

## 16.6: Stationary and Limiting Distributions of Discrete-Time Chains

In this section, we study some of the deepest and most interesting parts of the theory of discrete-time Markov chains, involving two different but complementary ideas: stationary distributions and limiting distributions. The theory of renewal processes plays a critical role.

### Basic Theory

As usual, our starting point is a (time homogeneous) discrete-time Markov chain  $\mathbf{X} = (X_0, X_1, X_2, \dots)$  with (countable) state space  $S$  and transition probability matrix  $P$ . In the background, of course, is a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  so that  $\Omega$  is the sample space,  $\mathcal{F}$  the  $\sigma$ -algebra of events, and  $\mathbb{P}$  the probability measure on  $(\Omega, \mathcal{F})$ . For  $n \in \mathbb{N}$ , let  $\mathcal{F}_n = \sigma\{X_0, X_1, \dots, X_n\}$ , the  $\sigma$ -algebra of events determined by the chain up to time  $n$ , so that  $\mathcal{F} = \{\mathcal{F}_0, \mathcal{F}_1, \dots\}$  is the natural filtration associated with  $\mathbf{X}$ .

### The Embedded Renewal Process

Let  $y \in S$  and  $n \in \mathbb{N}_+$ . We will denote the number of visits to  $y$  during the first  $n$  positive time units by

$$N_{y,n} = \sum_{i=1}^n \mathbf{1}(X_i = y) \quad (16.6.1)$$

Note that  $N_{y,n} \rightarrow N_y$  as  $n \rightarrow \infty$ , where

$$N_y = \sum_{i=1}^{\infty} \mathbf{1}(X_i = y) \quad (16.6.2)$$

is the total number of visits to  $y$  at positive times, one of the important random variables that we studied in the section on transience and recurrence. For  $n \in \mathbb{N}_+$ , we denote the time of the  $n$ th visit to  $y$  by

$$\tau_{y,n} = \min\{k \in \mathbb{N}_+ : N_{y,k} = n\} \quad (16.6.3)$$

where as usual, we define  $\min(\emptyset) = \infty$ . Note that  $\tau_{y,1}$  is the time of the first visit to  $y$ , which we denoted simply by  $\tau_y$  in the section on transience and recurrence. The times of the visits to  $y$  are stopping times for  $\mathbf{X}$ . That is,  $\{\tau_{y,n} = k\} \in \mathcal{F}_k$  for  $n \in \mathbb{N}_+$  and  $k \in \mathbb{N}$ . Recall also the definition of the hitting probability to state  $y$  starting in state  $x$ :

$$H(x, y) = \mathbb{P}(\tau_y < \infty \mid X_0 = x), \quad (x, y) \in S^2 \quad (16.6.4)$$

Suppose that  $x, y \in S$ , and that  $y$  is recurrent and  $X_0 = x$ .

1. If  $x = y$ , then the successive visits to  $y$  form a renewal process.
2. If  $x \neq y$  but  $x \rightarrow y$ , then the successive visits to  $y$  form a delayed renewal process.

**Proof**

Let  $\tau_{y,0} = 0$  for convenience.

1. Given  $X_0 = y$ , the sequence  $(\tau_{y,1}, \tau_{y,2}, \dots)$  is the sequence of arrival times of a renewal process. Every time the chain reaches state  $y$ , the process starts over, independently of the past, by the Markov property. Thus the interarrival times  $\tau_{y,n+1} - \tau_{y,n}$  for  $n \in \mathbb{N}$  are conditionally independent, and are identically distributed, given  $X_0 = y$ .
2. If  $x \neq y$  but  $x \rightarrow y$ , then given  $X_0 = x$ , the sequence  $(\tau_{y,1}, \tau_{y,2}, \dots)$  is the sequence of arrival times of a delayed renewal process. By the same argument as in (a), the interarrival times  $\tau_{y,n+1} - \tau_{y,n}$  for  $n \in \mathbb{N}$  are conditionally independent, given  $X_0 = x$ , and all but  $\tau_{y,1}$  have the same distribution.

As noted in the proof,  $(\tau_{y,1}, \tau_{y,2}, \dots)$  is the sequence of arrival times and  $(N_{y,1}, N_{y,2}, \dots)$  is the associated sequence of counting variables for the embedded renewal process associated with the recurrent state  $y$ . The corresponding renewal function, given  $X_0 = x$ , is the function  $n \mapsto G_n(x, y)$  where

$$G_n(x, y) = \mathbb{E}(N_{y,n} \mid X_0 = x) = \sum_{k=1}^n P^k(x, y), \quad n \in \mathbb{N} \quad (16.6.5)$$

Thus  $G_n(x, y)$  is the expected number of visits to  $y$  in the first  $n$  positive time units, starting in state  $x$ . Note that  $G_n(x, y) \rightarrow G(x, y)$  as  $n \rightarrow \infty$  where  $G$  is the potential matrix that we studied previously. This matrix gives the expected total number visits to state  $y \in S$ , at positive times, starting in state  $x \in S$ :

$$G(x, y) = \mathbb{E}(N_y | X_0 = x) = \sum_{k=1}^{\infty} P^k(x, y) \quad (16.6.6)$$

### Limiting Behavior

The limit theorems of renewal theory can now be used to explore the limiting behavior of the Markov chain. Let  $\mu(y) = \mathbb{E}(\tau_y | X_0 = y)$  denote the mean return time to state  $y$ , starting in  $y$ . In the following results, it may be the case that  $\mu(y) = \infty$ , in which case we interpret  $1/\mu(y)$  as 0.

