Quickest detection of intensity change for Poisson process in generalized Bayesian setting

Evgeniy Burnaev
Institute for Information Transmission Problems, Russian Academy of Sciences, Bolshoi Karetnyi per. 19, Moscow GSP-4, 127994, Russia.

Keywords: Poisson disorder problem. Quickest detection. Optimal stopping.
AMS: 60G40, 62M20

Abstract

The paper deals with the quickest detection of intensity change for Poisson process. We show that the generalized Bayesian formulation of the quickest detection problem can be reduced to the conditional-extremal optimal stopping problem for a piecewise-deterministic Markov process. For this problem the optimal procedure is described and its characteristics are found.

1. Introduction and problem formulation

This paper deals with the problem of the quickest detection of intensity change for Poisson process. Let $N_{\lambda_0} = (N_{\lambda_0}(t))_{t \geq 0}$ and $N_{\lambda_1} = (N_{\lambda_1}(t))_{t \geq 0}$ be Poisson processes with known intensities $\lambda_0$ and $\lambda_1$ correspondingly defined on a filtered probability space $(\Omega, F, (F_t)_{t \geq 0}, P)$. We observe a stochastic process $X_t = (X_t)_{t \geq 0}$ with representation

$$X_t = \int_0^t I(s < \theta) dN_{\lambda_0}^s + \int_0^t I(s \geq \theta) dN_{\lambda_1}^{s-\theta}, \quad X_0 = 0,$$

where $I(A)$ is an indicator function of some event $A$. We interpret nonrandom unknown parameter $\theta \in [0, \infty]$ as a time when the "disorder" appears (in accordance with the terminology of [1, 2, 3]) during the observation of $X$: for $t < \theta$ the observed process $X_t = N_{\lambda_0}^t$ and for $t \geq \theta$ the observed process $X_t = N_{\lambda_0}^t + N_{\lambda_1}^{t-\theta}$. The appearance of a disorder should be detected as soon as possible trying to avoid false alarms. In the sequel the case $\lambda_1 < \lambda_0$ will be considered.

Let $P_t = \text{Law}(X|t)$ be the distribution of the process $X$ under the assumption that the disorder happened at the deterministic time $\theta = t$. In particular, $P_\infty$ is the distribution of $X$ under the assumption that the disorder never happened, i.e. $P_\infty = \text{Law}(N_{\lambda_0}^t, t \geq 0)$, and $P_0$ is the distribution of the process $N_{\lambda_1}^t$, $t \geq 0$.

Let $\tau = \tau(\omega)$ be a finite stopping time with respect to the filtration $(F_t^X)_{t \geq 0}$ generated by the process $X_t$. We interpret $\tau$ as the decision that the disorder has happened at the time $\tau(\omega)$.
Quickest detection of intensity change

From the point of view of false alarms, the performance of the system controlled by the stopping time $\tau$ can be characterized, for example, by the mean time $\mathbb{E}_\infty \tau$ until the false alarm, where $\mathbb{E}_\theta$ denotes the expectation with respect to the probability $P_\theta$. For every $T > 0$ we denote by $\mathcal{M}_T = \{ \tau \geq 0 : \mathbb{E}_\infty \tau = T \}$ the class of stopping times with the mean time $T$ until the false alarm.

We will estimate the quality of the observation system, identified with the choice of $\tau \in \mathcal{M}_T$, by the quantity $B(T; \tau) = \frac{1}{T} \int_0^\infty \mathbb{E}_\theta (\tau - \theta)^+ d\theta$, where $(x)^+ = \max(0, x)$. The problem is to find a stopping time $\tau^*_T \in \mathcal{M}_T$, if it exists, such that $B(T) = B(T; \tau^*_T) = \inf_{\tau \in \mathcal{M}_T} B(T; \tau)$. Such stopping time $\tau^*_T$ is called optimal in the class $\mathcal{M}_T$. This variant of the quickest detection problem is called generalized Bayesian, since $\theta$ can be interpreted as a generalized random variable with the "uniform" distribution on $[0, \infty)$.

