

S. TRYBULA (Wrocław)

## MINIMAX AND BAYES ESTIMATION WHEN THE LOSS FUNCTION IS UNKNOWN

1. In this paper the problem of minimax estimation is considered in the situation when the loss function is unknown but belongs to a class for which a minimax estimator, the same for each loss function, can be determined. Numerous examples, also in the sequential case, are presented. Problems of Bayes estimation based partly on the above idea are considered. An example from game theory is examined from this point of view.

2. Suppose that the random variable  $X$  has the distribution  $P_\lambda$ , dependent on the parameter  $\lambda = (\lambda_1, \dots, \lambda_r) \in A \subset R^r$ .  $X$  is observed and an estimator  $d(X) = (d_1(X), \dots, d_r(X))$  is applied. Let  $\{L(\lambda, a), L \in \mathcal{L}\}$  be a family of loss functions and let  $\Gamma$  be a family of distributions  $\pi$  of parameter  $\lambda$ . Let  $r(\pi, d)$  be the Bayes risk associated with the distribution  $\pi$  and the estimator  $d$  and let  $D$  be the class of all estimators for which  $r(\pi, d)$  exists for each  $\pi \in \Gamma$  and  $L \in \mathcal{L}$ . The problem is to determine an estimator minimax for  $L$ , which means to determine an estimator  $d_0$  such that

$$\sup_{\pi \in \Gamma} r(\pi, d_0) = \inf_{d \in D} \sup_{\pi \in \Gamma} r(\pi, d)$$

for each  $L \in \mathcal{L}$ .

Let  $R(\lambda, d)$  be the risk function associated with estimator  $d$  and let now  $D$  be the class of estimators for which  $R(\lambda, d)$  exists and is finite for each  $\lambda \in A$  and  $L \in \mathcal{L}$ . If we do not impose any restrictions on the class of a priori distributions then we say that  $d_0$  is *minimax* for  $L \in \mathcal{L}$  if

$$\sup_{\lambda \in A} R(\lambda, d_0) = \inf_{d \in D} \sup_{\lambda \in A} R(\lambda, d)$$

for each  $L \in \mathcal{L}$ .

Obviously, if an estimator is minimax for  $L \in \mathcal{L}$  it is minimax for  $L \in \mathcal{L}' \subset \mathcal{L}$ .

In the situation when the minimax estimator  $d_0$  for  $L \in \mathcal{L}$  exists, the information about the loss function is not necessary to determine  $d_0$ . It is sufficient only to know that  $L \in \mathcal{L}$ .

We give some examples in which an estimator minimax for  $L \in \mathcal{L}$  can be determined.

3. Assume that the random variable  $X = (X_1, \dots, X_r)$  is normally distributed with mean vector  $\lambda = (\lambda_1, \dots, \lambda_r)$  and known positive definite covariance matrix  $\Sigma$ . Let  $a_i$  be an estimate of  $\lambda_i$ ,  $\hat{a} = (a_1, \dots, a_r)$ , and let the loss function be

$$L(\hat{a}, \lambda) = \sum_{i,j=1}^r c_{ij} (a_i - \lambda_i)(a_j - \lambda_j),$$

where matrix  $C = \|c_{ij}\|$  is positive definite. It is well known that  $d(X) = X$  is a minimax estimator of  $\lambda$  for each  $C$  positive definite. Then to determine a minimax estimator we have not to know the exact values of  $c_{ij}$  in  $L(\hat{a}, \lambda)$ .

However, the estimator  $d(X) = X$  is not admissible. Better is the Stein estimator

$$d(X) = \left(1 - \frac{r-2}{|X|^2}\right)X,$$

where  $|X|$  denotes the Euclidean norm of  $X$ . This estimator is also inadmissible if  $r \geq 3$ . The problem of admissibility of estimators of normal mean  $\lambda$  was considered by many authors. See, for example, Stein [11], Efron and Morris [4], Berger and Srinivasan [2].

4. Let  $X = (X_1, \dots, X_r)$  be normally distributed with mean vector  $\lambda = (\lambda_1, \dots, \lambda_r)$  and unknown positive definite matrix  $\Sigma$ . We observe  $X$  and also independently observe the  $r \times r$  random matrix  $W$  which has a Wishart distribution with  $n$  degrees of freedom and parameter  $\Sigma = n^{-1} E(W)$ .  $\lambda$  is estimated by an estimate  $d(X, W)$ . Let the loss function be

$$L(\lambda, \hat{a}, \Sigma) = \frac{1}{\text{tr}(\Sigma Q)} [(\hat{a} - \lambda) Q (\hat{a} - \lambda)],$$

where  $a_i$  is an estimate of  $\lambda_i$ ,  $\hat{a} = (a_1, \dots, a_r)$ . Matrix  $Q$  is known and positive definite.

Also in this case the estimator  $d(X, W) = X$  is minimax. In 1977 Berger, Bock, Brown, Casella and Gleser [1] found the first explicit examples of estimators which dominate  $X$  in risk when  $r \geq 3$ . But these estimators depend on the matrix  $Q$  and cannot be applied when  $Q$  is unknown.