If  $x, y \in S$  and  $y$  is recurrent then

$$\mathbb{P}\left(\frac{1}{n}N_{y,n} \rightarrow \frac{1}{\mu(y)} \text{ as } n \rightarrow \infty \mid X_0 = x\right) = H(x, y) \quad (16.6.7)$$

Proof

This result follows from the strong law of large numbers for renewal processes.

Note that  $\frac{1}{n}N_{y,n} = \frac{1}{n} \sum_{k=1}^n \mathbf{1}(X_k = y)$  is the average number of visits to  $y$  in the first  $n$  positive time units.

If  $x, y \in S$  and  $y$  is recurrent then

$$\frac{1}{n}G_n(x, y) = \frac{1}{n} \sum_{k=1}^n P^k(x, y) \rightarrow \frac{H(x, y)}{\mu(y)} \text{ as } n \rightarrow \infty \quad (16.6.8)$$

Proof

This result follows from the elementary renewal theorem for renewal processes.

Note that  $\frac{1}{n}G_n(x, y) = \frac{1}{n} \sum_{k=1}^n P^k(x, y)$  is the expected average number of visits to  $y$  during the first  $n$  positive time units, starting at  $x$ .

If  $x, y \in S$  and  $y$  is recurrent and aperiodic then

$$P^n(x, y) \rightarrow \frac{H(x, y)}{\mu(y)} \text{ as } n \rightarrow \infty \quad (16.6.9)$$

Proof

This result follows from the renewal theorem for renewal processes.

Note that  $H(y, y) = 1$  by the very definition of a recurrent state. Thus, when  $x = y$ , the law of large numbers [above](#) gives convergence with probability 1, and the [first](#) and [second](#) renewal theory limits above are simply  $1/\mu(y)$ . By contrast, we already know the corresponding limiting behavior when  $y$  is transient.

If  $x, y \in S$  and  $y$  is transient then

1.  $\mathbb{P}\left(\frac{1}{n}N_{y,n} \rightarrow 0 \text{ as } n \rightarrow \infty \mid X_0 = x\right) = 1$
2.  $\frac{1}{n}G_n(x, y) = \frac{1}{n} \sum_{k=1}^n P^k(x, y) \rightarrow 0 \text{ as } n \rightarrow \infty$
3.  $P^n(x, y) \rightarrow 0 \text{ as } n \rightarrow \infty$

Proof

1. Note that  $0 \leq \frac{1}{n}N_{y,n} \leq \frac{1}{n}N_y$ . But if  $y$  is transient,  $\mathbb{P}(N_y < \infty \mid X_0 = x) = 1$  and hence  $\mathbb{P}\left(\frac{1}{n}N_y \rightarrow 0 \text{ as } n \rightarrow \infty \mid X_0 = x\right) = 1$  so the result follows from the squeeze theorem for limits.
2. Similarly, note that

$$0 \leq \frac{1}{n} \sum_{k=1}^n P^k(x, y) \leq \frac{1}{n} \sum_{k=1}^{\infty} P^k(x, y) \quad (16.6.10)$$

If  $y$  is transient,  $G(x, y) = \sum_{k=1}^{\infty} P^k(x, y) < \infty$  and hence  $\frac{1}{n} G(x, y) \rightarrow 0$  as  $n \rightarrow \infty$ . Again the result follows from the squeeze theorem for limits.

3. Once more, if  $y$  is transient,  $G(x, y) = \sum_{k=1}^{\infty} P^k(x, y) < \infty$  and therefore  $P^n(x, y) \rightarrow 0$  as  $n \rightarrow \infty$ .

On the other hand, if  $y$  is transient then  $\mathbb{P}(\tau_y = \infty \mid X_0 = y) > 0$  by the very definition of a transience. Thus  $\mu(y) = \infty$ , and so the results in parts (b) and (c) agree with the corresponding results above for a recurrent state. Here is a summary.

For  $x, y \in S$ ,

$$\frac{1}{n} G_n(x, y) = \frac{1}{n} \sum_{k=1}^n P^k(x, y) \rightarrow \frac{H(x, y)}{\mu(y)} \text{ as } n \rightarrow \infty \quad (16.6.11)$$

If  $y$  is transient or if  $y$  is recurrent and aperiodic,

$$P^n(x, y) \rightarrow \frac{H(x, y)}{\mu(y)} \text{ as } n \rightarrow \infty \quad (16.6.12)$$

### Positive and Null Recurrence

Clearly there is a fundamental dichotomy in terms of the limiting behavior of the chain, depending on whether the mean return time to a given state is finite or infinite. Thus the following definition is natural.

Let  $x \in S$ .

1. State  $x$  is *positive recurrent* if  $\mu(x) < \infty$ .
2. If  $x$  is recurrent but  $\mu(x) = \infty$  then state  $x$  is *null recurrent*.

Implicit in the definition is the following simple result:

If  $x \in S$  is positive recurrent, then  $x$  is recurrent.

**Proof**

Recall that if  $\mathbb{E}(\tau_x \mid X_0 = x) < \infty$  then  $\mathbb{P}(\tau_x < \infty \mid X_0 = x) = 1$ .

