ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIX, 2 (1987), pp. 195–201

S. TRYBUŁA (Wrocław)

MINIMAX PREDICTION OF THE SUM OF PROCESSES

In the paper the minimax predictor of the sum of multinomial variables with different parameters is determined for the general quadratic loss function. Based on this result the minimax predictor of the sum of $m_2 n_2$ processes satisfying condition (26) is found, where m_2 is the number of processes with different distributions and n_2 is the number of their independent copies. The predictor is based on observations of $m_1 n_1$ processes satisfying condition (25) and similarly defined. These two families of processes are connected by condition (16). All processes considered are independent and the loss function, defined by (17), is also general quadratic.

1. Let $\bar{X}_1, \ldots, \bar{X}_{m_1}$, $\bar{X}_i = (X_{i1}, \ldots, X_{ir})$, be independent random variables having multinomial distributions with parameters n_1 and $(p_{i1}, \ldots, p_{ir})$, respectively, where

$$(p_{i1}, \ldots, p_{ir}) \in \Lambda = \{\lambda'_1 \geqslant 0, \ldots, \lambda'_r \geqslant 0, \sum_{j=1}^{r} \lambda'_j = 1\}.$$

Let $\bar{Y}_1, \ldots, \bar{Y}_{m_2}$, $\bar{Y}_i = (Y_{i1}, \ldots, Y_{ir})$, be independent random variables having multinomial distributions with parameters n_2 and $(q_{i1}, \ldots, q_{ir}) \in \Lambda$, respectively. Write

$$\bar{X} = (\bar{X}_1, \ldots, \bar{X}_{m_1}),\ p_j = \frac{1}{m_1} \sum_{i=1}^{m_1} p_{ij},\ p = (p_1, \ldots, p_r),\ \bar{p} = (p_{11}, \ldots, p_{m_1 r}),$$

$$\bar{Y} = (\bar{Y}_1, \ldots, \bar{Y}_{m_2}),\ q_j = \frac{1}{m_2} \sum_{i=1}^{m_2} q_{ij},\ q = (q_1, \ldots, q_r),\ \bar{q} = (q_{11}, \ldots, q_{m_2 r})$$

and suppose that the random variables $\bar{X}$ and Y are independent and that $p = q$. Let

$$X = (X_1, \ldots, X_r), \quad X_j = \frac{1}{m_1} \sum_{i=1}^{m_1} X_{ij}, \quad Y = (Y_1, \ldots, Y_r), \quad Y_j = \frac{1}{m_2} \sum_{i=1}^{m_2} Y_{ij}.$$

We want to predict Y observing $\bar{X}$. Let $d(\bar{X}) = (d_1(\bar{X}), \ldots, d_r(\bar{X}))$ be a predictor of Y. The problem is to determine a minimax predictor of Y for the loss function

$$L(Y, d) = \sum_{i,j=1}^{r} c_{ij}(d_i - Y_i)(d_j - Y_j), \tag{1}$$

where the matrix $\|c_{ij}\|_1^r$ is nonnegative definite.

The case $n_2 = 1$ is of special interest.

Let us consider a predictor $d = (d_1, \ldots, d_r)$ for which

$$d_i(\bar{X}) = \alpha X_i + (n_2 - \alpha n_1)\beta_i, \qquad \beta = (\beta_1, \ldots, \beta_r) \in \Lambda, \tag{2}$$

where

$$\alpha = \begin{cases} \dfrac{m_1}{n_1} \, \dfrac{n_1 n_2 - \sqrt{\dfrac{n_1 n_2}{m_1 m_2}(n_1 m_1 + n_2 m_2 - 1)}}{n_1 m_1 - 1} & \text{if } n_1 m_1 \neq 1, \\[2ex] \dfrac{n_2 m_2 - 1}{2m_2} & \text{if } m_1 = n_1 = 1. \end{cases} \tag{3}$$

Since α satisfies the condition

$$\alpha^2 \frac{n_1}{m_1} + \frac{n_2}{m_2} = (n_2 - \alpha n_1)^2, \tag{4}$$

the risk function for the predictor (2) is

$$R(\bar{p}, \bar{q}, d) = \mathrm{E}\big(L(Y, d(\bar{X}))\big) = \sum_{i,j=1}^{r} c_{ij} \Big\{ -\alpha^2 \frac{n_1}{m_1^2} \sum_{k=1}^{m_1} p_{ki} p_{kj} - \frac{n_2}{m_2^2} \sum_{k=1}^{m_2} q_{ki} q_{kj} + (n_2 - \alpha n_1)^2 (\beta_i - p_i)(\beta_j - p_j) \Big\} + (n_2 - \alpha n_1)^2 \sum_{i=1}^{r} c_{ii} p_i.$$

