



Soviet-era science, translated into English

Reports of the Academy of Sciences of the USSR

1963

SovietRxiv

View the original and related papers at <https://soviextrxiv.org/items/ru-196301.52486>

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

Abstract

Full Text

Reports of the Academy of Sciences of the USSR

1963. Volume 153, No. 3

MATHEMATICS

A. M. KAGAN

FAMILIES OF DISTRIBUTIONS AND SEPARATING PARTITIONS

(Presented by Academician A. N. Kolmogorov on 20 VI 1963)

The maximum-likelihood estimates introduced by Fisher possess a number of attractive properties, and the theory of these estimates has already become a classical chapter of mathematical statistics. However, the effective computation of maximum-likelihood estimates is almost always associated with great difficulties.

The separating partitions proposed in the present paper make it possible to obtain simple asymptotically efficient estimates for a broad class of families of distributions.

Let a family of probability measures $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$ be given on the space $\{\mathfrak{X}, \mathfrak{A}\}$; we shall assume that $P_\theta = P_{\theta'}$ implies $\theta = \theta'$. A partition of the space \mathfrak{X} into disjoint sets A_1, \dots, A_s , $A_1 \cup \dots \cup A_s = \mathfrak{X}$, will be called **separating the family** \mathcal{P} if from $P_\theta(A_i) = P_{\theta'}(A_i)$ for $i = 1, \dots, s$ it follows that $\theta = \theta'$.

We shall denote by $r(\mathcal{P})$ the minimal cardinality of a finite partition separating the family \mathcal{P} ; if no finite separating partition exists, we put $r(\mathcal{P}) = \infty$.

Separating partitions define simple consistent estimates of the parameter θ . Indeed, if the partition into the sets A_1, \dots, A_s is separating, $\Theta \subset R^1$, and the family \mathcal{P} satisfies certain natural smoothness requirements with respect to the parameter, then $\theta = \varphi(P_\theta(A_1), \dots, P_\theta(A_s))$, and the function $\varphi(x_1, \dots, x_s)$ is continuous in (x_1, \dots, x_s) . Now let X_1, \dots, X_N, \dots be independent random variables with values in \mathfrak{X} , identically distributed according to the law P_θ . Put

$$n_i(N) = \sum_{\substack{k: 1 \leq k \leq N \\ X_k \in A_i}} 1.$$

It is known that $(n_1(N)/N, \dots, n_s(N)/N)$ is a consistent estimate of the vector

$(P_\theta(A_1), \dots, P_\theta(A_s))$. Therefore

$$\tilde{\theta}_N(X_1, \dots, X_N) = \varphi(n_1(N)/N, \dots, n_s(N)/N)$$

will be a consistent estimate of θ . The loss of efficiency in the estimation problem that occurs under such an approach can be compensated by taking into account the corrections of Le Cam ⁽¹⁾. As a result we have a method for constructing best asymptotically normal estimates (BAN-estimates) ⁽²⁾ of the parameter of a family admitting a separating partition.

From the computational point of view, the indicated method of constructing BAN-estimates is in a number of cases preferable to the maximum-likelihood method, to various generalized minimum- χ^2 methods, etc. ⁽³⁻⁸⁾.

The study of families of distributions admitting separating partitions, and of the separating partitions themselves, is evidently also of independent interest.

In what follows we shall indicate various classes of families of distributions for which finite separating partitions exist.

1. Let the family \mathcal{P} be finite, $\mathcal{P} = \{P_1, \dots, P_q\}$. For each pair P_i, P_j there is a set $A_{ij} \in \mathfrak{A}$ such that $P_i(A_{ij}) \neq P_j(A_{ij})$. Consider the smallest algebra containing the sets A_{ij} for $1 \leq i < j \leq q$.

Obviously, the cardinality of this algebra does not exceed the absolute constant $B(q)$.

Thus, for a family \mathcal{P} consisting of q measures,

$$r(\mathcal{P}) \leq B(q).$$

- II. Let \mathcal{P} be a family of distributions on R^1 whose densities with respect to Lebesgue measure are of the form

$$p(x; \theta) = p(x - \theta); \quad \theta \in \Theta \subset R^1.$$

Suppose that for any interval $\Delta \subset R^1$

$$\int_{\Delta} p(x) dx > 0.$$

For such a family $r(\mathcal{P}) = 2$. The same is also true for families of distributions on R^1 whose densities have the form

$$p(x; \sigma) = \frac{1}{\sigma} p\left(\frac{x}{\sigma}\right).$$

- III. We now consider the family \mathcal{P} of distributions $P_{\theta, \sigma}$ on R^1 , whose densities with respect to Lebesgue measure have the form

$$p(x; \theta, \sigma) = \frac{1}{\sigma} p\left(\frac{x - \theta}{\sigma}\right),$$

