

## 16.12: Discrete-Time Queuing Chains

### Basic Theory

#### Introduction

In a *queuing model*, customers arrive at a station for service. As always, the terms are generic; here are some typical examples:

- The customers are persons and the service station is a store.
- The customers are file requests and the service station is a web server.
- The customers are packages and the service station is a processing facility.



Figure 16.12.1: Ten customers and a server

Queuing models can be quite complex, depending on such factors as the probability distribution that governs the arrival of customers, the probability distribution that governs the service of customers, the number of servers, and the behavior of the customers when all servers are busy. Indeed, queuing theory has its own lexicon to indicate some of these factors. In this section, we will study one of the simplest, discrete-time queuing models. However, as we will see, this discrete-time chain is embedded in a much more realistic continuous-time queuing process known as the *M/G/1 queue*. In a general sense, the main interest in any queuing model is the number of customers in the system as a function of time, and in particular, whether the servers can adequately handle the flow of customers.

Our main assumptions are as follows:

1. If the queue is empty at a given time, then a random number of new customers arrive at the next time.
2. If the queue is nonempty at a given time, then one customer is served and a random number of new customers arrive at the next time.
3. The number of customers who arrive at each time period form an independent, identically distributed sequence.

Thus, let  $X_n$  denote the number of customers in the system at time  $n \in \mathbb{N}$ , and let  $U_n$  denote the number of new customers who arrive at time  $n \in \mathbb{N}_+$ . Then  $\mathbf{U} = (U_1, U_2, \dots)$  is a sequence of independent random variables, with common probability density function  $f$  on  $\mathbb{N}$ , and

$$X_{n+1} = \begin{cases} U_{n+1}, & X_n = 0 \\ (X_n - 1) + U_{n+1}, & X_n > 0 \end{cases}, \quad n \in \mathbb{N} \quad (16.12.1)$$

$\mathbf{X} = (X_0, X_1, X_2, \dots)$  is a discrete-time Markov chain with state space  $\mathbb{N}$  and transition probability matrix  $P$  given by

$$P(0, y) = f(y), \quad y \in \mathbb{N} \quad (16.12.2)$$

$$P(x, y) = f(y - x + 1), \quad x \in \mathbb{N}_+, y \in \{x - 1, x, x + 1, \dots\} \quad (16.12.3)$$

The chain  $\mathbf{X}$  is the *queuing chain* with arrival distribution defined by  $f$ .

**Proof**

The Markov property and the form of the transition matrix follow from the construction of the state process  $\mathbf{X}$  in terms of the IID sequence  $\mathbf{U}$ . Starting in state 0 (an empty queue), a random number of new customers arrive at the next time unit, governed by the PDF  $f$ . Hence the probability of going from state 0 to state  $y$  in one step is  $f(y)$ . Starting in state  $x \in \mathbb{N}_+$ , one customer is served and a random number of new customers arrive by the next time unit, again governed by the PDF  $f$ . Hence the probability of going from state  $x$  to state  $y \in \{x - 1, x, x + 1, \dots\}$  is  $f[y - (x - 1)]$ .

#### Recurrence and Transience

From now on we will assume that  $f(0) > 0$  and  $f(0) + f(1) < 1$ . Thus, at each time unit, it's possible that no new customers arrive or that at least 2 new customers arrive. Also, we let  $m$  denote the mean of the arrival distribution, so that

$$m = \sum_{x=0}^{\infty} x f(x) \quad (16.12.4)$$

Thus  $m$  is the average number of new customers who arrive during a time period.

The chain  $\mathbf{X}$  is irreducible and aperiodic.

Proof

In a positive state, the chain can move at least one unit to the right and can move one unit to the left at the next step. From state 0, the chain can move two or more units to the right or can stay in 0 at the next step. Thus, every state leads to every other state so the chain is irreducible. Since 0 leads back to 0, the chain is aperiodic.

Our goal in this section is to compute the probability that the chain reaches 0, as a function of the initial state (so that the server is able to serve all of the customers). As we will see, there are some curious and unexpected parallels between this problem and the problem of computing the extinction probability in the branching chain. As a corollary, we will also be able to classify the queuing chain as transient or recurrent. Our basic parameter of interest is  $q = H(1, 0) = \mathbb{P}(\tau_0 < \infty \mid X_0 = 1)$ , where as usual,  $H$  is the hitting probability matrix and  $\tau_0 = \min\{n \in \mathbb{N}_+ : X_n = 0\}$  is the first positive time that the chain is in state 0 (possibly infinite). Thus,  $q$  is the probability that the queue eventually empties, starting with a single customer.

The parameter  $q$  satisfies the following properties:

1.  $q = H(x, x-1)$  for every  $x \in \mathbb{N}_+$ .
2.  $q^x = H(x, 0)$  for every  $x \in \mathbb{N}_+$ .

