ON THE FIDUCIAL APPROACH IN STATISTICS

MATHEMATICS

1970

SovietRxiv

Abstract

Full Text

UDC 519.21

MATHEMATICS

G. P. KLIMOV

ON THE FIDUCIAL APPROACH IN STATISTICS

(Presented by Academician A. N. Tikhonov, 4 IX 1969)

Let $\mathcal{P} = \{P_\theta,\ \theta \in \Omega\}$ be a family of distributions defined on one and the same sample space (X, B) and absolutely continuous with respect to a fixed measure μ. We denote the density of P_θ with respect to μ by $p(x \mid \theta)$. Consider the class $\mathcal{P}_\nu$ of (prior) distributions P, defined on the parameter space (Ω, F) and absolutely continuous with respect to the measure ν. Suppose that

$$0 < p(x) = \int p(x \mid \theta)\, d\nu(\theta) < \infty$$

for almost all (with respect to μ) x.

Definition. Let $\{\Omega_n\}_{n\geq 1}$ be an arbitrary monotonically increasing sequence of subsets of Ω such that $\bigcup \Omega_n = \Omega$, $\nu(\Omega_n) < \infty$. Denote by $\{P_n\}_{n\geq 1}$ a sequence of distributions from $\mathcal{P}_\nu$ such that the distribution P_n is concentrated on Ω_n and has maximal entropy $-\int p_n(\theta) \ln p_n(\theta)\, d\nu(\theta)$ among such distributions. Here $p_n(\theta)$ is the density of the distribution P_n with respect to the measure ν. Let $p_n(\theta \mid x)$ be the density of the posterior distribution (with respect to ν), computed by Bayes' formula starting from the prior density $p_n(\theta)$. Then the distribution P_x on (Ω, F), defined for almost all x and having density with respect to ν equal to $p(\theta \mid x) = \lim p_n(\theta \mid x)$, will be called the fiducial distribution (corresponding to the observation x). The density $p(\theta \mid x)$ is computed uniquely and is equal to $p(x \mid \theta)/p(x)$. Thus the fiducial distribution (fid. d.) P_x is defined uniquely up to the measure ν. We shall now present a statistical model for which the requirements of invariance of statistical conclusions uniquely single out the measure ν, and hence also the fid. d. P_x.

Let G be a group of measurable transformations of X onto X. We make the following assumptions.

P.1. $\theta_1 = \theta_2 \Longleftrightarrow P_{\theta_1}(E) = P_{\theta_2}(E)$ for every $E \in B$.

P.2. The family $\mathcal{P}$ is closed with respect to transformations from G. Thus, if the random variable (r.v.) x has distribution P_θ and $g \in G$, then there exists $\theta^* \in \Omega$ such that the r.v. gx has distribution P_{θ^*}. By P.1 such an element θ^* is unique. Thus to each $g \in G$ there corresponds a mapping g^* of the set Ω into itself. The set of such mappings g^* will be denoted by G^*.

P.3. Every element $g^* \in G^*$ is a mapping of Ω onto itself. Now the set G^* becomes a group.

P.4. The groups G and G^* are isomorphic.

P.5. The group G of transformations of the set X is strictly transitive, i.e., for any points $x, y \in X$ there exists a unique transformation $g \in G$ taking x into y: $y = gx$.

P.6. The group G^* of transformations of the set Ω is strictly transitive.

The elements of the group G can now be used to describe points $x \in X$ and points $\theta \in \Omega$. To this end choose arbitrary points x_0 and θ_0 from X and Ω, respectively. These points will play the role of reference points (scale elements), relative to which the remaining points of X and Ω will be "measured." Since for each $x \in X$ there exists a unique transformation $g_x \in G$ taking x_0 into $x = g_{xx}0$, it follows thereby ···

the elements $x \in X$ are parametrized by the elements $g_x \in G$. Similarly, since for each $\theta \in \Omega$ there exists a unique transformation $g^*_\theta \in G^*$ taking θ_0 into $\theta = g^*_\theta \theta_0$, the elements $\theta \in \Omega$ are described by the elements $g^*_\theta \in G^*$. By virtue of the isomorphism of the groups G and G^*, to the element $g^*_\theta \in G^*$ there corresponds an element $g_\theta \in G$. We have indicated a one-to-one correspondence between the elements of the sets X, Ω, G, G^*. In particular, this correspondence generates a σ-algebra F of subsets of the set Ω, starting from the σ-algebra B of subsets of the set X. Moreover, using such a correspondence, we shall denote by the same letter x both an element of X and the corresponding element g_x of G. We shall do the same for elements of Ω. For example, $\theta^{-1}x$ is understood to mean $g_\theta^{-1} g_x \in G$.

Suppose that the following two principles are fulfilled: the invariance principle of the fiducial distribution and the invariance principle of the entropy of the fiducial distribution with respect to the choice of the "scale elements" x_0 and θ_0. It turns out that if the left Haar measure on the group G is chosen as μ, then from the first principle it follows that the measure ν is relatively invariant (see ([1]), p. 257, problem 6). If the second principle is also fulfilled, then ν becomes a right-invariant measure, i.e. a right Haar measure, which is determined uniquely up to a constant positive factor. Thus the fiducial distribution P_x is determined uniquely. In this case

$$p(\theta \mid x) = \Delta(x) \cdot q(\theta^{-1}x),$$

where q is the density of the distribution $P_{\theta_0} \in \mathcal{P}$, and $\Delta(x)$ is the modular function (see ([1]), p. 256, problem 5).

