Записки научных
семинаров ПОМИ
Том 363, 2009 г.

R. Khasminskii

ESTIMATION OF NONLINEAR FUNCTIONALS REVISITED

Abstract. After the overview of known results concerning estimating the linear and nonlinear functionals of the density for i.i.d. observations and for functionals of the signal observed in the White Gaussian Noise (WGN) with small intensity we consider the similar problems for the observation the Poisson random process. Asymptotically efficient estimates of the once Frechet differentiable functionals are proposed.

1. Introduction

Estimation nonlinear functionals of unknown function related to the observations probability distribution is one of important problems in the nonparametric statistics. The natural minimax lower bound of the estimation risks for differentiable functionals of probability density function for i.i.d. observations was obtained by B. Levit in [7].

Let $X_1, \dots, X_n$ be i.i.d. observations with unknown distribution density $p(x) = \frac{dF}{dx}, x \in \mathbb{R}$, and $\Phi(p)$ be the differentiable functional with derivative $\Phi'(p) \in \mathbb{L}_2(F)$ (i.e., $\Phi'(p)(x)$ is a square integrable function with respect to the measure $F(dx)$). Then it was proven in [7] that for the wide class of the loss functions $l(x)$ and any $\delta > 0$ the lower bound

$$\lim_{n\to\infty} \inf \sup_{p\in\mathcal{O}_\delta(p_0)} E_p l(\sqrt{n}(\Phi_n - \Phi(p))) \geq El(\sigma(p_0)\xi); \qquad (1)$$

here ξ is the $\mathcal{N}(0,1)$ random variable, $\mathcal{O}_\delta(p_0)$ is the δ-vicinity of p_0 in the suitable metric,

$$\sigma^2(p) = \int_{\mathbb{R}} [\Phi'(p)(x) - E_p\Phi'(p)(X_1)]^2 p(x)dx,$$

is valid for any estimate Φ_n (see details in [7] or [4, Chap. 4]). Here and below E_p means the expectation with respect to the probability measure with the density $p(x)$. Asymptotically efficient estimates (i.e., estimates with the equality in (1)) for some classes of functionals were found in [8].

General method for estimation the functionals with various smoothness was proposed in [2].

Creation of asymptotically efficient nonparametric estimate for $\Phi(p)$ in [2] has two steps. The first one is the creation an estimate for $p(x)$. It is possible to construct uniformly consistent estimate in L_2-norm, if p belongs to some class of smooth functions. Let $\mathcal{W}(\beta, L)$ be the Sobolev space of densities with smoothness β, i.e.,

$$\mathcal{W}(\beta, L) = \left\{ p : \int\limits_{\mathbb{R}^1} |t|^{2\beta} |\widehat{p}(t)|^2 dt \leqslant L \right\};$$

here $\widehat{p}$ is the Fourier transform of p. Then there exist a kernel estimate $p_n(x, \omega)$ such that for any $r > 0$

$$\limsup_{n\to\infty} \sup_{p\in\mathcal{W}(\beta,L)} E_p \left[\|p_n(x,\omega) - p(x)\|_{L_2(\mathbb{R})} n^{\frac{\beta}{2\beta+1}} \right]^r < \infty. \tag{2}$$

(See, e.g., [1] or similar assertion in [3].)

The second step is a creation of the estimator Φ_n making use of the smoothness $\Phi(p)$.

Theorem 1.1. *Let $X_1, X_2, \dots, X_n$ be i.i.d. observations with the density $p(x) \in \mathcal{W}(\beta, L)$, the functional $\Phi(p)$ be Frechet differentiable and satisfies the conditions*

$$\begin{aligned} &\|\Phi'(p)(\cdot)\|_{L_2(F)} < C; \\ &\|\Phi'(p_2)(\cdot) - \Phi'(p_1)(\cdot)\|_{L_2(\mathbb{R})} < C\|p_2 - p_1\|^{\gamma}_{L_2(\mathbb{R})}, \end{aligned} \tag{3}$$

with

$$\gamma > (2\beta)^{-1}. \tag{4}$$

Then there exists estimate Φ_n, based on $X_1, X_2, \dots, X_n$, with property

$$\limsup_{n\to\infty} \sup_{p\in W(\beta,L)} \Big[nE_p|\Phi_n - \Phi(p)|^2 - \int [\Phi'(p)(x) - E_p\Phi'(p)(X_1)]^2 p(x)dx \Big] = 0.$$

