1989

Naive Asymptotics for Hitting Time Bounds in Markov Chains

Vernon J. Rego

Purdue University, rego@cs.purdue.edu

Report Number:

89-896

NAIVE ASYMPTOTICS FOR HITTING TIME
BOUNDS IN MARKOV CHAINS

Vernon Rego

CSD-TR-896
August 1989
(Revised April 1992)

NAIVE ASYMPTOTICS FOR HITTING TIME BOUNDS IN MARKOV CHAINS*

Vernon Rego †

Department of Computer Sciences
Purdue University
West Lafayette, IN 47906

Abstract

A set of sufficient conditions is obtained for Markov chains to yield upper and lower passage time bounds. While obtaining expected passage times is strictly a numerical procedure for general Markov chains, the results presented here outline a simple approach to bound expected passage times provided the chains satisfy certain easy to check criteria. The results may be useful in modelling situations, such as the analysis of algorithms, where simple ways of obtaining average complexity estimates are required.

* Research performed at the Mathematical Sciences Section of Oak Ridge National Laboratory under the auspices of the Faculty Research Participation Program of Oak Ridge Associated Universities, and supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. DOE, under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

† Also supported in part by the National Science Foundation, under award ASC-9002225 and NATO award CRG 900108.

1. INTRODUCTION

Hitting times in Markov chains are known to be useful in a variety of modelling contexts such as the analysis of algorithms, estimation of system reliability, combinatorial algorithms, queuing applications, etc. For example, if an algorithm's execution time can be modelled as a hitting time in a Markov chain, then average complexity bounds for the algorithm can readily be had by obtaining upper and lower bounds on the expected hitting time to some appropriate termination state. In this note we demonstrate the existence of a simple set of sufficient conditions which yield naive logarithmic type asymptotic bounds for hitting times in discrete-time Markov chains.

Let $\{Y_k ; k \geq 0\}$ be a discrete-time, homogeneous Markov chain with a finite or countable state space $\mathbf{S}$ and transition probability matrix $\mathbf{P}(i, j)$. For a subset of states $A \subset \mathbf{S}$, the first hitting time $T_n(A)$ to set A from some initial state $Y_0 = n \in \mathbf{S}$ is defined as

$$T_n(A) = \min \{ j \geq 0 ; Y_0 = n, Y_j \in A \}. \tag{1.1}$$

Let $P_n(\cdot)$ and $E_n(\cdot)$ denote probability and expectation, respectively, conditional on $Y_0 = n$. We take $P_n(T(A))$ and $E_n(T(A))$ to be notationally equivalent to $P(T_n(A))$ and $E(T_n(A))$, respectively. Assuming a positive-recurrent chain means that $E_n(T(A)) < \infty$, and the expected hitting times to set A from arbitrary initial states can be determined by solving the system

$$E_i(T(A)) = \begin{cases} 1 + \sum_j P(i, j) E_j(T(A)) & i \notin A \\ 0 & i \in A \end{cases} \tag{1.2}$$

for fixed $A \subset \mathbf{S}$.

An alternative view of the random variable $T_n(A)$ is as the time to absorption in a Markov chain. If the states in A are all lumped together into a single absorbing state, then $E_n(T(A))$ becomes the expected time to absorption for the chain, given the initial state n. In the sequel it is shown that under certain conditions, simple bounds for this expected time can be obtained without solving the system in (1.2). The results presented in Section **2** are summarized as follows. Let $\mathbf{S} = \{0,1,2, \ldots\}$, where state 0 is an absorbing state and all other states are transient. We assume that $\mathbf{S}$ is decomposed and ordered into mutually disjoint sets $S_0 = \{0\}$, S_1, $S_2, \ldots$, etc., and define

$$T_n = \min \{ j ; Y_0 \in S_n, Y_j \in S_0 \} \qquad \forall S_n \subset \mathbf{S}. \tag{1.3}$$

That is, T_n is the time required for the chain to reach state 0, given that it starts in some arbitrary state in set S_n. The sets $S_1, S_2, \ldots,$ etc. are defined in a specific way, as will be seen in the results that follow. Note that even though the time to absorption in a Markov chain depends on the specific initial state Y_0 that one chooses, the set containing the initial state is more relevant to our discussion of absorption times than the specific initial state itself. This is because our interest lies in obtaining upper and lower bound on expected hitting-times, where such bounds will hold for all states in a given subset.

Our objective is to provide a set of sufficient conditions which enable us to obtain bounds of the form

$$a \log_\nu^m n \;\le\; E(T_n^m) \;\le\; b \log_\beta^m n \tag{1.4}$$

where $E[T_n^m]$ is the mth moment of T_n. The quantities a, b are constants, and β, ν are parameters obtained from the transition probability matrix of the chain. More generally, it can also be shown that for a class of functions $\mathcal{F}$,

$$f(a \log_\nu n) \;\le\; E(f(T_n)) \;\le\; f(b \log_\beta n) \tag{1.5}$$

for $f \in \mathcal{F}$.

