

S. TRYBUŁA and K. SZAJOWSKI (Wrocław)

DECISION MAKING IN AN INCOMPLETELY KNOWN STOCHASTIC SYSTEM, II

1. Introduction. In this paper methods of decision theory are applied to determine a minimax control of the stochastic system defined by the equation (1). The situation in which the performance index depends on the parameter of disturbances is considered and is measured in terms of their variance. It is assumed that the horizon of control is a bounded random variable with known distribution and that the disturbances have a distribution belonging to the exponential family. The problem is considered in detail also in the case when a priori information about the parameter of disturbances, resulting in the fixing of an expectation parameter, is given.

The problem of determining a minimax control, often considered in analytical control theory, in the theory of control of stochastic systems is not treated so frequently [1], [8], [10]. Problems in which the disturbances have a distribution different from the normal one are also seldom considered [2], [4], [9]. On the other hand, there are situations in which the disturbances have distributions other than normal ones [2] and the theory assuming the disturbances to have a distribution belonging to the exponential family is more general. This paper is preceded by the paper [11] in which problems of Bayes control for disturbances belonging to the exponential family are considered.

Random variables and stochastic processes belonging to the exponential class are considered in [3], [5], [6], [7].

2. Statement of a minimax control problem. In the paper we use the notation introduced in [11].

Let us consider the discrete linear system with complete observations and random horizon

$$(1) \quad x_{n+1} = \alpha_n x_n + u_n + \gamma_n v_n, \quad x_0 = e, \quad n = 0, 1, \dots, N,$$

where x_n is the state variable, u_n is the control, v_0, v_1, \dots are independent random variables with the same distribution, N is a random variable independent of v_0, v_1, \dots , distributed according to the law

$$P(N = i) = p_i, \quad i = 1, \dots, M, \quad p_M > 0, \quad \sum_{i=1}^M p_i = 1,$$

α_n, γ_n, e are given constants, $\gamma_n \neq 0$.

It is assumed that the random variables v_0, v_1, \dots have the distribution P_λ belonging to the exponential family, i.e. its density with respect to a σ -finite measure μ_0 on $R = (-\infty, \infty)$ is

$$p(v, \lambda) = S(v, q) \exp [qA(\lambda) + vB(\lambda)],$$

where λ is a parameter. We suppose that the parametrization is chosen so as to satisfy the conditions of the paper [11]. Then we have

$$E_\lambda(v_n) = q\lambda$$

for some $q > 0$. Moreover, we additionally assume that

$$E_\lambda(v_n^2) = q_1 \lambda^2 + q_2 \lambda + q_3$$

for some constants q_1, q_2, q_3 . Here, $E_\lambda(\cdot)$ denotes the expectation with respect to the distribution P_λ .

We assume that $p(v, \lambda)$ is known with the only exception of the parameter $\lambda \in A_0$.

Let u_n be a (Borel) function of the data $X_n = (x_0, x_1, \dots, x_n)$ and $U_{n-1} = (u_0, u_1, \dots, u_{n-1})$ available at moment $n = 0, 1, \dots, M$. We call u_n the *control at moment n* .

The vector $U = (u_0, u_1, \dots, u_M)$ is called a *control policy*.

Given the initial state e and the a priori distribution P of the random variable N , choose a control policy $U^{(0)} \in \mathcal{U}$ such that

$$(2) \quad \sup_{\lambda \in A_0} R(\lambda, U^{(0)}) = \inf_{U \in \mathcal{U}} \sup_{\lambda \in A_0} R(\lambda, U),$$

where

$$(3) \quad R(\lambda, U) = B'(\lambda) E_p \left\{ E_\lambda \left[\sum_{i=0}^N (\xi_i x_i^2 + 2\eta_i x_i \lambda + \zeta_i \lambda^2 + k_i u_i) \mid X_0 \right] \right\}$$

and $E_p(\cdot)$ denotes the expectation with respect to the distribution P of the random variable N ; \mathcal{U} is the set of all control policies U for which $R(\lambda, U)$ exists for each $\lambda \in A_0$.

If the equation (2) holds, the policy $U^{(0)}$ is called a *minimax control policy*.

The function $R(\lambda, U)$ is called the *risk function* or shortly the *risk*.

We suppose that

$$\xi_i \zeta_i - \eta_i^2 \geq 0, \quad \xi_i > 0, \quad k_i > 0.$$

The derivative $B'(\lambda)$ is connected with the variance $D_\lambda^2(v_n)$ of the random variable v_n by the formula

$$(4) \quad B'(\lambda) = q/D_\lambda^2(v_n)$$

(see [11]). From (4) it follows that the performance index

$$J = \sum_{i=0}^N (\xi_i x_i^2 + 2\eta_i x_i \lambda + \zeta_i \lambda^2 + k_i u_i^2)$$

is measured in (3) in terms of the variance of the random variable v_n .

