KYBERNETIKA — VOLUME 14 (1978), NUMBER 1

On the Optimum Sequential Test of Two Hypotheses for Statistically Dependent Observations

JIŘÍ COCHLAR, IVAN VRANA

The paper deals with a derivation of the Bayes optimum sequential test of two hypotheses. A structure of the optimum sequential test for the case of independent but generally differently distributed observations is derived by the method of Chow and Robins [2]. This result is used for derivation of the Bayes optimum sequential test of two hypotheses for statistically dependent observations. Obtained results are applied to the case of detection of the known signal in Gaussian coloured noise.

1. INTRODUCTION

Here we shall formulate a problem of Bayes optimum sequential test of two single opposite hypotheses H_0, H_1. This problem was already solved in many works, e.g. in [1], [2], for statistically independent and identically distributed observations which are scalar random variables. The following definition of problem will be generalisation of the mentioned well known formulation to the case of vector observations which are generally statistically dependent and have different distributions.

We shall use these notations: $\mathbf{x}_i \triangleq (x_{i1}, \ldots, x_{iM})$ is the i-th vector of observations which has M real components x_{ij}, $j = 1, \ldots, M$, for $i = 1, 2, \ldots$ and for the given fixed $M > 0$. By the probability density function ${}^i\mathrm{w}(\mathbf{x}_i)$ of a random vector $\mathbf{x}_i$ we shall understand the probability density function ${}^i\tilde{\mathrm{w}}_M$ of its components, i.e.

$$ {}^i\mathrm{w}(\mathbf{x}_i) \triangleq {}^i\tilde{\mathrm{w}}_M(x_{i1}, \ldots, x_{iM}) \, . $$

Further

$$ \mathscr{X}_n \triangleq \{\mathbf{x}_i\}_1^n = \mathbf{x}_1, \ldots, \mathbf{x}_n $$

is the sequence of the first n vectors of the observation $\mathbf{x}_i$ for $i = 1, 2, \ldots, n$, $n > 0$. By the probability density function $f_n(\mathscr{X}_n)$ of the sequence $\mathscr{X}_n$ we shall understand a joint probability density function $\tilde{f}_{nM}$ of all nM components x_{ij}, i.e.

$$f_n(\mathscr{X}_n) \triangleq \tilde{f}_{nM}(x_{11}, \ldots, x_{1M}, \ldots, x_{n1}, \ldots, x_{nM}) .$$

Hypotheses H_0 and H_1 are defined by a conditional probability density functions. If H_k is the true hypothesis then $\mathscr{X}_n$ has the conditional probability density function $f_n(\mathscr{X}_n \mid k)$ for all $n = 1, 2, \ldots$ and $k = 0, 1$.

Let π be a priori probability that H_0 is true, $0 < \pi < 1$. Then unconditional probability density function $f_n(\mathscr{X}_n)$ of the sequence $\mathscr{X}_n$ is given by the relation

$$f_n(\mathscr{X}_n) = \pi\, f_n(\mathscr{X}_n \mid 0) + (1 - \pi)\, f_n(\mathscr{X}_n \mid 1) , \quad n = 1, 2, \ldots . \tag{1}$$

The loss owing to accepting H_1, when H_0 is true, we shall denote by $a > 0$, and vice versa the loss due to accepting H_0 when H_1 is true, we shall denote by $b > 0$. The cost of making one observation $\mathbf{x}_i$ is unity.

Analogically as in [2] we shall describe the sequential test by (δ, N). In this description N determines a generally random moment of a test termination and δ is a terminal decision. For a given sequential test (δ, N) the global risk $r(\delta, N)$ is described by the relation

$$r(\delta, N) = \pi[a\alpha_0 + E_0(N)] + (1 - \pi)[b\alpha_1 + E_1(N)] ,$$

where α_0 is a probability of accepting H_1 when H_0 is true and α_1 is a probability of accepting H_0 when H_1 is true; $E_k(N)$ is an average length of a test when H_k is true.

