APPLICATIONES MATHEMATICAE
27,4 (2000), pp. 437–444

S. TRYBUŁA (Wrocław)

MINIMAX MUTUAL PREDICTION

Abstract. The problems of minimax mutual prediction are considered for binomial and multinomial random variables and for sums of limited random variables with unknown distribution. For the loss function being a linear combination of quadratic losses minimax mutual predictors are determined where the parameters of predictors are obtained by numerical solution of some equations.

1. Introduction. Suppose that a number of statisticians are observing some random variables. Assume that the ith statistician is observing a random variable X_i. He wants to predict the random variables of his partners. Let $d_{ij}(X_i)$ be the predictor he applies to predict X_j. Let the loss connected with this prediction be $L_{ij}(X_j, d_{ij}(X_i))$. Then the total loss of all statisticians is

$$L(X,d)=\sum_{\substack{i,j=1\\ i\neq j}}^{l} L_{ij}(X_j,d_{ij}(X_i)) \tag{1}$$

where $X=(X_1,\ldots,X_l)$,

$$d=\begin{bmatrix} - & d_{12} & \ldots & d_{1l}\\ d_{21} & - & \ldots & d_{2l}\\ \ldots & \ldots & \ldots & \ldots\\ d_{l1} & d_{l2} & \ldots & - \end{bmatrix} =: [d_{ij}]_1^l.$$

Suppose that the random variable X has distribution depending on an un-

2000 *Mathematics Subject Classification*: Primary 62F15.

Key words and phrases: minimax mutual predictor, binomial, multinomial, Bayes.

known parameter μ. The risk function is defined as

$$R(\mu,d)=E_\mu(L(X,d))=\sum_{\substack{i,j=1\\ i\neq j}}^{l}E_\mu(L_{ij}(X_j,d_{ij}(X_i)))\tag{2}$$

where $E_\mu(\cdot)$ denotes the expected value.

A mutual predictor d_0 is *minimax* if

$$\sup_\mu R(\mu,d_0)=\inf_d\sup_\mu R(\mu,d).$$

In this paper we shall find minimax mutual predictors in some situations.

2. Mutual predictors for binomial random variables. Let the variables $X_1,\dots,X_l$ be independent and have binomial distributions

$$f_i(x_i\,|\,p)=\binom{n_i}{x_i}p^{x_i}(1-p)^{n_i-x_i}$$

where the f_i are densities with respect to the counting measure. Let the losses L_{ij} be quadratic. Then

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

where $k_{ij}\ge 0$, $\sum_{i\neq j}k_{ij}>0$. Hence the risk function $R(p,d)$ can be represented in the form

$$\begin{aligned}R(p,d)&=\sum_{\substack{i,j=1\\ i\neq j}}^{l}k_{ij}E_p(d_{ij}(X_i)-X_j)^2\\&=\sum_{\substack{i,j=1\\ i\neq j}}^{l}k_{ij}[E_p(d_{ij}(X_i)-n_jp)^2+n_jp(1-p)].\end{aligned}\tag{3}$$

We shall look for minimax mutual predictors. Consider predictors of the form

$$d_{ij}(X_i)=n_j\frac{X_i+\alpha}{n_i+\gamma},\qquad \alpha>0,\ \gamma>0.\tag{4}$$

In this case

$$R(p,d)=\sum_{\substack{i,j=1\\ i\neq j}}^{l}k_{ij}\left[n_j^2E_p\left(\frac{X_i+\alpha}{n_i+\gamma}-p\right)^2+n_jp(1-p)\right]\tag{5}$$

$$= \sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\left[\frac{n_j^2}{(n_i+\gamma)^2}(n_i p(1-p) + (\alpha-\gamma p)^2) + n_j p(1-p)\right].$$

The risk $R(p,d)$ will be constant if

$$\sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\left[\frac{n_j^2}{(n_i+\gamma)^2}(-n_i+\gamma^2) - n_j\right] = 0 \tag{6}$$

and

$$\sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\left[\frac{n_j^2}{(n_i+\gamma)^2}(n_i - 2\alpha\gamma) + n_j\right] = 0. \tag{7}$$

For $\gamma = 0$ the left side of the first equation is negative, and as $\gamma \to \infty$ it tends to

$$A = \sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij} n_j (n_j - 1) \geq 0. \tag{8}$$

Moreover it is an increasing function of the parameter γ. Therefore if $A > 0$ there always exists a unique solution of equation (6).

