The Annals of Probability
1974, Vol. 2, No. 1, 32-39

A RANDOM WALK WITH NEARLY UNIFORM N-STEP MOTION[1]

BY LAWRENCE E. MYERS

University of California at Berkeley

Let N be a strictly positive integer. Motivated by a certain discrete evasion game, we search for a $\{0, 1\}$-valued discrete time stochastic process whose conditional-on-the-past distributions of the sum of the next N terms are as close to uniform as possible. A process is found for which none of the sums ever occurs with conditional probability more than $2e/(N+1)$. The process is characterized by invariance under interchange of 0 and 1, and its waiting times between successive transitions, which are independently, identically, and uniformly distributed over $\{1, 2, \cdots, N+1\}$.

1. Introduction. Let N be a strictly positive integer, and let $\beta = \{\beta_n\}_{n=1}^\infty$ be a $\{0, 1\}$-valued discrete time process. Suppose that at time k the finite binary sequence $s = s_1 \cdots s_k$ has been realized, i.e. $\beta_1 = s_1, \cdots, \beta_k = s_k$. Let us inquire as to the conditional probability $P[\sum_{i=1}^N \beta_{k+i} = r \mid \beta_1 \cdots \beta_k = s]$ that the sum of the next N consecutive terms is r, where r ranges over the $N+1$ possible sums $0, 1, \cdots, N$. Can we find a β for which all these conditional distributions are uniform?

The quantity of interest to us is then $v(N)$, the smallest probability p for which there exists a $\{0, 1\}$-valued discrete time process with the property that conditional on the past at any time, none of the $N+1$ possible sums of the next N successive terms ever occurs with probability more than p. To be precise,

$$v(N) = \inf_{\beta \in B} \sup_{0 \leqq k < \infty, s \in \{0,1\}^k} \max_{0 \leqq r \leqq N} P[\textstyle\sum_{i=1}^N \beta_{k+i} = r \mid \beta_1 \cdots \beta_k = s] \,.$$

Here B is the collection of all $\{0, 1\}$-valued discrete time processes $\beta = \{\beta_n\}_{n=1}^\infty$. Also if $k = 0$ or $P[\beta_1 \cdots \beta_k = s] = 0$ we set

$$P[\textstyle\sum_{i=1}^N \beta_{k+i} = r \mid \beta_1 \cdots \beta_k = s] = P[\sum_{i=1}^N \beta_{k+i} = r] \,.$$

Our purpose is to investigate the asymptotic behavior of $v(N)$.

$v(N) \geqq 1/(N+1)$ for all $N \geqq 1$ since for any $\beta \in B$ there is some $r = r(\beta)$ for which $P[\sum_{i=1}^N \beta_i = r] \geqq 1/(N+1)$. By considering the process $\{\beta_n\}_{n=1}^\infty$ whose terms are independently and identically distributed with $P\{\beta_n = 1\} = \frac{1}{2}$ for all $n \geqq 1$, we see that $v(1) = \frac{1}{2}$ and, further, $v(N) < 1/N^{\frac{1}{2}}$ for all $N \geqq 1$. The upper bound is obtained by applying Stirling's formula to the central (modal) term of the binomial distribution with parameters N and $p = q = \frac{1}{2}$. We thus have

$$(1) \qquad 1/(N+1) \leqq v(N) < 1/N^{\frac{1}{2}} \quad \text{for all} \quad N \geqq 1 \,.$$

Received January 12, 1973; revised May 21, 1973.

[1] This research was done with the support of a National Science Foundation Graduate Traineeship, Grant GZ-1950.

AMS 1970 *subject classifications*. Primary 60G17, 60K99; Secondary 60C05.

Key words and phrases. Random walk, N-step motion, m-dependent process.

Our principal aim now is to improve (asymptotically) on this upper bound by showing $v(N) < 2e/(N+1)$ for all $N \geqq 1$.

REMARKS. Consider the following game between two players:

Player I observes Player II as the latter writes a sequence of 0's and 1's. At some point, unknown to Player II, Player I gives the referee a prediction of the sum of the next N numbers Player II writes (i.e. I's prediction should be an integer from $\{0, 1, 2, \cdots, N\}$).

