## 2. Appendix

> Page 6
> Example 1. State 0, when it rains, state 1 when it does not rain. Then the preceding is a two-state Markov chain.

## Page 7 <br> Example 2.

state 0: if it rained both today and yesterday,
state 1: if it rained today but not yesterday,
state 2: if it rained yesterday but not today,
state 3: if it did not rain either yesterday or today.

Page 8

$$
\begin{aligned}
P_{i j}^{n+m} & =P\left\{X_{n+m}=j \mid X_{0}=i\right\} \\
& =\sum_{k=0}^{\infty} P\left\{X_{n+m}=j, X_{n}=k \mid X_{0}=i\right\} \\
& =\sum_{k=0}^{\infty} P\left\{X_{n+m}=j \mid X_{n}=k, X_{0}=i\right\} P\left\{X_{n}=k \mid X_{0}=i\right\} \\
& =\sum_{k=0}^{\infty} P_{k j}^{m} P_{i k}^{n}
\end{aligned}
$$

Page 9, Example 4.

$$
\begin{gathered}
\mathbf{P}^{(2)}=\mathbf{P}^{2}=\left\|\begin{array}{ll}
0.7 & 0.3 \\
0.4 & 0.6
\end{array}\right\| \cdot\left\|\begin{array}{ll}
0.7 & 0.3 \\
0.4 & 0.6
\end{array}\right\|=\left\|\begin{array}{ll}
0.61 & 0.39 \\
0.52 & 0.48
\end{array}\right\| \\
\mathbf{P}^{(4)}=\left(\mathbf{P}^{2}\right)^{2}=\left\|\begin{array}{ll}
0.61 & 0.39 \\
0.52 & 0.48
\end{array}\right\| \cdot\left\|\begin{array}{ll}
0.61 & 0.39 \\
0.52 & 0.48
\end{array}\right\|=\left\|\begin{array}{ll}
0.5749 & 0.4251 \\
0.5668 & 0.4332
\end{array}\right\| .
\end{gathered}
$$

Page 9, Example 6. To determine $P(N \leq 8)$, define a Markov chain with states $0,1,2,3$ where for $i<3$ state $i$ means that we currently are on a run of $i$ consecutive heads, and where state 3 means that a run of three consecutive heads has already occurred. Thus, the transition probability matrix is

$$
\mathbf{P}=\left(\begin{array}{cccc}
1 / 2 & 1 / 2 & 0 & 0 \\
1 / 2 & 0 & 1 / 2 & 0 \\
1 / 2 & 0 & 0 & 1 / 2 \\
0 & 0 & 0 & 1
\end{array}\right)
$$

Because there would be a run of three consecutive heads within the first eight flips if and only if $X_{8}=3$, the desired probability is $P_{08}^{3}$. Squaring $\mathbf{P}$ to obtain $\mathbf{P}^{2}$, then squaring the result to obtain $\mathbf{P}^{4}$, and then squaring matrix gives the results

$$
\mathbf{P}^{8}=\left(\begin{array}{cccc}
81 / 256 & 44 / 256 & 24 / 256 & 107 / 256 \\
68 / 256 & 37 / 256 & 20 / 256 & 131 / 256 \\
44 / 256 & 24 / 256 & 13 / 256 & 175 / 256 \\
0 & 0 & 0 & 1
\end{array}\right)
$$

Hence, the probability that there will be a run of three consecutive heads within the first eight flips is $107 / 256 \approx .4180$.
(b) Noting that $N=8$ if the pattern has not yet occurred in the first 7 transitions, the state after 7 transitions is 2 , and the next flip lands heads, shows that

$$
P(N=8)=\frac{1}{2} \mathbf{P}_{0,2}^{7}
$$

## Page 13

suppose that state $i$ is transient. Hence, each time the process enters state $i$ there will be a positive probability, namely, $1-f_{i}$, that it will never again enter that state. Therefore, starting in state $i$, the probability that the process will be in state i for exactly $n$ time periods equals $f_{i}^{n-1}\left(1-f_{i}\right), n \geq 1$. In other words, if state $i$ is transient then, starting in state $i$, the number of time periods that the process will be in state $i$ has a geometric distribution with finite mean $1 /\left(1-f_{i}\right)$.