Suppose that there exists a solution γ of (6). In this case there exists a solution α of (7) and

$$\alpha = \gamma/2. \tag{9}$$

Equation (6) can be solved numerically.

When $n_1 = \ldots = n_l =: n > 1$, the solution of (6) is

$$\gamma = \frac{n}{n-1}(\sqrt{2n-1} + 1) \tag{10}$$

and it is independent of k_{ij}.

When all $k_{ij} = 0$ except, say k_{12}, equation (6) has a solution

$$\gamma = \frac{n_1}{n_2 - 1}\left(n_2\sqrt{\frac{1}{n_1} + \frac{1}{n_2} - \frac{1}{n_1 n_2}} + 1\right). \tag{11}$$

Suppose that there exists a solution $\gamma > 0$ of (6). It is easy to prove that the predictors given by (4) are Bayes with respect to the a priori distribution of the parameter p given by the density

$$g(p) = \frac{1}{B(\alpha, \gamma-\alpha)} p^{\alpha-1}(1-p)^{\gamma-\alpha-1} \quad \text{for } 0 < p < 1. \tag{12}$$

Thus the mutual predictor $d = [d_{ij}]_1^l$ defined by (4), being a constant risk Bayes predictor, is minimax if $\alpha = \gamma/2$, where γ is a solution of (6).

When $A = 0$ a minimax mutual predictor is given by

$$d_{ij}(X_i) = n_j/2. \tag{13}$$

It is obtained by letting $\gamma \to \infty$, $\alpha = \gamma/2$ in (4).

The problem of minimax prediction when only one $k_{ij} > 0$ was solved by Hodges and Lehmann in [1].

When $A > 0$ the minimax risk is

$$R(p, d_0) = \frac{1}{4} \sum_{\substack{i,j=1 \\ i \neq j}}^{l} \frac{k_{ij} n_j^2}{(n_i + \gamma)^2} \gamma^2 \tag{14}$$

where d_0 satisfies (4), (6) and (9).

When $A = 0$ the minimax risk can be obtained by letting $\gamma \to \infty$ in formula (14):

$$R(p, d_0) = \frac{1}{4} \sum_{\substack{i,j=1 \\ i \neq j}}^{l} k_{ij} n_j^2 \tag{15}$$

where d_0 is given by (13).

3. Minimax mutual predictors for multinomial random variables. Let now $X_i = (X_{i1}, \ldots, X_{ir})$, $i = 1, \ldots, 1$, be independent random variables distributed according to multinomial laws

$$f_i(x_i \mid p) = \frac{n_i!}{x_{i1}! \ldots x_{ir}!} p_1^{x_{i1}} \ldots p_r^{x_{ir}},$$

where $x_i = (x_{i1}, \ldots, x_{ir})$ is the value of X_i. Let the loss function be of the form

$$L(X, d) = \sum_{\substack{i,j=1 \\ i \neq j}}^{l} \sum_{k=1}^{r} k_{ij} (d_{ij}^{(k)}(X_i) - X_{jk})^2. \tag{16}$$

Let us consider the predictors

$$d_{ij}^{(k)}(X_i) = n_j \frac{X_{ik} + \alpha_k}{n_i + \gamma}, \qquad i, j = 1, \ldots, l,\ i \neq j,\ k = 1, \ldots, r. \tag{17}$$

