Igor Vajda

A discrete theory of search. II

Aplikace matematiky, Vol. 16 (1971), No. 5, 319–335

Persistent URL: http://dml.cz/dmlcz/103365

SVAZEK 16 (1971) APLIKACE MATEMATIKY ČÍSLO 5

A DISCRETE THEORY OF SEARCH II*)

IGOR VAJDA

(Received November 10, 1967)

3. ASYMPTOTIC PROPERTIES OF STRATEGIES

Let us consider the classical statistical decision problem with a (uniformly distributed random) parameter $\theta \in \Omega$ and with a sample $(\zeta_1, \zeta_2, \ldots, \zeta_N)$ which is supposed to be distributed by

$$\mathsf{P}_{\zeta_1 \ldots \zeta_N | \theta} = Q_\theta \otimes Q_\theta \otimes \ldots \otimes Q_\theta \quad (N \text{ times}), \tag{3.1}$$

where Q_θ is a probability distribution on a sample space $Z = \{1, 2, \ldots, k\}$ of the random variables ζ_j. Next let us consider Bayes' estimator $\hat{\theta}_N : Z^N \to \Omega$ of θ and denote by

$$e(N) = \frac{1}{n} \sum_{i=1}^{n} \mathsf{P}_{\zeta_1 \ldots \zeta_N | \theta}[\hat{\theta}_N \neq \theta_i]$$

the average probability of error corresponding to $\hat{\theta}_N$.

Define the α-entropy of two distributions $Q = (q_1, q_2, \ldots, q_k)$, $\tilde{Q} = (\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_k)$ by

$$H_\alpha(Q, \tilde{Q}) = \sum_{r=1}^{k} q_r^\alpha \tilde{q}_r^{1-\alpha}.$$

Obviously, $H_\alpha(Q, \tilde{Q})$ is an analytic and convex function of α in the domain $\alpha \in (0, 1)$. Let us denote

$$\lambda(Q, \tilde{Q}) = \inf_{\alpha \in (0,1)} H_\alpha(Q, \tilde{Q}). \tag{3.2}$$

It is easy to see that $0 \leqq \lambda(Q, \tilde{Q}) \leqq 1$, where $\lambda(Q, \tilde{Q}) = 0$ or 1 iff $Q \perp \tilde{Q}$ (singularity)

*) Part I of this paper has been published in the preceding issue of this journal. The sections, formulas and references are numbered accordingly.

or $Q = \tilde{Q}$ respectively. The following property of $\lambda(Q, \tilde{Q})$ introduced here for later references has been established in Th. 2 and (c) of [7].

Lemma 3.1. *It holds*

$$e(N) = \lambda^{N+o(N)}$$

where

$$\lambda = \max_{i \neq k} \lambda(Q_{\theta_i}, Q_{\theta_k}) \in [0, 1] . \tag{3.3}$$

In Th. 2.2 it has been shown that the probability of error $e_{\delta'}N$ corresponding to an optimum strategy δ' (if it exists) converges to zero at least exponentially (see (2.31)). According to the following theorem the convergence of $e_{\delta'}N$ to zero cannot exceed an exponential rate.

Theorem 3.1. *For any* $(\Omega, \mathscr{E}, \delta)$ *it holds*

$$e_\delta N \geqq [\max_{i \neq k} \prod_{l=1}^{m} \lambda(l \mid i, k)]^{N+o(N)} , \tag{3.4}$$

where

$$\lambda(l \mid i, k) = \lambda(P_l(. \mid \theta_i), P_l(. \mid \theta_k)) . \tag{3.5}$$

Proof. Let $Z = A^m$,

$$Q_\theta = \bigotimes_{l=1}^{m} P_l(. \mid \theta) ,$$

i.e. let ζ_j be a random vector, $\zeta_j = (\xi_{j1}, \xi_{j2}, \ldots, \xi_{jm}) \in A^m$, $j = 1, 2, \ldots$ According to Lemma 3.1, $e(N) = \lambda^{N+o(N)}$, where $\lambda = \max \lambda(Q_{\theta_i}, Q_{\theta_k})$. By (f) in [7],

$$\lambda(Q_{\theta_i}, Q_{\theta_k}) \geqq \prod_{l=1}^{m} \lambda(l \mid i, k)$$

and, consequently,

$$e(N) \geqq [\max_{i \neq k} \prod_{l=1}^{m} \lambda(l \mid i, k)]^{N+o(N)} . \tag{3.6}$$

If now $\delta_1, \delta_2, \ldots$ is a realization of strategy δ, then $\xi = (\xi_{1\delta_1}, \xi_{2\delta_2}, \ldots)$ satisfies (2.6) and, consequently, $\xi_j = \xi_{j\delta_j}$ for any j. Since $\eta_j = (\xi_{j\delta_j}, \delta_j)$ results from ζ_j (by an application of the statistic $d_1(\pi_1), d_2(., \pi_2), \ldots$), it holds $e(N) \leqq e_\delta N$. This together with (3.6) implies (3.4) Q.E.D.

Denote

$$H_\alpha(l \mid i, k) = H_\alpha(P_l(. \mid \theta_i), P_l(. \mid \theta_k)) . \tag{3.7}$$

Theorem 3.2. *If $\delta \equiv \mu$ is a random strategy, then*

$$e_\delta N = \lambda^{N+o(N)} \tag{3.8}$$

for any $(\Omega, \mathscr{E})$, where

$$\lambda = \max_{i \neq k} \left[\inf_{\alpha \in (0,1)} \sum_{l=1}^{m} \mu_l H_\alpha(l \mid i, k) \right] \in [0, 1] . \tag{3.9}$$

It holds $\lambda < 1$ iff $\{P_l(. \mid \theta) : \mu_l > 0\} \subset \mathscr{E}$ separates Ω.

Proof. If $\eta_1, \eta_2, \ldots$ are defined by $\delta \equiv \mu$ as described in Sec. 2 and if we put $\zeta_j = \eta_j$ then, by (2.10), assumptions of Lemma 3.1 are satisfied for

$$Q_\theta(l, r) = \mathsf{P}_{\eta_j \mid \theta}[\eta_j = (l, r)] = \mu_l P_l(r \mid \theta) . \tag{3.10}$$

Thus $e_\delta N = e(N)$ and (3.8) follows from Lemma 3.1. Relation (3.9) holds too, because Q_θ defined by (3.10) satisfies the following relation:

$$H_\alpha(Q_{\theta_i}, Q_{\theta_k}) = \sum_{l=1}^{m} \mu_l H_\alpha(l \mid i, k) \tag{3.11}$$

and it remains to apply (3.3), (3.4).

