Steven R. Dunbar
Department of Mathematics
203 Avery Hall
University of Nebraska-Lincoln
Lincoln, NE 68588-0130
http://www.math.unl.edu
Voice: 402-472-3731
Fax: 402-472-8466

Topics in Probability Theory and Stochastic Processes
Steven R. Dunbar

Waiting Time to Absorption

Rating

Mathematically Mature: may contain mathematics beyond calculus with proofs.

Section Starter Question

For a Markov chain with an absorbing state, describe the random variable for the time until the chain gets absorbed.

Key Concepts

1. Let $\{X_n\}$ be a finite-state absorbing Markov chain with a absorbing states and t transient states. Let the $(a+t) \times (a+t)$ transition probability matrix be P. Order the states so the absorbing states come first and non-absorbing, i.e. transient, states come last. Then the transition probability matrix has the block-matrix form

$$P = \begin{pmatrix} I_a & 0 \\ A & T \end{pmatrix}.$$

   Here I_a is an $a \times a$ identity matrix, A is the $t \times a$ matrix of single-step transition probabilities from the t transient states to the a absorbing states, T is a $t \times t$ submatrix of single-step transition probabilities among the transient states, and 0 is a $a \times t$ matrix of 0s representing the single-step transition probabilities from absorbing states to transient states.

2. The matrix $N = (I-T)^{-1}$ is the **fundamental matrix** for the absorbing Markov chain. The entries N_{ij} of this matrix have a probabilistic interpretation. The entries N_{ij} are the expected number of times that the chain started from state i will be in state j before ultimate absorption.

3. First-step analysis gives a compact expression in vector-matrix form for the waiting time $\mathbf{w}$ to absorption:

$$(I - T)\mathbf{w} = \mathbf{1}$$

   so $\mathbf{w} = (I - T)^{-1}\mathbf{1}$.

Vocabulary

1. Let $\{X_n\}$ be a finite-state absorbing Markov chain with a absorbing states and t transient states. Let the $(a+t) \times (a+t)$ transition probability matrix be P. Order the states so the absorbing states come first and non-absorbing, i.e. transient, states come last. Then $(P^n)_{ij} \to 0$ as $n \to \infty$ for i and j in the transient states, while for i in the absorbing states, $P_{ii} = 1$. Define the **absorption time** as the random variable

$$T = \min \{n \geq 0 : X_n \leq a\}.$$

2. The **absorption probability matrix** B is the probability of starting at state i and ending at absorbing state j.

3. For a Markov chain with a absorbing states and t transient states, if necessary, reorder the states so the absorbing states come first and non-absorbing, i.e. transient, states come last. Then the transition matrix has the canonical form:

$$P = \begin{pmatrix} I_a & 0 \\ A & T \end{pmatrix}.$$

The matrix $N = (I - T)^{-1}$ is the **fundamental matrix** for the absorbing Markov chain.

4. The Nth **harmonic number** H_N is

$$H_N = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{N}.$$

Mathematical Ideas

Theory

Let $\{X_n\}$ be a finite-state absorbing Markov chain with a absorbing states and t transient states. Let the $(a+t) \times (a+t)$ transition probability matrix be P. Order the states so the absorbing states come first and non-absorbing, i.e. transient, states come last. The states $a+1, 1, 2, \dots, a+t$ are transient in that $(P^n)_{ij} \to 0$ as $n \to \infty$ for $a+1 \le i, j \le a+t$, while states $1, \dots, a$ are absorbing, $P_{ii} = 1$ for $1 \le i \le a$. Then the transition probability matrix has the block-matrix form

$$P = \begin{pmatrix} I_a & 0 \\ A & T \end{pmatrix}.$$

Here I_a is an $a \times a$ identity matrix, A is the $t \times a$ matrix of single-step transition probabilities from the t transient states to the a absorbing states, T is a $t \times t$ submatrix of single-step transition probabilities among the transient states, and 0 is a $a \times t$ matrix of 0s representing the single-step transition probabilities from absorbing states to transient states.

Starting at one of the transient states i where $a+1 \le i \le a+t$ such a process will remain in the transient states for some duration. Ultimately, the process gets trapped in one of the absorbing states $i = 1, \dots a$. Before the Markov chain transitions to one of the absorbing states, the number of times it visits a transient state is a random variable. Let Y_{ij} denote the number of visits the system makes to transient state j before reaching an absorbing state, given the system started in transient state i. Thus, Y_{ij} is a discrete random variable that can take on any nonnegative integer value. The random variables Y_{ij} are the fundamental random variables of interest here. These fundamental random variables are the building blocks for constructing and investigating other random variables. The mean, variance and covariances of the Y_{ij} are the first statistics to investigate. Of special interest is the mean time until absorption. Define the **absorption time**

$$w_i = \min\left\{n \ge 0 : X_n \le a \mid X_0 = i\right\}.$$

Notice that $w_i = \sum_{k=a+1}^{a+t} Y_{ik}$, the total number of visits the process makes among the transient states. The expected value of this random time $\mathbb{E}\left[w_i\right]$, $i = a+1, \dots, a+t$ is a first measure of the random variable w_i. Also of interest is the probability distribution of the states into which absorption takes place. Using the fundamental random variables, it is possible to compute this probability too.

Indicator Bernoulli Random Variables

Let the indicator random variables be

$$U_{ij}^{(m)} = \begin{cases} 1 & \text{if the Markov chain is in transient state } j \\ & \text{after } m \text{ steps given that it starts in transient state } i \\ 0 & \text{if the Markov chain is \textit{not} in transient state } j \\ & \text{after } m \text{ steps given that it starts in transient state } i \end{cases}$$

for $m = 0, 1, 2, \ldots$. The case $m = 0$ simply indicates where the system starts:

$$U_{ij}^{(0)} = \begin{cases} 1 & \text{if the Markov chain starts in transient state } i \\ 0 & \text{if the Markov chain does \textit{not} start in transient state } i. \end{cases}$$

Using the usual notation, $U_{ij}^{(0)}$ is the Kronecker delta function, δ_{ij}. The indicator random variables connect to the fundamental random variables Y_{ij} through the sum

$$Y_{ij} = \sum_{m=0}^{\infty} U_{ij}^{(m)}.$$

Expected number of visits between states

The expected number of visits to transient state j given that the Markov chain starts in transient state i in terms of the indicator random variable is

$$\mathbb{E}\left[Y_{ij}\right] = \mathbb{E}\left[\sum_{m=0}^{\infty} U_{ij}^{(m)}\right] = \sum_{m=0}^{\infty} \mathbb{E}\left[U_{ij}^{(m)}\right].$$

Use mathematical induction to show

$$P^m = \begin{pmatrix} I_a & 0 \\ A & T \end{pmatrix}^{m-1} \begin{pmatrix} I_a & 0 \\ A & T \end{pmatrix} = \begin{pmatrix} I_a & 0 \\ (I_t + T + T^2 + \cdots + T^{m-1})A & T^m \end{pmatrix}.$$

(See the exercises.) The elements $(P^m)_{ij}$ of P^m are the m-step transition probabilities between all states. Since $\mathbb{E}\left[U_{ij}^{(m)}\right] = p_{ij}^{(m)}$, so

$$\mathbb{E}\left[Y_{ij}\right] = \sum_{m=0}^{\infty} \mathbb{E}\left[U_{ij}^{(m)}\right] = \sum_{m=0}^{\infty} (P^m)_{ij}.$$

When i and j are transient states, then we only need to consider the entries in the transient corner matrix T^m, so

$$\mathbb{E}\left[Y_{ij}\right] = \sum_{m=0}^{\infty} (T^m)_{ij} = \left(\sum_{m=0}^{\infty} T^m \right)_{ij}.$$

The basic theory of finite absorbing Markov chains ensures that the induced 2-norm (or operator norm) of T is less than 1 nder typical conditions, that is $\|T\| < 1$. (See the exercises.) Therefore $\sum_{m=0}^{\infty} T^m$ converges. Furthermore, it converges to the **fundamental matrix** $N = (I-T)^{-1}$. Thus $\mathbb{E}\left[Y_{ij}\right] = N_{ij}$.