5. Let  $X_1, \dots, X_n$  be independent random variables having the Poisson distribution

$$P(X_i = x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad (i = 1, \dots, n, x = 0, 1, 2, \dots).$$

Let

$$(1) \quad X^{(k)} = (X_1, \dots, X_k), \quad Y_k = \sum_{i=1}^k X_i.$$

The problem is to find a minimax estimator of parameter  $\lambda$  in the situation when we estimate systematically at step  $k = 1, 2, \dots$  and when the loss function is of the form

$$(2) \quad L(\lambda, \hat{a}) = \sum_{k=1}^n c_k (a_k - \lambda)^2,$$

where  $c_k \geq 0$  ( $k = 1, \dots, n$ ) are unknown and  $a_k$  is an estimate of  $\lambda$  in step  $k$ ,  $\hat{a} = (a_1, \dots, a_n)$ . Estimation in step  $k$  is based on  $X^{(k)}$ . Possible a priori distributions of parameter  $\lambda$  are that for which

$$(3) \quad E(\lambda^2) = m_2,$$

where  $m_2 > 0$  is given. We denote this set of distributions by  $\Gamma$ .

Let us consider the estimator  $d = (d_1, \dots, d_n)$  for which

$$(4) \quad d_k = \frac{Y_k + \alpha}{k + r}$$

for some  $\alpha > 0$  and some  $r > 0$ . For this estimator the risk function is

$$R(\lambda, d) = \sum_{k=1}^n c_k E_\lambda \left( \frac{Y_k + \alpha}{k + r} - \lambda \right)^2 = \sum_{k=1}^n \frac{c_k}{(k + r)^2} [k\lambda + (\alpha - r\lambda)^2].$$

For a distribution  $\pi$  of parameter  $\lambda$  for which equation (3) holds the Bayes risk is

$$r(\pi, d) = E(R(\lambda, d)) = \sum_{k=1}^n \frac{c_k}{(k + r)^2} [(k - 2\alpha r) E(\lambda) + \alpha^2 + r^2 m_2]$$

and it is independent of  $\pi$  when

$$(5) \quad \sum_{k=1}^n \frac{c_k}{(k + r)^2} (k - 2\alpha r) = 0.$$

Let

$$(6) \quad r = \sqrt{\alpha(\alpha + 1)/m_2}.$$

There always exists the solution  $\alpha^* > 0$  of the equation (5), where  $\alpha$  is given

by (6). On the other hand, the estimator  $d$  defined by (4) and (6) with  $\alpha = \alpha^*$  is a Bayes estimator with respect to the a priori distribution  $\pi \in \Gamma$  with density

$$g(\lambda) = \begin{cases} \frac{r^{\alpha^*}}{\Gamma(\alpha^*)} \lambda^{\alpha^*-1} e^{-r\lambda}, & \text{if } \lambda > 0, \\ 0, & \text{if } \lambda \leq 0. \end{cases}$$

Then the estimator  $d = d^0$  defined by (4) and (6) with  $\alpha = \alpha^*$  is minimax, that is

$$\sup_{\pi \in \Gamma} r(\pi, d^0) = \inf_{d \in D} \sup_{\pi \in \Gamma} r(\pi, d),$$

where  $D$  is the class of estimators for which  $r(\pi, d)$  exists and is finite for each  $\pi \in \Gamma$  and for each loss function of the form (2) for which equation (5) holds for  $\alpha = \alpha^*$ .

Then a minimax estimator (of the form (4)) may be determined for each set of vectors  $(c_1, \dots, c_n)$  for which equation (5) holds for some  $\alpha > 0$ ,  $r = \sqrt{\alpha(\alpha+1)/m_2}$ .

The "true" vector  $(c_1, \dots, c_r)$  may be unknown but the constants  $c_1 \geq 0, \dots, c_n \geq 0$  which determine the loss function  $L$  have to satisfy one condition in this problem.

6. Let  $X_1, \dots, X_{n+1}$  be independent random variables with the Poisson distribution with parameter  $\lambda$  and let  $Y_k$  be defined by (1). We want to predict (systematically) the random variable  $Y_{n+1}$  on the basis of  $Y_1, \dots, Y_k$  ( $k = 1, \dots, n$ ). Since in the  $k$ th step we know the values of random variables  $X_1, \dots, X_k$  and at this step  $Y_k$  is a sufficient statistic for  $\lambda$  it is sufficient to predict the values of

$$Y^{(k)} = \sum_{i=k+1}^{n+1} X_i, \quad k = 1, \dots, n,$$

on the basis of  $Y_k$  ( $k = 1, \dots, n$ ), respectively. We state the minimax problem in the same form as in Section 5 with the only difference that now the loss function is

$$L(Y, a) = \sum_{k=1}^n c_k (a_k - Y^{(k)})^2,$$

where  $c_1 \geq 0, \dots, c_n \geq 0$  are unknown.