On the other hand, it is possible to have  $\mathbb{P}(\tau_x < \infty \mid X_0 = x) = 1$ , so that  $x$  is recurrent, and also  $\mathbb{E}(\tau_x \mid X_0 = x) = \infty$ , so that  $x$  is null recurrent. Simply put, a random variable can be finite with probability 1, but can have infinite expected value. A classic example is the Pareto distribution with shape parameter  $a \in (0, 1)$ .

Like recurrence/transience, and period, the null/positive recurrence property is a *class property*.

If  $x$  is positive recurrent and  $x \rightarrow y$  then  $y$  is positive recurrent.

**Proof**

Suppose that  $x$  is positive recurrent and  $x \rightarrow y$ . Recall that  $y$  is recurrent and  $y \rightarrow x$ . Hence there exist  $i, j \in \mathbb{N}_+$  such that  $P^i(x, y) > 0$  and  $P^j(y, x) > 0$ . Thus for every  $k \in \mathbb{N}_+$ ,  $P^{i+j+k}(y, y) \geq P^j(y, x) P^k(x, x) P^i(x, y)$ . Averaging over  $k$  from 1 to  $n$  gives

$$\frac{G_n(y, y)}{n} - \frac{G_{i+j}(y, y)}{n} \geq P^j(y, x) \frac{G_n(x, x)}{n} P^i(x, y) \quad (16.6.13)$$

Letting  $n \rightarrow \infty$  and using renewal theory limit [above](#) gives

$$\frac{1}{\mu(y)} \geq P^j(y, x) \frac{1}{\mu(x)} P^i(x, y) > 0 \quad (16.6.14)$$

Therefore  $\mu(y) < \infty$  and so  $y$  is also positive recurrent.

Thus, the terms *positive recurrent* and *null recurrent* can be applied to equivalence classes (under the *to and from* equivalence relation), as well as individual states. When the chain is irreducible, the terms can be applied to the chain as a whole.

Recall that a nonempty set of states  $A$  is *closed* if  $x \in A$  and  $x \rightarrow y$  implies  $y \in A$ . Here are some simple results for a finite, closed set of states.

If  $A \subseteq S$  is finite and closed, then  $A$  contains a positive recurrent state.

**Proof**

Fix a state  $x \in A$  and note that  $P^k(x, A) = \sum_{y \in A} P^k(x, y) = 1$  for every  $k \in \mathbb{N}_+$  since  $A$  is closed. Averaging over  $k$  from 1 to  $n$  gives

$$\sum_{y \in A} \frac{G_n(x, y)}{n} = 1 \quad (16.6.15)$$

for every  $n \in \mathbb{N}_+$ . Note that the change in the order of summation is justified since both sums are finite. Assume now that all states in  $A$  are transient or null recurrent. Letting  $n \rightarrow \infty$  in the displayed equation gives the contradiction  $0 = 1$ . Again, the interchange of sum and limit is justified by the fact that  $A$  is finite.

If  $A \subseteq S$  is finite and closed, then  $A$  contains no null recurrent states.

**Proof**

Let  $x \in A$ . Note that  $[x] \subseteq A$  since  $A$  is closed. Suppose that  $x$  is recurrent. Note that  $[x]$  is also closed and finite and hence must have a positive recurrent state by the [previous result](#). Hence the equivalence class  $[x]$  is positive recurrent and thus so is  $x$ .

If  $A \subseteq S$  is finite and irreducible, then  $A$  is a positive recurrent equivalence class.

**Proof**

We already know that  $A$  is a recurrent equivalence class, from our study of transience and recurrence. From the previous theorem,  $A$  is positive recurrent.

In particular, a Markov chain with a finite state space cannot have null recurrent states; every state must be transient or positive recurrent.

### Limiting Behavior, Revisited

Returning to the limiting behavior, suppose that the chain  $\mathbf{X}$  is irreducible, so that either all states are transient, all states are null recurrent, or all states are positive recurrent. From the basic limit theorem [above](#), if the chain is transient or if the chain is recurrent and aperiodic, then

$$P^n(x, y) \rightarrow \frac{1}{\mu(y)} \text{ as } n \rightarrow \infty \text{ for every } x \in S \quad (16.6.16)$$

Note in particular that the limit is independent of the initial state  $x$ . Of course in the transient case and in the null recurrent and aperiodic case, the limit is 0. Only in the positive recurrent, aperiodic case is the limit positive, which motivates our next definition.

A Markov chain  $\mathbf{X}$  that is irreducible, positive recurrent, and aperiodic, is said to be *ergodic*.

In the ergodic case, as we will see,  $X_n$  has a limiting distribution as  $n \rightarrow \infty$  that is independent of the initial distribution.

The behavior when the chain is periodic with period  $d \in \{2, 3, \dots\}$  is a bit more complicated, but we can understand this behavior by considering the  $d$ -step chain  $\mathbf{X}_d = (X_0, X_d, X_{2d}, \dots)$  that has transition matrix  $P^d$ . Essentially, this allows us to trade *periodicity* (one form of complexity) for *reducibility* (another form of complexity). Specifically, recall that the  $d$ -step chain is aperiodic but has  $d$  equivalence classes  $(A_0, A_1, \dots, A_{d-1})$ ; and these are the *cyclic classes* of original chain  $\mathbf{X}$ .