Waiting time to absorption

Theorem 1. *The entries N_{ij} of the fundamental matrix are the expected number of times that the chain started from state i will be in state j before ultimate absorption and the vector of expected waiting times to absorption is $\mathbb{E}\left[\mathbf{w}\right] = N = (I-T)^{-1}\mathbf{1}$.*

Proof. The sum over all states j of the number of times that the chain started from state i will be in state j before ultimate absorption is the waiting time to absorption. □

Covariances of numbers of visits between states

Now use the indicator Bernoulli random variables to derive $\operatorname{Cov}\left[Y_{ij}, Y_{ik}\right]$ where i, j, k are transient states. Recall that

$$\operatorname{Cov}\left[Y_{ij}, Y_{ik}\right] = \mathbb{E}\left[Y_{ij} \cdot Y_{ik}\right] - \mathbb{E}\left[Y_{ij}\right] \cdot \mathbb{E}\left[Y_{ik}\right]$$

and since $\mathbb{E}\left[Y_{ij}\right]$ and $\mathbb{E}\left[Y_{ik}\right]$ are known, all that is necessary is $\mathbb{E}\left[Y_{ij} \cdot Y_{ik}\right]$. Start with

$$Y_{ij} Y_{ik} = \left(\sum_{\nu=0}^{\infty} U_{ij}^{(\nu)} \right) \left(\sum_{\mu=0}^{\infty} U_{ik}^{(\mu)} \right) = \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} U_{ij}^{(\nu)} U_{ik}^{(\mu)}$$

so that

$$\mathbb{E}\left[Y_{ij} Y_{ik}\right] = \mathbb{E}\left[\sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} U_{ij}^{(\nu)} U_{ik}^{(\mu)} \right] = \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} \mathbb{E}\left[U_{ij}^{(\nu)} U_{ik}^{(\mu)} \right].$$

Rearrange the double sum into a first term summing over the lattice points above the line $\mu = \nu$, a second term summing over the lattice points along the line $\mu = \nu$, and a third term summing over lattice points below the line $\mu = \nu$,

$$\sum_{\nu=0}^{\infty} \sum_{\mu=\nu+1}^{\infty} \mathbb{E}\left[U_{ij}^{(\nu)} U_{ik}^{(\mu)}\right] + \sum_{\nu=0}^{\infty} \mathbb{E}\left[U_{ij}^{(\nu)} U_{ik}^{(\nu)}\right] + \sum_{\mu=0}^{\infty} \sum_{\nu=\mu+1}^{\infty} \mathbb{E}\left[U_{ik}^{(\mu)} U_{ij}^{(\nu)}\right].$$

The first and third terms are symmetric, so only evaluate the first term.

The expression $\mathbb{E}\left[U_{ij}^{(\nu)} U_{ij}^{(\mu)}\right]$ is the probability that the system is in transient state j after exactly ν steps from the start in state i *and* the system is in transient state k after exactly μ steps from the start in state i. Recall that $\nu < \mu$. Using the Markov chain property, this is $\mathbb{E}\left[U_{ij}^{(\nu)} U_{ik}^{(\mu)}\right] = p_{ij}^{(\nu)} p_{jk}^{(\mu-\nu)}$. Therefore, the first term is

$$\begin{aligned}
\sum_{\nu=0}^{\infty} \sum_{\mu=\nu+1}^{\infty} \mathbb{E}\left[U_{ij}^{(\nu)} U_{ik}^{(\mu)}\right] &= \sum_{\nu=0}^{\infty} \sum_{\mu=\nu+1}^{\infty} p_{ij}^{(\nu)} p_{jk}^{(\mu-\nu)} = \sum_{\nu=0}^{\infty} \sum_{\mu-\nu=1}^{\infty} p_{ij}^{(\nu)} p_{jk}^{(\mu-\nu)} \\
= \sum_{\nu=0}^{\infty} \sum_{z=1}^{\infty} p_{ij}^{(\nu)} p_{jk}^{(z)} &= \left(\sum_{\nu=0}^{\infty} p_{ij}^{(\nu)}\right) \left(\sum_{z=1}^{\infty} p_{jk}^{(z)}\right) = \left(\sum_{\nu=0}^{\infty} p_{ij}^{(\nu)}\right) \left(\sum_{z=0}^{\infty} p_{jk}^{(z)} - p_{jk}^{(0)}\right) \\
= \left(\sum_{\nu=0}^{\infty} (T^{\nu})_{ij}\right) &\left(\sum_{z=0}^{\infty} (T^{z})_{jk} - \delta_{jk}\right) = \left(\sum_{\nu=0}^{\infty} T^{\nu}\right)_{ij} \left(\sum_{z=0}^{\infty} T^{z}\right)_{jk} - \delta_{jk} \\
&= \left((I_t - T)^{-1}\right)_{ij} \left((I_t - T)^{-1}\right)_{jk} - \delta_{jk} \\
&= N_{ij}(N_{jk} - \delta_{jk}).
\end{aligned}$$

The third term is the first term with j and k interchanged, so

$$\sum_{\mu=0}^{\infty} \sum_{\nu=\mu+1}^{\infty} \mathbb{E}\left[U_{ik}^{(\mu)} U_{ij}^{(\nu)}\right] = N_{ik}(N_{kj} - \delta_{kj}).$$

Finally, the second term is $\sum_{\nu=0}^{\infty} \mathbb{E}\left[U_{ij}^{(\nu)} U_{ik}^{(\nu)}\right]$ where each summand is the probability that the Markov chain is in transient state j after exactly ν steps after starting in transient state i and simultaneously in state k after exactly ν steps starting from transient state i. This is only possible if $j = k$ hence

$\mathbb{E}\left[U_{ij}^{(\nu)}U_{ik}^{(\nu)}\right] = p_{ij}^{(\nu)}\delta_{jk}$. Thus for the second term

$$\sum_{\nu=0}^{\infty}\mathbb{E}\left[U_{ij}^{(\nu)}U_{ik}^{(\nu)}\right] = \sum_{\nu=0}^{\infty} p_{ij}^{(\nu)}\delta_{jk} = \left(\sum_{\nu=0}^{\infty}(T^\nu)_{ij}\right)\delta_{jk} = \left(\sum_{\nu=0}^{\infty}T^\nu\right)_{ij}\delta_{jk} = N_{ij}\delta_{jk}.$$

Putting all terms together

$$\mathbb{E}\left[Y_{ij}Y_{ik}\right] = N_{ij}(N_{jk}-\delta_{jk})+N_{ij}\delta_{jk}+N_{ik}(N_{kj}-\delta_{kj}) = N_{ij}N_{jk}+N_{ik}N_{kj}-N_{ik}\delta_{kj}.$$

Then

$$\mathrm{Cov}\left[Y_{ij}, Y_{ik}\right] = \mathbb{E}\left[Y_{ij}Y_{ik}\right]-\mathbb{E}\left[Y_{ij}\right]\mathbb{E}\left[Y_{ik}\right] = N_{ij}N_{jk}+N_{ik}N_{kj}-N_{ij}N_{ik}-N_{ik}\delta_{kj}. \tag{1}$$

Note that the result is symmetric under the interchange of j and k as it should be.

