Optimal Choice of Observation Window for Poisson Observations

Yury Kutoyants
Laboratoire de Statistique & Processus,
Université du Maine, Le Mans, France
Vladimir Spokoiny
Weierstraß-Institut für Angewandte Analysis und Stochastik,
Berlin, Germany

December 16, 1996

[0] *1994 Mathematics Subject Classification.* 62L05; Secondary 62L12.
Key words and phrases. Poisson process, observation window, sequential design, van Trees' inequality, information bound.

Abstract

We consider the possibility of optimal choice of observation window in the problem of parameter estimation by the observations of an inhomogeneous Poisson process. A minimax lower bound is proposed for the risk of estimation under an arbitrary choice of observation window. Then the adaptive procedure is proposed which is asymptotically efficient in the sense of this bound.

Let $\mathcal{X}$ be a separable metric space, $\mathcal{B}$, the σ-algebra of its Borelian subsets, the set $A \in \mathcal{B}$ and a family of Poisson processes of mean measures Λ_ϑ, $\vartheta \in \Theta, \Theta \subset I\!R$ observed n times on the set A. We suppose that the value of the parameter ϑ is unknown to the observer and he have to estimate it by n independent observations of the Poisson process. If the set A is fixed then under the regularity conditions the maximum likelihood estimator (MLE) $\hat{\vartheta}_n$ is asymptotically normal with the limit variance σ^2 equal to the inverse Fisher information, i.e.,

$$\sqrt{n}(\hat{\vartheta}_n - \vartheta) \Longrightarrow \mathcal{N}(0, \sigma^2),$$
$$\sigma^{-2} = I(\theta) = \int_A \dot{S}(\vartheta, \vartheta, x)^2 \, \Lambda_\vartheta(dx)$$

where dot means the derivation with respect to ϑ and

$$S(\vartheta_1, \vartheta_2, x) = \Lambda_{\vartheta_1}(dx)/\Lambda_{\vartheta_2}(dx), \qquad \dot{S}(\vartheta, \vartheta, x) = \partial S(y, \vartheta, x)/\partial y\Big|_{y=\vartheta}$$

(see [5], Theorem 2.4).

Let us call the set A *an observation window* and consider the problem of its optimal choice. We write $I(\vartheta) = I(\vartheta, A)$ and note that a reasonable solution to this problem is to maximize $I(\vartheta, A)$ over some class of sets $\{A\}$. For instance, one may consider the class $\mathcal{A}$ defined by

$$\mathcal{A}_m = \{A : \; A \subseteq \mathbf{A}, \quad \Lambda(A) \leq m\}$$

where Λ is some measure on $\mathcal{X}$ (it can be one of the measures $\{\Lambda_\vartheta, \vartheta \in \Theta\}$ or, in finite-dimensional case, the Lebesgue measure), $\mathbf{A}$ is some (rather large) set from $\mathcal{B}$ and $m > 0$ is a given number. We see that the information matrix $I(\vartheta, A)$ depends generally on the unknown parameter ϑ and therefore there is no any universal optimal choice of the observation window A^*.

This motivates considering so-called sequential strategies which can adapt to unknown value of parameter.

