



Soviet-era science, translated into English

ON THE THEORY OF SEQUENTIAL ESTIMATION

MATHEMATICS

1970

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-197001.18510>

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

Abstract

Full Text

UDC 519.281

MATHEMATICS

Academician Yu. V. LINNIK, I. V. ROMANOVSKII

ON THE THEORY OF SEQUENTIAL ESTIMATION

The present note is adjacent to our preceding note ⁽¹⁾ and uses its terminology; we also rely substantially on the note of I. A. Ibragimov and R. Z. Khas' minskii ⁽²⁾. In the theory of sequential estimation one may distinguish results of an asymptotic character (see, for example, ⁽³⁻⁶⁾) and of an "exact" character (see ^(1,7-9)). We shall indicate here results in both directions.

We shall consider a repeated scalar sample x_1, x_2, \dots from a population belonging to a family of distributions P_θ . We shall assume $\theta \in \Theta$ to be a scalar parameter, specified in some interval Θ , and assume that there exists a density $f(x, \theta)$ with respect to Lebesgue or counting measure. Quantities of the form $\partial^k \ln f(x, \theta) / \partial \theta^k$, $k = 0, 1, 2, \dots$, if they exist, we shall call information quantities; as is known, $-E_\theta(\partial^2 \ln f(x, \theta) / \partial \theta^2)$ is the Fisher information quantity (E_θ denotes expectation for the value of the parameter θ).

In the note of I. A. Ibragimov and R. Z. Khas' minskii ⁽²⁾, under certain conditions imposed on $f(x, \theta)$ (mainly on the information quantities)—conditions which we shall henceforth call conditions I-X on the absence of discontinuities in the information quantities (they concern only quantities with $n \leq 20$)—there is given, in particular, the asymptotic behavior of the variance of Pitman's estimate $\tilde{\theta}_n$ for the parameter θ . Pitman's estimate $\tilde{\theta}_n$ has the form

$$\tilde{\theta}_n = \int \theta p_n(\theta), \quad \text{where } p_n(\theta) = \frac{\prod_{i=1}^n f(x_i, \theta)}{\int \left(\prod_{i=1}^n f(x_i, \theta) \right) d\theta}.$$

In the interesting works ⁽³⁻⁶⁾ on the asymptotic theory of processes of sequential estimation, a Bayesian approach to the question is adopted; in the present note we shall adopt a non-Bayesian approach. We shall consider a sequential estimation plan $S = \{\tau, T_\tau\}$ as a pair consisting of a Markov stopping time τ and a statistic T_τ , estimating a given function $g(\theta)$ unbiasedly or asymptotically unbiasedly. We consider the problem of finding an optimal plan S , minimizing the quantity $E_\theta(T_\tau - g(\theta))^2$ under the condition

$$E_\theta \tau \leq n, \tag{1}$$

where n is a given positive number. A plan S that is optimal for all values $\theta \in \Theta$ will exist very rarely, and we shall consider here only asymptotically optimal plans. The following theorems illustrate the point of view according to which, when conditions I–X on the absence of discontinuities in the information quantities are observed, the use of sequential analysis can give only an infinitely small relative gain in mean-square deviation in comparison with the fixed-sample-size method.

Theorem 1. *Let θ be a shift parameter ($f(x, \theta) = f(x - \theta)$), and suppose that conditions I–X on the absence of discontinuities in the information quantities are fulfilled. Then $\bar{\theta}_{[n]}$ is an unbiased estimate of θ , and for any sequential plan of sequential unbiased estimation we have*

$$D(T_\tau) \geq D(\bar{\theta}_{[n]}) (1 + O(1/n)). \quad (2)$$

Theorem 2. In the general case, under conditions II–X on the absence of discontinuities in the information quantities, the estimate $\bar{\theta}_{[n]}$ is asymptotically unbiased with accuracy $O(1/n)$, and for any such sequential estimation plan $S = \{\tau, T_\tau\}$, under condition (1), we have

$$E(T_\tau - \theta)^2 \geq E(\bar{\theta}_{[n]} - \theta)^2 (1 + O(1/n)). \quad (3)$$

Thus, the plan $\{\tau = [n], \bar{\theta}_{[n]}\}$ of sampling a fixed volume $[n]$ and applying the Pitman estimate $\bar{\theta}_{[n]}$ under the given conditions can be relatively improved by the use of sequential analysis by no more than $O(1/n)$.