If Player I's prediction turns out to be correct, he wins one dollar from Player II; otherwise, I wins nothing.

Players I and II are often referred to as the pursuer (predictor, bomber) and evader (sequence chooser, battleship), respectively, and the game is called the N-move lag bomber-battleship game.

David Blackwell [1] proved that the Nth game has value $v(N)$, and that in each case the evader possesses an optimal stationary strategy. Since the possible mixed strategies for Player II are exactly the members of B, it follows that each $v(N)$ is attained by some $\beta \in B$, and the infimum in the definition of $v(N)$ is in fact a minimum.

We saw earlier that $v(1) = \frac{1}{2}$. Lester Dubins [2], by way of solving the two-move lag game, proved $v(2) = (3 - 5^{\frac{1}{2}})/2$ and is attained by a Markov process. Hence $v(N) \neq 1/(N+1)$ for $N = 2$.

Thomas Ferguson discusses the extension of this game to the same game with a three-move information lag in [4]. He therein defines an *m-dependent process* as one which at any time depends on the past only through the preceding m states (so that a Markov process is 1-dependent). $v(3)$ is not known. However, Ferguson shows that $v(3)$ is not attained by a Markov process by exhibiting a two-dependent process which improves on the best Markov one.

2. The definition of the process and some of its elementary properties. The following notation will be used throughout. x and y and subscripted s's and t's will be generic members of $\{0, 1\}$. SEQ will denote the set of all finite sequences of 0's and 1's, including the null sequence $\langle\ \rangle$. $s = s_1 \cdots s_m$ and $t = t_1 \cdots t_n$ will be generic members of SEQ. st is that member of SEQ given by $s_1 \cdots s_m$, $t_1 \cdots t_n$. For $s = s_1 \cdots s_m \in \text{SEQ}$, $|s| = m$ is the *length* of s, with $|\langle\ \rangle| = 0$.

For $x \in \{0, 1\}$, $(0x)$ denotes the null sequence, and for $i \geqq 1$, (ix) is that member of SEQ of length i given by $xx \cdots x$, so that, for example, $sy(ix) = s_1 \cdots s_m yxx \cdots x \in \text{SEQ}$. Expressions such as $s\langle\ \rangle$, $\langle\ \rangle s$, and $s\langle\ \rangle t$ are to be interpreted as s, s, st respectively.

For any nonnull $s = s_1 \cdots s_m \in \text{SEQ}$ and $\beta = \{\beta_n\}_{n=1}^{\infty} \in B$, let $\boldsymbol{\beta}(s) = P[\beta_1 = s_1, \cdots, \beta_m = s_m]$, and let $S(\beta) = \{s \in \text{SEQ} : s \neq \langle\ \rangle \text{ and } \boldsymbol{\beta}(s) > 0\}$.

Let $N \geqq 1$ be fixed for the remainder of our discussion. To define the process $\{\alpha_n\}_{n=1}^{\infty}$ of interest to us, we first define a function $\alpha : SEQ \rightarrow \mathscr{R}(\{0, 1\})$ from SEQ into the space $\mathscr{R}(\{0, 1\})$ of $\{0, 1\}$-valued random variables.

For $x, y \in \{0, 1\}$, $x \neq y$, $1 \leqq k \leqq N+1$, $s \in SEQ$, let $\alpha(sy(kx)) = y$ with probability $1/(N+2-k)$, $\alpha(sy(kx)) = x$ with probability $(N+1-k)/(N+2-k)$. For any nonnull $tx \in SEQ$ which is not of the above form $sy(kx)$, let $\alpha(tx) = y$ with probability 1; tx may equal x here. Set the *predecessor* $\alpha_0 = 1$ with probability $\frac{1}{2}$, $\alpha_0 = 0$ with probability $\frac{1}{2}$, and let $\alpha(\langle\ \rangle) = \alpha(\alpha_0)$.

$\{\alpha_n\}_{n=1}^{\infty}$ is then defined inductively by $\alpha_1 = \alpha(\langle\ \rangle)$ and $\alpha_n = \alpha(\alpha_0 \cdots \alpha_{n-1})$ for $n \geqq 2$. We will often refer to $\{\alpha_n\}_{n=1}^{\infty}$ as α, though some abuse of notation is involved in this.