In particular, letting $j = k$ we get the variance of the number of visits to state j starting from i:

$$\mathrm{Var}\left[Y_{ij}\right] = 2N_{ij}N_{jj} - N_{ij}^2 - N_{ij}.$$

Note that this is the variance of the number of visits to transient state j when starting from state i, *not* the variance of the waiting time. Define $\mathrm{diag}(N)$ to be the diagonal matrix setting all off-diagonal elements of N to 0, and define N_{sq} to be the matrix resulting from squaring each entry, then

$$\mathrm{Var}\left[Y_{(i)}\right] = N(2\,\mathrm{diag}(N) - I) - N_{\mathrm{sq}}.$$

The covariance, for a fixed i, is the $t \times t$ matrix of entries $\mathrm{Cov}\left[Y_{ij}, Y_{ik}\right]$. Letting $N_{(i,\cdot)}$ denote the $1 \times t$ row-vector from row i and setting $(\mathrm{diag}\, N_{(i,\cdot)})$ to be the diagonal matrix with this vector along the diagonal this becomes

$$\mathrm{Cov}\left[Y_{(i)}, =\right](\mathrm{diag}\, N_{(i,\cdot)})N + N^T(\mathrm{diag}\, N_{(i,\cdot)}) - N_{(i,\cdot)}N_{(i,\cdot)}^T - \mathrm{diag}\, N_{(i,\cdot)}.$$

Note the outer product $N_{(i,\cdot)}N_{(i,\cdot)}^T$. The variance of the waiting time until absorption into any state from state i is then the sum of all t^2 entries in $\mathrm{Var}\left[Y_{(i)}\right]$.

First-step analysis

First-step analysis says the absorption time from state i is the first step to another transient state plus a weighted average, according to the transition probabilities over the transient states, of the absorption times from the other transient states. In symbols, first-step analysis says

$$w_i = 1 + \sum_{j=a+1}^{a+t} P_{ij} w_j.$$

As before, for a Markov chain with a absorbing states and t transient states, reorder the states so the absorbing states come first and non-absorbing, i.e. transient, states come last. Then the transition matrix has the canonical form:

$$P = \begin{pmatrix} I_a & 0 \\ A & T \end{pmatrix}.$$

Here I_a is an $a \times a$ identity matrix, A is the $t \times a$ matrix of single-step transition probabilities from the t transient states to the a absorbing states, T is a $t \times t$ submatrix of single-step transition probabilities among the transient states, and 0 is a $a \times t$ matrix of 0s representing the single-step transition probabilities from absorbing states to transient states.

Expressing the first-step analysis compactly in vector-matrix form as

$$\mathbf{w} = \mathbf{1} + T\mathbf{w}$$

or

$$(I - T)\mathbf{w} = \mathbf{1}.$$

Then $\mathbf{w} = (I - T)^{-1}\mathbf{1}$. The matrix $N = (I - T)^{-1}$ is the **fundamental matrix** for the absorbing Markov chain. The entries N_{ij} of this matrix have a probabilistic interpretation. The entries N_{ij} are the expected number of times that the chain started from state i will be in state j before ultimate absorption.

The $t \times a$ matrix of **absorption probabilities** B has as entries the probability of starting at state i and ending up at a given absorbing state j. The absorption probabilities come from N by the matrix product $B = NA = (I - T)^{-1}A$.

Theorem 2. *Let b_{ij} be the probability of the Markov chain starting in transient state i and ending in absorbing state j, then*

$$B = (b_{ij}) = NA.$$

Proof. The proof is by first-step analysis. Starting in state i, the process may be captured in j in one or more steps. The probability of capture in a single step is p_{ij}. If this does not happen, the process may move to another absorbing state, in which case it is impossible to reach j, or to a transient state k. In the latter case, the probability of being captured in j is b_{kj}. Hence

$$b_{ij} = p_{ij} + \sum_{k=a+1}^{a+t} p_{ik} b_{kj}$$

or in matrix form $B = A + TB$. Thus, $B = (I - T)^{-1} A = NA$. □

Kemeny and Snell [3, page 51] give an expression for the variance of the waiting time using first-step analysis.

Theorem 3.

$$\operatorname{Var}[\mathbf{w}] = (2N - I)N\mathbf{1} - (N\mathbf{1})_{sq}$$

where $(N\mathbf{1})_{sq}$ *is the element-wise squared vector.*

Proof.

1. Recall that $\mathbb{E}[\mathbf{w}] = N\mathbf{1}$. Let $\mathbf{w}_{sq}$ be the element-wise squared vector. Then use first-step analysis to evaluate $\mathbb{E}[\mathbf{w}_{sq}]$.

2. From its starting state i it can go to any state k with probability p_{ik}. If the new state is absorbing, then it can never reach another state, and the contribution is 1 step. If the new state is transient count the weighted average of the squares of the waiting times plus 1.

3. In symbols:

$$\begin{aligned}
\mathbb{E}[\mathbf{w}_{sq}] &= \sum_{k=1}^{a} p_{ik}\mathbf{1} + \sum_{k=a+1}^{a+t} p_{ik}(\mathbb{E}[(\mathbf{w}+\mathbf{1})_{\mathrm{sq}}])_k \\
&= \sum_{k=1}^{a} p_{ik}\mathbf{1} + \sum_{k=a+1}^{a+t} p_{ik}\Big((\mathbb{E}[\mathbf{w}]_{\mathrm{sq}})_k + 2(\mathbb{E}[\mathbf{w}])_k + \mathbf{1}\Big) \\
&= T\mathbb{E}[\mathbf{w}_{\mathrm{sq}}] + 2T\mathbb{E}[\mathbf{w}] + \mathbf{1}.
\end{aligned}$$

4. Rearrange using $N = (I_t - T)^{-1}$ and $NT = (I_t - T)^{-1}T = N - I_t$ (See

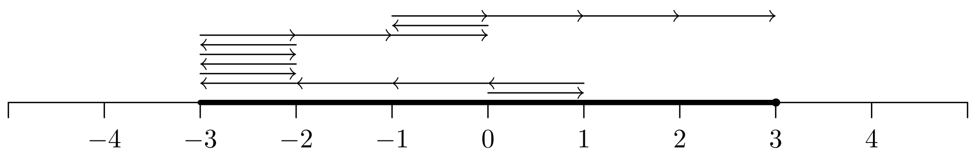


Figure 1: Image of a possible random walk in phase line after an odd number of steps.

the exercises.)

$$\begin{aligned}\mathbb{E}\left[\mathbf{w}_{sq}\right] &= (I_t - T)^{-1}(2T\mathbb{E}\left[\mathbf{w}\right] + \mathbf{1}) \\ &= 2NT\mathbb{E}\left[\mathbf{w}\right] + N\mathbf{1} \\ &= 2(N - I_t)\mathbb{E}\left[\mathbf{w}\right] + \mathbb{E}\left[\mathbf{w}\right] \\ &= (2N - I_t)\mathbb{E}\left[\mathbf{w}\right].\end{aligned}$$

5. Thus

$$\mathrm{Var}\left[\mathbf{w}\right] = (2N - I)N\mathbf{1} - (N\mathbf{1})_{\mathrm{sq}}.$$

□

Remark. Carchidi and Higgins [1] give an alternate proof and slightly different derivation of this formula for the variance of the waiting time until absorption using equation (1).