2. MAIN RESULTS

Let $\{Y_k\,; k \ge 0\}$ be a discrete-time, Markov chain with stationary transitions on a finite or countable state-space S. We assume that S is decomposed into mutually disjoint sets $S_0 = \{0\}$, $S_1, S_2, \ldots,$ etc. Define $\{X_k\,; k \ge 0\}$ to be a stochastic process associated with the chain $\{Y_k\,; k \ge 0\}$ through the relation

$$X_k \stackrel{\Delta}{=} j \cdot 1\{Y_k \in S_j\}, \;\; k \ge 0, \tag{2.1}$$

and observe that the hitting-time T_n can alternatively be defined as

$$T_n = \min\{\, j > 0\,;\, X_0 = n,\, X_j = 0\} \tag{2.2}$$

for all $n \ge 0$. At any step, the process $\{X_k\}$ takes on the value j if the Markov chain $\{Y_k\}$ is in the set S_j, for $j \ge 0$. In the sequel, the sets $S_1, S_2, \ldots,$ etc. are implicitly defined via conditional expectations for the process $\{X_k\}$. It turns out that it is easier to work with the associated process $\{X_k\}$ and its state-space, which is either S or a subset of S, than it is with $\{Y_k\}$ and the subsets S_n, $n \ge 1$.

The subsets S_0, S_1, S_2, ..., etc. which partition S are defined in the following manner. Define

$$N_i = \{j \in S \mid E(Y_{n+1} \mid Y_n = j) < i, \forall n > 0\} \tag{2.3}$$

be the set of states in S for which the expected next state is some state less than i. Observing that $N_i \subset N_j$ for $i \leq j$, define

$$S_i = N_i - N_{i-1} \tag{2.4}$$

for all $i \geq 2$, with $S_0 = \{0\}$ and $S_1 = N_1$. If the chain $\{Y_n\}$ is a stable M/GI/1 queueing chain [1] describing the number of customers queued in the system at departure epochs, then the above scheme for defining the subsets would yield $S_k = \{k\}$ for all $k \geq 1$. Here, T_n describes the amount of time required for the queue to become empty, given that n customers are found queued initially, at a departure epoch. In this case the associated process $\{X_k\}$ is a Markov chain identical to the chain $\{Y_k\}$.

On the other hand, if $\{Y_n\}$ takes uniformly random trajectories (i.e., has a doubly stochastic transition matrix) on a finite state space $S' = \{0, 1, 2, ... , n\}$, then the above scheme yields $S_k = \Phi$ for $1 \leq k < \lfloor (n+1)/2 \rfloor$, and $S_{\lfloor (n+1)/2 \rfloor} = S'$. In this case, while $\{Y_k\}$ jumps uniformly randomly over states in S', the associated process $\{X_k\}$ makes transitions on the space $\{0, \lfloor (n+1)/2 \rfloor\}$, with state $\lfloor (n+1)/2 \rfloor$ being the initial state, and state 0 the absorbing state. Here, $T_{\lfloor (n+1)/2 \rfloor}$ describes the amount of time required for $\{Y_k\}$ to hit state 0, given that it starts in some state in S'. The specific initial state is not important because all the states have been grouped together and relabelled as state $\lfloor (n+1)/2 \rfloor$ for the $\{X_k\}$ process, by virtue of the drift condition in (2.3). In both examples, a bound on the hitting time to state 0 for the $\{X_k\}$ process would give some useful information on the behaviour of the underlying Markov chain $\{Y_k\}$.

2.1 Upper Bounds

The first result, presented as a theorem, is one that motivates the results which succeed it in the text. In essence, it establishes sufficient conditions under which $E(T_n) \in O(\log n)$. Note that a function g is said [3] to be in $O(f)$ if, for some positive constant c, $c\,g(n)$ is at most $f(n)$ for all n greater than some integer n_0. Likewise, if $g(n)$ is at least $c' f(n)$ for all n greater than some integer n_0' and some positive constant c', then g is said to be in $\Omega(f)$. If g belongs to both $O(f)$ as well as $\Omega(f)$, then g is said to be in $\Theta(f)$.

The statement of the following theorem can be found in [5], where it was used to determine a bound on the expected time to absorption in a certain random field. Stavskaya and Pyatetskii-Shapiro [5] attribute the result to L. G. Mityushin, but however, do not provide a proof. The proof given in [4] is briefly recounted here for completeness. In words, if the expected next state of the chain $\{Y_k\}$ always lies in a subset whose index is less than the index of the set containing the current state, then a simple upper bound can be obtained for the hitting-time to state 0.

Theorem 1 (Mityushin's Theorem).

$$\text{If} \qquad E(X_1 \mid X_0 = n) < \frac{n}{\beta} \qquad \forall n \geq 1 \tag{2.5}$$

and some $\beta > 1$, then

$$E(T_n) < \lceil \log_\beta n \rceil + \frac{\beta}{\beta - 1}.$$

Proof:

Conditional on $X_0 = n$, the requirement in (2.5) establishes that

$$E_n(X_j) < E_n(X_{j-1}) \cdot \frac{1}{\beta} < \frac{n}{\beta^j} \tag{2.6}$$

for all $j \geq 1$. Since $P_n(X_j \neq 0) \leq E_n(X_j)$,

$$P(T_n > j) \leq E_n(X_j) < \frac{n}{\beta^j} \wedge 1 \tag{2.7}$$

so that

$$E(T_n) < \sum_{j=0}^{\lceil \log_\beta n \rceil} 1 + \frac{1}{\beta - 1} = \lceil \log_\beta n \rceil + \frac{\beta}{\beta - 1}. \tag{2.8}$$

□

The linear term $\frac{\beta}{\beta - 1}$ obtained above is different from, and slightly smaller than, the linear term given in the original statement of the theorem as reported in [5].