We follow the standard framework of sequential experimental design due to Chernoff (1959, 1972). Given n first observations $X^{(n)} = (X_1, \dots, X_n)$ of the Poisson process we construct an observation window $A_{n+1} \in \mathcal{A}_m$ for the next observation on the base of $X^{(n)}$. In the other words, if we denote by $\mathcal{F}_j = \sigma(X_1, \dots, X_j)$ the σ-algebra of the events up to time j, then A_j is $\mathcal{F}_{j-1}$-measurable with values in $\mathcal{A}_m$. The family of functions $A^{(n)} = (A_1, \dots, A_n)$ with $A_j : X^{(j-1)} \to \mathcal{A}_m$, $j = 1, \dots, n$, is called *a sequential design* or simply *a design*. Let now $\bar{\vartheta}_n$ be an estimator of ϑ by observations $X^{(n)}$. We call the couple $\left(A^{(n)}, \bar{\vartheta}_n\right)$ *an admissible strategy*. The problem of sequential experimental design is to select both the design $A^{(n)}$ and the estimator $\bar{\vartheta}_n$ in certain optimal way. We consider this problem in the asymptotic set-up when the number of observations n tends to infinity. First we state a lower bound for the risk of an arbitrary admissible strategy $\left(A^{(n)}, \bar{\vartheta}_n\right)_{n \geq 1}$. Then we show that this bound is sharp. Moreover, we present strategies $(A^{*(n)}, \hat{\vartheta}_n)$ which are asymptotically optimal in the sense of this bound. Here $A^{*(n)}$ are special so-called *two-stage* procedure and $\hat{\vartheta}_n$ is the MLE.

The described problem is a particular case of a general problem of sequential experimental design considered in Spokoiny (1992). Here we specify the results for the case of windowed observation of a Poisson process that allows to simplify and instructively verify the general regularity conditions. Also we present a relatively simple and clear proof of the lower bound on the base of van Tries' inequality for the case of squared losses.

Our regularity conditions will be the uniform on $\mathcal{A}_m$ versions of the conditions of Theorems 2.1 and 2.4 in [5].

C1. *For all $\vartheta \in \Theta$ and $A \in \mathcal{A}_m$ we have $\Lambda_\vartheta(A) < \infty$ and all measures $\Lambda_\vartheta, \vartheta \in \Theta$ are equivalent on the set $\mathbf{A}$.*

C2. *The function $S(\vartheta', \vartheta, x) = \Lambda_{\vartheta'}(dx)/\Lambda_\vartheta(dx), x \in \mathbf{A}, \vartheta', \vartheta \in \Theta$ is differentiable with respect to ϑ' for almost all $x \in \mathbf{A}$, the derivative $\dot{S}(\vartheta', \vartheta, x) \in L_2(\Lambda_\vartheta)$. Moreover, for all $\vartheta \in \Theta$*

$$\lim_{\delta_1, \delta_2 \to 0} \sup_{A \in \mathcal{A}_m} \left\| \dot{S}(\vartheta + \delta_1, \vartheta + \delta_2) - \dot{S}(\vartheta, \vartheta) \right\|_A = 0.$$

Here the norm $\|f\|_A$ in $L_2(\Lambda_\vartheta)$ is defined by integration over the set A, i.e.

$$\|f\|_A^2 = \int_A f^2(x)\,\Lambda_\vartheta(dx).$$

The Fisher information $I(\vartheta, A)$ is positive and bounded away from zero and from infinity uniformly on $A \in \mathcal{A}_m$, i.e.,

$$\sup_{A\in\mathcal{A}_m} I(\vartheta, A) = I^*(\vartheta) < \infty,$$
$$\inf_{A\in\mathcal{A}_m} I(\vartheta, A) = I_*(\vartheta) > 0.$$

Now we are about to formulate the main result concerning the lower bound of quadratic risk for an arbitrary sequence of strategies. We take therefore the loss function as $\ell(u) = |u|^2$.

Theorem 1 *Let the conditions* **C1** *through* **C2** *hold. Then for an arbitrary design $A^{(n)}$ and every estimator $\bar\vartheta_n$, the following inequality is fulfilled*

$$\lim_{\delta\to 0} \varliminf_{n\to\infty} \sup_{|\vartheta-\vartheta_0|<\delta} \mathbf{E}_\vartheta \ell\left(\sqrt{n}(\bar\vartheta_n - \vartheta)\right) \geq I^*(\vartheta_0)^{-1}. \tag{1}$$

Proof. This result and even more general results including the case of a multi-dimensional parameter and a loss function of the form $\ell(u) = |u|^p$ for some $p \leq 1$ can be obtained as a specification of the general statement from Spokoiny (1992, Theorem 3.1). However, the proof of this general theorem is rather sophisticated and it requires special rather involved technique. Note also that for the case of univariate parameter the method from Efroimovich (1980) based on random time-scaling can be applied to get the desirable lower bound.

