The Annals of Statistics
1975, Vol. 3, No. 4, 939-946

A FINITE MEMORY TEST OF THE IRRATIONALITY OF THE PARAMETER OF A COIN

BY PATRICK HIRSCHLER[1] AND THOMAS M. COVER[2]

Georgia Institute of Technology and Stanford University

Let $X_1, X_2, \cdots$ be a Bernoulli sequence with parameter p. An algorithm $T_{n+1} = f(T_n, X_n, n)$;

$$d_n = d(T_n)\,; \qquad f\colon \{1, 2, \cdots, 8\} \times \{0, 1\} \times \{0, 1, \cdots\} \to \{1, \cdots, 8\}\,;$$
$$d\colon \{1, 2, \cdots, 8\} \to \{H_0, H_1\}\,;$$

is found such that $d(T_n) = H_0$ all but a finite number of times with probability one if p is rational, and $d(T_n) = H_1$ all but a finite number of times with probability one if p is irrational (and not in a given null set of irrationals). Thus, an 8-state memory with a time-varying algorithm makes only a finite number of mistakes with probability one on determining the rationality of the parameter of a coin. Thus, determining the rationality of the Bernoulli parameter p does not depend on infinite memory of the data.

1. Introduction. Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed Bernoulli random variables with unknown mean p. We are interested in determining as much as possible about p with finite methods. Toward this end it has been shown in Cover [1] that there exists a four state finite memory algorithm of the type shown below that tests $H_0: p < p_0$ vs $H_1: p > p_0$ with only a finite number of errors with probability one. Hirschler [4] demonstrates that four states are sufficient to test $H_0: p = p_0$ vs $H_1: p \neq p_0$.

Without the restriction of finite memory, it is well known (see, e.g., Cover [2]) that there exists a test for the hypothesis $H_0: p$ is rational vs. $H_1: p$ is irrational, which makes a decision after each new observation and makes only a finite number of errors with probability one for any $p \in [0, 1] - N_0$, where N_0 is a null set of irrationals. In this paper we show that these results can be combined.

We consider algorithms of the form

$$T_{n+1} = f(T_n, X_n, n) \qquad n = 1, 2, \cdots; \quad T_n \in \{1, 2, 3, \cdots, m\}\,; \quad X_n \in \{0, 1\}\,, \tag{1}$$

with the interpretation that T_n is the state of memory at time n, and m is the number of states in memory. It is appropriate to designate f a time-varying algorithm, as opposed to time-invariant [3], because of its dependence on n. Let

$$d\colon \{1, 2, \cdots, m\} \to \{H_0, H_1\} \tag{2}$$

Received December 1972; revised October 1974.

[1] This work was supported under RADC Contract F30602-72-C-0118 and AFOSR Contract F44620-69-C-0101.

[2] This work was supported under AFOSR Contract F44620-69-C-0101.

AMS 1970 *subject classification*. 62C99.

Key words and phrases. Finite memory, coin, hypothesis testing, rationals.

be a decision rule making decision $d(T_n)$ at time n. We shall describe a deterministic 8-state time-varying algorithm (f, d) that makes only a finite number of mistakes with probability one on the above hypothesis testing problem. Thus 8 states of memory are sufficient for determining the rationality of the bias of a coin.

In other words, the infinite precision necessary to determine the irrationality of p does *not* imply the need for an infinite memory of the past data $X_1, X_2, \cdots$, but requires only the memory of an integer in $\{1, 2, \cdots, 8\}$ and knowledge of the index n of the current observation X_n.

2. Theorem and heuristic proof. We shall prove a generalized version of the aforementioned theorem that extends the test of the rationals to a test of any countable subset of the unit interval. Let (f, d) denote a finite memory decision rule of the form given in (1) and (2). Let S be a countable subset of $(0, 1)$. We shall say that an error is made at time n if the decision $d(T_n) \neq H_{\text{true}}$.

THEOREM. *Let $X_1, X_2, \cdots$ be a sequence of* i.i.d. *Bernoulli* rv's *with* $\Pr\{X_i = 1\} = p$. *Then there exists an 8-state algorithm (f, d) such that only a finite number of errors is made under either hypothesis for the two-hypothesis testing problem*

$$H_0 : p \in S = \{p_1, p_2, \cdots\} \quad \text{vs} \quad H_1 : p \in (0, 1) - S - N_0 , \tag{3}$$

where N_0 is a subset of $[0, 1] - S$ *of Lebesgue measure zero.*

An outline of a possible proof will now be given. A detailed proof involving error bounds and some simplifying (but unnecessary) randomization in f will be given in the next section.