## Page 14

From the preceding two paragraphs, it follows that state $i$ is recurrent if and only if, starting in state $i$, the expected number of time periods that the process is in state $i$ is infinite. But, letting

$$
I_{n}= \begin{cases}1, & \text { if } X_{n}=i \\ 0, & \text { if } X_{n} \neq i\end{cases}
$$

We have that $\sum_{n=0}^{\infty} I_{n}$ represents the number of periods that the process is in state $i$

$$
\mathrm{E}\left[\sum_{n=0}^{\infty} I_{n} \mid X_{0}=i\right]=\sum_{n=0}^{\infty} E\left[I_{n} \mid X_{0}=i\right]=\sum_{n=0}^{\infty} P\left(X_{n}=i \mid X_{0}=i\right)=\sum_{n=0}^{\infty} P_{i i}^{n}
$$

## Page 14, Corollary 1

We note that, since state $i$ communicates with state $j$, there exist integers $k$ and $m$ such that $P_{i j}^{k}>0, P_{j i}^{m}>0$. Now, for any integer n

$$
P_{j j}^{m+n+k} \geq P_{i j}^{k} P_{i i}^{n} P_{j i}^{m}
$$

So

$$
\sum_{n=1}^{\infty} P_{j j}^{m+n+k} \geq P_{j i}^{m} P_{i j}^{k} \sum_{n=1}^{\infty} P_{i i}^{n}=\infty .
$$

## Page 16 Example 10

- Since it is impossible to be even (using the gambling model interpretation) after an odd number of plays we must, of course, have that

$$
P_{00}^{2 n-1}=0, n=1,2, \ldots
$$

## Page 16 Example 10 continuous

- On the other hand, we would be even after $2 n$ trials if and only if we won $n$ of these and lost $n$ of these. Because each play of the game results in a win with probability p and a loss with probability $1-p$, the desired probability is thus the binomial probability

$$
P_{00}^{2 n}=C_{2 n}^{n} p^{n}(1-p)^{n}=\frac{(2 n)!}{n!n!} p^{n}(1-p)^{n} \approx \frac{4 p^{n}(1-p)^{n}}{\sqrt{\pi n}}, n=1,2,3, \ldots
$$

since $C_{2 n}^{n} \approx 2^{2 n} / \sqrt{n \pi}$. Then by showing

$$
\sum_{n=1}^{\infty} \frac{4 p^{n}(1-p)^{n}}{\sqrt{\pi n}}=\infty
$$

thus all states are recurrent.

## Page 17 Proposition 2

Because $i$ and $j$ communicate there is a value n such that $P^{i, j}>0$. Let $X_{0}=i$ and say that the first opportunity is a success if $X_{n}=j$, and note that the first opportunity is a success with probability $P_{i, j}^{n}>0$. If the first opportunity is not a success then consider the next time (after time $n$ ) that the chain enters state $i$. (Because state $i$ is recurrent we can be certain that it will eventually reenter state $i$ after time $n$.) Say that the second opportunity is a success if $n$ time periods later the Markov chain is in state $j$. If the second opportunity is not a success then wait until the next time the chain enters state $i$ and say that the third opportunity is a success if $n$ time periods later the Markov chain is in state $j$. Continuing in this manner, we can define an unlimited number of opportunities, each of which is a success with the same positive probability $P_{i, j}^{n}$. Because the number of opportunities until the first success occurs is geometric with parameter $P_{i, j}^{n}$, it follows that with probability 1 a success will eventually occur and so, with probability 1 , state $j$ will eventually be entered.

## Page 18 Proposition 3

Suppose that the Markov chain starts in state i , and let $T_{1}$ denote the number of transitions until the chain enters state $j$; then let $T_{2}$ denote the additional number of transitions from time $T_{1}$ until the Markov chain next enters state $j$; then let $T_{3}$ denote the additional number of transitions from time $T_{1}+T_{2}$ until the Markov chain next enters state $j$, and so on. Note that $T_{1}$ is finite because Proposition before tells us that with probability 1 a transition into $j$ will eventually occur. Also, for $n \geq 2$, because $T_{n}$ is the number of transitions between the $(n-1)$ th and the $n$th transition into state $j$, it follows from the Markovian property that $T_{2}, T_{3}, \ldots$ are independent and identically distributed with mean $m_{j}$. Because the $n$th transition into state $j$ occurs at time $T_{1}+\cdots+T_{n}$ we obtain that $\pi_{i}$, the long-run proportion of time that chain is in the state $j$ is

$$
\pi=\lim _{n \rightarrow \infty} \frac{n}{\sum_{i=1}^{n} T_{i}}=\lim _{n \rightarrow \infty} \frac{1}{\frac{1}{n} \sum_{i=1}^{n} T_{i}}=\frac{1}{m_{j}}
$$