3. Recurrence equations for the risk. Let $\pi_k = \sum_{i=k}^M p_i$ and let P_λ be the distribution of the random variables v_0, v_1, \dots . Define

$$R_n(\lambda, U) = B'(\lambda) E_\lambda \left[\sum_{i=n}^M \frac{\pi_i}{\pi_n} (\xi_i x_i^2 + 2\eta_i x_i \lambda + \zeta_i \lambda^2 + k_i u_i^2) | X_n, U_{n-1} \right],$$

where $E_\lambda(\cdot | X_n, U_{n-1})$ denotes the conditional expectation given X_n, U_{n-1} . Then

$$R(\lambda, U) = R_0(\lambda, U)$$

and $R_n(\lambda, U)$, $n = 0, 1, \dots$ satisfy the recurrence equation

$$(5) \quad \begin{aligned} R_n(\lambda, U) &= B'(\lambda) E_\lambda [(\xi_n x_n^2 + 2\eta_n x_n \lambda + \zeta_n \lambda^2 + k_n u_n^2) | X_n, U_{n-1}] + \\ &+ \frac{\pi_{n+1}}{\pi_n} E_\lambda \{ B'(\lambda) E_\lambda \left[\sum_{i=n+1}^M \frac{\pi_i}{\pi_{n+1}} (\xi_i x_i^2 + 2\eta_i x_i \lambda + \zeta_i \lambda^2 + k_i u_i^2) | X_{n+1}, U_n \right] X_n, U_{n-1} \} \\ &= B'(\lambda) (\xi_n x_n^2 + 2\eta_n x_n \lambda + \zeta_n \lambda^2 + k_n u_n^2) + \frac{\pi_{n+1}}{\pi_n} E_\lambda [R_{n+1}(\lambda, U) | X_n, U_{n-1}]. \end{aligned}$$

Let $U_{\beta, r}^* = (u_0^*, u_1^*, \dots, u_M^*)$ be the control policy for which

$$u_M^* = 0,$$

$$(6) \quad u_n^* = - \frac{\frac{\pi_{n+1}}{\pi_n} \alpha_n A_{n+1}}{k_n + \frac{\pi_{n+1}}{\pi_n} A_{n+1}} x_n - \frac{\frac{\pi_{n+1}}{\pi_n} (\gamma_n q A_{n+1} + \beta_{n+1} B_{n+1})}{k_n + \frac{\pi_{n+1}}{\pi_n} A_{n+1}} \cdot \frac{r_n}{\beta_n},$$

where

$$(7) \quad \beta_n = \beta + nq, \quad r_n = r + \sum_{i=0}^{n-1} v_i.$$

The constants A_n, B_n satisfy the equations

$$(8) \quad A_n = \xi_n + \frac{\frac{\pi_{n+1}}{\pi_n} k_n \alpha_n^2 A_{n+1}}{k_n + \frac{\pi_{n+1}}{\pi_n} A_{n+1}},$$

$$(9) \quad B_n = \frac{\eta_n}{\beta_n} + \frac{\pi_{n+1}}{\pi_n} \cdot \frac{\alpha_n k_n (\gamma_n q A_{n+1} + \beta_{n+1} B_{n+1})}{k_n + \frac{\pi_{n+1}}{\pi_n} A_{n+1}} \cdot \frac{1}{\beta_n}$$

and the boundary conditions $A_M = \xi_M, B_M = \eta_M/\beta_M$.

This control policy is Bayes with respect to the a priori distribution of the parameter λ with the density

$$(10) \quad g(\lambda; \beta, r) = D(\beta, r) \exp[\beta A(\lambda) + r B(\lambda)]$$

if $(\beta, r) \in S$ (see [11] for a corresponding result and the definition of the set S).

For the control u_n we obtain

$$\begin{aligned} E_\lambda(x_{n+1} | X_n, U_{n-1}) &= E_\lambda(\alpha_n x_n + u_n + \gamma_n v_n | X_n, U_{n-1}) = \alpha_n x_n + u_n + \gamma_n q \lambda, \\ E_\lambda(r_{n+1} | X_n, U_{n-1}) &= E_\lambda(r_n + v_n | X_n, U_{n-1}) = r_n + q \lambda, \\ E_\lambda(x_{n+1}^2 | X_n, U_{n-1}) &= (\alpha_n x_n + u_n)^2 + 2\gamma_n (\alpha_n x_n + u_n) q \lambda + \gamma_n^2 (q_1 \lambda^2 + q_2 \lambda + q_3), \\ (11) \quad E_\lambda(x_{n+1} r_{n+1} | X_n, U_{n-1}) &= (\alpha_n x_n + u_n + \gamma_n q \lambda) r_n + (\alpha_n x_n + u_n) q \lambda + \gamma_n (q_1 \lambda^2 + q_2 \lambda + q_3), \\ E_\lambda(r_{n+1}^2 | X_n, U_{n-1}) &= r_n^2 + 2q \lambda r_n + q_1 \lambda^2 + q_2 \lambda + q_3. \end{aligned}$$