This theorem follows immediately from (2) and Theorem 2 in [2]. Analogous approach was used in [5, 9] for the estimation of smooth functionals $\Phi(S)$ for the observation model "Signal plus WGN:"

$$dX_\varepsilon(t) = S(t)dt + \varepsilon dw(t), \quad 0 \leqslant t \leqslant 1;$$

here $w(t)$ is a standard Wiener process, $\varepsilon \to 0$. In [5, 9], the result, closed to the Theorem 1.1, was obtained if S belongs to the class of functions with smoothness β, and a functional $\Phi(S)$ is also smooth enough. It is necessary to emphasize that asymptotically efficient estimates were created in [2, 5] also for more smooth functionals under less restrictive assumptions concerning smoothness p and S. But here we consider for simplicity only the case of once differentiable functionals.

In this paper, we consider the functional estimation problem for the observation of the Poisson process with unknown intensity function.

2. Statement of the problem and preliminaries

We consider here the nonlinear functionals estimation problem for two observation models of the Poisson process.

Model 1. $X_1(t), X_2(t), \dots, X_n(t)$, $t \in [0, T]$ are i.i.d. Poisson processes with the intensity function $S(t)$. The problem is the estimation of differentiable functional $\Phi(S)$ on the base of $X_1(\cdot), X_2(\cdot), \dots, X_n(\cdot)$.

Model 2. $X_\varepsilon(t), t \in [0, T]$ is the Poisson process with the intensity function $\varepsilon^{-1}S(t)$, $\varepsilon \to 0$. The problem is the estimation of $S(t)$ on the base of observation $X_\varepsilon(t)$.

The linear functionals estimation problem for these models is very simple. Let

$$L(S) = \int_0^T f(t)S(t)dt.$$

Then the estimate

$$L_n = \frac{1}{n}\sum_{i=1}^{n}\int_0^T f(t)X_i(dt)$$

is evidently unbiased, asymptotically normal, as $n \to \infty$, and

$$E|L_n - L(S)|^2 = \frac{1}{n}\int_0^T f^2(t)S(t)dt,$$

see [6]. Here and below, E means expectation with respect to the measure generated by the Model 1 with the intensity $S(t)$ or by the model 2 with the intensity $S(t)/\varepsilon$.

Analogously, the estimate

$$L_\varepsilon = \varepsilon \int_0^T f(t) X_\varepsilon(dt)$$

is unbiased, asymptotically normal, as $\varepsilon \to 0$, and

$$E|L_\varepsilon - L(S)|^2 = \varepsilon \int_0^T f^2(t) S(t) dt.$$

It is easy to see that these estimates are asymptotically unimprovable in the minimax sense (see [4], [6] for details).

Theorem 2.1 below gives asymptotically minimax lower bound of risks for nonlinear functionals. Let $\mathcal{S}$ be the set of continuous and strictly positive functions on $[0, T]$. Assume that the functional $\Phi(S)$ is weakly (in Gateaux sense) differentiable: for any $h(t) \in L_2(0, T)$, $\lambda \to 0$

$$\Phi(S + \lambda h) = \Phi(S) + \lambda \int_0^T \Phi'(S, t) h(t) dt + o(\lambda). \tag{5}$$

Denote $\mathcal{O}_\delta(S)$ the δ-vicinity of S in uniform metric, and W the set of functions $l : \mathbb{R} \to \mathbb{R}^+$ such that $l(-x) = l(x)$, $l(0) = 0$ and $l(x)$ nondecreasing for $x > 0$ functions.

