Jiří Cochlar

The optimum sequential test of a finite number of hypotheses for statistically dependent observations

Kybernetika, Vol. 16 (1980), No. 1, (36)--47

Persistent URL: http://dml.cz/dmlcz/124261

KYBERNETIKA — VOLUME *16* (1980), NUMBER 1

The Optimum Sequential Test of a Finite Number of Hypotheses for Statistically Dependent Observations

Jiří Cochlar

The paper deals with Bayes optimum sequential tests of a finite number of hypotheses for independent and differently distributed observations. The obtained results are applied to the problem of Bayes optimum sequential test for distinguishing wheather one of a finite number of known signals is present in a coloured Gaussian noise. The paper generalizes the results of [1], which deals with a case of two hypotheses, to the case of a finite number of hypotheses when the cost c_n of an observation $\mathbf{x}_n$ depends on the index n. There is proved in Theorems 2.1 and 2.2 that for independent and generally differently distributed observations the Bayes optimum sequential test is always the test of a posteriori probability.

1. INTRODUCTION

In this paper we shall deal with the Bayes optimum sequential tests of a finite number of disjoint statistical hypotheses for statistically dependent vector observations. The present results are a generalisation of results of [1], where the same problem is introduced for the case of two hypotheses and for a constant cost of one observation. For the solution of our problem we shall use general results derived in [2]. In this chapter we shall define some necessary concepts and an exact formulation of the problem. The more detailed explanation of introduced concepts contain references [1] and [2].

For an arbitrary fixed integer H, $H \geqq 2$ we shall define

$$\mathscr{H} \triangleq \{1, 2, \ldots, H\}$$

as a *set of hypotheses* and

$$\mathscr{A} \triangleq \{1, 2, \ldots, H\}$$

as a *set of possible decisions* of a statistician. Let the $H \times H$ *matrix of losses* $\mathbf{L}$ be given with elements L_{ij}. Element L_{ij} corresponds to the loss of the statistician due to accepting the decision $j \in \mathscr{A}$ when hypothesis $i \in \mathscr{H}$ holds. We assume $0 < L_{ij} < < \infty$ for $i \neq j$, $L_{ii} = 0$.

Let a measurable space $(\Omega, \mathscr{F})$ and H probability measures $\boldsymbol{P}_i$ on this space $(i \in \mathscr{H})$ be given. Every triple $(\Omega, \mathscr{F}, \boldsymbol{P}_i)$ then represents a probability space corresponding to the validity of a hypothesis i for $i \in \mathscr{H}$. Let $\pi \triangleq ({}^1\pi, \ldots, {}^H\pi)$ be *a priori probability distribution* on $\mathscr{H}$, i.e. $0 \leqq {}^i\pi$, $i \in \mathscr{H}$, $\sum_{i=1}^{H} {}^i\pi = 1$. We shall denote by $\prod$ the set of all possible distributions π. For an arbitrary $\pi \in \prod$ we shall define a probability measure $\boldsymbol{P}$ on $(\Omega, \mathscr{F})$ by the relation

$$\boldsymbol{P}(A) \triangleq \sum_{i=1}^{H} {}^i\pi \, \boldsymbol{P}_i(A) \tag{1.1}$$

for every $A \in \mathscr{F}$. Let N be the set of all positive integers, i.e. $N \triangleq \{1, 2, \ldots\}$, and let an arbitrary fixed $M \in N$ be given. Let $\mathbf{E}$ be a M-dimensional Euclidean space and let $\mathscr{B}$ be a σ-algebra of Borel subsets of the set $\mathbf{E}$. We assume that a sequence of $\mathscr{F}/\mathscr{B}$-measurable functions $\mathbf{x}_n : \Omega \to \mathbf{E}$, $n \in N$ is given on the measurable space $(\Omega, \mathscr{F})$. We shall denote

$$\mathscr{X}_n \triangleq (\mathbf{x}_1, \ldots, \mathbf{x}_n)$$

$\mathscr{X}_n$ is then a $\mathscr{F}/\mathscr{B}^n$-measurable function: $\Omega \to \mathbf{E}^n$. We shall assume that $\mathbf{x}_n$ and $\mathscr{X}_n$ are random elements defined on $(\Omega, \mathscr{F}, \boldsymbol{P}_i)$ for all $n \in N$ and every $i \in \mathscr{H}$. Let a probability density ${}^iw_n(\mathscr{X}_n)$ exists for every $n \in N$ and $i \in \mathscr{H}$.

Random vectors $\mathbf{x}_n$ and n-tuples $\mathscr{X}_n$ can be understood as random elements defined on $(\Omega, \mathscr{F}, \boldsymbol{P})$ and it holds for the probability density $w_n(\mathscr{X}_n)$ on this space

$$w_n(\mathscr{X}_n) = \sum_{i=1}^{H} {}^i\pi \, . \, {}^iw_n(\mathscr{X}_n) \, . \tag{1.2}$$

In [2] there is made a construction of probability spaces $(\Omega, \mathscr{F}, \boldsymbol{P})$ and $(\Omega, \mathscr{F}, \boldsymbol{P}_i)$, with here required properties, rising from the given densities w_n, ${}^iw_n$ respectively.