The parameter λ in (3.8) is, in general, a function of $(\Omega, \mathscr{E}, \delta)$, $\lambda = \lambda(\Omega, \mathscr{E}, \delta)$. The following condition characterizes a very important case where $\lambda(\Omega, \mathscr{E}, \delta)$ can be relatively very easily evaluated or estimated.

We shall say that a random strategy $\delta \equiv \mu$ is homogeneous with a parameter β (relative to $(\Omega, \mathscr{E})$), if the μ-probability of the set of all experiments $P_l(. \mid \theta)$ non-separating $\theta_i \neq \theta_k$ does not depend on i, k, i.e.

$$\beta = \sum_{l \in M(i,k)} \mu_l \quad \text{for all} \quad i \neq k ,$$

where $M(i, k) = \{l \in M : P_l(. \mid \theta_i) = P_l(. \mid \theta_k)\}$.

Obviously, $0 \leqq \beta \leqq 1$ and $\beta = 0$ or 1 iff every experiment $P_l(. \mid \theta) \in \mathscr{E}$ separates Ω or no pair $\theta_i \neq \theta_k$ is separated by $\mathscr{E}$ respectively. In this sense $1 - \beta$ numerically measures how frequently a random design of experiments performed at every time $j = 1, 2, \ldots$ independently and in accordance with the ditsribution μ yields experiments $P_{\delta_j}(. \mid \theta)$ separating a pair of points from Ω. The homogeneity means that the frequency do not depend on the concrete pair considered.

Example 3.1. Let $(\Omega, \mathscr{E})$ be the same as in Example 1.1 and let $\delta \equiv \mu$ be a random strategy with $\mu = (2^{-n}, 2^{-n}, \ldots, 2^{-n})$. Here a half of the functions $f_l \in \mathscr{E}$ satisfies $f_l(\theta_i) \neq f_l(\theta_k)$ independently of $\theta_i \neq \theta_k$. Therefore δ is homogeneous with $\beta = 1/2$.

Example 3.2. Let $(\Omega, \mathscr{E})$ be the same as in Example 1.2 and let $\delta \equiv \mu$ be a random strategy with $\mu = (n^{-1}, n^{-1}, \ldots, n^{-1})$. Then δ is homogeneous with $\beta = (n - 2)/n$.

Example 3.3. Let $\Omega = \{\theta_1, \ldots, \theta_4\}$, $\mathscr{E} = \{f_1, \ldots, f_5\}$, where the matrix of values $f_l(\theta_i)$, $i = 1, \ldots, 4$, $l = 1, \ldots, 5$, is

$$\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

If $\delta \equiv \mu$ with $\mu = (1/5, 1/5, \ldots, 1/5)$, then δ is not homogeneous. If, however, $\mu = (1/3, 1/3, 0, 0, 1/3)$, then δ is homogeneous with $\beta = 2/3$.

Theorem 3.3. *If δ is a random homogeneous strategy with a parameter β, then λ which appears in (3.8) satisfies the inequality $\lambda \geqq \beta$. The equality holds iff all $P_l(.\mid\theta)$ such that $\mu_l > 0$ and all pairs $\theta_i \neq \theta_k$ satisfy either the relation $P_l(.\mid\theta) = P_l(.\mid\theta_k)$ or $P_l(.\mid\theta_i) \perp P_l(.\mid\theta_k)$. If $\mathscr{E} = \{f_1, f_2, \ldots, f_m\}$, then the last condition is satisfied and $\lambda = \beta$.*

Proof. (3.7) and (3.8) yield

$$\lambda = \max_{i \neq k} \left[\beta + \inf_{\alpha \in (0,1)} \sum_{l \in \overline{M}(i,k)} \mu_l H_\alpha(l \mid i, k)\right], \tag{3.12}$$

where $\overline{M}(i, k) = M - M(i, k)$. This immediately implies the inequality $\lambda \geqq \beta$ as well as the necessary and sufficient condition for the equality (notice that $H_\alpha(l \mid i, k) = 0$ or 1 iff $P_l(.\mid\theta_i) \perp P_l(.\mid\theta_k)$ or $P_l(.\mid\theta_i) = P_l(.\mid\theta_k)$ respectively).

4. CODING MODEL

If Ω is arbitrary and $\mathscr{E} = \{f_1, f_2, \ldots, f_m\}$ contains arbitrary mutually different functions $f_l : \Omega \to A$, we shall say that $(\Omega, \mathscr{E})$ defines a coding (noiseless) model. This terminology is motived by the fact that any strategy δ considered in the framework of this model defines a code

$$\begin{aligned} \bar{\varepsilon}_1 &= (f_{\delta_1}(\theta_1), f_{\delta_2}(\theta_1), \ldots, f_{\delta_N}(\theta_1)) \\ \bar{\varepsilon}_2 &= (f_{\delta_1}(\theta_2), f_{\delta_2}(\theta_2), \ldots, f_{\delta_N}(\theta_2)) \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \bar{\varepsilon}_n &= (f_{\delta_1}(\theta_n), f_{\delta_2}(\theta_n), \ldots, f_{\delta_N}(\theta_n)), \end{aligned} \tag{4.1}$$

The code of size n and length N ($N = 1, 2, \ldots$) will be denoted in the sequel by $\varepsilon(n, N)$. Since in this case it is supposed that the experimenter observes values which are equal to one of the code-words, namely $(\xi_1, \xi_2, \ldots, \xi_N) = (f_{\delta_1}(\theta), f_{\delta_2}(\theta), \ldots, f_{\delta_N}(\theta))$, we speak about a noiseless model.

A more general (noisy) coding model of the theory of search can be obtained if we suppose

$$\xi_j = f_{\delta_j}(\theta) + \zeta_j (\text{mod } a) \quad j = 1, 2, \dots, \tag{4.2}$$

where $\zeta_1, \zeta_2, \dots$ are mutually and on θ independent random variables assuming values $0, 1, \dots, a-1$ from A with the corresponding probabilities $1-p$, $p/(a-1), \dots$ $\dots, p/(a-1)$, $p \in [0, 1)$. In this case the experimenter observes the input code-word $(f_{\delta_1}(\theta), \dots, f_{\delta_N}(\theta))$ at the output of a symmetric memoryless channel defined by the following channel probabilities matrix

$$\begin{pmatrix} 1-p & \dfrac{p}{a-1} & \dots & \dfrac{p}{a-1} \\ \dfrac{p}{a-1} & 1-p & \dots & \dfrac{p}{a-1} \\ \dots & \dots & \dots & \dots \\ \dfrac{p}{a-1} & \dfrac{p}{a-1} & \dots & 1-p \end{pmatrix}. \tag{4.3}$$

Thus the coding model of the theory of search is defined by a triple $(\Omega, \mathscr{E}, p)$, where $\mathscr{E} = \{f_1, \dots, f_m\}$, $p \in [0, 1)$. If $p = 0$, the model reduces to the noiseless one. Now we will discuss in more detail the relation between strategies and coding for symmetric memoryless channels in the framework of this simple model. At the same time, this discussion will also indicate relations between the search theory and the information theory.

