APPLICATIONES MATHEMATICAE
30,4 (2003), pp. 371–377

Stanisław Trybuła (Wrocław)

MINIMAX MUTUAL PREDICTION OF MULTINOMIAL RANDOM VARIABLES

Abstract. The problem of minimax mutual prediction is considered for multinomial random variables with the loss function being a linear combination of quadratic losses connected with prediction of particular variables. The basic parameter of the minimax mutual predictor is determined by numerical solution of some equation.

1. Introduction. Suppose that m statisticians take part in a prediction process, $m \geq 2$. Let a random variable $X_i = (X_{i1}, \ldots, X_{ir})$ be observed by the ith statistician. The random variables X_i, $i = 1, \ldots, m$, are independent and have the multinomial distribution with parameters n_i, $p = (p_1, \ldots, p_r)$. The statisticians do not know the observations of their partners but they know all the numbers n_i. They cooperate with each other. The problem solved in this paper is to determine the minimax mutual predictor

$$d = \begin{bmatrix} - & d_{12} & \ldots & d_{1m} \\ d_{21} & - & \ldots & d_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ d_{m1} & d_{m2} & \ldots & - \end{bmatrix} =: [d_{ij}]_{i,j=1}^{m},$$

where $d_{ij}(X_i) = (d_{ij}^{(1)}(X_i), \ldots, d_{ij}^{(r)}(X_i))$ is the predictor of $X_j = (X_{j1}, \ldots, X_{jr})$ used by the ith statistician to predict this random variable, $i, j = 1, \ldots, m$, $i \neq j$.

Thus each statistician observes only his "own" random variable and predicts only the others.

2000 *Mathematics Subject Classification*: Primary 62F15.
Key words and phrases: minimax mutual prediction, multinomial distribution, Bayes predictor.

The total loss of all the statisticians is

$$L(X,d)=\sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij}\sum_{l=1}^{r} c_l(d_{ij}^{(l)}(X_i)-X_{jl})^2 \tag{1}$$

where $k_{ij}\geq 0$, $\sum_{i,j=1,\, i\neq j}^{m} k_{ij}>0$, $c_l\geq 0$ are constants. Without loss of generality we can assume that $c_1\geq\ldots\geq c_r\geq 0$.

Let $R(p,d)$ be the risk function connected with the predictor d,

$$R(p,d)=E_p(L(X,d))=\sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij}\sum_{l=1}^{r} c_l E_p(d_{ij}^{(l)}(X_i)-X_{jl})^2. \tag{2}$$

We then look for a mutual predictor d_0 for which

$$\sup_p R(p,d_0)=\inf_d\sup_p R(p,d).$$

2. Determining the minimax predictor in the main case. Let the random variables X_i, $i=1,\ldots,m$, be independent and distributed according to the multinomial law

$$P_p(X_{i1}=x_{i1},\ldots,X_{ir}=x_{ir})=\frac{n_i!}{x_{i1}!\ldots x_{ir}!}\,p_1^{x_{i1}}\ldots p_r^{x_{ir}}.$$

Then the risk (2) takes the form

$$R(p,d)=\sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij}\sum_{l=1}^{r} c_l[E_p(d_{ij}^{(l)}(X_i)-n_jp_l)^2+n_jp_l(1-p_l)].$$

Let us consider the predictors

$$d_{ij}^{(l)}(X_i)=n_j\,\frac{X_{il}+\alpha_l}{n_i+\gamma},\qquad i,j=1,\ldots,m,\ i\neq j,\ l=1,\ldots,r, \tag{3}$$

where $\alpha_l\geq 0$, $\gamma>0$ and

$$\sum_{l=1}^{r}\alpha_l=\gamma. \tag{4}$$

For this mutual predictor the risk is as follows:

$$\begin{aligned}(5)\quad & R(p,d)\\ &=\sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij}\sum_{l=1}^{r} c_l\left[E_p\left(n_j\,\frac{X_{il}+\alpha_l}{n_i+\gamma}-n_jp_l\right)^2+n_jp_l(1-p_l)\right]\\ &=\sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij}\sum_{l=1}^{r} c_l\left\{\frac{n_j^2}{(n_i+\gamma)^2}\,[n_ip_l(1-p_l)+(\alpha_l-\gamma p_l)^2]+n_jp_l(1-p_l)\right\}\end{aligned}$$

