}$$

The function f is analytic in the whole complex plane and we can apply the Cauchy theorem and integrate on the right along the circles of radii R^{-1} and R respectively. We find that for $R > 1$ because of the definition of the class $\mathbf{F}_2$

$$\begin{aligned}|b_k| &\le \pi^{-1} R^{-k} \int_{-\pi}^\pi |f(1/2(Re^{i\theta} + R^{-1}e^{-i\theta}))| d\theta \\ &\le 2R^{-k} \max_{|z|=R} |f(z)| \le 2MR^{-k} \exp\{cR^\rho\}.\end{aligned}$$

If we take here $R = \left(\frac{k}{c\rho}\right)^{\frac{1}{\rho}}$, we find that $|b_k| \le 2Me^{k/\rho}\left(\frac{k}{c\rho}\right)^{-k/\rho}$. But then

$$|a_N| \le \sqrt{\sum_N |a_k|^2} \le \sqrt{\sum_N |b_k|^2} \le (C\rho)^{N/\rho} N^{-N/\rho}. \tag{4.33}$$

The Legendre polynomials satisfy the inequality (see [17]): for all $x \in [-1,1]$

$$|P_k(x)| \le P_k(1) = \sqrt{\frac{2k+1}{2}}. \tag{4.34}$$

It follows from (4.33) and (4.34) that for all p the L_p-norm of the bias

$$\|b_N\|_p = \|\sum_{N+1}^{\infty} a_k P_k(x)\|_p \leq\leq c_1 \sum_{N+1} c_2^k k^{-k/\rho} \leq (c_3 N)^{-N/\rho}. \tag{4.35}$$

2.The random term. The random term is a Gaussian random polynomial

$$z_N(x) = \varepsilon \sum_0^N \xi_k P_k(x)$$

of degree N whose coefficients

$$\xi_k = \int_{-1}^{1} P_k(x) dw(x)$$

are iid standard Gaussian variables. It has been proved in [1], pp 195, 196 that the norms $\|z_N\|_p$ satisfy the following inequalities:

$$\begin{aligned} &\mathbf{E}\|z_N\|_p \leq c_p \varepsilon \sqrt{N}, && 1 \leq p < 4, \\ &\mathbf{E}\|z_N\|_4 \leq c_4 \varepsilon \sqrt{N} (\ln N)^{1/4}, \\ &\mathbf{E}\|z_N\|_p \leq c_p \varepsilon N^{1-2/p}, && 4 < p \leq \infty. \end{aligned} \tag{4.36}$$

Now we are ready to establish the upper bounds of the theorem. Combining inequalities (4.35, (4.36) we find that

$$\mathbf{E}_f \|f - f_N\|_p \leq c_1 N^{-N/\rho} + c_2 \varepsilon \mathbf{E}\|z_N\|_p.$$

Letting here $N \asymp (\ln \frac{1}{\varepsilon})(\ln\ln \frac{1}{\varepsilon})^{-1}$ with a proper constant we prove the upper bound of the theorem.

*Lower bounds.*We begin with the following lemma of Ibragimov and Khasminskii, see [18], ch. VI.

Lemma 4.10 *Let $S = \{f_j, j = 1, \dots, M\}$ be a family of functions $f_j \in \mathbf{F}$ such that $\|f_i - f_j\|_p \geq 2\delta$ for all $i, j, i \neq j$. Then for large M, ε^{-1}*

$$\inf_{\hat{f}_\varepsilon} \sup_{f \in \mathbf{F}} \mathbf{E}_f \|f - \hat{f}_\varepsilon\|_p \geq\geq \frac{1}{2}\delta \left(1 - \frac{1}{\ln M} \sup_{f \in S} \frac{\|f\|_2^2}{2\varepsilon^2}\right). \tag{4.37}$$

The inequality is true for any $p \geq 1$.

The construction of the set S depends on p and we have to consider separately three cases: $1 \leq p < 4, p = 4, p > 4$. All these constructions essentially coincide with those from [1]. The main difference is that now the norming factor in these constructions is not $e^{-\gamma N}$ as in [1] but $N^{-\gamma N}$. We outline shortly how to treat the case $p < 4$.