First we notice that if $\delta_1, \delta_2, \dots$ are not defined uniquely by the strategy δ, then also the code $\varepsilon(n, N)$ is not defined uniquely by δ. If, for example, $\delta \equiv \mu$ is a random strategy, then $\varepsilon(n, N)$ should also be interpreted as a random code defined by μ. Denote by $e(\varepsilon(n, N))$ an average probability of error (taken with respect to P_δ) corresponding to the code $\varepsilon(n, N)$. Obviously,

$$e(\varepsilon(n, N)) = e_\delta N\,. \tag{4.4}$$

Thus, for example, a necessary condition for $e_\delta N = 0$ is $p = 0$ and

$$n \leqq a^N\,. \tag{4.5}$$

Equality (4.4) means that if δ is an (asymptotically) "efficient" strategy, then $\varepsilon(n, N)$ is an (asymptotically) "efficient" code. The converse is not verbally true. Indeed, if $\varepsilon(n, N)$ is an efficient code (defined for $N = 1, 2, \dots$) it need not necessarily mean that there exist $\delta_1, \delta_2, \dots, \delta_N$ or a strategy $\delta = (\delta_1, \delta_2, \dots)$ such that (4.1) is satisfied.

If

$$\bar{\varepsilon}_1 = (\varepsilon_{11}, \varepsilon_{12}, \ldots, \varepsilon_{1N})$$
$$\bar{\varepsilon}_2 = (\varepsilon_{21}, \varepsilon_{22}, \ldots, \varepsilon_{2N}) \tag{4.6}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\bar{\varepsilon}_n = (\varepsilon_{n1}, \varepsilon_{n2}, \ldots, \varepsilon_{nN})$$

is a code of size n and length N with a code-alphabet A, then the converse is true only if for every $j = 1, 2, \ldots, N$ there exists $f_{l_j} \in \mathscr{E}$ such that

$$(f_{l_j}(\theta_1), f_{l_j}(\theta_2), \ldots, f_{l_j}(\theta_n)) = (\varepsilon_{1j}, \varepsilon_{2j}, \ldots, \varepsilon_{nj}) . \tag{4.7}$$

If this is the case, we can define $\delta_j = l_j$, $j = 1, 2, \ldots, N$. But if the class $\mathscr{E}$ is not ample enough, the condition (4.7) need not be satisfied for every j. Therefore, for a general $\mathscr{E}$, a strategy δ can be optimum (with respect to $(\Omega, \mathscr{E})$) even if the corresponding code $\varepsilon(n, N)$ is not optimum in the class of all codes at the input of the channel (4.3).

In view of this relation between the theory of search and the information theory, it is clear that every coding problem or coding theorem of the information theory can be interpreted in the framework of the theory of search (see a generalized coding model below) but not conversely. Because of various restrictions concerning $\mathscr{E}$ imposed by real experimental restrictions in the praxis, problems of the theory of search are more specific and cannot be solved by a direct application of coding theorems. Any concrete form of $\mathscr{E}$ (see Examples 1.1.–1.3.) restricts a structure of codes for which the "isomorphism" condition (4.7) can be satisfied. Thus some optimisation problems of the search theory seem to be not interesting from the point of view of the information theory (cf. Sec. 6, in particular Example 6.2), although in the information theory some codes with a specific structure are also studied (e.g. linear codes).

Let us remark that a further generalization of the coding model $(\Omega, \mathscr{E}, p)$ is also possible. Instead of the symmetric channel (4.3) with input and output alphabet A one can consider a general noiseless channel with input alphabet A, output alphabet $B = \{0, 1, \ldots, b - 1\}$ and a channel probabilities matrix

$$\mathscr{P} = \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{0b-1} \\ p_{10} & p_{11} & \cdots & p_{1b-1} \\ \cdots & \cdots & \cdots & \cdots \\ p_{a-10} & p_{a-11} & \cdots & p_{a-1b-1} \end{bmatrix} . \tag{4.8}$$

Here the noise is non-additive and (4.4) must be replaced by

$$\mathsf{P}_{\xi_j|\delta_j=l,\theta}[\xi_j = s] = P_l(s \mid \theta) = p_{rs}, \quad \text{where} \quad r = f_l(\theta), \quad s \in B . \tag{4.9}$$

Thus the generalized noisy coding model is defined either by a pair $(\Omega, \mathscr{E})$ where $P_l(. \mid \theta) \in \mathscr{E}$ satisfy (4.9), or by a triple $(\Omega, \mathscr{E}, \mathscr{P})$ where $\mathscr{E} = \{f_1, f_2, \ldots, f_m\}$ and $\mathscr{P}$ is a stochastic matrix (4.8). Clearly, the generalized model also admits the asymmetric "isomorphism" between strategies of search for $(\Omega, \mathscr{E})$ and codes for the channel $(A, \mathscr{P}, B)$ as described above. In the framework of this model the Shannon's problem of transmissibility of a source through a memoryless channel can be interpreted as a special problem of a more general statistical problem of the theory of search (see Sec. 6 below).

Example 4.1. If $(\Omega, \mathscr{E})$ is defined as in Example 1.1, then the condition (4.7) can be satisfied for any code $\{\bar{\varepsilon}_1, \bar{\varepsilon}_2, \ldots, \bar{\varepsilon}_n\}$. If this code is random, i.e. if ε_{ij} in (4.6) are independent realizations of a random variable assuming values 0 and 1 with probabilities 1/2, 1/2, then the strategy $\delta = (\delta_1, \delta_2, \ldots,)$ defined by (4.1) is random $\equiv \mu$, where $\mu_l = 2^{-n}$, $l = 1, 2, \ldots, 2^n$. Using (4.4), all results concerning random codes can be applied to random strategies δ in the framework of a coding model $(\Omega, \mathscr{E}, \mathscr{P})$ with

$$\mathscr{P} = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0b-1} \\ p_{10} & p_{11} & \cdots & p_{1b-1} \end{pmatrix}$$

arbitrary. In the special case

$$\mathscr{P} = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \tag{4.10}$$

the model reduces to $(\Omega, \mathscr{E}, p)$, where the corresponding communication channel is binary symmetric.

