Reports of the Academy of Sciences of the USSR

1963

SovietRxiv

Abstract

Full Text

Reports of the Academy of Sciences of the USSR

1963. Volume 150, No. 4

MATHEMATICS

A. M. KAGAN

ON ROBBINS' SCHEME

(Presented by Academician V. I. Smirnov on 7 I 1963)

Let X be a random variable whose distribution density (with respect to some measure) $p(x; a)$ depends in a known way on a parameter a, which is also a random variable with unknown distribution function. A number of problems lead to the need for consistent estimation of $E(a \mid x)$ from independent observations of X according to Robbins' scheme [1−4]. In the present note two theorems are given on the possibility of such estimation.

Theorem 1. *Suppose that the parameter set A is compact and that the following conditions are satisfied:*

1°. *$p(x; a)$ is continuous in x uniformly with respect to $a \in A$.*

2°. *For every function $g(a)$ continuous on A and every $\varepsilon > 0$, there exists a finite set of points $x_1, \dots, x_r$ and constants $c_0, c_1, \dots, c_r$ such that*

$$\left| c_0 + \sum_{i=1}^{r} c_i p(x_i; a) - g(a) \right| < \varepsilon \tag{1}$$

for $a \in A$.

Then, for all x, consistent estimation of $E(a \mid x)$ is possible in Robbins' scheme.

The proof of this theorem is based on results [5] on the estimation of an unknown distribution density.

If X assumes only a finite number of values $x_1, \dots, x_s$;
$\mathbf{P}\{X = x_i; a\} = p_i(a)$, then the situation is as follows.

Theorem 2. *Suppose that:*

1°. *The parameter set A contains some nondegenerate interval Δ.*

2°. *The $p_i(a)$ are continuous on Δ, $i = 1, \dots, s$.*

Then, for at least one i, consistent estimation of $E(a \mid x_i)$ in Robbins' scheme is impossible.

Proof of Theorem 2.

1. Suppose first that the system of functions $\gamma = \{1, p_1(a), \dots, p_s(a)\}$ is such that every proper subsystem γ is linearly independent on Δ. Consider the following system of $s-1$ functions:

$$\begin{gathered} \gamma_1^{(1)} = \{1, p_1(a), \dots, p_{s-1}(a), ap_1(a)\}, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \gamma_{s-1}^{(1)} = \{1, p_1(a), \dots, p_{s-1}(a), ap_{s-1}(a)\}. \end{gathered} \tag{2}$$

Suppose that one of these systems (we may always assume that it is $\gamma_1^{(1)}$) consists of functions linearly independent on Δ. Then, by the known theorem on a sufficient number of functionals ([(6)], p. 136), in the space $C(\Delta)$ of continuous functions on Δ there exists a linear functional Ψ such that

$$\Psi(1) = 0, \qquad \Psi(p_1(a)) = 0, \dots, \qquad \Psi(p_{s-1}(a)) = 0, \qquad \Psi(ap_1(a)) \neq 0. \tag{3}$$

By Riesz' s theorem ((6), p. 203), every linear functional on $C(\Delta)$ is defined by a signed measure B, $\operatorname{Var} B < \infty$:

$$\Psi(p(\alpha)) = \int_\Delta p(\alpha)\, dB(\alpha). \tag{4}$$

Consider the Hahn decomposition of the signed measure B:

$$B = B^+ - B^-,$$

where B^+ and B^- are already ordinary (unnormalized) measures on Δ. In view of (3),

$$B^+(\Delta) = B^-(\Delta); \tag{5}$$

$$\int_\Delta p_i(\alpha)\, dB^+(\alpha) = \int_\Delta p_i(\alpha)\, dB^-(\alpha), \qquad i = 1, \dots, s-1,$$

$$\int_\Delta \alpha p_1(\alpha)\, dB^+(\alpha) \neq \int_\Delta dp_1(\alpha)\, dB^-(\alpha). \tag{6}$$

We may assume that (6) holds for the measures B^+ and B^- with $B^+(\Delta) = B^-(\Delta) = 1$.

Since

$$p_s(\alpha) = 1 - \sum_{i=1}^{s-1} p_i(\alpha),$$

we have

$$\int_\Delta p_s(\alpha)\, dB^+(\alpha) = \int_\Delta p_s(\alpha)\, dB^-(\alpha).$$

Thus, on the parameter set A, two prior distributions B^+ and B^- have been constructed such that the corresponding unconditional distributions of X coincide, but

$$\int_\Delta \alpha p_1(\alpha)\, dB^+(\alpha) \neq \int_\Delta dp_1(\alpha)\, dB^-(\alpha).$$

It is clear that in this case consistent estimation of $E(\alpha \mid x_1)$ in the Robbins scheme is impossible.

