ON MINIMUM DISTANCE ESTIMATION
FOR SPATIAL POISSON PROCESSES

Yu.A. Kutoyants and F. Liese

Yerevan State University, Department of Mathematics, 375049, Yerevan, Mravian I, Armenia
Wilhelm-Pieck-Universität, Sektion Mathematik, Universitätsplatz 1, Rostock, Deutschland

The present paper is devoted to the problem of parameter estimation of the intensity function Λ_θ of a Poisson process defined on a bounded subset $U \subset R^d$, $d \geq 1$. We consider estimators based on n independent observations of this process. The minimum distance estimator θ_n^* is constructed and its asymptotical properties are described. Traditionally, in problems of this type one investigates the asymptotics of the maximum likelihood estimators $\hat\theta_n$, which are under regularity conditions also consistent, asymptotically normal and have the local asymptotic minimax property (LAM). In that case the LAM property means that the mean risk of the estimator attains the lower bound of Hajek–LeCam (see, for example [1], [2], [5]). Recall that this bound is derived under the condition of local asymptotic normality of the family of measures $\{P_\theta^{(n)}, \theta \in \Theta\}$ induced in the space of all realizations of a Poisson process. For a wide class of loss functions $l(\cdot)$ it has the following form:

$$\lim_{\delta \to 0} \varliminf_{n \to \infty} \inf_{\tilde\theta_n} \sup_{|\theta - \theta_0| < \delta} \mathbf{E}_\theta l\big(\sqrt{n} I(\theta_0)^{-1/2}(\tilde\theta_n - \theta)\big) \geq \mathbf{E}l(\xi),$$

where $I(\theta_0)$ is the Fisher information, $\mathscr{L}\{\xi\} = \mathscr{N}(0, J)$, J standing for the unit matrix, and the infimum is taken over all the estimates $\tilde\theta_n$ based on the first n observations. Note that the supremum in this inequality is taken over all the members $P_\theta^{(n)}$, $\theta \in \{\theta : |\theta - \theta_0| < \delta\}$ of the original family $P_\theta^{(n)}$, $\theta \in \Theta$.

It is also known that the maximum likelihood estimators usually are not robust, so if we take the supremum over a class of distributions, then LAM could be some other estimators. P.W. Millar [6–7] has shown, using the theory of L. LeCam [4], that for some "natural" neighbourhoods $\{P_h^{(n)}, h \in N_c\}$ of the true distribution $P_{\theta_0}^{(n)}$ a similar bound also exists, i.e.,

$$\lim_{c \uparrow \infty} \varliminf_{n \to \infty} \inf_{\tilde\theta_n} \sup_{h \in N_c} \mathbf{E}_h l\big(\sqrt{n}(\tilde\theta_n - \theta_{nh})\big) \geq \mathbf{E}l(\zeta),$$

where $\mathscr{L}\{\zeta\} = \mathscr{N}\big(0, \sigma(\theta_0)\big)$. The optimal estimators in this sense are not the maximum likelihood estimators but the so-called minimum distance estimators.

Our purpose in this article is to apply Millar's result, mentioned above, to the problem of parameter estimation in the case of a spatial Poisson process.

Without loss of generality we may suppose that the observed Poisson process is defined on the unit cube $U = [0,1]^d$. Let $\mathscr{L}_d$ stand for the σ-algebra of Borel sets in U and, respectively, M for the set of all measures φ defined on $(U, \mathscr{L}_d)$ taking values in the set $\{0,1,2,\ldots\}$. Let $\mathscr{M}$ be the σ-algebra of subsets of M induced by the mappings $B \to \varphi(B)$, $B \in \mathscr{L}_d$. Let Λ be a given finite measure defined on $(U, \mathscr{L}_d)$, and $(\Omega, \mathscr{F}, \mathbf{P})$ be the basic probability space. We call a random element $\phi(\omega, \cdot)$ a Poisson process with intensity Λ if $\phi(\omega, \cdot)$ is defined on subsets $B \in \mathscr{L}_d$ and the random variable $\phi(\omega, B)$, $\omega \in \Omega$, takes values in $(M, \mathscr{M})$ and

