

**CONDITIONS FOR  
OPTIMAL UNBIASED  
ESTIMATION OF  
PARAMETRIC  
FUNCTIONS FOR  
INCOMPLETE  
EXPONENTIAL  
FAMILIES WITH  
POLYNOMIAL  
CONSTRAINTS**

A. M. KAGAN, V. P. PALAMODOV

1967

SovietRxiv

---

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

**Abstract**

**Full Text**

UDC 519.281

**MATHEMATICS**

**A. M. KAGAN, V. P. PALAMODOV**

**CONDITIONS FOR OPTIMAL UNBIASED ESTIMATION OF PARAMETRIC FUNCTIONS FOR INCOMPLETE EXPONENTIAL FAMILIES WITH POLYNOMIAL CONSTRAINTS**

*(Presented by Academician Yu. V. Linnik on 27 X 1966)*

1°. We consider the problem of unbiased estimation of parametric functions from the results of  $n$  independent observations of a random variable with density, with respect to Lebesgue measure,

$$f(x; \alpha) = \exp\{t_0(x) + c_1(\alpha)t_1(x) + \dots + c_s(\alpha)t_s(x) + c_0(\alpha)\}, \quad (1)$$

where  $s \leq n$ ,  $\alpha \in A$  is an abstract parameter. Introduce the natural parameters

$$Q_i = c_i(\alpha), \quad i = 1, \dots, s,$$

and suppose that, as  $\alpha \in A$ , the point  $\theta = (\theta_1, \dots, \theta_s)$  runs through an everywhere dense subset of some algebraic subvariety of the domain  $\Omega \subset R^s$ . We shall write this subvariety in the form  $\Omega \cap \Pi$ , where  $\Pi$  is an algebraic variety in  $R^s$ , specified by the polynomial constraint equations

$$\Pi_1(\theta_1, \dots, \theta_s) = 0, \dots, \Pi_r(\theta_1, \dots, \theta_s) = 0, \quad (2)$$

where  $r < s$ .

The distribution of the repeated sample  $(x_1, \dots, x_n) = \mathbf{x}$  from the population (1) is given in  $R^n$  by the density with respect to Lebesgue measure

$$f^{(n)}(\mathbf{x}; \theta) = c(\theta)^n \exp \left\{ \sum_1^n t_0(x_i) + \theta_1 \sum_1^n t_1(x_i) + \dots + \theta_s \sum_1^n t_s(x_i) \right\}, \quad (3)$$

where  $c(\theta)$  is determined by the normalization condition. Sufficient statistics for the families (3) are

$$T_1 = \sum_1^n t_1(x_i), \dots, T_s = \sum_1^n t_s(x_i).$$

We shall assume that  $T_1, \dots, T_s$  are functionally independent; this holds under very general conditions, see (<sup>1</sup>). Then the distribution of the vector  $T = (T_1, \dots, T_s)$  is given in  $R^s$  by the density with respect to Lebesgue measure

$$p(T; \theta) = C(\theta)h(T) \exp(\theta_1 T_1 + \dots + \theta_s T_s), \quad (4)$$

where  $h(T) \geq 0$ ,  $\theta \in \Omega \cap \Pi$ . Denote by  $\text{supp } h$  the support of the function  $h(T)$ ; let  $\mathcal{T} = \text{int supp } h$ . Our condition on  $h(T)$  is as follows:

$$h(T) > 0, \quad T \in \mathcal{T}, \quad (5)$$

and  $h(T)$  is infinitely differentiable on  $\mathcal{T}$ .

**2°.** Let  $g(T)$  be an unbiased estimate of some function  $\gamma(\theta)$ ,  $\theta \in \Omega \cap \Pi$ , depending only on the vector of sufficient statistics  $T$  and having finite variance for all  $\theta \in \Omega \cap \Pi$ . We shall determine the conditions imposed on the variety  $\Pi$  and on the estimate  $g(T)$  itself by its property of being optimal for all  $\theta \in \Omega \cap \Pi$  in the class of unbiased estimates of the function  $\gamma(\theta)$  with finite variance. As the measure of the quality of an estimate we take its variance. Since the behavior of  $g(T)$  outside  $\mathcal{T}$  has no effect on its properties as an estimate of  $\gamma(\theta)$ , we shall regard  $g(T)$  as specified only on  $\mathcal{T}$ .