Denote

$$\alpha_n x_n + u_n^* = F_n x_n + G_n r_n,$$

i.e.

$$(12) \quad F_n = \frac{\alpha_n k_n}{k_n + \frac{\pi_{n+1}}{\pi_n} A_{n+1}}, \quad G_n = -\frac{\pi_{n+1}}{\pi_n} \cdot \frac{\gamma_n q A_{n+1} + \beta_{n+1} B_{n+1}}{k_n + \frac{\pi_{n+1}}{\pi_n} A_{n+1}} \cdot \frac{1}{\beta_n}.$$

From the recurrence equation (5) with the help of equations (11) we prove that

$$(13) \quad R_n(\lambda, U_{\beta, r}^*) = B'(\lambda)(a_n x_n^2 + c_n r_n^2 + 2d_n x_n \lambda + e_n \lambda^2 + 2f_n r_n \lambda + i_n \lambda + j_n),$$

where the constants a_n, \dots, j_n satisfy the recurrence equations

$$\begin{aligned}
a_n &= A_n, \\
c_n &= \frac{\pi_{n+1}}{\pi_n} (G_n^2 a_{n+1} + c_{n+1}) + k_n G_n^2, \\
d_n &= \frac{\pi_{n+1}}{\pi_n} (\gamma_n q F_n a_{n+1} + F_n d_{n+1}) + \eta_n, \\
(14) \quad e_n &= \frac{\pi_{n+1}}{\pi_n} (\gamma_n^2 q_1 a_{n+1} + q_1 c_{n+1} + 2\gamma_n d_{n+1} + e_{n+1} + 2q f_{n+1}) + \zeta_n, \\
f_n &= \frac{\pi_{n+1}}{\pi_n} (\gamma_n q G_n a_{n+1} + q c_{n+1} + G_n d_{n+1} + f_{n+1}), \\
i_n &= \frac{\pi_{n+1}}{\pi_n} (\gamma_n^2 q_2 a_{n+1} + q_2 c_{n+1} + i_{n+1}), \\
j_n &= \frac{\pi_{n+1}}{\pi_n} (\gamma_n^2 q_3 a_{n+1} + q_3 c_{n+1} + j_{n+1})
\end{aligned}$$

and the boundary conditions

$$(15) \quad a_M = \xi_M, \quad d_M = \eta_M, \quad e_M = \zeta_M, \quad c_M = f_M = i_M = j_M = 0.$$

4. A minimax theorem. We are looking for a minimax control policy defined by (2) and (3).

For this purpose let us consider the equations (8), (12) and (14). From (8) it follows that the A_n are independent of β . Then F_n do not depend on β , as well as a_n and d_n . For the remaining constants we write $c_n(\beta)$, $e_n(\beta)$, $f_n(\beta)$, $i_n(\beta)$, $j_n(\beta)$ instead of c_n , e_n , f_n , i_n , j_n to underline their dependence on parameter β . All these constants are independent of the parameters r and λ .

Write

$$\begin{aligned}
Z_1(\beta) &= e_0(\beta), \\
Z_2(\beta, r) &= 2d_0 x_0 + 2f_0(\beta)r + i_0(\beta), \\
Z_3(\beta, r) &= a_0 x_0^2 + c_0(\beta)r^2 + j_0(\beta).
\end{aligned}$$

Let $E_\pi(\cdot)$ denote the expectation with respect to the distribution π of the parameter λ and let U be a control policy. The Bayes risk connected with π and U is defined by

$$r(\pi, U) = E_\pi(R(\lambda, U)).$$

Let $\pi_{\beta,r}$, $(\beta, r) \in S$, be a distribution with density (10) and let $U_{\beta,r}^*$ be the control policy Bayes with respect to $\pi_{\beta,r}$ (see equations (6)–(9)). From (13) and the results of the paper [11] it follows that

$$r(\pi_{\beta,r}, U_{\beta,r}^*) = E_{\pi_{\beta,r}} [B'(\lambda)(Z_1(\beta)\lambda^2 + Z_2(\beta, r)\lambda + Z_3(\beta, r))]$$

$$= \frac{\beta q - q_1 + q^2}{\beta} \cdot \frac{Z_1(\beta) \left(\frac{r}{\beta}\right)^2 + Z_2(\beta, r) \frac{r}{\beta} + Z_3(\beta, r)}{(q_1 - q^2) \left(\frac{r}{\beta}\right)^2 + q_2 \frac{r}{\beta} + q_3} + \frac{Z_1(\beta)}{\beta}.$$

