



Soviet-era science, translated into English

Mathematics

Corresponding Member of the Academy of Sciences of the USSR
Yu. V. Linnik

1964

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196401.60243>

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

Abstract

Full Text

Mathematics

Corresponding Member of the Academy of Sciences of the USSR Yu. V. Linnik

STATISTICAL PROBLEMS WITH NUISANCE PARAMETERS

Many parametric problems of modern mathematical statistics pertain to exponential families of distributions. We shall consider families of distributions in an n -dimensional Euclidean space E_n , specified by a probability element in the space of sufficient statistics T_1, \dots, T_s :

$$dP_\theta = C(\theta_1, \dots, \theta_s) \exp(-\theta_1 T_1 + \dots + \theta_s T_s) p(T_1, \dots, T_s) dT_1 \dots dT_s. \quad (1)$$

Here $\theta_1, \dots, \theta_s$ are parameters; $p(T_1, \dots, T_s)$ is a density with respect to Lebesgue measure; we shall assume it to be Riemann integrable. The set of values of the parameters $\Omega = \{\theta_1, \dots, \theta_s\}$ will be taken from the Euclidean space E_s . We shall regard it as natural, i.e. consisting of all values that ensure the absolute convergence of $\int_{E_n} dP_\theta$. It is known (see (1)) that it is convex.

The usual statistical problems concerning the family (1), with nuisance parameters, reduce to testing for the existence among the parameters $\theta_1, \dots, \theta_s$ of a finite number of polynomial relations (for example, the well-known Behrens-Fisher problem reduces to testing the hypothesis $H_0 : \theta_1 \theta_4 - \theta_2 \theta_3 = 0$) under a corresponding choice of sufficient statistics and parametrization.

Let us dwell on the hypotheses that we shall test. Let $\Omega_0 \subset \Omega$ be a bounded parallelepiped. The hypothesis H_0 will consist in the assertion that among the parameters $\theta_1, \dots, \theta_s$ there exist $r < s$ relations:

$$P_1(\theta_1, \dots, \theta_s) = 0; \quad P_2(\theta_1, \dots, \theta_s) = 0; \quad \dots; \quad P_r(\theta_1, \dots, \theta_s) = 0, \quad (2)$$

where $P_i(\theta_1, \dots, \theta_s)$ are polynomials with real coefficients, homogeneous and absolutely irreducible (i.e. irreducible over the field C of all complex numbers).

Tests of the hypothesis H_0 against any simple or composite alternatives H_1 (of course, also concerning the behavior of the parameters $\theta_1, \dots, \theta_s$) will be conducted on a single sample corresponding to the family (1), with parameters $(\theta_1, \dots, \theta_s) \in \Omega_0$. In doing so we shall use test functions ϕ (tests) subject to the conditions:

A. The test ϕ may be randomized.

B. The test ϕ is defined in the space of sufficient statistics:

$$\phi = \phi(T_1, \dots, T_s).$$

C. The test ϕ excludes nuisance parameters, i.e. is a similar test under the hypothesis H_0 :

$$E(\phi | H_0) = \alpha, \tag{3}$$

where $\alpha \in (0, 1)$ is the level of the test, independent of the parameters.

We shall consider here the question of as complete as possible a description of such tests and of selecting from among them those that are optimal in one or another sense.

One can give a number of weighty statistical arguments to justify restrictions A, B, C, but we shall not dwell on them here.

Apparently, in the most general formulation one should consider not only tests ϕ without these restrictions, but also “generalized tests,” defined not as functions of a point of the sample space, but as functionals on families of sufficiently smooth “basic functions” of the L. Schwartz type, specified on the sample space. Such a point of view has its advantages, especially for exponential families. However, we shall consider this question in another paper.

The question of the optimal choice of the test ϕ , as is usual in statistics, will depend on what kind of optimization is assumed. In particular, when the alternative H_1 is a simple hypothesis, or, more generally, when we have a Bayesian situation with alternatives subject to a prior distribution with probability element $dB(\theta)$, vanishing on Ω_0 in some neighborhood of H_0 , it is natural to require the condition:

$$E(\phi | H_1) = \int_{E_n \times \Omega_0} \dots \int \phi dP_0 dB(\theta) = \max, \tag{4}$$

where the maximum is taken over all tests ϕ satisfying conditions A, B, C (henceforth we shall call such tests admissible). Of course, such a maximum may also fail to exist; therefore, putting

$$\beta = \sup_{\phi} E(\phi | H_1) \tag{5}$$

under the same conditions on the tests ϕ , we may consider the problem of constructing, for any $\varepsilon > 0$, an admissible test ϕ_ε for which $E(\phi_\varepsilon | H_1) \geq \beta - \varepsilon$. Such tests will be called ε -optimal, and henceforth we shall deal with the problem of constructing ε -optimal admissible tests.