## Page 18 Proposition 4

Suppose that $i$ is positive recurrent and that $i \leftrightarrow j$. Now, let $n$ be such that $P_{i, j}^{n}>0$. Because $\pi_{i}$ is the long-run proportion of time that the chain is in state $i$, and $P_{i, j}^{n}$ is the long-run proportion of time when the Markov chain is in state $i$ that it will be in state $j$ after $n$ transitions
$\pi_{i} P_{i, j}^{n}=$ long-run proportion of time the chain is in $i$ and will be in $j$ after $n$ transitions
$=$ long-run proportion of time the chain is in $j$ and was in $i n$ transitions ago
$\leq$ long-run proportion of time the chain is in $j$

Hence, $\pi_{j} \geq \pi_{i} P_{i, j}^{n}>0$, showing that $j$ is positive recurrent.

## Page 20. Example 11.

The long run proportions $\pi_{i}, i=0,1,2$ are obtained by solving the set of equation in Theorem 1. In this case these equations are

$$
\begin{aligned}
& \pi_{0}=0.5 \pi_{0}+0.3 \pi_{1}+0.2 \pi_{2} \\
& \pi_{1}=0.4 \pi_{0}+0.4 \pi_{1}+0.3 \pi_{2} \\
& \pi_{2}=0.1 \pi_{0}+0.3 \pi_{1}+0.5 \pi_{2}
\end{aligned}
$$

or

$$
\boldsymbol{\pi}=\mathbf{P}^{T} \boldsymbol{\pi}
$$

with $\pi_{1}+\pi_{2}+\pi_{3}=1$ and $\pi=\left(\pi_{0}, \pi_{1}, \pi_{2}\right)^{T}$ solving yields

$$
\pi_{0}=\frac{21}{62}, \pi_{1}=\frac{23}{62}, \pi_{2}=\frac{18}{62}
$$

Page 20. Example 12. The long run proportions $\pi_{i}$ thus satisfy

$$
\begin{aligned}
& \pi_{0}=0.45 \pi_{0}+0.05 \pi_{1}+0.01 \pi_{2} \\
& \pi_{1}=0.48 \pi_{0}+0.70 \pi_{1}+0.50 \pi_{2} \\
& \pi_{2}=0.07 \pi_{0}+0.25 \pi_{1}+0.49 \pi_{2}
\end{aligned}
$$

with $\pi_{0}+\pi_{1}+\pi_{2}=1$. Hence

$$
\pi_{0}=0.07, \pi_{1}=0.62, \pi-2=0.31
$$

In other words, a society in which social mobility between classes can be described by a Markov chain with transition probability matrix given by the transition probability matrix has, in the long run, 7 percent of its people in upper-class jobs, 62 percent of its people in middle-class jobs, and 31 percent in lower-class jobs.

Page 21. Example 13. To begin, note that randomly choosing a parent and then randomly choosing one of its genes is equivalent to just randomly choosing a gene from the total gene population. By conditioning on the gene pair of the parent, we see that a randomly chosen gene will be type $A$ with probability

$$
P\{A\}=P\{A \mid A A\} p_{0}+P\{A \mid a a\} q_{0}+P\{A \mid A a\} r_{0}=p_{0}+r_{0} / 2
$$

similarly, it will be the type $a$ with probability

$$
P\{a\}=q_{0}+r_{0}
$$

Thus, Thus, under random mating a randomly chosen member of the next generation will be type $A A$ with probability p , where

$$
p=P\{A\} P\{A\}=\left(p_{0}+r_{0} / 2\right)^{2}
$$

Similarly, the randomly chosen member will be type $a a$ with probability

$$
q=P\{a\} P\{a\}=\left(q_{0}+r_{0} / 2\right)^{2}
$$

## Page 21. Example 13. Continuous

and will be type $A a$ with probability

$$
r=2 P\{A\} P\{a\}=2\left(p_{0}+r_{0} / 2\right)\left(q_{0}+r_{0} / 2\right)
$$