PROOF (Outline). The case where S is a single point has already been proved (using 4 states) in [4] (see also [1]). The idea is to test $p = p_0$ by testing a sequence of n consecutive observations to see if the first np_0 terms are 1 and the last $n(1 - p_0)$ terms are 0. Only one bit of memory Q_1 is needed to test for such a block $B_{0,n}$. Suppose that, given $p = p_0$, the probability of this sequence of 1's and 0's is β_n. For large n, the probability of this event for any $p \notin [p_0 - \delta_n, p_0 + \delta_n]$ is some number $\tilde{\beta}_n \ll \beta_n$. Thus, by repeating this block test $m_n = (\beta_n \tilde{\beta}_n)^{-\frac{1}{2}}$ consecutive times, we have an expected number of successes (i.e., observations of the successful block $B_{0,n}$) given by $(\beta_n/\tilde{\beta}_n)^{\frac{1}{2}} \gg 1$, for $p = p_0$, and $(\tilde{\beta}_n/\beta_n)^{\frac{1}{2}} \ll 1$, for $p \notin [p_0 - \delta_n, p_0 + \delta_n]$. One additional bit of memory Q_2 keeps track of whether at least one success has occurred in the m_n blocks in the nth cycle. Let $Q_2 = 1$ denote at least one success. By Markov's inequality, we see that

$$\begin{aligned} \Pr\{Q_2 = 1\} &\approx 1 , & p &= p_0 \\ \Pr\{Q_2 = 0\} &\approx 0 , & p &\notin [p_0 - \delta_n, p_0 + \delta_n] . \end{aligned} \tag{4}$$

These probabilities can be made arbitrarily extreme for any δ_n by choice of large enough n and m_n. This is the object of Lemma 1.

Let $B(p_0, \delta_n)$ denote the above mentioned block test testing for $p = p_0$ with accuracy δ_n. The idea of the algorithm is to generate the sequence of block tests

$$
\begin{array}{l}
B(p_1, \delta_1) \\
B(p_1, \delta_2)B(p_2, \delta_2) \\
B(p_1, \delta_3)B(p_2, \delta_3)B(p_3, \delta_3) \\
B(p_1, \delta_4) \cdots . \\
\cdots
\end{array}
\tag{5}
$$

with the interpretation that the block test on line 1 is repeated m_1 times, the sequence of block tests on line 2, m_2 times, etc. The m_n consecutive tests of line n will be designated cycle n. At the end of each line, let a third memory variable T take on the value 0 if at least one block success has occurred in the line and 1 otherwise. The variable T denotes the current total decision of H_0 vs H_1, i.e., $d(T, Q_1, Q_2) = H_T$; $T, Q_1, Q_2 \in \{0, 1\}$. This entire procedure requires only 3 bits, i.e., 8 states. The probability of error in the hypothesis test $p \in \{p_1, p_2, \cdots, p_k\}$ vs $p \notin \bigcup_{i=1}^k [p_i - \delta_k, p_i + \delta_k]$ (at the end of the kth cycle) can be made less than any preassigned number $\nu_k > 0$ under either hypothesis.

For H_0, if $\sum \nu_k < \infty$ and $p \in S = \{p_1, p_2, \cdots\}$, then T will equal 0 all but a finite number of times with probability one. This follows, because $p = p_i$ will be tested from the ith line of blocks on, and the number of failures is finite with probability one from the Borel–Cantelli Lemma.

For H_1, by the construction of the test, the probability of the event $T = 0$ at the end of the kth cycle is less than ν_k for any $p \notin \bigcup_{i=1}^k [p_i - \delta_k, p_i + \delta_k] = E_k$. Since $\mu(E_k) \leqq 2k\delta_k$, where μ denotes Lebesgue measure, proper choice of δ_k yields $\sum \mu(E_k) < \infty$. This implies that the Lebesgue measure of $N_0 = \{p : p \in E_k, \text{ i.o.}\}$ is zero. Thus, $\sum \nu_k < \infty$ and $\sum k\delta_k < \infty$ imply $T_n = 0$ all but a finite number of times with probability one for $p \in (0, 1) - S - N_0$.

