Mathematics

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

1962

SovietRxiv

Abstract

Full Text

Mathematics

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

ON THE THEORY OF STATISTICALLY SIMILAR REGIONS

The theory of similar regions was founded by J. Neyman and E. Pearson in 1939. ([1]) (For definitions and results up to 1954, see ([2]).) Some properties of similar regions of linear type were given in the author' s preceding note ([3]). Here we indicate one rather broad class of distributions that allow the construction of similar regions.

We consider Euclidean n-dimensional space E_n and on it a family of probability measures P_θ, defined on one and the same σ-algebra. Let each measure P_θ have a continuous probability density $L(X, \theta)$ with respect to Lebesgue measure. We shall assume that the parameter θ is a number lying in the interval (a, b). A statistic t such that, for all ξ, θ, $P(t < \xi)$ exists and does not depend on θ, will be called zonal; the construction of similar regions and zonal statistics is essentially one and the same problem. If $\theta_1, \dots, \theta_s$ is any finite set of values of the parameter, then statistics t that are zonal for this set of parameter values and take, with positive probability, values lying in a finite number of prescribed intervals can be constructed with the aid of A. A. Lyapunov' s theorem (see ([4])). We shall consider those families of measures P_θ and the corresponding densities $L(X, \theta)$ for which, starting from the indicated construction for a finite set of parameter values, we can also carry out the construction for $\theta \in (a, b)$.

In what follows we shall assume that all functions introduced by us possess the required smoothness conditions, ensuring the validity of the analytic operations carried out below.

Consider a system of scalar statistics $V_1, V_2, \dots, V_r$ $(r < n)$. Suppose that, on any nonempty level surfaces $V_1 = C_1, \dots, V_r = C_r$, the measures P_θ generated by the expression $L(X, \theta)\, dX$ allow the introduction of a conditional distribution.

We shall suppose that a system of local coordinates $\xi_1, \dots, \xi_{n-r}$ can be introduced on the indicated level surfaces in such a way that throughout E_n we obtain the system of coordinates $(V_1, \dots, V_r; \xi_1, \dots, \xi_{n-r})$.

Now require that $L(X, \theta)$, for $\theta \in (a, b)$ and for all values of X, satisfy the differential equation:

$$p_0(V_1, \dots, V_r, \theta) \frac{\partial^k L(X, \theta)}{\partial \theta^k} + p_1(V_1, \dots, V_r, \theta) \frac{\partial^{k-1}}{\partial \theta^{k-1}} L(X, \theta) + \cdots$$

$$\cdots + p_k(V_1, \dots, V_r, \theta) L(X, \theta) + p_{k+1}(V_1, \dots, V_r, \theta) = 0. \tag{1}$$

The coefficients $p_j(V_1, \dots, V_r, \theta)$ are assumed to be sufficiently smooth; on $p_0(V_1, \dots, V_r, \theta)$ further analytic conditions are imposed, permitting the subsequent transformations.

Let, on the level surface corresponding to $V_1, \dots, V_r$, there exist a conditional probability density $l(\xi_1, \dots, \xi_{n-r}, V_1, \dots, V_r, \theta)$, so that $L(X, \theta) = g(V_1, \dots, V_r, \theta)\, l(\xi_1, \dots, \xi_{n-r}, V_1, \dots, V_r, \theta)\, J$, where $g(V_1, \dots, V_r, \theta)$ is the joint density of the distribution of the statistics $V_1, \dots, V_r$; J is the Jacobian of the transformation, of course not depending on θ.

Under sufficiently general analytic conditions, for given values $V_1, \dots, V_r$ the conditional probability density $l(\xi_1, \dots, \xi_{n-r}, V_1, \dots, V_r, \theta)$, which for brevity we shall denote by $l(\xi, \theta)$, will satisfy the equation

$$\rho_0(\theta) \frac{\partial^m}{\partial \theta^m} l(\xi, \theta) + \rho_1(\theta) \frac{\partial^{m-1}}{\partial \theta^{m-1}} l(\xi, \theta) + \cdots + \rho_{m-1}(\theta) \frac{\partial l(\xi, \theta)}{\partial \theta} = 0, \tag{2}$$

where $m = k + 2$.

When the analytic conditions are specified, the existence of equation (2) for the conditional probability density is, of course, equivalent to the existence of a finite basis for $l(\xi, \theta)$:

$$l(\xi, \theta) = c_1(\xi) s_1(\theta) + \cdots + c_m(\xi) s_m(\theta),$$

where $V_1, \dots, V_r$ are regarded as fixed, $s_j(\theta)$ depends only on θ, and $c_j(\xi)$ only on ξ. For what follows, however, it will be more convenient for us to use the language of differential equations.