Examples

Example. Consider a random walk of a particle which moves along a straight line in unit steps. Each step is 1 unit to the right with probability p and to the left with probability $q = 1 - p$. It moves until it reaches one of two extreme points which are **absorbing boundaries**. Assume that if the process reaches the boundary points, it remains there from that time on. Figure 1 has 9 states numbered from -4 to 4. The absorbing boundary states are -4 and 4.

The full transition probability matrix is

$$
P = \begin{array}{c} \\ -4 \\ -3 \\ -2 \\ -1 \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 & 4 \end{array} \\
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
q & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & q & 0 & p & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & q & 0 & p & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & q & 0 & p & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & q & 0 & p & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & q & 0 & p & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & q & 0 & p \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}.
$$

Reorder the states as $-4, 4, -3, -2, -1, 0, 1, 2, 3$ to bring the transition probability matrix to the standard form

$$
P_1 = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
q & 0 & 0 & p & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & q & 0 & p & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & q & 0 & p & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & q & 0 & p & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & q & 0 & p & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & q & 0 & p \\
0 & p & 0 & 0 & 0 & 0 & 0 & q & 0
\end{pmatrix}
$$

so

$$
T = \begin{pmatrix}
0 & p & 0 & 0 & 0 & 0 & 0 \\
q & 0 & p & 0 & 0 & 0 & 0 \\
0 & q & 0 & p & 0 & 0 & 0 \\
0 & 0 & q & 0 & p & 0 & 0 \\
0 & 0 & 0 & q & 0 & p & 0 \\
0 & 0 & 0 & 0 & q & 0 & p \\
0 & 0 & 0 & 0 & 0 & q & 0
\end{pmatrix}.
$$

A computer algebra system can compute the fundamental matrix $N(p) = (I - T)^{-1}$ but the general form in p and q is long and unhelpful. Two representative numerical examples suffice to show the possibilities.

$$N(1/2) = \begin{pmatrix} 7/4 & 3/2 & 5/4 & 1 & 3/4 & 1/2 & 1/4 \\ 3/2 & 3 & 5/2 & 2 & 3/2 & 1 & 1/2 \\ 5/4 & 5/2 & 15/4 & 3 & 9/4 & 3/2 & 3/4 \\ 1 & 2 & 3 & 4 & 3 & 2 & 1 \\ 3/4 & 3/2 & 9/4 & 3 & 15/4 & 5/2 & 5/4 \\ 1/2 & 1 & 3/2 & 2 & 5/2 & 3 & 3/2 \\ 1/4 & 1/2 & 3/4 & 1 & 5/4 & 3/2 & 7/4 \end{pmatrix}$$

Then the waiting times to absorption are $7, 12, 15, 16, 15, 12, 7$. For the variances, consider only the central state, originally labeled as 0, after reordering it is the sixth state in the middle of the transient states.

$$\operatorname{Cov}\left[Y_{0j}, Y_{0k}\right] = \begin{pmatrix} 3/2 & 5/2 & 2 & 1 & 0 & -1/2 & -1/2 \\ 5/2 & 6 & 13/2 & 4 & 3/2 & 0 & -1/2 \\ 2 & 13/2 & 21/2 & 9 & 9/2 & 3/2 & 0 \\ 1 & 4 & 9 & 12 & 9 & 4 & 1 \\ 0 & 3/2 & 9/2 & 9 & 21/2 & 13/2 & 2 \\ -1/2 & 0 & 3/2 & 4 & 13/2 & 6 & 5/2 \\ -1/2 & -1/2 & 0 & 1 & 2 & 5/2 & 3/2 \end{pmatrix}.$$

The variance of the number of visits to 0 until absorption is $\mathbf{1}^T V \mathbf{1} = 160$, the standard deviation is $4\sqrt{10}$. The negative covariance $-1/2$ of the number of visits from 0 to -3 until absorption with the number of visits from 0 to 3 until absorption indicates that the random number of visits from 0 to -3 until absorption has a somewhat opposite distribution to that of the random number of visits from 0 to 3 until absorption. That is, when the number of visits from 0 to -3 until absorption is large, the number of visits from 0 to 3 until absorption will be small. This makes sense, since if the number of visits from 0 to -3 is large, the likelihood of absorption into -4 increases, and the number of visits from 0 to 3 decreases.

The probabilities of absorption

$$N(1/2)A = \begin{pmatrix} 7/8 & 1/8 \\ 3/4 & 1/4 \\ 5/8 & 3/8 \\ 1/2 & 1/2 \\ 3/8 & 5/8 \\ 1/4 & 3/4 \\ 1/8 & 7/8 \end{pmatrix}$$

are symmetric as expected.

If $p = 2/3$ so the probability of moving to the right is twice the probability of moving to the left, then

$$N(2/3) = \begin{pmatrix} 127/85 & 126/85 & 124/85 & 24/17 & 112/85 & 96/85 & 64/85 \\ 63/85 & 189/85 & 186/85 & 36/17 & 168/85 & 144/85 & 96/85 \\ 31/85 & 93/85 & 217/85 & 42/17 & 196/85 & 168/85 & 112/85 \\ 3/17 & 9/17 & 21/17 & 45/17 & 42/17 & 36/17 & 24/17 \\ 7/85 & 21/85 & 49/85 & 21/17 & 217/85 & 186/85 & 124/85 \\ 3/85 & 9/85 & 21/85 & 9/17 & 93/85 & 189/85 & 126/85 \\ 1/85 & 3/85 & 7/85 & 3/17 & 31/85 & 63/85 & 127/85 \end{pmatrix}$$

and the waiting times to absorption are approximately 9.05, 12.07, 12.08, 10.59, 8.34, 5.72, 2.91. For the variances, consider only the central state, originally labeled as 0 and after reordering is the sixth state in the middle of the transient states.

$$\operatorname{Cov}\left[Y_{0j}, Y_{0k}\right] = \begin{pmatrix} \frac{462}{1445} & \frac{162}{289} & \frac{708}{1445} & \frac{72}{289} & 0 & -\frac{144}{1445} & -\frac{144}{1445} \\ \frac{162}{289} & \frac{2232}{1445} & \frac{2682}{1445} & \frac{324}{289} & \frac{504}{1445} & 0 & -\frac{144}{1445} \\ \frac{708}{1445} & \frac{2682}{1445} & \frac{5124}{1445} & \frac{882}{289} & \frac{1764}{1445} & \frac{504}{1445} & 0 \\ \frac{72}{289} & \frac{324}{289} & \frac{882}{289} & \frac{1260}{289} & \frac{882}{289} & \frac{324}{289} & \frac{72}{289} \\ 0 & \frac{504}{1445} & \frac{1764}{1445} & \frac{882}{289} & \frac{5838}{1445} & \frac{720}{289} & \frac{912}{1445} \\ -\frac{144}{1445} & 0 & \frac{504}{1445} & \frac{324}{289} & \frac{720}{289} & \frac{4068}{1445} & \frac{1728}{1445} \\ -\frac{144}{1445} & -\frac{144}{1445} & 0 & \frac{72}{289} & \frac{912}{1445} & \frac{1728}{1445} & \frac{1176}{1445} \end{pmatrix}.$$

The variance is $\mathbf{1}^T V \mathbf{1} = 15264/289 \approx 52.81661$, the standard deviation is 7.26750, less than the symmetric case where $p = 1/2 = q$, as expected.

