SOME CHARACTERIZATION PROBLEMS IN STATISTICS

YU. V. PROHOROV
V. A. STEKLOV INSTITUTE, MOSCOW

1. Introduction

In this paper we shall discuss problems connected with tests of the hypothesis that a theoretical distribution belongs to a given class, for instance, the class of normal distributions, or uniform distribution or Poisson distribution. The statistical data consist of a large number of small samples (see [1]).

2. Reduction to simple hypotheses

Let $(\mathcal{X}, \mathcal{A})$ be a measurable space ($\mathcal{X}$ is a set and $\mathcal{A}$ is a σ-algebra of subsets of $\mathcal{X}$). Let $\mathcal{P}$ be a set of probability distributions defined on $\mathcal{A}$, let $(\mathcal{Y}, \mathcal{B})$ be another measurable space, and let $Y = f(X)$, $X \in \mathcal{X}$, be a measurable mapping of $(\mathcal{X}, \mathcal{A})$ into $(\mathcal{Y}, \mathcal{B})$. With this mapping every distribution P induces on $\mathcal{B}$ a corresponding distribution which we shall denote by Q_P^Y. We will be interested in the mappings (statistics) Y which possess the following two properties:

(1) Q_P^Y is the same for all $P \in \mathcal{P}$; in this case we will simply write $Q_{\mathcal{P}}^Y$.

(2) If for some P' on $\mathcal{A}$ one has $Q_{P'}^Y = Q_{\mathcal{P}}^Y$, then $P' \in \mathcal{P}$.

Sometimes it is expedient to formulate requirement (2) in the weakened form:

(2a) If $P' \in \mathcal{P}' \supset \mathcal{P}$ and $Q_{P'}^Y = Q_{\mathcal{P}}^Y$, then $P' \in \mathcal{P}$. In other words, we can assert in this case only that the equation $Q_{P'}^Y = Q_{\mathcal{P}}^Y$ implies $P' \in \mathcal{P}$ for some a priori restrictions $(P' \in \mathcal{P}')$ on P'.

If Y is a statistic satisfying (1) and (2), then it is clear that the hypothesis that the distribution of X belongs to class $\mathcal{P}$ is equivalent to the hypothesis that the distribution of Y is equal to $Q_{\mathcal{P}}^Y$.

Let us consider some examples. In these examples $(\mathcal{X}, \mathcal{A})$ is an n-dimensional Euclidean space of points $X = (x_1, \cdots, x_n)$ with the σ-algebra of Borel sets. The distributions belonging to $\mathcal{P}$ have a probability density of the form

$$p(x_1, \theta)p(x_2, \theta) \cdots p(x_n, \theta) \tag{2.1}$$

where p is a one-dimensional density and θ a parameter taking values in a parameter space.

EXAMPLE 1 (I. N. Kovalenko [2]). *Translation parameter.* Let $p(x; \theta) = p(x - \theta)$, with $-\infty < \theta < \infty$ (*additive type*). Here obviously it is necessary to take the $(n - 1)$-dimensional statistic $Y = (x_1 - x_n, \cdots, x_{n-1} - x_n)$. Of course, we can take any uniquely invertible function, for example $Y' = (x_1 - \bar{x}, \cdots, x_n - \bar{x})$ where $\bar{x} = (1/n) \sum_1^n x_k$.