Theorem 2.1. *Let $\Phi(S)$ be differentiable in the sense (5) functional, and $\Phi'(S) \in L_2[0, T]$. Assume that $S \in \mathcal{S}$, and denote $\mathcal{F}_n$ the set of estimates based on $X_1(\cdot), X_2(\cdot), \ldots, X_n(\cdot)$ for the Model 1, and $\mathcal{F}_\varepsilon$ the set of estimates based on $X_\varepsilon(\cdot)$ for the Model 2. Then for any $l \in W$, $\delta > 0$,*

$$\liminf_{n\to\infty} \inf_{\Phi_n \in \mathcal{F}_n} \sup_{S \in \mathcal{O}_\delta(S_0)} El(\sqrt{n}(\Phi_n - \Phi(S))) \geqslant El(\sigma(S_0)\xi), \tag{6}$$

and

$$\liminf_{\varepsilon\to 0} \inf_{\Phi_\varepsilon \in \mathcal{F}_\varepsilon} \sup_{S \in \mathcal{O}_\delta(S_0)} El(\frac{1}{\sqrt{\varepsilon}}(\Phi_\varepsilon - \Phi(S))) \geqslant El(\sigma(S_0)\xi), \tag{7}$$

where ξ is the $\mathcal{N}(0,1)$ random variable,

$$\sigma^2(S_0) = \int_0^T (\Phi'(S_0, t))^2 S_0(t) dt. \tag{8}$$

Proof of (6) and (7) is analogous to the proof of similar assertions for the estimation differentiable functionals of probability density in [7, 4]. So, we only give a sketch of the proof for the lower bound (7).

Denote

$$\Phi'(S,t)_N := \Phi'(S,t)\mathbf{1}_{\{t:|\Phi'(S,t)|\leqslant N\}}$$

and consider the parametric family

$$S_N(\theta,t) = S_0(t) + (\theta-\theta_0)S_0(t)\Phi'(S,t)_N\Big(\int_0^T \Phi'(S_0,t)_N^2 S_0(t)dt\Big)^{-1}. \quad (9)$$

Then due to (5) we have, denoting $\theta_0 = \Phi(S_0)$,

$$\begin{aligned}\Phi(S_N(\theta,t)) &= \Phi(S_0) + (\theta-\theta_0)\frac{\int_0^T \Phi'(S_0,t)(\Phi'(S_0,t))_N S_0(t)dt}{\int_0^T (\Phi'(S_0,t))_N^2 S_0(t)dt} + o(\theta-\theta_0)\\ &= \theta + o(\theta-\theta_0), \quad (\theta-\theta_0\to 0). \qquad (10)\end{aligned}$$

It is easy to check that the Fisher information $I_\varepsilon(\theta)$ for the Model 2 with the intensity $S(\theta,t)/\varepsilon$ is equal to

$$I_\varepsilon(\theta) = \frac{1}{\varepsilon}\int_0^T \frac{(\frac{\partial S}{\partial\theta}(\theta,t))^2}{S(\theta,t)}dt.$$

Thus for $S(\theta,t) = S_N(\theta,t)$ we have

$$I_\varepsilon^{(N)}(\theta_0) = \frac{1}{\varepsilon}\left[\int_0^T (\Phi'(S_0,t)_N)^2 S_0(t)dt\right]^{-1}.$$

It is evident that for any $N>0$ and $\delta>0$, the family (9) belongs to $\mathcal{O}_\delta(S_0)$ for $|\theta-\theta_0|$ small enough. It follows from (10) that an estimation of the functional $\Phi(S)$ for the family (8) is equivalent, for $\theta\to\theta_0$, to the estimation of the parameter θ. The lower bound (7) follows from these facts and the minimax lower bound for the parameter estimation (see [6]). □

3. Asymptotically efficient estimates

3.1. Estimation of intensity function

Similarly to the functionals estimation for the statistical models mentioned above, the preliminary estimation of $S(t)$ is essential. Some a priori known smoothness of $S(t)$ is necessary for it.