For  $\alpha > 0$  the predictor  $d = (d_1, \dots, d_r)$  with

$$d_k = \frac{Y_k + \alpha}{k + r} (n - k + 1)$$

is minimax when it is known that the "true" vector  $(c_1, \dots, c_r)$  satisfies the

condition

$$\sum_{k=1}^n c_k \left[ \left( \frac{n-k+1}{k+r} \right)^2 (k-\alpha r) + n-k+1 \right] = 0,$$

where  $r$  is given by (6) (see [15]).

7. Let  $X = (X_1, \dots, X_r)$  be a random variable having the multinomial distribution

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

where the parameter  $p = (p_1, \dots, p_r)$  is unknown. Let  $\hat{a} = (a_1, \dots, a_r)$  be an estimate of  $p$  and let the loss function be

$$(7) \quad L(p, \hat{a}) = \sum_{i,j=1}^r c_{ij} (p_i - a_i)(p_j - a_j),$$

where the matrix  $C = \|c_{ij}\|$  is unknown but is assumed to be nonnegative definite. From [15] it follows that, for known  $C$ , the estimator  $d = (d_1, \dots, d_r)$ , where

$$(8) \quad d_i(X) = \frac{X_i + \beta_i \sqrt{n}}{n + \sqrt{n}} \quad (\beta_i \geq 0, \sum_{i=1}^r \beta_i = 1),$$

is minimax if there exist a constant  $v$  and a set  $A \subset R = \{1, \dots, r\}$ ,  $|A| \geq 2$ , such that

$$(9) \quad \begin{aligned} \sum_{j \in A} (c_{ii} - 2c_{ij}) \beta_j &= v, \quad \text{for } i \in A, \\ \sum_{j \in A} (c_{ii} - 2c_{ij}) \beta_j &\leq v, \quad \text{for } i \in R - A, \end{aligned}$$

$\beta_j > 0$  for  $j \in A$ ,  $\beta_j = 0$  for  $j \in R - A$ . In [17] it is proved that such a constant  $v$  and set  $A$  always exist.

Let, for example,

$$\beta_1 = \beta_2 = \dots = \beta_r = 1/r.$$

In this case the estimator (8) is a Steinhaus estimator. This estimator is minimax if  $C = I$  (see [12]). For  $\beta_i = 1/r$ , the constants  $c_{ij}$  in (7) should satisfy the condition

$$\sum_{j=1}^r (c_{ii} - 2c_{ij}) = w \quad (i = 1, \dots, r)$$

for some  $w$  (see (9)), and have to satisfy some inequalities which follow from the supposition that matrix  $C$  is nonnegative definite.

If  $C$  is diagonal (and known),  $c_{ij} = 0$  for  $i \neq j$ ,  $c_{ii} = c_i$ ,  $c_1 \geq c_2 \geq \dots \geq c_r \geq 0$ , the constants  $\beta_i$  in (8) are determined as follows [13]:

Let  $c_2 = 0$ . Let  $l_0$  be the greatest index  $i$  for which  $c_i \neq 0$  and let

$$(10) \quad L = \max_s \left[ s \leq l_0, \sum_{i=1}^s \frac{1}{c_i} > \frac{s-2}{c_s} \right],$$

$$(11) \quad \delta = (L-2) / \sum_{j=1}^L (1/c_j).$$

Then

$$(12) \quad \beta_i = \begin{cases} \frac{1}{2}(1 - \delta/c_i) & \text{when } i \leq L, \\ 0 & \text{when } i > L. \end{cases}$$

Let  $C$  be unknown. If  $L \geq 3$  then  $\delta > 0$  and  $\beta_i < 1/2$  for  $i = 1, \dots, r$ . Suppose then that

$$1/2 > \beta_1 \geq \dots \geq \beta_L, \quad \beta_{L+1} = \dots = \beta_r = 0, \quad L \geq 3.$$

Conditions (11) and (12) give

$$(1/c_i) / \sum_{j=1}^L (1/c_j) = (1 - 2\beta_i) / (L-2) \quad \text{for } i = 1, \dots, L.$$

Equation (10) gives

$$\frac{1}{c_i} / \sum_{j=1}^L \frac{1}{c_j} + \frac{1}{c_i} \geq \frac{1}{L-1} \quad \text{for } i = L+1, \dots, r$$

if  $c_i \neq 0$ . Then for given  $\beta_i$  ( $i = 1, \dots, r$ ) and a given multiplicative constant for  $c_i$ , the constants  $c_1, \dots, c_L$  are determined uniquely when it is known that matrix  $C$  is diagonal; the constants  $c_{L+1}, \dots, c_r$  are not uniquely determined.