Suppose that $p_1 \geqslant 0, \ldots, p_r \geqslant 0$ are fixed. The quadratic form

$$W = \sum_{i,j=1}^{r} c_{ij} \sum_{k=1}^{m} x_{ki} x_{kj} \tag{5}$$

of the variables x_{kj}, $m^{-1} \sum_{k=1}^{m} x_{kj} = p_j$, is nonnegative definite. It attains its minimum if $x_{kj} = p_j$. Then, by (4), we have

$$R(\bar{p}, \bar{q}, d) \leqslant (m_2 - \alpha n_1)^2 \Big\{ \sum_{i,j=1}^{r} c_{ij}(\beta_i \beta_j - 2\beta_i p_j) + \sum_{i=1}^{r} c_{ii} p_i \Big\} \overset{\mathrm{df}}{=} R_1(p, \beta) \tag{6}$$

with an equality if $p_{kj} = q_{kj} = p_j$.

Let $p \in \Lambda$. We have

$$\min_{\beta \in \Lambda} R_1(p, \beta) = R_1(p, p) = -\sum_{i,j=1}^{r} c_{ij} p_i p_j + \sum_{i=1}^{r} c_{ii} p_i. \tag{7}$$

Denote by $p_0 = (p_1^0, \ldots, p_r^0) \in \Lambda$ a point for which

(8) $$R_1(p_0, p_0) = \max_{p\in\Lambda} R_1(p, p)$$

and use the estimator $d_0 = (d_1^0, \ldots, d_r^0)$ defined in (2) and (3) with $\beta_i = p_i^0$. We have

(9) $$\sup_{(\bar{p},\bar{q})} R(\bar{p}, \bar{q}, d_0) \leqslant \sup_{p\in\Lambda} R_1(p, p_0) = \inf_{\beta\in\Lambda} R_1(p_0, \beta) = R_1(p_0, p_0).$$

The first equality results from the paper [6], where it was obtained by an application of Sion's theorem (the function $R_1(p, \beta)$ is convex and continuous in β and concave and continuous in p, and Λ is a convex compact set).

To determine p_0, Theorem 3.5.4 in Karmanov [2] is used to obtain the following corollary: A point $p_0 = (p_1^0, \ldots, p_r^0) \in \Lambda$ is a solution of (8) if and only if there are a set $A \in \{1, \ldots, r\} = R$ and a constant v such that

(10) $$\begin{aligned} &\sum_{j\in A} (c_{ii} - 2c_{ij}) p_j^0 = v \quad \text{if } i \in A, \\ &\sum_{j\in A} (c_{ii} - 2c_{ij}) p_j^0 \leqslant v \quad \text{if } i \in A, \end{aligned}$$

$p_j^0 > 0$ for $j \in A$, $\sum_{j\in A} p_j^0 = 1$.

$|A| = 1$ if and only if $c_{ij} = c$ for $i, j \in R$.

Suppose that $p_{kj} = q_{kj} = p_j$ and let $A = \{i_1, \ldots, i_s\}$. Applying formulae (10) to (6), we obtain

(11) $$R(\bar{p}, \bar{q}, d_0) = (n_2 - \alpha n_1)^2 \sum_{k,l=1}^{s} c_{i_k i_l} (p_{i_k}^0 p_{i_l}^0 + v) = R_1(p_0, p_0)$$

if $\bar{p}$ and $\bar{q}$ satisfy the conditions $p_j \geqslant 0$ and $\sum_{j\in A} p_j = 1$.