Consider, for example, a simple but useful application of Theorem 1. Let $\{Z_k\,; k \geq 0\}$ be a Markov chain on a finite set $S' = \{0, 1, 2, \cdots, n\}$, with state 0 being the single absorbing state and all other states transient. Assume that the associated process $\{X_k\}$ is defined as in (2.1) for the chain $\{Z_k\}$, and that the hypothesis of Theorem 1 is satisfied. Using $\{Z_{k,1}\}$, $\{Z_{k,2}\}$, $\cdots$, $\{Z_{k,m}\}$ to denote m independent versions of the chain $\{Z_k\}$, each version $\{Z_{k,j}\}$ defines an associated process $\{X_{k,j}\}$, for $1 \le j \le m$. Define the hitting-time $T_{n,j}$ from some initial state in set S_n to state 0 for the j-th chain as

$$T_{n,j} = \min \{ k > 0 ; X_{0,j} = n, X_{k,j} = 0 \} \tag{2.9}$$

and let $T_n(m) = \max(T_{n,1}, T_{n,2}, \cdots, T_{n,m})$ denote the maximum of these hitting-times. We are interested in computing an upper bound for $E[T_n(m)]$, where $m \gg n$.

Let $U_1, U_2, ..., U_m$ be independent geometric random variables with parameter $p = 1-q$, i.e., $P(U_j = k) = (1-p)^{k-1}p$, where p is chosen in such a way that $U_j \ge_{st} T_{n,j}$, for $1 \le j \le m$. Note that p is a function of the transition probability matrix for chain $\{Z_k\}$ and the index n of the set containing the initial state. With $R_m \triangleq \max(U_1, \cdots, U_m)$, it follows that $E[T_n(m)] \le E[R_m]$, so that by obtaining a bound for the expected maximum of m i.i.d geometric random variables with parameter p we also obtain a bound for $E[T_n(m)]$. It can be shown [4] that the random variable R_m is the time to absorption in a Markov chain on space $\mathbf{S}'$, with a single absorbing state 0. The transition probability matrix of this chain is upper triangular, and defined by

$$\mathbf{P}(i,j) = \binom{i}{j} p^{i-j} q^j \tag{2.10}$$

for $1 \le i \le m$ and $0 \le j \le i$. Since the hypothesis of Theorem 1 is satisfied for this matrix, it follows that

$$E(R_m) < \frac{\lceil \log m \rceil}{\log q^{-1}} + \frac{1}{1-q} \tag{2.11}$$

so that $E[T_n(m)] \in O(\log m)$.

The simple arguments used in the proof of Theorem 1 can also be used to obtain more general forms of Mityushin's inequality, such as upper bounds on moments of T_n, etc. Though these results generalize Theorem 1, they are given below as corollaries to the theorem, since it is in the original theorem that the essence of the bounding idea lies, linking negative drifts to passage times.

For $m > 1$, let $m(k)$ denote the greatest integer less than or equal to the mth root of k, i.e., $m(k) = \lfloor k^{1/m} \rfloor$. Then,

Corollary 1.1.

$$\text{If} \quad E(X_1^m \mid X_0 = n) < \frac{n^m}{\gamma} \qquad \forall n \ge 1, \tag{2.12}$$

for some $\gamma > 1$, then there exists a constant c, $0 \le c < 1$, such that

$$E(T_n^m) \; < \; \lceil \log_\alpha^m n \rceil \; + \; \frac{\beta}{\beta - 1}$$

where $\beta = \gamma^{1/m}$, and $1 < \alpha = \beta^{\lceil \log_\beta n / (\log_\beta n + c) \rceil} \le \beta$.

Proof:

Let the initial state be given by $X_0 = n$. The inequality in (2.12) yields the recursion

$$E_n(X_{k+1}^m) \; < \; E_n(X_k^m) \cdot \frac{1}{\gamma} \tag{2.13}$$

for fixed $m > 1$. It follows that

$$E_n(X_k^m) \; < \; \frac{n^m}{\gamma^k} \tag{2.14}$$

and consequently

$$E_n(X_{m(k)}^m) \; < \; \frac{n^m}{\gamma^{m(k)}} \; = \; \frac{n^m}{\beta^{m(k) \cdot m}} \tag{2.15}$$

for all $k \ge 1$. Since $P_n(X_k \ne 0) = P_n(X_k^m \ne 0)$ for each k,

$$P(T_n > m(k)) \; \le \; P_n(X_{m(k)}^m) \; \le \; E_n(X_{m(k)}^m). \tag{2.16}$$

Finally, using the relations

$$P(T_n^m > k) \; \le \; P(T_n > \lfloor k^{1/m} \rfloor) \tag{2.17}$$

and

$$E(T_n^m) \; = \; \sum_{k=0}^{\infty} P(T_n^m > k) \tag{2.18}$$

it follows that

$$E(T_n^m) \; < \; \sum_{k=0}^{\infty} (\frac{n^m}{\beta^{(m(k) \cdot m}} \; \wedge \; 1). \tag{2.19}$$