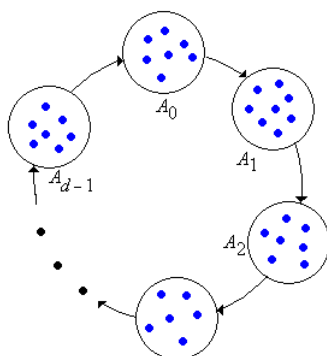


Figure 16.6.1: The cyclic classes of a chain with period  $d$

The mean return time to state  $x$  for the  $d$ -step chain  $\mathbf{X}_d$  is  $\mu_d(x) = \mu(x)/d$ .

Proof

Note that every single step for the  $d$ -step chain corresponds to  $d$  steps for the original chain.

Let  $i, j, k \in \{0, 1, \dots, d-1\}$ ,

1.  $P^{nd+k}(x, y) \rightarrow d/\mu(y)$  as  $n \rightarrow \infty$  if  $x \in A_i$  and  $y \in A_j$  and  $j = (i+k) \bmod d$ .
2.  $P^{nd+k}(x, y) \rightarrow 0$  as  $n \rightarrow \infty$  in all other cases.

Proof

These results follow from the previous theorem and the cyclic behavior of the chain.

If  $y \in S$  is null recurrent or transient then regardless of the period of  $y$ ,  $P^n(x, y) \rightarrow 0$  as  $n \rightarrow \infty$  for every  $x \in S$ .

### Invariant Distributions

Our next goal is to see how the limiting behavior is related to invariant distributions. Suppose that  $f$  is a probability density function on the state space  $S$ . Recall that  $f$  is *invariant* for  $P$  (and for the chain  $\mathbf{X}$ ) if  $fP = f$ . It follows immediately that  $fP^n = f$  for every  $n \in \mathbb{N}$ . Thus, if  $X_0$  has probability density function  $f$  then so does  $X_n$  for each  $n \in \mathbb{N}$ , and hence  $\mathbf{X}$  is a sequence of identically distributed random variables. A bit more generally, suppose that  $g: S \rightarrow [0, \infty)$  is invariant for  $P$ , and let  $C = \sum_{x \in S} g(x)$ . If  $0 < C < \infty$  then  $f$  defined by  $f(x) = g(x)/C$  for  $x \in S$  is an invariant probability density function.

Suppose that  $g: S \rightarrow [0, \infty)$  is invariant for  $P$  and satisfies  $\sum_{x \in S} g(x) < \infty$ . Then

$$g(y) = \frac{1}{\mu(y)} \sum_{x \in S} g(x)H(x, y), \quad y \in S \quad (16.6.17)$$

Proof

Recall again that  $gP^k = g$  for every  $k \in \mathbb{N}$  since  $g$  is invariant for  $P$ . Averaging over  $k$  from 1 to  $n$  gives  $gG_n/n = g$  for each  $n \in \mathbb{N}_+$ . Explicitly,

$$\sum_{x \in S} g(x) \frac{G_n(x, y)}{n} = g(y), \quad y \in S \quad (16.6.18)$$

Letting  $n \rightarrow \infty$  and using the limit theorem [above](#) gives the result. The dominated convergence theorem justifies interchanging the limit with the sum, since the terms are positive,  $\frac{1}{n}G_n(x, y) \leq 1$ , and  $\sum_{x \in S} g(x) < \infty$ .

Note that if  $y$  is transient or null recurrent, then  $g(y) = 0$ . Thus, an invariant function with finite sum, and in particular an invariant probability density function must be concentrated on the positive recurrent states.

Suppose now that the chain  $\mathbf{X}$  is irreducible. If  $\mathbf{X}$  is transient or null recurrent, then from the previous result, the only nonnegative functions that are invariant for  $P$  are functions that satisfy  $\sum_{x \in S} g(x) = \infty$  and the function that is identically 0:  $g = \mathbf{0}$ . In particular, the chain does not have an invariant distribution. On the other hand, if the chain is positive recurrent, then  $H(x, y) = 1$

for all  $x, y \in S$ . Thus, from the previous result, the only possible invariant probability density function is the function  $f$  given by  $f(x) = 1/\mu(x)$  for  $x \in S$ . Any other nonnegative function  $g$  that is invariant for  $P$  and has finite sum, is a multiple of  $f$  (and indeed the multiple is sum of the values). Our next goal is to show that  $f$  really is an invariant probability density function.

If  $\mathbf{X}$  is an irreducible, positive recurrent chain then the function  $f$  given by  $f(x) = 1/\mu(x)$  for  $x \in S$  is an invariant probability density function for  $\mathbf{X}$ .

Proof

Let  $f(x) = 1/\mu(x)$  for  $x \in S$ , and let  $A$  be a finite subset of  $S$ . Then  $\sum_{y \in A} \frac{1}{n} G_n(x, y) \leq 1$  for every  $x \in S$ . Letting  $n \rightarrow \infty$  using the basic limit [above](#) gives  $\sum_{y \in A} f(y) \leq 1$ . The interchange of limit and sum is justified since  $A$  is finite. Since this is true for every finite  $A \subseteq S$ , it follows that  $C \leq 1$  where  $C = \sum_{y \in S} f(y)$ . Note also that  $C > 0$  since the chain is positive recurrent. Next note that

$$\sum_{y \in A} \frac{1}{n} G_n(x, y) P(y, z) \leq \frac{1}{n} G_{n+1}(x, z) \quad (16.6.19)$$

for every  $x, z \in S$ . Letting  $n \rightarrow \infty$  gives  $\sum_{y \in A} f(y) P(y, z) \leq f(z)$  for every  $z \in S$ . It then follows that  $\sum_{y \in S} f(y) P(y, z) \leq f(z)$  for every  $z \in S$ . Suppose that strict inequality holds for some for some  $z \in S$ . Then

$$\sum_{z \in S} \sum_{y \in S} f(y) P(y, z) < \sum_{z \in S} f(z) \quad (16.6.20)$$

Interchanging the order of summation on the left in the displayed inequality yields the contradiction  $C < C$ . Thus  $f$  is invariant for  $P$ . Hence  $f/C$  is an invariant probability density function. By the uniqueness result noted earlier, it follows that  $f/C = f$  so that in fact  $C = 1$ .