DEFINITION. The process $\beta = \{\beta_n\} \in B$ is *invariant under interchange of* 0 *and* 1 provided $\boldsymbol{\beta}(s) = \boldsymbol{\beta}(s')$ for each $s \in SEQ$, where s' is the sequence obtained from s by interchanging 0 and 1.

PROPOSITION 1. *$\{\alpha_n\}_{n=1}^{\infty}$ is an $(N+1)$-dependent process which is invariant under interchange of* 0 *and* 1. *At any time, the process depends on the past only back to the latest change of state.*

PROOF. Clear from the definitions.

For the sake of convenience, set, for $s \in S(\alpha)$ and $A \subset \{0, 1\}^n$,

$$Q[A \mid s] = P[\alpha_{m+1} \cdots \alpha_{m+n} \in A \mid \alpha_1 \cdots \alpha_m = s]\,, \qquad Q[\langle\ \rangle \mid s] = 1\,,$$

and for $0 \leqq r \leqq n \leqq N$, define $A(r, n)$ to be the set of all finite binary sequences of length n having exactly r 1's. Also, let $S = S(\alpha)$. Then we obviously have $v(N) \leqq \sup_{s \in S} \max_{0 \leqq r \leqq N} Q[A(r, N) \mid s]$.

PROPOSITION 2. *Let $s \in SEQ$ and let x and y be distinct members of* $\{0, 1\}$. *Then the following assertions hold:*

(i) *If $1 \leqq k \leqq N+1$, then $(kx) \in S$ and $y(kx) \in S$.*
(ii) *If $s(kx) \in S$ or $sy(kx) \in S$, then $k \leqq N+1$.*
(iii) *The process $\{\alpha_n\}_{n=1}^{\infty}$ changes states at least once in each consecutive $N+2$ units of time.*
(iv) *$S \subset \{(kx) : 1 \leqq k \leqq N+1$ and $x \in \{0, 1\}\} \cup \{sy(kx) : s \in SEQ,\ 1 \leqq k \leqq N+1;\ x, y \in \{0, 1\},\ x \neq y\}$.*

PROOF. Parts (iii) and (iv) follow from (ii). Part (ii) follows, by contraposition, from the fact that α is $(N+1)$-dependent and $P[\alpha(s([N+1]x)) = x] = 0$, whatever be $s \in SEQ$.

Part (i) remains. We can assume without loss of generality (by the invariance of α) that $x = 0$ and $y = 1$. Then

$$\begin{aligned}\boldsymbol{\alpha}((kx)) &= \boldsymbol{\alpha}((k0)) \\ &= P[\alpha_1 \cdots \alpha_k = (k0)] \geqq P[\alpha_0 = 1 \text{ and } \alpha_1 \cdots \alpha_k = (k0)] \\ &= \tfrac{1}{2}1 \frac{N}{N+1} \frac{N-1}{N} \cdots \frac{N+1-(k-1)}{N+2-(k-1)} \\ &= \tfrac{1}{2} \times \frac{N+2-k}{N+1} > 0 \qquad \text{for all } 1 \leqq k \leqq N+1\,.\end{aligned}$$

The proof that $y(kx) \in S$ for $1 \leqq k \leqq N+1$ is similar.

COROLLARY. $v(N) \leqq \max_{1 \leq k \leq N+1} \max_{0 \leqq r \leq N} Q[A(r, N) \mid 0(k1)]$.

PROOF. Clear from Proposition 1, (iv) of Proposition 2, and the fact that the predecessor α_0 effectively incorporates a turn into each possible path.

PROPOSITION 3. *Given that the last change of state under α occurred k time units ago, then $k \leqq N+1$, and*

(i) *α remains in its present state for i more units of time with conditional probability*

$$\frac{N+2-k-i}{N+2-k} \quad \text{for } i \leqq N+1-k$$

$$0 \quad \text{for } i > N+1-k, \qquad \textit{and}$$

(ii) *with conditional probability* 1 *the process changes states at least once within the next $N+2-k$ units of time, and that first change occurs at each of the possible $N+2-k$ times with conditional probability $1/(N+2-k)$.*