(A) for any $B \in \mathscr{L}_d$ the random variable $\phi(\omega, B)$, $\omega \in \Omega$, has a Poisson distribution with parameter $\Lambda(B)$;

(B) for any collection of disjoint sets $B_1, \ldots, B_m \in \mathscr{L}_d$, the random variables $\phi(\cdot, B_1), \ldots, \phi(\cdot, B_m)$ are independent.

Let Θ be an open subset in R^k. Suppose for every $\theta \in \Theta$ a measure Λ_θ on $(U, \mathscr{L}_d)$ is defined. Furthermore, suppose that all the measures $\{\Lambda_\theta, \theta \in \Theta\}$ are continuous, i.e., for every $\theta \in \Theta$ and $t \in U$ one has $\Lambda_\theta(\{t\}) = 0$. Every measure Λ_θ corresponds to a Poisson process with intensity Λ_θ. The measure Λ_θ is called a leading measure or intensity because, according to the definition, for any $B \in \mathscr{L}_d$ $\mathbf{E}_\theta \phi(\omega, B) = \Lambda_\theta(B)$.

The main problem: it is necessary to estimate the parameter θ given n independent observations $\phi_1, \ldots, \phi_n$ of a Poisson process with intensity Λ_θ. Let θ_0 be the true value and μ some finite measure on U, $\mu(U) < \infty$. Let $L_2(\mu)$ be the space of all the square integrable functions $f(t)$, $t \in U$, with respect to μ with a norm defined by

$$\|f\| = \left(\int_U f^2(t) \, \mu(dt) \right)^{1/2}.$$

Let us introduce the empiric measure

$$\hat{\Lambda}_n(B) = \frac{1}{n} \sum_{j=1}^{n} \phi_j(B)$$

and the notation

$$L(\theta, t) = \Lambda_\theta([0,t]), \qquad \hat{L}_n(t) = \hat{\Lambda}_n([0,t]), \qquad t > 0.$$

We then define the minimum distance estimator (MDE) θ_n^* through the equality

$$\|L(\theta_n^*) - \hat{L}_n\| = \inf_{\theta \in \Theta} \|L(\theta) - \hat{L}_n\|,$$

and suppose that this equation has a unique solution.

We use the following conditions:

I. Identifiability: for any $\delta > 0$

$$\inf_{|\theta - \theta_0| > \delta} \|L(\theta) - L(\theta_0)\| > 0.$$

D. Differentiability: the function $L(\theta, \cdot)$, $\theta \in \Theta$, is differentiable in θ at the point θ_0 in the Frechet sense in the space $L_2(\mu)$, i.e., there exists such a vector-function $\dot{L}(\theta_0, t) = \big(\dot{L}_1(\theta_0, t), \ldots, \dot{L}_k(\theta_0, t)\big)$ that

$$\big\|\Lambda(\theta) - L(\theta_0) - \big(\dot{L}(\theta_0), (\theta - \theta_0)\big)\big\| = o\big(|\theta - \theta_0|\big).$$

Here $(\cdot, \cdot)$ and $|\cdot|$ stand for the inner product and, respectively, the norm in R^k.

Moreover, suppose the matrix

$$J(\theta) = \big\langle \dot{L}(\theta_0), \dot{L}(\theta_0)' \big\rangle \equiv \int_U \dot{L}(\theta_0, t)\dot{L}(\theta_0, t)' \, \mu(dt)$$

is positive definite.

