
Soviet-era science, translated into English

ON NONTRIVIAL SUFFICIENT STATISTICS

N. M. Mitrofanova

1967

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196701.18772>

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

Abstract

Full Text

UDC 519.281

MATHEMATICS

N. M. MITROFANOVA

ON NONTRIVIAL SUFFICIENT STATISTICS FOR OBSERVATIONS CONNECTED IN A MARKOV CHAIN

(Presented by Academician Yu. V. Linnik on 3 X 1966)

In papers ⁽¹⁻³⁾ the question was clarified of the functional form of a family of distributions $\gamma = \{P(x, \theta)\}$, $\theta \in \Omega$, admitting nontrivial sufficient statistics for the parameter θ from an independent sample. The analytic content of the main result consists in the fact that, under certain smoothness conditions on the family, from the factorization identity

$$\prod_{i=1}^n p(x_i, \theta) = p(x_1, \dots, x_n) g(\theta, \chi_1, (x_1, \dots, x_n), \dots, \chi_r(x_1, \dots, x_n)) \quad (1)$$

for any $n \geq r$ and under functional independence of the system of functions $\chi_1, \dots, \chi_r$, it follows that the family γ is exponential of rank r . In papers ⁽⁴⁻⁶⁾ various cases were considered in which the sequence of observations is connected in a Markov chain with a finite number of states. The present note is devoted to the case where the set of states is continuous and a parametric family of transition densities $\chi = \{f(x, y, \theta)\}$, $\theta \in \Omega$, is considered, where Ω is some parametric set. The factorization identity (1) is replaced in this case by the identity

$$\prod_{i=0}^n f(x_i, x_{i+1}, \theta) = \bar{f}(x_0, \dots, x_n) g(\theta, \chi_1, \dots, \chi_r)$$

for any $n \geq r$. It turns out that, when the sequence of observations is connected in a Markov chain, the presence of nontrivial sufficient statistics is not equivalent to the exponential form of the family of transition densities. In the note the general form is indicated of the family of transition densities of a sequence of observations connected in a simple Markov chain and admitting nontrivial sufficient statistics.

We use the method of describing sufficient statistics developed in ⁽²⁾, as well as the terminology and notation of that paper.

Let us pass to the precise exposition.

1°. Suppose there is a sequence $x_0, x_1, \dots, x_n$ of random variables connected in a simple Markov chain, and to each element θ of some parametric set Ω there is assigned a conditional distribution

$$P_\theta(x, A) = P_\theta(x_{k+1} \in A \mid x_k = x),$$

where $x \in E^1$, and A is a measurable subset of E^1 (E^s is s -dimensional Euclidean space). We shall denote the family of distributions $P_\theta(x, A)$ by γ .

Let Δ^s be a cube in E^s with side Δ .

Definition 1. We shall call the family γ **regular in Δ** if each distribution $P_\theta(x, A)$ is given by a transition density $f(x, y, \theta)$, which is a positive and continuous function of x and y in Δ^2 (A is a measurable subset of Δ),

$$P_\theta(x, A) = P_\theta(x_{k+1} \in A \mid x_k = x) = \int_A f(x, y, \theta) dy.$$

Definition 2. We shall call the family γ **smooth** in the domain G if, for all θ in Ω , for each conditional density $f(x, y, \theta)$ there exist continuous derivatives f'_x, f'_y, f''_{xy} in G .

The set of possible outcomes of observations over the vector $X = (x_0, x_1, \dots, x_n)$ is Δ^{n+1} , and the family γ induces in Δ^{n+1} the family of distributions γ^{n+1} . Let $q(x, \theta)$ be the initial density. Then the distributions of the family γ^{n+1} have densities

$$f_\theta^{(n+1)}(x_0, \dots, x_n) = q(x_0, \theta) f(x_0, x_1, \theta) \dots f(x_{n-1}, x_n, \theta).$$

Definition 3. A function $\chi(x_0, \dots, x_n)$, defined in the domain Δ^{n+1} , is called a **sufficient statistic** for the family γ^{n+1} if the density $f_\theta^{(n+1)}$ can be represented in the form

$$f_\theta^{(n+1)} = \bar{f}^{(n+1)}(\chi(x_0, \dots, x_n), \theta) g^{(n+1)}(x_0, \dots, x_n),$$

where $x_i \in \Delta$, $\theta \in \Omega$.

2°. Consider a family of transition densities of the form $\gamma = \{f(x, y, \theta)\}$. Denote by $L(\gamma, \Delta)$ the minimal linear space of functions defined on Δ^2 and containing the functions

$$g_{xy}(\theta) = \ln f(x, y, \theta) - \ln f(x, y, \theta_0)$$

for all $\theta \in \Omega$, where θ_0 is an arbitrary fixed element of Ω , and all constants. Let the dimension of $L(\gamma, \Delta)$ be equal to $r + 1$ (possibly $r = \infty$).

Lemma 1. If $\varphi(x, y) \in L(\gamma, \Delta)$, then

$$\chi(x_0, \dots, x_n) = \sum_{k=0}^{n-1} \varphi(x_k, x_{k+1})$$

is a necessary statistic.

Lemma 2. If the functions $1, \varphi_1, \dots, \varphi_s$ form a basis of $L(\gamma, \Delta)$, then the system of functions

$$\chi_0 = x_0, \quad \chi_i = \sum_{k=0}^{n-1} \varphi_i(x_k, x_{k+1}), \quad i = 1, \dots, s, \quad (2)$$

forms a sufficient statistic for a sample of size $n \geq r$.

These lemmas are proved analogously to the corresponding assertions from (2).