Denote by π_0 the a priori distribution of the parameter $(\bar{p}, \bar{q})$ defined as follows:

$$P(p_{ij} = q_{kj} = p_j,\ i = 1, \ldots, m_1,\ k = 1, \ldots, m_2,\ j = 1, \ldots, r) = 1,$$

where the density h of $(p_1, \ldots, p_r)$ is

(12) $$h(p_1, \ldots, p_r) = \begin{cases} \dfrac{\Gamma\left(m_1 \alpha^{-1} (n_2 - \alpha n_1)\right)}{\Gamma(\alpha_{i_1}) \ldots \Gamma(\alpha_{i_s})} p_{i_1}^{\alpha_{i_1} - 1} \ldots p_{i_s}^{\alpha_{i_s} - 1} & \\ \qquad\qquad \text{for } p_{i_j} \geqslant 0,\ j = 1, \ldots, s,\ \sum_{j=1}^{s} p_{i_j} = 1, \\ 0 \qquad\qquad \text{otherwise,} \end{cases}$$

$$\alpha_i = m_1 \alpha^{-1} (n_2 - \alpha n_1) p_i^0,$$

with the exception of the case $m_1 = m_2 = n_1 = n_2 = 1$ for which π_0 is defined by assuming that $P(p_i = p_i^0,\ i = 1, \ldots, r) = 1$.

The predictor d_0 is Bayes with respect to π_0 and, moreover, by (11), we have

(13) $$r(\pi_0, d_0) = R_1(p_0, p_0),$$

where r is the Bayes risk.

Formulae (9) and (13) prove that the predictor d_0 is minimax.

2. Let $\bar{X}_1, \ldots, \bar{X}_{m_1}$, $\bar{X}_i = (X_{i1}, \ldots, X_{ir})$, be independent discrete-time processes satisfying the conditions

(14) $$X_{ij} \geqslant 0,\ j = 1, \ldots, r, \quad \sum_{j=1}^{r} X_{ij} = z_1 \quad (i = 1, \ldots, m_1),$$

$z_1 > 0$, $r \geqslant 2$. Let $\bar{Y}_1, \ldots, \bar{Y}_{m_2}$, $\bar{Y}_i = (Y_{i1}, \ldots, Y_{ir})$, be independent processes satisfying

(15) $$Y_{ij} \geqslant 0,\ j = 1, \ldots, r, \quad \sum_{j=1}^{r} Y_{ij} = z_2 \quad (i = 1, \ldots, m_2).$$

Let $\bar{X} = (\bar{X}_1, \ldots, \bar{X}_{m_1})$, $\bar{Y} = (\bar{Y}_1, \ldots, \bar{Y}_{m_2})$, and suppose that $\bar{X}$ and $\bar{Y}$ are independent. Let $X^{(1)}, \ldots, X^{(n_1)}$ and $Y^{(1)}, \ldots, Y^{(n_2)}$ be independent processes having the same distributions as $\bar{X}$ and $\bar{Y}$, respectively. Put

$$\hat{X} = (X^{(1)}, \ldots, X^{(n_1)}), \quad X^{(k)} = (X_1^{(k)}, \ldots, X_{m_1}^{(k)}), \quad X_i^{(k)} = (X_{i1}^{(k)}, \ldots, X_{ir}^{(k)}),$$
$$\hat{Y} = (Y^{(1)}, \ldots, Y^{(n_2)}), \quad Y^{(k)} = (Y_1^{(k)}, \ldots, Y_{m_2}^{(k)}), \quad Y_i^{(k)} = (Y_{i1}^{(k)}, \ldots, Y_{ir}^{(k)}).$$

Let $\lambda_{ij} = \mathrm{E}(X_{ij})$, $\mu_{ij} = \mathrm{E}(Y_{ij})$ and put

$$\lambda_j = \frac{1}{m_1} \sum_{i=1}^{m_1} \lambda_{ij}, \quad \mu_j = \frac{1}{m_2} \sum_{i=1}^{m_2} \mu_{ij}.$$

The basic assumption connecting the processes $\bar{X}$ and $\bar{Y}$ is

(16) $$\lambda_i/z_1 = \mu_i/z_2 \quad (i = 1, \ldots, r).$$

Write

$$X_j' = \frac{1}{m_1} \sum_{i=1}^{m_1} \sum_{k=1}^{n_1} X_{ij}^{(k)}, \quad Y_j' = \frac{1}{m_2} \sum_{i=1}^{m_2} \sum_{k=1}^{n_2} Y_{ij}^{(k)}$$

and let $d(\hat{X}) = (d_1(\hat{X}), \ldots, d_r(\hat{X}))$ be a predictor of $Y' = (Y_1', \ldots, Y_r')$. We are looking for a minimax predictor of Y' for the loss function

(17) $$L(Y', d) = \sum_{i,j=1}^{r} c_{ij}(d_i - Y_i')(d_j - Y_j'),$$

where $\|c_{ij}\|_1^r$ is nonnegative definite.