$$= \sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij} \sum_{l=1}^{r} c_l \Big\{ \Big[\frac{n_j^2}{(n_i+\gamma)^2}\,(-n_i+\gamma^2) - n_j \Big] p_l^2 + \Big[\frac{n_j^2}{(n_i+\gamma)^2}\,(n_i - 2\alpha_l\gamma) + n_j \Big] p_l + \frac{n_j^2}{(n_i+\gamma)^2}\,\alpha_l^2 \Big\}.$$

Assume that

$$\varphi(\gamma) := \sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij} \Big[\frac{n_j^2}{(n_i+\gamma)^2}\,(-n_i+\gamma^2) - n_j \Big] = 0. \tag{6}$$

Since $\varphi(0) < 0$, the equation (6) always has a solution $\gamma > 0$ if

$$A =: \lim_{\gamma\to\infty} \varphi(\gamma) = \sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij} n_j (n_j - 1) > 0. \tag{7}$$

In this section we will suppose that the condition (7) holds and that γ is a solution of (6).

Applying the formula (6) to (5) we obtain

$$R(p,d) = \sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij}\, \frac{n_j^2}{(n_i+\gamma)^2} \sum_{l=1}^{r} c_l(\gamma^2 - 2\gamma\alpha_l) p_l + \sum_{\substack{i,j=1\\ i\neq j}}^{m} k_{ij}\, \frac{n_j^2}{(n_i+\gamma)^2} \sum_{l=1}^{r} c_l \alpha_l^2. \tag{8}$$

Assume that $c_2 \neq 0$. Let l_0 be the greatest index l for which $c_l \neq 0$ and let

$$L = \max_s \Big\{ s \le l_0 : \sum_{k=1}^{s} \frac{1}{c_k} > \frac{s-2}{c_s} \Big\}. \tag{9}$$

LEMMA. *Under the above notation, for $l = L+1, \ldots, r$,*

$$q := \frac{L-2}{\sum_{k=1}^{L} 1/c_k} \ge c_l. \tag{10}$$

Proof. Notice that the proof of the inequality (10) is only necessary for $l = L+1$. If $c_{L+1} \neq 0$ from (9) it follows that

$$L - 1 \ge c_{L+1} \sum_{k=1}^{L+1} \frac{1}{c_k} = 1 + c_{L+1} \sum_{k=1}^{L} \frac{1}{c_k}. \tag{11}$$

The inequality (10) follows from (11).

If $c_{L+1} = 0$ the inequality (10) obviously holds since $L \ge 2$. ∎

We shall prove the following theorem:

THEOREM 1. *If $A > 0$ then the game defined by the statistical decision problem considered has a value and the mutual predictor $d = [d_{ij}]_{i,j=1}^m$ defined by* (3) *for*

$$\alpha_l = \begin{cases} \frac{\gamma}{2}\left(1 - \frac{L-2}{c_l \sum_{k=1}^L 1/c_k}\right) & \text{for } l = 1, \ldots, L, \\ 0 & \text{for } l = L+1, \ldots, r, \end{cases} \tag{12}$$

where γ is the solution of (6), *is minimax for the loss function given by* (1).

Proof. Let the constants $\alpha_l \geq 0$ satisfy the equations

$$\gamma - 2\alpha_l = \frac{a}{c_l} \quad \text{for } l = 1, \ldots, L, \tag{13}$$

$$\alpha_l = 0 \quad \text{for } l = L+1, \ldots, r. \tag{14}$$

Taking into account the equations (4), (13) and (14) we obtain

$$(L-2)\gamma = a \sum_{l=1}^L \frac{1}{c_l}. \tag{15}$$

Since γ and L are known, the constant a is known and $a \geq 0$ because $\gamma > 0$ and $L \geq 2$. Moreover, from (10) and (15) it follows that

$$a = q\gamma \tag{16}$$

and from (13)–(15) we obtain the formula (12).