2. Sufficient statistics

Let $L_t = \frac{d(P_0|\mathcal{F}_t^X)}{d(P_\infty|\mathcal{F}_t^X)} = \exp \left( \log \left( \frac{\lambda_1}{\lambda_0} \right) X_t - (\lambda_1 - \lambda_0)t \right)$, $t \geq 0$ be a Radon-Nikodym derivative of $P_0|\mathcal{F}_t^X$ with respect to $P_\infty|\mathcal{F}_t^X$ and a process $\psi = (\psi_t)_{t \geq 0}$ has the form $\psi_t = \int_0^t L_\theta d\theta$. The process $(\psi_t)_{t \geq 0}$ satisfies the equation

$$d\psi_t = dt + \left( \frac{\lambda_1}{\lambda_0} - 1 \right) \psi_t \cdot (X_t - \lambda_0 t), \psi_0 = 0. \quad (1)$$

The following lemma holds (the proof is similar to the proof of the analogous lemma in [1]).

**Lemma 1.** The risk function

$$B(T) = \inf_{\tau \in \mathcal{M}_T} \frac{1}{T} \int_0^\infty \mathbb{E}_\theta (\tau - \theta)^+ d\theta$$

of the generalized Bayesian problem equals to the value of the conditional-extremal optimal stopping problem for the process $(\psi_t)_{t \geq 0}$:

$$B(T) = \inf_{\tau \in \mathcal{M}_T} \frac{1}{T} \mathbb{E}_\infty \int_0^\infty \psi_\theta d\theta.$$ 

This lemma clarifies why the process $(\psi_t)_{t \geq 0}$ is a sufficient statistics.

3. Sample-path properties of the process $(\psi_t)_{t \geq 0}$

Let us consider sample-path properties of the process $(\psi_t)_{t \geq 0}$ which we will use later.

Let $\sigma_1, \sigma_2, \ldots$ be the jumps times of the process $X$ ($\sigma_0 \equiv 0$) and $f(t, z) = \left( z - \frac{1}{\lambda_1 - \lambda_0} \right) \exp(- (\lambda_1 - \lambda_0)t) + \frac{1}{\lambda_1 - \lambda_0}$.
From the equation (1) it follows that
\[
\begin{aligned}
\psi_t &= f(t - \sigma_{n-1}, \psi_{\sigma_{n-1}}), \quad t \in [\sigma_{n-1}, \sigma_n) \\
\psi_{\sigma_n} &= \frac{\lambda_1}{\lambda_0} \psi_{\sigma_n-} 
\end{aligned}
\] , \( n = 1, 2, \ldots \)

First consider the case \( \lambda_1 > \lambda_0 \). It is obvious that for \( \hat{z} = \frac{1}{\lambda_1 - \lambda_0} \) the process \( \{\psi_t\}_{t \geq 0} \) can only jump upward, has a positive drift in \([0, \hat{z})\), a negative drift in \((\hat{z}, \infty)\), and a zero drift at \( \hat{z} \). Thus, if \( \{\psi_t\}_{t \geq 0} \) starts or ends up at \( \hat{z} \), it is trapped there until the first jump of the process \( X \) occurs. At that time \( \{\psi_t\}_{t \geq 0} \) finally leaves \( \hat{z} \) by jumping upward. This also shows that once \( \{\psi_t\}_{t \geq 0} \) leaves \([0, \hat{z})\) it never comes back. Hence for each \( A > 0 \) and \( \tau_A = \inf\{t \geq 0 : \psi_t \geq A\} \) an inequality \( \psi_{\tau_A} > A \) holds with nonzero probability.

Next consider the case \( \lambda_1 < \lambda_0 \). It is obvious that \( \{\psi_t\}_{t \geq 0} \) can only jump towards 0 (at times of the jumps of the process \( X \)). Moreover, the sign of the drift term is always positive. Thus \( \{\psi_t\}_{t \geq 0} \) always moves continuously upward and can only jump towards 0. In this case \( \psi_{\tau_A} = A \) (P.a.s.).

4. Optimal decision

The structure of the optimal stopping time in the class \( \mathcal{M}_T \) is described as follows.

**Theorem 1.** In the case \( \lambda_1 < \lambda_0 \) for each \( T > 0 \) the stopping time \( \tau_T^* = \inf\{t \geq 0 : \psi_t \geq T\} \) is optimal in the class \( \mathcal{M}_T \) for the generalized Bayesian problem.

**Remark 1.** In the case \( \lambda_1 > \lambda_0 \) the stopping time \( \tau_{A(T)}^* = \inf\{t \geq 0 : \psi_t \geq A(T)\} \) is optimal in the class \( \mathcal{M}_T \) for the generalized Bayesian problem if there exists some \( A(T) \) which satisfies an equation \( \mathbb{E}_\infty \tau_{A(T)}^* = T \).