The a posteriori probability π_n of validity H_0 for given $\mathscr{X}_n$ is

$$\pi_n = \pi_n(\mathscr{X}_n) = \frac{\pi\, f_n(\mathscr{X}_n \mid 0)}{f_n(\mathscr{X}_n)} , \quad n = 1, 2, \ldots , \tag{2}$$

$$\pi_0 \triangleq \pi . \tag{3}$$

Analogically like in [2] we can easy show that for all moments of test termination N there exists such a rule of terminal decision δ^* that it holds for arbitrary δ

$$r(\delta, N) \geqq r(\delta^*, N) \tag{4}$$

and this rule δ^* is given by relations

$$\delta^* : \begin{cases} \text{accept } H_1 \text{ if } N = n \text{ and } \pi_n(\mathscr{X}_n) \,.\, a \leqq [1 - \pi_n(\mathscr{X}_n)] \,.\, b , \\ \text{accept } H_0 \text{ if } N = n \text{ and } \pi_n(\mathscr{X}_n) \,.\, a > [1 - \pi_n(\mathscr{X}_n)] \,.\, b . \end{cases} \tag{5}$$

As it follows from (4), the problem of finding the Bayes optimum sequential test which minimizes the risk $r(\delta, N)$, is equivalent to determining such a termination moment N^* for which $r(\delta^*, N^*)$ is minimum, if such N^* exists.

According to works [2] and [3] and relation (5) it is valid

$$r(\delta^*, N) = \sum_{n=1}^{\infty} \int_{\{N=n\}} [h(\pi_n(\mathscr{X}_n)) + n] \, f_n(\mathscr{X}_n) \, d\mathscr{X}_n , \tag{6}$$

where

$h(t) \triangleq \min [at, b(1 - t)] ; \quad 0 \leqq t \leqq 1 ,$

$d\mathscr{X}_n \triangleq dx_{11} \ldots dx_{1M} \ldots dx_{n1} \ldots dx_{nM} ,$

$\{N = n\}$ is the set of all possible sequences of observations $\mathscr{X}_n$ for which it holds $N = n$.

Using the above discussion we shall define the concept of stopping rule and the concept of Bayes optimum stopping rule.

Definition 1. Let a sequence of $\mathscr{F}$-measurable random vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots$ on the probability space $(\Omega, \mathscr{F}, \mathbf{P})$ with elements ω be given. Each vector $\mathbf{x}_n$ has $M > 0$ real components. Probability density functions of sequences $\mathscr{X}_n = \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ exist and they are given by relation (1) for all $n > 0$. Let $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots$ be a non-decreasing sequence of sub-σ-algebras of $\mathscr{F}$ such that $\mathscr{F}_n$ is the minimum σ-algebra induced by $\mathscr{X}_n$. By the stopping rule (on the sequence $\mathscr{X}_n$) we shall understand each integer random variable t, defined on $(\Omega, \mathscr{F}, \mathbf{P})$ which can have only positive values and for which the event $\{t(\omega) = n\} \in \mathscr{F}_n$ for each $n = 1, 2, \ldots$.

Definition 2. Let a random sequence $\{y_n\}_1^\infty$ be given for elements of which y_n it is valid

$$y_n \triangleq -h(\pi_n(\mathscr{X}_n)) - n , \quad n = 1, 2, \ldots . \tag{7}$$

Let $\mathscr{C}$ be the set of all stopping rules t for which $E(y_t)$ exists. By the Bayes optimum stopping rule on the sequence $\mathscr{X}_n$ we shall understand a stopping rule t^*, for which it holds

$$E(y_{t^*}) = \sup_{t \in \mathscr{C}} E(y_t) . \tag{8}$$

Note 1. Problem of finding the Bayes optimum stopping rule is trivial for the case $a \leqq 1$ or $b \leqq 1$, because in this case it holds

$$h(t) < 1 ,$$

$$y_0 \triangleq -h(\pi) > y_n , \quad \text{for} \quad n = 1, 2, \ldots$$

and thus $E(y_0) > E(y_t)$ for any $t \in \mathscr{C}$. This case corresponds to making decision without any observation and we shall not deal with it further since it is of no interest from the point of view of practical application.

Theorem 1. If $\pi_n(\mathscr{X}_n)$ is defined for all $n = 1, 2, \ldots$, then for $a > 1$ and $b > 1$ the Bayes optimum stopping rule t^* exists.

Proof. The sufficient conditions of existence t^* are given by Theorem 2 in [2]. We shall verify satisfaction of these conditions in our case. Analogically as in [2], let us put

$$y_n \triangleq y'_n - y''_n \triangleq y^0_n - y^{00}_n ,$$

where

$$y'_n \triangleq y^0_n \triangleq -\mathrm{h}(\pi_n(\mathscr{X}_n)) ,$$

$$y''_n \triangleq y^{00}_n \triangleq n .$$

Since both y'_n and y^0_n are functions only of $\mathscr{X}_n$ and further y''_n and y^{00}_n are not random, $\mathscr{F}_n$-measurability of all these components is evident.