We present below another proof which is clear and straightforward and which makes heavily use of van Trees' inequality. This method can be generalized on the case of multi-dimensional parameter but the quadratic structure of the loss function is essential.

Let $A^{(n)}$ be some sequential design. Then the likelihood ratio can be written as

$$\begin{aligned} L(\vartheta, \vartheta_0, X^{(n)}) &= \exp\left\{\sum_{j=1}^n \int_{A_j} \ln S(\vartheta, \vartheta_0, x)\, X_j(dx)\right. \\ &\qquad \left. - \sum_{j=1}^n \left[\Lambda_\vartheta(A_j) - \Lambda_{\vartheta_0}(A_j)\right]\right\} \end{aligned}$$

where A_j are random $\mathcal{F}_{j-1}$ measurable sets, $j = 1, \ldots, n$. Let $p(u), u \in I\!R^1$ be a continuously differentiable density with support $B = [-1, 1]$, positive on B and vanishing outside B (i.e. $p(u) = 0$ for $|u| = 1$). Introduce also the rescaled density

$$p_n(\vartheta) = H n^{-1/2}\; p(\sqrt{n}(\vartheta - \vartheta_0)/H).$$

By the use of van Trees' inequality [4] we get for each $\delta > 0$ and $n > H^2\delta^{-2}$

$$\begin{aligned}
\sup_{|\vartheta-\vartheta_0|<\delta} \mathbf{E}_\vartheta \left(\bar\vartheta_n - \vartheta\right)^2 &\geq \sup_{\vartheta\in B_n} \mathbf{E}_\vartheta \left(\bar\vartheta_n - \vartheta\right)^2 \\
&\geq \int_{B_n} \mathbf{E}_\vartheta \left(\bar\vartheta_n - \vartheta\right)^2\; p_n(\vartheta)\, d\vartheta \\
&\geq \left\{ \sum_{j=1}^n \int_{B_n} \mathbf{E}_\vartheta \int_{A_j} S(\vartheta,\vartheta,x)^2\, \Lambda_\vartheta(dx)\; p_n(\vartheta)\, d\vartheta + n\, I(p)/H \right\}^{-1} \\
&= \left\{ \sum_{j=1}^n \int_{B_n} \mathbf{E}_\vartheta I(\vartheta, A_j)\; p_n(\vartheta)\, d\vartheta + n\, I(p)/H \right\}^{-1} \qquad (2)
\end{aligned}$$

where $B_n = \{\vartheta : \sqrt{n}|\vartheta - \vartheta_0| < 1\}$ and $I(p)$ is the Fisher information corresponding to the density $p(\cdot)$:

$$I(p) = \int_B \dot p^2(\vartheta)\, p(\vartheta)^{-1}\, d\vartheta.$$

Recall that the sets A_j are random and hence the quantities $I(\vartheta, A_j)$ are also random. However, by **C2**, with probability 1,

$$I(\vartheta, A_j) \leq I^*(\vartheta).$$

Next, condition **C2** easily implies that the function $I^*(\vartheta)$ is continuous in ϑ and hence, for an arbitrary $\varepsilon > 0$ and n large enough, we obtain

$$I(\vartheta, A_j) \leq I^*(\vartheta_0) + \varepsilon.$$

Now we get from (2)

$$\sup_{|\vartheta-\vartheta_0|<\delta} \mathbf{E}_\vartheta \left(\sqrt{n}(\bar\vartheta_n - \vartheta)\right)^2 \geq \left\{ I^*(\vartheta_0) + \varepsilon + I(p)/H^2 \right\}^{-1}.$$

Hence, letting $H \to \infty$ and $\varepsilon \to 0$ we obtain

$$\underline{\lim}_{n\to\infty} \sup_{|\vartheta-\vartheta_0|<\delta} \mathbf{E}_\vartheta \left(\sqrt{n}(\bar\vartheta_n - \vartheta)\right)^2 \geq I^*(\vartheta_0)^{-1}$$

as required.