Vector $\mathbf{x}_n$, for $n \in N$, will be called the *n-th observation* of a statistician. We shall assume that the statistician's decisions on validity of a hypothesis $i \in \mathscr{H}$ are based only on his knowledge of $\mathscr{X}_n$ for $n = 1, 2, \ldots$. We shall denote by c_n the cost which statistician pays for obtaining of value $\mathbf{x}_n$, i.e. c_n is a *cost of the n-th observation* $\mathbf{x}_n$. Everywhere further we shall assume

$$0 \leqq c_n \leqq +\infty \, , \quad \sum_{n=1}^{\infty} c_n = +\infty \, . \tag{1.3}$$

Let $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots \subset \mathscr{F}$ be a nondecreasing sequence of such σ-algebras that $\mathscr{F}_n$ is a minimum σ-algebra induced by $\mathscr{X}_n$ for every $n \in N$. By the *stopping rule* τ on a sequence $\{\mathbf{x}_n\}$ we shall understand every integer random variable τ defined on $(\Omega, \mathscr{F}, \boldsymbol{P})$ which has values from N and for which it holds

$$(1.4) \qquad \boldsymbol{P}(\{\tau < \infty\}) = 1\,, \quad \{\omega : \tau(\omega) = n\} \in \mathscr{F}_n \quad \text{for} \quad n \in N\,.$$

By the *rule of a terminal decision* d we understand every sequence of functions $d_n : \mathbf{E}^n \to \mathscr{A}$ for which it holds for an arbitrary $j \in \mathscr{A}$

$$(1.5) \qquad \{\mathscr{X} \in \mathbf{E}^n : d_n(\mathscr{X}) = j\} \in \mathscr{B}^n \quad \text{for} \quad n \in N\,.$$

Definition 1.1. By *the sequential test* of a set of hypotheses $\mathscr{H}$ we understand every pair (d, τ) (where d is an arbitrary rule of a terminal decision, τ is an arbitrary stopping rule on a sequence $\{\mathbf{x}_n\}$) when statistician accepts a decision $j = d_n(\mathscr{X}_n(\omega))$ then and only then if it holds $\tau(\omega) = n$ for $\omega \in \Omega$.

Let $\pi_n(\mathscr{X}_n) \triangleq ({}^1\pi_n(\mathscr{X}_n), \ldots, {}^H\pi_n(\mathscr{X}_n)) \in \prod$ be *a posteriori probability distribution* on $\mathscr{H}$ for a given n-tuple $\mathscr{X}_n$. It holds

$$(1.6) \qquad {}^i\pi_n(\mathscr{X}_n) = \frac{{}^i\pi \, . \, {}^iw_n(\mathscr{X}_n)}{w_n(\mathscr{X}_n)} \quad \text{for} \quad n \in N\,, \quad i \in \mathscr{H}\,, \quad \mathscr{X}_n \in \mathbf{E}^n\,.$$

Let h be a real nonnegative function defined on $\prod$ by the relation

$$(1.7) \qquad h(\mathbf{t}) \triangleq \min_{j \in \mathscr{A}} \{ \sum_{i=1}^{H} L_{ij} \, . \, {}^it\} \quad \text{for} \quad \mathbf{t} = ({}^1t, \ldots, {}^Ht) \in \prod\,.$$

For every $n \in N$ let us define on the probability space $(\Omega, \mathscr{F}, \boldsymbol{P})$ a real random variable y_n by the relation

$$(1.8) \qquad y_0(\omega) \triangleq -h(\pi_n(\mathscr{X}_n(\omega))) - \sum_{k=1}^{n} c_k \quad \text{for} \quad n \in N\,.$$

Note that y_n is $\mathscr{F}_n$-measurable function of ω for every $n \in N$.

By *the Bayes rule of a terminal decision* we shall understand a rule of a terminal decision $d^* \triangleq \{d_n^*\}_n$ defined by the relation

$$(1.9) \qquad d_n^*(\mathscr{X}) \triangleq \min_{j \in \mathscr{A}} \{j : \sum_{i=1}^{H} L_{ij} \, . \, {}^i\pi_n(\mathscr{X}) = h(\pi_n(\mathscr{X}))\} \quad \mathscr{X} \in \mathbf{E}^n\,, \quad n \in N$$

Definition 1.2. By the *Bayes optimum sequential test* of the set of hypotheses $\mathscr{H}$ (for the given $\pi \in \prod$) we shall understand the sequential test (d^*, τ^*), where d^* is the Bayes rule of a terminal decision and τ^* is a stopping rule on a sequence $\{\mathbf{x}_n\}$ if it holds

$$(1.10) \qquad M(y_{\tau^*}) = \sup_{\tau \in \mathscr{C}} M(y_\tau)$$

where $\boldsymbol{M}$ is the expected value on the space $(\Omega, \mathscr{F}, \boldsymbol{P})$ and $\mathscr{C}$ is the set of all possible stopping rules on a sequence $\{\mathbf{x}_n\}$.

