The Annals of Mathematical Statistics
1969, Vol. 40, No. 3, 828–835

HYPOTHESIS TESTING WITH FINITE STATISTICS

By Thomas M. Cover[1]

Stanford University

0. Summary. Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed random variables drawn according to a probability measure $\mathcal{P}$. The two-hypothesis testing problem $H_0: \mathcal{P} = \mathcal{P}_0$ vs. $H_1: \mathcal{P} = \mathcal{P}_1$ is investigated under the constraint that the data must be summarized after each observation by an m-valued statistic $T_n \varepsilon \{1, 2, \cdots, m\}$, where T_n is updated according to the rule $T_{n+1} = f_n(T_n, X_{n+1})$. An algorithm with a four-valued statistic is described which achieves a limiting probability of error zero under either hypothesis. It is also demonstrated that a four-valued statistic is sufficient to resolve composite hypothesis testing problems which may be reduced to the form $H_0: p > p_0$ vs. $H_1: p < p_0$ where $X_1, X_2, \cdots$ is a Bernoulli sequence with bias p.

1. Introduction and discussion of sufficient statistics. Let $X_1, X_2, \cdots$ be a sequence of independent identically distributed random variables (iid rv's) drawn according to some unknown measure $\mathcal{P}$. Throughout this paper we shall be interested in the hypothesis test $H_0: \mathcal{P} = \mathcal{P}_0$ vs. $H_1: \mathcal{P} = \mathcal{P}_1$. For a given decision procedure, which assigns each possible observation $(x_1, x_2, \cdots, x_n)$, $n = 0, 1, 2, \cdots$, to H_0 or H_1, we may define $\alpha_n = \Pr\{\text{Decide } H_1 \mid H_0\}$ and $\beta_n = \Pr\{\text{Decide } H_0 \mid H_1\}$. Thus α_n and β_n are the probabilities of error of each kind, based on the first n observations, for the given decision procedure.

It is well known that the standard likelihood-ratio decision procedure results in $\alpha_n \to 0$ and $\beta_n \to 0$ exponentially in n, with rates which depend on an information distance between $\mathcal{P}_0$ and $\mathcal{P}_1$. To apply this procedure at time n requires a memory capacity sufficient to store the observations $x_1, x_2, \cdots, x_n$. Observe that even in the simplest case the memory must grow indefinitely with time. Any truncation of memory to the last k observations, for example—as in the most familiar definition of finite memory (see [9], [10])—will preclude the convergence of α_n and β_n to 0, except in the singular case.

This paper is devoted to the hypothesis-testing problem under the constraint that the data be summarized after each observation by an m-valued statistic $T_n \varepsilon \{1, 2, \cdots, m\}$, where T_n is updated according to an algorithm of the general form $T_{n+1} = f_n(T_n, X_{n+1})$. The two-hypothesis testing problem, under the further constraint that f_n be independent of n, has been solved and is currently being submitted for publication [8]. Before proceeding to an analysis of this algorithm, we wish to discuss alternative formulations of a "reasonable" memory constraint for hypothesis testing problems.

Received 7 February 1968; revised 27 November 1968.

[1] Stanford University, December 1967. This work was supported under AFOSR contract AF49(638)1517.

It has been frequently observed that the data may be reduced by a sufficient statistic without loss of information. For example, $\bar{X}_n = (1/n) \sum_{i=1}^{n} X_i$ is a sufficient statistic for testing the mean of a normal distribution. However, while it is true the mapping corresponding to a sufficient statistic is many-to-one, and is in this sense data reducing, it is generally not true that the cardinality of the required memory is reduced. For example, in the case of the univariate normal, the mapping from $(x_1, x_2, \cdots, x_n) \varepsilon R^n$ to $\bar{x}_n \varepsilon R$ leaves the memory requirement uncountably infinite. A possible alternative information lossless mapping from R^n to R is given by the simple trick of interleaving the digits in the decimal expansions of $x_1, x_2, \cdots$, and x_n to form a single real number. This mapping is not continuous. However, even the nice continuity properties of the usual minimal sufficient statistic may be duplicated. The work of Denny [5] establishes the existence of a 1–1 uniformly continuous map of R^n into R, excluding a set of Lebesgue measure zero. Clearly, then, if one can store one real number, one may store any finite number of real numbers. We conclude that the statistic $\bar{x}_n$ has not decreased the memory requirement at all.