This bound justifies the following

Definition. Let conditions **C1, C2** be fulfilled. Strategies $\left(A^{(n)}, \bar\vartheta_n\right)_{n\geq 1}$ are *asymptotically efficient* if for all $\vartheta_0 \in \Theta$

$$\lim_{\delta\to 0} \lim_{n\to\infty} \inf_{A^{(n)},\bar\vartheta_n} \sup_{|\vartheta-\vartheta_0|<\delta} \mathbf{E}_\vartheta \ell\left(\sqrt{n}(\bar\vartheta_n - \vartheta)\right) = I^*(\vartheta_0)^{-1}. \qquad (3)$$

To state a result about existence of such strategies we need to strengthen our regularity conditions.

C3. *The functions $S(\vartheta,\vartheta_0,x)$, $l(\vartheta,\vartheta_0,x) = \ln S(\vartheta,\vartheta_0,x), x \in \mathbf{A}$ have two continuous bounded derivatives in ϑ for all $\vartheta_0 \in \Theta$.*

C4. *For any $\nu > 0$ and $\vartheta_0 \in \Theta$*

$$\inf_{A\in\mathcal{A}_m} \inf_{|\vartheta-\vartheta_0|>\nu} \int_A \left(\sqrt{S(\vartheta,\vartheta_0,x)} - 1\right)^2 \Lambda_{\vartheta_0}(dx) > 0$$

Note that the quantity $I^*(\vartheta)$ can be calculated as follows: put

$$A_{(r)} = \left\{x:\ \dot l(\vartheta,\vartheta,x)^2\,\lambda(\vartheta,x) \geq r\right\}$$

and define r as a root of the equation

$$\Lambda(A_{(r)}) = m. \qquad (4)$$

The sequence $A_{(r)}, r\geq 0$ is monotone, i.e., $A_{(r_1)} \subset A_{(r_2)}$ if $r_2 \leq r_1$. Denote this value of r as $r(\vartheta,m)$ and introduce the function

$$F(\vartheta,r) = \Lambda\left(\dot l(\vartheta,\vartheta,x)^2\lambda(\vartheta,x) - r \geq 0\right).$$

We suppose that

C5. *For all $\vartheta \in \bar\Theta$ the sets $A_{(r(\vartheta,m))} \subset \mathbf{A}$, the function $F(\vartheta,r)$ is strictly monotone on r and the Fisher information $I(\vartheta, A_{(r(\vartheta',m))})$ is continuous function of ϑ' at the point ϑ, i.e.,*

$$\lim_{\vartheta'\to\vartheta} I(\vartheta, A_{(r(\vartheta',m))}) = I^*(\vartheta)$$

This condition means that

$$\Lambda\left(x:\dot l(\vartheta,\vartheta,x)^2\lambda(\vartheta,x)=r(\vartheta,mp)\right)=0$$

and this simplifies the proof. If the function $f(\vartheta,x)=\dot l(\vartheta,\vartheta,x)^2\lambda(\vartheta,x)-r(\vartheta,m)$ is constant on some subsets of the set $\{x:\ f(\vartheta,x)=0\}$ of positive measure Λ, then the solution A^* is not unique. In this case we can take any subset $A_{(r)}$ solving (4) as A^* in the integral

$$I^*(\vartheta)=\int_{A^*}\dot l(\vartheta,\vartheta,x)^2\,\Lambda_\vartheta(dx)$$

and this will provide the optimal experiment design (see example with intensity (5) below).