Note 1.1. The stopping rule τ^* satisfying (1.10) will be called *the Bayes optimum stopping rule.*

Theorem 1.1. Let $\pi_n(\mathscr{X}_n)$ be defined for all $n \in N$ and let it hold

$$-h(\pi) < \sup_{\tau \in \mathscr{C}} \boldsymbol{M}(y_\tau) . \tag{1.11}$$

Then the Bayes optimum sequential test of a set of hypotheses $\mathscr{H}$ always exists for a given $\pi \in \prod$.

Proof. The existence of the Bayes optimum stopping rule follows directly from the Theorem 4.1 of [2]. If the condition (1.11) is not satisfied, then the stopping rule τ^* otherwise exists, as it follows e.g. from Theorem 2 of [3], but it has no practical importance, since an accepting of a decision $j \in \mathscr{A}$ without observations brings the minimum risk to the statistician, for which it holds

$$\sum_{i=1}^{H} L_{ij} \,.\, {}^i\pi = h(\pi) .$$

We shall not deal any more with the case when (1.11) does not hold.

In the next chapter of this paper (using Theorem 4.1 of [2]) we shall derive the Bayes optimum sequential test for the case of independent and differently distributed observations. Similarly as in [1] we shall use this result for some case of dependent observations in chapter 3. According to the results of chapter 3, in chapter 4 we shall solve a problem of the optimum sequential test of a presence of one signal from the finite set of known signals in the coloured Gaussian noise.

2. OPTIMUM SEQUENTIAL TEST FOR INDEPENDENT DIFFERENTLY DISTRIBUTED OBSERVATION

We shall assume in this chapter that for $n \in N$ and $i \in \mathscr{H}$ the following equation holds for the probability densities ${}^i w_n$

$${}^i w_n(\mathscr{X}_n) = \prod_{k=1}^{n} {}^i f_k(\mathbf{x}_k) , \quad \mathscr{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n) \tag{2.1}$$

where ${}^i f_n(\mathbf{x}) \geqq 0$ for $n \in N$, $\mathbf{x} \in \mathbf{E}$, $i \in \mathscr{H}$ and $\int_{\mathbf{E}} {}^i f_n(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 1$ for $n \in N$, $i \in \mathscr{H}$. This is evidently a case of statistically independent and generally differently distributed observations $\mathbf{x}_n$.

 It is easy to show that it holds in this case for every $n \in N$, $i \in \mathscr{H}$

$$ {}^i\pi_{n+1}(\mathscr{X}_{n+1}) = \frac{{}^i\pi_n(\mathscr{X}_n) \cdot {}^if_{n+1}(\mathbf{x}_{n+1})}{\sum_{s=1}^{H} {}^s\pi_n(\mathscr{X}_n) \cdot {}^sf_{n+1}(\mathbf{x}_{n+1})} \tag{2.2} $$

where $\mathscr{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, $\mathscr{X}_{n+1} = (\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{x}_{n+1})$.

Let $\mathscr{Z}$ be an arbitrary set. By *an expansion of the set* $\mathscr{Z}$ we shall understand any finite system of sets $\{{}^k\mathscr{Z}\}_k$, ${}^k\mathscr{Z} \subset \mathscr{Z}$, $k \in N_R$, where N_R is an arbitrary finite set of finite integers and it holds for the set ${}^k\mathscr{Z}$

$$ \bigcup_{k \in N_R} {}^k\mathscr{Z} = \mathscr{Z}\,, \quad {}^k\mathscr{Z} \cap {}^i\mathscr{Z} = \emptyset \quad \text{for} \quad k \neq i\,; \quad k, i \in N_R\,. $$

We shall prove the theorem.

Theorem 2.1. Let $\pi_n(\mathscr{X}_n)$ be defined for all $n \in N$. Then for every $n \in N$ there exists such an expansion $\{{}^j\prod_n\}_{j=0}^H \triangleq \{{}^0\prod_n, {}^1\prod_n, \ldots, {}^H\prod_n\}$ of the set $\prod$ that for ${}^0\prod_1 \neq \emptyset$ the Bayes optimum sequential test (d^*, τ^*) of the set of hypotheses $\mathscr{H}$ is given by relations:

$$ \tau^* = \inf_{n \in N} \{n : \pi_n(\mathscr{X}_n) \notin {}^0\textstyle\prod_n\} \tag{2.3} $$

$$ d_n^* = \sum_{j=1}^{H} j \cdot {}^j\varphi_n(\pi_n(\mathscr{X}_n)) + {}^0\varphi_n(\pi_n(\mathscr{X}_n)) \tag{2.4} $$

where ${}^j\varphi_n$ are indicators of sets ${}^j\prod_n$ for $j = 0, 1, \ldots, H$.