From (8) it follows that

$$\begin{aligned} R(p,d) &= \sum_{\substack{i,j=1 \\ i \neq j}}^m k_{ij} \frac{n_j^2}{(n_i+\gamma)^2} \Big[\sum_{l=1}^L c_l(\gamma^2 - 2\gamma\alpha_l)p_l + \sum_{l=L+1}^r c_l\gamma^2 p_l + \sum_{l=1}^r c_l \alpha_l^2 \Big] \\ &\overset{(13)}{=} \sum_{\substack{i,j=1 \\ i \neq j}}^m k_{ij} \frac{n_j^2}{(n_i+\gamma)^2} \Big[\sum_{l=1}^L a\gamma p_l + \sum_{l=L+1}^r c_l\gamma^2 p_l + \sum_{l=1}^L c_l \alpha_l^2 \Big] \\ &\overset{(16)}{=} \sum_{\substack{i,j=1 \\ i \neq j}}^m k_{ij} \frac{n_j^2}{(n_i+\gamma)^2} \Big[\sum_{l=1}^L q\gamma^2 p_l + \sum_{l=L+1}^r c_l\gamma^2 p_l + \sum_{l=1}^L c_l \alpha_l^2 \Big]. \end{aligned} \tag{17}$$

Thus $R(p,d) = \text{const} = C$ if $\sum_{l=1}^L p_l = 1$ and always, by the inequality (10), $R(p,d) \leq C$ for the mutual predictor d defined by (3) where γ and α_l are determined by (6) and (12). On the other hand, for any d and the loss function (1) the Bayes risk $r(\pi,d) = E_\pi(R(p,d))$ attains its minimum if

$$d_{ij}^{(l)}(X_i) = n_j E(p_l \mid X_i).$$

In the above formulae $E_\pi(\cdot)$ denotes the expectation with respect to the prior distribution π of the parameter $p = (p_1, \dots, p_r)$ and $E(p_l \mid X_i)$ is the conditional expectation of p_l for given X_i.

To prove that, assume that $p = (p_1, \dots, p_r)$ is a random variable. Let $X_i = (X_{i1}, \dots, X_{ir})$. The expression

$$E_\pi(E_p(d_{ij}^{(l)}(X_i) - X_{jl})^2) = E_\pi(E_p(d_{ij}^{(l)}(X_i) - n_j p_l)^2 + n_j p_l(1 - p_l))$$

attains its minimum when

$$d_{ij}^{(l)}(X_i) = n_j E(p_l \mid X_j) = E(p_l \mid (X_{i1}, \dots, X_{ir})).$$

Let the prior distribution π of $p = (p_1, \dots, p_r)$ be defined as follows:

$$\begin{aligned} &P(p_1 + \dots + p_L = 1) = 1, \quad p_i \ge 0, \\ &g(p_1, \dots, p_L) = \frac{\Gamma(\gamma)}{\Gamma(\alpha_1)\dots\Gamma(\alpha_L)} p_1^{\alpha_1 - 1} \dots p_L^{\alpha_L - 1}, \end{aligned} \tag{18}$$

where g is a density. For the prior density (18) and the loss function (1) the Bayes predictor is

$$\begin{aligned} &d_{ij}^{(l)}(x_{i1}, \dots, x_{iL}, 0, \dots, 0) \\ &\quad = n_j E(p_l \mid X_{i1} = x_{i1}, , \dots, X_{iL} = x_{iL},\ X_{il} = 0 \text{ for } l > L) \\ &\quad = \begin{cases} n_j \dfrac{x_{il} + \alpha_l}{n_i + \gamma} & \text{for } l = 1, \dots, L, \\ 0 & \text{for } l = L + 1, \dots, r;\ i, j = 1, \dots, m,\ i \ne j. \end{cases} \end{aligned}$$

Then $d = [d_{ij}]_{i,j=1}^m$ defined by (3) for γ and α_l satisfying (6) and (12) is the Bayes predictor and from the Hodges–Lehmann theorem (see [2]) it follows that the game defined by the statistical decision problem considered has a value and this predictor is minimax. ∎

3. Solution of the cases not solved in Section 2. Let $\varphi(\gamma)$ be defined by (6). Suppose that

$$A = \lim_{\gamma\to\infty} \varphi(\gamma) = \sum_{\substack{i,j=1 \\ i \ne j}}^{m} k_{ij} n_j (n_j - 1) = 0.$$