Of course we can not choose A^* as observation window because its definition depends on unknown parameter ϑ. This equality suggests to construct the optimal strategy in two stages. On the first step we estimate by the first $n_1=[\sqrt{n}\,]$ ($[q]$ is an integer part of q) observations $X_1,\ldots,X_{n_1}$ on some set $A\in\mathcal{A}_m$ the value of unknown parameter ϑ using the MLE $\hat\vartheta_{n_1}$ (or any other consistent estimator, say, MDE) and then we define the observation window $\hat A_n^*$ with the help of this estimator as

$$\hat A_n^*=\left\{x:\ |\dot l(\hat\vartheta_{n_1},\hat\vartheta_{n_1},x)|^2\,\lambda(\hat\vartheta_{n_1},x)\geq r(\hat\vartheta_n,m)\quad\right\}.$$

The next $n_2=n-n_1$ observations $X_{n_1+1},\ldots,X_n$ we realize on the set $\hat A_n^*$. We denoted this strategy as $(\hat A_n^*,\hat\vartheta_n)$.

Theorem 2. *Let the conditions* **C1-C5** *be satisfied then the strategy* $(\hat A_n^*,\hat\vartheta_n)$ *is asymptotically optimal in the sense of (3).*

Proof. The estimator $\hat\vartheta_{n_1}$ is consistent by the Theorem 2.10 in [5] and the function $I(\vartheta,A_{(r(\hat\vartheta_{n_1},m))})$ is continuous by condition **C5** hence

$$\mathbf{P}_\vartheta-\lim_{n\to\infty}I(\vartheta,\hat A_n^*)=I^*(\vartheta).$$

For the fixed experiment design $A_j=\hat A_n^*, j=n_1+1,\ldots,n$ and n_2 observations X_{n_1+1},

$\ldots,X_n$ we can study the asymptotic behavior of the maximum likelihood estimator $\hat\vartheta_{n_2}$ through the study of the likelihood ratio $Z_n(u)$ which is

$$\begin{aligned}Z_n(u)\ &=\ \exp\left\{\sum_{j=n_1+1}^{n}\int_{\hat A_n^*}\ln S(\vartheta+u/\sqrt{n},\vartheta,x)\;X_j(dx)-\right.\\ &\qquad\left.-n_2\,\Lambda_{\vartheta+u/\sqrt{n_2}}(\hat A_n^*)+n_2\;;\Lambda_\vartheta(\hat A_n^*)\right\}.\end{aligned}$$

We can check the conditions of Theorem 2.10 and verify that

- *The family of measures* $\{\mathbf{P}_\vartheta^{(n_2)}, \vartheta \in \Theta\}$ *generated by the observations* $X_{n_1+1}, \ldots, X_n$ *is uniformly on compacts* $\mathbf{K} \subset \Theta$ *LAN, i.e.,*

$$\begin{aligned} Z_n(u) &= \exp\left\{u\Delta_{n_2} - \frac{u^2}{2} I(\vartheta, \hat{A}_n^*) + r_n(u)\right\} = \\ &= \exp\left\{u\Delta_{n_2} - \frac{u^2}{2} I^*(\vartheta) + \tilde{r}_n(u)\right\}, \end{aligned}$$

where

$$\Delta_n = \frac{1}{\sqrt{n}} \sum_{j=n_1+1}^{n} \int_{A_n^*} \dot{l}(\vartheta, \vartheta, x)\, [X_j(dx) - \Lambda_\vartheta(dx)]$$

and uniformly on $\vartheta \in \mathbf{K}$

$$\mathcal{L}_\vartheta\left(\Delta_{n_2}\right) \Longrightarrow \mathcal{N}(0, I^*(\vartheta)), \quad \mathbf{P}_\vartheta - \lim_{n_2 \to \infty} \tilde{r}_n(u) = 0.$$

The proof. Note only that $\bar{I}_n(\vartheta) \to I^*(\vartheta)$ and the limit random variable $\Delta(\vartheta)$ is Gaussian $\mathcal{N}(0, I^*(\vartheta))$.