PROOF. The assertion $k \leqq N+1$ is just part (iii) of Proposition 2. For the rest, simple computations show that for $1 \leqq k \leqq N+1$, and x and y distinct members of $\{0, 1\}$,

$$\text{(i)} \quad Q[(ix) \mid y(kx)] = \frac{N+2-k-i}{N+2-k} \quad \text{for } i \leqq N+1-k$$

$$= 0 \quad \text{for } i > N+1-k, \qquad \text{and}$$

$$\text{(ii)} \quad Q[(ix)y \mid y(kx)] = \frac{1}{N+2-k} \quad \text{for } i \leqq N+1-k$$

$$= 0 \quad \text{for } i > N+1-k.$$

REMARK. Should the evader wish to employ the process α as his strategy in the N-move lag game, then it is clear from Proposition 3 that he needs only a fair coin and a fair $(N+1)$-sided die to do so. He begins with a flip of the coin and then makes repeated throws of the die. The coin determines whether he begins with 0 or 1 and the die determines the waiting times between transitions. To illustrate, suppose the coin lands heads and the die throws yield the sequence $k(1), k(2), \cdots$. Then his first $k(1)$ moves will be 1's, his next $k(2)$ moves will be 0's, and so on.

3. THEOREM. $1/(N+1) \leqq v(N) < 2e/(N+1)$ *for all* $N \geqq 1$.

PROOF. The lower bound was established in 1. By the corollary to Proposition 2 it suffices to show that $Q[A(r, N) \mid 0(k1)] < 2e/(N+1)$ for all $0 \leqq r \leqq N$ and $1 \leqq k \leqq N+1$.

Suppose first that $r = N$. Then by Proposition 3 $Q[A(N, N) \mid 0(k1)] = Q[(N1) \mid 0(k1)] \leqq 1/(N+1)$ for $1 \leqq k \leqq N+1$.

For the case $0 \leqq r < N$, set $W = \min(N+1-k, r)$. Then

$$\begin{aligned} &Q[A(r, N) \mid 0(k1)] \\ &= \textstyle\sum_{i=0}^{W} Q[(i1)0 \mid 0(k1)]Q[A(r-i, N-(i+1)) \mid 10] \\ &= \frac{1}{N+2-k} \textstyle\sum_{i=0}^{W} Q[A(r-i, N-i-1) \mid 10] \\ &= \frac{1}{N+2-k} \textstyle\sum_{i=0}^{W} Q[A(N-r-1, N-i-1) \mid 01], \end{aligned}$$

by the invariance of α.

Now suppose

(2) $$Q[A(r, n) \mid 01] \leqq \frac{2e-1}{N+1} \quad \text{for all} \quad 0 \leqq r < n \leqq N.$$

Then at most one of the above summands is of the form $Q[A(N-i-1, N-i-1) \mid 01] = [N+1-(N-i-1)]/(N+1) = (i+2)/(N+1)$ (Proposition 3), and each of the others is bounded above by $(2e-1)/(N+1)$ by the assumption (2). Hence

$$\begin{aligned} &\frac{1}{N+2-k} \textstyle\sum_{i=0}^{W} Q[A(N-r-1, N-i-1) \mid 01] \\ &\leqq \frac{1}{N+2-k}\left[(N+1-k)\frac{2e-1}{N+1} + \frac{N+1-k+2}{N+1}\right] \\ &< \frac{1}{N+2-k}\left[(N+2-k)\frac{2e-1}{N+1} + \frac{N+2-k}{N+1}\right] = \frac{2e}{N+1}, \end{aligned}$$

and the theorem is proved given the assumption (2).

We prove (2) through two lemmas by partitioning $A(r, n)$ into events which are simpler in the sense that any two members of one of them have approximately the same conditional probability given a past ending in 01; our breakdown of $A(r, n)$ into disjoint subsets will be based on the fact that any nonnull binary sequence is composed of alternating blocks of 0's and 1's. Roughly speaking, Lemmas 1 and 2 deal with the cases where conditional on a past ending in 01, the event $A(r, n)$ is realized through an even (Lemma 1) or odd (Lemma 2) number of changes of state.