In the presence of discontinuities in the information quantities the situation may change sharply, and the use of sequential analysis may considerably improve estimation. An example of such a situation may be found in the note of A. I. Shalyt⁽¹⁰⁾. A broader class of examples was calculated, at the authors' request, by A. I. Shalyt and I. I. Yaura; we shall present them here. We shall consider the case of a shift parameter and distributions having density $f(x - \theta)$ with respect to Lebesgue measure, with support $|x - \theta| \leq 1/2$, continuous there, symmetric, and such that $f(x - \theta) = 1$ for $||x - \theta| - 1/2| \leq \varepsilon_0$, $\varepsilon_0 > 0$. Since the Fisher information quantity has a discontinuity (it assumes two values: 0 and ∞), information inequalities of the Rao–Cramér–Wolfowitz type cannot be applied here. Instead, one must have some “exemplarily good” estimate of θ with which to compare others. The Pitman estimate $\bar{\theta}_{[n]}$ (where n is the right-hand side of (1)) may serve as such an estimate. For $[n] \geq 2$ it is unbiased and optimal, in the sense of variance, as an estimate of θ in the class of all “proper” estimates (proper estimates are statistics $T(x_1, \dots, x_n)$ such that $T(x_1 + c, \dots, x_n + c) = T(x_1, \dots, x_n) + c$ for any c). It follows from^(2,10) that there exists such a plan of unbiased sequential estimation $S = \{\tau, T_\tau\}$ that, as $n \rightarrow \infty$, $D(T_\tau)D(\bar{\theta}_{[n+1]})^{-1} \rightarrow 1/3$. Thus, here the use of sequential analysis

asymptotically reduces the variance by a factor of three in comparison with the “exemplarily good” estimate under sampling of fixed volume.

It is of interest to clarify in detail how the appearance of discontinuities in the information quantities affects the possibilities of improving estimation processes on the basis of the use of sequential analysis.

We shall now consider multinomial urn schemes with k balls ($k \geq 2$) and with replacement or without replacement (the latter case will be called the case of sampling without replacement). Let $p = (p_1, \dots, p_k)$ be the corresponding probability vector. We shall be interested in questions of completeness of the corresponding first-entry plans S for estimating functions of the vector p (for the terminology, see ⁽¹⁾). For the case $k = 2$ (the binomial process), questions of completeness (not yet completely resolved in the case of unrestricted plans) are connected with questions of uniqueness of the representation of polynomials in the form of expressions that we call generalized Bernstein polynomials. Suppose we have a binomial plan S with boundary ∂S ; let $\{(x, y)\}$ be its phase plane, and for $(x, y) \in \partial S$ let $K_{\alpha\beta}(x, y)$ be the number of possible trajectories inside S going from the point (α, β) to (x, y) . Then by a generalized Bernstein polynomial for a continuous function $f(\zeta)$, $\zeta \in [0, 1]$, we shall mean the polynomial

$$B_f^S(\zeta) = \sum_{(x,y) \in \partial S} K_{00}(x,y) f\left(\frac{K_{10}(x,y)}{K_{00}(x,y)}\right) \zeta^x (1-\zeta)^y.$$

Here the argument of $f(\cdot)$ is the unbiased estimate ξ . For the case of a fixed-sample-size plan with boundary ∂S : $x + y = n$, we have the usual Bernstein polynomial.

For the multinomial bounded first-entry plan, when $n \geq 2$ no geometric completeness conditions are known analogous to those found for $n = 2$. We can, however, indicate sufficient geometric conditions for completeness of S .

Theorem 3. *Suppose that for every $x \in \partial S$ there exists a number $i \leq k$ such that all points $y = (y_1, \dots, y_n)$ satisfying the condition*

$$y_i > x_i, \quad y_j \leq x_j \quad (j \neq i), \quad \sum_j y_j = 1 + \sum_j x_j$$

are unattainable. Then the plan S is complete.

It should be noted that for sampling without replacement in the present case no separate completeness problem arises. Namely, the following theorem holds.

Theorem 4. *In order that the plan S be complete under sampling without replacement, it is necessary and sufficient that it be complete under sampling with replacement.*

Let us also note that the asymptotic results (Theorems 1 and 2) for sampling with replacement can be carried over to the case of homogeneous processes with independent increments and continuous time.

Leningrad Branch
of the V. A. Steklov Mathematical Institute
of the Academy of Sciences of the USSR

Received
28 V 1970

CITED LITERATURE

- ¹ R. A. Zaidman, Yu. V. Linnik, I. V. Romanovskii, DAN, **185**, No. 6, 1222 (1969).
- ² I. A. Ibragimov, R. Z. Khas' minskii, DAN, **194**, No. 2 (1970).
- ³ P. I. Bickel, I. A. Yahav, Proc. V Berkeley Symposium on Math. Stat. and Prob., **1**, 1965, p. 401.
- ⁴ P. I. Bickel, I. A. Yahav, Ann. Math. Stat., **39**, No. 2, 442 (1968).
- ⁵ P. I. Bickel, I. A. Yahav, Wahrsch. u. verw. Geb., **11**, 257 (1969).
- ⁶ P. I. Bickel, I. A. Yahav, Ann. Math. Stat., **40**, No. 2, 417 (1969).
- ⁷ De Groot, Ann. Math., **30**, 80 (1959).
- ⁸ S. Trybuła, Rozprawy Matem., Warszawa, 1968.
- ⁹ R. A. Zaidman, Yu. B. Linnik, V. N. Sudakov, Soviet-Japanese Symposium on Probability Theory. Khabarovsk, August 1969, Publ. Acad. Sci. USSR, Novosibirsk, 1969, p. 122.
- ¹⁰ A. I. Shalyt, DAN, **189**, No. 1, 57 (1969).

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

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