Proof. It follows from the Theorem 4.1 and the Remark 4.2 of [2] that the Bayes optimum stopping rule τ^* is given by the relation

$$ \tau^* = \inf_{n \in N} \{n : r_n(\mathscr{X}_n) = h(\pi_n(\mathscr{X}_n))\}\,. \tag{2.5} $$

It holds for the function $r_n(\mathscr{X}_n)$:

$$ r_n(\mathscr{X}_n) = \min \{h(\pi_n(\mathscr{X}_n));\ \boldsymbol{M}\hat{\ }(r_{n+1}(\mathscr{X}_{n+1}) \mid \mathscr{X}_n) + c_{n+1}\} \tag{2.6} $$

$$ r_n(\mathscr{X}_n) = \lim_{k \to \infty} \tilde{Q}_n^k[h(\pi_n(\mathscr{X}_n))] \quad n \in N \tag{2.7} $$

where $\tilde{Q}_n^k$ is the k-th power of the operator $\tilde{Q}_n$ defined by the relation

$$ \tilde{Q}_n[\psi_n(\mathscr{X}_n)] \triangleq \min \{\psi_n(\mathscr{X}_n);\ \boldsymbol{M}\hat{\ }(\psi_{n+1}(\mathscr{X}_{n+1}) \mid \mathscr{X}_n) + c_{n+1}\} \quad n \in N \tag{2.8} $$

and $\boldsymbol{M}\hat{\ }(\cdot \mid \mathscr{X}_n)$ is a variant of the conditional mean value $\boldsymbol{M}(\cdot \mid \mathscr{F}_n)$, defined for all $\mathscr{X}_n$ on the space $(\Omega, \mathscr{F}, \boldsymbol{P})$, which is denoted in [2] by $\boldsymbol{M}_{n,\mathscr{X}_n}(\cdot)$.

It is clear from relations (2.2), (2.7) and (2.8) that the function $r_n(\mathscr{X}_n)$ depends on $\mathscr{X}_n$ only through $\pi_n(\mathscr{X}_n)$. We can then define a new real function g_n on $\prod$ by the relation

$$(2.9)\qquad g_n(\pi_n(\mathscr{X}_n)) \triangleq r_n(\mathscr{X}_n) \quad n \in N .$$

It is easy to show from relations (2.5) till (2.9) that it holds for the function g_n

$$(2.10)\qquad g_n(\mathbf{t}) = \min \{h(\mathbf{t});\ G_{n+1}(\mathbf{t}) + c_{n+1}\} \quad \mathbf{t} \in \prod , \quad n \in N$$

$$(2.11)\qquad G_n(\mathbf{t}) = \int_E g_n(\mathbf{t}_n(\mathbf{x})) \sum_{s=1}^{H} {}^s t \,.\, {}^s f_n(\mathbf{x})\, \mathrm{d}\mathbf{x} , \quad \mathbf{t} = ({}^1 t, \ldots, {}^H t) \in \prod$$

for

$$\mathbf{t}_n(\mathbf{x}) \triangleq \left(\frac{{}^1 t \,.\, {}^1 f_n(\mathbf{x})}{\sum_{s=1}^{H} {}^s t \,.\, {}^s f_n(\mathbf{x})} , \ldots , \frac{{}^H t \,.\, {}^H f_n(\mathbf{x})}{\sum_{s=1}^{H} {}^s t \,.\, {}^s f_n(\mathbf{x})} \right) \quad n \in N$$

$$(2.12)\qquad g_n(\mathbf{t}) = \lim_{k \to \infty} \hat{Q}_n^k[h(\mathbf{t})] \quad \mathbf{t} \in \prod , \quad n \in N$$

$$(2.13)\qquad \hat{Q}_n[\psi_n(\mathbf{t})] \triangleq \min \left\{ \psi_n(\mathbf{t});\ \int_E \psi_{n+1}(\mathbf{t}_{n+1}(\mathbf{x})) \,.\, \sum_{s=1}^{H} {}^s t \,.\, {}^s f_{n+1}(\mathbf{x})\, \mathrm{d}\mathbf{x} + c_{n+1} \right\} \quad n \in N .$$

Substituting (2.9) into (2.5) we obtain

$$(2.14)\qquad \tau^* = \inf_{n \in N} \{ n : g_n(\pi_n(\mathscr{X}_n)) = h(\pi_n(\mathscr{X}_n)) \} .$$

Let us define ${}^0\prod_n \subset \prod$ by the relation

$$(2.15)\qquad {}^0\prod_n \triangleq \{ \mathbf{t} \in \prod : g_n(\mathbf{t}) < h(\mathbf{t}) \} .$$

Relations (2.10), (2.14) and (2.15) then prove the relation (2.3).