Since $\varphi(0) < 0$ and $\varphi(\gamma)$ is an increasing function of γ, in this case there does not exist any finite $\gamma > 0$ for which $\varphi(\gamma) = 0$. But it is easy to prove that in this case the minimax mutual predictor is obtained by taking into account the formula (12) and letting $\gamma \to \infty$ in (3),

$$d_{ij}^{(l)}(X_i) = \lim_{\gamma\to\infty} n_j \frac{X_{il} + \alpha_l}{n_i + \gamma} \tag{19}$$

$$= \begin{cases} \dfrac{n_j}{2}\left(1 - \dfrac{L-2}{c_l \sum_{k=1}^{L} 1/c_k}\right) =: n_j w_l & \text{for } l = 1, \ldots, L, \\ 0 & \text{for } l = L+1, \ldots, r. \end{cases}$$

For the mutual predictor $d = [d_{ij}]_{i,j=1}^m$ obtained in the above formula by letting $\gamma \to \infty$ the risk is (see (17))

$$R(p,d) = \sum_{\substack{i,j=1 \\ i \neq j}}^{m} k_{ij} n_j^2 \Big(\sum_{l=1}^{L} q p_l + \sum_{l=L+1}^{r} c_l p_l + \sum_{l=1}^{L} c_l w_l^2 \Big).$$

The fact that the mutual predictor obtained in this way is minimax for the loss function (1) results from the following well known theorem (see [1, p. 90]) which can be adapted to the above situation.

Theorem. *If d_n is the Bayes rule with respect to π_n, if $r(\pi_n, d_n) \to C$, and if $R(\mu, d_0) \le C$ for all μ then the game has a value and d_0 is a minimax rule.*

To define π_n it is enough to put $\gamma = n$ in (12) for α_l in (18).

Then we have proved the theorem.

Theorem 2. *If $A = 0$ then the game defined by the statistical decision problem considered has a value and the mutual predictor $d = [d_{ij}]_{i,j=1}^m$ given by* (19) *is minimax for the loss function* (1).

Up to this point we have assumed that $c_2 \neq 0$. If only $c_1 \neq 0$ the problem considered reduces to the minimax mutual prediction problem of a random variable $X = (X_1, \ldots, X_m)$, where X_j's are independent and have binomial distributions with the parameters n_j, p_1. The loss function is now of the form

$$L(X,d) = \sum_{\substack{i,j=1 \\ i \neq j}}^{m} c_1 k_{ij} (d_{ij}^{(1)}(X_i) - X_j)^2.$$

In this case if $A > 0$ the minimax mutual predictor $d = [d_{ij}]_{i,j=1}^m$ is given by

$$d_{ij}^{(l)}(X_i) = n_j \, \frac{X_i + \gamma/2}{n_i + \gamma},$$

where the parameter γ satisfies the equation (6), whereas when $A = 0$ the minimax mutual predictor is given by the formula

$$d_{ij}^{(1)}(X_i) = n_j/2.$$

The problems of minimax estimation and prediction of binomial and multinomial random variables were considered by Hodges and Lehmann [2], Trybuła [3]–[5], Wilczyński [6] and others.

References

[1] T. S. Ferguson, *Mathematical Statistics. A Decision Theoretic Approach*, Academic Press, New York, 1967.

[2] J. L. Hodges and E. L. Lehmann, *Some problems in minimax point estimation*, Ann. Math. Statist. 21 (1950), 182–197.

[3] S. Trybuła, *Some problems of simultaneous minimax estimation*, ibid. 29 (1958), 245–253.

[4] —, *Minimax mutual prediction*, Appl. Math. (Warsaw) 27 (2000), 437–444.

[5] —, *Simultaneous minimax estimation of parameters of multinomial distribution*, ibid. 29 (2002), 307–311.

[6] M. Wilczyński, *Minimax estimation for multinomial and multivariate hypergeometric distribution*, Sankhyā Ser. A 47 (1985), 128–132.

Institute of Mathematics
Technical University of Wrocław
50-370 Wrocław, Poland

Received on 18.11.2002;
revised version on 18.4.2003 (1662)