8. Suppose that a random variable  $X' = (X'_1, \dots, X'_r)$  satisfies the condition

$$X'_1 \geq 0, \dots, X'_r \geq 0, \quad X'_1 + \dots + X'_r \leq s$$

( $s > 0$ ). Let  $X^{(1)}, \dots, X^{(n)}$ ,  $X^{(j)} = (X_1^{(j)}, \dots, X_r^{(j)})$ ,  $j = 1, \dots, n$ , be independent random variables having the same distribution as  $X'$ . Denote  $X = (X^{(1)}, \dots, X^{(n)})$ ,  $m_i = E(X_i)$ ,  $i = 1, \dots, r$ , and let  $d(X) = (d_1(X), \dots, d_r(X))$  be an estimator of parameter  $m = (m_1, \dots, m_r)$ . The problem is to find a minimax estimator of  $m$  for the loss function

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

where  $c_{ij}$  are unknown, when it is known that matrix  $C = \|c_{ij}\|$  is nonnegative definite.

Denote

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

In the paper [16] it is proved that when  $C$  is known there always exists the minimax estimator  $d = (d_1, \dots, d_r)$  of parameter  $m$  for which

$$(13) \quad d_i(X) = \frac{X_i + \beta_i \sqrt{n}}{n + \sqrt{n}},$$

where

$$(14) \quad \sum_{i=1}^r \beta_i \leq s, \quad \beta_1 \geq 0, \dots, \beta_r \geq 0.$$

Let  $\beta_1, \dots, \beta_r$  be fixed and let  $A$  be the set of those indices  $i$  for which  $\beta_i > 0$ . From the results of [16] it follows that if

$$(15) \quad \begin{aligned} \sum_{j \in A} (c_{ii} - 2c_{ij}) \beta_j &= v, & \text{when } i \in A, \\ \sum_{j \in A} (c_{ii} - 2c_{ij}) \beta_j &\leq v, & \text{when } i \in R - A, \end{aligned}$$

for some  $v$ , then the estimator (13) is minimax. Condition (15) determines the class of loss functions for which the estimator (13) is minimax. It looks similar to condition (9) but there is one important difference. Here,  $\beta_1, \dots, \beta_r$  must satisfy condition (14).

9. Let  $x(t)$ ,  $t \geq 0$  or  $t = 0, 1, 2, \dots$ , be a stochastic process defined on the probability space  $(\Omega, \mathcal{F}, P_\lambda)$ , where  $\Omega$  is the space of right continuous functions  $\omega = x(\cdot)$  for which the left-side limits exist and  $P_\lambda$  is probability measure defined on  $(\Omega, \mathcal{F})$  determined by parameter  $\lambda$ .

Let  $\mathcal{F}_t$  be the least  $\sigma$ -algebra with respect to which all  $x(s)$  are measurable for  $s \leq t$  and let  $P_{\lambda,t}$  be the restriction of the measure  $P_\lambda$  to the  $\sigma$ -algebra  $\mathcal{F}_t$ . Denote by  $Z(\omega, t): \Omega \rightarrow R$  a  $\mathcal{F}_t$ -measurable mapping that is right continuous with respect to  $t$ ,  $P_\lambda$ -almost surely. Let us suppose that for each  $t \geq 0$  the measures  $P_{\lambda,t}$ ,  $\lambda \in A$ , are absolutely continuous with respect to a measure  $P_{\lambda_0,t}$ ,  $\lambda_0 \in A$ , and that the density function is

$$\frac{dP_{\lambda,t}}{dP_{\lambda_0,t}} = g(t, Z(\omega, t), \lambda, \lambda_0),$$

where  $g$  is a continuous function. Then  $Z(\omega, t)$  is a sufficient statistic for  $\lambda$ .

Let  $T$  denote the set of values  $t$  and let  $\tau(\omega): \Omega \rightarrow T$  be a random variable such that for each  $t \in T$

$$\{\omega: \tau(\omega) \leq t\} \in \mathcal{F}_t$$

and

$$P_\lambda(0 \leq \tau(\omega) < \infty) = 1$$

for each  $\lambda \in \Lambda$ .

The random variable  $\tau(\omega)$  is called a *stopping rule*.

Let  $U = T \times R$  and let  $\mathcal{B}$  be the  $\sigma$ -field of Borel subsets  $U$ . On  $(U, \mathcal{B})$  we define measure  $m$  in the following way. For each  $A \in \mathcal{B}$

$$m_\lambda(A) = P_\lambda((\tau(\omega), Z(\omega, \tau(\omega))) \in A).$$

We observe the process  $x(t)$  up to the moment  $\tau(\omega)$  and we want to estimate parameter  $\lambda$ . An estimator  $d(u) = d(s, t)$  of  $\lambda$  is a  $\mathcal{B}$ -measurable function defined on  $U$  with the values in  $R$ . The pair  $\delta = (t, d)$  consisting of a stopping rule  $\tau$  and an estimator  $d$  is called a *sequential plan*.

Let  $L(\lambda, a)$  denote the loss function associated with the estimate  $a = d(u)$  and let  $c(t)$  be the cost function connected with the observation of the process  $x(s)$  up to the moment  $t$ . We assume that  $c(t)$  is nonnegative continuous and that it satisfies the condition  $\lim_{t \rightarrow \infty} c(t) = \infty$ .