The probabilities of absorption are

$$N(2/3)A = \begin{pmatrix} \frac{127}{255} & \frac{128}{255} \\ \frac{21}{85} & \frac{64}{85} \\ \frac{31}{255} & \frac{224}{255} \\ \frac{1}{17} & \frac{16}{17} \\ \frac{7}{255} & \frac{248}{255} \\ \frac{1}{85} & \frac{84}{85} \\ \frac{1}{255} & \frac{254}{255} \end{pmatrix}.$$

The absorption probabilities are strongly biased to the right, as expected.

Example. It's your 30th birthday, and your friends bought you a cake with 30 candles on it. You make a wish and try to blow them out. Every time you blow, you blow out a random number of candles between one and the number that remain, including one and that other number. How many times on average do you blow before you extinguish all the candles?

Let the states be the number of candles remaining lit. Order the $N = 31$ states as $0, 1, 2, 3, \dots, 29, 30$. Then the state 0 is absorbing and all other states are transient. Interpret "blow out a random number of candles between one and the number that remain" as uniform distribution on the number of candles remaining. Instead of solving this full problem at once, try a smaller problem first, with the number of candles $N = 5$. Then the transition probability matrix in canonical form is

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 1/5 & 1/5 & 1/5 & 1/5 & 1/5 & 0 \end{pmatrix}.$$

Then the first-step equations for the waiting time, that is the number of attempts needed to blow out the candles are

$$\begin{aligned} w_1 &= 1 \\ w_2 &= 1 + \frac{1}{2} w_1 \\ w_3 &= 1 + \frac{1}{3} w_1 + \frac{1}{3} w_2 \\ w_4 &= 1 + \frac{1}{4} w_1 + \frac{1}{4} w_2 + \frac{1}{4} w_3 \\ w_5 &= 1 + \frac{1}{5} w_1 + \frac{1}{5} w_2 + \frac{1}{5} w_3 + \frac{1}{5} w_4 \end{aligned}$$

Solving this recursively, $w_1 = 1$, $w_2 = 1 + \frac{1}{2}$, $w_3 = 1 + \frac{1}{3} + \frac{1}{3} \cdot (1 + \frac{1}{2})$, so $w_3 = 1 + \frac{1}{2} + \frac{1}{3}$. Continuing to solve recursively

$$\begin{aligned} w_4 &= 1 + 1/2 + 1/3 + 1/4 \\ w_5 &= 1 + 1/2 + 1/3 + 1/4 + 1/5. \end{aligned}$$

The inductive pattern is clear. The waiting time with N candles is the Nth **harmonic number** H_N

$$w_N = H_N = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots + \frac{1}{N}.$$

For the original problem with 30 candles the transient states are $1, 2, 3, \ldots, 30$ and the absorbing state is 0. Then the canonical form transition probability matrix is

$$P = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 1/2 & 1/2 & 0 & \dots & 0 & 0 \\ 1/3 & 1/3 & 1/3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1/30 & 1/30 & 1/30 & \dots & 1/30 & 0 \end{pmatrix}.$$

The quantity of interest is the expected waiting time until absorption into the state 0. Expressing the first-step analysis compactly in vector-matrix form as

$$(I - T)\mathbf{W} = \mathbf{1},$$

substituting in the values from the transition matrix and solving with a computer, this is $w_{30} = H_{30} \approx 3.9950$. By calculating the covariance matrix and taking the sum of all entries, the variance of the waiting time from the transient state of all candles lit is approximately 2.3828.

Example. An urn contains two unpainted balls. At a sequence of times, choose a ball at random, then paint it either red or black, and put it back. For an unpainted ball, choose a color at random. For a painted ball, change its color. Form a Markov chain by taking as a state the triple (x, y, z) where x is the number of unpainted balls, y the number of red balls, and z the number of black balls. The transition matrix is then

$$\begin{array}{c} \\ (0,1,1) \\ (0,2,0) \\ (0,0,2) \\ (2,0,0) \\ (1,1,0) \\ (1,0,1) \end{array} \begin{array}{c} \begin{array}{cccccc} (0,1,1) & (0,2,0) & (0,0,2) & (2,0,0) & (1,1,0) & (1,0,1) \end{array} \\ \begin{pmatrix} 0 & 1/2 & 1/2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 0 & 0 & 0 & 1/2 \\ 1/4 & 0 & 1/4 & 0 & 1/2 & 0 \end{pmatrix} \end{array}.$$

In this case, there is no absorbing state, that is, a state which once entered remains the same thereafter. However, the first three states together are an irreducible set, that is, once the process enters that set, it continues in that set. So lump those states together as a single absorbing state, and the transient states are $(2,0,0)$, $(1,1,0)$, and $(1,0,1)$. Then

$$T = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{pmatrix}$$

and the fundamental matrix is

$$N = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 4/3 & 2/3 \\ 0 & 2/3 & 4/3 \end{pmatrix}.$$

With this, we can compute

$$\mathbb{E}\left[\mathbf{w}\right] = N\mathbf{1} = (3,2,2)^T$$

and

$$\operatorname{Var}\left[\mathbf{w}\right] = (2N - I)\mathbb{E}\left[\mathbf{w}\right] - \left(\mathbb{E}\left[\mathbf{w}\right]\right)_{\text{sq}} = (2,2,2)^T.$$

Finally,

$$\operatorname{Var}\left[\mathbf{X}\right] = N(2\operatorname{diag}(N) - I) - N_{\text{sq}} = \begin{pmatrix} 0 & 2/3 & 2/3 \\ 0 & 4/9 & 2/3 \\ 0 & 2/3 & 4/9 \end{pmatrix}.$$

Since the process must immediately leave state $(2,0,0)$ and cannot go back, the variance is 0 for the number of times in this state.

Example. The following is a larger example of the painting the balls puzzle.

You play a game with four balls in a box: One ball is red, one is blue, one is green and one is yellow. You draw a ball out of the box at random and note its color. Without replacing the first ball, you draw a second ball and then paint it to match the color of the first. Replace both balls, and repeat the process. The game ends when all four balls have become the same color. What is the expected number of turns to finish the game?