We denote by B_η the space spanned by the functions $\eta_j(t) = \dot{L}_j(\theta_0, t)$, $t \in U$, and by π the projection operator from $L_2(\mu)$ to B_η. We also use the notation T for the linear operator from R^k to B_η defined by

$$T_t(a) = \sum_{j=1}^{k} a_j \eta_j(t), \qquad a \in R^k.$$

As usually, T^{-1} will stand for the inverse operator which exists and is continuous and linear provided the matrix $J(\theta)$ is nonsingular, i.e., under the assumption just made.

We also define $\xi_n(\theta, t) = \hat{L}_n(t) - L(\theta, t)$, and introduce the a.s. continuous Gaussian field $W(\cdot)$ with the following properties: $W(0) = 0$, $\mathbf{E}_\theta W(t) = 0$, $\mathbf{E}_\theta W(t)W(s) = L(\theta, t \wedge s)$. Here $t \wedge s$ is defined as the coordinate-wise minimum of the vectors t and s.

Theorem 1. *If the conditions* I *and* D *are fulfilled, then*
1) $\xi_n(\theta_n^*) - \xi_n(\theta_0) \Rightarrow \pi \circ W$, *i.e., the weak convergence in* $L_2(\mu)$ *takes place;*
2) $\sqrt{n}(\theta_n^* - \theta_0) \Rightarrow -T^{-1} \circ \pi \circ W$, *i.e., the weak convergence in* R^k *takes place.*

Our proof consists of verifying the conditions (3.2), (3.4) and (3.5) of Theorem 3.6 in Millar [7], which we shall carry out below.

The difference $\xi_n(\theta, t) - \xi_n(\theta_0, t) = L(\theta_0, t) - L(\theta, t)$ is a deterministic function, so the condition (3.2) follows immediately from our condition I. The same argument shows that the condition (3.5) coincides with our condition D.

We denote by $\mathscr{D}$ the Skorohod space of real functions on U, which are left continuous and possess right-hand limits at every point. The notions "left" and "right" used here are defined through the semiordering of the points in the unit cube. Let $\mathscr{D}_d$ denote the σ-algebra of Borel sets in $\mathscr{D}$. The random process

$$\sqrt{n}\xi_n(\theta_0,t) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left[\phi_j([0,t]) - L(\theta_0,t)\right], \qquad t \in U,$$

converges weakly in $(\mathscr{D},\mathscr{D}_d)$ to the Gaussian process $W(t)$, $t \in U$. The proof of this fact is contained, for example, in the proof of Theorem 3 in [3].

From the definition of the operators T^{-1} and π it follows that the estimator θ_n^* is asymptotically normal, i.e.,

$$\mathscr{L}_{\theta_0}\left\{\sqrt{n}(\theta_n^* - \theta_0)\right\} \Rightarrow \mathscr{N}(0,\sigma(\theta_0)),$$

where

$$\sigma(\theta_0) = J(\theta_0)^{-1}A(\theta_0)J(\theta_0)$$

and

$$A(\theta_0) = \int_U\!\!\int_U \dot{L}(\theta_0,t)\dot{L}(\theta_0,t)'L(\theta_0,t\wedge s)\,\mu(dt)\,\mu(ds).$$

To prove the LAM property of the MDE θ_n^* we first introduce a minimax-type lower bound on the risk function of an arbitrary estimator and then show that it is attained by the risk of the MDE.

We first introduce some additional notation:

$$L_2(\Lambda_{\theta_0}) = \left\{h : \int_U h^2(t)\Lambda_{\theta_0}(dt) < \infty\right\},$$

$$\Lambda_n(h,B) = \Lambda_{\theta_0}(B) + \frac{1}{\sqrt{n}}\int_B h(t)\Lambda_{\theta_0}(dt), \qquad B \in \mathscr{L}_d,$$

$$H_0 = \left\{h : h \in L_2(\Lambda_{\theta_0}), \Lambda_n(h,B) \geq 0, B \in \mathscr{L}_d\right\},$$

$$N_c = \left\{h : h \in H_0, \sup\left|h(t)\right| < c\right\},$$

$$\tau h(t) = \int_{[0,t]} h(s)\Lambda_{\theta_0}(ds).$$

In the definition of H_0 it is also possible to suppose that the inequality $\Lambda_n(h,B) \geq 0$ holds for sufficiently large n.