In summary, an irreducible, positive recurrent Markov chain  $\mathbf{X}$  has a unique invariant probability density function  $f$  given by  $f(x) = 1/\mu(x)$  for  $x \in S$ . We also now have a *test* for positive recurrence. An irreducible Markov chain  $\mathbf{X}$  is positive recurrent if and only if there exists a positive function  $g$  on  $S$  that is invariant for  $P$  and satisfies  $\sum_{x \in S} g(x) < \infty$  (and then, of course, normalizing  $g$  would give  $f$ ).

Consider now a general Markov chain  $\mathbf{X}$  on  $S$ . If  $\mathbf{X}$  has no positive recurrent states, then as noted earlier, there are no invariant distributions. Thus, suppose that  $\mathbf{X}$  has a collection of positive recurrent equivalence classes  $(A_i : i \in I)$  where  $I$  is a nonempty, countable index set. The chain restricted to  $A_i$  is irreducible and positive recurrent for each  $i \in I$ , and hence has a unique invariant probability density function  $f_i$  on  $A_i$  given by

$$f_i(x) = \frac{1}{\mu(x)}, \quad x \in A_i \quad (16.6.21)$$

We extend  $f_i$  to  $S$  by defining  $f_i(x) = 0$  for  $x \notin A_i$ , so that  $f_i$  is a probability density function on  $S$ . All invariant probability density functions for  $\mathbf{X}$  are mixtures of these functions:

$f$  is an invariant probability density function for  $\mathbf{X}$  if and only if  $f$  has the form

$$f(x) = \sum_{i \in I} p_i f_i(x), \quad x \in S \quad (16.6.22)$$

where  $(p_i : i \in I)$  is a probability density function on the index set  $I$ . That is,  $f(x) = p_i f_i(x)$  for  $i \in I$  and  $x \in A_i$ , and  $f(x) = 0$  otherwise.

Proof

Let  $A = \bigcup_{i \in I} A_i$ , the set of positive recurrent states. Suppose that  $f$  has the form given in the theorem. Since  $f(x) = 0$  for  $x \notin A$  we have

$$(fP)(y) = \sum_{x \in S} f(x) P(x, y) = \sum_{i \in I} \sum_{x \in A_i} p_i f_i(x) P(x, y) \quad (16.6.23)$$

Suppose that  $y \in A_j$  for some  $j \in I$ . Since  $P(x, y) = 0$  if  $x \in A_i$  and  $i \neq j$ , the last sum becomes

$$(fP)(y) = p_j \sum_{x \in A_j} f_j(x) P(x, y) = p_j f_j(y) = f(y) \quad (16.6.24)$$

because  $f_j$  is invariant for the  $P$  restricted to  $A_j$ . If  $y \notin A$  then  $P(x, y) = 0$  for  $x \in A$  so the sum above becomes  $(fP)(y) = 0 = f(y)$ . Hence  $f$  is invariant. Moreover,

$$\sum_{x \in S} f(x) = \sum_{i \in I} \sum_{x \in A_i} f(x) = \sum_{i \in I} p_i \sum_{x \in A_i} f_i(x) = \sum_{i \in I} p_i = 1 \quad (16.6.25)$$

so  $f$  is a PDF on  $S$ . Conversely, suppose that  $f$  is an invariant PDF for  $\mathbf{X}$ . We know that  $f$  is concentrated on the positive recurrent states, so  $f(x) = 0$  for  $x \notin A$ . For  $i \in I$  and  $y \in A_i$

$$\sum_{x \in A_i} f(x) P(x, y) = \sum_{x \in S} f(x) P(x, y) = f(y) \quad (16.6.26)$$

since  $f$  is invariant for  $P$  and since, as noted before,  $f(x) P(x, y) = 0$  if  $x \notin A_i$ . It follows that  $f$  restricted to  $A_i$  is invariant for the chain restricted to  $A_i$  for each  $i \in I$ . Let  $p_i = \sum_{x \in A_i} f(x)$ , the normalizing constant for  $f$  restricted to  $A_i$ . By uniqueness, the restriction of  $f/p_i$  to  $A_i$  must be  $f_i$ , so  $f$  has the form given in the theorem.

### Invariant Measures

Suppose that  $\mathbf{X}$  is irreducible. In this section we are interested in general functions  $g: S \rightarrow [0, \infty)$  that are invariant for  $\mathbf{X}$ , so that  $gP = g$ . A function  $g: S \rightarrow [0, \infty)$  defines a positive measure  $\nu$  on  $S$  by the simple rule

$$\nu(A) = \sum_{x \in A} g(x), \quad A \subseteq S \quad (16.6.27)$$

so in this sense, we are interested in invariant positive measures for  $\mathbf{X}$  that may not be probability measures. Technically,  $g$  is the density function of  $\nu$  with respect to counting measure  $\#$  on  $S$ .