Let us further define ${}^j\prod_n \subset \prod$ by the relation

$$(2.16)\qquad {}^j\prod_n \triangleq \{ \mathbf{t} \in \prod : \sum_{i=1}^{H} L_{ij} \,.\, {}^i t = h(\mathbf{t}) , \quad \sum_{i=1}^{H} L_{ik} \,.\, {}^i t > h(\mathbf{t}) , \quad k = 1, \ldots, j-1 \} \cap (\prod \setminus {}^0\prod_n) \quad j \in \mathscr{A} , \quad n \in N .$$

Relations (1.9) and (2.16) prove the relation (2.4). Let us note that, in accordance with the Definition 1.1, the value of d_n^* in (2.4) can be defined quietly arbitrarily for $\pi_n(\mathscr{X}_n) \in {}^0\prod_n$. But according to the definition of the rule of terminal decision it must hold $d_n^* \in \mathscr{A}$ also in this case. The condition ${}^0\prod_1 \neq \emptyset$ is equivalent to the condition (1.11) of our Theorem 1.1 (see relation (4.15) of [2]). By this the proof of theTheorem 2.1 is completed.

Remark 2.1. As it follows from the proof of the Theorem 2.1, some elements of an expansion $\{{}^j\prod_n\}$ can be empty sets.

Theorem 2.2. Let $\mathscr{P}$ be a σ-algebra of all Borel subsets of the set $\prod$. ${}^j\prod_n$ are convex sets and it holds ${}^j\prod_n \in \mathscr{P}$, ${}^0\prod_n \in \mathscr{P}$ for every $n \in N$ and all $j \in \mathscr{A}$.

Proof. First we shall prove that ${}^j\prod_n$ are convex sets for $j \in \mathscr{A}$. The proof is trivial for ${}^j\prod_n = \emptyset$. Let us thus assume that ${}^j\prod_n \neq \emptyset$. Let us define a function h_j on $\prod$ for $j \in \mathscr{A}$ by the relation

$$h_j(\mathbf{t}) \triangleq \sum_{i=1}^{H} L_{ij} \,.\, {}^i t \,, \quad \mathbf{t} = ({}^1 t, \ldots, {}^H t) \in \prod . \tag{2.17}$$

According to relations (2.10), (2.15) and (2.16) it holds for ${}^j\prod_n$, $j \in \mathscr{A}$, $n \in N$

$$\begin{aligned} {}^j\prod_n = \{\mathbf{t} : G_{n+1}(\mathbf{t}) + c_{n+1} \geqq h(\mathbf{t}),\ h_j(\mathbf{t}) = h(\mathbf{t}), \\ h_k(\mathbf{t}) > h(\mathbf{t}) \quad \text{for} \quad k = 1, \ldots, j-1\} . \end{aligned} \tag{2.18}$$

Let $\mathbf{t}_1 \triangleq ({}^1 t_1, \ldots, {}^H t_1)$, $\mathbf{t}_2 \triangleq ({}^1 t_2, \ldots, {}^H t_2)$ be two arbitrary elements of the set ${}^j\prod_n$. According to [4] we must prove that it holds for every $\lambda \in (0, 1)$

$$\lambda \mathbf{t}_1 + (1 - \lambda)\, \mathbf{t}_2 \in {}^j\prod_n . \tag{2.19}$$

It holds for $\mathbf{t}_{1,2} \in {}^j\prod_n$ and arbitrary $\lambda \in (0, 1)$

$$\begin{aligned} h(\lambda \mathbf{t}_1 + (1 - \lambda)\, \mathbf{t}_2) = \min_{j \in \mathscr{A}} \{ \sum_{i=1}^{H} L_{ij}(\lambda \,.\, {}^i t_1 + (1 - \lambda) \,.\, {}^i t_2)\} = \\ = \lambda\, h(\mathbf{t}_1) + (1 - \lambda)\, h(\mathbf{t}_2) . \end{aligned} \tag{2.20}$$

Further it holds for every $j \in \mathscr{A}$

$$h_j(\lambda \mathbf{t}_1 + (1 - \lambda)\, \mathbf{t}_2) = \lambda\, h_j(\mathbf{t}_1) + (1 - \lambda)\, h_j(\mathbf{t}_2) . \tag{2.21}$$

Since the function $h(\mathbf{t})$ is concave, functions $g_n(\mathbf{t})$ and $G_n(\mathbf{t})$ are also concave for every $n \in N$ according to [4] and relations (2.11), (2.12). Then it holds for every $\lambda \in (0, 1)$

$$G_{n+1}(\lambda \mathbf{t}_1 + (1 - \lambda)\, \mathbf{t}_2) \geqq \lambda\, G_{n+1}(\mathbf{t}_1) + (1 - \lambda)\, G_{n+1}(\mathbf{t}_2) . \tag{2.22}$$

The verified relation (2.19) directly follows from (2.18) together with (2.20) till (2.22)

Now let us show that it holds for $j \in \{0, 1, \ldots, H\}$

$$ {}^j\prod_n \in \mathscr{P} \quad \text{for} \quad n \in N . \tag{2.23}$$

It is clear that functions $h(\mathbf{t})$ and $h_j(\mathbf{t})$ are $\mathscr{P}$-measurable functions. It is easy to show that $g_n(\mathbf{t})$ and $G_n(\mathbf{t})$ are also $\mathscr{P}$-measurable. Relation (2.23) then follows directly from relations (2.15) and (2.18). Theorem 2.2 is proved.