Define $L_n(h,t) = \Lambda_n(h,[0,t])$. We are interested in the risk of an arbitrary estimator when the true distribution is not necesarily contained in the original family $\{\Lambda_\theta, \theta \in \Theta\}$ but in a larger class $\left\{\Lambda_n(h), h \in N_c\right\}$.

Let us also introduce the following notation: Let $P_h^{(n)}$ be the measure corresponding to the sequence $\phi_1,\ldots,\phi_n$ of Poisson processes of intensity $\Lambda_n(h)$ and

let Q_0 be the measure of the process $W(t)$, $t \in U$, in the space $(\mathscr{C}(U), \mathscr{B}(U))$ of continuous functions on U. Introduce the family of measures $\{Q_h, h \in H_0\}$ through $Q_h(B) = Q_0(B - \tau h)$ for any $B \in \mathscr{L}_d$.

The Radon–Nikodym derivatives of this family are given by

$$\frac{dQ_h}{dQ_0}(x) = \exp\left\{I_h(x) - \tfrac{1}{2}\int_U h^2(t)\Lambda_{\theta_0}(dt)\right\},$$

where $I_h(x) = \int_U h(t)\,x(dt)$ is a stochastic integral (see, for example, [3]).

In [3] (Theorem 2) we have proved that the finite dimensional distributions of the process $\{dP_h^{(n)}/dP_0^{(n)}, h \in H_0\}$ converge to the finite dimensional distributions of the process $\{dQ_h/dQ_0, h \in H_0\}$.

This implies the convergence of the experiments $\{P_h^{(n)}, h \in H_0\}$ to the experiment $\{Q_h, h \in H_0\}$, as defined by LeCam [4].

We denote by $\mathbf{B}$ the closure in $L_2(\mu)$ of the set $\{\tau h, h \in H_0\}$ and note that the triplet $(\tau, L_2(\Lambda_{\theta_0}), \mathbf{B})$ defines an abstract Wiener space. We introduce the sequence θ_{nh} by the rule

$$\inf_{\theta \in \Theta} \|L_n(h) - L(\theta)\| = \|L_n(h) - L(\theta_{nh})\|,$$

i.e., $L(\theta_{nh}, \cdot)$ is the closest member, in the $L_2(\mu)$ norm, to the true $L_n(h)$ in the family $\{L(\theta, \cdot), \theta \in \Theta\}$.

Let $g(\cdot)$ be a nonnegative increasing function on $[0, \infty)$. Suppose the mean risk of the estimator $\tilde{\theta}_n$ is defined by the formula

$$\int g\big(\sqrt{n}|\tilde{\theta}_n - \theta_{nh}|\big)\,dP_h^{(n)}.$$

Here $P_h^{(n)}$ stands for the true distribution of the observations.

Our proof of the following theorem practically coincides with the proof of Theorem 5.12 in [7].

Theorem 2. *Suppose that*

(1) $$g(x) \le C\exp\{-\varepsilon_0 x\}$$

for all $x \ge 0$ and $\varepsilon_0 < 3/104$. Then

(2) $$\lim_{c\uparrow\infty} \varliminf_{n\to\infty} \inf_{\tilde{\theta}_n} \sup_{h \in N_c} \int g\big(\sqrt{n}|\tilde{\theta}_n - \theta_{nh}|\big)\,dP_h^{(n)} \ge \mathbf{E}g(\zeta),$$

where $\mathscr{L}\{\zeta\} = \mathscr{N}(0, \sigma(\theta_0))$.

This inequality allows us to define local asymptotic minimax estimators $\hat{\theta}_n$ as the estimators for which one has an equality in (2) for all $\theta_0 \in \Theta$.