From the corresponding theorem of decision theory (see, for example, [12], p. 374), we obtain the following theorem:

THEOREM 1. Suppose that there is a sequence $\{\pi_{\beta^{(k)}, r^{(k)}}\}$, $k = 1, 2, \dots$, of a priori distributions with density (10), $(\beta^{(k)}, r^{(k)}) \in S$, for which the corresponding sequence $\{U_{\beta^{(k)}, r^{(k)}}^*\}$ of Bayes control policies satisfies the condition

$$\lim_{k \rightarrow \infty} r(\pi_{\beta^{(k)}, r^{(k)}}, U_{\beta^{(k)}, r^{(k)}}^*) = c.$$

If there is a control policy $U^{(0)}$ such that

$$R(\lambda, U^{(0)}) = c$$

for each $\lambda \in \Lambda_0$, then $U^{(0)}$ is a minimax control policy.

From (14) we obtain that

$$(16) \quad c_n(\beta) = \sum_{i=n}^{M-1} t_{in} \frac{1}{\beta_i^2},$$

$$(17) \quad e_n(\beta) = t_n + \sum_{i=n+1}^{M-1} \left(t_{in} \frac{1}{\beta_i} + t'_{in} \frac{1}{\beta_i^2} \right),$$

$$(18) \quad f_n(\beta) = \sum_{i=n}^{M-1} x_{in} \frac{1}{\beta_i} + \sum_{i=n+1}^{M-1} x'_{in} \frac{1}{\beta_i^2},$$

$$(19) \quad i_n(\beta) = q_2 \left(y_n + \sum_{i=n+1}^{M-1} y_{in} \frac{1}{\beta_i^2} \right),$$

$$j_n(\beta) = q_3 \left(z_n + \sum_{i=n+1}^{M-1} z_{in} \frac{1}{\beta_i^2} \right),$$

where only the variables β_i depend on β .

From (9) and (14) it follows that

$$d_n = \beta_n B_n.$$

For the normal distribution (with variance 1)

$$p(v, \lambda) = \frac{1}{\sqrt{2\pi}} e^{-(v-\lambda)^2/2},$$

$$D_\lambda^2(v_n) = 1, \quad S = \{(\beta, r): \beta > 0\}.$$

For the Poisson distribution

$$p(v, \lambda) = \frac{\lambda^v}{v!} e^{-\lambda},$$

$$D_\lambda^2(v_n) = \lambda, \quad S = \{(\beta, r): \beta > 0, r > 0\}.$$

For the gamma distribution

$$p(v, \lambda) = \frac{1}{\Gamma(q) \lambda^q} v^{q-1} e^{-v/\lambda} I_{(0, \infty)}(v),$$

$$D_\lambda^2(v_n) = q\lambda^2, \quad S = \{(\beta, r): \beta > 1, r > 0\},$$

where $I_C(v)$ denotes the characteristic function of the set C .

For the binomial distribution

$$p(v, \lambda) = \binom{q}{v} \lambda^v (1-\lambda)^{q-v},$$

$$D_\lambda^2(v_n) = q\lambda(1-\lambda), \quad S = \{(\beta, r): r > 0, \beta - r > 0\}.$$

For the negative binomial distribution

$$p(v, \lambda) = \frac{\Gamma(q+v)}{\Gamma(q) v!} \cdot \frac{\lambda^v}{(1+\lambda)^{q+v}},$$

$$D_\lambda^2(v_n) = q\lambda(1+\lambda), \quad S = \{(\beta, r): \beta > 1, r > 0\}.$$

Let the random variables v_0, v_1, \dots have a gamma distribution. Then $q_1 = q(q+1)$, $q_2 = q_3 = 0$ and

$$r(\pi_{\beta,r}, U_{\beta,r}^*) = \frac{\beta-1}{\beta} \cdot \frac{e_0(\beta) \left(\frac{r}{\beta}\right)^2 + 2(d_0 x_0 + f_0(\beta)r) \frac{r}{\beta} + a_0 x_0^2 + e_0(\beta)r^2}{\left(\frac{r}{\beta}\right)^2} + \frac{e_0(\beta)}{\beta}.$$

If $x_0 = 0$, r is fixed, $\beta \rightarrow 1+$, then from the above and the conditions (16)–(18) we obtain

$$(20) \quad r(\pi_{\beta,r}, U_{\beta,r}^*) \rightarrow e_0(1).$$

Denote by $U_{1,0}$ the control policy defined by equations (6)–(9) with $\beta = 1$ and $r = 0$. $(1, 0) \notin S$ and the policy $U_{1,0}$ is not Bayes (the distribution (10) does not exist for $\beta = 1$ and $r = 0$) but for this policy equations (11)–(15) hold. Then from (13) and (19) we have

$$(21) \quad R(\lambda, U_{1,0}) = e_0(1)$$

if $x_0 = 0$. From (20), (21) and Theorem 1 it follows that if $x_0 = 0$ then the policy $U_{1,0}$ is a minimax control policy.