Let us consider a predictor $d=(d_1, \ldots, d_r)$ for which

$$d_i(\hat{X})=\frac{z_2}{z_1}[\alpha X_i'+(n_2-\alpha n_1)\beta_i], \tag{18}$$

where α is given by (3), and $\beta/z_1=(\beta_1/z_1, \ldots, \beta_r/z_1)\in \Lambda$. For this predictor the risk function is

(19)

$$R(F, d)=\mathrm{E}\big(L(Y', d(\hat{X}))\big)$$

$$=\sum_{i,j=1}^{r} c_{ij}\left\{\left(\frac{z_2}{z_1}\right)^2\left[\alpha^2\frac{n_1}{m_1^2}\sum_{k=1}^{m_1}\mathrm{E}(X_{ki}-\lambda_{ki})(X_{kj}-\lambda_{kj})+\right.\right.$$

$$\left.\left.+(n_2-\alpha n_1)^2(\beta_i-\lambda_i)(\beta_j-\lambda_j)\right]+\frac{n_2}{m_2^2}\sum_{k=1}^{m_2}\mathrm{E}(Y_{ki}-\mu_{ki})(Y_{kj}-\mu_{kj})\right\},$$

where F is the distribution of $(\bar{X}, \bar{Y})$.

But in view of (14) we have

$$\sum_{i,j=1}^{r} c_{ij}X_{ki}X_{kj}$$

$$=-\frac{1}{2}\sum_{i,j=1}^{r}(c_{ii}+c_{jj}-2c_{ij})X_{ki}X_{kj}+z_1\sum_{i=1}^{r}c_{ii}X_{ki}\leqslant z_1\sum_{i=1}^{r}c_{ii}X_{ki}.$$

Similarly,

$$\sum_{i,j=1}^{r}c_{ij}Y_{ki}Y_{kj}\leqslant z_2\sum_{i=1}^{r}c_{ii}Y_{ki}.$$

Then

(20) $R(F, d)$

$$\leqslant\sum_{i,j=1}^{r}c_{ij}\left\{\left(\frac{z_2}{z_1}\right)^2\left[-\alpha^2\frac{n_1}{m_1^2}\sum_{k=1}^{m_1}\lambda_{ki}\lambda_{kj}+(n_2-\alpha n_1)^2(\beta_i-\lambda_i)(\beta_j-\lambda_j)\right]-\right.$$

$$\left.-\frac{n_2}{m_2^2}\sum_{k=1}^{m_2}\mu_{ki}\mu_{kj}\right\}+\frac{z_2^2}{z_1}\left(\alpha^2\frac{n_1}{m_1^2}+\frac{n_2}{m_2}\right)\sum_{i=1}^{r}c_{ii}\lambda_i$$

$$\leqslant\left(\frac{z_2}{z_1}\right)^2(n_2-\alpha n_1)^2\left\{\sum_{i,j=1}^{r}c_{ij}(\beta_i\beta_j-2\beta_i\lambda_j)+z_1\sum_{i=1}^{r}c_{ii}\lambda_i\right\}\stackrel{\text{df}}{=}c(\beta, \lambda),$$

$$\lambda=(\lambda_1, \ldots, \lambda_r).$$

Let $e_1=(1, 0, \ldots, 0)$, $e_2=(0, 1, \ldots, 0)$, $\ldots$, $e_r=(0, 0, \ldots, 1)$, and let $U=(U_1, \ldots, U_r)$ and $V=(V_1, \ldots, V_r)$ be processes distributed according to the laws

$$P(U=z_1 e_i)=\lambda_i/z_1\stackrel{\text{df}}{=}p_i, \tag{21}$$

$$P(V = z_2 e_i) = \mu_i / z_2 = p_i, \tag{22}$$

$p = (p_1, \ldots, p_r) \in \Lambda$. Let the processes $\bar{X}_1, \ldots, \bar{X}_{m_1}$ and $\bar{Y}_1, \ldots, \bar{Y}_{m_2}$ be independent and have the same distributions as U and V, respectively. In this case, for the predictor (18) we have

