Adv. Appl. Prob. **15,** 216–218 (1983)
Printed in N. Ireland

THE MEAN WAITING TIME TO A REPETITION

GUNNAR BLOM*, *University of Lund*

Abstract

Let $X_1, X_2, \cdots$ be a stationary sequence of random variables and $E_1, E_2, \cdots, E_N$ mutually exclusive events defined on k consecutive X's such that the probabilities of the events have the sum unity. In the sequence $E_{j_1}, E_{j_2}, \cdots$ generated by the X's, the mean waiting time from an event, say E_{j_1}, to a repetition of that event is equal to N (under a mild condition of ergodicity). Applications are given.

STATIONARITY; BERNOULLI TRIALS; MARKOV CHAIN; UPCROSSINGS AND DOWNCROSSINGS

1. Introduction

Johnson (1968) considered independent trials with m mutually exclusive outcomes $E_1, E_2, \cdots, E_m$, where $P(E_j) = p_j$. Let S be the result of k consecutive trials and T the waiting time, that is, the number of additional trials required to repeat the pattern S, where $T = k, k+1, \cdots$. Johnson showed that, averaging over all possible patterns of length k, the mean waiting time is $m^k + k - 1$, and hence does not depend on the probabilities p_j.

Johnson's result depends crucially on the fact that overlaps are not allowed. For example, in the binary sequence 1 0 1 0 1 0 0 1 0 1, repetition of the first 1 0 1 occurs after $T = 7$ trials.

In the present note we change the definition of T so that overlaps are allowed; hence in the above example $T = 2$. Also, the problem is given a somewhat more general formulation.

2. Main result

Let $X_1, X_2, \cdots$ be a stationary sequence of random variables and $E_1, E_2, \cdots, E_N$ mutually exclusive events defined as follows on k consecutive X's: The event E_j occurs if $(X_1, \cdots, X_k) \in A_j$, where $\{A_1, \cdots, A_N\}$ is a partition of R_k. Set

$$P(E_j) = P[(X_1, \cdots, X_k) \in A_j], \qquad j = 1, \cdots, N.$$

The P's sum to unity.

The sequence $X_1, X_2, \cdots$ generates a sequence $E_{j_1}, E_{j_2}, \cdots$ of events. In what follows, we shall suppose that the following condition is fulfilled:

Received 16 June 1982.

* Postal address: Department of Mathematical Statistics, University of Lund, Box 725, S-220 07 Lund, Sweden.

In the sequence $E_{j_1}, E_{j_2}, \cdots$, each possible event E_j, $j = 1, \cdots, N$, occurs at least once, with probability 1.

Consider the waiting time T from an event in a certain position, say E_{j_1}, to the first repetition of this event. For each given $E_j = E_i$, the conditional mean of T is given by $M_i = 1/P(E_i)$; cf. Breiman (1968), p. 123. Averaging over all possible E_i, we immediately obtain the overall mean

$$E(T) = \sum_{i=1}^{N} M_i P(E_i) = N.$$

Hence the mean waiting time is equal to the number N of events E_j.

3. Examples

We give several examples, to show the wide applicability of the result.

Example 1. Patterns in Bernoulli trials. Consider a sequence of Bernoulli trials $1\,0\,1\,0\,1\,0\,0\,1\,0\,1 \cdots$, where the probability of 1 and 0 is p and $q = 1-p$, respectively. Take $k = 3$ and consider all possible $N = 8$ patterns 000, 001, $\cdots$, 111. The mean waiting time from the pattern in the three first positions to a repetition of this pattern is 8. (If overlaps are not allowed, the mean distance is 10, as shown by Johnson.)

Example 2. Small prison. A prison receives N types of prisoners, one prisoner each morning. The probability is p_j that the prisoner is of type j. The prison is small; there is only one one-person cell for each type of prisoner. Therefore, as soon as a prisoner of type j arrives, the prisoner in cell j is made free. The mean time of imprisonment is N days.

Example 3. Poisson distribution. Consider a sequence of independent Poisson variables with the same mean. The mean waiting time, until the value assumed by the first variable is repeated, is infinite.

Example 4. Markov chain. Consider an N-state ergodic Markov chain that has reached equilibrium. Let $E_1, \cdots, E_N$ be the states of the chain. Consider a certain time-point. The mean waiting time to a repetition of the state assumed at this time-point is equal to the number of states N, averaging over all possible states.

Example 5. Upcrossings and downcrossings. Consider a stationary sequence of continuous random variables $X_1, X_2, \cdots$. We are interested in studying whether $k = 2$ successive values lie above or below a certain level a, according to the following classification into $N = 4$ events:

E_1: $X_1 < a, X_2 > a$ (upcrossing)
E_2: $X_1 > a$, $X_2 < a$ (downcrossing)
E_3: $X_1 > a$, $X_2 > a$ (both values at high level)
E_4: $X_1 < a$, $X_2 < a$ (both values at low level).

For example, if $a = 0$ and the values are

$$0{\cdot}51 \quad -0{\cdot}23 \quad -0{\cdot}34 \quad 0{\cdot}05 \quad 0{\cdot}16 \quad -0{\cdot}60,$$

the sequence of events is $E_2\ E_4\ E_1\ E_3\ E_2$. The mean waiting time from one event to a repetition is 4.

Acknowledgement

Daniel Thorburn made several thought-provoking comments on the first draft of this letter.

References

BREIMAN, L. (1968) *Probability*. Addison-Wesley, Reading, Mass.
JOHNSON, N. L. (1968) Repetitions. *Amer. Math. Monthly* **75,** 382–383.