For a given  $\lambda$  the risk connected with the sequential plan  $\delta = (\tau, d)$  is defined by

$$R(\lambda, \delta) = \int_U [L(\lambda, d(u)) + c(t)] m_\lambda(du).$$

In further considerations we restrict ourselves to those sequential plans  $\delta$  for which  $R(\lambda, \delta)$  exists and is finite for each  $\lambda \in \Lambda$ .

For the a priori distribution  $\pi$  of parameter  $\lambda$  the Bayes risk is defined in the usual way

$$r(\pi, \delta) = \int_\Lambda R(\lambda, \delta) \pi(d\lambda),$$

assuming the integral does exist.

Let  $\delta^t = (t, d^t)$  be a plan for which  $\tau(\omega) = t$  for each  $\omega \in \Omega$ . Under this plan, for given  $\pi$ , let  $\pi(\cdot | z(t) = z)$  be the a posteriori distribution of  $\lambda$  for given  $z(\omega, t) = z$  and let

$$(16) \quad r_1(\pi(\cdot | Z(t) = z), d^t) = \int_\Lambda L(\lambda, d^t(z)) \pi(d\lambda | Z(t) = z)$$

be the a posteriori risk associated with  $\pi$  and  $d^t$ . An estimator  $d^{*t}$  is called a *t-Bayes estimator* if for  $d^t = d^{*t}$  the risk (16) attains its minimum.

Let  $\Gamma$  be the set of a priori distributions. Sometimes we have the information that  $\pi \in \Gamma$ . Let  $\Delta$  be the class of all sequential plans  $\delta = (\tau, d)$  for which the Bayes risk  $r(\pi, \delta)$  exists and is finite for all  $\pi \in \Gamma$ . A sequential plan  $\delta_0 = (\tau_0, d_0)$  is called *minimax* if

$$\sup_{\pi \in \Gamma} r(\pi, d_0) = \inf_{\delta \in \Delta} \sup_{\pi \in \Gamma} r(\pi, \delta).$$



Let

$$r_0(\pi, \delta^t) = \int_A R_0(\lambda, \delta^t) \pi(d\lambda),$$

where

$$R_0(\lambda, \delta^t) = \int_U L(\lambda, d(u)) m_\lambda(du).$$

The following theorem holds (see [15]):

**THEOREM.** Suppose that, for every  $t \in T$ , there is a sequence of distributions  $\pi_k$  ( $k = 1, 2, \dots$ ) of parameter  $\lambda$ ,  $\pi_k \in \Gamma$ , for which there are the corresponding  $t$ -Bayes estimators  $d_k^{*t}$  such that the a posteriori risk  $r_1(\pi_k(\cdot | Z(t) = z), d_k^{*t})$  associated with  $\pi_k$  and  $d_k^{*t}$  is independent of  $z$ . Suppose that there is a fixed time plan  $\delta^{t_0} = (t, \hat{d}^{t_0})$ ,  $t_0 \in T$ , such that for each  $\varepsilon > 0$  there is a  $k$  such that for each  $t \in T$

$$r(t_0) + c(t_0) = \sup_{\pi \in \Gamma} r_0(\pi, \delta^{t_0}) + c(t_0) \leq r_1(\pi_k(\cdot | Z(t) = z), d_k^{*t}) + c(t) + \varepsilon.$$

Then the fixed time plan  $\delta^{t_0}$  is a minimax plan.

If we admit all a priori distributions  $\pi$  then we can write in the above  $\sup_{\lambda \in I} R_0(\lambda, \delta^{t_0})$  instead of  $\sup_{\pi \in \Gamma} r_0(\pi, \delta^{t_0})$ .

For a wider and more precise formulation of the problem and for definitions see [15].

**Example 1.** Let  $x(t)$ ,  $t \geq 0$ , be Poisson process

$$P(x(t) = x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t} \quad (x = 0, 1, 2, \dots),$$

where  $\lambda \in (0, \infty)$  is unknown. It is well known that for this process the statistic  $x(t)$  is sufficient for  $\lambda$ . Let the loss function be

$$L(\lambda, a) = k \frac{(a - \lambda)^2}{\lambda} \quad (k > 0),$$

and let the cost function be  $c(t)$ . We assume that  $k$  and  $c(t)$  are unknown. Then applying the theorem (for  $\pi_k$  being gamma distributions) we obtain that if we restrict ourselves to the pairs  $(k, c(t))$  for which

$$k/t_0 + c(t_0) = \inf_{t > 0} [1/t + c(t)]$$

for given  $t_0$ , then the plan  $\delta^{t_0} = (t_0, \hat{d}(x)) = (t_0, x/t_0)$  is minimax (see [3], see also [8] for the exponential family of processes).

