ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIX, 1 (1987), pp. 85–91

S. TRYBUŁA (Wrocław)

MINIMAX ESTIMATION AND PREDICTION FOR RANDOM VARIABLES WITH BOUNDED SUM

1. In this paper the form of a minimax estimator $d = (d_1, \ldots, d_r)$ of the parameter $m = (m_1, \ldots, m_r)$, $m_i = \mathrm{E}(X'_i)$, is determined for the loss function (2), in the case when the random variables $X'_1, \ldots, X'_r$ satisfy the conditions

$$X'_1 \geqslant 0, \ldots, X'_r \geqslant 0, \quad X'_1 + \ldots + X'_r \leqslant s. \tag{1}$$

An application to random variables with hierarchical structure (see (8) and (9)) is given. A prediction problem for random variables with bounded sum is considered.

2. Let $X' = (X'_1, \ldots, X'_r)$ be a random variable satisfying the conditions

$$X'_1 \geqslant 0, \ldots, X'_r \geqslant 0, \quad X'_1 + \ldots + X'_r = s, \quad s > 0,\ r \in \{2, 3, \ldots\}.$$

Let $X^{(1)}, \ldots, X^{(n)}$, $X^{(j)} = (X_1^{(j)}, \ldots, X_r^{(j)})$, $j = 1, \ldots, n$, be independent random variables having the same distribution as X'. Write $X = (X^{(1)}, \ldots, X^{(n)})$, $m_i = \mathrm{E}(X'_i)$, $i = 1, \ldots, r$, and let $d(X) = \big(d_1(X), \ldots, d_r(X)\big)$ be an estimator of the parameter $m = (m_1, \ldots, m_r)$. The problem is to find a minimax estimator of m for the loss function

$$L(m, \hat{a}) = \sum_{i,j=1}^{r} c_{ij}(a_i - m_i)(a_j - m_j), \tag{2}$$

where $\hat{a} = (a_1, \ldots, a_r)$ is an estimate of m and the matrix $C = \|c_{ij}\|_1^r$ is nonnegative definite.

Denote

$$X_i = \sum_{j=1}^{n} X_i^{(j)}.$$

Let us consider an estimator $d = (d_1, \ldots, d_r)$ for which

$$d_i(X) = \frac{X_i + \beta_i \sqrt{n}}{n + \sqrt{n}}, \tag{3}$$

where $\beta_i \geqslant 0$, $i = 1, \ldots, r$, and $\sum_{i=1}^{r} \beta_i = s$. For such an estimator

$$(4)\quad R(m, d) = \mathrm{E}\big(L(m, d(X))\big)$$
$$= \frac{1}{(\sqrt{n}+1)^2} \sum_{i,j=1}^{r} c_{ij}[\mathrm{E}(X_i'-m_i)(X_j'-m_j)+(m_i-\beta_i)(m_j-\beta_j)]$$

$$= \frac{1}{(\sqrt{n}+1)^2} \sum_{i,j=1}^{r} c_{ij}[\mathrm{E}(X_i' X_j')-2\beta_j m_i+\beta_i \beta_j]$$

is the risk function.

But

$$\sum_{i,j=1}^{r} c_{ij} X_i' X_j' - s \sum_{i=1}^{r} c_{ii} X_i'$$
$$= \sum_{i,j=1}^{r} c_{ij} X_i' X_j' - \tfrac{1}{2} \sum_{i,j=1}^{r} c_{ii} X_i' X_j' - \tfrac{1}{2} \sum_{i,j=1}^{r} c_{jj} X_i' X_j'$$
$$= -\tfrac{1}{2} \sum_{i,j=1}^{r} (c_{ii}+c_{jj}-2c_{ij}) X_i' X_j' \leqslant 0,$$

because matrix C is nonnegative definite and $X_i' \geqslant 0$, $i = 1, \ldots, r$. Thus we obtain

$$(5)\quad R(m, d) \leqslant \frac{1}{(\sqrt{n}+1)^2}\Big[\sum_{i,j=1}^{r} c_{ij}(\beta_i \beta_j - 2\beta_j m_i) + s \sum_{i=1}^{r} c_{ii} m_i\Big].$$

Let $e_1 = (s, 0, \ldots, 0), \ldots, e_r = (0, 0, \ldots, s)$,

$$(6)\quad P(X' = e_i) = \frac{m_i}{s} \stackrel{\text{def}}{=} p_i.$$

Then

$$\mathrm{E}(X_i') = m_i, \quad \mathrm{E}(X_i' X_j') = 0 \quad \text{for} \quad i \neq j,$$
$$\mathrm{E}(X_i'^2) = s m_i$$

and for each estimator (3)

$$R(m, d) = \frac{1}{(\sqrt{n}+1)^2}\Big[\sum_{i,j=1}^{r} c_{ij}(\beta_i \beta_j - 2\beta_j m_i) + s \sum_{i=1}^{r} c_{ii} m_i\Big].$$