Denote by $\Sigma(\beta, L)$ the set of positive on $[0,T]$ k times differentiable functions with the property: for any t, $t+h \in [0,T]$

$$|S^{(k)}(t+h) - S^{(k)}(t)| \leqslant L|h|^\alpha; \quad 0 < \alpha \leqslant 1, \quad \beta = k + \alpha.$$

Lemma 3.1. *Let $S(t) \in \Sigma(\beta, L)$ and $g_i(t)$, $i = 1, 2$, be $\mathbb{R} \to \mathbb{R}$ functions with a compact support and properties*

$$\int\limits_{\mathbb{R}} g_i(t)dt = 1; \quad \int\limits_{\mathbb{R}} t^j g_i(t)dt = 0; \quad j = 1, \dots, k; \tag{11}$$

$g_1(t) = 0$, as $t > 0$; $g_2(t) = 0$ as $t < 0$.

Consider the estimates

$$S_n(t) = \frac{1}{nh_n} \sum_{i=1}^{n} \int\limits_0^T \Big[g_1\Big(\frac{t-s}{h_n}\Big) \mathbf{1}_{\{t<T/2\}} + g_2\Big(\frac{t-s}{h_n}\Big) \mathbf{1}_{\{t\geqslant T/2\}} \Big] X_i(ds), \tag{12}$$

$$S_\varepsilon(t) = \frac{\varepsilon}{h_\varepsilon} \int\limits_0^T \Big[g_1\Big(\frac{t-s}{h_\varepsilon}\Big) \mathbf{1}_{\{t<T/2\}} + g_2\Big(\frac{t-s}{h_\varepsilon}\Big) \mathbf{1}_{\{t\geqslant T/2\}} \Big] X_\varepsilon(ds) \tag{13}$$

with $h_n = n^{-1/(2\beta+1)}$ and $h_\varepsilon = \varepsilon^{1/(2\beta+1)}$ for the Models 1 and 2, respectively. Then for any loss function $l(x) : \mathbb{R} \to \mathbb{R}^+$ with the property $l(x) \leqslant c_1 \exp(c_2|x|)$ ($c_i > 0$ are the constants) the upper bound

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} El\big(n^{\beta/(2\beta+1)}(\widehat{S_n}(t) - S(t))\big) < \infty$$

$$\limsup_{\varepsilon\to 0} \sup_{t\in[0,T]} El\big(\varepsilon^{-\beta/(2\beta+1)}(\widehat{S_\varepsilon}(t) - S(t))\big) < \infty$$

are valid.

Proof of this lemma for the Model 1 and the case where the sup over $[0,T]$ replaced by the sup over $[a,b]$, $0 < a < b < T$, is done in [6], see Proposition 6.3 there. (It was enough in [6] to consider similar (12) estimate with the same function $g(.)$ for all $t \in [a,b]$). Proposed in (12), (13) modification of the kernel estimates allows to make the upper bound for risks uniform in $t \in [0,T]$ with help of the same reasoning, as in [6], see also [4]. Let us prove, for instance, that the bias $\widehat{S}_\varepsilon(t)$ has the order $\varepsilon^{\beta/(2\beta+1)}$ uniformly in $t \in [0,T]$. Making use of well-known formula for the expectation of the stochastic integral with respect to the Poisson process (see, e.g. [6], Lemma 1.1) we have for $t > T/2$

$$\operatorname{bias} \widehat{S}_\varepsilon(t) = E\widehat{S}_\varepsilon(t) - S(t) = \frac{1}{h_\varepsilon}\int_0^T g_2\left(\frac{t-s}{h_\varepsilon}\right)S(s)ds - S(t)$$

$$= \int_{\mathbb{R}} g_2(z)[S(t-h_\varepsilon z) - S(t)]dz = \int_0^\infty g_2(z)[S(t-h_\varepsilon z) - S(t)]dz.$$

Applying the Taylor formula and (11), we have

$$\sup_{T/2\leqslant t\leqslant T} |\operatorname{bias} \widehat{S}_\varepsilon(t)| \leqslant Ch_\varepsilon^\beta = C\varepsilon^{\beta/(2\beta+1)}.$$