Proof

1. The critical observation is that if  $x \in \mathbb{N}_+$  then  $P(x, y) = P(1, y-x+1) = f(y-x+1)$  for  $y \in \{x-1, x, x+1, \dots\}$ . Thus, the chain, starting in  $x$ , and up until the time that it reaches  $x-1$  (if it does), behaves stochastically like the chain starting in state 1, and up until it reaches 0.
2. In order to reach 0, starting in state  $x \in \mathbb{N}_+$ , the chain must first reach  $x-1$  and then from  $x-1$  must reach  $x-2$ , until finally reaching 0 from state 1. Each of these intermediate trips has probability  $q$  by part (a) and are independent by the Markov property.

The parameter  $q$  satisfies the equation:

$$q = \sum_{x=0}^{\infty} f(x) q^x \quad (16.12.5)$$

Proof

This follows from the previous theorem by conditioning on the first state.

$$\mathbb{P}(\tau_0 < \infty \mid X_0 = 1) = \sum_{x=0}^{\infty} \mathbb{P}(\tau_0 < \infty \mid X_0 = 1, X_1 = x) \mathbb{P}(X_1 = x \mid X_0 = 1) \quad (16.12.6)$$

Note first that  $\mathbb{P}(\tau_0 < \infty \mid X_0 = 1, X_1 = 0) = 1 = q^0$ . On the other hand, by the Markov property and the previous result,

$$\mathbb{P}(\tau_0 < \infty \mid X_0 = 1, X_1 = x) = \mathbb{P}(\tau_0 < \infty \mid X_1 = x) = q^x, \quad x \in \mathbb{N}_+ \quad (16.12.7)$$

Of course  $\mathbb{P}(X_1 = x \mid X_0 = 1) = P(1, x) = f(x)$  for  $x \in \mathbb{N}$ .

Note that this is exactly the same equation that we considered for the branching chain, namely  $\Phi(q) = q$ , where  $\Phi$  is the probability generating function of the distribution that governs the number of new customers that arrive during each period.

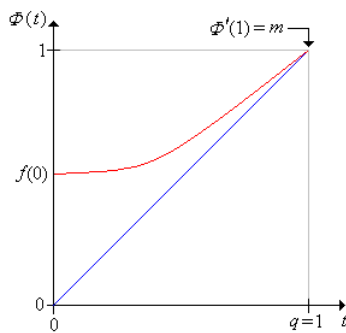


Figure 16.12.2 The graph of  $\phi$  in the recurrent case

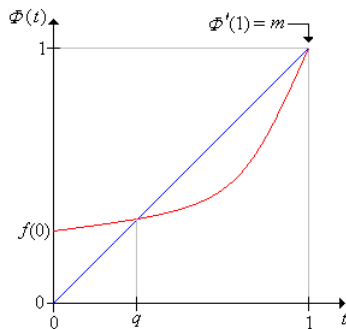


Figure 16.12.3 The graph of  $\phi$  in the transient case

$q$  is the smallest solution in  $(0, 1]$  of the equation  $\Phi(t) = t$ . Moreover

1. If  $m \leq 1$  then  $q = 1$  and the chain is recurrent.
2. If  $m > 1$  then  $0 < q < 1$  and the chain is transient..

Proof

This follows from our analysis of branching chains. The graphs above show the two cases. Note that the condition in (a) means that on average, one or fewer new customers arrive for each customer served. The condition in (b) means that on average, more than one new customer arrives for each customer served.

### Positive Recurrence

Our next goal is to find conditions for the queuing chain to be positive recurrent. Recall that  $m$  is the mean of the probability density function  $f$ ; that is, the expected number of new customers who arrive during a time period. As before, let  $\tau_0$  denote the first positive time that the chain is in state 0. We assume that the chain is recurrent, so  $m \leq 1$  and  $\mathbb{P}(\tau_0 < \infty) = 1$ .

Let  $\Psi$  denote the probability generating function of  $\tau_0$ , starting in state 1. Then

1.  $\Psi$  is also the probability generating function of  $\tau_0$  starting in state 0.
2.  $\Psi^x$  is the probability generating function of  $\tau_0$  starting in state  $x \in \mathbb{N}_+$ .

Proof

1. The transition probabilities starting in state 1 are the same as those starting in state 0:  $P(0, x) = P(1, x) = f(x)$  for  $x \in \mathbb{N}$ .
2. Starting in state  $x \in \mathbb{N}_+$ , the random time to reach 0 is the sum of the time to reach  $x - 1$ , the additional time to reach  $x - 2$  from  $x - 1$ , and so forth, ending with the time to reach 0 from state 1. These random times are independent by the Markov property, and each has the same distribution as the time to reach 0 from state 1 by our argument [above](#). Finally, recall that the PGF of a sum of independent variables is the product of the corresponding PGFs.

$$\Psi(t) = t\Phi[\Psi(t)] \text{ for } t \in [-1, 1].$$

Proof

Once again, the trick is to condition on the first state:

$$\Psi(t) = \mathbb{E}(t^{\tau_0} \mid X_0 = 1) = \sum_{x=0}^{\infty} \mathbb{E}(t^{\tau_0} \mid X_0 = 1, X_1 = x) \mathbb{P}(X_1 = x \mid X_0 = 1) \quad (16.12.8)$$