Setting $\frac{n^m}{\beta^{m(k) \cdot m}} = 1$ yields

$$k^{1/m} \; = \; \log_\beta n \; + \; c, \quad \text{for} \quad 0 \le c < 1. \tag{2.20}$$

Defining $c = \log_\alpha n - \log_\beta n$, for $1 < \alpha \le \beta$, gives

$$\log_\alpha \beta = 1 + \frac{c}{\log_\beta n}, \tag{2.21}$$

and thus $\alpha = \beta^{[\log_\beta n/(\log_\beta n + c)]}$. With $k = \lceil \log_\alpha^m n \rceil$ from (2.20), the inequality in (2.19) reduces to

$$E(T_n^m) < \sum_0^{\lceil \log_\alpha^m n \rceil} 1 + \frac{1}{\beta - 1} = \lceil \log_\alpha^m n \rceil + \frac{\beta}{\beta - 1} \tag{2.22}$$

for $\beta \ge \alpha > 1$.

□

The above result can be shown to hold for any invertible function $f(\cdot)$ of T_n satisfying $f(k) \ge 1$ for all $k \ge 1$, and $f(0) = 0$. Using $\mathrm{f}(k)$ to denote $\lfloor f(k) \rfloor$ for $k \ge 1$, and $h \equiv f^{-1}$,

Corollary 1.2.

$$\text{If} \quad E[f(X_1) | X_0 = n)] < \frac{g(n)}{\beta} \qquad \forall n \ge 1, \tag{2.23}$$

some $\beta > 1$, and an arbitrary function $g(\cdot)$, then there exists a constant c, $0 \le c < 1$ such that

$$E[f(T_n)] < \lceil f(\log_\alpha g(n)) \rceil + \frac{\beta}{\beta - 1}$$

where $1 < \alpha = \beta^{[\log_\beta g(n)/(\log_\beta g(n) + c)]} \le \beta$, $f(k) \ge 1 \;\; \forall\, k \ge 1$, and $f(0) = 0$.

Proof:

Repeating inequalities (2.13) through (2.15) with $f(X_k)$ and $g(n)$ in place of X_k^m and n^m, respectively, one obtains

$$P(T_n > h(k)) \le P(T_n > \mathrm{h}(k)) < \frac{g(n)}{\beta^{\mathrm{h}(k)}} \wedge 1, \tag{2.24}$$

for all $k \ge 1$. The last inequality in (2.24) uses the fact that $f(k) \ge 1$ for all $k \ge 1$. Since $h \equiv f^{-1}$ and

$$h(k) = \log_\beta g(n) + c, \qquad 0 \le c < 1, \tag{2.25}$$

if the rightmost term in (2.24) must equal unity for all n above some value, it follows that

$$E(f(T_n) > k) \;<\; \sum_{0}^{\lceil f(\log_\alpha g(n)) \rceil} 1 \;+\; \frac{1}{\beta - 1} \;=\; \lceil f(\log_\alpha g(n)) \rceil \;+\; \frac{\beta}{\beta - 1} \tag{2.26}$$

where $1 < \alpha = \beta^{\lceil \log_\beta g(n) / (\log_\beta g(n) + c) \rceil} \leq \beta$.

□

Remark

As a special case of the above result, let $g(\cdot)$ be an arbitrary but bounded function, and $f(k) = k$ for $k \geq 0$. If $E[f(X_1) | X_0] < \frac{g(n)}{\beta}$ for all $n \geq 1$ and some $\beta > 1$, then it follows that

$$E(T_n) \;<\; \lceil \log_\beta g(n) \rceil \;+\; \frac{\beta}{\beta - 1} \tag{2.27}$$

□

It should be clear that if the negative drift condition for the process $\{X_k\}$ (or equivalently, $\{Y_k\}$) was violated for all initial states n, then $E(T_n)$ would be infinite. However, what should one expect if the negative drift condition was violated for only a finite number of states **H**? Intuitively, one would expect that a finite set **H** should not effect the form of the upper bound, and indeed, it can be shown that the logarithmic upper bound can still be had in such a situation. This result is given as a theorem below, and it should be recognized that the corollaries to Theorem 1 can be generalized via the following theorem.

Let **H** be a subset of **S** satisfying

$$\frac{j}{\beta} \;<\; E(Y_1 | Y_0 = j) \;<\; g(j) \qquad \forall j \in \mathbf{H} \tag{2.28}$$

where $\beta > 1$ and $g(\cdot)$ is a bounded function. Using the transition probability matrix **P**, define

$$p_i \;=\; \sum_{j \in \mathbf{H}} \mathbf{P}(i, j) \qquad \forall i \in \mathbf{S} \tag{2.29}$$

and set $p = \max_{i \in \mathbf{S}} p_i$. That is, the quantity p describes the maximum probability that the process $\{X_k\}$ will next visit a state in **H**, given that it is currently in any state in **S**. For convenience, assume (initially) that $0 < p < 1$.

Theorem 2.

$$\text{If} \qquad E(X_1 \mid X_0 = n) \; < \; \frac{n}{\beta} \qquad \forall n \in \mathbf{S} - \mathbf{H} \tag{2.30}$$

where **H** is a finite subset of **S**, then

$$E(T_n) \; < \; \lceil \log_\alpha g(n) \rceil \; + \; \frac{\alpha}{\alpha - 1}$$

where $1 < \alpha = \dfrac{\beta}{1 + p(\beta - 1)} < \beta$, and p is defined through (2.29).