Applying the same reasoning for $0 < t < T/2$, we arrive at the upper bound

$$\sup_{0\leqslant t\leqslant T} |\operatorname{bias} \widehat{S}_\varepsilon(t)| \leqslant C\varepsilon^{\beta/(2\beta+1)}.$$

□

3.2. Functional estimation for the Model 1

Let $X_1(t), \ldots, X_n(t), t \in [0,T]$ be i.i.d. Poisson processes with bounded intensity function $S(t) \in \Sigma(\beta, L)$. Assume that $\Phi(S)$ is the Frechet differentiable functional with derivative $\Phi'(S,t)$, satisfying the conditions

$$\|\Phi'(S,\cdot)\| < C; \quad \|\Phi'(S_2,\cdot) - \Phi'(S_1,\cdot)\| \leqslant C\|S_2 - S_1\|^\gamma. \qquad (14)$$

Here and below, the notation $\|f\|$ is used for $\mathbb{L}_2(0,T)$-norm of f, by C, C_i we denote a the constants, independent of n, ε, may be different in different appearances. Similar the models, described above, the creation of an estimate Φ_n for $\Phi(S)$ has two steps. On the first step we use $[n^\delta], 0 < \delta < 1$, first observations for the estimation $S(t)$. Due to Lemma 3.1 and Cauchy–Schwarz inequality this estimate $S_n(t)$ is consistent, and

$$\limsup_{n\to\infty} \sup_{S\in\Sigma(\beta,L)} E\|S_n(\cdot)-S(\cdot)\|^r \leqslant C_r n^{-\beta\delta r/(2\beta+1)} \tag{15}$$

for any $r > 0$.

Assume now that $\gamma > (2\beta)^{-1}$ and consider the estimate

$$\Phi_n = \Phi(S_n) + \frac{1}{n-[n^\delta]} \sum_{i=[n^\delta]+1}^{n} \int_0^T \Phi'(S_n,t)(X_i(dt) - S_n(t)dt) \tag{16}$$

with $\delta \in (\frac{1+1/(2\beta)}{1+\gamma}, 1)$.

Theorem 3.2. *Let $S(t) \in \Sigma(\beta,L), t \in [0,T]$, and the functional $\Phi(S)$ satisfies conditions (14) with $\gamma > (2\beta)^{-1}$. Then the normalized estimate (16) $\zeta_n = \sqrt{n}(\Phi_n - \Phi(S))$ is asymptotically normal with parameters $(0, \sigma^2(S))$ and asymptotically efficient uniformly in $\Sigma(\beta,L)$ for any loss function $l(x)$ admitting a polynomial majorant.*

Proof. Denote by $\mathcal{G}_n$ the σ-algebra generated by $X_1(\cdot), \dots, X_{[n^\delta]}(\cdot)$. Then we have from (16):

$$E\{\Phi_n|\mathcal{G}_n\} = \Phi(S_n) + \int_0^T \Phi'(S_n,t)(S(t)-S_n(t))dt. \tag{17}$$

On the other hand, it follows from Lagrange mean value theorem and (14) that

$$\Phi(S) = \Phi(S_n) + \int_0^T \Phi'(S_n,t)(S(t)-S_n(t))dt + \mathcal{O}(\|S(\cdot)-S_n(\cdot)\|^{1+\gamma}). \tag{18}$$

Hence $E\{\Phi_n|\mathcal{G}_n\} = \Phi(S) + \mathcal{O}(\|S(\cdot) - S_n(\cdot)\|^{1+\gamma})$ and we have, making use of (15) and $\delta > [1+1/(2\beta)]/(1+\gamma)$, that uniformly in $S(t) \in \Sigma(\beta,L)$

$$|E\Phi_n - \Phi(S)| \leqslant CE\|S(\cdot)-S_n(\cdot)\|^{1+\gamma} \leqslant Cn^{-\delta\beta(1+\gamma)/(2\beta+1)} = o(n^{-1/2}). \tag{19}$$

