STRONGLY SYMMETRIC FAMILIES

MATHEMATICS

1970

SovietRxiv

Abstract

Full Text

UDC 519.281

MATHEMATICS

A. L. RUKHIN

STRONGLY SYMMETRIC FAMILIES

(Presented by Academician Yu. V. Linnik on 2 VI 1969)

The aim of the present note is to describe a class of families with a shift parameter for which the optimal homogeneous estimator does not essentially depend on the choice of a symmetric quality measure.

Let P be a probability measure on the real line R_1 with the σ-algebra of Borel subsets $\mathfrak{B}$, and let

$$\{P_\theta\} = \{P_\theta : \theta \in R,\ P_\theta(B) = P(B - \theta),\ B \in \mathfrak{B}\}.$$

Definition 1. We shall say that an estimator of the shift parameter g is no worse than an estimator f if, for all $s \in R_1,\ \theta \in R_1$,

$$E_\theta \left| e^{isg(x)} - e^{is\theta} \right|^2 \leq E_\theta \left| e^{isf(x)} - e^{is\theta} \right|^2. \tag{1}$$

Here $f(x) = f(x_1, \dots, x_n)$, where $x = (x_1, \dots, x_n) \in R_n$ is a repeated sample from the indicated population.

Proposition. *If f is a homogeneous estimator, i.e.*

$$f(x_1 + c, \dots, x_n + c) = f(x_1, \dots, x_n) + c, \qquad (x_1, \dots, x_n) \in R_n, \quad c \in R_1,$$

and the estimator g is defined by the relation

$$e^{isg(x)} = e^{isf(x)} \frac{E\{e^{-isf(x)}/\mathfrak{A}\}}{\left|E\{e^{-isf(x)}/\mathfrak{A}\}\right|}, \tag{2}$$

where $\mathfrak{A} = \langle x_2 - x_1, \dots, x_n - x_1 \rangle,\ s \in R_1$, then g is homogeneous and is no worse than f in the sense of (1). Moreover, in the case of existence, g does not depend on f and, consequently, is optimal in the class of homogeneous estimators.

If

$$P(B) = \int_B p(t)\, dt, \qquad B \in \mathfrak{B},$$

then the estimator (2) can be written in the more convenient form

$$e^{isg(x_1, \dots, x_n)} = \int_{-\infty}^{\infty} e^{ist} \prod_1^n p(x_j - t)\, dt \Big/ \left| \int_{-\infty}^{\infty} e^{ist} \prod_1^n p(x_j - t)\, dt \right|. \tag{3}$$

It follows from (3) that the following is a necessary condition for the existence of g. The set

$$\left\{(x_1, \dots, x_n) : \prod_1^n p(x_j - t) = \prod_1^n p(x_j + t) \text{ for almost all } t\right\}$$

is a set of zeros of a homogeneous estimator. This condition is also sufficient under the assumption of positivity of the Fourier transform of the function

$$\prod_1^n p(x_j - t)$$

for almost all $(x_1, \dots, x_n)$.

Definition 2. The family $\{p_\theta\}$ (or the density p itself) is called **strongly symmetric** if the set

$$\left\{(x_1, \dots, x_n) \prod_1^n p(x_i - t) = \prod_1^n p(x_j + t) \text{ for all } t\right\} \tag{4}$$

is a set of zeros of a homogeneous estimator. It is clear that in this case also $p(c - t) = p(c + t)$ for all t and some constant c, which in what follows we shall take to be equal to zero.

If W is a differentiable loss function satisfying the symmetry condition $W(t) = W(-t)$, then the optimal homogeneous estimate has the set of zeros

$$\left\{(x_1, \dots, x_n) : \int W'(t) \prod_1^n p(x_j - t)\, dt = 0\right\}$$

(see (1)). It is clear that if p is a strongly symmetric density, then the two estimates coincide. The estimate defined by (3) also coincides with the maximum likelihood estimate. We also note that strongly symmetric families may be defined as those for which the optimal confidence interval for θ of constant length $2b$ has the form $(g(x) - b,\ g(x) + b)$, independently of b.

Let $p(t) > 0$ and $\varphi(t) = -\log p(t)$. Then (3) for $n = 3$ (to which the general case reduces) means

$$R(s, u) + R(t, u) = R(\psi(s, t), u), \tag{5}$$

where $R(s, u) = \varphi(s + u) - \varphi(s - u)$ and $g(s, t, \psi(s, t)) = 0$. For arbitrary s, t, and v, we further have

$$R(s, u) + R(t, u) + R(v, u) = R(\psi(\psi(s, t), v), u).$$

On the other hand,

$$R(s,u)+R(t,u)+R(v,u)=R(\psi(s,\psi(t,v)),u),$$

so that

$$R(\psi(\psi(s,t),v),u)=R(\psi(s,\psi(t,v)),u),$$

$$\psi(\psi(s,t),v)=\psi(s,\psi(t,v)).$$

But it is known ((2), p. 176) that then there exists a monotone function h such that

$$\psi(s,t)=h\big(h^{-1}(s)+h^{-1}(t)\big).$$

Returning to (5), we obtain

$$R(h(s),u)+R(h(t),u)=R(h(s+t),u);$$

hence $R(h(s),u)=ah(s)$, whence we derive $R(s,t)=h^{-1}(s)h^{-1}(t)$. Thus we have arrived at the functional equation

$$\varphi(s+t)-\varphi(s-t)=h^{-1}(s)h^{-1}(t),$$

all measurable solutions of which that have probabilistic meaning are equal to ((2), p. 131)

$$\varphi(t)=\begin{cases}\alpha\ \mathrm{ch}\,\beta t, & \alpha,\ \beta>0,\\ \gamma t^2, & \gamma>0.\end{cases} \tag{6}$$