Let us show that \( \tau_T^* \in \mathcal{M}_T \). It follows from the previous section that \( \psi_{\tau_T^*} = T \) (P.a.s.). Under the measure \( \mathbb{P}_\infty \) the stochastic differential equation (1) has the form
\[
d\psi_t = dt + \left( \frac{\lambda_1}{\lambda_0} - 1 \right) \psi_t d \left( N^\lambda_0 - \lambda_0 t \right), \quad \psi_0 = 0.
\]

Since \( \psi_{\min(\tau_T^*)} \leq T \), by the martingale properties of stochastic integrals \( \mathbb{E}_\infty \int_0^{\tau_T^*} \psi_{\tau_T^*} d \left( N^\lambda_0 - \lambda_0 t \right) = 0 \). So, \( T = \mathbb{E}_\infty \psi_{\tau_T^*} = \mathbb{E}_\infty \tau_T^* \).

In order to prove theorem 1 we have to consider a conditional-extremal problem
\[
\inf_{\tau \in \mathcal{M}} \left[ \int_0^\tau \psi_g d\theta - c\tau \right],
\]
where \( c > 0 \) is a Lagrange multiplier and a class of stopping times \( \mathcal{M} = \{ \tau \geq 0 : \mathbb{E}_\infty \tau < \infty \} \). Two important lemmas can be proved on the basis of the technique developed in the section 24 of [2].
Lemma 2. For each \( c > 0 \) there exists \( x^*(c) \geq c \) such that \( \tau^*(c) = \inf\{t \geq 0 : \psi_t \geq x^*(c)\} \) is optimal for the conditional-extremal problem (2).

Lemma 3. For each \( T > 0 \) there exists \( c > 0 \) such that \( x^*(c) = T \).

According to the lemma 3 for each \( T > 0 \) there exists \( c^* = c^*(T) > 0 \) such that \( x^*(c^*) = T \). Hence it follows from the lemma 2 that \( \tau^*(c^*) \in \mathcal{M}_T \). Take any \( \tau \in \mathcal{M} \) with \( \mathbb{E}_\infty \tau \geq T \). By definition of the conditional-extremal problem (2) it follows that \( \mathbb{E}_\infty \int_0^{\tau^*(c^*)} \psi_t d\theta \leq \mathbb{E}_\infty \int_0^\tau \psi_t d\theta + c^* \mathbb{E}_\infty (\tau^*(c^*) - \tau) \). Thus \( \mathbb{E}_\infty \int_0^{\tau^*(c^*)} \psi_t d\theta \leq \mathbb{E}_\infty \int_0^\tau \psi_t d\theta \) for any \( \tau \in \mathcal{M} \) with \( \mathbb{E}_\infty \tau \geq T \). This proves theorem 1.

Remark 2. From the previous considerations it follows that \( \tau^*_T = \inf\{t \geq 0 : \psi_t \geq T\} \) is optimal in a wider class of stopping times \( \mathcal{M}_T = \{\tau \geq 0 : \mathbb{E}_\infty \tau \geq T\} \supset \mathcal{M}_T \).

5. Risk function \( B(T) \) and its asymptotic

The structure of the risk function \( B(T) \) for the generalized Bayesian problem is described in the following theorem.

Theorem 2. For each \( T > 0 \) the risk function \( B(T) = \frac{1}{T} V(0; T) \), where \( V(x; T) \) is a solution of the differential-difference equation

\[
(1 - (\lambda_1 - \lambda_0)x)V'(x; T) + \lambda_0 \left( V \left( \frac{\lambda_1}{\lambda_0} \cdot x; T \right) - V(x; T) \right) = -x, \quad (3)
\]

\[
V(0; T) < \infty, V(T; T) = 0.
\]

Unfortunately it is impossible to find an explicit formula for the solution of the differential-difference equation (3) (see also analogous equation in [2]).