Denote $(x)^+ = \max[x, 0]$; $(x)^- = \max[-x, 0]$. Evidently $(y'_n)^+ \equiv 0$, thus

$$\mathrm{E}[\sup_n (y'_n)^+] = 0 < \infty .$$

The sequence y''_n is evidently increasing and it holds

$$\lim_{n\to\infty} y''_n = \infty .$$

Further it is valid that

$$0 \leqq (y^0_n)^- = \mathrm{h}(\pi_n(\mathscr{X}_n)) \leqq \frac{ab}{a+b}$$

and the sequence $\{(y^0_n)^-\}$ is uniformly integrable for all n, see also [4].

We can see that all assumptions of Theorem 2 in [2] are satisfied and thus the Bayes optimum stopping rule t^* exists.

In the next section of this paper we shall generalize Wald's result [1] to the case of independent but generally non-identically distributed observations. In the third section we shall utilize this generalization for derivation of Bayes optimum stopping rule when observations are statistically dependent. In the fourth section we shall apply these general results to the practically important problem of the optimum sequential detection in coloured Gaussian noise.

2. OPTIMUM SEQUENTIAL TEST FOR INDEPENDENT AND GENERALLY DISTRIBUTED OBSERVATIONS

We shall find the Bayes optimum stopping rule t^* for the case when the conditional joint probability density function $\mathrm{f}_n(\mathscr{X}_n \mid k)$ is given by the relation

$$\mathrm{f}_n(\mathscr{X}_n \mid k) = \prod_{i=1}^{n} {}^i\mathrm{w}(\mathbf{x}_i \mid k) , \quad k = 0, 1; \quad n = 1, 2, \ldots . \tag{9}$$

Further we shall denote by $\lambda_n(\mathscr{X}_n)$ the likelihood ratio defined by

(10) $$\lambda_n(\mathscr{X}_n) \triangleq \mathrm{f}_n(\mathscr{X}_n \mid 1)/\mathrm{f}_n(\mathscr{X}_n \mid 0)\,, \quad n = 1, 2, \ldots .$$

Under these conditions we shall prove a theorem.

Theorem 2. Let a, $b > 1$ and let (9) hold. Then the Bayes optimum stopping rule t^* is given by the relation

(11) $$t^* = \min_{n>0} \{n : \lambda_n(\mathscr{X}_n) \notin (A_n, B_n)\}\,,$$

where $\{A_n\}_1^\infty$, $\{B_n\}_1^\infty$ are some sequences of real numbers.

Proof. We shall use a method described in Sec. 6 of [2]. According the Theorem 2′ of [2] it holds for t^*

(12) $$t^* = \min_{n>0} \{n : y_n = \beta_n \triangleq \lim_{N\to\infty} \beta_n^N\}$$

if $\lim_{N\to\infty} \beta_n^N$ exists. Here β_n^N is given by

(13) $$\beta_n^N \triangleq \max\left[y_n, \mathrm{E}(\beta_{n+1}^N \mid \mathscr{F}_n)\right], \beta_{N+1}^N = -\infty\,, \quad n = 1, 2, \ldots, N\,.$$

As it is valid

$$\pi_{n+1} = \frac{\pi_n \,.\, {}^{n+1}\mathrm{w}(\mathbf{x}_{n+1} \mid 0)}{\pi_n \,.\, {}^{n+1}\mathrm{w}(\mathbf{x}_{n+1} \mid 0) + (1 - \pi_n) \,.\, {}^{n+1}\mathrm{w}(\mathbf{x}_{n+1} \mid 1)}\,,$$

the value of the random variable β_n^N is determined only by the value π_n. Then we can define a new auxiliary random variable

$$\mathrm{g}_n^N(\pi_n) = -\beta_n^N(\mathscr{X}_n) - n\,, \quad n = 1, 2, \ldots, N$$
$$\mathrm{g}_{N+1}^N(\pi_{N+1}) = \infty\,.$$

According to (13)