Since each member of the next generation will independently be of each of the three gene types with probabilities $p, q, r$, it follows that the percentages of the members of the next generation that are of type $A A, a a$, or $A a$ are respectively $p, q$, and $r$. If we now consider the total gene pool of this next generation, then $p+r / 2$, the fraction of its genes that are $A$, will be unchanged from the previous generation. This follows either by arguing that the total gene pool has not changed from generation to generation or by the following simple algebra:

$$
\begin{aligned}
p+r / 2 & =\left(p_{0}+r_{0} / 2\right)^{2}+\left(p_{0}+r_{0}\right)\left(q_{0}+r_{0} / 2\right) \\
& =\left(p_{0}+r_{0} / 2\right)\left(p_{0}+r_{0} / 2+q_{0}+r_{0} / 2\right) \\
& =p_{0}+r_{0} / 2=P\{A\}
\end{aligned}
$$

since $p_{0}+r_{0}+q_{0}=1$. Thus, the fractions of the gene pool that are A and a are the same as in the initial generation. From this it follows that, under random mating, in all successive generations after the initial one the percentages of the population having gene pairs $A A, a a$, and $A a$ will remain fixed at the values $\mathrm{p}, \mathrm{q}$, and r . This is known as the Hardy-Weinberg law.

## Page 21. Example 13. Continuous

Suppose now that the gene pair population has stabilized in the percentages $\mathrm{p}, \mathrm{q}, \mathrm{r}$, and let us follow the genetic history of a single individual and her descendants. (For simplicity, assume that each individual has exactly one offspring.) So, for a given individual, let $X_{n}$ denote the genetic state of her descendant in the nth generation. The transition probability matrix of this Markov chain, namely,

$$
\begin{array}{cccc} 
& A A & a a & A a \\
A A & p+\frac{r}{2} & 0 & q+\frac{r}{2} \\
a a & 0 & q+\frac{r}{2} & p+\frac{r}{2} \\
A a & \frac{p}{2}+\frac{r}{4} & \frac{q}{2}+\frac{r}{4} & \frac{p}{2}+\frac{q}{2}+\frac{r}{2}
\end{array}
$$

is easily verified by conditioning on the state of the randomly chosen mate. It is quite intuitive (why?) that the limiting probabilities for this Markov chain (which also equal the fractions of the individual's descendants that are in each of the three genetic states) should just be $p, q$, and $r$.

## Page 21. Example 13. Continuous

Because one of the equations in Theorem 1 is redundant, it suffices to show that

$$
\begin{aligned}
& p=p\left(p+\frac{r}{2}\right)+r\left(\frac{p}{2}+\frac{r}{4}\right)=\left(p+\frac{r}{2}\right)^{2} \\
& q=q\left(q+\frac{r}{2}\right)+r\left(\frac{q}{2}+\frac{r}{4}\right)=\left(q+\frac{r}{2}\right)^{2}
\end{aligned}
$$

But this follows from the above equation, and thus the result is established.

## Page 23. Example 14.

(a) To find $P_{i, j}$, suppose there are $i$ families checked into the hotel at the beginning of a day. Because each of these $i$ families will stay for another day with probability $q=1-p$ it follows that $R_{i}$, the number of these families that remain another day, is a binomial $(i, q)$ random variable. So, letting $N$ be the number of new families that check in that day, we see that

$$
P_{i, j}=P\left(R_{i}+N=j\right)
$$

Conditioning on $R_{i}$ and using that $N$ is Poisson with mean $\lambda$, we obtain

$$
\begin{aligned}
P_{i, j} & =\sum_{k=0}^{i} P\left(R_{i}+N=j \mid R_{i}=k\right) C_{i}^{k} q^{k} p^{i-k} \\
& =\sum_{k=0}^{i} P\left(N=j-k \mid R_{i}=k\right) C_{i}^{k} q^{k} p^{i-k} \\
& =\sum_{k=0}^{\min (i, j)} P\left(N=j-k \mid R_{i}=k\right) C_{i}^{k} q^{k} p^{i-k} \\
& =\sum_{k=0}^{\min (i, j)} e^{-\lambda} \frac{\lambda^{j-k}}{(j-k)!} C_{i}^{k} q^{k} p^{i-k}
\end{aligned}
$$

## Page 23. Example 14. Continuous

(b) Using the preceding representation $R_{i}+N$ for the next state from state $i$, we see that

$$
\mathrm{E}\left[X_{n} \mid X_{n-1}=i\right]=\mathrm{E}\left[R_{i}+N\right]=i q+\lambda
$$