- *For all* $u, v \in U_n = \{u : \vartheta + u/\sqrt{n} \in \Theta\}$

$$\sup_{\vartheta \in \mathbf{K}} \mathbf{E}_\vartheta \left(Z_n^{1/2}(u) - Z_n^{1/2}(v)\right)^2 \leq C\,|u - v|^2.$$

Indeed by Lemma 1.3 [5]

$$\begin{aligned} &\mathbf{E}_\vartheta \left(Z_n^{1/2}(u) - Z_n^{1/2}(v)\right)^2 = \mathbf{E}_\vartheta \left(\mathbf{E}_\vartheta \left(Z_n^{1/2}(u) - Z_n^{1/2}(v)\right)^2 |\mathcal{F}_{n_1}\right) \leq \\ &\leq n_2 \mathbf{E}_\vartheta \int_{A_n^*} \left(\sqrt{S(\vartheta + u/\sqrt{n}, \vartheta, x)} - \sqrt{S(\vartheta + v/\sqrt{n}, \vartheta, x)}\right)^2 \Lambda_\vartheta(dx) \leq \\ &\leq \frac{n_2\,(u-v)^2}{4n} \int_0^1 \mathbf{E}_\vartheta I(\vartheta + u/\sqrt{n} + s(u-v)/\sqrt{n}, A_n^*)\, ds \leq C\,(u-v)^2. \end{aligned}$$

- *For all compacts* $\mathbf{K} \in \Theta$

$$\sup_{\vartheta \in \mathbf{K}} \mathbf{E}_\vartheta Z_n^{1/2}(u) \leq \exp\left\{-\kappa |u|^\mu\right\}$$

with some positive κ *and* μ.

We have by the same Lemma 1.3

$$\mathbf{E}_\vartheta Z_n^{1/2}(u) = \mathbf{E}_\vartheta \left(\mathbf{E}_\vartheta Z_n^{1/2}(u)|\mathcal{F}_{n_1}\right) =$$
$$= \mathbf{E}_\vartheta \exp\left\{\frac{1}{2}\int_{A_n^*}\left(\sqrt{S(\vartheta+u/\sqrt{n},x)}-1\right)^2 \Lambda_\vartheta(dx)\right\}.$$

The integral

$$\int_{A_n^*}\left(\sqrt{S(\vartheta+u/\sqrt{n},x)}-1\right)^2 \Lambda_\vartheta(dx) \geq 2\kappa\,|u|^\nu$$

because the conditions **C** are uniform on $A\in\mathcal{A}_m$.

These properties of the likelihood ratio provide by Theorem 2.10 [5] the uniform asymptotic normality

$$\mathcal{L}_\vartheta\left(\sqrt{n_2}(\hat\vartheta_{n_2}-\vartheta)\right) \Longrightarrow \mathcal{N}(0, I^*(\vartheta)^{-1})$$

and the uniform convergence of moments

$$\lim_{n\to\infty} n^{p/2}\mathbf{E}_\vartheta\left|\hat\vartheta_{n_2}-\vartheta\right|^p = I^*(\vartheta)^{-p/2}\mathbf{E}|\xi|^p.$$

The last equalities provide the asymptotic efficiency (3) of this two-stage strategy.

Examples.

Example 1. Let us consider the periodic Poisson process X of intensity

$$\Lambda_\vartheta([0,t]) = \int_0^t [b+\vartheta\,\sin(\omega x)]^2\,dx \quad 0\leq t\leq T,$$

proposed by J.Tukey. Here b and $\omega>0$ are known parameters and we have to estimate $\vartheta\in(\alpha,\beta)$. The asymptotics corresponds to $T\to\infty$. We introduce the number of periods n on $[0,T]$ as the integer part of $(2\pi)^{-1}T\omega$. Using independence of increments of the Poisson process we can say that this model is equivalent to the n independent observations on the one period $\tau=(2\pi)/\omega$ of the process and $n\to\infty$. Let $m\in(0,\tau)$ and Λ be a Lebesgue measure ($\Lambda(dx)=dx$). The Fisher information