It is not difficult to show that if $\beta\to 0$, $\alpha\beta^2\to 1/\sigma^2$, then

$$p(t)=\frac{2\beta}{K_0(\alpha)}\exp\{-\alpha\ \mathrm{ch}\,\beta t\}\to\frac{1}{\sqrt{2\pi\sigma}}\exp\left\{-\frac{t^2}{2\sigma^2}\right\}.$$

Here

$$K_0(\alpha)=\int_0^\infty e^{-\alpha\ \mathrm{ch}\,t}\,dt.$$

If, however, $\beta \to \infty$, $-\frac{1}{\beta} \ln \alpha \to \delta$, then the indicated density converges to the density of the rectangular law on the interval $[-\delta, \delta]$. It can also be proved that the latter is the only strongly symmetric law among those whose density vanishes on the union of a finite number of intervals.

Thus, the following is true:

Theorem 1. Every strongly symmetric law whose density $\cdots$

tends to zero on the union of a finite number of intervals, is determined from (6), or is rectangular.

The optimal homogeneous estimator for strongly symmetric families has the form

$$g_\beta(x_1, \dots, x_n) = \begin{cases} \bar{x}, & \beta = 0, \\ \dfrac{1}{2\beta} \ln \left(\sum_1^n e^{\beta x_j} \Big/ \sum_1^n e^{-\beta x_j} \right), & 0 < \beta < \infty, \\ \dfrac{x_{\max} + x_{\min}}{2}, & \beta = \infty. \end{cases}$$

Let us note that the property of this estimator of being a maximum-likelihood estimator of the parameter θ characterizes our distribution. An analogous result is also true for confidence estimation of θ and of certain other loss functions.

Theorem 2. If P is some symmetric measure for which

$$\int_{-\infty}^{\infty} e^{2\beta t} \, dP(t) < \infty$$

and $0 \leq \beta < \infty$, then

$$P_\theta\{\sqrt{n}(g_\beta(x_1, \dots, x_n) - \theta) \in B\} \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}\sigma} \int_B e^{-t^2/2\sigma^2} \, dt,$$

where

$$B \in \mathfrak{B}, \qquad \sigma^2 = \int \operatorname{sh}^2 \beta t \, dP(t) \Big/ \beta^2 \left(\int \operatorname{ch} \beta t \, dP(t) \right)^2.$$

We observe that under the assumptions of Theorem 2 the asymptotic variance of the estimator g_β, for sufficiently small β, is smaller than that of $\bar{x}$ if the excess of the distribution $\tilde{E} = \alpha_4/\alpha_2^2 - 3$ is negative.

Let us also note that the question of describing strongly symmetric families can be posed for other groups as well. Thus, for the group of rotations of the circle, all strongly symmetric families have the form

$$p(t) = \frac{1}{I_0(a)} \exp\{a \cos t\}, \qquad 0 \le t < 2\pi,\ a > 0.$$

The corresponding estimator is equal to

$$g(x_1, \dots, x_n) = \arg(e^{ix_1} + \dots + e^{ix_n}).$$

The analogue of Theorem 2 in this case is

Theorem 3. If

$$c_1 = \int_0^{2\pi} e^{-it}\, dP(t) \ne 0,$$

then

$$P\{[\sqrt{n}](\arg(e^{ix_1} + \dots + e^{ix_n}) - \arg c_1) \in B\} \underset{n\to\infty}{\longrightarrow} \int_B \vartheta(t)\, dt. \tag{7}$$

Here

$$\vartheta(t) = \sum_{k=-\infty}^{\infty} e^{ikt - \sigma^2 k^2},$$

and in (7)

$$\sigma^2 = \int \sin^2 t\, dP(t) \Big/ \left(\int \cos t\, dP(t)\right)^2.$$

Theorem 3 can be used for constructing asymptotically most powerful confidence intervals for the shift parameter of distributions on the circle, which is important in biological and geological applications ([3], p. 60).

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

Received
27 V 1969

REFERENCES

1. R. H. **Farrell**, Ann. Math. Stat., **35**, 3 (1964).

2. J. **Aczel**, *Vorlesungen uber Funktionalgleichungen und ihre Anwendungen*, Basel, 1961.

3. A. B. **Vistelius**, *Structural Diagrams*, Publishing House of the Academy of Sciences of the USSR, 1958.

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

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