(14) $$\mathrm{g}_n^N(t) = \min\left[\mathrm{h}(t)\,, \quad \mathrm{G}_{n+1}^N(t) + 1\right], \quad 0 \leqq t \leqq 1\,,$$

where

$$\mathrm{G}_n^N(t) \triangleq \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \mathrm{g}_n^N \left(\frac{t \,.\, {}^n\mathrm{w}(\mathbf{x}_n \mid 0)}{t \,.\, {}^n\mathrm{w}(\mathbf{x}_n \mid 0) + (1 - t) \,.\, {}^n\mathrm{w}(\mathbf{x}_n \mid 1)}\right) .$$
$$.\, \left[t \,.\, {}^n\mathrm{w}(\mathbf{x}_n \mid 0) + (1 - t)\, {}^n\mathrm{w}(\mathbf{x}_n \mid 1)\right] \mathrm{d}\mathbf{x}_n\,,$$
$$\mathrm{d}\mathbf{x}_n \triangleq \mathrm{d}x_{n1} \ldots \mathrm{d}x_{nM}\,.$$

We can easy show by induction backward that it holds

(15) $$g_n^N(t) \geqq g_n^{N+1}(t) \quad \text{for} \quad n = 1, 2, \ldots, N + 1 .$$

From (15) and from the fact that $g_n^N(t) \geqq 0$ the existence of the limit

(16) $$g_n(t) \triangleq \lim_{N \to \infty} g_n^N(t) , \quad n = 1, 2, \ldots,$$

follows. By the Lebesgue theorem of the dominated convergence

$$g_n(t) = \min \left[h(t) , \quad G_{n+1}(t) + 1 \right] ,$$

where

$$G_n(t) \triangleq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_n \left(\frac{t \,.\, {}^n w(\mathbf{x}_n \mid 0)}{t \,.\, {}^n w(\mathbf{x}_n \mid 0) + (1 - t) \,.\, {}^n w(\mathbf{x}_n \mid 1)} \right) .$$
$$. \left[t \,.\, {}^n w(\mathbf{x}_n \mid 0) + (1 - t) \,.\, {}^n w(\mathbf{x}_n \mid 1) \right] d\mathbf{x}_n .$$

Substituting into (12) we obtain

(17) $$t^* = \min_{n > 0} \left[n : g_n(\pi_n) = h(\pi_n) \right] .$$

Similarly as in [2] it is possible to show that both $g_n(t)$ and $G_n(t)$ are concave functions of t. Further

$$g_n(0) = G_n(0) = g_n(1) = G_n(1) = 0$$

is valid. Modifying the method of [2] we shall discuss condition $g_n(\pi_n) = h(\pi_n)$ from (17). Denote

$$\alpha_{1n}(t) \triangleq at - G_{n+1}(t) - 1 ,$$
$$\alpha_{2n}(t) \triangleq b(1 - t) - G_{n+1}(t) - 1 .$$

Then it holds

$$\alpha_{1n}(0) = \alpha_{2n}(1) = -1 < 0 ,$$
$$\alpha_{1n}(1) = a - 1 > 0 ,$$
$$\alpha_{2n}(0) = b - 1 > 0 .$$

Since $G_n(t)$ is concave, $G_n(0) = G_n(1) = 0$ and at is linear, there exists one and only one number $\pi_n' = \pi_n'(a, b)$ such that

(18) $$\alpha_{1n}(t) \begin{cases} <0 & \text{for} \quad t < \pi_n' , \\ =0 & \text{for} \quad t = \pi_n' , \quad n = 1, 2, \ldots \\ >0 & \text{for} \quad t > \pi_n' , \end{cases}$$

and analogically there exists one and only one number $\pi''_n = \pi''_n(a, b)$ such that

$$(19) \qquad \alpha_{2n}(t) \begin{cases} >0 & \text{for} \quad t < \pi''_n, \\ =0 & \text{for} \quad t = \pi''_n, \quad n = 1, 2, \ldots . \\ <0 & \text{for} \quad t > \pi''_n, \end{cases}$$

It follows from (18) and (19) that (17) could be written in the form

$$(20) \qquad t^* = \min_{n>0} \left[n : \pi_n \notin (\pi'_n, \pi''_n) \right].$$

Using likelihood ratio $\lambda_n(\mathscr{X}_n)$ it is possible to rewrite (20) in the form

$$t^* = \min_{n>0} \left[n : \lambda_n(\mathscr{X}_n) \notin (A_n, B_n) \right],$$

where, for $n = 1, 2, \ldots, A_n$ and B_n are given by relations

$$A_n = \frac{\pi}{1 - \pi} \cdot \frac{1 - \pi''_n}{\pi''_n},$$

$$B_n = \frac{\pi}{1 - \pi} \cdot \frac{1 - \pi'_n}{\pi'_n}.$$

Theorem 2 is then proved.