Example 2. Let  $x(t)$ ,  $t \geq 0$ , be the Poisson process and let the loss function  $L(\lambda, a)$  be the same as before. Suppose that the family of a priori distributions is restricted to the set  $\Gamma$  for which

$$E(\lambda) = m \quad (m > 0).$$

Then applying the theorem we obtain that if all pairs  $(k, c(t))$  considered (and unknown) satisfy the condition

$$\frac{k}{t_0 + 1/m} + c(t_0) = \inf_{t \geq 0} \left[ \frac{k}{t + 1/m} + c(t) \right]$$

for some  $t_0 \geq 0$ , then the plan  $\hat{\delta}^{t_0} = (t_0, x/(t_0 + 1/m))$  is minimax (see [15]).

Example 3. Let  $x(t)$ ,  $t \geq 0$ , be the Poisson process and suppose now that the a priori distributions are restricted to the set  $\Gamma$  for which

$$E(\lambda) = m, \quad E(1/\lambda) = m_{-1} \quad (m > 0, mm_{-1} > 1).$$

Let  $L(\lambda, a)$  be as before. If all pairs  $(k, c(t))$  considered satisfy the condition

$$\frac{k}{t_0 + m_{-1}/(mm_{-1} - 1)} + c(t_0) = \inf_{t \geq 0} \left[ \frac{k}{t + m_{-1}/(mm_{-1} - 1)} + c(t) \right],$$

then the plan

$$\hat{\delta}^{t_0} = \left( t_0, \frac{x + 1/(mm_{-1} - 1)}{t_0 + m_{-1}/(mm_{-1} - 1)} \right)$$

is minimax (see [15]).

Example 4. Let  $x(t)$ ,  $t \geq 0$ , be the Poisson process and let the loss function be

$$L(\lambda, a) = k\lambda^2(a - 1/\lambda)^2$$

and let the cost function  $c(x)$  be a function of the value  $x$  of the process  $x(t)$  at the moment of stopping,  $c(x) \geq 0$ ,  $c(x) \rightarrow \infty$  if  $x \rightarrow \infty$ .  $k$  and  $c(x)$  are unknown. If all pairs  $(k, c(x))$  considered satisfy the condition

$$\frac{1}{x_0 + 1} + c(x_0) = \inf_{x \in N} \left[ \frac{1}{x + 1} + c(x) \right] \leq 1 + c(0)$$

for some  $x_0 \in N$ ,  $N = \{1, 2, \dots\}$ , then the plan  $\hat{\delta}^{x_0} = (\hat{\tau}, \hat{d}^{x_0})$ , where

$$\hat{\tau} = \inf_{t > 0} \{t: x(t) = x_0\}, \quad \hat{d}^{x_0}(t) = t/(x_0 + 1)$$

(an inverse plan) is minimax (see [15]).

Example 5. Let  $x(t)$  be the Ornstein-Uhlenbeck process.  $x(t)$  is a stationary, Gaussian, Markov process with mean  $\lambda$  and the correlation function

$$R(s) = e^{-\beta|s|} \quad (\beta > 0).$$

It is assumed that  $\beta$  is known. It is well known that for this process the statistic

$$Z(t, \omega) = x(0) + x(t) + \beta \int_0^t x(s) ds$$

is a sufficient statistic for  $\lambda$ . Let the loss function be

$$L(\lambda, a) = k(a - \lambda)^2,$$

where  $a$  is the value of the estimator  $d(z)$ , and let  $c(t)$  be the cost function. The problem is to determine the minimax plan for  $\lambda$  when  $k$  and  $c(t)$  are unknown.

The Radon-Nikodym derivative  $dP_{\lambda,t}/dP_{0,t}$  is

$$\frac{dP_{\lambda,t}}{dP_{0,t}} = \exp \left[ -\frac{\lambda^2}{2} \left( 1 + \frac{\beta t}{2} \right) + \frac{\lambda}{2} Z(\omega, t) \right].$$

Let the a priori distributions  $\pi_k$  ( $k = 1, 2, \dots$ ) in the theorem be defined as follows:  $\pi_k$  has the density

$$g_k(\lambda) = \frac{1}{2} \sqrt{\frac{\beta}{\pi k}} e^{-\beta \lambda^2 / 4k}.$$

Applying the theorem one obtains that if the pair  $(k, c(t))$  considered (and unknown) satisfies the condition

$$\frac{2k}{\beta t_0 + 2} + c(t_0) = \min_{t \geq 0} \left( \frac{2k}{\beta t + 2} + c(t) \right)$$

for some  $t_0 \geq 0$ , then the plan  $\delta^{t_0} = (t, d^{t_0})$  with

$$d^{t_0}(z) = z/(\beta t_0 + 2)$$

is minimax. For a corresponding result (and theorem) see [10].