If $p < 4$, we define the set S as such a set of functions

$$f_{\bar{a}}(x) = N^{-\gamma N} \sum_0^N a_j P_j(x), \qquad \bar{a} = (a_0, \dots, a_N), \qquad a_j = \pm 1,$$

that for any two functions $f_{\bar{a}}, f_{\bar{a}'} \in \mathbf{F}$ the distance

$$\|f_{\bar{a}} - f_{\bar{a}'}\|_p \geq \frac{1}{2}\sqrt{N} N^{-\gamma N}. \tag{4.37}$$

It is shown in [1] that the number M of the points of the set S is bigger than $e^{N/8}$.

The polynomials $P_k(z)$ satisfy the inequality (see [16]): for all complex z

$$|P_k(z)| \leq \sqrt{\frac{2k+1}{2}} |z + \sqrt{z^2 - 1}|^k.$$

Hence for $|z| = R$

$$|f_{\bar{a}}(z)| \leq C N^2 N^{-\gamma N} (1 + 2R)^N$$

and it is easy to see that for sufficiently large γ the functions $f_{\bar{a}} \in \mathbf{F}_2$.

Because of (4.36), (4.37) for any estimator $\hat{f}$ of f

$$\sup_{f \in \mathbf{F}_2} \mathbf{E}_f \|f - \hat{f}\|_p \geq \frac{1}{2}\sqrt{N} N^{-\gamma N} \cdot (1 - c\varepsilon^{-2} N^{-2\gamma N}).$$

If we take here N as the minimal integer for which $c\varepsilon^{-2} N^{-2\gamma N} \leq 1/2$, we find that

$$\sup_{f \in \mathbf{F}^2} \|f - \hat{f}\|_p \geq c\varepsilon \sqrt{\frac{\ln(1/\varepsilon)}{\ln\ln(1/\varepsilon)}}, \qquad c > 0.$$

The theorem is proved. ∎

5 A problem of extrapolation

An analytic function $f(z)$ possesses a remarkable property: being observed on an interval it becomes immediately known throughout its domain of analyticity. Of course the problem of recovering $f(z)$ from such observations is an ill posed problem and it would be interesting to know what will happen if the observations are noisy. We consider below an example of this problem.

Denote $\mathbf{F} = \mathbf{F}(M, \sigma)$ the class of integer functions $f(z)$ such that f are integer functions of exponential type $\leq \sigma$, real valued on the real line and such that $\|f\|_2 \leq M$ where $\|\cdot\|_2$ denotes the $L_2(R^1)$ norm. By the Paley-Wiener theorem (see, for example, [20]) functions $f \in \mathbf{F}$ admit the following representation

$$f(z) = \frac{1}{2\pi} \int_{-\sigma}^{\sigma} e^{izu} g(u) du \tag{5.1}$$

and hence

$$(5.2)\qquad |f(z)| \le \frac{1}{2\pi}\|g\|_2\sqrt{\int_{-\sigma}^{\sigma} e^{2|z||u|}du} \le e^{\sigma|z|}\|f\|_2 \min(\sqrt{2\sigma}, |z|^{-1}).$$

Suppose now that a function $f \in \mathbf{F}$ is observed on the interval $[a,b]$, where $-\infty < a < b < \infty$, in the Gaussian white noise of intensity ε. It means that one observes $X_\varepsilon(t)$ where

$$(5.3)\qquad dX_\varepsilon(t) = f(t)dt + \varepsilon dw(t), \qquad a \le t \le b,$$

$w(t)$ is the standard Wiener process. Consider the following problem: estimate the value $f(z)$ of a function $f \in \mathbf{F}$ at the point z on the base of observations (5.3).