Denote by  $N$  the smallest (complex) algebraic variety in  $C^s$  containing  $\Omega \cap \Pi$ . Since  $\Pi$  itself is an algebraic variety,  $\Pi \cap \Omega = N \cap \Omega$ .

**Theorem.** In order that a statistic  $g(T)$ ,  $T \in \mathcal{T}$ , for which  $E_\theta g^2 < \infty$ ,  $\theta \in \Omega \cap \Pi$ , be the best unbiased estimate of the function  $E_\theta g = \gamma(\theta)$  for the exponential family (4),  $\theta \in \Omega \cap \Pi$ , with conditions (5), it is necessary and sufficient that:

1. In the space  $C^s$  of the variables  $\theta_1, \dots, \theta_s$  there exist a linear system of coordinates  $\theta'_1, \dots, \theta'_s$  in which  $N$  is a cylinder of the form  $L \times \nu$ , where  $L$  is the coordinate subspace  $\theta'_1 = \dots = \theta'_m = 0$ , and  $\nu$  is some set in the subspace  $\theta'_{m+1} = \dots = \theta'_s = 0$ ,  $0 \leq m \leq s$ .
2. In the corresponding coordinate system  $T'_1, \dots, T'_s$  in the space  $R^s$  of values of  $T_1, \dots, T_s$ , the function  $g(T)$  depends only on  $T'_{m+1}, \dots, T'_s$ .

This theorem generalizes the result of paper (2) to the case of arbitrary estimates of parametric functions; in (2) it was proved only for polynomials in sufficient statistics.

It is interesting to compare the situation with optimal unbiased estimation of parametric functions for full and incomplete exponential families. For the former, by the Rao–Blackwell–Kolmogorov theorem, every function  $g(T)$  depending on sufficient statistics is an optimal unbiased estimate of its mathematical expectation  $E_\theta g$ . For incomplete exponential families with polynomial constraints, the optimality of  $g(T)$  as an estimate of  $E_\theta g$  means, according to the theorem formulated above, the quasi-completeness of the family with respect to some of the parameters (the representation  $N = L \times \nu$  is analogous to completeness). The optimal estimate itself depends only on sufficient statistics of the same names as the parameters for which quasi-completeness holds.

Thus, for the “majority” of incomplete exponential families with polynomial constraints, parametric functions do not admit optimal unbiased estimation. We note that for the families under consideration there are examples of estimates depending on sufficient statistics and nevertheless inadmissible in the class of unbiased estimates of the corresponding parametric functions (3).

3°. Let  $\chi(T)$  be an unbiased estimate of zero with finite variance, i.e.

$$E_\theta \chi = 0, \quad E_\theta \chi^2 < \infty, \quad \theta \in \Omega \cap \Pi.$$

We have the following simple necessary and sufficient condition for the optimality of the estimate  $g(T)$  (see, for example, (4)).

**Lemma 1.** In order that  $g(T)$  be an optimal unbiased estimate of  $\gamma(\theta)$ ,  $\theta \in \Omega \cap \Pi$ , it is necessary and sufficient that for every unbiased estimate of zero with finite variance

$$E_\theta(g\chi) = 0, \quad \theta \in \Omega \cap \Pi.$$

The sufficiency of the conditions of the theorem is proved with the aid of Lemma 1 without particular difficulty. The necessity of these conditions is established considerably more complicatedly. In the proof, a sufficiently extensive and convenient stock of unbiased estimates of zero is used.