Consequently

$$
\mathrm{E}\left[X_{n} \mid X_{n-1}\right]=X_{n-1} q+\lambda
$$

Iterating the preceding gives

$$
\begin{aligned}
\mathrm{E}\left[X_{n}\right] & =\lambda+q \mathrm{E}\left[X_{n-1}\right]=\lambda+q\left(\lambda+q \mathrm{E}\left[X_{n-2}\right]\right) \\
& =\cdots \cdots=\lambda\left(1+q+q^{2}+\cdots+q^{n-1}\right)+q^{n} \mathrm{E}\left[X_{0}\right]
\end{aligned}
$$

and yielding the result

$$
\mathrm{E}\left[X_{n} \mid X_{0}=i\right]=\frac{\lambda\left(1-q^{n}\right)}{p}+q^{n} i
$$

## Page 23. Example 14. Continuous

(c) To find the stationary probabilities we will not directly use the complicated transition probabilities derived in part (a). Rather we will make use of the fact that the stationary probability distribution is the only distribution on the initial state that results in the next state having the same distribution. Now, suppose that the initial state $X_{0}$ has a Poisson distribution with mean $\alpha$. That is, assume that the number of families initially in the hotel is Poisson with mean $\alpha$. Let $R$ denote the number of these families that remain in the hotel at the beginning of the next day. Then, using the result that if each of a Poisson distributed (with mean $\alpha$ ) number of events occurs with probability $q$, then the total number of these events that occur is Poisson distributed with mean $\alpha q$, it follows that $R$ is a Poisson random variable with mean $\alpha q$. In addition, the number of new families that check in during the day, call it $N$, is Poisson with mean $\lambda$, and is independent of $R$. Hence, since the sum of independent Poisson random variables is also Poisson distributed, it follows that $R+N$, the number of guests at the beginning of the next day, is Poisson with mean $\lambda+\alpha q$. Consequently, if we choose $\alpha$ so that

$$
\alpha=\lambda+\alpha q
$$

then the distribution of $X_{1}$ would be the same as that of $X_{0}$. But this means that when the initial distribution of $X_{0}$ is Poisson with mean $\alpha=\frac{\lambda}{p}$, then so is the distribution of $X_{1}$, implying that this is the stationary distribution. That is, the stationary probabilities are

$$
\pi_{i}=e^{-\lambda / p}(\lambda / p)^{i} / i!, i \geq 0
$$

## Page 24. Proposition 3.

Proof: If we let $a_{j}(N)$ be the amount of time the Markov chain spends in state $j$ during time periods $1, \ldots, N$, then

$$
\sum_{n=1}^{N} r\left(X_{n}\right)=\sum_{j=0}^{\infty} a_{j}(N) r(j)
$$

Since $a_{j}(N) / N \rightarrow \pi_{j}$ the result follows from the preceding upon dividing by $N$ and then letting $N \rightarrow \infty$.

## page 27. The Gambler's Ruin Problem

If we let $X_{n}$ denote the player's fortune at time $n$, then the process $\left\{X_{n}, n=\right.$ $0,1,2, \ldots\}$ is a Markov chain with transition probabilities

$$
\begin{gathered}
P_{00}=P_{N N}=1 \\
P_{i, i+1}=p=1-P_{i, i-1}, i=1, \ldots, N-1
\end{gathered}
$$

This Markov chain has three classes, namely, $\{0\},\{1,2, \ldots, N-1\}$, and $\{N\}$; the first and third class being recurrent and the second transient. Since each transient state is visited only finitely often, it follows that, after some finite amount of time, the gambler will either attain his goal of N or go broke.

Let $P_{i}, i=0,1, \ldots, N$, denote the probability that, starting with $i$, the gambler's fortune will eventually reach $N$. By conditioning on the outcome of the initial play of the game we obtain

$$
P_{i}=p P_{i+1}+q P_{i-1}, i=1,2, \ldots, N-1
$$

or equivalently since $p+q=1$
$p P_{i}+q P_{i}=p P_{i+1}+q P_{i-1} \quad$ or $\quad P_{i+1}-P_{i}=\frac{q}{p}\left(P_{i}-P_{i-1}\right), i=1, \ldots, N-1$
page 27. The Gambler's Ruin Problem, continuous Hence, since $P_{0}=0$, we obtain from the preceding line that

$$
\begin{gathered}
P_{2}-P_{1}=\frac{q}{p}\left(P_{1}-P_{0}\right)=\frac{q}{p} p_{1} \\
P_{3}-P_{2}=\frac{q}{p}\left(P_{2}-P_{1}\right)=\left(\frac{q}{p}\right)^{2} p_{1} \\
\ldots \ldots \\
P_{i}-P_{i-1}=\frac{q}{p}\left(P_{i-1}-P_{i-2}\right)=\left(\frac{q}{p}\right)^{i-1} p_{1} \\
\ldots \ldots \\
P_{N}-P_{N-1}=\frac{q}{p}\left(P_{N-1}-P_{N-2}\right)=\left(\frac{q}{p}\right)^{N-1} p_{1}
\end{gathered}
$$