Proof:

Conditional on $X_0 = n$, combining (2.28) and (2.30) gives

$$E(X_1 \mid X_0 = n) \; < \; p \, g(n) \; + \; (1-p) \frac{g(n)}{\beta} \tag{2.31}$$

whence the recursion

$$E(X_k) \; < \; \frac{g(n)}{\alpha^k} \qquad \forall k \geq 1 \tag{2.32}$$

is obtained, for $\alpha = \dfrac{\beta}{1 + p(\beta - 1)}$. Applying the steps shown in (2.6) through (2.8) yields

$$E(T_n) \; < \; \lceil \log_\alpha g(n) \rceil \; + \; \frac{\alpha}{\alpha - 1}. \tag{2.33}$$

□

Remarks:

1. The assumption that $0 < p < 1$ can be dropped. The case $p = 0$ is merely Theorem 1. If $p = 1$, then transitions from certain states in **S** take the process $\{X_k\}$ to a state in **H** with probability 1. However, since $\{Y_k\}$ is an absorbing chain with a finite expected time to absorption, the average number of steps required for $\{X_k\}$ to leave **H** once it enters this set is finite. That is, since **H** is finite, the absorbing chain can exhibit nonnegative drift for only a finite number of steps before leaving **H** to reach a state in **S**–**H** which yields negative drift. Denoting the smallest integer greater than or equal to this finite expectation by r, it follows that (2.31) generalizes to

$$E(X_{kr}) \; < \; \frac{g(n)}{\alpha^k} \qquad \forall k \geq 1 \tag{2.34}$$

with $\alpha = \beta$, so that

$$E(T_n) \ < \ r \lceil \log_\alpha g(n) \rceil \ + \ \frac{\alpha}{\alpha - 1}. \tag{2.35}$$

2. For $0 < p \leq 1$, Theorem 2 can be generalized to yield results similar to those shown in Corollaries 1.1 and 1.2. Thus the logarithmic upper bound remains, in spite of the negative drift violation over a finite subset of **S**.

3. The above result can also be used to generalize the example, shown below Theorem 1, involving the maximum of m i.i.d hitting-times.

□

2.2 Lower Bounds

In the case of lower bounds on expected times to absorption, similar logarithmic type bounds can be obtained, but only at the expense of sufficient conditions of a somewhat different form. Instead of working with drift, actual transition probabilities come into play. Consider first an analogue of Theorem 1 for a lower bound on time to absorption. Define $r_{n,k} = [P_n(X_k = 0)]^{1/n}$ for each $n > 0$ and $k \geq 1$. Then,

Theorem 3.

$$\text{If} \qquad \frac{r_{n,k}}{(1-\delta^k)} \leq 1, \qquad \forall k \geq 1, \ 0 < \delta < 1, \tag{2.36}$$

and given n, then there exist constants $\varepsilon > 0$, $0 < a < 1$, such that $\forall n > n_0$, n_0 an integer,

$$E(T_n) \ \geq \ \log_\nu n \tag{2.37}$$

where $\nu = \left(\frac{1}{\delta}\right)^{\frac{1}{(1-\varepsilon)(1-a)}} > 1$.

Proof:

If the transition matrix **P** is such that (2.36) is satisfied, then

$$r_{n,k} \ \leq \ (1 - \delta^k), \tag{2.38}$$

$$\text{implying that} \qquad P_n(X_k = 0) \ \leq \ (1 - \delta^k)^n, \qquad \forall k \geq 1. \tag{2.39}$$

Equivalently,

$$P(T_n > k) \ = \ P_n(X_k \neq 0) \ > \ 1-(1-\delta^k)^n \qquad \forall k > 1. \tag{2.40}$$

Observing that $r_{n,k}$ is nondecreasing in k,

$$E(T_n) \ > \ \sum_{k=0}^{\infty} [1-(1-\delta^k)^n] \ \geq \ k\,[\,1-(1-\delta^k)^n\,] \tag{2.41}$$

for all $k \geq 1$. Setting

$$k \ = \ \frac{(1-a)\log n}{\log \delta^{-1}}, \qquad 0 \ < a < 1. \tag{2.42}$$

yields

$$E(T_n) \ \geq \ (1-a)(1-\varepsilon)\,\frac{\log n}{\log \delta^{-1}} \tag{2.43}$$

for any $\varepsilon > 0$ and $n > n_0$ satisfying

$$(1 - n_0^{a-1})^{n_0} < \varepsilon. \tag{2.44}$$

Finally, defining $\nu = \left(\frac{1}{\delta}\right)^{\frac{1}{(1-\varepsilon)(1-a)}}$ puts (2.43) in the required form (2.37).

