MATHEMATICS

1961

SovietRxiv

Abstract

Full Text

MATHEMATICS

A. N. SHIRYAEV

THE PROBLEM OF THE EARLIEST DETECTION OF A VIOLATION OF A STATIONARY REGIME

(Presented by Academician A. N. Kolmogorov, 21 I 1961)

1. Suppose that, for $t \geqslant 0$, a random process $\eta(t)$ satisfying the stochastic equation

$$d\eta(t) = \chi(t-\theta)\,dt + d\xi(t), \tag{1}$$

is observed continuously, where $\xi(t)$ is a Gaussian process with independent increments, $\xi(0) = 0$,

$$\mathbf{M}\Delta\xi = 0, \qquad \mathbf{M}(\Delta\xi)^2 = \Delta t,$$

and the function χ has the form

$$\chi(s) = \begin{cases} 0, & s \leqslant 0, \\ 1, & s > 0. \end{cases}$$

The moment θ at which the "discharge" appears is unknown.

The purpose of the present note is to give a brief exposition of the solution of the following problem, posed by A. N. Kolmogorov.

It is required to find such a method of observation*, that after the appearance of the discharge, on the basis of observations of the course of $\eta(t)$, a corresponding signal of its presence be given as soon as possible. At the same time, the appearance of erroneous signals given before the moment θ must occur, in some sense, rarely.

2. The problem under consideration is related to the problem of choosing between the hypotheses $\chi = 0$ and $\chi = 1$ on the basis of observations of the course of the function

$$d\eta(t) = \chi\, dt + d\xi(t), \tag{2}$$

considered in § 3 of work (^{2}).

For comparison, let us recall the well-known results pertaining to this latter problem.

Let α and β denote the probabilities of erroneous decisions under the hypotheses $\chi = 0$ and $\chi = 1$; let $\mathbf{M}_i\nu$ be the mathematical expectations of the duration of observations under the hypotheses $\chi = i$. If α and β are fixed, then

$$\mathbf{M}_0\nu \geqslant 2\omega(\alpha, \beta), \qquad \mathbf{M}_1\nu \geqslant 2\omega(\beta, \alpha),$$

where

$$\omega(\alpha, \beta) = (1-\alpha)\ln\frac{1-\alpha}{\beta} + \alpha\ln\frac{\alpha}{1-\beta}.$$

Both inequalities simultaneously turn into equalities for the optimal Wald procedure: the process $\zeta(t)$, $\zeta(0) = 0$, is observed,

$$d\zeta(t) = -\frac{1}{2}\, dt + d\eta(t) \tag{3}$$

until the first occurrence of one of the equalities $\zeta(t) = A$ or $\zeta(t) = B$, where $B > 0 > A$. In the first case the hypothesis $\chi = 0$ is accepted, in the second $\chi = 1$.

The problem formulated in item 1, although in it the number of hypotheses (possible values of the parameter θ) is infinite, admits an analogous solution, not requiring an a priori weighting of hypotheses. Such a formulation of the problem

---

* Concerning the concepts used, "method of observation" and others, see (^{1}).

(hereafter—variant A), but in essence it is applicable only in the case when the onset of the disorder is preceded by a long period of observation of the process $\eta(t)$ (with $\chi(t) = 0$), during which a stationary observation regime is established.

Variant A. Find a method of issuing signals such that, for a given T—the mathematical expectation (m.e.) of the time between two false alarms—the corresponding mean delay time $\tau = \tau(T)$, calculated under the assumption that the disorder appears against the background of an established stationary regime arising when $\chi(t) = 0$, takes the minimum value.

The following two variants (B and C) of the formulation of the problem (whose discrete analogue is set forth in (^{1})) assume that an a priori distribution is adopted for θ

$$\mathbf{P}(\theta < t) = 1 - e^{-\lambda t}, \tag{4}$$

where λ is a known constant.

In other words, it is assumed that $\chi(t)$ is a Markov process with two states 0 and 1, $\chi(0) = 0$, and the only transition $0 \to 1$ has probability density equal to λ.

Variant B. For a given N—the m.e. number of false signals issued before the moment θ—find an observation method for which

$$\tau_\chi = \mathbf{M}\{v_1 + \dots + v_{\chi+1} - \theta\},$$

where the random variables v_i are the durations of the i-th stages of observation, is minimized.