Note 2. Our result for generally distributed independent observations differs from the case when they are identically distributed [1] by the fact that instead of having one pair of tresholds π', π'' we have two sequences of tresholds $\{\pi'_n\}$, $\{\pi''_n\}$ or two sequences $\{A_n\}$, $\{B_n\}$ respectively. This is caused due to the fact that the function $g_n(t)$ in (16) generally depends on n.

3. OPTIMUM SEQUENTIAL TEST FOR DEPENDENT OBSERVATIONS

In this section we shall discuss the case when the conditional joint probability density functions $f_n(\mathscr{X}_n \mid k)$, $k = 0, 1$ can not be expressed as a product (9).

We shall express the sequence of observations $\mathscr{X}_n$ by a row-vector $\mathbf{X}_n$ with components x_{ij}, i.e.

$$(21) \qquad \mathbf{X}_n \triangleq (x_{11}, \ldots, x_{1M}, \ldots, x_{n1}, \ldots, x_{nM}).$$

Let for all $n = 1, 2, \ldots$ the following relation on the probability space $(\Omega, \mathscr{F}, \mathbf{P})$ hold:

$$(22) \qquad \mathbf{X}'^T_n = \mathbf{D}_n \mathbf{X}^T_n,$$

 where

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

$\mathbf{D}_n$ is the $Mn \times Mn$ matrix with real components for which it holds

$$\mathbf{D}_{n+1} = \left[\begin{array}{c|c} \mathbf{D}_n & \mathbf{d}'''_n \\ \hline \mathbf{d}'_n & \mathbf{d}''_n \end{array}\right], \tag{23}$$

where (for all $n = 1, 2, \ldots$) $\mathbf{d}'_n$, $\mathbf{d}''_n$, $\mathbf{d}'''_n$ are $M \times Mn$, $M \times M$, $Mn \times M$ matrices and matrices $\mathbf{D}_1$ and $\mathbf{d}''_n$ are regular matrices and $\mathbf{d}'''_n$ is the zero matrix.

We can express $\mathbf{X}'_n$ in (22) as a sequence $\mathscr{X}'_n$ of vectors $\mathbf{x}'_i$, i.e.

$$\mathscr{X}'_n \triangleq (\mathbf{x}'_1, \ldots, \mathbf{x}'_n), \quad n = 1, 2, \ldots,$$
$$\mathbf{x}'_i \triangleq (x'_{i1}, \ldots, x'_{iM}), \quad i = 1, \ldots, n .$$

This sequence has the probability density function $f'_n(\mathscr{X}'_n)$. The linear transform (22) must satisfy a condition that we can write the conditional probability density function in the form (9), i.e.

$$f'_n(\mathscr{X}'_n \mid k) = \prod_{i=1}^{n} {}^i w'(\mathbf{x}'_i \mid k), \quad k = 0, 1; \quad n = 1, 2, \ldots . \tag{24}$$

It follows from (22) and (23) that the sequence $\mathscr{X}'_n$ is $\mathscr{F}_n$-measurable for each n and thus we can interpret it as a sequence of observations $\mathbf{x}'_i$ $(i = 1, 2, \ldots, n)$. According to (24) this sequence further is created by independent and generally differently distributed observations.

It further follows from (23) that matrix $\mathbf{D}_n$ is regular for each n and thus there exists its inverse $\mathbf{C}_n$, i.e.

$$\mathbf{C}_n = \mathbf{D}_n^{-1} . \tag{25}$$

For the above described case we shall prove the following theorem.

Theorem 3. Let the probability space $(\Omega, \mathscr{F}, \mathbf{P})$ and on it defined sequence of observations $\mathscr{X}_n$ satisfy conditions (22), (23) and (24). Then for a, $b > 1$ the Bayes optimum stopping rule t^* is given by the equation

$$t^* = \min_{n>0} [n : \lambda_n(\mathscr{X}_n) \notin (A_n, B_n)] , \tag{26}$$

where

$$\lambda_n(\mathscr{X}_n) \triangleq f_n(\mathscr{X}_n \mid 1)/f_n(\mathscr{X}_n \mid 0) \quad n = 1, 2, \ldots$$

is the likelihood ratio and $\{A_n\}_1^\infty$, $\{B_n\}_1^\infty$ are some sequences of real numbers.