It is easy to show that for the loss function given by (16) with some $k_{ij} > 0$ these Bayes predictors satisfy the equations

$$\sum_{k=1}^{r} d_{ij}^{(k)}(X_i) = n_j, \qquad j = 1, \ldots, l. \tag{18}$$

All these equations will surely be satisfied by $d_{ij}^{(k)}$ given in (17) when

$$\alpha_k = \gamma/r, \qquad k = 1, \ldots, r. \tag{19}$$

For predictors satisfying (17) and (19) the risk function will take the form

$$R(p,d) = \sum_{\substack{i,j=1\\ i\neq j}}^{l} \sum_{k=1}^{r} k_{ij}\Big[\frac{n_j^2}{(n_i+\gamma)^2}(n_i p_k(1-p_k) + (\alpha_k - \gamma p_k)^2) + n_j p_k(1-p_k)\Big]. \tag{20}$$

Notice that in (20) the coefficients of p_k^2 for $k = 1,\dots,r$ are the same. They are zero when γ satisfies (6).

Let α_k satisfy (19). In this case for γ given by (6) the expresion (20) will take the form

$$\begin{aligned} R(p,d) &= \sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\Big[\frac{n_j^2}{(n_i+\gamma)^2}\Big(n_i - \frac{\gamma^2}{r}\Big) + n_j\Big] \\ &\overset{(6)}{=} \sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\Big[\frac{n_j^2}{(n_i+\gamma)^2}\Big(n_i - \frac{\gamma^2}{r}\Big) \\ &\qquad + n_j + \frac{n_j^2}{(n_i+\gamma)^2}(-n_i+\gamma^2) - n_j\Big] \\ &= \frac{r-1}{r}\sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\frac{n_j^2}{(n_i+\gamma)^2}\gamma^2. \end{aligned} \tag{21}$$

From the above it follows that $d = [(d_{ij}^{(1)},\dots,d_{ij}^{(r)})]_1^l$, where $d_{ij}^{(k)}$ are given by (17), (19) and (6), is a constant risk mutual predictor. Let $p = (p_1,\dots,p_r)$ be a random variable. The expression

$$E(E_p(d_{ij}^{(k)}(X_i) - X_{jk})^2) = E(E_p(d_{ij}^{(k)}(X_i) - n_j p_k)^2 + n_j p_k(1-p_k)) \tag{22}$$

attains its minimum when

$$d_{ij}^{(k)}(X_i) = n_j E(p_k \mid X_i) = n_j E(p_k \mid (X_{i1},\dots,X_{ir})). \tag{23}$$

Here $E(p_k \mid X_i)$ denotes the conditional expectation of the random variable p_k under the condition that X_i is given. For the a priori distribution given by the density

$$g(p_1,\dots,p_k) = \frac{\Gamma(\gamma)}{[\Gamma(\gamma/r)]^r}(p_1\dots p_r)^{\gamma/r-1} \tag{24}$$

we find that

$$d_{ij}^{(k)}(X_i) = n_j E(p_k \mid X_i) = n_j\frac{X_{ik}+\gamma/r}{n_i+\gamma} \tag{25}$$

is a Bayes predictor.

We have shown that the mutual predictor d given by (17) and (19) is a Bayes predictor. Thus for γ satisfying (6) it is a minimax predictor for the loss function (16).

Equation (6) has a solution γ when $A > 0$ (see (8)). When $A = 0$ it is easy to show that a minimax mutual predictor is independent of $X = (X_1, \ldots, X_l)$ and is given by

$$d_{ij}^{(k)}(X_i) = n_j/r, \quad i, j = 1, \ldots, l, \ i \neq j, \ k = 1, \ldots, r, \tag{26}$$

and

$$R(p, d) = \frac{r-1}{r} \sum_{\substack{i,j=1 \\ i \neq j}}^{l} k_{ij} n_j^2. \tag{27}$$

If only one $k_{ij} \neq 0$, the results of this section follow from the paper of Wilczyński [4].

