



Soviet-era science, translated into English

Mathematics

1961

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196101.32633>

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

Abstract

Full Text

Mathematics

A. N. Shiryaev

Detection of Spontaneously Arising Effects

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

1. For $t \geq 1$ a random process $\xi(t)$ with discrete time ($t = 1, 2, \dots$) is observed. It is assumed that up to some time $\theta > 1$ the quantities $\xi(1), \dots, \xi(\theta-1)$ are independent and identically distributed with distribution function $F_0(x)$. The quantities $\xi(\theta), \xi(\theta+1), \dots, \theta \geq 1$, are also independent (both among themselves and of all preceding quantities) and identically distributed, but already with another distribution function $F_1(x)$. The time θ at which the “disorder” appears is unknown.

It is required to find such a method of observation that, on the basis of the data observed after the appearance of the disorder, a corresponding “signal” of its presence would be given as soon as possible. At the same time it is naturally desirable that erroneous signals, given before the time θ , be avoided as far as possible.

It is assumed that each time after a signal has been given about the presence of a disorder, the correctness of this decision is checked. If it is established that the signal of a disorder was given correctly, then the observations are stopped. Otherwise, the observation process is resumed and is continued, possibly with the giving of erroneous signals, up to the time at which the disorder is detected.

The latter circumstance is a distinctive feature of the present formulation in comparison with the usual problems of testing statistical hypotheses, in which the nature of the distribution does not change during the entire time of observation and the observation process consists of a single stage.

We note that the order of observation described fits into the general scheme of controlled processes proposed by A. N. Kolmogorov.

In what follows, for the parameter θ the prior distribution is assumed to be

$$\mathbf{P}(\theta = t) = (1 - p)^{t-1}p, \quad (1)$$

where p is a constant known to us. This assumption means that, under the supposition of the absence of a disorder up to time $t - 1$, its conditional probability of appearing at time t is equal to p .

Assumption (1), taking into account the distributions F_0 and F_1 , generates, in a known way, a definite probability distribution in the space of numerical sequences $x = \{x_1, x_2, \dots\}$ ⁽¹⁾.

2. Each method of observation is specified by the collection of conditional distributions for the ν -time of giving a signal about the presence of a disorder:

$$\mathbf{P}\{\nu \leq t \mid \xi(s) = x(s)\} = \mathfrak{F}(t \mid x(s)),$$

where the functional \mathfrak{F} is subject to the condition

$$\mathfrak{F}(t \mid x(s)) = \mathfrak{F}(t \mid x^t(s)),$$

meaning that, for fixed t , the value of \mathbf{P} depends on the values of the function $x(s)$ for $s \leq t$ ($x^t(s)$ denotes the function equal to $x(s)$, but defined only for $1 \leq s \leq t$).

In the case where the probability \mathbf{P} takes only two values, 0 and 1, the observation method is called nonrandomized.

The distribution (1) for the parameter θ , together with the conditional distribution for $\xi(t)$ for given θ and the conditional distribution for ν for fixed $\xi(t)^*$, uniquely determines the joint distribution of θ , $\xi(t)$, and ν . Thus the probability

$$\omega = \mathbf{P}(\nu < \theta) \tag{2}$$

of the occurrence of an erroneous signal is determined, as is the conditional mathematical expectation of the delay

$$\tau = \mathbf{M}(\nu - \theta \mid \nu \geq \theta) \tag{3}$$

in the case where the signal is given correctly.

The distribution (1) has the property that the conditional distribution of the time of occurrence of maladjustment, under the condition of its absence up to some random time, is given by the same formula (1). This observation shows that, after carrying out the check and ascertaining the fact that no maladjustment has occurred, we are in the same situation as before the beginning of the observation, with the sole difference that one erroneous signal has been given. We resume observation according to the very same rule by which the first stage of observation was carried out, and independently of the results of the observations obtained at the first stage. We proceed analogously at the subsequent stages, if such are necessary.

Let, further,

$$N = \mathbf{M}\chi$$

be the mathematical expectation of χ , the number of false signals given before the random time θ , and

$$\tau_\chi = \mathbf{M}(\nu_1 + \dots + \nu_{\chi+1} - \theta),$$

where ν_i is the duration of the i -th stage of observation.

It is not difficult to see that, under assumption (1),

$$\tau_\chi = \tau,$$

and therefore we shall omit the index χ .