Let the random variables v_0, v_1, \dots have the normal distribution and let $U_{\beta,0}^*$ be the control policy defined by (6)–(9) for $r = 0$ and given $\beta > 0$. We have

$$\lim_{\beta \rightarrow 0+} r(\pi_{\beta,0}, U_{\beta,0}^*) = \lim_{\beta \rightarrow 0+} (a_0 x_0^2 + j_0(\beta) + e_0(\beta)/\beta).$$

From (17) and (19) it follows that this limit is finite if $\lim_{\beta \rightarrow 0+} (e_0(\beta)/\beta)$ is finite which holds only for special performance indices.

Let, for example, $P(N = 2) = 1$. In this case from (14) we obtain

$$\begin{aligned} e_0(\beta) &= \gamma_0^2 \left(\xi_1 + \frac{k_1 \alpha_1^2 \xi_2}{k_1 + \xi_2} \right) + 2\gamma_0 \left((\gamma_1 \xi_2 + \eta_2) \frac{\alpha_1 k_1}{k_1 + \xi_2} + \eta_1 \right) \\ &\quad + \frac{(\gamma_1 \xi_2 + \eta_2)^2}{k_1 + \xi_2} \left(\frac{1}{\beta_1^2} - \frac{2}{\beta_1} \right) + \gamma_1^2 \xi_2 + 2\gamma_1 \eta_2 + \zeta_2 + \zeta_1 + \zeta_0 \\ &= \frac{(\gamma_1 \xi_2 + \eta_2)^2}{k_1 + \xi_2} \left(\frac{\beta}{\beta_1} \right)^2 + \frac{1}{\xi_1} (\gamma_0 \xi_1 + \eta_1)^2 + \frac{k_1}{(k_1 + \xi_2) \xi_2} [(\gamma_0 \alpha_1 + \gamma_1) \xi_2 + \eta_2]^2 \\ &\quad + \frac{\xi_1 \zeta_1 - \eta_1^2}{\xi_1} + \frac{\xi_2 \zeta_2 - \eta_2^2}{\xi_2} + \zeta_0. \end{aligned}$$

Then $\lim_{\beta \rightarrow 0+} (e_0(\beta)/\beta)$ is finite if

$$\begin{aligned} \gamma_0 \xi_1 + \eta_1 &= 0, \quad (\gamma_0 \alpha_1 + \gamma_1) \xi_2 + \eta_2 = 0, \\ \zeta_0 &= 0, \quad \xi_1 \zeta_1 - \eta_1^2 = 0, \quad \xi_2 \zeta_2 - \eta_2^2 = 0. \end{aligned}$$

If this is not satisfied then for each K there is an a priori distribution of the parameter λ such that the Bayes risk is greater than K for each control policy.

5. Use of previous experience. Sometimes we have additional information about the parameter λ which results from previous experience.

Let us assume that we know that the a priori distribution of the parameter λ belongs to the class Γ . In this case we define the minimax policy as follows: $U^{(0)} \in \mathcal{U}_\Gamma$ is a minimax control policy if

$$\sup_{\pi \in \Gamma} r(\pi, U^{(0)}) = \inf_{U \in \mathcal{U}_\Gamma} \sup_{\pi \in \Gamma} r(\pi, U).$$

Here \mathcal{U}_Γ denotes the class of all control policies U for which $r(\pi, U)$ exists for each $\pi \in \Gamma$.

The following theorem holds:

THEOREM 2. Suppose that there is a pair $(\beta, r) \in S$ such that the Bayes risk $r(\pi, U_{\beta,r}^*) = \text{const}$ for each $\pi \in \Gamma$, where $U_{\beta,r}^*$ is defined by (6)–(9). If the a

priori distribution $\pi_{\beta,r}$ with the density (10) belongs to the class Γ then the control policy $U_{\beta,r}^*$ is a minimax control policy.

Proof. It is well known from decision theory that a constant Bayes risk strategy which is Bayes with respect to an a priori distribution $\pi \in \Gamma$ is a minimax strategy. Then $U_{\beta,r}^*$ is a minimax control policy.