From our work above, We know the situation if  $\mathbf{X}$  is positive recurrent. In this case, there exists a unique invariant probability density function  $f$  that is positive on  $S$ , and every other nonnegative invariant function  $g$  is a nonnegative multiples of  $f$ . In particular, either  $g = 0$ , the zero function on  $S$ , or  $g$  is positive on  $S$  and satisfies  $\sum_{x \in S} g(x) < \infty$ .

We can generalize to chains that are simply recurrent, either null or positive. We will show that there exists a positive invariant function that is unique, up to multiplication by positive constants. To set up the notation, recall that  $\tau_x = \min\{k \in \mathbb{N}_+ : X_k = x\}$  is the first positive time that the chain is in state  $x \in S$ . In particular, if the chain starts in  $x$  then  $\tau_x$  is the time of the first return to  $x$ . For  $x \in S$  we define the function  $\gamma_x$  by

$$\gamma_x(y) = \mathbb{E} \left( \sum_{n=0}^{\tau_x-1} \mathbf{1}(X_n = y) \mid X_0 = x \right), \quad y \in S \quad (16.6.28)$$

so that  $\gamma_x(y)$  is the expected number of visits to  $y$  before the first return to  $x$ , starting in  $x$ . Here is the existence result.

Suppose that  $\mathbf{X}$  is recurrent. For  $x \in S$ ,

1.  $\gamma_x(x) = 1$
2.  $\gamma_x$  is invariant for  $\mathbf{X}$
3.  $\gamma_x(y) \in (0, \infty)$  for  $y \in S$ .

Proof

1. By definition, given  $X_0 = x$ , we have  $X_0 = x$  but  $X_n \neq x$  for  $n \in \{1, \dots, \tau_x - 1\}$ . Hence  $\gamma_x(x) = 1$ .
2. Since the chain is recurrent, with probability 1 we have  $\tau_x < \infty$  and  $X_{\tau_x} = x$ . Hence for  $y \in S$ ,

$$\gamma_x(y) = \mathbb{E} \left( \sum_{n=0}^{\tau_x-1} \mathbf{1}(X_n = y) \mid X_0 = x \right) = \mathbb{E} \left( \sum_{n=1}^{\tau_x} \mathbf{1}(X_n = y) \mid X_0 = x \right) \quad (16.6.29)$$

(Note that if  $x = y$  then with probability 1, the  $n = 0$  term in the first sum and the  $n = \tau_x$  term in the second sum are 1 and the remaining terms are 0. If  $x \neq y$ , the  $n = 0$  term in the first sum and the  $n = \tau_x$  term in the second sum are 0 with probability 1, so again the two sums are the same.) Hence

$$\gamma_x(y) = \mathbb{E} \left( \sum_{n=1}^{\infty} \mathbf{1}(X_n = y, \tau_x \geq n) \mid X_0 = x \right) = \sum_{n=1}^{\infty} \mathbb{P}(X_n = y, \tau_x \geq n \mid X_0 = x) \quad (16.6.30)$$

Next we partition on the values of  $X_{n-1}$  in the sum to obtain

$$\begin{aligned} \gamma_x(y) &= \sum_{n=1}^{\infty} \sum_{z \in S} \mathbb{P}(X_n = y, X_{n-1} = z, \tau_x \geq n \mid X_0 = x) \\ &= \sum_{n=1}^{\infty} \sum_{z \in S} \mathbb{P}(X_n = y \mid X_{n-1} = z, \tau_x \geq n, X_0 = x) \mathbb{P}(X_{n-1} = z, \tau_x \geq n \mid X_0 = x) \end{aligned}$$

But  $\{X_0 = x, \tau_x \geq n\} \in \mathcal{F}_{n-1}$  (that is, the events depend only on  $(X_0, \dots, X_{n-1})$ ). Hence by the Markov property, the first factor in the last displayed equation is simply  $\mathbb{P}(X_n = y \mid X_{n-1} = z) = P(z, y)$ . Substituting and re-indexing the sum gives

$$\begin{aligned} \gamma_x(y) &= \sum_{n=1}^{\infty} \sum_{z \in S} P(z, y) \mathbb{P}(X_{n-1} = z, \tau_x \geq n \mid X_0 = x) = \sum_{z \in S} P(z, y) \mathbb{E} \left( \sum_{n=1}^{\tau_x} \mathbf{1}(X_{n-1} = z) \mid X_0 = x \right) \\ &= \sum_{z \in S} P(z, y) \mathbb{E} \left( \sum_{m=0}^{\tau_x-1} \mathbf{1}(X_m = z) \mid X_0 = x \right) = \sum_{z \in S} P(z, y) \gamma_x(z) = \gamma_x P(y) \end{aligned}$$

3. By the invariance in part (b),  $\gamma_x = \gamma_x P^n$  for every  $n \in \mathbb{N}$ . Let  $y \in S$ . Since the chain is irreducible, there exists  $j \in \mathbb{N}$  such that  $P^j(x, y) > 0$ . Hence

$$\gamma_x(y) = \gamma_x P^j(y) \geq \gamma_x(x) P^j(x, y) = P^j(x, y) > 0 \quad (16.6.31)$$

Similarly, there exists  $k \in \mathbb{N}$  such that  $P^k(y, x) > 0$ . Hence

$$1 = \gamma_x(x) = \gamma_x P^k(x) \geq \gamma_x(y) P^k(y, x) \quad (16.6.32)$$

and therefore  $\gamma_x(y) \leq 1/P^k(y, x) < \infty$ .