Proof. On the sequence $\{\mathbf{x}'_n\}_1^\infty$ we shall define a sequence of random variables $\{y'_n\}_1^\infty$ by the relation

$$y'_n(\mathscr{X}'_n) \triangleq -\mathrm{h}(\pi'_n(\mathscr{X}'_n)) - n\,, \quad n = 1, 2, \ldots, \tag{27}$$

where

$$\pi'_n(\mathscr{X}'_n) = \frac{\pi\, \mathrm{f}'_n(\mathscr{X}'_n \mid 0)}{\mathrm{f}'_n(\mathscr{X}'_n)}\,, \quad n = 1, 2, \ldots\,. \tag{28}$$

According to [5] and relations (22), (25) it holds

$$\mathrm{f}'_n(\mathscr{X}'_n \mid k) = J_n\, \mathrm{f}_n(\mathscr{X}_n \mid k)\,, \quad k = 0, 1\,; \quad n = 1, 2, \ldots, \tag{29}$$

where J_n is the absolute value of Jacobian of the linear regular transform (22), i.e.

$$J_n = |\det \mathbf{C}_n| \neq 0\,, \quad n = 1, 2, \ldots\,.$$

From (29) it follows

$$\mathrm{f}'_n(\mathscr{X}'_n) = J_n\, \mathrm{f}_n(\mathscr{X}_n)\,, \quad n = 1, 2, \ldots\,. \tag{30}$$

Substituting (29) and (30) into (28) we obtain

$$\pi'_n(\mathscr{X}'_n) = \pi_n(\mathscr{X}_n)\,, \quad n = 1, 2, \ldots \tag{31}$$

and thus also

$$y'_n(\mathscr{X}'_n) = y_n(\mathscr{X}_n)\,, \quad n = 1, 2, \ldots\,. \tag{32}$$

According to the Theorem 2 of this work, there exists the Bayes optimum stopping rule τ^* on the sequence $\mathscr{X}'_n$ for which it holds

$$\tau^* = \min_{n>0} \left[n : \lambda'_n(\mathscr{X}'_n) \notin (A_n, B_n)\right], \tag{33}$$

where

$$\lambda'_n(\mathscr{X}'_n) \triangleq f'_n(\mathscr{X}'_n \mid 1)/f'_n(\mathscr{X}'_n \mid 0)\,, \quad n = 1, 2, \ldots$$

and $\{A_n\}_1^\infty$, $\{B_n\}_1^\infty$ are sequences introduced in Theorem 2. Owing to (22) and (23), the stopping rule τ^* on the sequence $\mathscr{X}'_n$ can be interpreted as the stopping rule on the sequence $\mathscr{X}_n$ and thus according to Definition 2 and equation (32) it must hold

$$\mathrm{E}(y_{\tau^*}) \leqq \mathrm{E}(y_{t^*}) \tag{34}$$

where t^* is the Bayes optimum stopping rule on the sequence $\mathscr{X}_n$, existence of which is guaranted by Theorem 1.

Due to the fact that (22) is a one-one transform, we can interpret the sequence $\mathscr{X}_n$

 to be derived from the given srquence $\mathscr{X}'_n$. By the similar approach, as in the previous part of the proof, we can show that it must hold

$$\text{(35)} \qquad \mathrm{E}(y_{\tau^*}) \geqq \mathrm{E}(y_{t^*}) .$$

From (34) and (35) then follows

$$\text{(36)} \qquad \mathrm{E}(y_{\tau^*}) = \mathrm{E}(y_{t^*}) .$$

From equations (32), (33) and (36) it follows that the stopping rule t^* on the sequence $\mathscr{X}'_n$ can be given by the equation

$$\text{(37)} \qquad t^* = \min_{n>0} \left[n : \lambda'_n(\mathscr{X}'_n) \notin (A_n, B_n) \right] .$$

From the equation (29) we obtain

$$\lambda_n(\mathscr{X}_n) = \frac{\mathrm{f}_n(\mathscr{X}_n \mid 1)}{\mathrm{f}_n(\mathscr{X}_n \mid 0)} = \lambda'_n(\mathscr{X}'_n) , \quad n = 1, 2, \ldots .$$

Substituting into (37) we shall obtain the relation which we wanted to prove:

$$t^* = \min_{n>0} \left[n : \lambda_n(\mathscr{X}_n) \notin (A_n, B_n) \right] .$$

By this, Theorem 3 is proved.