Adding the first $i-1$ of these equations yields

$$
P_{i}-P_{1}=P_{1}\left[\left(\frac{q}{p}\right)+\left(\frac{q}{p}\right)^{2}+\cdots+\left(\frac{q}{p}\right)^{i-1}\right]
$$

## page 27. The Gambler's Ruin Problem, continuous

 or$$
P_{i}= \begin{cases}\frac{1-(q / p)^{i}}{1-(q / p)} P_{1}, & \text { if } \frac{q}{p} \neq 1 \\ i p_{1}, & \text { if } \frac{q}{p}=1\end{cases}
$$

Now, using the fact that $P_{N}=1$, we obtain

$$
P_{1}= \begin{cases}\frac{1-(q / p)}{1-(q / p)^{N}}, & \text { if } \frac{q}{p} \neq 1 \\ \frac{1}{N}, & \text { if } \frac{q}{p}=1\end{cases}
$$

Hence

$$
P_{i}= \begin{cases}\frac{1-(q / p)^{i}}{1-(q / p)^{N}}, & \text { if } \frac{q}{p} \neq 1 \\ \frac{i}{N}, & \text { if } \frac{q}{p}=1\end{cases}
$$

Note that, as $N \rightarrow \infty$,

$$
P_{i}= \begin{cases}1-\left(\frac{q}{p}\right)^{i}, & \text { if } \frac{q}{p}>\frac{1}{2} \\ 0, & \text { if } \frac{q}{p} \leq \frac{1}{2}\end{cases}
$$

## Page 27. Example 15.

(a) The desired probability is obtained from the equation above by letting $i=5, N=15$, and $p=0.6$. Hence, the desired probability is

$$
\frac{1-\left(\frac{2}{3}\right)^{5}}{1-\left(\frac{2}{3}\right)^{15}} \approx 0.87
$$

(b) the desired probability is

$$
\frac{1-\left(\frac{2}{3}\right)^{10}}{1-\left(\frac{2}{3}\right)^{30}} \approx 0.98
$$

## Page 28. A Model for Algorithm Efficiency

To obtain a feel for whether or not the preceding statement is surprising, let us consider a simple probabilistic (Markov chain) model as to how the algorithm moves along the extreme points. Specifically, we will suppose that if at any time the algorithm is at the $j$ th best extreme point then after the next pivot the resulting extreme point is equally likely to be any of the $j-1$ best. Under this assumption, we show that the time to get from the $N$ th best to the best extreme point has approximately, for large $N$, a normal distribution with mean and variance equal to the logarithm (base e) of N .

## Page 28. A Model for Algorithm Efficiency, continuous

Consider a Markov chain for which $P_{11}=1$ and

$$
P_{i j}=\frac{1}{i-1}, j=1, \ldots, i-1, i>1
$$

and let $T_{i}$ denote the number of transitions needed to go from state $i$ to state 1. A recursive formula for $\mathrm{E}\left[T_{i}\right]$ can be obtained by conditioning on the initial transition:

$$
\mathrm{E}\left[T_{i}\right]=1+\frac{1}{i-1} \sum_{j=1}^{i-1} \mathrm{E}\left[T_{i}\right]
$$

Starting with $\mathrm{E}\left[T_{1}\right]=0$, we successively see that

$$
\mathrm{E}\left[T_{2}\right]=1, \mathrm{E}\left[T_{3}\right]=1+\frac{1}{2}, \mathrm{E}\left[T_{4}\right]=1+\frac{1}{3}\left(1+1+\frac{1}{2}\right)=1+\frac{1}{2}+\frac{1}{3}
$$

and it is not difficult to guess and then prove inductively that

$$
\mathrm{E}\left[T_{i}\right]=\sum_{j=1}^{i-1} 1 / j
$$

## Page 28. A Model for Algorithm Efficiency, continuous

However, to obtain a more complete description of $T_{N}$, we will use the representation