Remark 2.2. Let us denote by $\{{}^j\prod_0\}_{j=0}^H$ an expansion of the set $\prod$ given by relations

$$
\begin{aligned}
{}^0\textstyle\prod_0 &\triangleq \emptyset\,, \\
{}^j\textstyle\prod_0 &\triangleq \{\mathbf{t} : h(\mathbf{t}) = h_j(\mathbf{t}),\ h_k(\mathbf{t}) > h(\mathbf{t}) \text{ for } k = 1, \ldots, j-1\}\,.
\end{aligned}
$$

Then ${}^j\prod_0$ are convex sets for $j \in \mathscr{A}$ and it holds for all $n \in N$

$$
{}^j\textstyle\prod_n \subset {}^j\prod_0 \in \mathscr{P} \quad j \in \mathscr{A}\,.
$$

Remark 2.3. Let some sequence of expansions $\{\{{}^j\prod_n\}_{j=0}^H\}_n$ satisfying assertion of Theorem 2.2 for $n \in N$ be given. We shall denote by *a sequential test of a posteriori probability*, defined by the mentioned sequence of expansions, every sequential test (d, τ) for which relations (2.3) and (2.4) hold for this sequence of expansions. It is known that the optimum sequential test of a posteriori probability defined by a sequence of expansions (2 18) is equivalent for $H = 2$ to the sequential likelihood ratio test with some sequence of thresholds which was introduced in [1].

3. OPTIMUM SEQUENTIAL TEST FOR DEPENDENT OBSERVATIONS

In this chapter we shall discuss one special case of statistically dependent observations $\mathbf{x}_n$. In explanation we shall follow chapter 3 of [1], results of which we shall generalize for $H \geqq 2$ of statistical hypotheses and for the cost c_n depending on n.

We shall express the n-tuple $\mathscr{X}_n$ of observations $\mathbf{x}_1, \ldots, \mathbf{x}_n$ by a row vector defined by the relation

$$
\mathbf{X}_n \triangleq (x_{11}, \ldots, x_{1M}, \ldots, x_{n1}, \ldots, x_{nM}) \quad n \in N
$$

where

$$
\mathbf{x}_n = (x_{n1}, \ldots, x_{nM}) \quad n \in N\,.
$$

Let us define a new row vector $\mathbf{X}'_n$

$$
\mathbf{X}'_n \triangleq (x'_{11}, \ldots, x'_{1M}, \ldots, x'_{n1} \ldots, x'_{nM})
$$

by the relation

$$
\mathbf{X}'^T_n = \mathbf{D}_n \mathbf{X}^T_n \quad n \in N \tag{3.1}
$$

where $\mathbf{Y}^T$ denotes a transposed (column) vector to the row vector $\mathbf{Y}$ and $\mathbf{D}_n$ is a $Mn \times Mn$ matrix with real elements, given by the recursive relation

$$(3.2) \qquad \mathbf{D}_{n+1} = \left[\begin{array}{c:c} \mathbf{D}_n & \mathbf{d}_n''' \\ \hdashline \mathbf{d}_n' & \mathbf{d}_n'' \end{array} \right] \quad n \in N .$$

For every $n \in N$ the $\mathbf{d}_n'$, $\mathbf{d}_n''$, $\mathbf{d}_n'''$ are $M \times Mn$, $M \times M$, $Mn \times M$ matrices where $\mathbf{D}_1$ and $\mathbf{d}_n''$ are regular matrices and $\mathbf{d}_n'''$ is a zero matrix. It follows from these assumptions that the matrix $\mathbf{D}_n$ is regular for every $n \in N$ and there exists its inverse $\mathbf{C}_n$, $\mathbf{C}_n = \mathbf{D}_n^{-1}$ for $n \in N$.

We shall express vector $\mathbf{X}_n'$ as the n-tuple $\mathscr{X}_n'$ of vectors $\mathbf{x}_i'$, i.e.

$$\begin{aligned} \mathscr{X}_n' &= (\mathbf{x}_1', \ldots, \mathbf{x}_n') \\ \mathbf{x}_n' &= (x_{n1}', \ldots, x_{nM}') \end{aligned} \quad n \in N ,$$

According to (3.1) and (3.2), $\mathscr{X}_n'$ is a random element defined on probability spaces $(\Omega, \mathscr{F}, \mathbf{P})$ and $(\Omega, \mathscr{F}, \mathbf{P}_i)$, $i \in \mathscr{H}$. Let us denote the probability density of n-tuple $\mathscr{X}_n'$ on $(\Omega, \mathscr{F}, \mathbf{P}_i)$ by ${}^i w_n'(\mathscr{X}_n')$. This probability density always exists due to the regularity of the matrix $\mathbf{D}_n$ and it holds

$$(3.3) \qquad {}^i w_n'(\mathscr{X}_n') = J_n \,.\, {}^i w_n(\mathscr{X}_n) \quad i \in \mathscr{H} , \quad n \in N$$

where J_n is the absolute value of Jacobian of the linear regular transform (3.1), i.e.

$$(3.4) \qquad J_n = |\det \mathbf{C}_n| \neq 0 \quad n \in N ,$$

We shall prove the following theorem:

Theorem 3.1. Let for every $n \in N$ there exists a matrix $\mathbf{D}_n$, satisfying (3.2) and such that it holds for every $i \in \mathscr{H}$ and for every $n \in N$

$$(3.5) \qquad {}^i w_n'(\mathscr{X}_n') = \prod_{k=1}^{n} {}^i f_k'(\mathbf{x}_k')$$

where ${}^i f_n'(\mathbf{x}') \geqq 0$ for $\mathbf{x}' \in \mathbf{E}$, $i \in \mathscr{H}$, $n \in N$ and $\int_{\mathbf{E}} {}^i f_n'(\mathbf{x}')\, \mathrm{d}x' = 1$ for $i \in \mathscr{H}$, $n \in N$. Let $\pi_n(\mathscr{X}_n)$ be defined for all $n \in N$. Then for every $n \in N$ there exists an expansion $\{{}^j\prod_n\}_{j=0}^{H} \triangleq \{{}^0\prod_n, {}^1\prod_n, \ldots, {}^H\prod_n\}$ of the set $\prod$ with the following properties

a) ${}^j\prod_n \in \mathscr{P}$ for $j \in \{0, 1, \ldots, H\}$, $n \in N$;

b) ${}^j\prod_n$ are convex sets for every $j \in \mathscr{A}$ and $n \in N$;

c) The following relations hold for the Bayes optimum sequential test (d^*, τ^*) of the set of hypotheses $\mathscr{H}$ for ${}^0\prod_1 \neq \emptyset$

$$(3.6) \qquad \tau^* = \inf_{n \in N} \{n : \pi_n(\mathscr{X}_n) \notin {}^0\textstyle\prod_n\}$$

$$d_n^* = \sum_{j=1}^{H} j \,.\, {}^j\varphi_n(\pi_n(\mathscr{X}_n)) + {}^0\varphi_n(\pi_n(\mathscr{X}_n)) \tag{3.7}$$

where ${}^j\varphi_n$ are indicators of sets ${}^j\prod_n$.

Proof. Since Bayes optimum sequential test (d^*, τ^*) exists, according to Theorem 1.1, the proof of the statement c) of our theorem is an easy generalization of a proof of Theorem 3 in [1] for the case $H \geqq 2$ using Theorem 2.1 of the proceeding chapter. Relations (2.15) and (2.16) hold for elements ${}^j\prod_n$ of an expansion $\{{}^j\prod_n\}_{j=0}^H$. Statements a), b) then follow from Theorem 2.2.

Remark 3.1. It follows from Theorem 3.1 that the Bayes optimum sequential test (d^*, τ^*) is the sequential test of a posteriori probability for here assumed type of dependence of observations $\mathbf{x}_n$.

4. OPTIMUM SEQUENTIAL TEST FOR DISTINGUISHING OF KNOWN SIGNALS IN A COLOURED GAUSSIAN NOISE

Theorem 3.1 gives us a possibility to solve a problem which of H possible known signals is present at the output of the transmission channel with a coloured Gaussian noise. We shall deal with this problem which is important from the point of view of practical applications.

Let it hold for the n-th observation

$$\mathbf{x}_n \triangleq \mathbf{n}_n + {}^i\mathbf{s}_n \quad i \in \mathscr{H}, \quad n \in N \tag{4.1}$$

where ${}^i\mathbf{s}_n \triangleq ({}^is_{n1}, \ldots, {}^is_{nM})$ is a given vector of signal and $\mathbf{n}_n \triangleq (n_{n1}, \ldots, n_{nM})$ is Gaussian random vector. Let the vector

$$\mathbf{X}_n \triangleq \mathbf{N}_n + {}^i\mathbf{S}_n \triangleq (x_{11}, \ldots, x_{1M}, \ldots, x_{n1}, \ldots, x_{nM}) \quad i \in \mathscr{H}, \quad n \in N \tag{4.2}$$

be a Gaussian random vector with a mean ${}^i\mathbf{S}_n$ and with a covariance matrix $\mathbf{R}_n$ and let it hold for every $n \in N$

$$\begin{aligned} \mathbf{N}_n &\triangleq (n_{11}, \ldots, n_{1M}, \ldots, n_{n1}, \ldots, n_{nM}) \\ {}^i\mathbf{S}_n &\triangleq ({}^is_{11}, \ldots, {}^is_{1M}, \ldots, {}^is_{n1}, \ldots, {}^is_{nM}) \quad i \in \mathscr{H} \\ \mathbf{R}_n &\triangleq M_i((\mathbf{X}_n - {}^i\mathbf{S}_n)^{\mathrm{T}} \,.\, (\mathbf{X}_n - {}^i\mathbf{S}_n)) = \\ &= M_j((\mathbf{X}_n - {}^j\mathbf{S}_n)^{\mathrm{T}} \,.\, (\mathbf{X}_n - {}^j\mathbf{S}_n)) \quad i, j \in \mathscr{H} \end{aligned} \tag{4.3}$$

where M_i is the expected value on the probability space $(\Omega, \mathscr{F}, \mathbf{P}_i)$, $i \in \mathscr{H}$. We shall assume that symetric $Mn \times Mn$ matrix $\mathbf{R}_n$ is positive definite.