Particular attention ([6], [11], [1]) has been paid to the existence of finite-dimensional sufficient statistics (such as $\bar{x}_n$ for the normal). Here, by implication, it would seem that the memory is bounded in some sense by the dimension of the minimal sufficient statistic. Again, from the standpoint of memory capacity, there is no resultant saving in memory. The previously mentioned interleaving decimal expansion yields a mapping of an arbitrary number of univariate observations into a 1-dimensional sufficient statistic, thus accomplishing the same task as the perhaps nonexistent finite-dimensional sufficient statistic.

A first step toward defining a statistic with a realistic memory constraint might be to consider rounding off the statistic at each stage. Hopefully, the infinite-accuracy theory would apply directly, and α_n, β_n would still tend to 0, although at slower rates. Even this is not the case, as the following simple example will show: Let $X_1, X_2, \cdots$ be iid rv's drawn according to a normal $N(\mu, \sigma^2)$ distribution with unknown mean μ and known variance σ^2. We wish to test $\mu = 1$ vs. $\mu = -1$. Observe that the new statistic $\bar{x}_{n+1}$ may be expressed in terms of the old statistic $\bar{x}_n$ and the current observation x_{n+1} in the form

$$\bar{x}_{n+1} = n\bar{x}_n/(n+1) + x_{n+1}/(n+1). \tag{1.1}$$

Suppose now that $\bar{x}_n$ may be recalled only up to some roundoff error. Let $[\bar{\bar{x}}_n]$ denote the sequentially rounded off version of $\bar{x}_n$. Rounding off at each stage results in the algorithm

$$[\bar{\bar{x}}_{n+1}] = [n[\bar{\bar{x}}_n]/(n+1) + x_{n+1}/(n+1)]. \tag{1.2}$$

To what random variable does $[\bar{\bar{x}}_n]$ converge? The best hope is that $[\bar{\bar{x}}_n] \to [\mu]$ wpl. Thus the decision procedure that decides $\mu = \pm 1$ accordingly as $[\bar{\bar{x}}_n] \gtrless 0$ would result in α_n, $\beta_n \to 0$. Instead, the worst possible situation occurs. As we show elsewhere, $[\bar{\bar{x}}_n]$ converges wpl to a random variable which has strictly positive probability mass on each of the countably infinite number of lattice roundoff values. In particular, there is positive mass on both sides of the origin; consequently, α_n, β_n converge to nonzero limits. So the first realistic approximation to the data reduction problem will not resolve the hypotheses. This is true despite the fact that we have used a reasonable procedure and a countably infinite memory. From this example it may be seen that hypothesis testing problem must be approached from first principles.

To our surprise, we find that this same problem may be solved with a two-state memory. Consider the sequence of statistics $\{T_n\}_1^\infty$, $T_n \in \{-1, 1\}$, defined recursively by

$$\begin{aligned} T_n &= 1, && x_n > (2\sigma^2 \log n)^{\frac{1}{2}}, \\ &= -1, && x_n < -(2\sigma^2 \log n)^{\frac{1}{2}}, \\ &= T_{n-1}, && \text{otherwise} \end{aligned} \tag{1.3}$$

$$T_0 \text{ arbitrary } \in \{-1, 1\}.$$

These thresholds are suggested by the fact (see, for example, [2], [7]) that, for $X_1, X_2, \cdots$ iid $\sim N(\mu, \sigma^2)$, $\max\{X_1, X_2, \cdots, X_n\} - (2\sigma^2 \log n)^{\frac{1}{2}} \to \mu$, in probability. It may be shown that $\sum \Pr\{X_n > (2\sigma^2 \log n)^{\frac{1}{2}}\} = \infty$ or $< \infty$ accordingly as $\mu > 0$ or $\mu < 0$. Consequently, by the Borel zero-one law $X_n > (2\sigma^2 \log n)^{\frac{1}{2}}$ infinitely often wpl for $\mu > 0$ and finitely often wpl for $\mu < 0$. A corresponding statement holds for $X_n < -(2\sigma^2 \log n)^{\frac{1}{2}}$. Therefore $T_n \to 1$ or -1 accordingly as $\mu > 0$ or $\mu < 0$. We have thus furnished an example of a 2-state memory which resolves the composite hypothesis testing problem $\mu > 0$ vs. $\mu < 0$ with probabilities of error α_n, $\beta_n \to 0$. This example will be generalized to arbitrary distributions in Section 2.