The more detailed proof in the next section is accomplished in two steps. Lemma 2 first studies the steady-state probability distribution $\boldsymbol{\nu}_n$ on (H_0, H_1) at the "end" of cycle n (i.e., m_n infinite). It is shown that the probability of the state associated with the incorrect hypothesis can be made less than $1/n^2$ by proper choice of δ_n. Finally, the true probability distribution μ_n on (H_0, H_1) can be made very close to $\boldsymbol{\nu}_n$ by proper choice of the duration m_n of cycle n. A possible choice for m_n is exhibited in Lemma 3.

This concludes the outline of the construction of a deterministic algorithm that achieves the goal of Theorem 1.

3. Detailed proof of theorem.

PROOF. For a given enumeration $\{p_i\}$, choose $\delta_n > 0$, $\delta_n \to 0$, such that

$$0 < p_j - 2\delta_n < p_j + 2\delta_n < 1 , \qquad j = 1, \cdots, n . \tag{6}$$

Define

$$p_{j,n} = p_j - \delta_n , \qquad \text{and} \qquad p'_{j,n} = p_j + \delta_n . \tag{7}$$

Thus, $p_{j,n} \nearrow p_j$ and $p'_{j,n} \searrow p_j$. Let $q = 1 - p$ throughout, and define

$$
\begin{aligned}
&a_{j,n} = \log (q_{j,n}/q'_{j,n}) \qquad b_{j,n} = \log (p'_{j,n}/p_{j,n}) \\
(8) \qquad &H_{j,n} = (p_{j,n})^{a_{j,n}}(q_{j,n})^{b_{j,n}} \\
&r_n(p_j, p) = (a_{j,n} \log p + b_{j,n} \log q)/(a_{j,n} \log p_{j,n} + b_{j,n} \log q_{j,n}) .
\end{aligned}
$$

It can be seen that $a_{j,n}$ and $b_{j,n}$ converge to 0 as n tends to infinity. In addition, $r_n(p_j, p)$ satisfies the relations

$$
(9) \qquad r_n(p_j, p_{j,n}) = r_n(p_j, p'_{j,n}) = 1 , \qquad \forall j, \forall n .
$$

Moreover, $r_n(p_j, p)$ is strictly convex function of p with a minimum < 1 achieved in the interval $[p_{j,n}, p'_{j,n}]$. Let $\{m_n\}_{n=1}^{\infty}$ be a sequence of positive integers. Divide the sequence of observations into m_n consecutive superblocks P_n, each of which consists of a sequence of blocks $P_{1,n}, P_{2,n}, \cdots, P_{n,n}$. A successful block consists of $[a_{j,n} t_{j,n}]$ 1's followed by $[b_{j,n} t_{j,n}]$ 0's. (The symbol $[a]$ denotes the least integer greater than or equal to a.)

The proof of the general case, i.e., $S = \{p_1, p_2, \cdots\}$ relies heavily on the proof given here for the point test. See [4] for a different proof. An algorithm involving randomization will be used. The block $B_{j,n}$ has the length of $P_{j,n}$.

Lemma 1. *In the test of m_n consecutive blocks $B_{0,n}$, the probabilities of at least one success can be made arbitrarily near one and zero under hypotheses H_0 and H_1, respectively, for any δ_n, by choosing n and m_n sufficiently large.*

Proof. To achieve this behavior, let $W_1, W_2, \cdots$ be i.i.d. Bernoulli rv's with $\Pr\{W_i = 1\} = \varepsilon_n$. Let the state variable T equal 0 at the end of the kth block if the block is a success. Then, if the result of the experiment W_k is 1, we let T equal 1.

Clearly, for fixed n, the steady-state probability for $(T = 0, T = 1)$ is

$$
(10) \qquad \boldsymbol{\nu}_n = \left(\frac{\beta_n}{\beta_n + \varepsilon_n}, \frac{\varepsilon_n}{\beta_n + \varepsilon_n} \right) .
$$

Let $\lambda_n = \min (r_n(p_0, p_0 - 2\delta_n) - 1, r_n(p_0, p_0 + 2\delta_n) - 1)$ and $\varepsilon_n = (1/n)^{3+6/\lambda_n}$. Under H_0, we have

$$
(11) \qquad \frac{\beta_n}{\varepsilon_n} \geqq p_0 q_0 \frac{1}{\varepsilon_n} \left(\frac{1}{n} \right)^{6r_n(p_0,p_0)/\lambda_n} ,
$$

i.e., $\beta_n/\varepsilon_n \geqq p_0 q_0 n^3$ or $\beta_n/\varepsilon_n > n^2$, for n sufficiently large. Under H_1,

$$
(12) \qquad \frac{\varepsilon_n}{\beta_n} \geqq \exp_n((6r_n(p_0, p) - 1)/\lambda_n - 3) .
$$