4. Minimax predictors of limited random variables. Suppose that the ith statistician is observing n_i random variables $X_{i1}, \ldots, X_{in_i}$ with values in the interval $[0,1]$ and let $X_i = \sum_{k=1}^{n_i} X_{ik}$, $X = (X_1, \ldots, X_l)$, where the X_{ik} are independent. The statistician wants to predict the random variables X_j, $j = 1, \ldots, l$, $j \neq i$. Let the total loss function of all statisticians be of the form

$$L(X, d) = \sum_{\substack{i,j=1 \\ i \neq j}}^{l} k_{ij} (d_{ij}(X_i) - X_j)^2. \tag{28}$$

Let the random variables X_{ik}, $i = 1, \ldots, l$, $k = 1, \ldots, n_i$, have the same distribution function F and let $\mu = E_F(X_{ik})$. The risk function is

$$\begin{aligned} R(\mu, d) &= \sum_{\substack{i,j=1 \\ i \neq j}}^{l} k_{ij} E_F (d_{ij}(X_i) - X_j)^2 \\ &= \sum_{\substack{i,j=1 \\ i \neq j}}^{l} k_{ij} [E_F (d_{ij}(X_i) - n_j \mu)^2 + E_F (X_j - n_j \mu)^2]. \end{aligned} \tag{29}$$

We look for minimax predictors.

Let us try predictors of the form

$$d_{ij} = n_j \frac{X_i + \alpha}{n_i + \gamma}. \tag{30}$$

Then

$$R(\mu,d)=\sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\frac{n_j^2}{(n_i+\gamma)^2}[E_F(X_i-n_i\mu)^2 + (\alpha-\gamma\mu)^2+E_F(X_j-n_j\mu)^2] \le \sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\Big[\frac{n_j^2}{(n_i+\gamma)^2}(n_i\mu(1-\mu) + (\alpha-\gamma\mu)^2)+n_j\mu(1-\mu)\Big]. \tag{31}$$

Let equations (6) and (7) be satisfied. Then $\alpha=\gamma/2$ and

$$R(\mu,d)=\frac{1}{4}\sum_{\substack{i,j=1\\ i\neq j}}^{l} k_{ij}\frac{n_j^2}{(n_i+\gamma)^2}\gamma^2=:c \tag{32}$$

and for any random variable $X=(X_1,\ldots,X_l)$ satisfying the conditions given at the beginning of this section,

$$R(\mu,d)\le c. \tag{33}$$

Let X_{ik} have two-point distribution

$$P(X_{ik}=0)=1-p,\qquad P(X_{ik}=1)=p. \tag{34}$$

Then $\mu=p$ and equality holds in (31). From formulae (31)–(34) and the results of Section 2 it follows that for γ satisfying (6) and $\alpha=\gamma/2$, the predictor $d=[d_{ij}]_1^l$ given by (30) is a minimax mutual predictor. This holds when $A>0$, where A is defined in (8).

For $A=0$ a minimax mutual predictor is given by the formula

$$d_{ij}(X_i)=n_j/2. \tag{35}$$

When only one $k_{ij}\neq 0$ the problem considered in this section was solved by Hodges and Lehmann in [1].

Analogous results can also be obtained for mutual prediction of sample cumulative distribution functions for a properly chosen loss function. For minimax estimation of cumulative distribution functions, see Phadia [2].

For minimax estimators and predictors of many parameters and random variables, see [3], [4].

References

[1] J. L. Hodges and E. L. Lehmann, *Some problems in minimax point estimation*, Ann. Math. Statist. 21 (1950), 182–191.

[2] E. G. Phadia, *Minimax estimation of cumulative distribution functions*, Ann. Statist. 1 (1973), 1149–1157.
[3] S. Trybuła, *Some problems of simultaneous minimax estimation*, Ann. Math. Statist. 29 (1958), 245–253.
[4] M. Wilczyński, *Minimax estimation for multinomial and multivariate hypergeometric distribution*, Sankhyā, Ser. A 47 (1985), 128–132.

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

Received on 30.11.1999;
revised version on 17.3.2000