$$I(\vartheta,A) = 4\int_A (\sin(\omega x))^2\;dx$$

does not depend on ϑ and the set

$$A_{(r)} = \{x: \ |\sin(\omega x)| \geq r\} =$$
$$= \left(\frac{1}{\omega}\arcsin(\frac{r}{2}), \pi - \arcsin(\frac{r}{2})\right) \bigcup \left(\frac{1}{\omega}\pi + \arcsin(\frac{r}{2}), 2\pi - \arcsin(\frac{r}{2})\right).$$

Therefore the value $r(\vartheta, m)$ can be chosen as

$$r(\vartheta, m) = 2\sin(\pi - m/2)\omega \equiv r^*$$

In this particular case we need not preliminary estimate ϑ and the strategy $(A_{(r^*)}, \hat{\vartheta}_n)$ will be optimal in the s ens (2), i.e., we observe the process X over the sets

$$\bigcup_{i=0}^{n-1}\left(\frac{(2i+1)\pi - m}{2\omega}, \frac{(2i+1)\pi + m}{2\omega}\right).$$

The conditions **C1 -C3** of course satisfied. To check **C4** we write (supposing for simplicity that $b > \beta$)

$$\int_A \left(\sqrt{S(\vartheta, \vartheta_0, x)} - 1\right)^2 \Lambda_{\vartheta_0}(dx) \geq \frac{(\vartheta - \vartheta_0)^2}{(b+\beta)^2} \int_A [\sin(\omega x)]^2 \ dx \geq \nu^2 \ C(A)$$

and $\inf_{A \in \mathcal{A}_m} C(A) > 0$.

In the case

$$\Lambda_\vartheta(A) = \int_A [\vartheta + b\sin(\omega x)]^2 \ dx \tag{5}$$

the Fisher information is $I(\vartheta, [0,t]) = 4\ t$ and any subset $A_{(r)}$ of Lebesgue measure m can be taken for asymptotically efficient strategy $(A_{(r)}, \hat{\vartheta}_n)$.

The case of *frequency modulation*:

$$\Lambda_\vartheta(A) = \int_A [c + b\sin(\vartheta x)]^2 \ dx$$

is more complicate because the Fisher information is

$$I(\vartheta, A) = b^2 \int_A x^2 \ [\sin(\vartheta x)]^2 \ dx$$

and to introduce $\hat{A}_n^*$ we need at first to estimate ϑ and knowing $\hat{\vartheta}_{n_1}$ we construct $\hat{A}_n^*$ as a sequence of increasing intervals of total Lebesgue measure $m\,n$. This type of modulation is not a particular case of theorems 4.8 and

4.9 because we do not know the period and we can not replace this model by repeated observations of the same process, but the consideration of this section can be generalized to include such situations too (scheme of series).

Example 2. We consider a two-dimensional Poisson process with the intensity function $S(\vartheta, x, y), x, y \in A \subset I\!R^2$ with respect to Lebesgue measure. Suppose that $S(\vartheta, x, y) = S(\vartheta, \varrho),\ \varrho^2 = x^2 + y^2$ and consider at first the linear case: $S(\vartheta, \rho) = \vartheta\, f(\varrho)$ where $f(\varrho)$ is a continuous, positive function. Then (in polar coordinates) the Fisher information

$$I(\vartheta, A) = \int_A \frac{\dot{S}(\vartheta, \varrho)^2}{S(\vartheta, \varrho)}\, \varrho d\varrho\, d\varphi = \vartheta^{-1} \int_A f(\varrho)\varrho\, d\varrho\, d\varphi$$

depends on ϑ, but the optima l set A^* can be chosen without preliminary estimation. The set

$$A_{(r^*)} = \{\varrho, \varphi:\ f(\varrho)\ \geq r\}$$

with

$$\int_{A_{(r^*)}} \varrho d\varrho\, d\varphi = m$$

is a disc or a disc and rings and their choice does not depend on ϑ.