$$R(F, d) = c(\beta, \lambda). \tag{23}$$

Denote by $\mathscr{F}$ the family of distributions F_p, $p \in \Lambda$, of the process

$$(\bar{X}, \bar{Y}) = (\bar{X}_1, \ldots, \bar{X}_{m_1}, \bar{Y}_1, \ldots, \bar{Y}_{m_2})$$

by assigning the distribution (21) to the first m_1 components and the distribution (22) to the last m_2 ones. Let $\bar{\pi}_0$ be the a priori distribution on $\mathscr{F}$ determined by the distribution π_0 defined in (12). Let the vector $\beta_0 = (\beta_1^0, \ldots, \beta_r^0)$, $\beta^0/z_1 \in \Lambda$, be a solution of the equation

$$-\sum_{i,j=1}^{r} c_{ij} \beta_i^0 \beta_j^0 + z_1 \sum_{i=1}^{r} c_{ii} \beta_i^0 = \max_{\beta/z_1 \in \Lambda} \Big(-\sum_{i,j=1}^{r} c_{ij} \beta_i \beta_j + z_1 \sum_{i=1}^{r} c_{ii} \beta_i \Big). \tag{24}$$

The predictor $d_0 = (d_1^0, \ldots, d_r^0)$ defined in (18) with $\beta = \beta^0 = z_1 p_0$ is Bayes with respect to $\bar{\pi}_0$. Then from the results of Section 1 and formulae (20) and (23) it follows that d_0 is minimax.

Let the processes $X_1, \ldots, X_{m_1}$, $X_i = (X_{i1}, \ldots, X_{ir})$, satisfy the conditions

$$X_{ij} \geqslant 0, \; j = 1, \ldots, r, \quad \sum_{i=1}^{r} X_{ij} \leqslant z_1 \quad (i = 1, \ldots, m_1), \tag{25}$$

$z_1 > 0$, let $\bar{Y}_1, \ldots, \bar{Y}_{m_2}$, $\bar{Y}_i = (Y_{i1}, \ldots, Y_{ir})$, be processes satisfying

$$Y_{ij} \geqslant 0, \; j = 1, \ldots, r, \quad \sum_{i=1}^{r} Y_{ij} \leqslant z_2 \quad (i = 1, \ldots, m_2), \tag{26}$$

$z_2 > 0$, and suppose that $\bar{X}_1, \ldots, \bar{X}_{m_1}, \bar{Y}_1, \ldots, \bar{Y}_{m_2}$ are independent. Let the loss function be given by (17). Let us put

$$X_{ir+1} = z_1 - \sum_{j=1}^{r} X_{ij}, \quad i = 1, \ldots, m_1,$$

$$Y_{ir+1} = z_2 - \sum_{j=1}^{r} Y_{ij}, \quad i = 1, \ldots, m_2,$$

and let $c_{ir+1} = 0$, $i = 1, \ldots, r+1$. Then we are in the position considered in this section and there is a minimax predictor $d = (d_1, \ldots, d_r)$ of the form (18) with $\beta_i \geqslant 0$, $i = 1, \ldots, r$, $\sum_{i=1}^{r} \beta_i \leqslant z_1$, where the β_i are obtained by applying equality (24) to the new situation.

In the paper we generalized the results obtained in [1], [3], [4] and [6]. Some estimation problems for the sum of processes are considered in [5].

References

[1] J. L. Hodges and E. L. Lehmann, *Some problems in minimax point estimation*, Ann. Math. Statist. 21 (1950), pp. 182–196.
[2] V. G. Karmanov, *Mathematical programming* (in Russian), 2nd edition, Nauka, Moscow 1980.
[3] S. Trybuła, *Some problems in simultaneous minimax estimation*, Ann. Math. Statist. 29 (1958), pp. 245–253.
[4] – *Some investigations in minimax estimation theory*, Dissertationes Math. 240 (1985), pp. 1–44.
[5] – *Minimax estimation for the mean of the sum of processes*, submitted to Sankhyā.
[6] M. Wilczyński, *Minimax estimation for multinomial and multivariate hypergeometric distribution*, Sankhyā, Ser. A, 47 (1985), pp. 128–132.

INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF WROCŁAW
50-370 WROCŁAW

Received on 1985.02.15