Analogously, making use of (14) again, we have

$$\sqrt{n}(\Phi_n - \Phi(S)) = \frac{1}{\sqrt{n}} \sum_{i=[n^\delta]+1}^{n} \int_0^T \Phi'(S,t)(X_i(dt) - S(t)dt) + \xi_n. \quad (20)$$

Here $\xi_n \to 0$ in probability uniformly in $S(t) \in \Sigma(\beta, L)$. The asymptotic normality of Φ_n with parameters $(0, \sigma^2(S))$ follows from (19), (20), and the central limit theorem. The asymptotic efficiency of Φ_n for the bounded loss functions also follows. In order to finish the proof of theorem, it is enough to prove that

$$\limsup_{n\to\infty} \sup_{S\in\sum(\beta,L)} E|\sqrt{n}(\Phi_n - E(\Phi_n|\mathcal{G}_n))|^{2r} < \infty.$$

Making use of Lemma 1.2 in [6], we have for $r \geqslant 1$,

$$E|\sqrt{n}(\Phi_n - E(\Phi_n|\mathcal{G}_n))|^{2r} = E\{E|\sqrt{n}(\Phi_n - E(\Phi_n|\mathcal{G}_n))|^{2r}|\mathcal{G}_n\}$$

$$= E\biggl\{E\biggl|\frac{\sqrt{n}}{n-[n^\delta]}\int_0^T \Phi'(S_n,t)\biggl(\sum_{i=[n^\delta]+1}^{n} X_i(dt) - (n-[n^\delta])S(t)dt\biggr)\biggr|^{2r}\biggl|\mathcal{G}_n\biggr\}$$

$$\leqslant \frac{1}{n^r}E\biggl\{\int_0^T |\Phi'(S_n,t)|^{2r}(n-[n^\delta])S(t)dt$$

$$+ \biggl(\int_0^T |\Phi'(S_n,t)|^2(n-[n^\delta])S(t)dt\biggr)^r\biggr\}(1+o(1)) \leqslant C.$$

□

3.3. Functional estimation for the Model 2

We consider the estimation of a smooth functional $\Phi(S)$ on the base of the $X_\varepsilon(t)$, $t \in [0,T]$. As before, $X_\varepsilon(t)$ is the Poisson process with intensity function $S(t)/\varepsilon$. In order to create estimate for $\Phi(S)$, we need the following elementary result, known as the thining of a Poisson process, see, e.g., [10, Proposition 5.2]. Let $X(t)$, $t \in [0,T]$ be the Poisson process with intensity function $\lambda(t)$. Let τ_i, $i = 1, 2, \ldots$ be the times of "jumps" for $X(t)$, so that

$$X(t) = \sum_i 1_{\{\tau_i < t\}}.$$

Let $\xi_1, \xi_2, \ldots$ be i.i.d. and independent of $X(t)$ random variables such that $P\{\xi_i = 1\} = p = 1 - P\{\xi_i = 0\}$. Create the new process $Y(t)$ as follows

$$Y(t) = \sum_i \xi_i \mathbf{1}_{\{\tau_i < t\}} \tag{21}$$

Lemma 3.3. *The process $Y(t)$ is a Poisson process with the intensity function $p\lambda(t)$. The process $X(t) - Y(t)$ is also Poissonian with intensity function $(1-p)\lambda(t)$. The processes $Y(t)$ and $X(t)-Y(t)$ are independent.*

As above, the creation of the estimate Φ_ε for $\Phi(S)$ will consist of several steps.

Step 1. Create the process $Y_\varepsilon(t)$ by (21) with $X(t) = X_\varepsilon(t)$, $p = \varepsilon^\delta$, $0 < \delta < 1$. Note that $Y_\varepsilon(t)$ is created on the base the observable process $X_\varepsilon(t)$ and the auxiliary i.i.d. random variables $\xi_1, \xi_2, \ldots$. Hence $Y_\varepsilon(t)$ is also observable one.