Theorem 4.1. *If in the framework of a coding model $(\Omega, \mathscr{E}, p)$, δ is a homogeneous random strategy with a parameter β, then λ in (3.8) is given by*

$$\lambda = \beta + (1-\beta)\left[2\sqrt{\left(\frac{p(1-p)}{a-1}\right)} + \frac{(1-p)(a-2)}{a-1}\right]. \tag{4.11}$$

Proof. By (4.9) and (4.3), the distributions $P_l(\,.\mid\theta)$ are of the form

$$\left(\frac{p}{a-1}, \ldots, \frac{p}{a-1}, 1-p, \frac{p}{a-1}, \ldots, \frac{p}{a-1}\right).$$

The position of $1-p$ in this probability vector depends on θ and l. If $\theta_i \neq \theta_k$ and $l \in \overline{M}(i, k)$, then the position corresponding to l, θ_i and l, θ_k is different and consequently (see (3.7))

$$H_\alpha(l \mid i, k) = \frac{p(a-2)}{a-1} + \left(\frac{p}{a-1}\right)^\alpha (1-p)^{1-\alpha} + (1-p)^\alpha \left(\frac{p}{a-1}\right)^{1-\alpha}.$$

This yields

$$\inf_{\alpha\in(0,1)} H_\alpha(l \mid i, k) = 2\sqrt{\left[\frac{p(1-p)}{a-1}\right]} + \frac{p(a-2)}{a-1},$$

where the infimum is attained for $\alpha = 1/2$. This together with (3.12) yields (4.11).

Corollary. *If* $(\Omega, \mathscr{E})$ *is a noiseless coding model and* δ *is homogeneous with a parameter* β, *then*

$$e_\delta N = \beta^{N+o(N)}\,. \tag{4.12}$$

This result also follows from Th. 3.3., because in the noiseless case the condition $P_l(.\mid\theta_i) = P_l(.\mid\theta_k)$ or $P_l(.\mid\theta_i) \perp P_l(.\mid\theta_k)$ is satisfied.

In the rest of this section we shall consider the noiseless coding model. Our aim will be to give more precision to the expression $o(N)$ in (4.12).

Lemma 4.1. *Let us consider a noiseless coding model and let* δ *be an arbitrary strategy of search. If* $M^N_{ik} \subset M^N$ *is the set of all* $(\delta, \delta_2, \ldots, \delta_N) \in M^N$ *such that there exist exactly k values of* θ *such that*

$$(f_{\delta_1}(\theta), f_{\delta_2}(\theta), \ldots, f_{\delta_N}(\theta)) = (f_{\delta_1}(\theta_i), f_{\delta_2}(\theta_i), \ldots, f_{\delta_N}(\theta_i)) \tag{4.13}$$

then

$$e_\delta N = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{n}\left(1 - \frac{1}{k}\right)\mathsf{P}_{\delta_1\ldots\delta_N\mid\theta_i}[(\delta_1 \ldots \delta_N) \in M^N_{ik}]\,. \tag{4.14}$$

Proof. Let $\theta = \theta_i$ and let $(\eta_1, \eta_2, \ldots, \eta_N) = ((\delta_1, \xi_1), \ldots, (\delta_N, \xi_N))$. If $(\delta_1, \ldots, \delta_N) \in M^N_{ik}$, then the maximum in (2.26) is attained for all θ satisfying (4.13) because for all these θ it is $\xi_j = f_{\delta_j}(\theta)$, $j = 1, 2, \ldots, N$ with probability 1. If the value of $\hat{\theta}_N$ is chosen from these θ randomly with a uniform distribution, the probability of error is $1 - 1/k$. In symbols,

$$\mathsf{P}_{\xi_1\ldots\xi_N\mid\theta_i,\delta_1\ldots\delta_N}[\hat{\theta}_N \neq \hat{\theta}_i] = 1 - \frac{1}{k}$$

for all $(\delta_1, \ldots, \delta_N) \in M^N_{ik}$. Hence

$$\mathsf{P}_{\eta_1\ldots\eta_N\mid\theta_i}[\hat{\theta}_N \neq \theta_i] = \sum_{k=1}^{n}\left(1 - \frac{1}{k}\right)\mathsf{P}_{\delta_1\ldots\delta_N\mid\theta_i}[(\delta_1 \ldots \delta_N) \in M^N_{ik}]\,.$$

This together with (2.27) implies (4.17).

Corollary. *It holds*

$$\frac{1}{2n}\sum_{i=1}^{n} p(\delta, N, i) \leqq e_\delta N \leqq \frac{1}{n}\sum_{i=1}^{n} p(\delta, N, i) \quad \textit{for every } \delta\,, \tag{4.15}$$

where

$$p(\delta, N, i) = \mathsf{P}_{\delta_1\ldots\delta_N\mid\theta_i}[(\delta_1 \ldots \delta_N) \in M^N - M^N_{i1}]\,. \tag{4.16}$$

Example 4.2. Let $(\Omega, \mathscr{E}, \delta)$ be the same as in Example 3.1 or 4.1. It follows from the definition of δ that $\mathsf{P}_{\delta_1 \ldots \delta_N | \theta_i}[(\delta_1 \ldots \delta_N) \in M_{i1}^N] = (1 - 2^{-N})^{n-1}$ independently of i so that, by (4.15),

$$\tfrac{1}{2}[1 - (1 - 2^{-N})^{n-1}] \leqq e_\delta N \leqq 1 - (1 - 2^{-N})^{n-1} \,. \tag{4.17}$$

Hence $e_\delta N = 2^{-N + o(N)}$. This is in accordance with (4.12) and with the fact that δ is homogeneous with $\beta = 1/2$.

Example 4.3. Let $(\Omega, \mathscr{E}, \delta)$ be the same as in Example 3.2 and suppose that $\delta_1, \delta_2, \ldots, \delta_N$ is a realization of a strategy δ and that $\theta = \theta_i$ where θ_i is fixed. Two different cases will be considered: (a) $f_i \in \{f_{\delta_1}, f_{\delta_2}, \ldots, f_{\delta_N}\}$, (b) $f_i \notin \{f_{\delta_1}, f_{\delta_2}, \ldots, f_{\delta_N}\}$. In the case (a), $(\xi_1, \xi_2, \ldots, \xi_N) = (0, 0, \ldots, 0, 1, 0, \ldots, 0)$ where the position of 1 depends on the position (or positions) of the experiment f_i in the vector $(f_{\delta_1}, f_{\delta_2}, \ldots \ldots, f_{\delta_N})$. In the case (b), ξ_j are identically zero. If (2.26) is respected, then $\hat{\theta}_N \neq \theta$ may appear (with probability $1 - n^{-1}$) in the case (b) only. The probabilityof the case (b) is $(1 - n^{-1})^N$, so that

$$e_\delta N = \left(1 - \frac{1}{n}\right)^{N+1} . \tag{4.18}$$

This is also in accordance with (4.12) and with the result of Example 3.2.