It remains only to prove the following

Theorem 3. *The* MDE θ_n^* *is* LAM.

Proof. Following [7] we introduce two processes

$$\xi_n^{(1)}(\theta, t) = \left[L_n(h, t) - L(\theta_0, t)\right]\sqrt{n},$$
$$\xi_n^{(2)}(\theta, t) = \left[\hat{L}_n(t) - L(\theta_0, t)\right]\sqrt{n}.$$

The conditions I and D are fulfilled for each $\xi_n^{(i)}$. Moreover, both of the processes $|\xi_n^{(1)}|$ and $|\xi_n^{(2)}|$ are bounded in probability. Hence, according to Theorem 3.6 in [7] the representations

$$\sqrt{n}(\theta_n^* - \theta_0) = -T^{-1} \circ \pi \circ \xi_n^{(2)} + o(1),$$
$$\sqrt{n}(\theta_{nh} - \theta_0) = -T^{-1} \circ \pi \circ \xi_n^{(1)} + o(1)$$

are valid, implying

$$\sqrt{n}(\theta_n^* - \theta_{nh}) = -T^{-1} \circ \pi \circ (\xi_n^{(2)} - \xi_n^{(1)}) + o(1).$$

It is proved in [3, Theorem 3] that the difference

$$\xi_n^{(2)}(\theta_0, t) - \xi_n^{(1)}(\theta_0, t) = \sqrt{n}(\hat{L}_n(t) - L_n(h, t)), \qquad t \in U,$$

converges weakly in $(\mathscr{D}, \mathscr{D}_d)$ to the process $W(t)$, $t \in U$, uniformly on $h \in N_c$. Hence for any sequence $h_n \in N_c$

$$\sqrt{n}(\theta_n^* - \theta_{nh}) \Rightarrow -T^{-1} \circ \pi \circ W.$$

Thus, the same statement is true also for the sequence h_n^* defined through the equality

$$\sup_{h \in N_c} \int g\left(\sqrt{n}|\theta_n^* - \theta_{nh}|\right) dP_h^{(n)} = \int g\left(\sqrt{n}|\theta_n^* - \theta_{nh_n^*}|\right) dP_{h_n^*}^{(n)}.$$

This gives the LAM property for the MDE θ_n^* for all the bounded uniformly continuous functions $g(\cdot)$. For functions $g(\cdot)$ with bounded growth rate (cf.(1)) the equality in (2) also holds because the sequence of random variables $\{g(\sqrt{n}|\theta_n^* - \theta_{nh_n^*}|)\}$ is uniformly integrable (see [3, Theorem 3]).

References

[1]   KRICKEBERG, K.: Processus ponctuels en statistique. - Lecture Notes in Mathematics 929. Springer-Verlag, Berlin, 1982, 205–313.

[2]   KUTOYANTS, YU.A.: Parameter estimation for stochastic processes. - Heldermann Verlag, Berlin, 1984.

[3]   KUTOYANTS, YU.A., and F. LIESE: Minimax bounds in the problem of intensity estimation of spatial Poisson process. - Soviet J. Contemporary Math. Anal. 26:2, 1991 (English translation of Izv. Akad. Nauk Armyan SSR ser. Mat. 26:2, 1991).

[4]   LECAM, L.: Limits of experiments. - Proceedings of the Sixth Berkeley Symposium of Mathematical Statistics and Probability, University of California Press, Berkeley, 1972, Volume I, 245–261.

[5]   LORZ, U.: Beiträge zur Statistik unbegränzt teilbarer Felder mit unabhängigen Zuwächsen. - Dissertation, Rostock, 1987.

[6]   MILLAR, P.W.: The minimax principle in asymptotic statistical theory. - Lecture Notes in Mathematics 976. Springer-Verlag, Berlin, 1983, 75–265.

[7]   MILLAR, P.W.: A general approach to the optimality of minimum distance estimates. - Trans. Amer. Math. Soc. 286, 1984, 377–418.