2. Suppose now that all the systems $\gamma_1^{(1)}, \dots, \gamma_{s-1}^{(1)}$ are composed of functions linearly dependent on Δ; then identically on Δ we have:

$$\begin{gathered} c_{11}p_1(\alpha) + \cdots + c_{1,s-1}p_{s-1}(\alpha) + c_1\alpha p_1(\alpha) = c_{1,0}, \\ \vdots \\ c_{s-1,1}p_1(\alpha) + \cdots + c_{s-1,s-1}p_{s-1}(\alpha) + c_{s-1}\alpha p_{s-1}(\alpha) = c_{s-1,0}. \end{gathered} \tag{7}$$

We shall regard (7) as a system of equations in $p_1(\alpha), \dots, p_{s-1}(\alpha)$. Its determinant is

$$D(\alpha) = \begin{vmatrix} c_{11} + c_1\alpha & \cdots & c_{1,s-1} \\ \cdots & \cdots & \cdots \\ c_{s-1,1} & \cdots & c_{s-1,s-1} + c_{s-1}\alpha \end{vmatrix}, \tag{8}$$

$$D(\alpha) = c_1 \cdots c_{s-1}\alpha^{s-1} + \cdots \tag{9}$$

If the leading coefficient of $D(\alpha)$ is zero, then for some r,

$$1 \le r \le s-1, \qquad c_{r1}p_1(\alpha) + \cdots + c_{r,s-1}p_{s-1}(\alpha) = c_{r,0}, \tag{10}$$

which is excluded by assumption.

Therefore, for $\alpha \in \Delta$ different from the zeros of $D(\alpha)$,

$$p_i(\alpha) = \frac{D_i(\alpha)}{D(\alpha)}, \qquad i = 1, \dots, s-1. \tag{11}$$

Here the degree of $D(\alpha)$ is equal to $(s-1)$, and the degree of $D_i(\alpha)$ is not greater than $(s-2)$.

Let us consider the systems

$$\gamma_1^{(2)} = \{1, p_1(\alpha), \dots, p_{s-2}(\alpha), p_s(\alpha), \alpha p_1(\alpha)\},$$

$$\dots \tag{12}$$

$$\gamma_{s-1}^{(2)} = \{1, p_1(\alpha), \dots, p_{s-2}(\alpha), p_s(\alpha), \alpha p_s(\alpha)\}.$$

If one of the systems $\gamma_i^{(2)}$, $i = 1, \dots, s-1$, consists of linearly independent functions, then Theorem 2 is proved (see item 1); otherwise, from the system

$$\begin{aligned} d_{11} p_1(\alpha) + \dots + d_{1,s-1} p_s(\alpha) + d_1 \alpha p_1(\alpha) &= d_{1,0}, \\ &\dots \\ d_{s-1,1} p_1(\alpha) + \dots + d_{s-1,s-1} p_s(\alpha) + d_{s-1} \alpha p_s(\alpha) &= d_{s-1,0}, \end{aligned} \tag{13}$$

we shall have, for $\alpha \in \Delta$ that are not zeros of the polynomial

$$\widetilde{D}(\alpha) = \begin{vmatrix} d_{11} + d_1\alpha & \cdots & d_{1,s-1} \\ \cdots & \cdots & \cdots \\ d_{s-1,1} & \cdots & d_{s-1,s-1} + d_{s-1}\alpha \end{vmatrix}, \tag{14}$$

($\widetilde{D}(\alpha)$ has degree $(s-1)$ for the same reason as $D(\alpha)$),

$$p_s(\alpha) = \frac{\widetilde{D}_s(\alpha)}{\widetilde{D}(\alpha)}, \tag{15}$$

where the degree of $\widetilde{D}_s(\alpha)$ is not greater than $(s-2)$.

But from (11) we have

$$p_s(\alpha) = 1 - \frac{\sum_{i=1}^{s-1} D_i(\alpha)}{D(\alpha)} = \frac{D_s(\alpha)}{D(\alpha)}, \tag{16}$$

where the degree of $D_s(\alpha)$ is equal to $(s-1)$.

The contradiction between (15) and (16) proves Theorem 2 under the sole assumption of item 1.

3. Let us now abandon the assumption that every proper subsystem γ consists of linearly independent functions. Obviously, we can always choose from the functions $1, p_1(\alpha), \dots, p_s(\alpha)$ a system of functions

$$\tilde{\gamma} = \{1, p_{i_1}(\alpha), \dots, p_{i_r}(\alpha)\}$$

so that every proper subsystem of the system $\tilde{\gamma}$ consists of linearly independent functions.

Indeed, let $p_{i_1}(\alpha)$ be the first, in order, among the functions $p_1(\alpha), \dots, p_s(\alpha)$ that is linearly independent with 1 (in the event that no such function is found, Theorem 2 is true in a trivial way); let $p_{i_2}(\alpha)$ be the first among the remaining functions that is linearly independent with $\{1, p_{i_1}(\alpha)\}$, and so on up to $p_{i_{r-1}}$.

To the system $\{1, p_{i_1}(\alpha), \dots, p_{i_{r-1}}(\alpha)\}$ we add any one of the remaining functions that is linearly independent of $\{p_{i_1}(\alpha), \dots, p_{i_{r-1}}(\alpha)\}$, and denote it by $p_{i_r}(\alpha)$. Such a function must exist, for otherwise, as is easy to see, the system $\tilde{\gamma}$ cannot consist of linearly independent functions. To this system $\tilde{\gamma}$ all the arguments of items 1–2 are applicable.

Received
2 I 1963

REFERENCES

[1] H. Robbins, Proc. III Berkeley Symp. Math. Stat. Prob., **1**, 1956.
[2] K. Miyasawa, Bull. Inst. Intern. Statistique, **38** (1961).
[3] J. Neyman, *Two Breakthroughs in the Theory of Statistical Decision Making*, Univ. California Preprint, 1961.
[4] A. M. Kagan, DAN, **147**, No. 5 (1962).
[5] E. Perzen, Ann. Math. Statistics, **33**, 3 (1962).
[6] L. V. Kantorovich, G. P. Akilov, *Functional Analysis in Normed Spaces*, 1959.

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