[2] ARYEH DVORETZKY, "On stochastic approximation," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 39-56, University of California Press, 1956.
[3] J. KIEFER and J. WOLFOWITZ, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Stat.*, Vol. 23 (1952), pp. 462–466.
[4] MICHEL LOÈVE, *Probability Theory*, D. Van Nostrand Company, New York, 1955.
[5] HERBERT ROBBINS and SUTTON MONRO, "A stochastic approximation method," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 400–407.
[6] JEROME SACKS, "Asymptotic Distribution of Stochastic Approximation Procedures," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 373–405.
[7] J. WOLFOWITZ, "On Stochastic Approximation Methods," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 1151–1155.

ON A PROBLEM OF ROBBINS

BY J. R. ISBELL

University of Washington

1. Introduction. This note concerns a sequential decision problem raised by Herbert Robbins [2]. The problem is not solved; in fact, it is not known if there is a uniformly best procedure. A procedure is given here which is uniformly better than the one proposed in [2] and is best at least in a special case.

The nature of the problem is this: given two coins with unknown probabilities p_1, p_2, of coming up heads, to prescribe a rule for making an infinite sequence of tosses, choosing the coin for the nth toss as a function of the history of the sequence since the $(n - r)$-th toss (inclusive). The memory length r is fixed. The aim is to maximize the frequency of heads.

The rule proposed here is best in case p_1 or p_2 is 0. We cannot say *the* best, since many rules have the same effects in this case. The rule may be briefly stated: "*Change coins when one coin shows tails r successive times, or when $r - 1$ tails with one coin are followed by a single toss with the other coin, which is tails*". Robbins' rule [2] calls for changing in these cases and further whenever the first toss with a new coin is tails. For $r \leqq 2$, the rules coincide. Otherwise the present rule is better except in two trivial cases, $p_1 = p_2$ and $\max(p_1, p_2) = 1$.

2. Formulation. The description of the memory requires some amplification for the case $n < r$. (None is given in [2].) Here we shall regard the sequence of tosses as a Markov process with 4^r states, namely the states of the memory. We consider that the process may begin in any state, and we propose to evaluate any procedure according to the results it yields starting from the worst possible state.

This is an artificial description which one might prefer to avoid. On the other hand, any decision procedure which might be optimal according to some other version of the problem but disqualified by our artificial start could be described

Received July 8, 1958; revised January 5, 1959.