Next is the uniqueness result.

Suppose again that  $\mathbf{X}$  is recurrent and that  $g: S \rightarrow [0, \infty)$  is invariant for  $\mathbf{X}$ . For fixed  $x \in S$ ,

$$g(y) = g(x) \gamma_x(y), \quad y \in S \quad (16.6.33)$$

Proof

Let  $S_x = S - \{x\}$  and let  $y \in S$ . Since  $g$  is invariant,

$$g(y) = gP(y) = \sum_{z \in S} g(z) P(z, y) = \sum_{z \in S_x} g(z) P(z, y) + g(x) P(x, y) \quad (16.6.34)$$

Note that the last term is  $g(x) \mathbb{P}(X_1 = y, \tau_x \geq 1 \mid X_0 = x)$ . Repeating the argument for  $g(z)$  in the sum above gives

$$g(y) = \sum_{z \in S_x} \sum_{w \in S_x} g(w) P(w, z) P(z, y) + g(x) \sum_{z \in S_x} P(x, z) P(z, y) + g(x) P(x, y) \quad (16.6.35)$$

The last two terms are

$$g(x) [\mathbb{P}(X_2 = y, \tau_x \geq 2 \mid X_0 = x) + \mathbb{P}(X_1 = y, \tau_x \geq 1 \mid X_0 = x)] \quad (16.6.36)$$

Continuing in this way shows that for each  $n \in \mathbb{N}_+$ ,

$$g(y) \geq g(x) \sum_{k=1}^n \mathbb{P}(X_k = y, \tau_x \geq k \mid X_0 = x) \quad (16.6.37)$$

Letting  $n \rightarrow \infty$  then shows that  $g(y) \geq g(x) \gamma_x(y)$ . Next, note that the function  $h = g - g(x) \gamma_x$  is invariant, since it is a difference of two invariant functions, and as just shown, is nonnegative. Also,  $h(x) = g(x) - g(x) \gamma_x(x) = 0$ . Let  $y \in S$ . Since the chain is irreducible, there exists  $j \in \mathbb{N}$  such that  $P^j(y, x) > 0$ . Hence

$$0 = h(x) = hP^j(x) \geq h(y)P^j(y, x) \geq 0 \quad (16.6.38)$$

Since  $P^j(y, x) > 0$  it follows that  $h(y) = 0$ .

Thus, suppose that  $\mathbf{X}$  is null recurrent. Then there exists an invariant function  $g$  that is positive on  $S$  and satisfies  $\sum_{x \in S} g(x) = \infty$ . Every other nonnegative invariant function is a nonnegative multiple of  $g$ . In particular, either  $g = \mathbf{0}$ , the zero function on  $S$ , or  $g$  is positive on  $S$  and satisfies  $\sum_{x \in S} g(x) = \infty$ . The section on reliability chains gives an example of the invariant function for a null recurrent chain.

The situation is complicated when  $\mathbf{X}$  is transient. In this case, there may or may not exist nonnegative invariant functions that are not identically 0. When they do exist, they may not be unique (up to multiplication by nonnegative constants). But we still know that there are no invariant probability density functions, so if  $g$  is a nonnegative function that is invariant for  $\mathbf{X}$  then either  $g = \mathbf{0}$  or  $\sum_{x \in S} g(x) = \infty$ . The section on random walks on graphs provides lots of examples of transient chains with nontrivial invariant functions. In particular, the non-symmetric random walk on  $\mathbb{Z}$  has a two-dimensional space of invariant functions.

## Examples and Applications

### Finite Chains

Consider again the general two-state chain on  $S = \{0, 1\}$  with transition probability matrix given below, where  $p \in (0, 1)$  and  $q \in (0, 1)$  are parameters.

$$P = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} \quad (16.6.39)$$

1. Find the invariant distribution.
2. Find the mean return time to each state.
3. Find  $\lim_{n \rightarrow \infty} P^n$  without having to go to the trouble of diagonalizing  $P$ , as we did in the introduction to discrete-time chains.

Answer

1.  $f = \left( \frac{q}{p+q}, \frac{p}{p+q} \right)$
2.  $\mu = \left( \frac{p+q}{q}, \frac{p+q}{p} \right)$
3.  $P^n \rightarrow \frac{1}{p+q} \begin{bmatrix} q & p \\ q & p \end{bmatrix}$  as  $n \rightarrow \infty$ .

Consider a Markov chain with state space  $S = \{a, b, c, d\}$  and transition matrix  $P$  given below:

$$P = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \quad (16.6.40)$$

1. Draw the state diagram.
2. Determine the equivalent classes and classify each as transient or positive recurrent.
3. Find all invariant probability density functions.
4. Find the mean return time to each state.
5. Find  $\lim_{n \rightarrow \infty} P^n$ .