□

Granted that ascertaining (2.36) for a given Markov chain requires more work than verifying (2.5), it is worthwhile pointing out that one needs to check (2.36) only for state n, given that $X_0 = n$. This is done by taking powers $\mathbf{P}^k$ of the transition matrix to determine if $\mathbf{P}^k(n, 0)$ is bounded from above by $(1-\delta^k)^n$, for some appropriately chosen δ, $0 < \delta < 1$. Again, as was the case with the upper bound, a similar argument can be used to obtain lower bounds for higher order moments of T_n. In this case let $m[k]$ denote the smallest integer greater than or equal to the mth root of k, i.e., $m[k] = \lceil k^{1/m} \rceil$. For each $m \geq 1$ define $r_{n,k}^m = [P_n(X_k^m = 0)]^{1/n}$, where $n \geq 1$, and $k \geq 1$. Under the condition

$$\frac{r_{n,k}^m}{(1-\delta^k)} \leq 1 \quad \forall k \geq 1, \ 0 \ < \delta < 1, \tag{2.45}$$

which is an obvious extension of requirement (2.36), one obtains

$$P_n(X_{m[k]}^m \neq 0) \ \geq \ 1-(1-\delta^{m[k]})^n \tag{2.46}$$

$$\geq 1 - (1 - \delta^{k^{1/m}})^n.$$

Since $P(T_n > k^{1/m}) \geq P(T_n > m(k))$,

$$E(T_n^m) \geq k[1 - (1 - \delta^{k^{1/m}})^n] \qquad \forall k \geq 1. \tag{2.47}$$

Setting $k^{1/m} = \dfrac{(1-a)\log n}{\log \delta^{-1}}$ and repeating previously used arguments gives

$$E(T_n^m) \geq \log_\nu^m n \tag{2.48}$$

where $\nu = \left(\dfrac{1}{\delta}\right)^{\frac{1}{(1-\varepsilon)(1-a)^m}}$.

Corollary 3.1.

$$\text{If} \qquad \frac{r_{n,k}^m}{(1-\delta^k)} \leq 1 \quad \forall k \geq 1,\ 0 < \delta < 1,$$

then there exist constants $\varepsilon > 0$, $0 < a < 1$ such that $\forall n > n_0$, n_0 an integer,

$$E(T_n^m) \geq \log_\nu^m n$$

with $\nu = \left(\dfrac{1}{\delta}\right)^{\frac{1}{(1-\varepsilon)(1-a)^m}} > 1$.

□

In order to obtain a lower bound analogous to that provided by Corollary 1.2, let $f(\cdot)$ be any invertible function of T_n satisfying $f(0) = 0$. Using f(k) to denote $\lceil f(k) \rceil$ here, for $k \geq 1$, and $h \equiv f^{-1}$,

Corollary 3.2.

$$\text{If} \quad [P(f(X_k) = 0 \mid X_0 = n)]^{1/n} \leq (1 - \delta^k) \qquad \forall k \geq 1, 0 < \delta < 1, \tag{2.49}$$

then there exist constants $\varepsilon > 0$, $0 < a < 1$ such that $\forall n > n_0$, n_0 an integer,

$$E(f(T_n)) \geq f(\log_\nu n) \tag{2.50}$$

with $\nu = \left(\dfrac{1}{\delta}\right)^{\frac{1}{(1-\varepsilon)(1-a)}} > 1$.

Proof:

Arguing along the same lines as before,

$$P_n(X_{\mathbf{h}(k)} \neq 0) \;\geq\; 1-(1-\delta^{\mathbf{h}(k)})^n$$
$$\geq\; 1-(1-\delta^{h(k)})^n. \qquad (2.51)$$

Since $P(T_n > h(k)) \geq P(T_n > \mathbf{h}(k))$,

$$E(f(T_n)) \;\geq\; k[1-(1-\delta^{h(k)})^n], \qquad \forall k \geq 1. \qquad (2.52)$$

Setting $h(k) = (1-a)\log \dfrac{n}{\log \delta^{-1}}$ and following steps (2.42) through (2.44) gives

$$E(f(T_n)) \;\geq\; f(\log_\nu n) \qquad (2.53)$$

with $\nu = \left[\dfrac{1}{\delta}\right]^{\frac{1}{(1-\varepsilon)(1-a)}} > 1.$

□

A lower bound analogous to the upper bound given in the example following Theorem 1 is easily had. Let $\{X_{k,1}\}$, $\{X_{k,2}\}$, $\cdots$, $\{X_{k,m}\}$ be independent processes, with the same law as the process $\{X_k\}$, and define $X_k(m) \stackrel{\Delta}{=} \max(X_{k,1}, \cdots, X_{k,m})$, for $k \geq 1$. Note that $T_{n,j}$ is defined as in (2.9) for $1 \leq j \leq m$, and $T_n(m) = \max(T_{n,1}, T_{n,2}, \ldots, T_{n,m})$ is the maximum hitting-time to the absorbing state over all m independent processes.

Corollary 3.3.

If $\quad [\,P(X_k(m)=0 \mid X_0(m)=n)]^{1/n} \;\leq\; (1-\delta^k)^m \qquad \forall k \geq 1,\; 0 < \delta < 1 \quad (2.54)$

then there exist constants $\varepsilon > 0$, $0 < a < 1$, such that $\forall m > \dfrac{n_0}{n}$, n_0 an integer.

$$E(T_n(m)) \;\geq\; \log_\nu mn \qquad (2.55)$$

with $\nu = \left[\dfrac{1}{\delta}\right]^{\frac{1}{(1-\varepsilon)(1-a)}} > 1.$

Proof:

The inequality (2.54) implies that

$$P_n(X_k(m)=0) \;\le\; (1-\delta_k)^{mn} \qquad \forall k \ge 1. \tag{2.56}$$

Using the same arguments as before,

$$E(T_n(m)) \;\ge\; k\,[1-(1-\delta^k)]^{mn} \qquad \forall k \ge 1, \tag{2.57}$$

and setting $k = \dfrac{(1-a)\log nm}{\log \delta^{-1}}$, one obtains, for $\nu = \left(\dfrac{1}{\delta}\right)^{\frac{1}{(1-\varepsilon)(1-a)}}$

$$E(T_n(m)) \;\ge\; \log_\nu nm \tag{2.58}$$

for $\varepsilon > 0$ and $m > \dfrac{n_0}{n}$ satisfying

$$(1 - n_0^{a-1})^{n_0/n} < \varepsilon \tag{2.59}$$

□

3. APPLICATIONS

a. Geometric random variables

Let $\{Y_k\}$ be a Markov chain on the space $\{0,1\}$, with transitions defined by $P(Y_{k+1}=1 \mid Y_k=1)=q=1-p>0$, and $P(Y_{k+1}=0 \mid Y_k=0)=1$ for $k\ge 0$. With state 1 as transient and state 0 as absorbing, $S_0=\{0\}$, and $S_1=\{1\}$. Define the process $\{X_k\}$ as in (2.1), and note that T_1 is a geometric random variable with parameter p. Let $T_{1,1}, T_{1,2}, \ldots, T_{1,m}$ be m i.i.d random variables with the same law as T_1, and let $T_n(m)$ denote the maximum of these m random variables.

As indicated in the example following Theorem 1, the random variable $T_n(m)$ is the time to absorption in an $(m+1)$-state Markov chain with absorbing state 0 and transition matrix given by (2.10). Using Corollary 3.1, or Corollary 3.3 (in which case $n=1$), it follows that

$$E(T_n(m)) \;\in\; \Omega(\log_\nu m), \tag{3.1}$$

and from Corollary 3.1,

$$E(T_n^k(m)) \;\in\; \Omega(\log_\nu^k m) \tag{3.2}$$

$\forall k \geq 1$. Applying Corollary 1.1 to the same Markov chain,

$$E(T_n^k(m)) \in O(\log_\alpha^k n) \tag{3.3}$$

$\forall k \geq 1$, so that

$$E(T_n^k(m)) \in \Theta(\log^k m). \tag{3.4}$$

b. Coupon collector's problem

Consider the following simple sequential occupancy problem known as the coupon collector's problem [2]. Assume an urn contains N balls, labelled as 1 through N. If one ball is randomly drawn at a time, with replacement, let T_N denote the number of draws required until each ball has been drawn at least once. How does $E(T_N^m)$ behave for large N?

Defining the sets $S_k = \{k\}$ for $k \geq 0$, it follows that the process $\{X_k\}$ defined via (2.1) is also a Markov chain, identical to the process $\{Y_k\}$. The random variable T_N is the hitting-time to the absorbing state 0 in the (N+1)-state Markov chain $\{X_k\}$. The transition probability matrix **P** for this chain is given by

$$P(X_{k+1} = j \mid X_k = i) = \begin{cases} \dfrac{N-i}{N} & j = i \\ \dfrac{i}{N} & j = i-1 \\ 0 & \text{otherwise} \end{cases} \tag{3.5}$$

for $1 \leq i \leq N-1$ and $k \geq 1$. Here X_k describes the number of balls that have not been sampled at the end of the kth draw. Given $X_0 = N$, the first draw must sample a ball, so that $P(X_1 = N-1 \mid X_0 = N) = 1$.

By Theorem 2, we may ignore the (negative) drift condition satisfied by row N without affecting the logarithmic form of the upper bound. Since,

$$E(X_1^m \mid X_0 = j) < \frac{j^m}{\beta} \tag{3.6}$$

for $1 \leq j \leq N-1$ and $1 < \beta \leq \frac{N}{N-1}$, Corollary 1.1 yields,

$$E(T_N^m) < \log_\alpha^m N + \frac{\beta}{\beta-1} = \log^m N + c\,\frac{\log^m N}{\log_\beta N} + \frac{\beta}{\beta-1}, \tag{3.7}$$

where $\alpha = \beta^{[\log_\beta N/(\log_\beta N + c)]}$, $0 \le c < 1$. Finally, since $1 < \beta \le \frac{N}{N-1}$ and $\log_\mu(\frac{N}{N-1}) \ge \frac{1}{N}$, for $1 < \mu < \left[\frac{N}{N-1}\right]^N$,

$$E(T_N^m) \; < \; c\, N \log_\mu^m N \; + \; O(N), \tag{3.8}$$

so that $E(T_N^m) \;\in\; O(N \log^m N)$.

For the lower bound, using $\delta = (N-1)/N$ and Corollary 3.1,

$$E(T_N^m) \;\ge\; \log_\nu^m N \tag{3.9}$$

for N sufficiently large. With the relation

$$\frac{1}{\log \nu} \;=\; \delta(1-\varepsilon)(1-a) \;=\; (1-\varepsilon)(1-a)\left[\frac{N-1}{N}\right] \tag{3.10}$$

for $0 < \varepsilon < 1$ and $0 < a < 1$,

$$E(T_N^m) \;\ge\; c' N \log^m N \tag{3.11}$$

for some constant c', so that $E(T_N^m) \;\in\; \Omega(N \log^m N)$ and thus

$$E(T_N^m) \;\in\; \Theta(\log^m N). \tag{3.12}$$

Consider now a simple variant of the coupon collector's problem. Suppose that at each step k, $k > 1$, if the state (i.e., number of balls not yet sampled) of the chain $\{X_n\}$ is neither 0 nor N, then there is a nonzero probability that the next state is N. In other words, if the number of balls sampled before the kth draw is neither 0 nor N, then there is a nonzero probability that the state at step k is "forgotten", so that the process must start all over again from state N at step $k+1$, for $k > 1$.