To summarize our point of view, we admit the utility of sufficient statistics in computation, but doubt their utility in the problem of memory reduction. First, for multivariate data, there exist trivial data preserving mappings into the unit interval. Denny's work provides uniformly continuous such mappings. Second, the straightforward sequential rounding off of sufficient statistics generally fails to yield α_n, $\beta_n \to 0$. Thus a memory constraint requires a more careful approach than the simple rounding off of sufficient statistics. Third, we might add that the existence of a finite dimensional sufficient statistic is destroyed by a slight distortion of the distributions in the family (without greatly affecting the resolving power of the old statistic). In this sense, the existence of a finite-dimensional sufficient statistic is a "measure-zero" phenomenon, not to be taken too seriously.

In the next sections we shall provide the beginnings of a theory of hypothesis testing with finite statistics in which the hypotheses are resolved despite nontrivial data reduction at each stage. Specifically we shall demonstrate the existence of a four-valued (and in some cases two-valued) sequentially updated statistic achieving $\alpha_n \to 0$ and $\beta_n \to 0$. Sole concern will be with the algorithm

$$T_{n+1} = f_n(T_n, x_{n+1}), \tag{1.4}$$

where the memory (or statistic) T_n takes values in the set $\{1, 2, \cdots, m\}$. Thus the new value of T depends explicitly only on the old value of T, the current observation and the number of trials. Discussion of the naturalness of this formulation will be delayed until Section 4.

2. Learning in the unbounded likelihood ratio case. Let $X_1, X_2, \cdots$ be a sequence of iid rv's drawn according to a probability measure $\mathcal{P}$. Consider the hypothesis test $H_0: \mathcal{P} = \mathcal{P}_0$ vs. $H_1: \mathcal{P} = \mathcal{P}_1$. Let $\mathcal{P}_0$ and $\mathcal{P}_1$ possess probability densities f_0 and f_1 respectively with respect to a measure μ. There is no loss in generality in this assumption since one can take $\mu = \mathcal{P}_0 + \mathcal{P}_1$. Define the likelihood ratio $l(x) = f_1(x)/f_0(x)$. Let us consider f_1 and f_0 such that l is unbounded above and unbounded away from zero, with probability one, under each hypothesis. We shall refer to this as the unbounded likelihood ratio case.

THEOREM 1. *In the unbounded likelihood ratio case, there exist sequences of thresholds* $\{\bar{l}_n\}$, $\{\underline{l}_n\}$ *such that the algorithm*

$$T_n = 1, \qquad l(x_n) > \bar{l}_n, \tag{2.1}$$
$$= -1, \qquad l(x_n) < \underline{l}_n,$$
$$= T_{n-1}, \qquad \textit{otherwise}$$

results in $T_n \to 1$ wpl *under* H_1 *and* $T_n \to -1$ wpl *under* H_0. *Thus* $\alpha_n \to 0$ *and* $\beta_n \to 0$ *with a 2-state memory, under either hypothesis,* i.e., *with probability one, only a finite number of mistakes will be made by* $\{T_n\}$.

PROOF. Let $G_1(l)$ and $G_0(l)$ be the distribution functions of l under hypotheses H_1 and H_0 respectively. Since $l(x) = f_1(x)/f_0(x)$, it follows that $dG_1(l) = l\,dG_0(l)$. Hence

$$G_1(l) = \int_0^l dG_1(l') = \int_0^l l'\,dG_0(l') \leqq lG_0(l). \tag{2.2}$$

We wish to demonstrate a sequence of thresholds $\{\underline{l}_n\}$ such that

$$\sum_{n=1}^{\infty} G_1(\underline{l}_n) < \infty, \tag{2.3}$$
$$\sum_{n=1}^{\infty} G_0(\underline{l}_n) = \infty,$$

for it would then follow by the Borel zero-one law that transitions would be made to the -1 state infinitely often wpl under H_0 but only finitely often under H_1. The following is a simple construction of a suitable sequence $\{\underline{l}_n\}$. Let

$$\{\underline{l}_n\} = \{\underbrace{l_1, l_1, \cdots, l_1}_{\nu_1}, \underbrace{l_2, l_2, \cdots, l_2}_{\nu_2}, l_3, \cdots\} \tag{2.4}$$

where

$$0 < l_n \leqq (\tfrac{1}{4})^{n+1} \tag{2.5}$$

and

$$G_1(l_n) \leqq (\tfrac{1}{2})^n. \tag{2.6}$$

Such a sequence may always be chosen because l is not bounded away from zero, by hypothesis. Let ν_n be the integer satisfying

$$\nu_n G_1(\underline{l}_n) \leqq (\tfrac{1}{2})^n < 2\nu_n G_1(\underline{l}_n). \tag{2.7}$$

