APPLICATIONES MATHEMATICAE
29,3 (2002), pp. 307–311

Stanisław Trybuła (Wrocław)

SIMULTANEOUS MINIMAX ESTIMATION OF PARAMETERS OF MULTINOMIAL DISTRIBUTION

Abstract. The problem of minimax estimation of parameters of multinomial distribution is considered for a loss function being the sum of the losses of the statisticians taking part in the estimation process.

1. Introduction. Let $X_i = (X_{i1}, \ldots, X_{ir})$, $i = 1, \ldots, m$, be observed by the ith statistician. The random variables X_i, $i = 1, \ldots, r$, have multinomial distribution with parameters n_i, $p = (p_1, \ldots, p_r)$, and are independent. The statisticians do not know the observations of their colleagues but they know all the numbers n_i. They cooperate with each other. The problem is to determine the simultaneous minimax estimator $d = (d_1, \ldots, d_m)$ of the parameter p where $d_i(X_i) = (d_{i1}(X_i), \ldots, d_{ir}(X_i))$ is the estimator of this parameter, used by the ith statistician. Let the loss function be

$$L(p, d) = \sum_{i=1}^{m} k_i \sum_{j=1}^{r} c_j (d_{ij}(X_i) - p_j)^2 \tag{1}$$

where $k_i > 0$, $c_j \geq 0$ are constants.

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

$$R(p, d) = E_p(L(p, d))$$

where $E_p(\cdot)$ is the operator of expected value with respect to the distribution of the random variable $X = (X_1, \ldots, X_m)$. Then we have to find the estimator d^0 for which

$$\sup_p R(p, d^0) = \inf_d \sup_p R(p, d).$$

2000 *Mathematics Subject Classification*: Primary 62F15.
Key words and phrases: minimax estimation, multinomial distribution, Bayes estimator.

2. Solution of the problem. Without loss of generality we can suppose that $c_1 \geq \ldots \geq c_r \geq 0$. For the moment assume also that $c_2 \neq 0$. Consider the estimator d for which

$$d_{ij}(X_i) = \frac{X_{ij} + \alpha_j}{n_i + \gamma}, \qquad i = 1, \ldots, m, \ j = 1, \ldots, r. \tag{2}$$

For this estimator the risk will take the form

$$\begin{aligned} R(p, d) &= \sum_{i=1}^{m} k_i \sum_{j=1}^{r} c_j E_p \left(\frac{X_{ij} + \alpha_j}{n_i + \gamma} - p_j \right)^2 \\ &= \sum_{i=1}^{m} k_i \sum_{j=1}^{r} \frac{c_j}{(n_i + \gamma)^2} \left[E_p (X_{ij} - n_i p_j)^2 + (\alpha_j - \gamma p_j)^2 \right] \\ &= \sum_{j=1}^{r} c_j \sum_{i=1}^{m} \frac{k_i}{(n_i + \gamma)^2} \left[(\gamma^2 - n_i) p_j^2 + (n_i - 2\alpha_j \gamma) p_j + \alpha_j^2 \right]. \end{aligned} \tag{3}$$

Assume that

$$\sum_{i=1}^{m} \frac{k_i}{(n_i + \gamma)^2} (\gamma^2 - n_i) = 0. \tag{4}$$

It is easy to see that equation (4) always has a solution with respect to the constant $\gamma > 0$.

Moreover assume that the constants $\alpha_j \geq 0$ satisfy the equations

$$c_j \sum_{i=1}^{m} \frac{k_i}{(n_i + \gamma)^2} (n_i - 2\alpha_j \gamma) = c \qquad \text{for } j \leq L, \tag{5}$$

for some integer L and

$$\alpha_j = 0 \qquad \text{for } L < j \leq r. \tag{6}$$

Finally, let

$$\sum_{j=1}^{r} \alpha_j = \gamma. \tag{7}$$

From (4) and (5) we obtain for $j \leq L$, if $c_j \neq 0$,

$$\sum_{i=1}^{m} \frac{k_i}{(n_i + \gamma)^2} (\gamma^2 - 2\alpha_j \gamma) = \frac{c}{c_j}. \tag{8}$$

Then from (6)–(8), if $c_L \neq 0$, we obtain

$$(L - 2)\gamma^2 \sum_{i=1}^{m} \frac{k_i}{(n_i + \gamma)^2} = \sum_{j=1}^{L} \frac{c}{c_j},$$

or by (4),

$$(L-2)\sum_{i=1}^{m}\frac{k_i n_i}{(n_i+\gamma)^2}=\sum_{j=1}^{L}\frac{c}{c_j}. \tag{9}$$

Thus the constant c is determined for given L and γ.