Lemma 3. If the system of functions $1, \varphi_1(x, y), \dots, \varphi_s(x, y)$ is linearly independent and the functions φ_i , $i = 1, \dots, s$, are twice continuously differentiable, then the system of functions (2) is functionally dependent if and only if the functions φ_i have the form

$$\varphi_i(x, y) = P_i(x) - P_i(y) + \psi_i(x, y) + C_i \quad (3)$$

and any $s - 1$ functions from the set $\psi_1, \dots, \psi_s$ are linearly dependent with the unit function.

3°. Lemma 3 makes it possible to describe the class of transition densities admitting nontrivial sufficient statistics.

If the dimension of the space $L(\gamma, \Delta)$ is finite, then nontrivial sufficient statistics always exist, and their number does not exceed r . If $r = \infty$ and the basis of the space $L(\gamma, \Delta)$ is composed of functions of the form (3), then the transition density $f(x, y, \theta)$ admits the representation

$$f(x, y, \theta) = \exp \left\{ P(y, \theta) - P(x, \theta) + \sum_1^r c_i(\theta) \psi_i(x, y) + C_0(\theta) + \psi_0(x, y) \right\}, \quad (4)$$

where

$$P(x, \theta) = \sum_1^\infty c_i(\theta) P_i(x).$$

It follows from this:

Theorem. If the family γ of transition densities is smooth and regular, its generating densities have the form (4), and the systems of functions $(c_1 \dots c_r)$, $(\psi_1 \dots \psi_r)$ are linearly independent, then for $n \geq r+2$ it admits $r+2$ functionally independent sufficient statistics equivalent to the vector

$$\left(x_0; \sum_{j=0}^n \psi_i(x, x_{j+1}), i = 1, \dots, r; x_n \right).$$

Any sufficient statistic for a sample of size $n < r+2$ is trivial.

Conversely: if the family γ admits nontrivial sufficient statistics, then its generating functions have the form (4).

Remark 1. If the family γ admits only one nontrivial sufficient statistic, then it is an exponential family of rank 1, since among the functions $\chi_0, \dots, \chi_s$ there are always at least two functionally independent ones if the basis of the space $L(\gamma, \Lambda)$ has the form (3).

Remark 2. If the density $f(x, y, \theta)$ depends only on the difference $x - y$ and is representable in the form (4), then $P(x, \theta) = xC(\theta)$, and the family of transition densities is an exponential family. It is not difficult to verify that the number of nontrivial sufficient statistics is equal to r .

Let us show that transition densities of the form (4) exist. This question reduces to the question of the existence of an integral equation of the form

$$v(x, \theta) = \int_0^\infty K(x, y, \theta) v(y, \theta) dy, \quad (5)$$

where

$$K(x, y, \theta) = \exp \left\{ \sum \alpha_1(\theta) \psi_i(x, y) + \psi_0(x, y) - a_0(\theta) \right\},$$

having a nonnegative solution that does not lead to an exponential family. Let the functions α_i, ψ_i be such that the kernel K is a symmetric difference kernel $K = K(x - y, \theta)$. Make the change of variables

$$v(x, \theta) = u(x, \theta) e^{\beta(\theta)x}. \quad (6)$$

Then equation (5) takes the form

$$u(x, \theta) = \int_0^\infty K e^{\beta(\theta)(y-x)} u(y, \theta) dy. \quad (7)$$

Choose $\beta(\theta) > 0$ so that

$$\int_{-\infty}^{\infty} K(t, \theta) e^{\beta(\theta)t} dt = 1. \quad (8)$$

For this it is enough to take $\beta(\theta)$ equal to the positive root of the equation $1 - L(s)$, where $L(s)$ is the Laplace transform of the kernel $K(t, \theta)$.

Since the kernel $K(t, \theta)$ is symmetric and nonnegative and $\beta(\theta) > 0$, we have

$$\int_{-\infty}^{\infty} tK(t, \theta) e^{\beta(\theta)t} dt = \int_0^{\infty} tK(t, \theta) e^{\beta(\theta)t} dt - \int_0^{\infty} tK(t, \theta) e^{-\beta(\theta)t} dt < 0. \quad (9)$$

The fulfillment of conditions (8), (9), on the basis of Lindley's theory [7], guarantees the existence of a nondecreasing solution $u(x, \theta)$, continuous from the right, of equation (7), such that $u(x, \theta) = 0$ for $x < 0$ and $\lim_{x \rightarrow 0} u(x, \theta) = 1$.

The fact that solution (6) of equation (5) does not lead to an exponential se- follows from the theory of Wiener-Hopf equations. Namely, equation (5) with a symmetric difference nonnegative kernel has, in the strip $-c < \operatorname{Re} s < c$, n solutions, where $2n$ is the number of roots of the equation $1 - L(s) = 0$, and these solutions have the form

$$u(x) = \sum_{i=1}^n Q_i(x) e^{s_i x} + O(e^{-hx}),$$

where $Q_i(x)$ is a polynomial, and the constant $h > 0$ does not depend on c .

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

Received
19 IX 1966

REFERENCES

1. W. O. K o o p m a n, Trans. Am. Math. Soc., **39**, 399 (1936).
2. E. B. D y n k i n, UMN, **6**, no. 1 (41) (1951).
3. G. B r o w n, Ann. Math. Stat., **36**, no. 3 (1965).
4. G a n i, J. Biometrika, **42**, 3 (1955).

5. J. G a n i, *Biometrika*, **43**, 3 (1956).
6. B. R. B h a t, J. G a n i, *Biometrika*, **47**, 3 (1960).
7. D. V. L i n d l e y, *Proc. Cambr. Phil. Soc.*, **78**, 1 (1955).
8. V. I. S m i r n o v, *Course of Higher Mathematics*, vol. 6, 3rd ed., Moscow, 1953.

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

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