Then

$$\sum G_1(\underline{l}_n) = \sum \nu_n G_1(\underline{l}_n) \leqq \sum (\tfrac{1}{2})^n = 1 < \infty. \tag{2.8}$$

On the other hand, from Eqs. (2.2) and (2.7)

$$\begin{aligned} \sum G_0(\underline{l}_n) &= \sum \nu_n G_0(\underline{l}_n) \\ &\geqq \sum \nu_n \underline{l}_n^{-1} G_1(\underline{l}_n) \\ &\geqq \sum (\tfrac{1}{2})^{n+1} \underline{l}_n^{-1} \\ &= \sum 2^{n+1} = \infty. \end{aligned} \tag{2.9}$$

Thus the existence of a sequence $\{\underline{l}_n\}$ satisfying Eq. (2.3) has been established. We note in passing that Eq. (2.8) implies that the expected number of transitions to $T = +1$ is less than 1 under H_0. We may choose this expected number as small as desired by appropriate choices of $\{\underline{l}_n\}$, $\{\bar{l}_n\}$.

A precisely parallel argument establishes the existence of a sequence of upper thresholds $\{\bar{l}_n\}$ such that

$$\sum (1 - G_1(\bar{l}_n)) = \infty \quad \text{and} \quad \sum (1 - G_0(\bar{l}_n)) < \infty. \tag{2.10}$$

Therefore, under H_1, $l(X_n) > \bar{l}_n$ i.o., while $l(X_n) < \underline{l}_n$ only finitely often. Consequently $T_n \to 1$ wpl. Similarly, under H_0, $T_n \to -1$ wpl. Thus $\alpha_n \to 0$, $\beta_n \to 0$, as desired.

3. The bounded likelihood ratio case. There are certain heuristic considerations which make it appear unlikely that learning is possible in the bounded likelihood ratio case. In the Bayesian formulation, for example, where prior probabilities are associated with the two hypotheses, the rule which stores the Bayes decision at each stage will not learn. Eventually the posterior probabilities will be such that no single observation will yield a change in the decision.

Fortunately, experiments of arbitrarily large information may be compounded from experiments of bounded information by the artifice of looking for sequences of events before changing the state of the memory. (This point of view yields some interesting comments [3] on the 2-armed bandit problem with finite memory [9].)

Consider the basic problem of testing the hypothesis that a coin with bias p has bias $p > p_0$ vs. $p < p_0$. Note that the general two-hypothesis testing problem with $\{\tilde{X}_i\}$ iid may be put in this framework under the correspondence

$$\begin{aligned} X_i &= 1, \quad l(\tilde{X}_i) \geqq 1 \\ &= 0, \quad l(\tilde{X}_i) < 1 \end{aligned}$$

and

$$p_0 = \tfrac{1}{2}(\Pr\{X_i = 1 \mid H_1\} + \Pr\{X_i = 1 \mid H_0\}).$$

THEOREM 2. *Let* $X_1, X_2, \cdots$ *be a sequence of* iid *Bernoulli* rv's *with* $\Pr\{X_i = 1\} = p$. *There exists an algorithm with a* 4*-state memory for which the hypothesis* $p > p_0$ vs. $p < p_0$ *is resolved with limiting probability of error zero under either hypothesis.*

PROOF. We shall exhibit one such scheme. Let the memory consist of the pair (T, Q) where T and Q both take values in $\{0, 1\}$. Thus the memory has 4 states (or 2 bits). In the proposed scheme T will keep track of the currently favored hypothesis and Q will keep track of the successfulness of the current run test.

Consider two sequences $\{s_i\}_1^\infty$ and $\{r_i\}_1^\infty$ of positive integers. Divide the sequence of observations into blocks $S_1, R_1, S_2, R_2, \cdots$ with the first s_1 observations being denoted by S_1, the next r_1 by R_1, etc. We shall always know where we are in which block from the knowledge of the sample number n. An S_i block will be considered a success if all s_i observations result in 1's, while an R_i block will be considered a success if all r_i observations result in 0's.