where $\theta \in \Theta$, $\sigma \in \Sigma$, $\sigma > 0$. Suppose that Θ and Σ contain nondegenerate intervals Θ_1 and Σ_1 , that $P_{\theta,\sigma}(A)$ is continuous on $\Theta_1 \times \Sigma_1$ for all $A \in \mathfrak{A}$, and that for any interval Δ

$$\int_{\Delta} p(x) dx > 0.$$

Then it follows from simple topological considerations that $r(\mathcal{P}) \geq 3$. Consider the partition (A_1, A_2, A_3) , where $A_1 = (-\infty, -1)$, $A_2 = (-1, 0)$, $A_3 = (0, +\infty)$. We have

$$P_{\theta,\sigma}(A_2) = \int_{-1/\sigma - \theta/\sigma}^{-\theta/\sigma} p(x) dx; \quad (1)$$

$$P_{\theta,\sigma}(A_3) = \int_{-\theta/\sigma}^{+\infty} p(x) dx. \quad (2)$$

Take two distinct parameter points (θ_1, σ_1) and (θ_2, σ_2) . If $\theta_1/\sigma_1 \neq \theta_2/\sigma_2$, then P_{θ_1, σ_1} and P_{θ_2, σ_2} are distinct, in view of (2), on the set A_3 ; but if $\theta_1/\sigma_1 = \theta_2/\sigma_2$, then $\sigma_1 \neq \sigma_2$, and P_{θ_1, σ_1} and P_{θ_2, σ_2} are distinct on the set A_2 , by (1). Thus, for family III, $r(\mathcal{P}) = 3$.

IV. Let us pass to the Koopman-Dynkin family \mathcal{P} of distributions on R^1 , specified with respect to Lebesgue measure by densities

$$p(x; \theta) = \exp\{\varphi_0(x) + \varphi_1(x)\psi_1(\theta) + \dots + \varphi_N(x)\psi_N(\theta) + \psi_{N+1}(\theta)\}.$$

We shall assume the functions $\psi_i(x)$, $i = 1, \dots, N$, to be continuous; the parameter $\theta \in \Theta$ is of arbitrary nature.

Undoubtedly, for distribution families IV, $r(\mathcal{P}) < \infty$, although we can prove this only for $N = 2^*$. For the proof we need the additional assumption of mutual absolute continuity of the measures P_{θ} . Then

$$\begin{aligned} \frac{dP_{\theta}}{dP_{\theta_0}} = \tilde{p}(x; \theta) = & \exp\{\varphi_1(x)[\psi_1(\theta) - \psi_1(\theta_0)] + \varphi_2(x)[\psi_2(x) - \psi_2(\theta_0)] + \\ & + [\psi_3(\theta) - \psi_3(\theta_0)]\}. \end{aligned}$$

* *Note added in proof.* Recently S. M. Vishik and A. A. Rozental proved that, for arbitrary N , $r(\mathcal{P}) \leq N\%2$.

Take some finite interval Δ on which $\varphi_1(x)$ is not identically constant (such an interval necessarily exists, since the functions $1, \varphi_0(x), \varphi_1(x), \varphi_2(x)$ may always be assumed linearly independent). Let

$$\int_{\Delta} \varphi_1(x) dx = c_1; \quad \tilde{\varphi}_1(x) = \varphi_1(x) - c_1 \quad \text{on } \Delta; \quad \tilde{\varphi}_1(x) = \tilde{\varphi}_1^+(x) + \tilde{\varphi}_1^-(x).$$

Put $\tilde{\varphi}_2(x) = \varphi_2(x) - d_1 - d_2\tilde{\varphi}_1(x)$, and choose the constants d_1 and d_2 from the conditions

$$\int_{\Delta} \tilde{\varphi}_2(x)\tilde{\varphi}_1^+(x) dx = 0; \quad \int_{\Delta} \tilde{\varphi}_2(x)\tilde{\varphi}_1^-(x) dx = 0.$$

The constants d_1 and d_2 will then be determined from a system of linear equations whose determinant is positive.

Now write $\tilde{p}(x; \theta)$ in the form

$$\tilde{p}(x, \theta) = \exp\{\tilde{\varphi}_1(x)\tilde{\psi}_1(\theta) + \tilde{\varphi}_2(x)\tilde{\psi}_2(\theta) + \tilde{\psi}_3(\theta)\},$$

where

$$\begin{aligned} \tilde{\psi}_1(\theta) &= \psi_1(\theta) - \psi_1(\theta_0) + d_2[\psi_2(\theta) - \psi_2(\theta_0)], \\ \tilde{\psi}_2(\theta) &= \psi_2(\theta) - \psi_2(\theta_0), \\ \tilde{\psi}_3(\theta) &= \psi_3(\theta) - \psi_3(\theta_0) + c_1[\psi_1(\theta) - \psi_1(\theta_0)] + d_1[\psi_2(\theta) - \psi_2(\theta_0)]. \end{aligned}$$