Take the states as the number of balls of each different color without regard for the colors themselves. For example, partition 4 balls in 5 different ways:

$$1+1+1+1, \quad 2+1+1, \quad 2+2, \quad 3+1, \quad 4.$$

The partition $2+1+1$, for example, consists of cases where two of the balls are the same color and the other two balls are two other colors. For example, the cases “red & red & green & blue” and “blue & blue & yellow & red’ correspond to the partition $2+1+1$. By using these five partitions as states in a Markov Chain, we can compute the transition probabilities to go from one state to the next. For example, the probability of transition from $2+1+1$ to $3+1$ is $1/3$ because in order for this transition to occur, we must first choose one of the two identically colored balls with probability $2/4$, then we must choose one of the other two balls out of the remaining three with probability $2/3$. The joint probability is $2/4 \cdot 2/3 = 1/3$. As another example, the probability of transition from $2+2$ to $3+1$ is the probability of first picking a ball of either color, leaving 1 ball of that color and 2 balls of the other color, then from those 3 balls picking the second color with probability $2/3$. Calculate all remaining transition probabilities in the same way.

$$P = \begin{array}{c} \\ 1+1+1+1 \\ 2+1+1 \\ 3+1 \\ 2+2 \\ 4 \end{array} \begin{array}{c} \begin{array}{ccccc} 1+1+1+1 & 2+1+1 & 3+1 & 2+2 & 4 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 1/2 & 1/3 & 1/6 & 0 \\ 0 & 0 & 1/2 & 1/4 & 1/4 \\ 0 & 0 & 2/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

The absorbing state is 4. Rearranging into the canonical block-matrix form

$$P = \begin{array}{c} \\ 4 \\ 1+1+1+1 \\ 2+1+1 \\ 3+1 \\ 2+2 \end{array} \begin{array}{c} \begin{array}{ccccc} 4 & 1+1+1+1 & 2+1+1 & 3+1 & 2+2 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/3 & 1/6 \\ 1/4 & 0 & 0 & 1/2 & 1/4 \\ 0 & 0 & 0 & 2/3 & 1/3 \end{pmatrix} \end{array}.$$

Then $N = (I - T)^{-1}$ is

$$\begin{pmatrix} 1 & 2 & 4 & 2 \\ 0 & 2 & 4 & 2 \\ 0 & 0 & 4 & 3/2 \\ 0 & 0 & 4 & 3 \end{pmatrix}$$

and the waiting time from the initial state is $1+2+4+2=9$. From the state $1+1+1+1$ (labeled as state 2 in the canonical format), the covariance matrix is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 12 & 6 \\ 0 & 0 & 6 & 6 \end{pmatrix}.$$

Then the variance of the number of visits is 32 and the standard deviation of the number of visits is $4\sqrt{2} \approx 5.66$.

Example. The following example is from the December 22, 2017 Riddler at Fivethirtyeight.com. It concerns a game of chance called Left, Right, Center. In this game, everyone sits in a circle and starts with some number of \$1 bills. You take turns, in order around the circle, rolling three dice. For each die, if it comes up 1 or 2, you give a dollar to the person on your left. If it comes up 3 or 4, you give a dollar to the person on your right. And if it comes up 5 or 6, you put a dollar in the center. Assume the following: First, if a player has no dollars, then her turn is skipped. Second, if a player has one or two dollars, then the player rolls only one or two dice, respectively. The game ends as soon as only a single person has any money left. How long is the game expected to last for six players each starting with three \$1 bills? For X players each starting with Y \$1 bills?

Consider the allocation of the total of \$18 among the 6 players and the center, that is 7 places, as the states of a Markov chain. Using the "stars and bars" argument, choosing $7 - 1 = 6$ bars from $18 + 6$ objects the number of states is $\binom{18+6}{6} = 134{,}596$. This also counts all the money in the center as one state which would not strictly be a state of the game. This makes 134,595 actual game states. This is the same as the number of states dividing any number of dollars from 1 to 18 among 6 players, ignoring the money in the center which would be $\sum_{j=6}^{23} \binom{j}{5} = 134{,}595$. This is essentially the conclusion of the hockey-stick identity for binomial coefficients.

The game ends when exactly 1 of the 6 players has any amount of money from \$1 to \$18. That is, there are $6 \cdot (1 + \cdots + 18) = 6 \cdot 18 \cdot 19/2 = 1026$ terminal states that can be considered as the absorbing states of the game.

The shortest possible game would be for each player in turn to throw all 5's and 6's, thereby placing all the player's starting money in the center. This shortest game would be 6 turns with probability $(1/27)^6$. Consider the total number of dollars in the circle at each turn. The expected loss of dollars

	6	7	8	9
3	25.7096	32.3160	38.9660	45.3878
4	31.7890	39.7598	47.3006	55.3804
5	38.3244	47.4416	56.3536	65.4438
6	44.8394	54.9732	65.5688	75.3548

Table 1: Mean number of turns in 5000 simulations of the game with X players (columns) each with Y dollars (rows).

from the circle to the center at each turn is

$$0 \cdot \left(\frac{2}{3}\right)^3 + 1 \cdot \left(3 \cdot \left(\frac{2}{3}\right)^2 \cdot \left(\frac{1}{3}\right)\right) + 2 \cdot \left(3 \cdot \left(\frac{2}{3}\right) \cdot \left(\frac{1}{3}\right)^2\right) + 3 \cdot \left(\frac{1}{3}\right)^3 = 1.$$

That means it takes should take about 17 turns to end the game. More precisely, consider the lumped system where each state is the number of dollars in the circle, from \$18 to \$1. Let D_j be the duration of the game from j dollars with $j = 1, \dots, 18$. Then a first-step analysis gives the system of equations $D_j = (8/27)D_j + (12/27)D_{j-1} + (6/27)D_{j-2} + (1/27)D_{j-3} + 1$, with $D_j = 0$ for $j = 1, 0, -1$. The solution of this system with Octave gives $D_{18} = 17.3333$ (see the scripts for the solution method) which is consistent with the previous crude estimate.

The number of states and absorbing states are too large to be effectively handled by matrix methods. Although the conceptual set-up of the game is clearly the waiting time until absorbtion in a Markov chain, simulation seems to be the best way to answer the question. Results of simulating games with 6 to 9 players each starting with 3 to 6 dollars are in the tables. The Riddler says that the game with X players each starting with Y dollars will last $2(X-2)Y$ turns but does not give a proof. The simulations give means which are roughly consistent with this value.

	6	7	8	9
3	6.546804	7.009249	7.564145	7.898221
4	7.751729	8.176348	8.594918	9.099816
5	8.989401	9.367077	9.878952	10.387840
6	10.161804	10.468601	11.080443	11.654668

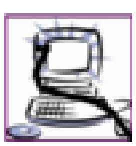
Table 2: Variance of number of turns in 5000 simulations of the game with X players each with Y dollars.

Section Ending Answer

The number of times the Markov chain visits a transient state is a random variable. Let Y_{ij} denote the number of visits the system makes to transient state j before reaching an absorbing state, given the system started in transient state i. Thus, Y_{ij} is a discrete random variable that can take on any nonnegative integer value.

Sources

The subsection on covariances is adapted from Carchidi and Higgins [1]. The section on first-step analysis is adapted from *Finite Markov Chains* by Kemeny and Snell [3]. Other ideas are from *An Introduction to Stochastic Modeling* by Taylor and Karlin and *Random Walks and Electrical Networks* by Doyle and Snell [2]. The birthday candle example is adapted from the January 13, 2017 "Riddler" at Fivethirtyeight.com/ The two colored balls example is adapted from *Finite Markov Chains* by Kemeny and Snell [3]. The four colored balls example is adapted from http://www.laurentlessard.com/bookproofs/colorful-balls-puzzle/ The Left-Center-Right game example is from the December 22, 2017 Riddler at Fivethirtyeight.com.