At the beginning of an S_i block, set $Q = 1$ (success) if the initial observation is a 1. Subsequently, in that block, let

$$Q_n = 0, \qquad X_n = 0,$$
$$= Q_{n-1}, \qquad X_n = 1. \tag{3.1}$$

Similarly, for an R_i block, set $Q = 1$ if the initial observation is a 0. Subsequently, in that block, let

$$Q_n = 0, \qquad X_n = 1,$$
$$= Q_{n-1}, \qquad X_n = 0. \tag{3.2}$$

Thus R_i checks for r_i consecutive 1's and S_i checks for s_i consecutive 0's; and $Q_n = 1$ at the end of the block if and only if the desired run has occurred.

The currently favored hypothesis T is updated by the rule

$$\begin{aligned} T_n &= 1, \qquad Q_n = 1, \qquad n \,\varepsilon\, N_1, \\ &= 0, \qquad Q_n = 1, \qquad n \,\varepsilon\, N_2, \\ &= T_{n-1}, \qquad \text{otherwise}, \end{aligned} \tag{3.3}$$

where $N_1 = \{s_1, s_1 + r_1 + s_2, \cdots\}$ and $N_2 = \{s_1 + r_1, s_1 + r_1 + s_2 + r_2, \cdots\}$. Thus changes in T occur only at the ends of test blocks.

Under this rule, the transition probabilities are $\Pr(T_n = 1 \mid T_{n-1} = 0) = p^{s_i}$ at the end of an S_i block and $\Pr(T_n = 0 \mid T_{n-1} = 1) = q^{r_i}$ at the end of an R_i block. It follows, by the independence of the blocks and the zero-one law, that $T_n \to 1$ wpl if

$$\sum p^{s_i} = \infty$$

(3.4) and

$$\sum q^{r_i} < \infty.$$

Similarly, $T_n \to 0$ wpl if

$$\sum p^{s_i} < \infty$$

(3.5) and

$$\sum q^{r_i} = \infty.$$

That is, T_n will be wrong only a finite number of times wpl.

Thus the compound hypothesis test $p > p_0$ versus $p < p_0$ is resolved by this scheme if we can demonstrate sequences of integers $\{s_i\}_1^\infty$, $\{r_i\}_1^\infty$ such that (3.4) holds for $p > p_0$ and (3.5) holds for $p < p_0$. To accomplish this we shall let s_i be the integer defined by

$$\log_{p_0} (1/i) \leqq s_i < \log_{p_0} (1/i) + 1,$$

(3.6) or, for $0 < p < 1$,

$$p^{\log_{p_0}(1/i)} \geqq p^{s_i} > p^{\log_{p_0}(1/i)+1}.$$

Consequently

$$(3.7) \qquad (1/i)^\alpha \geqq p^{s_i} > p(1/i)^\alpha, \qquad \text{where} \quad \alpha = \log_{p_0} p.$$

Now, for $0 < p, p_0 < 1$, $\alpha = \log_{p_0} p$ is $>$ or $\leqq 1$ accordingly as $p < p_0$ or $p \geqq p_0$. Thus $\sum p^{s_i} = \infty$ for $p \geqq p_0$; and $\sum p^{s_i} < \infty$ for $p < p_0$. Similarly, by selecting the integer r_i to satisfy $\log_{q_0} (1/i) \leqq r_i < \log_{q_0} (1/i) + 1$, we may conclude that $\sum q^{r_i} < \infty$, for $p > p_0$; and $\sum q^{r_i} = \infty$, for $p \leqq p_0$. Thus (3.4) and (3.5) are satisfied and the theorem is proved.

This test actually resolves the compound hypothesis testing problem $\mathcal{P} \varepsilon \mathcal{F}_1$ vs. $\mathcal{P} \varepsilon \mathcal{F}_2$ when there exists a set S in the space of outcomes of X such that $\inf_{\mathcal{P}\varepsilon\mathcal{F}_1} \mathcal{P}(S) > \sup_{\mathcal{P}\varepsilon\mathcal{F}_2} \mathcal{P}(S)$. (Define the new random variable $\tilde{X}_i = \begin{Bmatrix} 1, & X_i \,\varepsilon\, S \\ 0, & X_i \not\varepsilon\, S \end{Bmatrix}$ and apply Theorem 2 to $\{\tilde{X}_i\}$.)