Exactly as in (^{2}), it is shown that the latter problem is equivalent to finding an optimal method in the following formulation.

Variant C. For a given probability $\omega = \mathcal{P}(v < \theta)$, devise an observation method with minimal $\tau(\omega) = \mathbf{M}(v - \theta \mid v \geqslant \theta)$, where v is the moment at which a signal about the disorder is issued.

3. Denote by

$$\pi(t) = \mathcal{P}\{\chi(t) = 1 \mid \eta^t(s)\} \tag{5}$$

the posterior probability of the appearance of the disorder before time t. In (5), as usual, $\eta^t(s)$ denotes the function equal to $\eta(s)$, but defined only for $0 \leqslant s \leqslant t$.

Theorem 1. *The optimal (in the sense of variants B and C) method consists in observing the process $\pi(t)$, $\pi(0) = 0$, until it first reaches a certain value L, which is computed from the condition fixing N or ω. For $\pi(t)$ the following stochastic equation holds:*

$$d\pi(t) = (\lambda - \pi^2)(1 - \pi)\, dt + \pi(1 - \pi)\, d\eta(t). \tag{6}$$

Sometimes, instead of $\pi(t)$, it is convenient to consider

$$\psi(t) = \frac{\pi(t)}{1 - \pi(t)}, \tag{7}$$

$$\varphi(t) = \ln \psi(t), \tag{8}$$

for which from (6) we obtain

$$d\psi(t) = \lambda(1+\psi)\,dt + \psi\,d\eta(t), \tag{9}$$

$$d\varphi(t) = \left[\lambda(1+e^{-\varphi}) - \frac{1}{2}\right] dt + d\eta(t) \tag{10}$$

under the conditions $\psi(0) = 0,\ \varphi(0) = -\infty$.

Naturally, as $\lambda \to 0$, equation (10) passes into the known equation (3) for the logarithm of the likelihood ratio.

4. With the aid of the limiting transition indicated to us by A. N. Kolmogorov, from the preceding theorem one can obtain the optimal method of observation in the formulation of the problem in variant A. Moreover, we shall consider only such observation methods for which $\tau(T)$ depends continuously on T, and the distributions for v_i when $\chi(t) = 0$ are nonlattice and stationary—

regime, which is a consequence of the indecipherability of v_i, is established in such a way that for any T

$$\lim_{t\to\infty} \sup\,[\tau^t(T) - \tau(T)] = 0.$$

Here $\tau^t(T)$ is the mathematical expectation of the delay time under the assumption that the disorder appeared at time t.

For simplicity of the subsequent formulas we shall everywhere assume $M_1\Delta\eta = \sqrt{2}\Delta t$, which can always be achieved by changing the scale of time.

Theorem 2. *Among the observation methods described above, for a given T the optimal method is the one based on observing the random process $\rho(t)$, $\rho(0) = 0$,*

$$d\rho(t) = \frac{1}{T}\,dt + \sqrt{2}\rho\,d\eta(t) \tag{11}$$

until the first attainment of the unit level. In this case

$$\tau(T) = e^{\gamma}(-\operatorname{Ei}(-\gamma)) - 1 + \gamma \int_0^\infty e^{-t}\frac{\ln(1+t/\gamma)}{t}\,dt, \tag{12}$$

where $\gamma = 1/T$,

$$-\operatorname{Ei}(-x) = \int_x^\infty \frac{e^{-t}}{t}\,dt, \qquad x \geqslant 0,$$

is the tabulated exponential integral function.

From (12) we immediately obtain

$$\tau(T) = \begin{cases} \ln T - 1 - C + o(1), & T \to \infty, \\ T/2 + O(T^2), & T \to 0, \end{cases} \tag{13}$$

where $C = 0.577\ldots$ is Euler' s constant.

5. Of great interest is a comparison of the optimal observation method* with the known methods of sequential analysis of Wald and Neyman–Pearson.

As applied to the present problem, the sequential analysis method consists in observing the process $\zeta(t)$, $\zeta(0) = 0$, $d\zeta(t) = \frac{1}{2}dt + d\eta(t)$, with a decision being made at the moment ν of first exit to a boundary. If $\zeta(\nu) = B > 0$, where B is the "upper" boundary, then a decision is made that a disorder is present and this decision is checked. If the disorder actually exists, the observations are terminated. In the absence of a disorder, and also when $\zeta(\nu) = A < 0$, A being the "lower" boundary, the observation process is begun again from the start (reflection of the process from the boundaries to zero occurs).