For $j \geqq 0$ and $0 \leqq r < n \leqq N$ let $A'(j, r, n) = \{(l_0 1)(l_1 0) \cdots (l_{2j} 1) \in A(r, n): \sum_{i=0}^{2j} l_i = n, \sum_{i=0}^{j} l_{2i} = r, \text{ and } l_i > 0 \text{ for } i \geqq 1\}$, and let $A''(j, r, n) = \{(l_0 1)(l_1 0) \cdots (l_{2j} 1)(l_{2j+1} 0) \in A(r, n): \sum_{i=0}^{2j+1} l_i = n, \sum_{i=0}^{j} l_{2i} = r, \text{ and } l_i > 0 \text{ for } i \geqq 1\}$. Then since $r < n$, $A'(0, r, n)$ is empty and

(3) $$A(r, n) = \textstyle\sum_{j \geqq 1} A'(j, r, n) + \sum_{j \geqq 0} A''(j, r, n).$$

LEMMA 1. *For* $0 \leqq r < n \leqq N$

$$Q[\textstyle\sum_{j \geqq 1} A'(j, r, n) \mid 01] \leqq \frac{e-1}{N+1}.$$

PROOF. By Proposition 3, if $(l_0 1)(l_1 0)\cdots(l_{2j}1)\in A'(j,r,n)$

$$\begin{aligned} &Q[(l_0 1)(l_1 0)\cdots(l_{2j}1)\,|\,01] \\ &\quad = Q[(l_0 1)0\,|\,01]Q[([l_1-1]0)1\,|\,10]\cdots \\ &\qquad Q[([l_{2j-1}-1]0)1\,|\,10]Q[([l_{2j}-1]1)\,|\,01] \\ &\quad = \left(\frac{1}{N+1}\right)^{2j} Q[([l_{2j}-1]1)\,|\,01] \leqq \left(\frac{1}{N+1}\right)^{2j}. \end{aligned}$$

Hence $Q[A'(j,r,n)\,|\,01] \leqq (1/(N+1))^{2j}\sharp(A'(j,r,n))$, where $\sharp(A'(j,r,n))$ denotes the cardinality of $A'(j,r,n)$.

Now, the number of distinct realizations of the event $A'(j,r,n)$ is just ab, where a is the number of different ways of placing r indistinguishable balls into $j+1$ boxes in such a way that every box, except possibly the first, contains at least one ball ($l_i > 0$ for all $i \geqq 1$, but l_0 may equal 0). Similarly, b is the number of ways of placing $n-r$ balls into j boxes so that none of the boxes is empty.

Hence a is the number of different $(j+1)$-tuples of nonnegative integers satisfying $m_1+\cdots+m_{j+1}=r-j$ and b is the number of j-tuples of nonnegative integers satisfying $m_1+\cdots+m_j=n-r-j$. We have ([3] page 38)

$$a = \binom{(r-j)+(j+1)-1}{(j+1)-1} = \binom{r}{j} \quad \text{and}$$

$$b = \binom{(n-r-j)+j-1}{j-1} = \binom{n-r-1}{j-1}, \quad \text{so that}$$

$$\begin{aligned} &\textstyle\sum_{j\geqq 1} Q[A'(j,r,n)\,|\,01] \\ &\quad \leqq \sum_{j\geqq 1}\binom{n-r-1}{j-1}\binom{r}{j}\left(\frac{1}{N+1}\right)^{2j} \\ &\quad \leqq \frac{1}{N+1}\left(\sum_{j\geqq 1}\binom{n-r-1}{j-1}\left(\frac{1}{N+1}\right)^{j-1}\right)\left(\sum_{j\geqq 1}\binom{r}{j}\left(\frac{1}{N+1}\right)^{j}\right) \\ &\quad = \frac{1}{N+1}\left(1+\frac{1}{N+1}\right)^{n-r-1}\left(\left(1+\frac{1}{N+1}\right)^{r}-1\right) \\ &\quad = \frac{1}{N+1}\left(\left(1+\frac{1}{N+1}\right)^{n-1}-\left(1+\frac{1}{N+1}\right)^{n-r-1}\right) \\ &\quad < \frac{1}{N+1}\left(\left(1+\frac{1}{N+1}\right)^{N+1}-1\right) < \frac{e-1}{N+1}. \end{aligned}$$

This proves Lemma 1.