If, say,

$$S(\vartheta, \varrho) = \exp\left\{-(\vartheta - \varrho)^2/2\right\}, \quad \vartheta \in \Theta = (0, \beta)$$

then the MLE $\hat{\vartheta}_{n_1}$ constructed by n_1 observations on some set A (say, $A = \{x, y : x^2 + y^2 \leq m/\pi\}$) is consistent, Fisher information

$$I^*(\vartheta) = \int_{A^*} e^{-(\vartheta - \varrho)^2/2}\, (\vartheta - \varrho)^2\, \varrho\, d\varrho\, d\varphi$$

and the observation window $\hat{A}_n^*$ is defined as follows. Let us denote $x_1(r)$ and $x_2(r)$ two solutions of the equation $2xe^{-x} = r,\ (x_1(r) < x_2(r))$. There are two possibilities. If r satisfies the equality

$$\left(\hat{\vartheta}_{n_1} + \sqrt{x_1(r)}\right)^2 - \left(\hat{\vartheta}_{n_1} + \sqrt{x_2(r)}\right)^2 = m/\pi$$

and $\hat{\vartheta}_{n_1} \leq x_1(r)$ then $\hat{A}_n^*$ is the ring

$$\hat{A}_n^* = \left\{\varrho, \varphi:\ 0 \leq \varphi < 2\pi,\ \hat{\vartheta}_{n_1} + \sqrt{x_1(r)} \leq \varrho \leq \hat{\vartheta}_{n_1} + \sqrt{x_2(r)}\right\}.$$

If $\hat{\vartheta}_{n_1} > x_1(r)$ and

$$\left(\hat{\vartheta}_{n_1} - \sqrt{x_1(r)}\right)^2 + \left(\hat{\vartheta}_{n_1} + \sqrt{x_1(r)}\right)^2 - \left(\hat{\vartheta}_{n_1} + \sqrt{x_2(r)}\right)^2 = m/\pi$$

then $\hat{A}_n^*$ is the disc and the ring

$$\begin{aligned} \hat{A}_n^* \quad &= \quad \left\{\varrho, \varphi :\ 0 \leq \varphi < 2\pi,\ 0 \leq \varrho \leq \hat{\vartheta}_{n_1} - \sqrt{x_1(r)},\right\} \bigcup \\ &\quad \bigcup \left\{\varrho, \varphi :\ 0 \leq \varphi < 2\pi,\ \hat{\vartheta}_{n_1} + \sqrt{x_1(r)} \leq \varrho \leq \hat{\vartheta}_{n_1} + \sqrt{x_2(r)}\right\}. \end{aligned}$$

References

[1] Chernoff, H. (1959) Sequential design of experiment. *Ann. Math. Stat.*, **30**, 735–750.

[2] Chernoff, H. (1972) *Sequential analysis and optimal design.* SIAM, Philadelphia, Paris.

[3] Efroimovich, S. (1980) On sequential estimation under local asymptotical normality condition. *Theory Probab. Appl.*, **25**, 30–43.

[4] Gill, R.D., Levit, B.Y. (1995) *Applications of the van Trees inequality: a Bayessian Cramér-Rao bound.* Bernoulli, 1/2, 59-79.

[5] Kutoyants, Yu.A. (1996) *Statistical Inference for Spatial Poisson Processes* submitted.

[6] Ibragimov, I.A., Khasminskii R.Z.(1981) *Statistical Estimation. Asymptotic Theory.* Springer, N.Y.

[7] Spokoiny, V. (1992) On Asymptotically Optimal Sequential Experimental Design. *Advance in Soviet Mathematics*, **12** (1992), 135–150.