Let  $D = D(\mathcal{T})$  be the space of infinitely differentiable functions of  $T$  with compact supports lying in  $\mathcal{T}$ . For every function  $\Phi(T) \in D$  the Laplace transform

$$\tilde{\Phi}(\theta) = \int \exp(\theta, T) \Phi(T) dT$$

is an entire function in  $C^s$ . Directly from Lemma 1 it follows that if  $g(T)$  is the best unbiased estimate of the function  $\gamma(\theta)$ ,  $\theta \in \Omega \cap \Pi$ , and  $\chi(T) \in D$  is such that  $\tilde{\chi}(\theta) = 0$ ,  $\theta \in \Omega \cap \Pi$ , then

$$\tilde{g\chi}(\theta) = 0, \quad \theta \in \Omega \cap \Pi. \tag{6}$$

Let  $\mathcal{A}$  be the ring of all polynomials in  $\theta \in C^s$  with complex coefficients. If  $J$  is some ideal in  $\mathcal{A}$ , and  $\xi \in C^s$ , then by  $J_\xi$  we shall denote the ideal generated by the polynomials  $p(\theta + \xi)$ , where  $p(\theta) \in J$ . Consider the set  $\mathfrak{G} \in \mathcal{A}$  of those polynomials  $p(\theta)$  for which  $g(T)$ —the best unbiased estimate of the function  $\gamma(\theta)$ —is a generalized solution in  $\mathcal{T}$  of the equation

$$p(D)g = 0, \quad D = (\partial/\partial T_1, \dots, \partial/\partial T_s).$$

It is easy to see that  $\mathfrak{G}$  is an ideal. Denote by  $I$  the ideal in  $\mathcal{A}$  generated by all polynomials that vanish on  $N$ . The starting point in the proof of necessity is

**Lemma 2.** If  $\xi \in N$ , then  $I_\xi \subset \mathfrak{G}$ .

**Proof of Lemma 2.** Let  $p(\theta) \in I_\xi$ , i.e.  $p(\theta - \xi) \in I$ . If  $g(T)$  is regarded as a generalized function in  $\mathcal{T}$ , then for an arbitrary  $\varphi \in \mathcal{D}$  we shall have

$$\begin{aligned} (p(D)g, \varphi \exp(\xi, T)) &= (g, p(-D)\varphi \exp(\xi, T)) = \\ &= (g, \exp(\xi, T)p(-D - \xi)\varphi) = \int g(T) \exp(\xi, T)p(-D - \xi)\varphi(T) dT. \end{aligned} \quad (7)$$

Since  $p(\theta - \xi) \in I$ , the function

$$p(-\widetilde{D} - \xi)\varphi = p(\theta - \xi)\tilde{\varphi}(\theta)$$

vanishes on  $\Omega \cap \Pi$ . But then, according to (6),

$$\int g(T) \exp(\xi, T)p(-D - \xi)\varphi(T) dT = 0, \quad \theta \in \Omega \cap \Pi,$$

since the expression on the left is equal to  $\widetilde{g\chi}(\theta)$ , where  $\chi = p(-D - \xi)\varphi$ , and satisfies the condition  $\tilde{\chi}(\theta) = 0$ ,  $\theta \in \Omega \cap \Pi$ . Now from (7) we obtain that the generalized function  $p(D)g$  is equal to zero on functions of the form  $\varphi \exp(\xi, T)$ . Since any function in  $\mathcal{D}$  can be represented in this form,  $p(D)g = 0$  in  $\mathcal{T}$ , i.e.  $p(\theta) \in \mathfrak{G}$ .

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

Moscow State University  
named after M. V. Lomonosov

Received  
18 IX 1966

## REFERENCES

1. E. B. Dynkin, *Uspekhi Mat. Nauk*, 6, 1 (1951).
2. A. M. Kagan, V. P. Palamodov, *Theory of Probability and Its Applications*, 12, 1 (1967).
3. Yu. V. Linnik, *Statistical Problems with Nuisance Parameters*, “Nauka,” 1966.
4. C. R. Rao, *Linear Statistical Inference and its Applications*, N. Y., 1965.

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

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