$$
T_{N}=\sum_{j=1}^{N-1} I_{j}
$$

where $I_{j}=1$ if the process ever enters $j$ otherwise 0 . The importance of the preceding representation stems from the following: $I_{1}, \ldots, I_{N-1}$ are independent and

$$
P_{I_{j}=1}=1 / j, 1 \leq j \leq N-1
$$

Given $I_{j+1}, \ldots, I_{N}$, let $n=\min \left\{i: i>j, I_{i}=1\right\}$ denote the lowest numbered state, greater than j , that is entered. Thus we know that the process enters state n and the next state entered is one of the states $1,2, \ldots, j$. Hence, as the next state from state $n$ is equally likely to be any of the lower number states $1,2, \ldots, n-1$ we see that

$$
P\left\{I_{j}=1 \mid I_{j+1}, \ldots, I_{N}\right\}=\frac{1 /(n-1)}{j /(n-1)}=1 / j
$$

Hence, $P\left\{I_{j}=1\right\}=1 / j$, and independence follows since the preceding conditional probability does not depend on $I_{j+1}, \ldots, I_{N}$.

## Page 28. A Model for Algorithm Efficiency, continuous

Since $\log N \approx \sum_{j=1}^{N-1} 1 / j$ when $N \rightarrow \infty$, we have
(i) $\mathrm{E}\left[T_{N}\right]=\sum_{j=1}^{N-1} 1 / j$
(ii) $\operatorname{Var}\left(T_{N}\right)=\sum_{j=1}^{N-1}(1 / j)(1-1 / j)$
(iii) For $N$ large, $T_{N}$ has approximately a normal distribution with mean $\log N$ and variance $\log N$.

## Page 32 Example 16.

The matrix PT, which specifies $P_{i j}, i, j \in\{1,2,3,4,5,6\}$, is as follows:

$$
\mathbf{P}_{T}=\left\|\begin{array}{cccccc}
0 & 0.4 & 0 & 0 & 0 & 0 \\
0.6 & 0 & 0.4 & 0 & 0 & 0 \\
0 & 0.6 & 0 & 0.4 & 0 & 0 \\
0 & 0 & 0.6 & 0 & 0.4 & 0 \\
0 & 0 & 0 & 0.6 & 0 & 0.4 \\
0 & 0 & 0 & 0 & 0.6 & 0
\end{array}\right\|
$$

Inverting $\mathbf{I}-\mathbf{P}_{T}$ gives

$$
\mathbf{S}=\left(\mathbf{I}-\mathbf{P}_{T}\right)^{-1}=\left[\begin{array}{llllll}
1.6149 & 1.0248 & 0.6314 & 0.3691 & 0.1943 & 0.0777 \\
1.5372 & 2.5619 & 1.5784 & 0.9228 & 0.4857 & 0.1943 \\
1.4206 & 2.3677 & 2.9990 & 1.7533 & 0.9228 & 0.3691 \\
1.2458 & 2.0763 & 2.6299 & 2.9990 & 1.5784 & 0.6314 \\
0.9835 & 1.6391 & 2.0763 & 2.3677 & 2.5619 & 1.0248 \\
0.5901 & 0.9835 & 1.2458 & 1.4206 & 1.5372 & 1.6149
\end{array}\right]
$$

Hence (a) $s_{3,5}=0.9228$, (b) $S_{3,2}=2.3677$

## Page 32 Example 16. Continuous

(c) Since $s_{3,1}=1.4206$ and $s_{1,1}=1.6149$, then

$$
f_{3,1}=\frac{s_{3,1}}{s_{1,1}}=0.8797
$$

As a check, note that $f_{3,1}$ is just the probability that a gambler starting with 3 reaches 1 before 7 . That is, it is the probability that the gambler's fortune will go down 2 before going up 4; which is the probability that a gambler starting with 2 will go broke before reaching 6 . Therefore,

$$
f_{3,1}=1-\frac{1-(0.6 / 0.4)^{2}}{1-(0.6 / 0.4)^{6}}=0.8797
$$

which checks with our earlier answer.

## Page 38 Example 17.

(a) Since $\mu=\frac{3}{4} \leq 1$, it follows that $\pi_{0}=1$.
(b) $\pi_{0}$ satisfies

$$
\pi_{0}=\frac{1}{4}+\frac{1}{4} \pi_{0}+\frac{1}{2} \pi_{0}^{2}
$$

or

$$
2 \pi_{0}^{2}-3 \pi_{0}+1=0
$$

The smallest positive solution of this quadratic equation is $\pi_{0}=\frac{1}{2}$.
(c) Since the population will die out if and only if the families of each of the members of the initial generation die out, the desired probability is $\pi_{0}^{n}$. For (a) this yields $\pi_{0}^{n}=1$, and for (b), $\pi_{0}^{n}=\frac{1}{2^{n}}$.