Algorithms, Scripts, Simulations

Algorithm

Turn

Initialize Mean and Variance matrices to hold results
Initialize number of simulations (5000)
Loop over number of players for each simulation (6:9)
Loop over the number of bills each player starts with (3:6)
Initialize empty vector to hold waiting time for each simulation
Record results of waiting times for each simulation
return Mean and Variance of Waiting Times to Game End
Modular division to number the players in a circle
Initialize the player circle to hold equal number of bills
while more than one player has more that one bill
Current player takes a Turn
Move to next player in circle
turns = turns + 1

return turns
Get 3 random integers from 1 to 6 into a vector
for three places in the vector
if current player has bills
if roll (vector entry) is 1 or 2
Player to left gets bill, current player loses a bill,
if roll (vector entry) is 3 or 4
Player to right gets bill, current player loses a bill,
if roll (vector entry) is 5 or 6
Center gets bill (don't keep track), current player loses a bill
return State of players (bills each holds)

Scripts

R R script for Waiting Time in Left-Center-Right.

```
player_index <- function(j) {
    ((j-1) %% player_count) + 1
}

```

```r

game <- function(bills, player_count){

    turn <- function(player) {
        rolls <- sample(6, 3, replace=TRUE)
        for (roll in rolls) {
            if (players[player] > 0) {      #player has money
                if (roll <= 2) {            #pass dollar to Left
                    players[player_index(player - 1)] =
                        players[player_index(player - 1)] + 1
                    players[player] = players[player] - 1
                } else if (roll <= 4) {     #pass dollar to right
                    players[player_index(player + 1)] =
                        players[player_index(player + 1)] + 1
                    players[player] = players[player] - 1
                } else {                    #dollar goes to center
                    players[player] = players[player] - 1
                }
            }
        }
        return(players)
    }

    players <- rep(bills, player_count)
    turns <- 1
    current_player <- 1

    while ( length( players[ players > 0] ) > 1 ) {
        players <- turn(current_player)
        current_player <- player_index(current_player + 1)
        turns <- turns + 1
    }

    return(turns)
}

M <- matrix(0, 4, 4)
V <- matrix(0, 4, 4)
```

```r
simulations <- 5000

for (player_count in c(6:9)) {
    for (bills in c(3:6)) {

        sim_results <- c()
        for (i in 1:simulations) {
            sim_results <- c( sim_results, game(bills, player_count) )
        }

        M[player_count-5, bills-2] <- mean(sim_results)
        V[player_count-5, bills-2] <- sd(sim_results)
    }
}
```

Octave

```octave
    V = [(8/27)*ones(17,1), (12/27)*ones(17,1), (6/27)*ones(17,1),
         (1/27)*ones(17,1)]
    A = spdiags(V, [0,1,2,3], eye(17))
    (eye(17) - A)\ones(17,1)
```

Problems to Work for Understanding

1: A law firm employs three types of lawyers: junior lawyers, senior lawyers, and partners. During a given year, there is probability 0.15 of promoting a junior lawyer to senior lawyer and probability 0.5 that the junior lawyer will leave the firm. There is probability 0.20 of promoting the senior lawyer to partner and probability 0.10 that the senior lawyer will leave the firm. Finally, there is probability 0.05 that a partner will leave the firm, see Table 3. The firm never demotes a lawyer or a partner.

Table 3: The transition probabilities

	leave firm	junior lawyer	senior lawyer	partner
leave firm	1	0	0	0
junior lawyer	0.05	0.80	0.15	0
senior lawyer	0.10	0	0.70	0.20
partner	0.05	0	0	0.95

1. What is the average number of years that a newly hired junior lawyer stays with the firm?

2. What is the variance of the number of years a newly hired junior lawyer stays with the firm?

2: Consider the random walk of a particle along a straight line from -4 to 4 with probability $2/3$ of moving to the right and probability $1/3$ of moving to the left. The absorbing boundary states are -4 and 4. Compute the vector of variances of waiting times until absorbtion using Theorem 3. Compare the result for the walk started at 0 with the variance computed in the text using the sum of the entries in the covariance matrix.

3: Use mathematical induction to show

$$P^m = \begin{pmatrix} I_a & 0 \\ A & T \end{pmatrix}^{m-1} \begin{pmatrix} I_a & 0 \\ A & T \end{pmatrix} = \begin{pmatrix} I_a & 0 \\ (I_t + T + T^2 + \cdots + T^{m-1})A & T^m \end{pmatrix}.$$

4:

(a) Use computer software to show that the transient state transition probability matrix T in the random walk example has induced 2-norm (or operator norm) less than one for $p = 1/2$ and $p = 2/3$.

(b) Use computer software to show that the transient state transition probability matrix T in the birthday candle example with $N = 5$, has induced 2-norm (or operator norm) less than 1.

(c) Use computer software to show that the transient state transition probability matrix T in the two unpainted balls example has induced 2-norm (or operator norm) less than 1.

(d) (Mathematicians only) Show that if T is irreducible and row sub-stochastic, with at least one row having sum less than 1, then the induced 2-norm (or operator norm) of T is less than 1, $\|T\| < 1$.

(e) Use computer software to show that the transient state transition probability matrix T in the four painted balls example has induced 2-norm (or operator norm) *greater* than 1. Explain why the previous theorem is not violated. Nevertheless, show that the eigenvalues of the matrix are all less than 1, so that T is contractive and $(I - T)^{-1}$ still exists.

5: Show that $NT = (I_t - T)^{-1}T = N - I_t$

6: Under the assumption that the induced 2-norm (or operator norm) of T is less than 1, $\|T\| < 1$ prove that

$$\sum_{\nu=0}^{\infty}(T^{\nu}) = (I - T)^{-1}.$$

7: For the birthday candle problem, prove by induction that the waiting time with N candles is the Nth **harmonic number** H_N.

Reading Suggestion:

References

[1] Michael A. Carchidi and Robert L. Higgins. Covariances between transient states in finite absorbing markov chains. *The College Mathematics Journal*, 48(1):42–50, Janaury 2017.

[2] Peter G. Doyle and J. Laurie Snell. *Random Walks and Electric Networks*. The Mathematical Association of America, 1984.

[3] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. D. Van Nostrand Company, Inc., 1960.

Outside Readings and Links:

1. http://www.laurentlessard.com/bookproofs/colorful-balls-puzzle/
2.
3.
4.

Solutions

1: This leads to the single-step transition matrix P partitioned into one absorbing state for "leave firm" and three transient states: junior lawyer, senior lawyer, partner

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.05 & 0.80 & 0.15 & 0 \\ 0.10 & 0 & 0.70 & 0.20 \\ 0.05 & 0 & 0 & 0.95 \end{pmatrix}.$$

The transient transition matrix is

$$T = \begin{pmatrix} 0.80 & 0.15 & 0 \\ 0 & 0.70 & 0.20 \\ 0 & 0 & 0.95 \end{pmatrix}.$$

The corresponding fundamental matrix is

$$N = (I_3 - T)^{-1} = \begin{pmatrix} 5 & 5/2 & 10 \\ 0 & 10/3 & 40/3 \\ 0 & 0 & 20 \end{pmatrix}.$$

The waiting times to absorption are then

$$(I-T)^{-1}\mathbf{1} = (35/2, 50/3, 20)^T.$$

Now compute the covariances for the first transient state, junior lawyer. Using

$$\operatorname{diag} N_{(1)} N + N^T \operatorname{diag} N_{(1)} - N_{(1)}^T N_{(1)} - \operatorname{diag} N_{(1)}$$

we get

$$\operatorname{Cov}[Y_{1j}, Y_{1k}] = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 95/12 & 25/3 \\ 0 & 25/3 & 290 \end{pmatrix} = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 7.91667 & 8.33333 \\ 0 & 8.33333 & 290 \end{pmatrix}.$$

Since $w_i = \sum_{j=a+1}^{t} Y_{ij}$, using that $\operatorname{Var}[w_i] = \sum_{j=a+1}^{t} \sum_{k=a+1}^{t} \operatorname{Cov}[Y_{1j}, Y_{1k}]$, the variance is the sum of the entries in the covariance matrix. Hence $\operatorname{Var}[w_1] = 334.58$ (in units of year^2). The standard deviation is 18.29 years. As a point of special interest, note that the standard deviation is larger than the mean so that knowing the mean is not enough to understand how long a junior lawyer will be with the firm.