Using relation (4.15), asymptotic formulas for $e_\delta N$ can be obtained from analogous formulas for $p(\delta, N, i)$. We will show that a somewhat strengthened condition of homogeneity of random strategies makes it possible to find out more precise asymptotic formulas for $p(\delta, N, i)$ than those of the form (4.12).

We shall say that a random strategy $\delta \equiv \mu$ is homogeneous with parameters β, γ if it is homogeneous with the parameter β and

$$\gamma = \sum_{l \in M(i,k,r)} \mu_l \tag{4.19}$$

holds for all mutually different i, k, r, where $M(i, k, r) = \{l \in M : f_l(\theta_i) = f_l(\theta_k) = f_l(\theta_r)\}$. Obviously, $1 - \gamma$ numerically measures how frequently a random design of experiments performed at every time $j = 1, 2, \ldots$ independently and in accordance with the distribution μ yields experiments $f_{\delta_j} \in \mathscr{E}$ separating triplets of points from Ω.

Lemma 4.2. *If $\delta \equiv \mu$ is homogeneous with parameters β, γ, then $\beta \geqq \gamma$. If $A = \{0, 1\}$, $\mu = (1/m, 1/m, \ldots, 1/m)$ is uniform and δ is homogeneous with a parameter β, then it is homogeneous with paarmeters β, γ as well. In this case $\beta \geqq (n - 2)/2(n - 1)$ and if $\beta \leqq 1/2$, then*

$$\gamma \leqq \beta^2 \,. \tag{4.20}$$

Proof. The first statement is evident, the other one has been proved in Lemma 3a and Lemma 4 of [2].

Example 4. 3. The strategy from Example 3.1 is homogeneous with parameters β, γ. This follows from Lemma 4.2 and from the fact that $\beta = 1/2$. It is easy to see that $\gamma = 1/4$. The strategy from Example 3.2 is homogeneous with parameters $\beta = (n - 2)/n$, $\gamma = (n - 3)/n$.

Lemma 4.3. *If δ is a homogeneous random strategy with parameters β, γ, then*

$$(n - 1)\,\beta^N - \binom{n-1}{2}\gamma^N \leqq p(\delta, N, i) \leqq (n - 1)\,\beta^N\,, \quad i = 1, 2, \ldots, n\,. \tag{4.21}$$

Proof. Let i, N be arbitrarily fixed and define $M_k = \{f_{\delta_j}(\theta_i) = f_{\delta_j}(\theta_k),\ j = 1, 2, \ldots, N\} \subset M^N$. Obviously

$$\bigcup_{k \neq i} M_k = M^N - M_{i1}^{N} \tag{4.22}$$

and by the definition*) of β, γ,

$$\mathsf{P}_{\delta_1 \ldots \delta_N}[(\delta_1 \ldots \delta_N) \in M_k] = \beta^N\,, \quad k \neq i\,, \tag{4.23}$$

$$\mathsf{P}_{\delta_1 \ldots \delta_N}[(\delta_1 \ldots \delta_N) \in M_k \cap M_r] = \gamma^N\,, \quad k \neq i \neq r\,. \tag{4.24}$$

Since every class of events M_k from an algebra on which a measure P is defined satisfies the inequalities

$$\sum_{k \neq i} P(M_k) - \sum_{k \neq r \neq i} P(M_k \cap M_r) \leqq P(\bigcup_{k \neq i} M_k) \leqq \sum_{k \neq i} P(M_k)\,,$$

(4.21) follows from (4.22)–(4.24).

Theorem 4.2. *Let $(\Omega, \mathscr{E})$ be a noiseless coding model and δ a random strategy of search homogeneous with a parameter β. Then*

$$e_\delta N \leqq (n - 1)\,\beta^N \quad N = 1, 2, \ldots \tag{4.25}$$

holds. If moreover δ is homogeneous with parameters β, γ, then

$$e_\delta N \geqq \frac{n-1}{2}\,\beta^N - \frac{1}{2}\binom{n-1}{2}\gamma^N \quad N = 1, 2, \ldots \tag{4.26}$$

Proof. See Lemma 4.3 and its proof and (4.15).

5. MODEL WITH INCREASING Ω

As we emphasized in the preceding sections, a large number n of elements in Ω is one of the implicit assumptions of our model. In our exposition above n fixed and N tending to infinity have been considered. In this section we will study the asymptotic

*) For random strategies $\mathsf{P}_{\delta_1 \ldots \delta_N | \theta} = \mathsf{P}_{\delta_1 \ldots \delta_N}$ holds.

behaviour of $e_\delta N$ under the assumption that both n and N tend to infinity, but N functionally depends on $n, N = N_n$ (we suppose $N_1 \leqq N_2 \leqq \ldots$, $\lim_n N_n = +\infty$).

From the point of view of the practice, there is a need for small rates of convergence of N_n to infinity but, on the other hand, if the rate were too small, the probability of error $e_\delta N_n$ might converge to zero slowly or might not converge to zero at all.

In general, we shall assume that $\mathscr{E}$ as well as δ may depend on n, $\mathscr{E} = \mathscr{E}(n)$, $\delta = \delta(n)$. More precisely, we shall assume that $\mathscr{E}(n) = \{P_l^{(n)}(. \mid \theta), l = 1, 2, \ldots, m_n\}$ and $\delta(n) \equiv \mu(n) = (\mu_1^{(n)}, \mu_2^{(n)}, \ldots, \mu_{m_n}^{(n)})$, where $P_l^{(n)}(. \mid \theta)$ are conditional probability distributions on $A = \{0, 1, \ldots, a_n - 1\}$ and $\mu_l^{(n)} > 0$ are elements of a probability distribution $\mu(n)$ (we emphasize that in the rest of this paper only random strategies $\delta = \delta(n)$ will be considered).

Any triple $(\Omega, \mathscr{E}(n), \delta(n))$, $n = 1, 2, \ldots$ (or $n = n_0, n_0 + 1, \ldots$) defines a model of search with increasing Ω (i.e. with $n \to \infty$). In the same manner as above we can define a probability of error $e_{\delta(n)} N_n$. In what follows we shall be interested in the behaviour of $e_{\delta(n)} N_n$ (rate of convergence to zero) for $n \to \infty$. In particular, we shall be interested in the problem of the minimum rate of $N_n \to \infty$ for which

$$\lim_n e_{\delta(n)} N_n = 0 \tag{5.1}$$

holds or for which the convergence in (5.1) is exponential. As it will be shown in Sec. 6, under some assumptions concerning $\mathscr{E}(n)$ this problem reduces to the well-known Shannon's problem of transmissibility of information sources through communication channels. In the present section some general results will be stated.