Suppose that there are a set $A \subset R = \{1, 2, \ldots, r\}$, $|A| \geqslant 2$, and constants $\beta_1, \ldots, \beta_r, v$ such that

$$
\begin{aligned}
&\sum_{j \in A} (c_{ii} - 2c_{ij}) \beta_j = v \quad \text{if} \quad i \in A, \\
&\sum_{j \in A} (c_{ii} - 2c_{ij}) \beta_j \leqslant v \quad \text{if} \quad i \in R - A,
\end{aligned} \tag{7}
$$

$\beta_i > 0$ for $i \in A$, $\beta_i = 0$ for $i \in R - A$, $\sum_{i=1}^{r} \beta_i = s$. It follows from [4] that such a set A and such constants always exist. For $\beta_1, \ldots, \beta_r, v$ chosen in such a way we have

$$R(m, d) = \frac{1}{(\sqrt{n}+1)^2} \Big(\sum_{i,j=1}^{r} c_{ij} \beta_i \beta_j + v \Big) = c,$$

if X' is distributed according to (6) with $m_i = 0$ if $i \in R - A$, and

$$R(m, d) \leqslant c$$

for any distribution of X'.

One can view the problem of finding a minimax estimator of the parameter $m = (m_1, \ldots, m_r)$ as the problem of determining a minimax strategy in a game against nature: the nature chooses a distribution of the random variable X', the statistician chooses an estimator d of $m = \mathrm{E}(X')$, the payoff is the risk function $R(m, d)$. Choose a mixed strategy of the nature in the following way:

(S) At first choose the parameter $p = (p_1, \ldots, p_r)$ according to the density

$$
g(p_1, \ldots, p_r) = \begin{cases} \dfrac{\Gamma\big(\sum_{i=1}^{r} \alpha_i\big)}{\Gamma(\alpha_{i_1}) \ldots \Gamma(\alpha_{i_s})} p_{i_1}^{\alpha_{i_1}-1} \ldots p_{i_s}^{\alpha_{i_s}-1} & \text{if } p_{i_k} > 0, \ \sum_{k=1}^{s} p_{i_k} = 1, \\ 0 & \text{otherwise} \end{cases}
$$

$$(A = \{i_1, \ldots, i_s\}, \ \alpha_i = (\sqrt{n}/s)\beta_i).$$

and later, choose the distribution P of X' according to (6).

It is not difficult to verify that the estimator defined by (3) and (7) is a Bayes estimator with respect to such a mixed strategy of nature, and we have proved

THEOREM 1. *Each estimator $d = (d_1, \ldots, d_r)$ with d_i defined by* (3), *where the β_i are chosen according to* (7), *is a minimax estimator of the parameter $m = (m_1, \ldots, m_r)$ for the loss function* (2). *Such a minimax estimator always exists.*

In [4] it is proved that $\beta_1^0, \ldots, \beta_r^0$ satisfying (7) are solutions to the equation

$$s\sum_{i=1}^{r} c_{ii}\beta_i^0 - \sum_{i,j=1}^{r} c_{ij}\beta_i^0\beta_j^0 = \max\left(s\sum_{i=1}^{r} c_{ii}\beta_i - \sum_{i,j=1}^{r} c_{ij}\beta_i\beta_j\right),$$

where the maximum is taken over the set of $(\beta_1, \ldots, \beta_r)$ such that $\beta_i \geqslant 0$ for $i = 1, \ldots, r$, and $\sum_{i=1}^{r} \beta_i = s$.

Taking into account the maximin strategy of nature defined in (S) one can notice that each estimator (3) with β_i satisfying (7) is a minimax estimator of the parameter $p = (p_1, \ldots, p_r)$ of the multinomial distribution

$$P(X_1 = x_1, \ldots, X_r = x_r) = \frac{n!}{x_1! \ldots x_r!} p_1^{x_1} \ldots p_r^{x_r}$$

for the loss function

$$L(p, \hat{a}) = \sum_{i,j=1}^{r} c_{ij}(a_i - p_i)(a_j - p_j)$$

if matrix C is nonnegative definite. This was proved in [4]. Our considerations are partly based on this result.