Since $T = T(A, B)$ and $\tau = \tau(A, B)$, for a given T the method in fact depends on one parameter, say A. Thus the problem reduces to finding the optimal A.

Recall that τ is computed under the assumption of the established stationary regime arising in the absence of a disorder. In the present case one actually finds the one-dimensional stationary distribution for the process $\zeta(t)$ with the boundary conditions described, which is needed in order to compute τ.

Theorem 3. *There exists a $\widetilde{T} < \infty$ such that for all $T \geqslant \widetilde{T}$ the optimal choice of the parameter is $A \equiv 0$. In this case*

$$\tau(T) = \frac{1}{T}\left\{B\left(e^B - \frac{B}{2} - e^{-B}\right) - \frac{3}{2}\left(e^B - 2 + e^{-B}\right)\right\}, \tag{14}$$

where $T = e^B - B - 1$.

From (14) we obtain

$$\tau(T) = \begin{cases} \ln T - \frac{3}{2} + o(1), & T \to \infty, \\ \frac{5}{6}T + O(T^2), & T \to 0. \end{cases} \tag{15}$$

* We note that equation (11) for $\rho(t)$ is a typical example of equations with feedback.

We have no doubt that in fact, for all T, the choice $A \equiv 0$ is optimal. Our method of proof allows us only to establish that $\widetilde{T} \leqslant 3000$. Thus, in the class of Wald' s sequential analysis, the optimal method (at least for $T > \widetilde{T}$) consists in observing the process $\zeta(t)$, $\zeta(0) = 0$, with a reflecting barrier $A = 0$. We shall call this method **degenerate sequential analysis**.

Let us also note that the condition $A = 0$ corresponds to errors $\alpha = 0$, $\beta = 1$.

The Neyman-Pearson method applied to the present problem is as follows. The decision that a disorder is present is made at the moment m, if $\zeta(m) \geq h$, where m and h are certain constants. If $\zeta(m) \geq h$ and it turns out that a disorder is in fact present, then the observations are terminated. If, however, it is not present, and also in the case $\zeta(m) < h$, the observation process begins anew. Again here $T = T(m, h)$ and $\tau = \tau(m, h)$, and for fixed T the observation method is entirely determined by a single parameter, say m. As for the established stationary regime, here this is understood in the following sense: for any fixed m, the distribution of the appearance of a disorder on the interval $[0, m]$, conditional on its appearance on this interval, is uniform. This assumption will not seem strange if one notes that the distribution (4) has this property as $\lambda \to 0$.

In view of the cumbersome nature of the formulas obtained for determining the optimal m, we give only the asymptotic result.

Theorem 4. *As $T \to \infty$, the optimal choice is $m \sim \ln T$, and in this case*

$$\tau(T) \sim \frac{3}{2} \ln T. \tag{16}$$

If $T \to 0$, then the optimal choice is $m \sim T$, and then

$$\tau(T) \sim \frac{1}{2} T. \tag{17}$$

It follows from Theorems 2, 3, and 4 that for large T, degenerate sequential analysis gives a surprisingly good approximation to the optimal method. For small T, however, the Neyman-Pearson method gives a good approximation to the optimal method.

We give the results of numerical computation for these three methods:

T	0.1	1	10	10^2	10^3	10^4
Optimal method	0.04746	0.34153	1.37173	3.16015	5.34728	7.63502
Sequential analysis method	0.06324	0.38892	1.44096	3.25994	5.43759	7.71529

T	0.1	1	10	10^2	10^3	10^4
N.-P. method	0.05	0.44101	1.76845	4.35794	7.73121	11.45836

I express my gratitude to A. N. Kolmogorov for the help rendered to me in working on this problem.

Mathematical Institute named after V. A. Steklov
Academy of Sciences of the USSR

Received
18 I 1961

REFERENCES

1. A. N. Shiryaev, DAN, **138**, No. 4 (1961).
2. A. Dvoretzky, J. Kiefer, J. Wolfowitz, Ann. of Math. Stat., **24**, No. 1 (1953).

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.