Let us define (cf. (3.7), (3.9))

$$H_\alpha^{(n)}(l \mid i, k) = H_\alpha(P_l^{(n)}(. \mid \theta_i), \quad P_l^{(n)}(. \mid \theta_k)), \quad \alpha \in (0, 1), \tag{5.2}$$

$$\lambda_n = \max_{i \neq k} \left[\inf_{\alpha \in (0,1)} \sum_{l=1}^{m_n} \mu_l^{(n)} H_\alpha^{(n)}(l \mid i, k) \right], \tag{5.3}$$

$$\lambda = \liminf_n \lambda_n \,. \tag{5.4}$$

The case

$$\lambda = \lim_n \lambda_n \tag{5.5}$$

will be of special interest for us.

We shall say that $e_{\delta(n)} N_n$ converges to zero exponentially if there exists $\lambda_0 \in [0, 1)$ such that $e_{\delta(n)} N_n \leqq \lambda_0^{N_n}$ for all sufficiently large N_n.

Theorem 5.1. *The sequence $e_{\delta(n)} N_n$ converges to zero exponentially for all sufficiently fast increasing sequences N_n iff $\lambda < 1$. If $\lambda = 1$, $\varepsilon_{\delta(n)} N_n$ does not converge to zero exponentially for any N_n satisfying the following condition*

$$\lim_n \frac{N_n}{\log n} = +\infty \,. \tag{5.6}$$

If $\lambda < 1$, then for every N_n satisfying the conditions

$$\lim_n \frac{m_n a_n \log N_n}{N_n} = 0 , \tag{5.7}$$

$$\lim_n N_n \mu_l^{(n)} P_l(k \mid \theta) \geqq c > 0 \quad \textit{uniformly with respect to } k, l, \theta , \tag{5.8}$$

$e_{\delta(n)} N_n$ *converges to zero exponentially. If* (5.5) *holds, then*

$$e_{\delta(n)} N_n = \lambda^{N_n} + {}^{o(N_n)} . \tag{5.9}$$

Proof. For the sake of simplicity we shall suppose that (5.5) holds. The corresponding modification will be obvious. Let i_n, k_n be the value of i, k for which $\lambda_n = \lambda(Q_{i_n}, Q_{k_n})$ holds in (5.3), where Q_i's are probability distributions defined on $\{1, 2, \ldots, m_n\} \otimes A$ by $Q_i(l, s) = \mu_l^{(n)} P_l^{(n)}(s \mid \theta)$ (cf. (3.10)). According to (h) in [7],

$$1 - [\tfrac{1}{2} \operatorname{Var} (Q_{i_n} Q_{k_n})]^2 \geqq \lambda_n .$$

Thus, if $\lim_n \lambda_n = 1$ then for any $\varepsilon > 0$ there exists n_0 such that for $n > n_0$, $\operatorname{Var} (Q_{i_n}, Q_{k_n}) \leqq 2\varepsilon/(2 + \varepsilon)$. According to the second inequality on p. 493 in [7],

$$1 - \tfrac{1}{2} \operatorname{Var} (Q_{i_n}^{N_n}, Q_{k_n}^{N_n}) \geqq [1 - \tfrac{1}{2} \operatorname{Var} (Q_{i_n}, Q_{k_n})]^{N_n}$$

(by Q^N we denote the Cartesian product $Q \otimes \ldots, \otimes Q$ (N times)), so that

$$- \frac{1}{N_n} \log (1 - \tfrac{1}{2} \operatorname{Var} (Q_{i_n}^{N_n}, Q_{k_n}^{N_n})) \leqq -\log \left(1 - \frac{\varepsilon}{2 + \varepsilon}\right) < \frac{\varepsilon}{2} . \tag{5.10}$$

By (2.12) in [6] it holds

(5.11)

$$\frac{1}{2n} \max_{i \neq k} [1 - \tfrac{1}{2} \operatorname{Var} (Q_i^{N_n}, Q_k^{N_n})] \leqq e_{\delta(n)} N_n \leqq n \max_{i \neq k} [1 - \tfrac{1}{2} \operatorname{Var} (Q_i^{N_n}, Q_k^{N_n})] .$$

The left-hand part of this inequality and (5.10) yield

$$- \frac{1}{N_n} \log e_{\delta(n)} N_n \leqq \frac{\varepsilon}{2} + \frac{\log 2n}{N_n} .$$

By (5.6) there exists n_1 such that for $n > n_1$, $(\log 2n)/N_n < \varepsilon/2$ and, consequently, $e_{\delta(n)} N_n$ converges to zero exponentially Q.E.D. The second statement of Th. 5.1. easily follows from the following fact proved in § 8 of [1]: If (5.6)–(5.8) hold, then it holds for every i, k

$$\lim_N \left| - \frac{1}{N} \log [1 - \tfrac{1}{2} \operatorname{Var} (Q_i^N, Q_k^N)] + \log \lambda(Q_i, Q_k) \right| = 0 .$$

Analogously as Th. 3.3, the following theorem can be proved.

Theorem 5.2. *If $\delta(n)$ is homogeneous with a parameter β_n, then for λ_n* in (5.3) *or λ in* (5.4) *the following inequalities hold*

$$\lambda \geqq \beta_n\,, \quad \lambda \geqq \liminf_n \beta_n\,. \tag{5.12}$$

The inequalities take place iff for all $P_l^{(n)}(.\mid\theta) \in \mathscr{E}(n)$ and $i \neq k$ it holds either $P_l^{(n)}(.\mid\theta_i) = P_l^{(n)}(.\mid\theta_k)$ or $P_l^{(n)}(.\mid\theta_i) \perp P_l^{(n)}(.\mid\theta_k)$.

The applicability of this theorem can be illustrated by the following

Example 5.1. Let $\mathscr{E}(n) = \mathscr{E}$, $\delta(n) = \delta$ be defined as in Example 3.2 for $n = 1, 2, \ldots$ Then $\delta(n)$ is homogeneous with $\beta_n = (n-2)/n$ and, by (5.12), $\lambda = 1$. Hence, by Th. 5.1 $\varepsilon_{\delta(n)}N_n$ does not converge to zero exponentially (cf. (4.18)).