But $r_n(p_0, p) > 1 + \lambda_n$ for $p \notin (p_0 - 2\delta_n, p_0 + 2\delta_n)$, for n sufficiently large. Thus, we have

$$
(13) \qquad \nu_n^{1-i} < \frac{1}{n^2} \qquad \text{under} \quad H_i \quad (i = 0, 1) .
$$

Since this fully regular Markov chain approaches its steady-state distribution, it is clear now that in the test of $p = p_0$ vs $p \neq p_0$, the probabilities $\mu_n^i(m_n)$ can be made arbitrarily small (e.g., less than $2/n^2$) under any hypothesis by choosing m_n large enough.

Let the memory consist of the triple (T, Q_1, Q_2) where $T, Q_1, Q_2 \in \{0, 1\}$. Consider the automaton A described by the following algorithm:

$$
\begin{array}{lll}
\text{Start} & n: = 2\,; & \\
\text{Cycle} & n: = n + 1\,; \quad m: = 0\,; & \\
L_1 & m: = m + 1\,; \quad j: = 0\,; \quad Q_2 = 0\,; & \\
L_2 & j: = j + 1\,; \quad Q_1: = 0\,; & \\
 & \text{If} \quad Q_1(B_{j,n}, P_{j,n}) = 1\,, \quad \text{set} \quad Q_2 = 1\,; & \\
 & \text{Otherwise} \quad Q_2 \quad \text{stays unchanged;} & \\
 & \text{If} \quad j < n\,, \quad \text{go to} \quad L_2\,; & \\
 & \text{If} \quad Q_2 = 1\,, \quad \text{set} \quad T = 0\,; & \\
 & \text{Set} \quad T = 1 \quad \text{with probability} \quad \varepsilon_n\,; & \\
 & \text{If} \quad m < m_n\,, \quad \text{go to} \quad L_1\,; & \\
 & \text{Go to Cycle; End.} &
\end{array}
\tag{14}
$$

In other words, the blocks are tested sequentially in the order of appearance. When a block $B_{j,n}$ in B_n is successful, the memory T takes the value 0. At the end of each superblock, if $T = 0$, a random mechanism sets $T = 1$ with conditional probability ε_n. This updating procedure is repeated similarly m_n consecutive times before the new cycle $n + 1$ starts. Within each cycle the process constitutes a fully regular two-state Markov chain with transition probabilities $P_{01} = \varepsilon_n$ and $P_{10} = \alpha_n$. The decision rule chooses H_i if $T = i$ $(i = 0, 1)$. Let d_n be the decision taken at the end of cycle n. Let e_n denote the event that the decision is incorrect. The probability of error at the end of cycle n is $\Pr\{e_n \mid H_i\} = \Pr\{d_n \neq H_i \mid H_i\}$. By the Borel–Cantelli Lemma, if $\sum_{n=1}^{\infty} \Pr\{e_n \mid H_i\}$ is finite under each hypothesis, the above algorithm will make a finite number of errors w.p. 1.

If the blocks $B_{j,n}$ are too long, transitions to state 0 will occur too rarely. On the other hand, if the blocks $B_{j,n}$ are too short, transitions to state 1 will occur too easily. We propose to show that the length of the blocks $B_{j,n}$ can be adjusted in such a way that $\Pr\{e_n \mid H_i\} \leqq 1/n^2$, for $i = 0, 1$.

First, consider the transition probabilities. Let

$$\beta_{j,n} = \Pr\{B_{j,n} \text{ succeeds}\} = p^{[a_{j,n}t_{j,n}]} q^{[b_{j,n}t_{j,n}]}\,. \tag{15}$$

We have

$$\alpha_n = 1 - \textstyle\prod_{j=1}^{n} (1 - \beta_{j,n})\,. \tag{16}$$

From the inequalities $a \leqq [a] < a + 1$, we conclude

$$pq(p^{a_{j,n}} q^{b_{j,n}})^{t_{j,n}} < \beta_{j,n} \leqq (p^{a_{j,n}} q^{b_{j,n}})^{t_{j,n}}\,. \tag{17}$$