Relations (4.1) till (4.3) define a transmission channel, to the input of which one from H possible known signals $\{{}^i\mathbf{s}_n\}$ is led. In the following theorem we shall derive the Bayes optimum sequential test which knowing the output of the channel $\{\mathbf{x}_n\}$ estimates what signal $\{{}^i\mathbf{s}_n\}$ was led to the input.

Theorem 4.1. Let $\mathscr{X}_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, where $\mathbf{x}_n$ is given by the relation (4.1) and let a priori probability distribution on the set of all possible signals $\{\{{}^1\mathbf{s}_n\}_n, \ldots, \{{}^H\mathbf{s}_n\}_n\}$ be $\pi \in \prod$. Then for every $n \in N$ there exists an expansion $\{\prod_n\}_{j=0}^H$ of a set $\prod$ with the following properties:

a) ${}^j\prod_n \in \mathscr{P}$ for $j \in \{0, 1, \ldots, H\}$, $n \in N$;

b) ${}^j\prod_n$ are convex sets for every $j \in \mathscr{A}$ and $n \in N$;

c) For Bayes optimum sequential test (d^*, τ^*) of the set of hypotheses $\{H_i\}_{i=1}^H \triangleq \triangleq \{\{\mathbf{x}_n\}_n = \{(\mathbf{n}_n + {}^1\mathbf{s}_n)\}_n, \ldots, \{\mathbf{x}_n\}_n = \{(\mathbf{n}_n + {}^H\mathbf{s}_n)\}_n\}$ it holds for ${}^0\prod_1 \neq \emptyset$

$$\tau^* = \inf_{n \in N} \{n : \pi_n(\mathscr{X}_n) \notin {}^0\prod_n\} \tag{4.4}$$

$$d_n^* = \sum_{j=1}^H j \cdot {}^j\varphi_n(\pi_n(\mathscr{X}_n)) + {}^0\varphi_n(\pi_n(\mathscr{X}_n)) \tag{4.5}$$

where ${}^j\varphi_n$ are indicators of sets ${}^j\prod_n$.

Proof. Analogically as in discussions of chapter 4 in [1] we can show that for our case there always exists such matrix $\mathbf{D}_n$ with a property (3.2) that $\mathbf{X}_n'$ in (3.1) is a Gaussian vector with uncorrelated Gaussian components x_{kl}' for $k = 1, \ldots, n$ and $l = 1, \ldots, M$. Since uncorrelated Gaussian components are statistically independent, thus condition (3.5) is satisfied. Theorem 4.1 is then a consequence of Theorem 3.1, since $\pi_n(\mathscr{X}_n)$ exists for all $n \in N$ in our Gaussian case.

Remark 4.1. Theorem 4.1 contains, as a special case for $H = 2$ and for ${}^1\mathbf{s}_n \triangleq 0$, ${}^2\mathbf{s}_n \triangleq \mathbf{s}_n$, $n \in N$, the assertion of Theorem 4 of [1].

It is clear from Theorem 4.1 that Bayes optimum sequential test for coloured (statistically dependent) Gaussian observations is a sequential test of a posteriori probability if Gaussian observations differ in their means according to the finite number of possible hypotheses.

5. CONCLUSIONS

Theorems 2.1, 3.1 and 4.1 determine the Bayes optimum sequential test of the finite set of hypotheses as a sequential test of a posteriori probability defined by some sequence of expansions of a set $\prod$ of all possible probability distributions on the set of hypotheses $\mathscr{H}$. In connection with this there arises a very interesting and not yet solved problem of finding sufficient and necessary conditions when the sequential test of a posteriori probability is at the same time the Bayes optimum sequential test.

Further not yet solved problem of a great importance in practical applications of Bayes optimum sequential test according to Theorems 2.1, 3.1 and 4.1 is finding the constructive methods how to determine expansions $\{{}^j\prod_n\}_{j=o}^{H}$.

(Received October 9, 1978.)

REFERENCES

[1] J. Cochlar, I. Vrana: On the Optimum Sequential Test of Two Hypotheses for Statistically Dependent Observations. Kybernetika *14* (1978), 1, 57–69.

[2] J. Cochlar: Optimum Stopping Rules on the Sequence of Statistically Dependent Vectors. Kybernetika *16*, (1980), 1, 13 – 35.

[3] Y. S. Chow, H. Robbins: On Optimal Stopping Rules. Zeitschr. f. Wahrscheinlichkeitstheorie u. verw. Geb. *2* (1963/64), 1, 33–49.

[4] R. T. Rockafellar: Convex Analysis. Princeton University Press, Princeton 1970.

Ing. Jiří Cochlar, CSc., katedra radioelektronických zařízení a soustav elektrotechnické fakulty ČVUT (Czech Technical University — Department of Radioelectronic Devices and Systems), Suchbátarova 2, 166 27 Praha 6. Czechoslovakia.