## Page 43 Example 18.

If $X$ represents the amount of time that the customer spends in the bank, then the first probability is just

$$
P\{X>15\}=e^{-15 \lambda}=e^{-3 / 2} \approx 0.223
$$

The second question asks for the probability that a customer who has spent ten minutes in the bank will have to spend at least five more minutes. However, since the exponential distribution does not "remember" that the customer has already spent ten minutes in the bank, this must equal the probability that an entering customer spends at least five minutes in the bank. That is, the desired probability is just

$$
P\{X>5\}=e^{-5 \lambda}=e^{-1 / 2} \approx 0.607
$$

## Page 46 Example 19.

(a) Because the number of arrivals in 7 days is Poisson with mean $7 \lambda=14$, it follows that the probability there will be 10 arrivals is $e^{-14}(14)^{10} / 10$ !.
(b) $\mathrm{E}\left[S_{20}\right]=20 / \lambda=10$.

## Page 49 Example 21.

We start by deriving an equation for $M(t+h)$ by conditioning on $X(t)$. This yields

$$
M(t+h)=\mathrm{E}[X(t+h)]=\mathrm{E}[\mathrm{E}[X(t+h) \mid X(t)]]
$$

Now, given the size of the population at time t then, ignoring events whose probability is $o(h)$, the population at time $t+h$ will either increase in size by 1 if a birth or an immigration occurs in $(t, t+h)$, or decrease by 1 if a death occurs in this interval, or remain the same if neither of these two possibilities occurs. That is, given $X(t)$
$X(t+h)= \begin{cases}X(t)+1, & \text { with probability }[\theta+X(t) \lambda] h+o(h) \\ X(t)-1, & \text { with probability } X(t) \mu h+o(h) \\ X(t), & \text { with probability } 1-[\theta+X(t) \lambda+X(t) \mu] h+o(h)\end{cases}$

## Page 49 Example 21. continuous

Therefore

$$
\mathrm{E}[X(t+h) \mid X(t)]=X(t)+[\theta+X(t) \lambda-X(t) \mu] h+o(h)
$$

Taking expectations yields

$$
M(t+h)=M(t)+(\lambda-\mu) M(t) h+\theta h+o(h)
$$

or, equivalently

$$
\frac{M(t+h)-M(t)}{h}=(\lambda-\mu) M(t)+\theta+\frac{o(h)}{h}
$$

Taking the limit as $h \rightarrow 0$ yields the differential equation

$$
M^{\prime}(t)=(\lambda-\mu) M(t)+\theta
$$

## Page 49 Example 21. continuous

If we now define the function $h(t)$ by

$$
h(t)=(\lambda-\mu) M(t)+\theta
$$

then

$$
h^{\prime}(t)=(\lambda-\mu) M^{\prime}(t)
$$

Therefore, the differential equation above can be rewritten as

$$
\frac{h^{\prime}(t)}{\lambda-\mu}=h(t) \quad \text { or } \quad \frac{h^{\prime}(t)}{h(t)}=\lambda-\mu
$$

Integration yields

$$
\log [h(t)]=(\lambda-\mu) t+c, \quad \text { or } \quad h(t)=K e^{(\lambda-\mu) t}
$$

Putting this back in terms of $M(t)$ gives

$$
\theta+(\lambda-\mu) M(t)=K e^{(\lambda-\mu) t}
$$

## Page 49 Example 21. continuous

To determine the value of the constant $K$, we use the fact that $M(0)=i$ and evaluate the preceding at $t=0$. This gives

$$
\theta+(\lambda-\mu) i=K
$$

Substituting this back in the preceding equation for $M(t)$ yields the following solution for $M(t)$ :

$$
M(t)=\frac{\theta}{\lambda-\mu}\left[e^{(\lambda-\mu) t}-1\right]+i e^{(\lambda-\mu) t}
$$

Note that we have implicitly assumed that $\lambda=\mu$. If $\lambda=\mu$, then differential equation above reduces to

$$
M^{\prime}(t)=\theta
$$

Integrating and using that $M(0)=i$ gives the solution

$$
M(t)=\theta t+i
$$