Answer

- 1.
2.  $\{a, b\}$  recurrent;  $\{c\}$  recurrent;  $\{d\}$  transient.
3.  $f = \left( \frac{3}{5}p, \frac{2}{5}p, 1-p, 0 \right), 0 \leq p \leq 1$
4.  $\mu = \left( \frac{5}{3}, \frac{5}{2}, 1, \infty \right)$

State graph  


$$5. P^n \rightarrow \begin{bmatrix} \frac{3}{5} & \frac{2}{5} & 0 & 0 \\ \frac{3}{5} & \frac{2}{5} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{2}{5} & \frac{4}{15} & \frac{1}{3} & 0 \end{bmatrix} \text{ as } n \rightarrow \infty$$

Consider a Markov chain with state space  $S = \{1, 2, 3, 4, 5, 6\}$  and transition matrix  $P$  given below:

$$P = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{1}{4} & 0 & \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \end{bmatrix} \quad (16.6.41)$$

1. Sketch the state graph.
2. Find the equivalence classes and classify each as transient or positive recurrent.
3. Find all invariant probability density functions.
4. Find the mean return time to each state.
5. Find  $\lim_{n \rightarrow \infty} P^n$ .

Answer

- 1.
2.  $\{1, 3, 5\}$  recurrent;  $\{2, 6\}$  transient;  $\{4\}$  recurrent.
3.  $f = \left( \frac{2}{19}p, 0, \frac{8}{19}p, 1-p, \frac{9}{19}p, 0 \right), \quad 0 \leq p \leq 1$
4.  $\mu = \left( \frac{19}{2}, \infty, \frac{19}{8}, 1, \frac{19}{8}, \infty \right)$

State graph  


$$5. P^n \rightarrow \begin{bmatrix} \frac{2}{19} & 0 & \frac{8}{19} & 0 & \frac{9}{19} & 0 \\ \frac{1}{19} & 0 & \frac{4}{19} & \frac{1}{2} & \frac{9}{38} & 0 \\ \frac{2}{19} & 0 & \frac{8}{19} & 0 & \frac{9}{19} & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{2}{19} & 0 & \frac{8}{19} & 0 & \frac{9}{19} & 0 \\ \frac{1}{19} & 0 & \frac{4}{19} & \frac{1}{2} & \frac{9}{38} & 0 \end{bmatrix} \text{ as } n \rightarrow \infty.$$

Consider a Markov chain with state space  $S = \{1, 2, 3, 4, 5, 6\}$  and transition matrix  $P$  given below:

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad (16.6.42)$$

1. Sketch the state graph.
2. Find the equivalence classes and classify each as transient or positive recurrent.
3. Find all invariant probability density functions.
4. Find the mean return time to each state.
5. Find  $\lim_{n \rightarrow \infty} P^n$ .

Answer

1.

State graph



2.  $\{1, 2\}$  recurrent;  $\{3, 4\}$  transient;  $\{5, 6\}$  recurrent.

3.  $f = \left(\frac{1}{3}p, \frac{2}{3}p, 0, 0, \frac{1}{2}(1-p), \frac{1}{2}(1-p)\right), \quad 0 \leq p \leq 1$

4.  $\mu = \left(3, \frac{3}{2}, \infty, \infty, 2, 2\right)$

$$5. P^n \rightarrow \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 & 0 \\ \frac{4}{15} & \frac{8}{15} & 0 & 0 & 0 & 0 \\ \frac{1}{5} & \frac{2}{5} & 0 & 0 & \frac{1}{5} & \frac{1}{5} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \text{ as } n \rightarrow \infty$$

Consider the Markov chain with state space  $S = \{1, 2, 3, 4, 5, 6, 7\}$  and transition matrix  $P$  given below:

$$P = \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{3}{4} & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (16.6.43)$$

1. Sketch the state digraph, and show that the chain is irreducible with period 3.

2. Identify the cyclic classes.

3. Find the invariant probability density function.

4. Find the mean return time to each state.

5. Find  $\lim_{n \rightarrow \infty} P^{3n}$ .

6. Find  $\lim_{n \rightarrow \infty} P^{3n+1}$ .

7. Find  $\lim_{n \rightarrow \infty} P^{3n+2}$ .

Answer

1.

State graph



2. Cyclic classes:  $\{1, 2\}, \{3, 4, 5\}, \{6, 7\}$

3.  $f = \frac{1}{1785} (232, 363, 237, 58, 300, 333, 262)$

4.  $\mu = 1785 \left( \frac{1}{232}, \frac{1}{363}, \frac{1}{237}, \frac{1}{58}, \frac{1}{300}, \frac{1}{333}, \frac{1}{262} \right)$

$$5. P^{3n} \rightarrow \frac{1}{585} \begin{bmatrix} 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \end{bmatrix} \text{ as } n \rightarrow \infty$$

$$\begin{aligned}
 6. P^{3n+1} &\rightarrow \frac{1}{585} \begin{bmatrix} 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \end{bmatrix} \text{ as } n \rightarrow \infty \\
 7. P^{3n+2} &\rightarrow \frac{1}{585} \begin{bmatrix} 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 0 & 0 & 0 & 0 & 0 & 333 & 262 \\ 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 232 & 363 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \\ 0 & 0 & 237 & 58 & 300 & 0 & 0 \end{bmatrix} \text{ as } n \rightarrow \infty
 \end{aligned}$$

### Special Models

Read the discussion of invariant distributions and limiting distributions in the Ehrenfest chains.

Read the discussion of invariant distributions and limiting distributions in the Bernoulli-Laplace chain.

Read the discussion of positive recurrence and invariant distributions for the reliability chains.

Read the discussion of positive recurrence and limiting distributions for the birth-death chain.

Read the discussion of positive recurrence and for the queuing chains.

Read the discussion of positive recurrence and limiting distributions for the random walks on graphs.

This page titled [16.6: Stationary and Limiting Distributions of Discrete-Time Chains](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist \(Random Services\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.