Our task is to find an optimal observation method for which, for a given N , the mean $\tau = \tau(N)$ attains its minimal value.

In finding optimal methods we immediately exclude the case where the distribution F_0 is singular with respect to F_1 , since then error-free detection is possible.

The following lemma shows that the problem just formulated is equivalent to the problem of finding a method for which, for fixed ω , the corresponding $\tau = \tau(\omega)$ assumes its minimal value.

Lemma. If the parameter θ has distribution (1), then

$$N = \frac{\omega}{1 - \omega}. \quad (4)$$

3. Let $\pi(t) = \mathcal{P}\{\theta \leq t \mid \xi^t(s)\}$ be the posterior distribution for the parameter θ .

Theorem 1. If, for every t , the distribution of the random variable $\pi(t)$ is continuous, then the optimal method is nonrandomized and consists in observing the process $\pi(t)$ up to the first time ν for which $\pi(\nu) \geq \pi_1$, where π_1 is computed from the condition of fixing ω .

* This distribution is assumed, naturally, to be independent of the unknown parameter θ .

Remark. We required the continuity of the distribution of the random variables $\pi(t)$ in order that the prescribed probability ω be attainable. If the continuity condition is abandoned, then the optimal observation method, generally speaking, will be randomized.

The assertion of Theorem 1 is a consequence of the theorem formulated below on the form of the Bayesian solution in the problem of minimizing a certain risk.

In doing this we follow the method for constructing Bayesian solutions set forth (in another setting) in the work of Wald and Wolfowitz ⁽²⁾.

We shall assume that “disorder,” with some probability, occurs even before the beginning of observation:

$$\mathbf{P}(\theta = 0) = \pi, \quad \mathbf{P}(\theta = t \mid \theta > 0) = (1 - p)^{t-1}p, \quad t \geq 1,$$

in connection with which the possibility is also allowed of taking a decision at the moment $\nu = 0$, i.e., without observations.

Let the nonnegative function $W(t, s)$ be such that

$$W(t, s) = \begin{cases} W(t-1, s-1), & t < s, \\ a_1(t-s) + a_2, & t \geq s, \end{cases}$$

where a_i are positive constants and $W(0, s) < \infty$ for $s < \infty$.

Theorem 2. *If the distributions F_0 and F_1 are non-atomic, then the Bayesian solution in the problem of minimizing the risk*

$$\mathfrak{R} = \mathbf{M}W(\nu, \theta) \tag{5}$$

is nonrandomized and consists in observing the process $\pi(t)$ ($\pi(0) = \pi$) up to the first moment ν for which $\pi(\nu) \geq \pi_1$, where π_1 is a certain constant.

Remark 1. In the case when the distributions F_0 and F_1 contain atomic components, one may, as is known, instead of observing the original variables $\xi(t)$ with distributions F_0 or F_1 , carry out observation of certain functions of them $\tilde{\xi}(t)$ with non-atomic distributions \tilde{F}_0 and \tilde{F}_1 , respectively, in such a way that the risk $\tilde{\mathfrak{R}} = \mathfrak{R}$ constructed for them according to (5). From this point of view, in all cases one may regard the distributions F_0 and F_1 as non-atomic from the very beginning and, consequently, the Bayesian solution as nonrandomized.

Remark 2. Theorem 2 remains valid if, instead of (5), one considers the somewhat more general risk

$$\mathfrak{R} = \mathbf{M}W(\nu, \theta) + c_1 \mathbf{M}(\nu - \theta \mid \nu \geq \theta) + c_2 \mathbf{M}(\theta - \nu \mid \nu \leq \theta),$$

where $c_i \geq 0$.

The case of continuous observation will be studied in the next note ⁽³⁾.

The author expresses gratitude to A. N. Kolmogorov for posing the problem and for advice received in the course of its solution.

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

Received
18 I 1961

REFERENCES

1. A. N. Kolmogorov, *Basic Concepts of Probability Theory*, Moscow, 1936.
2. A. Wald, J. Wolfowitz, *Ann. of Math. Stat.*, **21**, 82 (1950).
3. A. N. Shiryaev, *DAN*, **138**, No. 5 (1961).

* In the case when it has been stipulated in advance that the observation time is bounded, the assertion of the theorem also remains in force with the sole change that at each step the comparison threshold (π_1) depends, generally speaking, on the number of the step.

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

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