Let the random variable $X' = (X'_1, \ldots, X'_r)$ satisfy the conditions

$$X'_1 \geqslant 0, \ldots, X'_r \geqslant 0, \quad X'_1 + \ldots + X'_r \leqslant s, \quad s > 0,\ r = 1, 2, \ldots,$$

and let the loss function be given by (2). Let us define $X'_{r+1} = s - \sum_{i=1}^{r} X'_i$ and $c_{i,r+1} = 0$ for $i = 1, \ldots, r+1$. Then we are in the situation considered in this section and there exists a minimax estimator $d = (d_1, \ldots, d_r)$ of the parameter $m = (m_1, \ldots, m_r)$ of the form (3) with $\beta_i \geqslant 0$, $i = 1, \ldots, r$ and $\sum_{i=1}^{r} \beta_i \leqslant s$. In the case $r = 1$ the problem was solved in [1] ($\beta_1 = s/2$).

3. Let $X' = (X'_{11}, \ldots, X'_{1s_1}, \ldots, X'_{r1}, \ldots, X'_{rs_r})$ be a random variable satisfying the conditions

$$\sum_{i=1}^{r}\sum_{k=1}^{s_i} X'_{ik} = s, \qquad X'_{ik} \geqslant 0. \tag{8}$$

Let $m_{ik} = \mathrm{E}(X'_{ik})$ and let $X_{ik}^{(j)}$, $X^{(j)}$, X, X_{ik} ($i = 1, \ldots, r$, $k = 1, \ldots, s_i$; $j = 1, \ldots, n$) be defined as the corresponding random variables is Section 2. Let the loss function be of the form

$$L(m, \hat{a}) = \sum_{i=1}^{r} c_i(a_i - m_i)^2 + \sum_{i=1}^{r}\sum_{k=1}^{s_i} c_{ik}(a_{ik} - m_{ik})^2, \tag{9}$$

where

$$m_i = \sum_{k=1}^{s_i} m_{ik}, \quad a_i = \sum_{k=1}^{s_i} a_{ik},$$

$c_i \geq 0$, $c_{ik} > 0$ and $\hat{a} = (a_{11}, \ldots, a_{1s_1}, \ldots, a_{r1}, \ldots, a_{rs_r})$ is an estimate of $m = (m_{11}, \ldots, m_{1s_1}, \ldots, m_{r1}, \ldots, m_{rs_r})$. Consider the estimator $d = (d_{11}, \ldots \ldots, d_{1s_1}, \ldots, d_{r1}, \ldots, d_{rs_r})$ of m for which

$$d_{ik} = \frac{X_{ik} + \beta_{ik}\sqrt{n}}{n+\sqrt{n}}, \tag{10}$$

where

$$\beta_{ik} \geq 0, \quad \sum_{i=1}^{r} \sum_{k=1}^{s_i} \beta_{ik} = s.$$

Denote

$$X_i = \sum_{k=1}^{s_i} X_{ik}, \quad \beta_i = \sum_{k=1}^{s_i} \beta_{ik}.$$

Then

$$d_i(X) = \sum_{k=1}^{s_i} d_{ik}(X) = \frac{X_i + \beta_i\sqrt{n}}{n+\sqrt{n}}$$

is the corresponding estimator of m_i. From Theorem 1 it follows that there exists an estimator d of m, with d_{ik} given by (10), which is minimax. In paper [3] a method of determining the constants in the case $s = 1$ is given (it is done for the multinomial distribution). I think that a modification of this method may be used when

$$\sum_{i=1}^{r} \sum_{k=1}^{s_i} X'_{ik} \leq s, \quad X'_{ik} \geq 0.$$

When all $c_{ik} = 0$ in (9) a simple method to determine β_i is given in [2]. This was also found for the multinomial distribution.

4. Let $X' = (X'_1, \ldots, X'_r)$ be a random variable satisfying the conditions (1) and let $X^{(1)}, \ldots, X^{(n_1)}; Y^{(1)}, \ldots, Y^{(n_2)}, X^{(j)} = (X_1^{(j)}, \ldots, X_r^{(j)})$, $j = 1, \ldots, n_1$, $Y^{(k)} = (Y_1^{(k)}, \ldots, Y_r^{(k)})$, $k = 1, \ldots, n_2$, be independent random variables having the same distribution as X'. Let $X = (X^{(1)}, \ldots, X^{(n_1)})$, $Y = (Y^{(1)}, \ldots, Y^{(n_2)})$,

$$X_i = \sum_{j=1}^{n_1} X_i^{(j)}, \quad Y_i = \sum_{k=1}^{n_2} Y_i^{(k)},$$

$$Y = (Y_1, \ldots, Y_r).$$

The problem is to find a minimax predictor of Y, based on X, for the loss function

$$L(Y, \hat{a}) = \sum_{i,j=1}^{r} c_{ij}(a_i - Y_i)(a_j - Y_j), \tag{11}$$

where $\hat{a} = (a_1, \ldots, a_r)$ is a prediction of Y and the matrix $C = \|c_{ij}\|_1^r$ is nonnegative definite.