LEMMA 2. *For* $0 \leqq r < n \leqq N$

$$Q\left[\textstyle\sum_{j\geqq 0} A''(j,r,n)\,|\,01\right] \leqq \frac{e}{N+1}.$$

PROOF. After the fashion of Lemma 1, $s\in A''(j,r,n)$ implies $Q[s\,|\,01]\leqq (1/(N+1))^{2j+1}$, so that $Q[A''(j,r,n)\,|\,01] \leqq (1/(N+1))^{2j+1}\,\sharp(A''(j,r,n))$ and

$$\textstyle\sum_{j\geqq 0} Q[A''(j,r,n)\,|\,01] \leqq \sum_{j\geqq 0}\binom{n-r-1}{j}\binom{r}{j}\left(\frac{1}{N+1}\right)^{2j+1} \leqq \frac{e}{N+1}.$$

The assumption (2) now follows immediately from (3) and the two lemmas, and we are done.

The following are immediate consequences of the theorem, and the proofs are omitted.

COROLLARY 1. $\{N \cdot v(N)\}_{N=1}^{\infty}$ *is a bounded sequence.*

COROLLARY 2. $\lim_{N\to\infty} N^a \cdot v(N) = 0$ *for each* $a \in [0, 1)$.

4. Concluding remarks. It was mentioned in the remarks at the end of Section 1 that each $v(N)$ is actually attained by some process, but that $v(3)$ is not attained by any Markov process. One may ask, however, whether an order of $1/N$ may be achieved if we restrict attention to some subset of B, such as the Markov processes, or even the processes whose terms are independent.

It is easily seen that each $v(N)$ may be attained with a process invariant under interchange of 0 and 1. Blackwell [1] shows that we may also restrict attention to stationary processes.

For each $m \geqq 0$ let B_m be that subset of B consisting of all the m-dependent processes. (B_0 then contains only processes whose terms are independent; B_1 consists of Markov processes, etc.)

Also, for each $N \geqq 1$ and $m \geqq 0$ let

$$v_m(N) = \inf_{\beta\in B_m} \sup_{0\leq k<\infty, s\in\{0,1\}^k} \max_{0\leq r\leq N} P[\textstyle\sum_{i=1}^{N} \beta_{k+i} = r \mid \beta_1 \cdots \beta_k = s] .$$

Then $v(N) \leqq v_m(N)$ for each m and N. The reader may have noted that the process used in 1 to show $v(N) < 1/N^{\frac{1}{2}}$ for N sufficiently large also yields the result

$$1/(2N)^{\frac{1}{2}} < v_0(N) < 1/N^{\frac{1}{2}} \qquad \text{for } N \text{ sufficiently large.}$$

Does there exist $a > 0$ such that

$$v_1(N) < a/(N+1) \quad \text{for all} \quad N \geqq 1? \tag{4}$$

That is, can we attain order $1/N$ with Markov processes? The author conjectures the answer to be in the negative. In fact, he suspects that for each $m \geqq 0$ there exists strictly positive $b(m)$ such that

$$b(m)/N^{\frac{1}{2}} < v_m(N) < 1/N^{\frac{1}{2}} \quad \text{for all} \quad N \geqq 1 .$$

If, for example, (4) is satisfied by some $a > 0$, then the process α used in 3. to show $\{N \cdot v(N)\}_{N=1}^{\infty}$ is a bounded sequence is unnecessarily complicated in that it is $(N+1)$-dependent. At this point we can only say that α is computationally convenient, and invite the reader to compute (the order of) $v_1(N)$.

REFERENCES

[1] BLACKWELL, D. (1955). The prediction of sequences. The RAND Corporation, Research Memorandum RM 1570.

[2] DUBINS, L. (1957). A discrete evasion game. Contributions to the Theory of Games III, *Annals of Mathematics Studies* No. 39, 231–255.

[3] FELLER, W. (1968). *An Introduction to Probability Theory and its Applications*, **1** 3rd ed. Wiley, New York.

[4] FERGUSON, T. (1966). On discrete evasion games with a two-move information lag. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* **1** 453–462.

[5] MYERS, L. (1972). An N-move prediction game. Ph. D. Dissertation, Univ. of California, Berkeley.

DEPARTMENT OF MATHEMATICS
DUKE UNIVERSITY
DURHAM, NORTH CAROLINA