Define

(18) $$\lambda_n = \min_{j=1,\cdots,n} \min \{r_n(p_j, p_j - 2\delta_n) - 1, r_n(p_j, p_j + 2\delta_n) - 1\},$$

and choose $t_{j,n}$ such that

(19) $$t_{j,n} = \log (H_{j,n})[n^{-6/\lambda_n}].$$

From (17) we obtain

(20) $$pq\gamma_{j,n} < \beta_{j,n} \leqq \gamma_{j,n},$$

where

(21) $$\gamma_{j,n} = n^{-6r_n(p_j,p)/\lambda_n}.$$

In addition, choose the probability ε_n to be

(22) $$\varepsilon_n = n^{-3-6/\lambda_n}.$$

Next, consider the asymptotic behavior. Let $\boldsymbol{\mu}_n(0) = (\mu_n{}^0(0), \mu_n{}^1(0))$ be the probability vector on the states 0 and 1 at the beginning of cycle n. Let $\boldsymbol{\mu}_n = (\mu_n{}^0(m), \mu_n{}^1(m))$ be that same probability vector after m iterations within cycle n, and $\boldsymbol{\nu}_n = (\nu_n{}^0, \nu_n{}^1)$ be the steady-state probability vector. Then,

(23) $$\boldsymbol{\nu}_n = \left(\frac{\alpha_n}{\alpha_n + \varepsilon_n}, \frac{\varepsilon_n}{\alpha_n + \varepsilon_n}\right),$$

and by a simple computation,

(24) $$\boldsymbol{\mu}_n(m) = \left(\frac{\alpha_n - \Delta_n(m)}{\alpha_n + \varepsilon_n}, \frac{\varepsilon_n + \Delta_n(m)}{\alpha_n + \varepsilon_n}\right),$$

where

(25) $$\Delta_n(m) = (1 - \alpha_n - \varepsilon_n)^m[\alpha_n\mu_n{}^1(0) - \varepsilon_n\mu_n{}^0(0)].$$

We study now the steady-state probability vector for cycle n, and show the following.

LEMMA 2. *Within a given cycle, the steady-state probability of the state associated with the incorrect hypothesis can be made less than $1/n^2$ by proper choice of δ_n.*

PROOF. Under H_0, $p = p_l$ for some fixed l. This implies $r_n(p_l, p_l) < 1$. But

(26) $$\frac{\alpha_n}{\varepsilon_n} = \frac{1}{\varepsilon_n}[1 - \prod_{j=1}^n (1 - \beta_{j,n})] \geqq \frac{1}{\varepsilon_n}[1 - \exp(-\sum_{j=1}^n \beta_{j,n})],$$

and since $\beta_{l,n} \to 0$ as $n \to \infty$, we have

(27) $$\frac{\alpha_n}{\varepsilon_n} \geqq \frac{1}{2}\frac{\beta_{l,n}}{\varepsilon_n} = \frac{1}{2}p_l q_l \frac{\gamma_{l,n}}{\varepsilon_n} > \frac{1}{2}p_l q_l n^3.$$

Hence, $\alpha_n/\varepsilon_n > n^2$, and consequently $\nu_n{}^1 < 1/n^2$ under H_0, for sufficiently large n.

Under H_1, we have

(28) $$\begin{aligned}\frac{\varepsilon_n}{\alpha_n} &= \varepsilon_n[1 - \prod_{j=1}^n (1 - \beta_{j,n})]^{-1} \geqq \varepsilon_n[\sum_{j=1}^n \beta_{j,n}]^{-1} \\ &\geqq \varepsilon_n[\sum_{j=1}^n \gamma_{j,n}]^{-1} = [\sum_{j=1}^n \gamma_{j,n}/\varepsilon_n]^{-1}.\end{aligned}$$

But,

$$\frac{\gamma_{j,n}}{\varepsilon_n} = \left(\frac{1}{n}\right)^{6[r_n(p_j,p)-1-(\frac{1}{2})\lambda_n]/\lambda_n}. \tag{29}$$

Let

$$E_n = \{p \in (0, 1) \mid \min_{j=1,\cdots,n} r_n(p_j, p) \leqq 1 + \lambda_n\}. \tag{30}$$

From the definition of λ_n following (10), the Lebesgue measure of the set E_n is less than $4n\delta_n$. Let $\delta_n = 1/n^3$. Thus,

$$\sum_{n=1}^{\infty} \mu(E_n) \leqq \sum_{n=1}^{\infty} 4n\delta_n < \infty. \tag{31}$$