Consider a predictor $d = (d_1, \ldots, d_r)$, where

$$d_i(X) = aX_i + b_i \quad (i = 1, \ldots, r). \tag{12}$$

In this case

$$\begin{aligned} R(m, d) &= \mathrm{E}\big(L(Y, d(X))\big) \\ &= \sum_{i,j=1}^{r} c_{ij}\{(a^2 n_1 + n_2)\mathrm{E}(X_i' - m_i)(X_j' - m_j) \\ &\qquad + [b_i - (n_2 - an_1)m_i][b_j - (n_2 - an_1)m_j]\}. \end{aligned}$$

Assume that

$$a^2 n_1 + n_2 = (n_2 - an_1)^2, \tag{13}$$

$$b_i = (n_2 - an_1)\beta_i, \tag{14}$$

where $\beta_i \geqslant 0$, $i = 1, \ldots, r$, and $\sum_{i=1}^{r} \beta_i = s$. For the predictor d satisfying these conditions we have

$$\begin{aligned} R(m, d) &= (n_2 - an_1)^2 \sum_{i,j=1}^{r} c_{ij}[\mathrm{E}(X_i' - m_i)(X_j' - m_j) + (m_i - \beta_i)(m_j - \beta_j)] \\ &\leqslant (n_2 - an_1)^2 \Big[\sum_{i,j=1}^{r} c_{ij}(\beta_i \beta_j - 2\beta_j m_i) + s\sum_{i=1}^{r} c_{ii} m_i\Big] \end{aligned}$$

(see (4) and (5)).

Equation (13) holds surely if

$$a = \begin{cases} \dfrac{n_1 n_2 - \sqrt{n_1 n_2 (n_1 + n_2 - 1)}}{n_1(n_1 - 1)} & \text{for} \quad n_1 > 1, \\ (n_2 - 1)/2 & \text{for} \quad n_1 = 1. \end{cases} \tag{15}$$

Let us notice that $a = 0$ if $n_2 = 1$.

On the other hand, for any predictor d the risk function may be presented as follows

$$R(m, d) = \mathrm{E}\Big[\sum_{i,j=1}^{r} c_{ij}(d_i(X) - Y_i)(d_j(X) - Y_j)\Big]$$

$$= \mathrm{E}\Big[\sum_{i,j=1}^{r} c_{ij}\big(d_i(X)-n_2 m_i\big)\big(d_j(X)-n_2 m_j\big)\Big] + \sum_{i,j=1}^{r} c_{ij}\,\mathrm{E}(Y_i-n_2 m_i)(Y_j-n_2 m_j),$$

where the second term is independent of d. Taking this into account one can prove that for $n_2 > 1$ the predictor d, determined by (12), (14), and (15), with β_i satisfying conditions (7), is Bayesian with respect to the mixed (maximin) strategy of nature defined by (S) with

$$\alpha_i = \frac{n_2 - an_1}{a}\beta_i \quad (i = 1, \ldots, r).$$

For $n_2 = 1$, to define the strategy of nature, one can choose with probability 1 in (S) the parameter $p = (p_1, \ldots, p_r)$ equal to $(1/s)(\beta_1, \ldots, \beta_r)$ obtaining the same conclusion. Then, similarly as in Section 2, we obtain

THEOREM 2. *Each predictor* $d = (d_1, \ldots, d_r)$ *with*

$$d_i(X) = aX_i + (n_2 - an_1)\beta_i \quad (i = 1, \ldots, r),$$

where a *is given by* (15) *and* β_i *are chosen according to* (7), *is a minimax predictor of the random variable* $Y = (Y_1, \ldots, Y_r)$ *for the loss function* (11). *Such a minimax predictor always exists.*

In a similar way as in Section 2 one can find the conditions for the minimax predictor of Y in the case

$$X'_1 \geqslant 0, \ldots, X'_r \geqslant 0, \quad X'_1 + \ldots + X'_r \leqslant s.$$

References

[1] J. L. Hodges and E. L. Lehmann, *Some problems in minimax point estimation*, Ann. Math. Statist. 21 (1950), p. 182-196.

[2] S. Trybuła, *Some problems in simultaneous minimax estimation*, ibid. 29 (1958), p. 245-253.

[3] —, *Some investigations in minimax estimation theory*, Dissertationes Mathematicae 240, Warszawa 1985, 42 pp.

[4] M. Wilczyński, *Minimax estimation for multinomial and multivariate hypergeometric distribution*, Sankhyā 47, Series A, Pt. 1, p. 128–132.

INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF WROCŁAW
50-370 WROCŁAW

Received on 1983.02.08