Let the random variables v_0, v_1, \dots have the normal distribution and let Γ be the class of all a priori distributions π for which

$$E_{\pi}(\lambda^2) = m_2,$$

where $m_2 > 0$ is given. For the a priori distribution $\pi_{\beta,r}$ with density

$$g(\lambda; \beta, r) = \sqrt{\frac{\beta}{2\pi}} \exp \left[-\frac{\beta}{2} \left(\lambda - \frac{r}{\beta} \right)^2 \right]$$

we have

$$E_{\pi_{\beta,r}}(\lambda^2) = \left(\frac{r}{\beta} \right)^2 + \frac{1}{\beta}.$$

Then $\pi_{\beta,r} \in \Gamma$ if

$$(22) \quad \left(\frac{r}{\beta} \right)^2 + \frac{1}{\beta} = m_2.$$

For (β, r) satisfying equation (22) with $\beta > 0$, we obtain

$$r(\pi, U_{\beta,r}^*) = a_0 x_0^2 + c_0(\beta)r + j_0(\beta) + e_0(\beta)m + 2(d_0 x_0 + f_0(\beta)r)E_{\pi}(\lambda)$$

for each $\pi \in \Gamma$. The Bayes risk $r(\pi, U_{\beta,r}^*)$ does not depend on π if

$$(23) \quad d_0 x_0 + f_0(\beta)r = 0.$$

From Theorem 2 it follows that if conditions (22) and (23) hold with $\beta > 0$ then the control policy $U_{\beta,r}^*$ is a minimax control policy.

We prove now a theorem for which Theorems 1 and 2 can be considered as particular cases.

THEOREM 3. Suppose that there is a sequence $\{\pi_{\beta^{(k)}, r^{(k)}}\}$, $k = 1, 2, \dots$, of a priori distributions belonging to Γ and having the density (10), $(\beta^{(k)}, r^{(k)}) \in S$, for which the corresponding sequence $\{U_{\beta^{(k)}, r^{(k)}}^*\}$ of Bayes control policies satisfies the condition

$$(24) \quad \lim_{k \rightarrow \infty} r(\pi_{\beta^{(k)}, r^{(k)}}, U_{\beta^{(k)}, r^{(k)}}^*) = c.$$

If there is a control policy $U^{(0)} \in \mathcal{U}_{\Gamma}$ such that

$$r(\pi, U^{(0)}) = c \quad \text{for each } \pi \in \Gamma,$$

then $U^{(0)}$ is a minimax control policy.

The theorem results from the general theorems of decision theory. Since we could not find it in the literature, we present a formal proof of the theorem.

Suppose that $U^{(0)}$ is not a minimax control policy. Then there is control policy U' such that

$$\sup_{\pi \in \Gamma} r(\pi, U^{(0)}) > \sup_{\pi \in \Gamma} r(\pi, U') = c - \varepsilon,$$

where $\varepsilon > 0$. But, on the other hand,

$$c - \varepsilon = \sup_{\pi \in \Gamma} r(\pi, U') \geq r(\pi_{\beta(k), r(k)}, U') \geq r(\pi_{\beta(k), r(k)}, U_{\beta(k), r(k)}^*)$$

in contrary to (24).

Assume that the random variables v_0, v_1, \dots are distributed according to the Poisson law. Let the class Γ be the class of all a priori distributions π of the parameter λ for which

$$E_{\pi}(\lambda) = m,$$

where $m > 0$ is given. Consider the a priori distribution $\pi_{\beta, r}$ with the density

$$g(\lambda; \beta, r) = \frac{\beta^{r+1}}{\Gamma(r+1)} \lambda^r e^{-\beta\lambda} I_{(0, \infty)}(\lambda),$$

where $\beta = (1 + \varepsilon)/m$ and $r = \varepsilon$. We obtain $E_{\pi_{\beta, r}}(\lambda) = (r+1)/\beta = m$; thus $\pi_{\beta, r} \in \Gamma$. Moreover, for the same parameters $\beta = (1 + \varepsilon)/m$, $r = \varepsilon$ we obtain that the Bayes risk

$$\begin{aligned} & r(\pi_{\beta, r}, U_{\beta, r}^*) \\ &= \frac{\beta}{r} \left[e_0(\beta) \left(\frac{r}{\beta} \right)^2 + (2d_0 x_0 + 2f_0(\beta)r + i_0(\beta)) \frac{r}{\beta} + a_0 x_0^2 + c_0(\beta)r^2 \right] + \frac{e_0(\beta)}{\beta} \end{aligned}$$

has the finite limit

$$c \stackrel{\text{df}}{=} e_0(1/m)m + i_0(1/m)$$

for $\varepsilon \rightarrow 0$ if $x_0 = 0$.

Let $U_{\beta, 0}$ be defined as $\lim_{r \rightarrow 0} U_{\beta, r}^*$. Assuming $x_0 = 0$ we obtain from (13)

$$R(\lambda, U_{\beta, 0}) = e_0(\beta)\lambda + i_0(\beta).$$

Then for $\beta = 1/m$ and any $\pi \in \Gamma$

$$r(\pi, U_{\beta, 0}) = e_0(1/m)m + i_0(1/m) = c$$

which proves that the policy $U_{1/m, 0}$ is a minimax control policy.

In the next two examples we also assume that $x_0 = 0$.