2: Theorem 3 gives the variances of the waiting times as

$$\operatorname{Var}[\mathbf{w}] = (2N - I)N\mathbf{1} - (N\mathbf{1})_{\text{sq}}.$$

Using the matrix N from the test example, this is

$$\begin{pmatrix} \frac{551592}{7225} \\ \frac{476496}{7225} \\ \frac{83352}{1445} \\ \frac{15264}{289} \\ \frac{330312}{7225} \\ \frac{247536}{7225} \\ \frac{5400}{289} \end{pmatrix}$$

or approximately

$$\begin{pmatrix} 76.34491349480969 \\ 65.95100346020762 \\ 57.68304498269896 \\ 52.81660899653979 \\ 45.71792387543253 \\ 34.26103806228374 \\ 18.68512110726644 \end{pmatrix}$$

The entry for 0 rounds to 52.81661, the same as the variance in the text example.

3: The equality is trivial for the base case $m = 1$. Assume the equality holds for $m - 1$. Check the multiplication for the lower left block, the other three blocks are easily seen to hold. The multiplication for the lower left block is

$$\big((I_t + T + T^2 + \cdots + T^{m-2})A\big)\, I_a + T^{m-1}A = (I_t + T + T^2 + \cdots + T^{m-1})A.$$

Note that $AI_a = A$. Using associative and distributive properties for matrix multiplication completes the induction step.

4:

(a) With R, for $p = 1/2$, `norm(T, type=c("2")) = 0.9238795`. With R, for $p = 2/3$, `norm(T, type=c("2")) = 0.9326442`.

(b) With R, `norm(T, type=c("2")) = 0.7941471`.

(c) With R, `norm(T, type=c("2")) = 0.8660254`.

(d) Proof lightly adapted from https://math.stackexchange.com/questions/36828/substochastic-matrix-spectral-radius#666603 For any state i and integer $n \ge 0$, let $r_i^n = \sum_k (T^n)_{ik}$ denote the ith row sum of T^n. For $n = 1$, for convenience write r_i rather than r_i^1. Since T is a substochastic transition probability matrix, $0 \le r_i^n \le 1$.

Let k^* be an index with $r_{k^*} < 1$, and note that for $n \ge 1$

$$r_{k^*}^n = \sum_k T_{k^*k} r_k^{n-1} \le \sum_k T_{k^*k} = r_{k^*} < 1.$$

By irreducibility, for any i, there is an m with $(T^m)_{ik^*} > 0$. In fact, if T is an $N \times N$ matrix, and $i \ne k^*$ then take $m < N$. (Take the shortest path from i to k^* with positive transition probability). Since $(T^m)_{ik}$ puts positive weight on the index $k = k^*$,

$$r_i^N = \sum_k T_{ik}^m r_k^{N-m} < r_i^m \le 1.$$

That is, every row sum of T^N is strictly less than 1. Now it is possible to show that $T^{jN} \to 0$ as $j \to \infty$ and this shows that T^N (and hence T) cannot have any eigenvalue with modulus 1.

(e) Note that if $\mathbf{x} = (0,1,0,0)^T$, then $T\mathbf{x} = (0,1,1/2,0)$ with $\|T(x)\| = \sqrt{5}/2 \approx 1.118 > 1$. In fact, With R, `norm(T, type=c("2")) = 1.162299` but the eigenvalues are 0.83333, 0.50000, 0 and 0. The previous theorem is not violated because the transient probability transition matrix is not irreducible, it is not possible to reach state $1+1+1+1$. Note also that the converse of the previous theorem is not true: There is an example of a matrix with operator norm less than 1, but it is not irreducible

5: First note that $NT = TN$ because

$$\begin{aligned} T(I_t - T) &= (I_t - T)T \\ T &= (I_t - T)T(I_t - T)^{-1} \\ (I_t - T)^{-1}T &= T(I_t - T)^{-1} \\ NT &= TN \end{aligned}$$

Start with $N = (I_t - T)^{-1}$. Then

$$\begin{aligned} I_t &= N(I_t - T) \\ 0 &= N - NT - I_t \\ T &= N - TN - I_t + T \\ T &= (I_t - T)N - (I_t - T) \\ T &= (I_t - T)(N - I_t) \\ (I_t - T)^{-1}T &= N - I_t \end{aligned}$$

6: Consider

$$\begin{aligned} (I-T)(I+T+T^2+\cdots+T^m) = (I+T+T^2+\cdots+T^m)- \\ (T+T^2+\cdots+T^{m+1}) = I - T^{m+1} \end{aligned}$$

Then

$$\|(I-T)\sum_{\nu=0}^{m}(T^{\nu}) - I\| = \|T^{m+1}\| \to 0$$

so

$$\sum_{\nu=0}^{\infty}(T^{\nu}) = (I-T)^{-1}.$$

7: With 1 candle, and the assumption you blow out a random number of candles between one and the number that remain, w_1 must be $1 = H_1$ and the base case is established. Next assume

$$w_N = H_N = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{N} = \sum_{\nu=1}^{N} \frac{1}{\nu}$$

and

$$\begin{aligned}
w_{N+1} &= 1 + \frac{1}{N+1} w_1 + \frac{1}{N+1} w_2 + \cdots + \frac{1}{N+1} w_N \\
&= 1 + \frac{1}{N+1} \sum_{\nu=1}^{N} w_\nu \\
&= 1 + \frac{1}{N} \sum_{\nu=1}^{N+1} \sum_{j=1}^{\nu} \frac{1}{j} \\
&= 1 + \frac{1}{N+1} \sum_{j=1}^{N} \frac{N-(j-1)}{j} \\
&= 1 + \frac{1}{N+1} \sum_{j=1}^{N} \left(\frac{N+1}{j} - 1 \right) \\
&= 1 + \sum_{j=1}^{N} \frac{1}{j} - \frac{1}{N+1}(N) \\
&= \sum_{j=1}^{N-1} \frac{1}{j} + \frac{1}{N+1} \\
&= w_{N+1}
\end{aligned}$$

I check all the information on each page for correctness and typographical errors. Nevertheless, some errors may occur and I would be grateful if you would alert me to such errors. I make every reasonable effort to present current and accurate information for public use, however I do not guarantee the accuracy or timeliness of information on this website. Your use of the information from this website is strictly voluntary and at your risk.

I have checked the links to external sites for usefulness. Links to external websites are provided as a convenience. I do not endorse, control, monitor, or guarantee the information contained in any external website. I don't guarantee that the links are active at all times. Use the links here with the same caution as you would all information on the Internet. This website reflects the thoughts, interests and opinions of its author. They do not explicitly represent official positions or policies of my employer.

Information on this website is subject to change without notice.

Steve Dunbar's Home Page, `http://www.math.unl.edu/~sdunbar1`

Email to Steve Dunbar, `sdunbar1 at unl dot edu`

Last modified: Processed from LaTeX source on July 27, 2020