The Annals of Mathematical Statistics
1968, Vol. 39, No. 2, 674-675

NOTES

AN EXAMPLE IN DENUMERABLE DECISION PROCESSES

BY LLOYD FISHER AND SHELDON M. ROSS[1]

University of Washington and Stanford University

1. Introduction and summary. This note uses the terminology and notation of Ross [3]. The example presented here is a denumerable Markovian decision process which has an optimal nonstationary rule which is better than every stationary rule. This answers a question of Derman [1].

2. The example.

$$I = \{0, 1, 1', 2, 2', 3, 3', \cdots\}, \qquad K_0 = K_{i'} = 1, \qquad K_i = 2,$$

$$P(0, i{:}1) = P(0, i'{:}1) = \tfrac{3}{2}(\tfrac{1}{4})^i, \qquad i > 0,$$

$$P(i, 0{:}1) = (\tfrac{1}{2})^i \qquad = 1 - P(i, i'{:}1),$$

$$P(i, 0{:}2) = \tfrac{1}{2} \qquad = 1 - P(i, i + 1{:}2),$$

$$P(i', 0{:}1) = (\tfrac{1}{2})^i \qquad = 1 - P(i', i'{:}1),$$

$$C(0, \cdot) = 1, \qquad \text{all other costs are zero, i.e.,}$$

$$C(i, \cdot) = 0 = C(i', \cdot), \qquad i > 0.$$

Let R_n be the stationary deterministic rule which takes action 2 at states $0<i<n$ and action 1 elsewhere

$$M_{00}(R_n) = 1 + \textstyle\sum_{j=1}^{\infty} \tfrac{3}{2}(\tfrac{1}{4})^j M_{j0}(R_n) + \sum_{j=1}^{\infty} \tfrac{3}{2}(\tfrac{1}{4})^j 2^j.$$

Now $j \geqq n \Rightarrow M_{j0}(R_n) = 2^j$, whereas

$$j < n \Rightarrow M_{j0}(R_n) = \tfrac{1}{2} + 2(\tfrac{1}{2})^2 + \cdots + (n-j)(\tfrac{1}{2})^{n-j} + (\tfrac{1}{2})^{n-j}[n - j + 2^n]$$
$$= 2 + 2^j - (\tfrac{1}{2})^{n-j-1}.$$

Therefore

$$M_{00}(R_n) = \tfrac{5}{2} + \textstyle\sum_{j=1}^{\infty} \tfrac{3}{2}(\tfrac{1}{4})^j(2 + 2^j) - \sum_{j=1}^{n-1} \tfrac{3}{2}(\tfrac{1}{4})^j(\tfrac{1}{2})^{n-j-1}$$
$$- 2 \textstyle\sum_{j=n}^{\infty} \tfrac{3}{2}(\tfrac{1}{4})^j$$
$$= 5 - \textstyle\sum_{j=1}^{n-1} \tfrac{3}{2}(\tfrac{1}{4})^j(\tfrac{1}{2})^{n-j-1} - 3 \sum_{j=n}^{\infty} (\tfrac{1}{4})^j.$$

Hence $M_{00}(R_n) < 5$ for all n, and $M_{00}(R_n) \to 5$ as $n \to \infty$.

Received 26 May 1967; revised 6 November 1967.
[1] This work was supported in part by the Army, Navy, Air Force and NASA under contract Nonr 225(53) (NR-042-002) with the Office of Naval Research.

Now let R be any stationary rule, let P_i be the probability that R takes action 1 when in state i. Now

$$M_{00}(R) = 1 + \sum_{j=1}^{\infty} \tfrac{3}{2}(\tfrac{1}{4})^j M_{j0}(R) + \sum_{j=1}^{\infty} \tfrac{3}{2}(\tfrac{1}{4})^j 2^j$$

but

$$M_{j0}(R) = \sum_{n=j}^{\infty} [P_n \prod_{k=j}^{n-1} (1 - P_k)] M_{j0}(R_n) + 2 \prod_{k=j}^{\infty} (1 - P_k)$$
$$< (2 + 2^j) [\sum_{n=j}^{\infty} P_n \prod_{k=j}^{n-1} (1 - P_k) + \prod_{k=j}^{\infty} (1 - P_k)] = 2 + 2^j.$$

Consequently

$$M_{00}(R) < 5 \text{ for all stationary rules } R, \text{ and}$$
$$\varphi(i, R) > \tfrac{1}{5} \text{ for all stationary rules } R, \text{ for all } i.$$

However if we consider the nonstationary rule R^* which uses

$$R_1 \quad \text{for} \quad t = 1, 2, \cdots, N_1,$$
$$R_2 \quad \text{for} \quad t = N_1 + 1, \cdots, N_1 + N_2,$$
$$\vdots$$
$$R_n \quad \text{for} \quad t = \sum_{i=1}^{n-1} N_i + 1, \cdots, \sum_{i=1}^{n} N_i,$$
$$\vdots$$

it can be shown (as in Theorem 4.3 of Ross [3]) that there exists N_j's such that $\varphi(i, R^*) = \lim_n \varphi(i, R_n) = \frac{1}{5}$. It also follows from Theorem 3.1 of Ross [3] that R^* is optimal.

Since every stationary rule gives rise to a recurrent Markov chain the results of Fisher [2] may also be used to show this result.

REFERENCES

[1] Derman, C. (1966). Denumerable state Markovian decision processes—average cost criterion. *Ann. Math. Statist.* **37** 1545–1553.

[2] Fisher, L. (1968). On recurrent denumerable decision processes. *Ann. Math. Statist.* **39** 424–432.

[3] Ross, S. M. (1968). Non-discounted denumerable Markovian decision models. *Ann. Math. Statist.* **39** 412–423.