Let $\psi(\xi)$ be a continuous function of $\xi_1, \dots, \xi_{n-r}$ such that the conditional mathematical expectation exists:

$$\int \cdots \int \psi(\xi) l(\xi, \theta)\, d\xi = u_\psi(\theta) = u_\psi. \tag{3}$$

Put $du_\psi / d\theta = w_\psi$; under sufficiently general conditions one may interchange the operators of differentiation with respect to θ and integration with respect to the parameter ξ, so that the equation obtained is

$$\rho_0(\theta) \frac{d^{m-1}}{d\theta^{m-1}} w_\psi + \rho_1(\theta) \frac{d^{m-2}}{d\theta^{m-2}} w_\psi + \cdots + \rho_{m-1}(\theta) w_\psi = 0. \tag{4}$$

Let, for $\theta \in (a, b)$, $w_1(\theta), w_2(\theta), \dots, w_{m-1}(\theta)$ be a set of its fundamental solutions. We have:

$$w_\psi = w_\psi(\theta) = c_1^{(\psi)} w_1(\theta) + \dots + c_{m-1}^{\psi} w_{m-1}(\theta). \tag{5}$$

Under sufficiently general conditions there will be segments $[\theta_i^0 - \varepsilon, \theta_i^0 + \varepsilon] \subset (a, b)$ $(i = 1, 2, \dots, m-1)$ such that if $\theta_i \in [\theta_i^0 - \varepsilon, \theta_i^0 + \varepsilon]$ and $w_\psi(\theta_i) = 0$ $(i = 1, 2, \dots, m-1)$, then, by virtue of relation (5), $c_j^{(\psi)} = 0$ $(j = 1, 2, \dots, m-1)$ and $w_\psi(\theta) = 0$ for $\theta \in (a, b)$.

Take such segments and in each of them take two distinct points θ_i' and θ_i''. Consider $2m-2$ conditional probability measures generated by $l(\xi, \theta)$, where $\theta = \theta_i'$ or $\theta = \theta_i''$ $(i = 1, 2, \dots, m-1)$. For this finite family of measures, according to A. A. Lyapunov's theorem ([4]), construct a conditional zonal statistic $t = t(\xi)$. Put $\psi = \psi(\xi) = \exp i\tau t$, where τ is a scalar parameter. Then $u_\psi(\tau, \theta)$ is the characteristic function of t. At the same time we have: $\Re u_\psi(\tau, \theta_i') = \Re u_\psi(\tau, \theta_i'')$; $\Im u_\psi(\tau, \theta_i') = \Im u_\psi(\tau, \theta_i'')$.

By Rolle' s theorem there will be points $\theta_i \in [\theta_i', \theta_i'']$ such that

$$\frac{d}{d\theta} \Re u_\psi(\tau, \theta) = 0$$

at $\theta = \theta_i$; from (5) and the preceding it follows that then

$$\frac{d}{d\theta} \Re u_\psi(\tau, \theta) = 0$$

for $\theta \in (a, b)$, and the same holds for

$$\frac{d}{d\theta} \Im u_\psi(\tau, \theta),$$

so that $u_\psi(\tau, \theta)$ does not depend on θ, and $t = t(\xi)$ is a conditional zonal statistic on the level surface with given $V_1, \dots, V_r$. Under sufficiently general analytic assumptions we can carry out such a construction on every level surface and cut out similar zones there, and then paste them together over all level surfaces, as is done in J. Neyman' s theory of structures (see ([5])).

Let us consider some particular cases.

1. The case $k = 1$. In (1) this is a differential equation of the first order. With a special choice of the coefficients in equation (1) we obtain

$$L(X, \theta) = \exp(T_1(X) + \theta T_2(X) + h(\theta)).$$

This is the case of the existence of sufficient statistics under a natural parametrization.

2.

$$\frac{\partial}{\partial\theta}\frac{L(X,\theta)}{g(V_1,\dots,V_r,\theta)}=0.$$

This is the general case of sufficient statistics $V_1,\dots,V_r$ and of the existence of J. Neyman structures.

3. The quantity $L(X,\theta)/g$ satisfies a linear differential equation of some order with coefficients depending only on θ. In this case, which includes the preceding one, the statistics $V_1,\dots,V_r$ may be called quasi-sufficient. It is possible to describe the construction of similar regions quite fully.

4. The case of repeated sampling $x_1,\dots,x_r$ for x_j having the (one-dimensional) distribution density

$$\mu(x,\theta)=\left(\exp\sum_{j=1}^{M}h_j(\theta)T_j(x)\right)\left(\sum_{i=0}^{N}\varphi_i(x)g_i(\theta)\right). \tag{6}$$

This is the case of a family that previously had sufficient statistics and was then "spoiled" by perturbation terms of a fairly general kind. It is useful for studying the power of statistical tests based on similar regions.

Received
11 VI 1962

REFERENCES

[1] J. Neyman, E. Pearson, Phil. Trans. Roy. Soc., Ser. A, 289 (1933).
[2] J. Neyman, Current Problems of Mathematical Statistics, Intern. Math. Congress Amsterdam, 1954.
[3] Yu. V. Linnik, DAN, 144, No. 5 (1962).
[4] A. A. Lyapunov, Izv. AN SSSR, ser. matem., 4, 467 (1940).
[5] E. Lehmann, Testing Statistical Hypotheses, N. Y., 1959.

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