DISCRETE ANALOGS OF FISHER INFORMATION AND LEAST FAVORABLE DISTRIBUTIONS IN ROBUST ESTIMATION OF LOCATION PARAMETERS

N. O. Vil'chevskii and G. L. Shevlyakov (St. Petersburg)

UDC 519.2

The analog of Fisher information for discrete lattice distributions is introduced. The variational problems of minimization of Fisher information are solved for the discrete analogs of the classes of finite distributions, nonsingular densities and ε-contaminated distributions.

One of the basic approaches to the synthesis of robust estimation procedures is the minimax principle or what Germeyer [1] calls "the guaranteed result principle." In this case, in a given class of densities the least favorable one which minimizes the Fisher information is determined. The unknown parameters are then estimated by means of the maximum likelihood method for this density [2]. Such an approach makes it possible to construct robust statistical procedures that are stable with respect to deviations from a priori assumptions about the distribution in the case of prior uncertainty of the probability models being used.

The form of the solution, obtained by the minimax approach, essentially depends upon the class of densities. As a rule, continuous symmetric distributions are considered. Many applications of data processing deal, however, with groups of equal-valued data obtained from measurements. Furthermore, measuring instruments usually output rounded data due to their discrete scales. In this case continuous models are inadequate, so we shall consider their discrete analogs.

Consider some classes of continuous densities used in robust estimation. The problem of synthesis of the robust estimation algorithm for the location parameter of the density \( f(x, \theta) = f(x - \theta) \) is reduced to minimization of the Fisher information

\[
I(f) = \int_{-\infty}^{\infty} \left( \frac{f'}{f} \right)^2 f \, dz.
\]

For all classes of distributions \( \mathcal{F} \) the following conditions are common [2]:

\[
f(x) \geq 0, \quad f(-x) = f(x), \quad \int_{-\infty}^{\infty} f(x) \, dx = 1. \tag{2}
\]

Depending on additional restrictions on the class \( \mathcal{F} \), different forms of the density \( f^* \) may result.

It follows from [2, 3] that, for a class of finite distributions,

\[
\mathcal{F}_1 = \{ f : \int_{-a}^{a} f(x) \, dx = 1 \},
\]

\[
f_1^*(x) = \begin{cases} a^{-1} \cos^2(\pi x/(2a)), & |x| \leq a, \\ 0, & |x| > a, \end{cases}
\]

for a class of nonsingular densities,

\[
\mathcal{F}_2 = \{ f : f(0) \geq 1/(2s) > 0 \}, \quad f_2^*(x) = (2s)^{-1} \exp\{-|x|/s\},
\]

and for a class of ε-contaminated densities.
\( \mathcal{F}_3 = \{ f : f(x) \geq (1 - \epsilon)p(x) \}, \quad 0 \leq \epsilon \leq 1, \) \( f_3^*(x) = \begin{cases} (1 - \epsilon)p(x), & |x| \leq \Delta, \\ A f_2^*(Bx), & |x| > \Delta, \end{cases} \) (5)

where \( p(x) \) is a given density; the constants \( A, B, \Delta \) are chosen to satisfy the conditions of normalization and the sewing smoothness at the point \( x = \Delta \).

Note that \( f \geq 0 \) and therefore we change variables \( f = g^2 \). The problem of constructing the density \( f^* \) is reduced to minimizing the functional

\[
\begin{align*}
I &= \int_{-\infty}^{\infty} (g'(x))^2 \, dx \to \min, \\
\int_{-\infty}^{\infty} g^2(x) \, dx &= 1.
\end{align*}
\] (6)

The following theorem provides the discrete analog of this problem.

**Theorem 1.** For the class of discrete distributions

\[
f(x) = \sum_i p_i \delta(x - i\Delta), \quad \sum_i p_i = 1, \quad p_i \geq 0,
\] (7)

where \( \Delta \) is the quantization pitch and \( \Delta(\cdot) \) is the Dirac delta function, the problem of minimization of (6) is equivalent to

\[
\begin{align*}
\sum_i \lambda_i \lambda_{i+1} &\to \max, \\
\sum_i \lambda_i^2 &= 1, \quad \lambda_i = p_i.
\end{align*}
\] (8)

**Proof.** Introduce a prelimit approximation to (7), e.g.,

\[
f_h(x) = g_h^2(x), \quad g_h(x) = \sum_i \frac{P_i^{1/2}}{(2\pi h^2)\Delta^{1/4}} \exp \left\{ -(x - i\Delta)^2/(4h^2) \right\}.
\]

Introducing the notation \( p_i = \lambda_i^2 \), we put functional (6) and the normalization condition in the form

\[
I_h = 1/h^2 - \Delta^2(4h^2)^{-1} \sum_i \sum_j \lambda_i \lambda_j (i - j)^2 \exp \left\{ -(i - j)^2\Delta^2/(8h^2) \right\},
\]

\[
\sum_i \sum_j \lambda_i \lambda_j \exp \left\{ -(i - j)^2\Delta^2/(8h^2) \right\} = 1.
\]

By passage to the limit as \( h \to 0 \) and taking the principal part of the functional, we obtain the statement of the theorem.

Note that in (8) we do not assume the density \( f \) to be symmetric. Imposing such restrictions, we can obtain discrete analogs of problems (3)-(5).

**Theorem 2.** For the class of discrete, finite, symmetric densities the solution of (8) has the form

\[
p_{-i} = p_i = \frac{\cos^2(\pi/(2(n + 1)))}{n + 1}, \quad i = 1, \ldots, n.
\] (9)

**Proof.** Introduce the extended Lagrange functional for (8)

\[
2 \sum_{i=0}^{n-1} \lambda_i \lambda_{i+1} - \mu \left( \lambda_0^2 + 2 \sum_{i=1}^{n} \lambda_i^2 - 1 \right) \to \max_{\lambda_i, \mu}.
\] (10)