Then consider the partition formed by the sets

$$\begin{aligned} A_1 &= \{x \in \Delta : \tilde{\varphi}_1(x) \geq 0, \tilde{\varphi}_2(x) \geq 0\}, \\ A_2 &= \{x \in \Delta : \tilde{\varphi}_1(x) \geq 0, \tilde{\varphi}_2(x) < 0\}, \\ A_3 &= \{x \in \Delta : \tilde{\varphi}_1(x) < 0, \tilde{\varphi}_2(x) \geq 0\}, \\ A_4 &= \{x \in \Delta : \tilde{\varphi}_1(x) < 0, \tilde{\varphi}_2(x) < 0\}, \\ A_5 &= \overline{\Delta}, \end{aligned}$$

and show that it is separating.

Each of the sets A_1, A_2, A_3, A_4 has positive Lebesgue measure. Let θ_1 and θ_2 be two distinct parameter points and, for definiteness, let $\tilde{\psi}_3(\theta_1) \geq \tilde{\psi}_3(\theta_2)$. As is easy to see, on at least one of the sets A_1, A_2, A_3, A_4 one always has $\tilde{p}(x; \theta_1)/\tilde{p}(x; \theta_2) > 1$. For example, if $\tilde{\psi}_1(\theta_1) > \tilde{\psi}_1(\theta_2)$, $\tilde{\psi}_2(\theta_1) \geq \tilde{\psi}_2(\theta_2)$, then such a set is A_3 . Thus, for families of distributions IV with $N = 2$, $r(\mathfrak{P}) \leq 5$.

In conclusion we give one example.

Consider the family \mathfrak{P} of uniform distributions on $(0, \theta)$, with the right endpoint θ as the family parameter; $0 < \theta < 1$. We shall show that for this family $r(\mathfrak{P}) > 2$.

Suppose that there exists a set A such that the function

$$\Psi(\theta) = \frac{1}{\theta} \int_0^\theta \chi_A dx$$

is different for different θ ; here χ_A is the characteristic function of the set A . Obviously, then $0 < \text{mes } A < 1$. By the well-known theorem of Luzin, at almost all points of A ,

$$\frac{d}{d\theta} \int_0^\theta \chi_A dx = 1$$

and at almost all points of \bar{A}

$$\frac{d}{d\theta} \int_0^\theta \chi_A dx = 0;$$

$$\frac{d}{d\theta} \Psi(\theta) = \begin{cases} -\frac{1}{\theta^2} \int_0^\theta \chi_A dx + \frac{1}{\theta} & \text{at almost all points of } A, \\ -\frac{1}{\theta^2} \int_0^\theta \chi_A dx & \text{at almost all points of } \bar{A}. \end{cases}$$

But

$$-\frac{1}{\theta^2} \int_0^\theta \chi_A dx < 0, \quad -\frac{1}{\theta^2} \int_0^\theta \chi_A dx + \frac{1}{\theta} > 0,$$

and therefore $\Psi(\theta)$ cannot be a one-to-one function of θ .

Remark. It is interesting to investigate the connection between separating partitions and sufficient statistics (^{11,12}).

Let the family \mathfrak{P} have a nontrivial sufficient subalgebra $\mathfrak{B} \subset \mathfrak{A}$ and admit a separating partition. It is unknown whether in such a case there always exists a separating partition formed by sets of the sufficient algebra \mathfrak{B} .

Received
20 VI 1963

REFERENCES

- ¹ L. Le Cam, Proc. III Berkeley Symp. on Probability and Math. Statistics, 1, 1956.
- ² J. Neyman, Proc. Berkeley Symp. on Math. Statistics and Probability, 1949.
- ³ G. Kramer, *Mathematical Methods of Statistics*, IL, 1949.
- ⁴ E. Barankin.
- ⁵ J. Gurland, Univ. California Publ. Statistics, 1, 89 (1950).
- ⁵ W. Taylor, Ann. Math. Statistics, 24, 1 (1953).
- ⁶ C. Chiang, Ann. Math. Statistics, 27, 2 (1956).

- ⁷ Th. Ferguson, Ann. Math. Statistics, 29, 4 (1958).
⁸ R. Wijsman, Ann. Math. Statistics, 30, 1 (1959).
⁹ W. Koopman, Trans. Am. Math. Soc., 39, 399 (1936).
¹⁰ E. B. Dynkin, UMN, 6, issue 1 (1951).
¹¹ P. Haimos, L. Savage, Ann. Math. Statistics, 20, 1 (1949).
¹² R. Bahadur, Ann. Math. Statistics, 25, 3 (1954).

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

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