Step 2. Create a kernel estimate $S_\varepsilon(t)$ for $S(t)$ on the base of $Y_\varepsilon(t)$. Due to the Lemma 3.3 the process $Y_\varepsilon(t)$ has the intensity function $S(t)/\varepsilon^{1-\delta}$. Hence, applying Lemma lm3.1, we have

$$E\|\widehat{S}_\varepsilon(t) - S(t)\|^r \leqslant C\varepsilon^{\frac{(1-\delta)r\beta}{2\beta+1}}. \tag{22}$$

Step 3. Create the estimate

$$\Phi_\varepsilon = \Phi(S_\varepsilon) + \int_0^T \Phi'(S_\varepsilon, t)\left[\frac{\varepsilon}{1-\varepsilon^\delta}\big(X_\varepsilon(dt) - Y_\varepsilon(dt)\big) - S_\varepsilon(t)dt\right]. \tag{23}$$

Let $\sigma^2(S)$ be defined by (8).

Theorem 3.4. *Let $S(\cdot) \in \sum(\beta, L)$, the functional $\Phi(S)$ satisfies conditions* (14) *with $\gamma > (2\beta)^{-1}$. Let the estimate S_ε and δ in Step* 2 *are chosen so that*

$$(1 + (2\beta)^{-1})/(1+\gamma) < 1 - \delta < 1. \tag{24}$$

Then the normalized estimate $\zeta_\varepsilon = \varepsilon^{-\frac{1}{2}}(\Phi_\varepsilon - \Phi(S))$ is asymptotically Gaussian with parameters $(0, \sigma^2(S))$ and asymptotically efficient uniformly in $\sum(\beta, L)$ for any loss function $l(x)$ with polynomial majorant.

Proof. Denote $\mathcal{G}_\varepsilon$ the σ-algebra generated by the process $Y_\varepsilon(t)$, $t \in [0,T]$. Due to Lemma 3.3, the process $X_\varepsilon(t) - Y_\varepsilon(t)$ is independent of this σ-algebra. Hence

$$E(\Phi_\varepsilon|\mathcal{G}_\varepsilon) = \Phi(S_\varepsilon) + \int_0^T \Phi'(S_\varepsilon, t)(S(t) - S_\varepsilon(t))dt. \tag{25}$$

Then, analogously to (19), we obtain from (25), (22), and (24) that for $\gamma > (2\beta)^{-1}$

$$|E\Phi_\varepsilon - \Phi(S)| = o(\sqrt{\varepsilon}) \quad (\varepsilon \to 0). \tag{26}$$

Analogously to (20), we find also that

$$\frac{1}{\sqrt{\varepsilon}}(\Phi_\varepsilon - \Phi(S))$$

$$= \frac{\sqrt{\varepsilon}}{1-\varepsilon^\delta}\int_0^T \Phi'(S,t)\Big[X_\varepsilon(dt) - Y_\varepsilon(dt) - \frac{1-\varepsilon^\delta}{\varepsilon}S(t)dt\Big] + \zeta_\varepsilon := Z_\varepsilon + \zeta_\varepsilon;$$

here $\zeta_\varepsilon \to 0$ in probability, as $\varepsilon \to 0$. Note that

$$E(X_\varepsilon(t) - Y_\varepsilon(t)) = \frac{1-\varepsilon^\delta}{\varepsilon}S(t),$$

so Z_ε is the integral with respect to the centered Poisson process. So, making use of Lemma 1.1 in [6] again, we have, as $\varepsilon \to 0$

$$E\exp(i\lambda Z_\varepsilon)$$

$$= \exp\Big\{\int_0^T \Big[e^{i\frac{\lambda\varepsilon^{1/2}}{1-\varepsilon^\delta}\Phi'(S,t)} - 1 - \frac{i\lambda\varepsilon^{1/2}}{1-\varepsilon^\delta}\Phi'(S,t)\Big]\frac{S(t)(1-\varepsilon^\delta)}{\varepsilon}ds\Big\}$$

$$= \exp\left\{\int_0^T -\frac{1}{2}\frac{\lambda^2\varepsilon}{(1-\varepsilon^\delta)^2}(\Phi'(S,t))^2\frac{S(t)(1-\varepsilon^\delta)}{\varepsilon}dt(1+o(1))\right\}$$

$$= \exp\left\{-\frac{\lambda^2}{2}\int_0^T(\Phi'(S,t))^2 S(t)dt\right\}(1+o(1)).$$