Theorem 5.1 *There exist an estimator $\hat{f}_\varepsilon(z)$ of $f(z)$ such that uniformly in $\{z : |z| \le (\ln\frac{1}{\varepsilon})^\alpha\}, \alpha < 1$,*

$$(5.4)\qquad \sup_{f\in\mathbf{F}} \mathbf{E}_f|f(z) - \hat{f}_\varepsilon(z)|^2 \le C_\alpha \varepsilon^{2(1-\alpha)}.$$

Moreover

$$(5.5)\qquad \sup_{f\in\mathbf{F}} \mathbf{E}_f\{ \sup_{|z|\le(\ln\frac{1}{\varepsilon})^\alpha} |f(z) - \hat{f}_\varepsilon(z)|^2\} \le C_\alpha \varepsilon^{2(1-\alpha)}.$$

Proof We may and will suppose that the interval $[a,b] = [-1,1]$. Expand the function $f(t)$ into the Fourier series with respect to the orthonormal Legendre polynomials

$$(5.6)\qquad f(z) = \sum_0^\infty a_k P_k(z).$$

It follows from (5.2) (see (4.33)) that

$$(5.7)\qquad |a_k| \le C_1 e^{k+1/2}\left(\frac{\sigma}{k+1/2}\right)^{k+1/2} \le C_2\frac{\sigma^k}{k!\sqrt{k}}.$$

We have seen also that

$$(5.8)\qquad |P_k(z)| \le \sqrt{\frac{2k+1}{2}}(|z| + \sqrt{|z|^2+1})^k.$$

Hence the series (5.6) converges in the whole complex plane.

As before consider estimators

$$f_N(z) = \sum_0^N \hat{a}_k P_k(z), \qquad \hat{a}_k = \int_{-1}^1 P_k(t)dX_\varepsilon(t)$$

and study separately their bias and variance.

It follows from (5.7), (5.8) that for $|z| = R$ the bias satisfies

(5.9)

$$|f(z) - \mathbf{E}f_N(z)| \le \sum_{N+1}^{\infty} |a_k||P_k(z)| \le$$
$$\le C_1 \sum_{N+1}^{\infty} \frac{\sigma^k}{k!}(R + \sqrt{R^2+1})^k \le C_2 e^{c_1 R}(c_2 R)^N N^{-N}.$$

The random term is a Gaussian random polynomial

$$r_N(z) = \varepsilon \sum_0^N \xi_k P_k(z)$$

where ξ_k are iid standard Gaussian random variables. Hence

$$\mathbf{E}|r_N(z)|^2 = \varepsilon^2 \sum_0^N |P_k(z)|^2 \tag{5.10}$$
$$\le \varepsilon^2 \sum_0^N k|R + \sqrt{R^2+1}|^{2k} \le \varepsilon^2 N^2 (cR)^{2N}.$$

Moreover

$$\mathbf{E} \max_{|z| \le R} |r_N(z)| \le \mathbf{E} \sum_0^N |\xi_k|(R + \sqrt{R^2+1})^k \le (cR)^N. \tag{5.11}$$

Combining (5.9) and (5.10) we find that

$$\mathbf{E}|f(z) - f_N(z)|^2 \le e^{c_1 R}(c_2 R)^{2N}(N^{-2N} + \varepsilon^2).$$

It follows that if we let $N \sim (\ln \frac{1}{\varepsilon})(\ln\ln \frac{1}{\varepsilon})^{-1}$ and denote f_N with this N trough f_ε, then for $|z| = R$

$$\sup_{f \in \mathbf{F}} \mathbf{E}|f(z) - f_\varepsilon(z)|^2 \le c_1 \varepsilon^2 e^{c_2 R} \exp\{\frac{\ln \frac{1}{\varepsilon}}{\ln\ln \frac{1}{\varepsilon}} \ln R\}.$$

The last inequality yields the inequality (5.4). In the same way the inequalities (5.9, (5.11) imply the inequality (5.5). The theorem is proved. ■

Theorem 5.2 *Let $|z| > (\ln \frac{1}{\varepsilon})^\alpha, \alpha > 1$. Then for all sufficiently small $\varepsilon, \varepsilon < c_\alpha$,*

$$\inf_{f_\varepsilon} \sup_{f \in \mathbf{F}} \mathbf{E}_f |f(z) - f_\varepsilon|^2 \ge 1/2. \tag{5.12}$$