Let us proceed to describe the solution of this problem for the class of exponential families satisfying certain additional requirements (exponential families of type \mathcal{E}_0).

Let the density $p(T_1, \dots, T_s)$ satisfy the condition:

$$p(T_1, \dots, T_s) < \exp A_0 (|T_1| + \dots + |T_s|) \quad (6)$$

for $|T_i| \geq 1$ ($A_0 > 0$ is a constant).

Such a condition ensures the nondegeneracy of the set Ω of values of the parameters Ω in the space E_s . Further, suppose that the set where $p(T_1, \dots, T_s)$ vanishes consists of those and only those points T_1, \dots, T_s where at least one of the k inequalities

$$T_i \leq c_i, \quad (7)$$

where $i \leq k$ and c_i are finite numbers, is satisfied. We shall also require that the inequality

$$p(T_1, \dots, T_s) > |(T_1 - c_1) \dots (T_k - c_k)|^{K_0} \quad (8)$$

be satisfied for sufficiently small values of $T_i - c_i$; $K_0 > 0$ is a constant.

Such families will be called families of type \mathcal{E}_0 . They include most of the known problems of testing hypotheses with nuisance parameters in exponential families. In what follows, we shall assume that we are dealing with the indicated family \mathcal{E}_0 ; the theorem stated below will also apply to it.

Along with admissible tests ϕ it is convenient to consider cotests. A cotest for a given test ϕ will be a collection of functions of the form

$$\psi = K(\phi - \alpha), \quad (9)$$

where α is the level of the test (the right-hand side of relation (3)), and K is any constant. For every cotest we have:

$$E(\psi | H_0) = 0. \quad (10)$$

To describe the ε -optimal tests it is sufficient to describe the corresponding ε -optimal cotests. Consider the absolutely irreducible polynomials (2) and multiply them. We obtain a homogeneous polynomial $\Pi(\theta_1, \dots, \theta_s)$ of degree h . Divide it by $(T_1 \dots T_s)^{h+2}$. We obtain a new polynomial $F\left(\frac{1}{\theta_1}, \dots, \frac{1}{\theta_s}\right)$.

In it, replace each power $\theta_j^{-s_j}$ by $(-1)^{s_j-1}T_j^{s_j-1}$. We obtain a new polynomial $F_0(T_1, \dots, T_s)$.

Theorem. For every $\varepsilon > 0$ one can effectively specify such a bounded parallelepiped \mathfrak{S}_ε in the space of sufficient statistics (T_1, \dots, T_s) that the ε -optimal cotests will be found among the family of cotests defined by the formula:

$$\Psi = F_0 * H = \int_{\mathfrak{S}_\varepsilon} \dots \int F_0(T_1 - x_1, \dots, T_s - x_s) H(x_1, \dots, x_s) dx_1 \dots dx_s, \quad (11)$$

where H is any function in L_1 on \mathfrak{S}_ε .

By virtue of this theorem, the search for ε -optimal tests is reduced to a variational problem for the function H : it is sought under the conditions

$$\int_{E_0 \times \Omega_0} \dots \int (F_0 * H) dB(\theta) = \max;$$

$$-\alpha p(T_1, \dots, T_s) \leq F_0 * H \leq (1 - \alpha) p(T_1, \dots, T_s). \quad (12)$$

The method of proof was suggested to the author by the very interesting work of R. A. Wijsman ⁽²⁾, where special cases of the construction of such tests are given. The method of proof of the theorems is based on the construction and study of ideals formed by cotests.

Received
13 IV 1964

REFERENCES

- ¹ E. Lehmann, *Testing Statistical Hypotheses*, N. Y., 1959.
² R. A. Wijsman, *Ann. Math. Statistics*, **29**, No. 4 (1958).

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

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