However we estimated an asymptotic of the solution. Let \( \rho = \frac{1}{2} \left( \frac{\lambda_1 - \lambda_0}{\sqrt{\lambda_0}} \right)^2 \),

\[
\beta = \frac{1}{6} \left( \frac{\lambda_1 - \lambda_0}{\sqrt{\lambda_0}} \right)^3, \quad -Ei(-x) = \int_x^{\infty} e^{-t} \cdot \frac{dt}{t} \quad (x > 0)
\]

be the integral exponential function [4], \( C = 0.577 \ldots \) be Euler’s constant. If \( \lambda_1 \to \infty, \lambda_0 \to \infty \) such that \( \lambda_1 = \lambda_0 + \sqrt{\lambda_0} C \) for each fixed constant \( C < 0 \), then with \( \varepsilon = \frac{1}{\sqrt{\lambda_0}} \) the solution of (3) and the risk function \( B(T) \) can be decomposed in powers of \( \varepsilon \), that is \( V(x; T) = V_0(x; T) + \varepsilon V_1(x; T) + \ldots \) and \( B(T) = B_0(T) + \varepsilon B_1(T) + \ldots \), where \( B_0(T) = \frac{1}{T} V_0(0; T), B_1(T) = \frac{1}{T} V_1(0; T) \) and \( V_0(x; T), V_1(x; T) \) are solutions of the differential equations

\[
V_0'(x; T) + \rho x^2 \cdot V_0''(x; T) = -x, \quad V_0(0; T) < \infty, V_0(T; T) = 0, \quad (4)
\]

\[
V_1'(x; T) + \rho x^2 \cdot V_1''(x; T) = -\beta x^2, \quad V_1(0; T) < \infty, V_1(T; T) = 0. \quad (5)
\]

Solving (4) and (5) we obtain

\[
B_0(T) = \frac{1}{\rho} \left[ e^{\frac{\lambda_1}{\rho T}} \left( -Ei \left( -\frac{1}{\rho T}\right) \right) - \left( 1 - \frac{1}{\rho T} \int_0^{\infty} e^{-t} \log \left( 1 + \rho t \right) \frac{dt}{t} \right) \right].
\]
Quickest detection of intensity change

\[ B_1(T) = \frac{\beta}{\rho^2} \left[ \frac{3}{\rho T} \int_0^\infty \frac{e^{-t} \log (1 + \rho T t)}{t} dt + e^{\frac{1}{\rho T}} Ei \left( -\frac{1}{\rho T} \right) \left( \frac{1}{2\rho T} - 1 \right) - \frac{5}{2} \right]. \]

Moreover,

\[ B_0(T) = \begin{cases} \frac{1}{\rho} \left[ \frac{\rho T}{2} + O \left( (\rho T)^2 \right) \right] & \text{for } T \to 0, \\ \frac{1}{\rho} \left[ \log(\rho T) - 1 - C + O \left( \frac{\log^2 \rho T}{\rho T} \right) \right] & \text{for } T \to \infty, \end{cases} \]

\[ B_1(T) = \begin{cases} \frac{\beta}{\rho^2} \left[ \frac{(\rho T)^3}{2} + O \left( (\rho T)^4 \right) \right] & \text{for } T \to 0, \\ \frac{\beta}{\rho^2} \left[ \log(\rho T) - \frac{5}{2} - C + O \left( \frac{\log^2 \rho T}{\rho T} \right) \right] & \text{for } T \to \infty. \]

**Remark 3.** Under our assumptions for \( \lambda_1 \to \infty, \lambda_0 \to \infty \) such that \( \lambda_1 = \lambda_0 + \sqrt{\lambda_0} \cdot C \) for each fixed constant \( C < 0 \) the likelihood

\[ L_t = \frac{d(P_0 | F_t^X)}{d(P_\infty | F_t^X)} \xrightarrow{\text{Law}} \exp \left( \log \left( \frac{\lambda_1}{\lambda_0} \right) X_t - (\lambda_1 - \lambda_0)t \right), \]

where \( X_t = \begin{cases} \sigma B_t, & \text{for } t < \theta, \\ \mu t + \sigma B_t, & \text{for } t \geq \theta, \end{cases} \) \( B_t \) is a Brownian motion and \( \mu = \sigma \sqrt{2\rho}. \) This means that in the limit our problem converge to the analogous quickest detection problem of drift change for Brownian motion in generalized Bayesian setting considered in [1].

**Acknowledgements:** The research was partially supported by the analytical departmental target programme RNP.2.2.1.1.2467, the Program of State Support of Leading Scientific Schools HS-5379.2006.1, the Fundamental Research Program of the RAS Presidium 15, grant RGNF 06-02-91821 a/G, grant RFFI 05-01-00944.

6. Bibliography