Note 3. Transformation (22) with properties (23), (24) exists e.g. in the case when the vector $\mathbf{X}_n$ in (21) has Gaussian joint probability density function with the same covariance matrix for bòth hypotheses H_0 and H_1.

Theorem 3 shows that the optimum character of the sequential probability ratio test is also valid for the above described coloured case. It is further clear from the Note 3, that Theorem 3 gives us the possibility to solve the technically important case of sequential detection of signal in coloured Gaussian noise.

4. OPTIMUM SEQUENTIAL DETECTION OF KNOWN SIGNAL IN COLOURED GAUSSIAN NOISE

Let us assume that the observed vector $\mathbf{x}_n$ is given by the equation

$$\text{(38)} \qquad \mathrm{H}_k : \mathbf{x}_n \triangleq \mathbf{n}_n + k\mathbf{s}_n , \quad k = 0, 1 ; \quad n = 1, 2, \ldots ,$$

where H_k denotes the valid hypothesis,

$\mathbf{s}_n \triangleq (s_{n1}, \ldots, s_{nM})$ is a given vector of signal and

$\mathbf{n}_n \triangleq (n_{n1}, \ldots, n_{nM})$ is the Gaussian random vector.

Let the vector

(39) $$\mathbf{X}_n \triangleq \mathbf{N}_n + k \,.\, \mathbf{S}_n \triangleq (x_{11}, \ldots, x_{1M}, \ldots, x_{n1}, \ldots, x_{nM})$$

be the Gaussian random vector with the mean $k \,.\, \mathbf{S}_n$ and with the covariance matrix $\mathbf{R}_n$

$$\mathbf{S}_n \triangleq (s_{11}, \ldots, s_{1M}, \ldots, s_{n1}, \ldots, s_{nM}),$$
$$\mathbf{N}_n \triangleq (n_{11}, \ldots, n_{1M}, \ldots, n_{n1}, \ldots, n_{nM}),$$
$$\mathbf{R}_n \triangleq \mathrm{E}[(\mathbf{X}_n - k \,.\, \mathbf{S}_n)^T \,.\, (\mathbf{X}_n - k \,.\, \mathbf{S}_n)],$$

where $\mathbf{R}_n$ is $Mn \times Mn$ positive-definite symmetric matrix.

Now we shall show that for the above described case, the transformation with properties (22), (23) and (24) always exists. For symmetric positive-definite matrices there always exists the expansion (see theorem V 19.1 in [6])

(40) $$\mathbf{R}_n = \mathbf{C}_n \,.\, \mathbf{C}_n^T, \quad n = 1, 2, \ldots,$$

where $\mathbf{C}_n$ is $Mn \times Mn$ lower triangular regular matrix with real components. From equation for components of matrix $\mathbf{C}_n$ in accordance with Choleski method, we can easy show validity of relation (23) for the matrix $\mathbf{C}_n$. In accordance with (25), let us define matrix $\mathbf{D}_n = \mathbf{C}_n^{-1}$. Then $\mathbf{D}_n$ is also the $Mn \times Mn$ lower triangular regular matrix with real components and it is easy to check validity of equation (23) using expression of elements of an inverse matrix to the triangular matrix (see [6] p. 93).

According to (22) let us now define

(41) $$\mathbf{X}_n'^T = \mathbf{D}_n \mathbf{X}_n^T, \quad n = 1, 2, \ldots.$$

We shall determine the covariance matrix $\mathbf{R}_n'$ of the vector $\mathbf{X}_n'$. From (41) it follows that the mean value of $\mathbf{X}_n'$ is given by

(42) $$\mathrm{E}(\mathbf{X}_n') = k \,.\, \mathbf{S}_n \mathbf{D}_n^T, \quad n = 1, 2, \ldots.$$