Let the random variables v_0, v_1, \dots have the binomial distribution and

Γ be the class of all a priori distributions π of the parameter λ for which

$$(25) \quad E_{\pi} \left(\frac{1}{1-\lambda} \right) = m,$$

where $m > 1$ is given. For the a priori distribution $\pi_{\beta,r}$ with the density

$$g(\lambda; \beta, r) = \frac{1}{B(r+1, \beta-r+1)} \lambda^r (1-\lambda)^{\beta-r} L_{(0,1)}(\lambda)$$

we have

$$E_{\pi_{\beta,r}} \left(\frac{1}{1-\lambda} \right) = \frac{\beta+1}{\beta-r}.$$

If

$$\beta = 1/(m-1) + m\varepsilon, \quad r = (m-1)\varepsilon,$$

the equation (25) holds and

$$\begin{aligned} r(\pi_{\beta,r}, U_{\beta,r}^*) &= \frac{\beta+1}{\beta} \cdot \frac{e_0(\beta) \left(\frac{r}{\beta} \right)^2 + (2f_0(\beta)r + i_0(\beta)) \frac{r}{\beta} + e_0(\beta)r^2}{-\left(\frac{r}{\beta} \right)^2 + \left(\frac{r}{\beta} \right)} + \frac{e_0(\beta)}{\beta} \\ &\xrightarrow{\varepsilon \rightarrow 0} e_0 \left(\frac{1}{m-1} \right) (m-1) + i_0 \left(\frac{1}{m-1} \right) m \stackrel{\text{df}}{=} c. \end{aligned}$$

On the other hand, for the control policy $U_{1/(m-1),0}$, defined as in the previous example, we have

$$\begin{aligned} R(\lambda, U_{1/(m-1),0}) &= \frac{e_0 \left(\frac{1}{m-1} \right) \lambda^2 + i_0 \left(\frac{1}{m-1} \right) \lambda}{\lambda(1-\lambda)} \\ &= -e_0 \left(\frac{1}{m-1} \right) + \frac{e_0 \left(\frac{1}{m-1} \right) + i_0 \left(\frac{1}{m-1} \right)}{1-\lambda}. \end{aligned}$$

Thus

$$r(\pi, U_{1/(m-1),0}) = c$$

for each $\pi \in \Gamma$ and the policy $U_{1/(m-1),0}$ is a minimax control policy.

Let the random variables v_0, v_1, \dots have the negative binomial distribution and let Γ be the class of all a priori distributions π of the parameter λ for which

$$E_{\pi} \left(\frac{1}{1+\lambda} \right) = m,$$

where $m < 1$ is given. For the distribution $\pi_{\beta,r}$ with the density

$$g(\lambda; \beta, r) = \frac{1}{B(\beta-1, r+1)} \cdot \frac{\lambda^r}{(1+\lambda)^{\beta+r}} I_{(0,\infty)}(\lambda)$$

we have

$$E_{\pi_{\beta,r}} \left(\frac{1}{1+\lambda} \right) = \frac{\beta-1}{\beta+r}.$$

For $\beta = 1/(1-m) + m\varepsilon$, $r = (1-m)\varepsilon$ the distribution $\pi_{\beta,r} \in \Gamma$ and

$$r(\pi_{\beta,r}, U_{\beta,r}^*) \xrightarrow{\varepsilon \rightarrow 0} e_0 \left(\frac{1}{1-m} \right) (1-m) + i_0 \left(\frac{1}{1-m} \right) m \stackrel{\text{df}}{=} c.$$

On the other hand

$$R(\lambda, U_{1/(1-m),0}) = e_0 \left(\frac{1}{1-m} \right) + \frac{i_0 \left(\frac{1}{1-m} \right) - e_0 \left(\frac{1}{1-m} \right)}{1+\lambda}.$$

Thus for any $\pi \in \Gamma$

$$r(\pi, U_{1/(1-m),0}) = c$$

and the policy $U_{1/(1-m),0}$ is a minimax control policy.

Finally we find the minimax controls in the case when the class Γ consists of a priori distributions on which two special kinds of conditions are imposed.

Let the random variables v_0, v_1, \dots have the normal distribution and let the class Γ be the class of all distributions π for which

$$(26) \quad E_{\pi}(\lambda) = m, \quad E_{\pi}(\lambda^2) = m_2,$$

where m, m_2 are given and $m_2 - m^2 > 0$. In this case for each fixed policy $U_{\beta,r}^*$ with $\beta > 0$

$$r(\pi, U_{\beta,r}^*) = \text{const} \quad \text{for each } \pi \in \Gamma.$$

From the above we obtain that if $\pi_{\beta,r} \in \Gamma$ then $U_{\beta,r}^*$ is a minimax control policy.