10. Let  $P_\lambda$ ,  $\lambda = (\lambda_1, \dots, \lambda_r) \in A \subset R$  be a distribution of the random variable  $X$  dependent on parameter  $\lambda$ .  $X$  is observed and an estimate  $d(X) = (d_1(X), \dots, d_r(X))$  of  $\lambda$  is applied. Let  $\{L(\lambda, a), L \in \mathcal{L}\}$  be a family of loss functions. Another problem which can be considered is the following: Find an estimate  $d_1^0(X)$  of parameter  $\lambda_1$  such that there exists an estimator  $d^0(X) = (d_1^0(X), \dots, d_r^0(X))$  which is a minimax estimator of  $\lambda = (\lambda_1, \dots, \lambda_r)$  for each  $L \in \mathcal{L}$ .

Let, for example,  $X = (X_1, \dots, X_r)$  be distributed according to the multinomial law with parameter  $p = (p_1, \dots, p_r)$  and let the loss function be

$$L(p, \hat{a}) = \sum_{i=1}^r c_i (a_i - p_i)^2,$$

where the  $c_i > 0$  are unknown. Consider the estimator

$$d_1^0(X) = X_1 / (n + \sqrt{n}).$$

Then for the existence of the solution corresponding to  $d_1^0(X)$  in the sense given above, it is sufficient to satisfy

$$c_2 \geq \dots \geq c_r \geq c_1 > 0,$$

$$\sum_{i=1}^r (1/c_i) \leq (r-2)/c_1$$

(see Section 7). Notice that the set of vectors  $(c_1, \dots, c_r)$  is here determined only by inequalities.

11. Suppose that the random variable  $X$  has the distribution  $P_\lambda$ , dependent on parameter  $\lambda \in A \subset R^r$ .  $X$  is observed and an estimate  $d(X)$  of  $\lambda$  is applied. Let  $\{L(\lambda, a), L \in \mathcal{L}\}$  be a family of loss functions and let  $\mathcal{P}$  be a family of distributions  $\pi$  of parameter  $\lambda$ . The problem is to find an estimator  $d^0$ , of parameter  $\lambda$ , which minimizes the Bayes risk  $r(\pi, d)$  for each  $(L, \pi) \in \mathcal{R}$ . We call this estimator *Bayes for  $\mathcal{R}$* .

Examples of situations in which such estimators exist are numerous. The problem can also be formulated in the sequential case. We restrict ourselves to a presentation of the solution of the problem in a situation often considered.

Let  $X$  be a random variable with the Poisson distribution and let the loss function be

$$(17) \quad L(\lambda, a) = c(\lambda)(a - \lambda)^2,$$

where  $c(\lambda)$  is unknown; it is known that  $c(\lambda) > 0$  for each  $\lambda \in A = (0, \infty)$ .

Let the Bayes estimator be of the form

$$(18) \quad d(x) = ax + b.$$

It is well known that for the loss function (17) the Bayes estimator is

$$(19) \quad d(x) = \frac{E(\lambda c(\lambda) | X = x)}{E(c(\lambda) | X = x)},$$

assuming the expectations do exist together with  $E(\lambda^2 c(\lambda) | X = x)$ .

Let the a priori distribution  $\pi$  have the density  $g(\lambda)$ . To determine a Bayes estimator we can apply the method used in [14] (for another approach see [5]). From (18) and (19) it follows that

$$\int_0^\infty \lambda^{x+1} e^{-\lambda} c(\lambda) g(\lambda) d\lambda = (ax + b) \int_0^\infty \lambda^x e^{-\lambda} c(\lambda) g(\lambda) d\lambda$$

for  $x = 0, 1, 2, \dots$ . Let us denote

$$B(\lambda) = c(\lambda) g(\lambda) e^{-\lambda}.$$

Then

$$(20) \quad \int_0^{\infty} \lambda^{x+1} B(\lambda) d\lambda = (ax+b) \int_0^{\infty} \lambda^x B(\lambda) d\lambda.$$

But

$$(x+1) \int_0^{\infty} \lambda^x B(\lambda) d\lambda = \lambda^{x+1} B(\lambda)|_0^{\infty} - \int_0^{\infty} \lambda^{x+1} B'(\lambda) d\lambda.$$

Let us suppose that for each  $x = 0, 1, 2, \dots$

$$(21) \quad \lambda^{x+1} B(\lambda)|_0^{\infty} = 0.$$

Then equation (20) takes the form

$$\int_0^{\infty} \lambda^x [a\lambda B'(\lambda) + (\lambda - b + a) B(\lambda)] d\lambda = 0$$

and it surely holds if

$$a\lambda B'(\lambda) + (\lambda - b + a) B(\lambda) = 0.$$

Solving this equation we obtain

$$B(\lambda) = C\lambda^{(b-a)/a} e^{-\lambda/a}$$

which gives

$$(22) \quad c(\lambda)g(\lambda) = C\lambda^{b/a} e^{-\lambda/a}.$$

The integrals considered exist and condition (21) holds if

$$(23) \quad a > 0, \quad b/a > -1.$$

From the above discussion there follows that if all pairs  $(c(\lambda), g(\lambda))$  considered satisfy equation (22) and condition (23) and if  $g(\lambda)$  is the density for which the expectations (19) exist, then the estimator  $d(X) = aX + b$  is Bayes for  $\mathcal{R}$ .