Then it holds

(43) $$\mathbf{R}_n' = \mathrm{E}[(\mathbf{X}_n' - k \,.\, \mathbf{S}_n \mathbf{D}_n^T)^T \,.\, (\mathbf{X}_n' - k \,.\, \mathbf{S}_n D_n^T)] = \mathbf{D}_n \mathbf{R}_n \mathbf{D}_n^T, \quad n = 1, 2, \ldots$$

and substituting (40) we then obtain

(44) $$\mathbf{R}_n' = \mathbf{D}_n \mathbf{C}_n \mathbf{C}_n^T \mathbf{D}_n^T = \mathbf{E}_n, \quad n = 1, 2, \ldots,$$

where $\mathbf{E}_n$ is $Mn \times Mn$ unit matrix. From (44) it follows that components of the vector $\mathbf{X}_n'$ are uncorrelated. Vector $\mathbf{X}_n$ is Gaussian and according to (41) vector $\mathbf{X}_n'$ must be also Gaussian. Due to (44), components of the vector $\mathbf{X}_n'$ are statistically independent. By this we have verified validity of equation (24). Thus all assumptions of Theorem 3 are satisfied. The obtained result can be formulated in the following theorem.

Theorem 4. Let hypotheses H_0 and H_1 be defined by equation (38) and let vector $\mathbf{X}_n$ in (39) be Gaussian random vector with the mean $k \, . \, \mathbf{S}_n$, for $k = 0, 1$ and with a real positive-definite covariance matrix $\mathbf{R}_n$. Then the Bayes optimum sequential test of hypothesis H_0 versus H_1 is given by relations

$$(45) \qquad \begin{cases} \textit{accept} \;\; H_0 & \textit{if} \;\; \lambda_n(\mathscr{X}_n) \leqq A_n \,, \\ \textit{accept} \;\; H_1 & \textit{if} \;\; \lambda_n(\mathscr{X}_n) \geqq B_n \,, \\ \textit{make the next observation} & \textit{if} \;\; \lambda_n(\mathscr{X}_n) \in (A_n, B_n) \end{cases}$$

sequentially for $n = 1, 2, \ldots$ *till accepting one of hypotheses.*

$\lambda_n(\mathscr{X}_n)$ is the likelihood ratio and A_n, B_n are thresholds from Theorem 3.

Proof is clear from the Theorem 3 and the rule of terminal decision δ^* defined by relations (5).

Note 4. From equations (42) and (44) it follows that single statistically independent observations $\mathbf{x}'_i$ of the "whitened" vector $\mathbf{X}'_n$ have unit covariance matrices, but generally they will differ by their means even in the case when $\mathbf{s}_i \triangleq \mathbf{s} \triangleq$ **const. vector** for all $i = 1, 2, \ldots$. Thus even in this case, due to the Theorem 2, we shall have two sequences of thresholds instead of a pair of thresholds.

Theorem 4 principially solves the Bayes optimum sequential detection of the known signal in coloured Gaussian noise. The question of determining concrete values of thresholds A_n, B_n remains the open problem.

5. CONCLUSIONS

Theorems 2, 3 and 4 determine a structure of the Bayes optimum sequential test of two single hypotheses for the case when observations need not be statistically independent or to have identical distributions. As to the thresholds A_n, B_n, these theorems unfortunately guarantee only their existence but don't give a constructive description how to determine their values. Besides the application of these theorems to the case when their assumptions are satisfied, one heuristic statement still follows from them: In general case when the sequential likelihood ratio test is the optimum one, we can achieve that this test will have two sequences of thresholds instead of a pair of thresholds. From this point of view, the Wald's result [1], when the optimum sequential likelihood ratio test has only a pair of thresholds, we can assume as a special case of the test having two sequences of thresholds.

(Received July 13, 1977.)

[1] A. Wald, J. Wolfowitz: Optimum character of the sequential probability ratio test. Ann. of Math. Stat. *19* (1948), 3, 326—339.

[2] Y. S. Chow, H. Robbins: On optimal stopping rules. Zeitsch. für Wahrscheinlichkeitstheorie u. verw. Geb. *2* (1963/64), 1, 33—49.

[3] J. Cochlar: Formulace problému sekvenční detekce radiolokačních cílů v korelovaném šumu. Res. rep. III-1-4/4, 1976, ČVUT - FEL Praha.

[4] J. Neveu: Bases mathématiques du calcul des probabilités. Masson et Cie, Paris 1964.

[5] M. H. DeGroot: Optimal statistical decisions. McGraw-Hill, New York 1970.

[6] J. Schmidtmayer: Maticový počet a jeho použití v elektrotechnice. SNTL, Praha 1954.

Ing. Jiří Cochlar, CSc., Ing. Ivan Vrana, 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.