Remark. For the fiducial distribution so defined, the equality (of fiducial and confidence probabilities)

$$P_\theta\{\theta \in S(x)\} = P_x\{\theta \in S(x)\}$$

holds for every class $\{S(x), x \in X\}$ of measurable (confidence) sets $S(x) \subseteq \Omega$, invariant with respect to the group G, i.e. such that $g^* S(x) = S(gx)$ for any $g \in G$. This equality may be taken as the basis for the definition of the fiducial distribution P_x: it uniquely determines P_x, if one does not require the fulfillment of the two principles indicated above. For the definition of the fiducial distribution on the basis of the so-called central function and the frequency interpretation of the fiducial distribution, see ([2]).

Now one may introduce the fiducial distribution of a sample variable as the distribution of the sample variable when the unobserved parameter θ has the fiducial distribution corresponding to some observation x.

Example. Let $x_1, \dots, x_n$ be independent observations on a random variable from an r-dimensional normal population $N(\mu, A)$.

Case 1. A is known, μ is unknown. $\mathcal{P}$ is the family of distributions of sufficient statistics

$$\bar{x} = \frac{1}{n}(x_1 + \dots + x_n),$$

corresponding to different values of μ; X is the Euclidean space E_r; G is the group of translations in E_r.

Case 2. A is unknown, μ is known and $\mu = 0$. Denote by G^- the group of lower triangular matrices of dimension $r \times r$ with positive elements on the main diagonal. Let $\mathcal{P}$ be the family of distributions of sufficient statistics t, corresponding to different values $a \in G^-$; here t and a are determined uniquely (almost everywhere) by the requirement:

$$aa' = A, \quad tt' = T = \sum_1^n x_k x_k'; \qquad a, t \in G^-; \quad n \geq r. \qquad X = \Omega = G = G^-.$$

Case 3. A is unknown, μ is unknown; $\mathcal{P}$ is the family of distributions of sufficient statistics $(s, \bar{x})$, corresponding to different $(a, \mu) \in G^- \times E_r$; here a and s are determined uniquely (almost everywhere) by the requirement:

$$aa' = A, \quad ss' = S = \frac{1}{n-1} \sum_1^n (x_k - \bar{x})(x_k - \bar{x})'; \qquad a, \ s \in G^-,$$

$X = \Omega = G^- \times E_r$. $G = \{[a, \mu] \mid a \in G^-, \ \mu \in E_r\}$ is a group with group operation $[a, \mu][s, x] = [as, ax + \mu]$.

Let $W^-(r, n, B)$ be a distribution concentrated on the set of positive definite matrices of dimension $r \times r$, with density

$$p(A) = \gamma_0(r, n) \frac{|B|^{(n-r-1)/2} d_-(B)}{|A|^{n/2} d_-(A)} \exp\left\{-\frac{n}{2} \operatorname{tr}(A^{-1} B)\right\},$$

where $d\_(A)$ is the product of the principal minors. Denote by $S^{-}(r,n)$ the distribution on E_r with density

$$p(t) = \gamma_1(r,n)\left[1+\frac{(t,t)}{n}\right]^{-(n+1)/2} \frac{[1+(t,t)/n]^{(r-1)/2}}{\prod_1^{r-1}[1+(t,t)_k/n]},$$

where $t = (t_1, \dots, t_r)$; $(t,t)_k = t_1^2 + \dots + t_k^2$. If the random variable t has distribution $S^{-}(r,n)$ and $a \in G^{-}$, then the distribution of the random variable at will be denoted by $K^{-}(r,n,A)$, where $A = aa'$. A fiducial random unobservable parameter and a fiducial sample variable will be marked by an asterisk above. Then, in case 1,

$$\mu^* \in N\left(\bar{x}, \frac{1}{n}A\right); \qquad x^* \in N\left(\bar{x}, \frac{n+1}{n}A\right)$$

(the sign $\in$ means that, for example, x^* has distribution $N\left(\bar{x}, \frac{n+1}{n}A\right)$).

In case 2

$$A^* \in W^{-}(r,n,\hat{A}); \quad x^* - \mu \in K^{-}(r,n,A); \quad \hat{A} = \frac{1}{n}\sum_1^n (x_k-\mu)(x_k-\mu)'.$$

In case 3

$$\sqrt{n}(\mu^* - \bar{x}) \in K^{-}(r, n-1, S); \qquad A^* \in W^{-}(r, n-1, S);$$

$$x^* - \bar{x} \in K^{-}\left(r, n-1, \frac{n+1}{n}S\right).$$

Moscow State University
named after M. V. Lomonosov

Received
1 IX 1969

CITED LITERATURE

[1] P. Halmos, *Measure Theory*, Moscow, 1953.
[2] D. A. S. Fraser, *Biometrika*, **48**, 261 (1961).

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