6. CODING MODEL WITH INCREASING Ω

Throughout this section we shall consider a model of search given by $(\Omega, \mathscr{E}(n), \delta(n), N_n)$. Here $\mathscr{E}(n)$ is described by a class $\{f_l^{(n)}, l = 1, 2, \ldots, m_n\}$, $f_l^{(n)} : \Omega \to A = \{0, 1, \ldots \ldots, a-1\}$ (i.e. A does not depend on n), and by a stochastic matrix defined either by (4.3) (model with additive noise) or by (4.8) (general coding model), which is also supposed to be independent of n. In other words,

$$P_l^{(n)}(s \mid \theta) = p_{rs} \qquad \text{where} \quad r = f_l^{(n)}(\theta) \quad (\text{see } (4.9)) \tag{6.1}$$

holds for every $P_l^{(n)}(.\mid\theta) \in \mathscr{E}(n)$. The strategy $\delta(n)$ is supposed to be random, $\delta(n) \equiv \mu(n)$ and N_n is a non-decreasing with $\lim_n N_n = +\infty$. Under these, assumptions the following variant of Th. 1 obviously holds:

Theorem 6.1. *Let us consider a model with an additive noise $(\Omega, \mathscr{E}(n) = \{f_l^{(n)}, l = 1, 2, \ldots, m_n\}, p)$. If $\delta(n)$ is homogeneous with a parameter β_n, then λ_n defined in* (5.3) *satisfies the relation*

$$\lambda_n = \beta_n + (1-\beta_n)\left[2\sqrt{\left(\frac{p(1-p)}{a-1}\right)} + \frac{p(a-2)}{a-1}\right]. \tag{6.2}$$

This theorem combined with (5.9) enables us usually to find out an asymptotic expression for $e_{\delta(n)}N_n$.

Example 6.1. Let Ω, $\mathscr{E}(n) = \mathscr{E}$, $\delta(n) = \delta$ be the same as in Example 4.1, $n = 1, 2, \ldots$ Since $\delta(n)$ is homogeneous with $\beta_n = 1/2$, (6.2) and (5.4) imply

$$\lambda = \tfrac{1}{2}[1 + 2\sqrt{(p(1-p))}]\,. \tag{6.3}$$

Thus, by (5.9) for all sufficiently fast increasing N_n the following asymptotic formula holds:

$$e_{\delta(n)}N_n = 2^{-N_n[1-\log(1+2\sqrt{(p(1-p))})+o(1)]}\,. \tag{6.4}$$

Let us now consider the general model $(\Omega, \{f_l^{(n)}, l = 1, 2, \ldots, m_n\}, \mathscr{P}, \delta(n))$. In view of the correspondence between the random strategies and the random codes discussed in Sec. 4, one of the problems of Sec. 5, namely, for which N_n relation (5.1) holds, can be considered as a generalization of Shannon's problem on random coding at the input of a memoryless channel $(A, \mathscr{P}, B)$. The generalization consists in the fact in our case the structure of the class $\{f_l^{(n)}, l = 1, 2, \ldots, m_n\}$ must be respected (see the condition (4.7)). In the special case when $m_n = a^n$ (i.e. when $\{f_l^{(n)}, l = 1, 2, \ldots, m_n\}$ contains all mappings $\Omega \to A$) both the problems are identical. In this case a theorem of Shannon asserts that for an appropriately chosen $\mu(n)$ there exists a constant $C(\mathscr{P}) > 0$ such that (5.1) holds or not depending on whether

$$\text{(6.5)} \qquad \liminf_n \frac{N_n}{\log n} > C(\mathscr{P})^{-1}$$

or

$$\text{(6.6)} \qquad \limsup_n \frac{N_n}{\log n} < C(\mathscr{P})^{-1}$$

respectively.

To compare this result with what was said above let us consider a simple model with an additive noise where $a = 2$, $\mu(n) = (2^{-n}, 2^{-n}, \ldots, 2^{-n})$ and (4.10) holds. As it is shown for example in Chap. 9 of [8], in this case the above stated result holds with $C(\mathscr{P}) = 1 - H(p)$, where $H(p) = -p \log p - (1 - p) \log (1 - p)$. If

$$\text{(6.7)} \qquad \lim_n \frac{N_n}{\log n} = R^{-1},$$

then at the same place formulas for $\alpha(R)$, $R \in (0, +\infty)$ can be found such that

$$\text{(6.8)} \qquad e_{\delta(n)} N_n = 2^{-N_n[\alpha(R) + o(1)]}\,.$$

$\alpha(R)$ is a non-decreasing function of the parameter R called a rate of transmission in the information theory. If $R > C(\mathscr{P})$, then $\alpha(R) = 0$. Let us notice that it follows from (6.4) (relation (6.4) has been proved under the assumption (5.6), i.e. for $R = 0$) that

$$\text{(6.9)} \qquad \alpha(0) = 1 - \log [1 + 2 \sqrt{(p(- p))}]\,.$$

In the noiseless case (see Example 4.1) we obtain from (4.17) for $N = N_n$ satisfying (6.7) that a necessary and sufficient condition for (5.1) is that $R < 1$. This is in accordance with what was said above, because in the noiseless case $C(\mathscr{P}) = 1$.

Remark at this place that our asymptotic formula (5.9), analogous to (6.8), has been derived under the condition of zero transmission rate. This is obviously a very strong condition and the case $R = 0$ is not very interesting for the information theory itself. But in the theory of search, in view of various restrictions concerning the class $\mathscr{E}$, there exist situations where (5.1) can be satisfied only under the zero transmission rate condition. This is illustrated by the following

Example 6.2. Let $(\Omega, \mathscr{E}(n), \delta(n))$ be the same as in Example 5.1. According to (4.18),

$$\frac{1}{2}\left(1-\frac{1}{n}\right)^{N_n} \leqq e_{\delta(n)}N_n \leqq \left(1-\frac{1}{n}\right)^{N_n}$$

holds for $N = N_n$. Hence (5.1) holds iff

$$\lim_n \frac{N_n}{n} = +\infty \,. \tag{6.10}$$

Hence if $R > 0$, (5.1) does not hold.

In the following theorem Lemma 4.2 and Th. 4.2 are applied to the coding model with increasing Ω.