Therefore, for $p \in E_n^c$, we have $r_n(p_j, p) > 1 + \lambda_n$. This implies, for n sufficiently large,

$$\varepsilon_n/\alpha_n \geqq [n(1/n)^3]^{-1} = n^2, \qquad \text{i.e.,} \quad \nu_n^0 \leqq n^{-2} \quad \text{under} \quad H_1. \tag{32}$$

Finally,

$$\nu_n^{1-i} \leqq 1/n^2 \qquad \text{under} \quad H_i \quad (i = 0, 1). \tag{33}$$

The last step of the proof is to show by proper choice of the duration of cycle n that it is possible to have $\sum_1^\infty \Pr\{e_n \mid H_i\} < \infty$. This then results in a finite number of failures with probability one.

Lemma 3. *There exists a sequence $\{m_n\}_{n=1}^\infty$ such that $\mu_n^{1-i}(m_n) \leqq 2/n^2$ under H_i $(i = 0, 1)$.*

Proof. We shall exhibit a sequence $\{m_n\}$ for which $\mu_n^{1-i}(m_n) \leqq 2\nu_n^{1-i}$ under H_i. Equation (25) can be rewritten in the form

$$\Delta_n(m) = (1 - \alpha_n - \varepsilon_n)^m[(\alpha_n + \varepsilon_n)\mu_n^1(0) - \varepsilon_n]. \tag{34}$$

Since $0 \leqq \mu_n^1(0) \leqq 1$, (34) implies

$$-(1 - \alpha_n - \varepsilon_n)^m\varepsilon_n \leqq \Delta_n(m) \leqq (1 - \alpha_n - \varepsilon_n)^m\alpha_n. \tag{35}$$

Under H_0, since $\alpha_n/\varepsilon_n > n^2$, we have $|\Delta_n(m)| \leqq (1 - n^2\varepsilon_n)^m\alpha_n$. Thus, if

$$m \geqq [\log \varepsilon_n/\log(1 - n^2\varepsilon_n)], \tag{36}$$

then $|\Delta_n(m)| \leqq \varepsilon_n$, for H_0.

Under H_1, $|\Delta_n(m)| \leqq (1 - n^2\alpha_n)^m\varepsilon_n$, by (35). But for any integer $s \in \{1, 2, \cdots, n\}$,

$$\varepsilon_n/\alpha_n = \varepsilon_n[1 - \textstyle\prod_{j=1}^n (1 - \beta_{j,n})]^{-1} \leqq \dfrac{2\varepsilon_n}{\beta_{s,n}} \tag{37}$$
$$\leqq \frac{2}{pq}\left(\frac{1}{n}\right)^{6[1-r_n(p_s,p)]/\lambda_n+3}.$$

Consider integers s and n_0 such that $r_n(p_s, p) \in (3/2, 2)$, $\forall n > n_0$. Then,

$$\varepsilon_n/\alpha_n < (2/pq)n^{6/\lambda_n}, \qquad \text{for} \quad n > \max\{s, n_0\}. \tag{38}$$

If we choose m greater than $[[-(1 + 6/\lambda_n)\log n][\log(1 - n^2\alpha_n)]^{-1}]$, then

$|\Delta_n(m)| \leqq \alpha_n$, for H_1. Let $m_n = (\log \varepsilon_n)(\log(1 - n^2\alpha_n))^{-1}$. Thus, we have shown that

$$\mu_n^{1-i}(m_n) \leqq 2\nu_n^{1-i} \quad \text{under} \quad H_i \quad (i = 0, 1), \tag{39}$$

and the lemma is proved.

REFERENCES

[1] COVER, THOMAS M. (1969). Hypothesis testing with finite statistics. *Ann. Math. Statist.* **40** 828–835.

[2] COVER, THOMAS M. (1973). On determining the irrationality of the mean of a random variable. *Ann. Math. Statist.* **1** 862–871.

[3] HELLMAN, M. E. and COVER, THOMAS M. (1970). Learning with finite memory. *Ann. Math. Statist.* **41** 765–782.

[4] HIRSCHLER, P. (1974). Finite memory algorithms for testing means of Bernoulli random variables. *Information and Control* **24** 11–19.

E.N.S.T. LABORATOIRE D'AUTOMATIQUE
PIECE B 214
46, RUE BARRAULT
75634 PARIS CEDEX 13 FRANCE

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305