Conditions (26) give for the distribution $\pi_{\beta,r}$

$$\frac{r}{\beta} = m, \quad \left(\frac{r}{\beta} \right)^2 + \frac{1}{\beta} = m_2,$$

i.e.

$$\beta^* = 1/(m_2 - m^2), \quad r^* = m/(m_2 - m^2)$$

determine the minimax policy U_{β^*,r^*}^* .

In a similar way one can determine minimax control policies U_{β^*, r^*}^* for other distributions of random variables v_0, v_1, \dots belonging to an exponential family.

Let the class Γ be determined by the conditions:
for the Poisson distribution

$$E_{\pi}(\lambda) = m, \quad E_{\pi}(1/\lambda) = m_{-1} \quad (m > 0, mm_{-1} > 1);$$

for the gamma distribution

$$E_{\pi}(1/\lambda) = m_{-1}, \quad E_{\pi}(1/\lambda^2) = m_{-2} \quad (m_{-1} > 0, m_{-2} - m_{-1}^2 > 0);$$

for the binomial distribution

$$E_{\pi}(1/\lambda) = m_{-1}, \quad E_{\pi}(1/(1-\lambda)) = \bar{m}_{-1} \quad (m_{-1} > 1, \bar{m}_{-1} > 1);$$

for the negative binomial distribution

$$E_{\pi}(1/\lambda) = m_{-1}, \quad E_{\pi}(1/(1+\lambda)) = \hat{m}_{-1} \quad (0 < \hat{m} < 1, m_{-1} > \hat{m}/(1-\hat{m})).$$

This leads to the solutions:
for the Poisson distribution

$$\beta^* = m_{-1}/(mm_{-1} - 1), \quad r^* = 1/(mm_{-1} - 1);$$

for the gamma distribution

$$\beta^* = m_{-2}/(m_{-2} - m_{-1}^2), \quad r^* = m_{-1}/(m_{-2} - m_{-1}^2);$$

for the binomial distribution

$$\beta^* = (m_{-1} + \bar{m}_{-1})/(m_{-1}\bar{m}_{-1} - m_{-1} - \bar{m}_{-1}),$$

$$r^* = \bar{m}_{-1}/(m_{-1}\bar{m}_{-1} - m_{-1} - \bar{m}_{-1});$$

for the negative binomial distribution

$$\beta^* = (m_{-1} - \hat{m}_{-1})/(m_{-1} - \hat{m}_{-1} - m_{-1}\hat{m}_{-1}),$$

$$r^* = \hat{m}_{-1}/(m_{-1} - \hat{m}_{-1} - m_{-1}\hat{m}_{-1});$$

U_{β^*, r^*}^* being a minimax control policy.

References

- [1] M. Aoki, *Optimal Bayesian and min.-max. control of class of stochastic and adaptive dynamic systems*, Proc. IFAC Symposium System Engineering for Control System Design, Tokyo 1965, p. 77-84.
- [2] —, *Optimization of stochastic systems*; in: *Topics in discrete time systems*, Academic Press, 1967.
- [3] J. Franz and W. Winkler, *Sequential estimation problems for the exponential class of processes with independent increments*, Scand. J. Statist. 6 (1979), p. 129-139.

- [4] I. Ja. Kac and A. B. Kuržanskiĭ, *Minimax multistep filtration in statistically undetermined situations* (in Russian), *Avtomat. i Telemekh.* 2 (1978), p. 79-87.
- [5] I. Kūchler and U. Kūchler, *Analytical aspects of exponential families of distribution functions*, *Math. Nachr.* 101 (1981), p. 153-164.
- [6] —, *An analytical treatment of exponential families of stochastic processes with independent stationary increments*, *ibid.* 103 (1981), p. 21-30.
- [7] R. Magiera, *On sequential minimax estimation of the exponential class of processes*, *Zastos. Mat.* 15 (1977), p. 445-454.
- [8] D. D. Sworder, *Minimax control of discrete time stochastic systems*, *J. Soc. Ind. Appl. Math. Ser. A, Control* 2 (1964), p. 433-449.
- [9] S. Trybuła, *Sterowanie dualne przy samoreprodukujących się rozkładach*, *Prace V Krajowej Konferencji Automatyki, Tom I, Teoria sterowania*, 1971, p. 163-169.
- [10] S. Trybuła and K. Szajowski, *Minimax control of a stochastic system with disturbances belonging to the exponential family*, *Zastos. Mat.* 18 (1985), p. 525-539.
- [11] —, —, *Decision making in an incompletely known stochastic system, I*, *ibid.*, this issue, p. 31-41.
- [12] S. Zaks, *The theory of statistical inference*, John Wiley and Sons, 1971.

INSTITUTE OF MATHEMATICS
TECHNICAL UNIVERSITY OF WROCLAW
50-370 WROCLAW

Received on 1984.07.05