Let j_0 be the greatest index j for which $c_j \neq 0$ and let

$$L=\max_s\left\{s\le j_0 : \sum_{l=1}^{s}\frac{1}{c_l}>\frac{s-2}{c_s}\right\}. \tag{10}$$

We shall prove

LEMMA. *For $j=L+1,\ldots,r$,*

$$q:=\frac{L-2}{\sum_{l=1}^{L}1/c_l}\ge c_j. \tag{11}$$

Proof. Notice that the proof is only necessary for $j=L+1$. If $c_{L+1}=0$ the lemma obviously holds. If $c_{L+1}\neq 0$, from (10) it follows that

$$L-1\ge c_{L+1}\sum_{l=1}^{L+1}\frac{1}{c_l}=1+c_{L+1}\sum_{l=1}^{L}\frac{1}{c_l}.$$

The lemma is a consequence of this inequality.

Taking into account (4) we obtain, from (3),

$$\begin{aligned}R(p,d)&=\sum_{j=1}^{r}c_j\sum_{i=1}^{m}\frac{k_i}{(n_i+\gamma)^2}\left[(n_i-2\alpha_j\gamma)p_j+\alpha_j^2\right]\\&\overset{(5),(6)}{=}\sum_{j=1}^{L}cp_j+\sum_{j=1}^{L}c_j\sum_{i=1}^{m}\frac{k_i\alpha_j^2}{(n_i+\gamma)^2}+\sum_{j=L+1}^{r}c_j\sum_{i=1}^{m}\frac{k_i n_i}{(n_i+\gamma)^2}p_j\\&\overset{(9),(11)}{=}\sum_{i=1}^{m}\frac{k_i n_i}{(n_i+\gamma)^2}\Big(\sum_{j=1}^{L}qp_j+\sum_{j=L+1}^{r}c_jp_j\Big)+\sum_{j=1}^{L}c_j\sum_{i=1}^{m}\frac{k_i\alpha_j^2}{(n_i+\gamma)^2}.\end{aligned}$$

Thus $R(p,d)=\text{const}=C$ if $\sum_{j=1}^{L}p_j=1$ and always, by the Lemma, $R(p,d)\le C$ for the simultaneous estimator d defined by (2) and satisfying (4)–(7), (9) and (10). On the other hand, for any d and the loss function (1) the expected risk $r(\pi,d)=E(R(p,d))$ attains its minimum if

$$d_{ij}(X_i)=E(p_j\mid X_i).$$

Here $E(\cdot)$ is the expectation for a prior distribution π of the parameter p and $E(p_j\mid X_i)$ is the conditional expectation of p_j given X_i.

Let a prior distribution of $p = (p_1, \ldots, p_r)$ be defined as follows:

$$
(12) \qquad \begin{aligned} &P(p_1 + \ldots + p_L = 1) = 1, \\ &g(p_1, \ldots, p_L) = \frac{\Gamma(\gamma)}{\Gamma(\alpha_1) \ldots \Gamma(\alpha_L)} \, p_1^{\alpha_1 - 1} \ldots p_L^{\alpha_{L-1}} \end{aligned}
$$

where g is a density. For the prior distribution (12) we obtain

$$
\begin{aligned} d_{ij}(x_{i1}, \ldots, x_{iL}, 0, \ldots, 0) &= E(p_j \mid X_{i1} = x_{i1}, \ldots, X_{iL} = x_{iL}, \\ &\qquad\qquad X_{i,L+1} = \ldots = X_{ir} = 0) \\ &= \begin{cases} \dfrac{x_{ij} + \alpha_j}{n_i + \gamma} & \text{for } j = 1, \ldots, L, \\ 0 & \text{for } j = L+1, \ldots, r;\ i = 1, \ldots, m. \end{cases} \end{aligned}
$$

Then the estimator d defined by (2) and satisfying (4)–(7), (9), (10) is Bayes and from the Hodges and Lehmann theorem it follows that it is minimax.

For $r = 2$ always $L = 2$, $c = q = 0$ and the relevant estimator is a constant risk estimator.

Up to this point we have assumed that $c_2 \neq 0$. If only $c_1 \neq 0$ the problem reduces to simultaneous estimation of the parameter p of binomial distribution for the loss function

$$
L(p, d) = \sum_{i=1}^{m} k_i (d_i(X_i) - p)^2
$$

where X_i is observed by the ith statistician. In this case the simultaneous minimax estimator is given by the formula

$$
d_i(X_i) = \frac{X_i + \gamma/2}{n_i + \gamma}
$$

where γ satisfies (4).

The problem considered in this paper may be generalized by introducing the loss function

$$
L(p, d) = \sum_{i=1}^{m} k_i \sum_{j,l=1}^{r} c_{jl} (d_{ij}(X_i) - p_j)(d_{il}(X_i) - p_l)
$$

where the matrix $\|c_{jl}\|_1^r$ is nonnegative definite. To solve the problem for this loss function one can apply linear programming methods as done by Wilczyński [2] for $m = 1$.

For problems of minimax estimation of many parameters by one statistician see Trybuła [1].

References

[1] S. Trybuła, *Some problems of simultaneous minimax estimation*, Ann. Math. Statist. 29 (1958), 245–253.

[2] 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 6.2.2002;
revised version on 1.7.2002 (1612)