12. At the end let us examine an example from game theory. Player  $A$  has one bullet which he fires in the interval  $[0, 1]$ . Player  $B$  is in the same situation. They fire their bullets independently but the shot of a player is heard by the other player. Let  $P(t)$  be the probability of hitting player  $B$  under the assumption that the bullet is fired at the moment  $t$  and let  $Q(t)$  be the probability of hitting player  $A$  when firing at  $t$ . The functions  $P(t)$  and  $Q(t)$  are called *accuracy functions*. It is assumed that  $P(t)$  and  $Q(t)$  are continuous in  $[0, 1]$ , they increase in this interval and  $P(0) = Q(0) = 0$ ,  $P(1) = Q(1) = 1$ . The game is over if one of the players is hit or at the moment  $t = 1$ . The player  $A$  obtains 1 if he only hits the other player,  $-1$  if he only is hit and 0 in the other cases. It is the simplest example of the so-called noisy duel (see [7]).

If player  $A$  fires his bullet at the moment  $x < 1$  and  $B$  has still his bullet, the best what he can do is to fire his bullet at the moment  $x = 1$ . Assuming that  $A$  and  $B$  play in such a way the payoff function is

$$K(x, y) = \begin{cases} 2P(x) - 1 & \text{if } x < y, \\ P(x) - Q(x) & \text{if } x = y, \\ 1 - 2Q(y) & \text{if } x > y, \end{cases}$$

where  $K(x, y)$  is the expected payoff in the case when  $A$  and  $B$  intend to fire their bullets at the moments  $x, y$ , respectively.

Let  $x_0 \in (0, 1)$  be the moment such that

$$(24) \quad P(x_0) + Q(x_0) = 1.$$

Let us consider the strategy defined as follows: to wait to the moment  $x_0$  and if up to this moment there was no shot of the opponent, to fire at  $x_0$ . Otherwise fire at moment  $x = 1$ . It is easy to verify that such a strategy is minimax.

But a minimax strategy can also be determined if  $P(x)$  and  $Q(x)$  are known only in some neighbourhood of point  $x_0$  satisfying (24) or at point  $x_0$ .

Questions of information play an important role in game theory and in decision theory.

### References

- [1] J. Berger, M. E. Bock, L. Brown, G. Casella and L. J. Gleser, *Minimax estimation of a normal mean vector for arbitrary quadratic loss and unknown covariance matrix*, Ann. Statist. 4 (1977), p. 763-771.
- [2] J. Berger and C. Srinivasan, *Generalized Bayes estimators in multivariate problems*, ibid. 6 (1978), p. 783-801.
- [3] A. Dvoretzky, J. Kiefer and J. Wolfowitz, *Sequential decision problems for processes with continuous time parameter. Problem of estimation*, Ann. Math. Statist. 24 (1953), p. 403-415.
- [4] B. Efron and C. Morris, *Stein's estimation rule and its competitors — an empirical Bayes approach*, J. Amer. Statist. Assoc. 68 (1973), p. 117-130.
- [5] M. Ghosh and G. Meeden, *Admissibility of linear estimators in the one parameter exponential family*, Ann. Statist. 5 (1977), p. 774-778.
- [6] L. J. Gleser, *Minimax estimation of a normal mean vector when the covariance matrix is unknown*, ibid. 7 (1979), p. 838-846.
- [7] S. Karlin, *Mathematical Methods and Theory in Games, Programming and Economics*, Pergamon Press, London 1959.
- [8] R. Magiera, *On sequential minimax estimation for the exponential class of processes*, Zastos. Mat. 15 (1977), p. 445-454.
- [9] R. Róžański, *A modification of Sudakov lemma and efficient sequential plans for the Ornstein-Uhlenbeck process*, ibid. 17 (1980), p. 73-86.

- [10] —, *On minimax sequential estimation of the mean value of a stationary Gaussian-Markov process*, *ibid.* 17 (1981), p. 25-32.
- [11] C. Stein, *A necessary and sufficient condition for admissibility*, *Ann. Math. Statist.* 26 (1955), p. 751-768.
- [12] H. Steinhaus, *A problem of estimation*, *ibid.* 28 (1957), p. 633-648.
- [13] S. Trybuła, *Some problems in simultaneous minimax estimation*, *ibid.* 29 (1958), p. 245-253.
- [14] —, *Two problems of minimax estimation*, *Zastos. Mat.* 14 (1974), p. 195-204.
- [15] —, *Some investigations in minimax estimation theory*, *Dissertationes Mathematicae* 240, Warszawa 1985, 42 pp.
- [16] —, *Minimax estimation and prediction for random variables with bounded sum*, *Zastos. Mat.*, this issue, p. 85-91.
- [17] 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 WROCLAW  
50-370 WROCLAW

*Received on 1983.05.04*

---