Proof We reduce our problem to a one dimensional estimation problem of a parameter θ and establish the lower bound for this new problem. Consider the following set of functions $f_\theta(t)$ depending on a real parameter $\theta, |\theta| \le 1$,

$$(5.13) \quad f_{\theta,T}(z) = f_\theta(z) = \theta \frac{\sin z}{z} \prod_{n=1}^{\infty} \left(\frac{\sin(\alpha-1)(z-T)n^{-\alpha}}{(z-T)(\alpha-1)} n^{\alpha} \right) = \theta\phi(z)$$

where $\alpha > 1, T$ are real numbers. Because of $\left|\frac{sinz}{z}\right| \le e^{|z|}$ the product on the right converges uniformly on compacts of the complex plane and determines an integer function. Moreover this integer function has an exponential growth ≤ 1. Indeed

$$|f_\theta(z)| \le \theta \exp\{|z| + |z-T|(\alpha-1)\sum_1^\infty n^{-\alpha}\} \le e^T e^{2|z|}.$$

In the same time because on the real line $|\frac{\sin x}{x}| \le 1$

$$\|f\|_2^2 \le \int_{-\infty}^{\infty} \frac{\sin^2 t}{t^2} dt = \pi.$$

Thus all $f_\theta \in \mathbf{F}(\sqrt{\pi}, 2)$.

Let $\hat{f}$ be an estimator for $f_\theta(T)$. We have

$$(5.14) \qquad \mathbf{E}_\theta |f_\theta(T) - \hat{f}|^2 = \mathbf{E}_\theta |\theta - \hat{f}|^2.$$

Consider now the problem of estimation of the parameter θ on the base of observation $X_\varepsilon(t)$. It is easy to see that the statistic

$$Y_\varepsilon = \int_{-1}^1 \phi(t) dX_\varepsilon(t) = \theta \int_{-1}^1 \phi^2(t) dt + \varepsilon \int_{-1}^1 dw(t)$$

is a sufficient statistic of the θ estimation problem. Hence

$$(5.15) \quad \sup_\theta \mathbf{E}_\theta |\theta - \hat{f}_\varepsilon|^2 \ge \sup_\theta \mathbf{E}_\theta |\theta - \mathbf{E}\{\hat{f}_\varepsilon | Y_\varepsilon\}|^2 \ge \inf_\psi \sup_\theta \mathbf{E}_\theta |\theta - \psi(Y_\varepsilon)|^2$$

where

$$Y_\varepsilon = \theta + \varepsilon \frac{\int_{-1}^1 \phi(t) dw(t)}{\int_{-1}^1 \phi^2(t) dt} = \theta + \eta$$

and inf is taken over all statistics $\psi(Y_\varepsilon)$.

We use the following well known result, see [19]:

Lemma 5.3 *Suppose one is estimating a parameter $|\theta| \le a$ on the base of observation*

$$Y = \theta + \eta$$

where η is a Gaussian random variable with mean zero and variance s^2. Then

$$\inf_{\psi} \sup_{\theta} \mathbf{E}_\theta |\theta - \psi(Y)|^2 \geq s^2 h(a/s) \tag{5.16}$$

where the function h is an increasing function such that $h(x) \sim x^2$ when $x \to 0$ and $h(\infty) = 1$

It follows from the last result together with (5.14) and (5.15) that

$$\inf_{f_\varepsilon} \sup_{f \in \mathbf{F}} \mathbf{E}|f(T) - f_\varepsilon|^2 \geq c\varepsilon \left(\int_{-1}^{1} \phi^2(t) dt \right)^{-1}, \qquad c > 0.$$

Let us estimate the integral on the right. We have

$$\int_{-1}^{1} \phi^2(t) dt \leq \prod_{1}^{\infty} \min(1, n^{2\alpha}(T-1)^{-2}(\alpha-1)^{-2}) \leq ((T-1)(\alpha-1))^{-2L}(L!)^{2\alpha}$$

where L is the largest integer $\leq ((T-1)(\alpha-1))^{1/\alpha}$. Hence for large T

$$\int_{-1}^{1} \phi^2(t) dt \leq c_2 (T(\alpha - 1))^\alpha e^{-2\alpha T^{1/\alpha}}.$$