Theorem 6.2. *Let $(\Omega, \mathscr{E}(n) = \{f_l^{(n)}, l = 1, 2, \ldots, m_n\})$ be an arbitrary coding noiseless model and let $\delta(n) \equiv \mu(n)$ be homogeneous with a parameter β_n. If*

$$\lim_n \left(N_n \log \frac{1}{\beta_n} - \log n\right) = +\infty \,, \tag{6.11}$$

then (5.1) *holds. If $A = \{0, 1\}$, $\mu(n) = (m_n^{-1}, m_n^{-1}, \ldots, m_n^{-1})$, and*

$$\lim_n \left(N_n \log \frac{1}{\beta_n} - \log n\right) = \varrho \,, \tag{6.12}$$

then

$$\begin{aligned} \exp(-\varrho) - \tfrac{1}{4}\exp(-2\varrho) &\leqq \liminf_n e_{\delta(n)}N_n \leqq \\ &\leqq \limsup_n e_{\delta(n)}N_n \leqq \exp(-\varrho)\,. \end{aligned} \tag{6.13}$$

7. CONCLUDING REMARKS

a) Let N_n satisfy (6.7) and let us consider the noiseless model $(\Omega, \mathscr{E}(n))$ with $A = \{0, 1\}$. As it has been said above, for $R > 1$, (5.1) cannot be satisfied by any random strategy, $\delta(n) \equiv \mu(n) = (m_n^{-1}, m_n^{-1}, \ldots, m_n^{-1})$. The question is what is the maximum R, say $R_0 = R_0(\mathscr{E}(n))$ for which (5.1) holds or, more precisely, under which conditions such $R_0 \in [0, 1]$ exists. For example, if $m_n = 2^n$, i.e. if $\mathscr{E}(n)$ contains all mappings $\Omega \to \{0, 1\}$, then $R_0 = 1$. If $\mathscr{E}(n)$ is such that $\delta(n)$ is homogeneous with parameters $\beta, \gamma, < \beta$, then $R_0 = -\log \beta$. The R_0 is an analogy of the capacity $C(\mathscr{P})$ in the noisy case.

b) As it has been shown in Sec. 6, if $\mathscr{E}(n) = \{f_l^{(n)}, l = 1, 2, \ldots, m_n\}$ is ample enough, the random search gives asymptotically the same probability of error as the "best possible" systematic (sequential) search. In particular, it follows from (6.13) that if $\beta_n \approx 1/2$, $\log n + \varrho$ random experiments in question make it possible to determine θ with the probability $1 - \exp(-\varrho)$ while the best systematic strategy requires at least $\log n$ experiments. It can be seen that for relatively small values of ϱ the probability of error $\leqq \exp(-\varrho)$ is satisfactorily small.

c) Two basic concepts of the paper, namely, the separability of Ω by $\mathscr{E}$ and the homogeneity of random strategies have been used for the first time, in a somewhat different form and in the framework of the coding noiseless model, by A. Rényi [2]. Their generalization given in the present paper seems to be fruitful.

d) Theorems 3.2, 3.3, 4.1, 5.1, 5.2 and 6.1 have been first stated (without proofs) in [1]. Lemma 4.2 and 4.3 are due to A Rényi [2] (Th. 6.2 is also a modification of a result in [2]). The remainder is new.

References

[1] *I. Vajda:* On random design of experiments (in Russian). Information Transmission Problems 4 (1968), N. 4, 26—36.

[2] *A. Rényi:* On the theory of random search. Bull. Amer. Math. Soc. 71 (1965), N. 1, 809—828.

[3] *J. Nedoma:* Coding of finite sets of messages. Presented at 2nd Czechoslovak Conf. on Cybernetics, Prague 16.—19. XI. 1965.

[4] *A. Feinstein:* Foundations of Information Theory, McGraw Hill, N. Y. 1958.

[5] *I. Vajda:* Rate of convergence of the information in a sample concerning a parameter. Czechoslovak Math. Journal, 17 (1967), N. 2, 225—231.

[6] *I. Vajda:* On statistical decision problems with finite parameter space. Kybernetika 3 (1967), (N. 5, 451—466.

[7] *I. Vajda:* On the convergence of information contained in a sequence of observations. Proc. Coll. on Inf. Theory, Debrecen 1967, Budapest 1969.

[8] *R. M. Fano:* Transmission of information. M.I.T. Press & Wiley, N.Y. 1961.

Souhrn

DISKRÉTNÍ TEORIE VYHLEDÁVÁNÍ II

IGOR VAJDA

První část této práce byla uveřejněna v předchozím čísle Aplikací matematiky. Tam byla hlavní pozornost věnována formulacím základních modelů a studiu strategií, jejichž optimálnost je měřena středním počtem pozorování, nutných k bezchybnému vyhledání hodnoty neznámého parametru θ. Zde se studují vlastnosti strategií, jejichž optimálnost se posuzuje podle asymptotického chování bayesovské chyby.

Mějme m jednoparametrických populací výběrových rozložení $\mathscr{E} = \{P_1(. \mid \theta), \ldots$ $\ldots, P_m(. \mid \theta)\}$. Předpokládá se, že při strategii $\delta = (\delta_1, \delta_2, \ldots)$ je posloupnost pozorování $\xi_1, \acute{a}_2, \ldots$ rozložena podle $P_{\delta_1}(. \mid \theta) \otimes P_{\delta_2}(. \mid \theta) \otimes \ldots$ Kvalita strategií se

posuzuje podle chování průměrné chyby $e_\delta N$ Bayesova estimátoru $\hat{\theta}(\delta_1, \ldots, \delta_N, \xi_1, \ldots$ $\ldots, \xi_N)$ při $N \to \infty$.

V kapitole 3 je dokázáno, že pro žádnou δ nemůže $e_\delta N$ konvergovat k nule rychleji než exponenciálně a že exponenciální rychlosti se dosahuje při náhodných strategiích. Pro náhodné strategie je odvozen vzorec (3.9) pro výpočet parametru $\lambda(\mathscr{E}, \delta)$ exponenciální konvergence. Dále je tam vyčleněna třída homogenních strategií, pro které se velmi snadno odhadne anebo vypočte $\lambda(\mathscr{E}, \delta)$. V kapitole 5 jsou tyto výsledky zobecněny na případ, kdy $N = N_n$ roste v závislosti na počtu hodnot n parametru θ.

V kapitole 4 je definován kódovací model teorie vyhledávání, kde $P_t(. \mid \theta) \in \mathscr{E}$ mají jistou speciální strukturu. Zde se za jistých podmínek statistické úlohy nalézt optimální náhodnou strategii, resp. náhodnou strategii, při kterých N_n roste, za podmínky $e_\delta N_n \to 0$, co nejpomaleji do nekonečna, redukují na Shannonovy úlohy o optimálním kódování, resp. o přenesitelnosti informačních zdrojů kanály. Jsou také nalezeny jednoduché vzorce pro $\lambda(\mathscr{E}, \delta)$ příslušné homogenním náhodným strategiím ((4.11), (6.2)). Ve (4.25), (4.26), (6.13) jsou upřesněny asymptotické výrazy pro $e_\delta N$ za předpokladu, že δ vyhovuje silnější podmínce homogenity. V rámci kódovacího modelu se rovněž ukazuje, že kvalita náhodných strategií může být asymptoticky blízká anebo dokonce identická kvalitě nejlepších systematických strategií.

Author's address: Igor Vajda, CSc., Ústav teorie informace a automatisace ČSAV, Vyšehradská 49, Praha 2.