Given that the chain is in state i before draw k, $1 \le i \le N-1$, let η_i be the probability that on draw k the state of the chain is forgotten and it lands back in state N, for $k > 1$. The transition probability matrix for $\{X_n\}$ becomes

$$P(X_{k+1} = j \,|\, X_k = i) = \begin{cases} \frac{N-i}{N} - \eta_i & j = i \\ \frac{i}{N} & j = i-1 \\ \eta_i & j = N \\ 0 & \text{otherwise} \end{cases} \tag{3.13}$$

for $0 < i < N$, and $k > 1$. If $X_k = N$ for any k, then $P(X_{k+1} = N-1) = 1$, as before. What can we expect of $E(T_N)$ for this memory-loss type coupon collector's problem ? By placing the restriction that

$$N\eta_i + [\frac{N-i}{N} - \eta_i]\, i + \frac{i}{N}(i-1) \;<\; \frac{i}{\beta} \tag{3.14}$$

for $0 < i < N$, one obtains a condition on η_i that preserves (2.5), ensuring $E(T_N) \in O(\log n)$. At an extreme, if $\eta_i = \eta$ for $0 < i < N$, then (3.14) reduces to the requirement

$$(N-1)(\eta + \frac{1}{N}) \;<\; \frac{1}{\beta} \tag{3.15}$$

where we use the fact that β is defined by the entries in row 1 of $\mathbf{P}$, since this row exhibits the smallest negative drift in this case. For the lower bound, one need only note the fact that the "forgetfulness" of the process must increase the expected time to absorption, so that $E(T_N) \in \Omega(\log n)$, as before.

c. A single server queue

Consider a single server queue where customers arrive according to a Poisson process of rate $\lambda > 0$. Service times are assumed to be independent with some common distribution function $F(\cdot)$. The process $\{Y_k ; k \geq 0\}$ describing queue size at service completion epochs is a well-known [1] Markov chain, with transition matrix $\mathbf{P} = [P_{i,j}]$,

$$P_{i,j} = \begin{cases} \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^{j-i+1}}{(j-i+1)!}\, dF(t) & (j \geq i-1,\ i \geq 1) \\ 0 & (j < i-1,\ i \geq 1) \end{cases} \tag{3.16}$$

Defining $S_k = \{k\}$ for $k \geq 0$, as done earlier, it follows that the process $\{X_k\}$ is stochastically identical to the process $\{Y_k\}$. Hence $\{X_k\}$ is a Markov chain with transition probability matrix $\mathbf{P}$ defined above. Let T_n denote the time it takes the server to empty the queue, given that the queue initially contains n customers. The random variable T_n is a passage time which has a complicated distribution in general. For a stable queue, the chain $\{X_n\}$ will exhibit negative drift, so that the upper bound results of Section 2 are applicable. That is,

$$E[X_1 \mid X_0 = n] \;=\; n - \rho \qquad \forall\; n \geq 1 \tag{3.17}$$

for $\rho = \sum_{j=0}^{\infty} P_{0,j}$, with $0 < \rho < 1$. Choosing β, $1 < \beta < \frac{n}{n-\rho}$, the requirement in (2.5) is satisfied, and

$$E(T_n) < n \log_\mu n + \frac{n}{\rho} \tag{3.18}$$

for $1 < \mu < \left[\frac{n}{n-\rho} \right]^n$. Using $\delta = 1 - k_0$ in (2.36),

$$E(T_n) \geq \log_\nu n \tag{3.19}$$

where ν is defined in Theorem 2. Asymptotic bounds for higher order moments can also be had from the results of Section 2.

Acknowledgements

I am grateful to Professor Prem Puri for the many helpful, pleasant and memorable discussions on these and various aspects of probability. His untimely death in August 1989 deprived us of a most valued probabilist, colleague and a dear friend. I am also indebted to Professor Burgess Davis of Purdue and Dr. Uppuluri Rao of ORNL for their most helpful comments.

References

[1] Gross, D., and Harris, C. M., *Fundamentals of Queueing Theory*, John Wiley & Sons Inc., N. Y., 1985.

[2] Johnson, N. L., and Kotz, S., *Urn Models and Their Application*, John Wiley & Sons Inc., N. Y., 1977.

[3] Knuth, D. E., "Big omicron and big omega and big theta," SIGACT News, vol. 8, no.2, pp. 18-24, April-June, 1976.

[4] Rego, V., "A Band and Bound Technique for Simple Random Algorithms," *Probability in the Engineering and Informational Sciences*, Vol. **4**, pp. 333–344, 1990.

[5] Stavskaya, O. N., and Pyatetskii-Shapiro, I. I., "On certain properties of homogeneous nets of spontaneously active elements," *Problemi Cibernetiki*, 20, M. Nauka, pp. 91–106, 1968.