Example 1. Let $X_1, X_2, \cdots$ be iid real valued normal random variables with mean zero and unknown variance σ^2. Let $\tilde{X}_i = \begin{cases} 1, & X_i^2 > c^2 \\ 0, & X_i^2 \leqq c^2 \end{cases}$. Note that $\Pr\{\tilde{X}_i = 1 \mid \sigma^2\} = 2\Phi(c/\sigma)$. Let $p_0 = 2\Phi(c/\sigma)$. Then the 4-state test described in this section will test $\sigma^2 > c^2$ versus $\sigma^2 < c^2$, with limiting probability of error zero. In the final analysis, we are testing the hypothesis $\Pr\{X \,\varepsilon\, S\} > p_0$ vs. $\Pr\{X \,\varepsilon\, S\} < p_0$.

Example 2. In the univariate case, let $\tilde{X}_i = \begin{cases} 1, & X_i \geqq c \\ 0, & X_i < c \end{cases}$. The test in this section resolves $F(c) \gtrless p_0$, where F is the unknown cdf of X. If $p_0 = \frac{1}{2}$, this provides a nonparametric finite-memory test of whether or not the median is greater than c.

4. Concluding remarks. Now that it has been shown that the two-hypothesis testing problem may be resolved with a four-state memory, we wish to reinvestigate the naturalness of the algorithm $T_{n+1} = f_n(T_n, X_{n+1})$. Note that f_n is specified independently of the data and that $T_n \varepsilon \{1, 2, 3, 4\}$ is a finite statistic sequentially summarizing the past observations. However, the dependence of f_n on n requires external specification of n if the algorithm is to be considered to have truly finite memory. It is clear that requiring f_n to be independent of n will preclude $\alpha_n \to 0$, $\beta_n \to 0$ except in the singular case. Thus, requiring the probability of error to approach zero requires an "infinite" algorithm. The variation of f_n with n is a natural way to meet this requirement. The algorithm has been factored into two parts—that dealing with the data is finite, while the part concerned with the data processing is unbounded. In contrast, in the theory of computation, Turing machines have an essentially infinite memory (an infinite tape) with finite computation (i.e., an f_n which is independent of n).

Fortunately, time-independent algorithms of the form $T_{n+1} = f(T_n, X_{n+1})$ have an interesting theory of their own. The determination of all ϵ-admissible time-independent finite-memory algorithms for the two-hypothesis testing problems will be given in [8]. Solutions of some problems in the sequential design of experiments under the finite memory constraints of this paper and [8] may be found in [3] and [4].

Acknowledgment. I wish to thank J. MacQueen for the many key ideas he has graciously contributed to the formulation of this problem.

REFERENCES

[1] Abramson, N., and Braverman, D. (1962). Learning to recognize patterns in a random environment. *IEEE Trans. Information Theory.* **IT-8** 58–63.

[2] Berman, Simeon M. (1962). A law of large numbers for the maximum in a stationary gaussian sequence. *Ann. Math. Statist.* **33** 93–97.

[3] Cover, Thomas M. (1968). A note on the two-armed bandit problem with finite memory. *Information and Control.* **12** 371–377.

[4] Cover, Thomas M., and Hellman, Martin E. The two-armed bandit with time-independent finite memory. Submitted to *IEEE Trans. Information Theory,* April 1969.

[5] Denny, J. L. (1964). A continuous real-valued function on E^n almost everywhere $1 - 1$. *Fund. Math.* **55** 95–99.

[6] Dynkin, E. B. (1961). Necessary and sufficient statistics for a family of probability distributions. *Selected Translations in Mathematical Statistics and Probability.* **1** 17–40.

[7] Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d'une serie aléatorie. *Ann. of Math.* **44** 423–453.

[8] Hellman, Martin E. and Cover, Thomas M. Learning with finite memory. Submitted to *Ann. Math. Statist.*, December 1968.

[9] Robbins, H. (1956). A sequential decision problem with finite memory. *Proc. Nat. Acad. Sci., U.S.A.* **42** 920-923.

[10] Smith, C. V., and Pyke, R. (1965). The Robbins-Isbell two-armed bandit problem with finite memory. *Ann. Math. Statist.* **36** 1375–1386.

[11] Spragins, J. (1966). Learning without a teacher. *IEEE Trans. on Information Theory.* **IT-12** 223–230.