Hence $\dfrac{1}{\sqrt{\varepsilon}}(\Phi_\varepsilon - \Phi(S))$ is asymptotically Gaussian $\mathcal{N}(0,\sigma^2(S))$ for $\varepsilon \to 0$. The rest part of the proof is completely analogous to the proof of Theorem 3.2. □

4. Concluding remarks

1. The Model 1 can be reduced to the Model 2: for independent Poisson processes $X_1(t), \dots, X_n(t)$ with the intensity function $S(t)$ the process

$Z_n(t) = \sum_{i=1}^{n} X_i(t)$ is a Poisson process with the intensity $nS(t)$. Thus $Z_n(t)$ is a particular case of the Model 2 with $\varepsilon = 1/n$. Nevertheless we have considered the estimation of nonlinear functionals for the Models 1 and 2 separately because estimate (16) for the Model 1 is essentially simpler: the auxiliary Poisson process $Y_\varepsilon(t)$ is not needed for this case.

2. We have considered here the estimation nonlinear functionals of the intensity function for the Poisson random process $X(t)$, $t \in [0, T]$. The analogous problem for the case of observation of Poisson random measure $X(A)$, $A \subset D$, D is the subset of $\mathbb{R}^l$, can be considered by the similar approach.

3. The estimation problem of once continuously differentiable functional has been considered here. Other estimates, analogous to the ones proposed in [2, 5, 9] for more smooth functionals, can be applied for the models considered here, too. I suppose that these estimates will be asymptotically efficient under less restrictive assumptions concerning the smoothness of intensity function.

Acknowledgments

I would like to thank Ildar Ibragimov for his collaboration during almost 40 years and for very useful discussions, not only concerning this paper. I also thankful to Sergey Pirogov and Anatolii Puhalskii for their useful remarks.

References

1. G. K. Gobulev, *Nonparametric estimation of smooth densitie s in L_2.* — Problems Inform. Transmission **28**(1), (1992), 44–54.
2. R. Hasminskii, I.Ibragimov, *On the nonparametric estimation of functionals.* — In: Proc. II Prague Symp. Asymptotic Statistics (1978), pp. 41–51.
3. R. Hasminskii, I. Ibragimov, *On density estimation in the view of Kolmogorov's ideas in approximation theory.* — Annals Statist. **18** (1990), 999–1010.
4. I. Ibragimov, R. Khasminskii, *Statistical Estimation: Asymptotic Theory.* Springer Verlag, 1981.
5. I. Ibragimov, A. Nemirovski, R Khasminskii, *Some problems of nonparametric estimation in Gaussian white noise.* — Theory Probab. Appl. **31** (1986), no. 3, 391–406.
6. Yu. Kutoyants, *Statistical Inference for Spatial Poisson Processes.* — Springer, Lect. Notes Statistics **134**, 1998.
7. B. Levit, *On optimality of some statistical estimates.* — In: Proc. Prague Symp. Asymptotic Statistics (1973), pp. 215–238.
8. B. Levit, *On the efficiency of a class of nonparametric estimates.* — Theor. Probab. Applic. **20** (1975), 723–74.

9. A. Nemirovski, *Topics in Nonparametric Statistics*, — In: M. Emery, A. Nemirovski, D. Voiculescu, Lectures on Probability Theory and Statistics, Ecole d'ete de Probabilities de Saint-Flour XXVIII – 1998, Editor: P. Bernard, Lect. Notes Math. **1738**, Springer (2000).
10. S. Ross, *Introduction to Probability Models.* Academic Press, 9th edition (2006).

Mathematics Department
Wayne State University
Detroit, MI 48202, USA

E-mail: rafail@math.wayne.edu

Поступило 5 октября 2007 г.