It follows that for any f_ε

$$\sup_{f \in \mathbf{F}(2,1)} \mathbf{E}_f \{ |f(T) - f_\varepsilon|^2 \} \geq 1/2$$

if $T > c(\alpha)(\ln \frac{1}{\varepsilon})^\alpha$. With slight changes the same arguments will work for any $z, |z| > c(\alpha)(\ln \frac{1}{\varepsilon})^\alpha$. The theorem is proved. ∎

Remark 5.1 The Theorems 5.1 and 5.2 mean that roughly speaking when $\varepsilon \to 0$ consistent estimation of $f(z)$ is possible on the disks $\{|z| << \ln \frac{1}{\varepsilon}\}$ and impossible outside larger disks, namely on $\{|z| >> \ln \frac{1}{\varepsilon})\}$.

Acknowledgements I wish to thank Professors Chris Klaassen and Aad van der Vaart who read the manuscript and made many helpful comments.

REFERENCES

[1] I. Ibragimov, (1998). On estimation of analytic functions. *Studia Sci Math Hungarica* **34**, 191–210.

[2] I. Ibragimov, (1998). *Estimation of analytic spectral density of Gaussian stationary process.* ANU preprint No.SRR 006–98, Australian Nat Univ.

[3] I. Ibragimov, R. Khasminskii, (1983). On the estimation of distribution density. *J. of Soviet Math* **24**, 40–57.

[4] I. Ibragimov, R. Khasminskii, (1977). A problem of statistical estimation in Gaussian white noise. *Soviet Math Dokl* **236**, 1053–1056.

[5] M. Pinsker, (1980). Optimal filtering of square integrable signals in Gaussian white noise. *Problems of Information Transmission* **16**, 120–133.

[6] E. Guerre, A. Tsybakov, (1998). Exact asymptotic minimax constants for the estimation of analytical function in L_p. *Probab Theory Relat Fields* **112**, 33–51.

[7] G. Golubev, B. Levit, A. Tsybakov, (1996). Asymptotically efficient estimation of analytic functions in Gausian white noise. *Bernoulli* **2**, 167–181.

[8] E. Titchmarsh, (1937). *Introduction to the theory of Fourier integrals.* Oxford Univ Press.

[9] H. Rosenthal, (1970). On the subspaces of $L_p(p > 2)$ spanned by sequences of independent random variables. *Israil J Math* **8**, 273–303.

[10] V.Petrov, (1995). *Limit theorems of probability theory.* Clarendon Press, Oxford.

[11] A. Timan, (1963). *Theory of approximation of functions of a real variable.* McMillan, N.Y..

[12] I. Ibragimov, R. Khasminskii, (1989). On density estimation in the view of Kolmogorov's ideas in approximation theory. *Ann.Stat.* **18**, 999–1010.

[13] I. Ibragimov, R. Khasminskii, (1982). On density estimation within a class of entire functions *(in Russian)*. *Theory Prob Appl* **27**, 514–524.

[14] I. Ibragimov, R. Khasminskii, (1982). Bounds for the quality of nonparametric estimation of regression. *Theory Prob Appl* **27**, 84–99.

[15] I. Ibragimov, R. Khasminskii, (1984). On nonparametric estimation of regression in L_p norms. *J of Soviet Math* **25**, 540–550.

[16] S. Bernstein, (1937). *Extremal properties of polynomials (in Russian).* ONTI, Moscow.

[17] Szegö, (1959). *Orthogonal polinomials,* AMS Colloq.Publ vol.XXIII.

[18] I. Ibragimov, R. Khasminskii, (1981). *Statistical estimation: asymptotic theory.* Springer, N.Y..

[19] I. Ibragimov, R. Khasminskii, (1984). On nonparametric estimation of the value of a linear functional in Gaussian white noise. *Theory Probab Appl* **29**, 18–32.

[20] R. Boas, (1954). *Entire Functions.* New York Academic Press.

St.Petersburg Branch of
Steklov Mathematical Institute Russian Ac.Sci.
Fontanka 27
St.Petersburg
191011, Russia
ibr32@pdmi.ras.ru