First note that  $\mathbb{E}(t^{\tau_0} \mid X_0 = 1, X_1 = 0) = t^1 = t\Psi^0(t)$ . On the other hand, by the Markov property and the [previous theorem](#),

$$\mathbb{E}(t^{\tau_0} \mid X_0 = 1, X_1 = x) = \mathbb{E}(t^{1+\tau_0} \mid X_0 = x) = t\mathbb{E}(t^{\tau_0} \mid X_0 = x) = t\Psi^x(t), \quad x \in \mathbb{N}_+ \quad (16.12.9)$$

Of course  $\mathbb{P}(X_1 = x \mid X_0 = 1) = P(1, x) = f(x)$ . Hence we have

$$\Psi(t) = \sum_{x=0}^{\infty} t\Psi^x(t)f(x) = t\Phi[\Psi(t)] \quad (16.12.10)$$

The PGF of any variable that takes positive integer values is defined on  $[-1, 1]$ , and maps this interval back into itself. Hence the representation is valid at least for  $t \in [-1, 1]$ .

The derivative of  $\Psi$  is

$$\Psi'(t) = \frac{\Phi[\Psi(t)]}{1 - t\Phi'[\Psi(t)]}, \quad t \in (-1, 1) \quad (16.12.11)$$

Proof

Recall that a PGF is infinitely differentiable on the open interval of convergence. Hence using the result in the [previous theorem](#) and the product and chain rules,

$$\Psi'(t) = \Phi[\Psi(t)] + t\Phi'[\Psi(t)]\Psi'(t) \quad (16.12.12)$$

Solving for  $\Psi'(t)$  gives the result.

As usual, let  $\mu_0 = \mathbb{E}(\tau_0 \mid X_0 = 0)$ , the mean return time to state 0 starting in state 0. Then

1.  $\mu_0 = \frac{1}{1-m}$  if  $m < 1$  and therefore the chain is positive recurrent.
2.  $\mu_0 = \infty$  if  $m = 1$  and therefore the chain is null recurrent.

Proof

Recall that  $\Psi$  is the probability generating function of  $\tau_0$ , starting at 0. From basic properties of PGFs we know that  $\Phi(t) \uparrow 1$ ,  $\Psi(t) \uparrow 1$ ,  $\Phi'(t) \uparrow m$ , and  $\Psi'(t) \uparrow \mu_0$  as  $t \uparrow 1$ . So letting  $t \uparrow 1$  in the result of the previous theorem, we have  $\mu_0 = 1/(1-m)$  if  $m < 1$  and  $\mu_0 = \infty$  if  $m = 1$ .

So to summarize, the queuing chain is positive recurrent if  $m < 1$ , null recurrent if  $m = 1$ , and transient if  $m > 1$ . Since  $m$  is the expected number of new customers who arrive during a service period, the results are certainly reasonable.

## Computational Exercises

Consider the queuing chain with arrival probability density function  $f$  given by  $f(0) = 1 - p$ ,  $f(2) = p$ , where  $p \in (0, 1)$  is a parameter. Thus, at each time period, either no new customers arrive or two arrive.

1. Find the transition matrix  $P$ .
2. Find the mean  $m$  of the arrival distribution.
3. Find the generating function  $\Phi$  of the arrival distribution.
4. Find the probability  $q$  that the queue eventually empties, starting with one customer.
5. Classify the chain as transient, null recurrent, or positive recurrent.
6. In the positive recurrent case, find  $\mu_0$ , the mean return time to 0.

Answer

1.  $P(0, 0) = 1 - p$ ,  $P(0, 2) = p$ . For  $x \in \mathbb{N}_+$ ,  $P(x, x-1) = 1 - p$ ,  $P(x, x+1) = p$ .
2.  $m = 2p$ .

3.  $\Phi(t) = pt^2 + (1-p)$  for  $t \in \mathbb{R}$ .
4.  $q = 1$  if  $0 < p \leq \frac{1}{2}$  and  $q = \frac{1-p}{p}$  if  $\frac{1}{2} < p < 1$ .
5. The chain is transient if  $p > \frac{1}{2}$ , null recurrent if  $p = \frac{1}{2}$ , and positive recurrent if  $p < \frac{1}{2}$ .
6.  $\mu_0 = \frac{1}{1-2p}$  for  $p < \frac{1}{2}$ .

Graphs of  $t \mapsto \Phi(t)$  and  $t \mapsto t$  when  $p = \frac{1}{3}$

 Graphs

Graphs of  $t \mapsto \Phi(t)$  and  $t \mapsto t$  when  $p = \frac{2}{3}$

 Graphs

Consider the queuing chain whose arrival distribution is the geometric distribution on  $\mathbb{N}$  with parameter  $1-p$ , where  $p \in (0, 1)$ . Thus  $f(n) = (1-p)p^n$  for  $n \in \mathbb{N}$ .

1. Find the transition matrix  $P$ .
2. Find the mean  $m$  of the arrival distribution.
3. Find the generating function  $\Phi$  of the arrival distribution.
4. Find the probability  $q$  that the queue eventually empties, starting with one customer.
5. Classify the chain as transient, null recurrent, or positive recurrent.
6. In the positive recurrent case, find  $\mu_0$ , the mean return time to 0.

Answer

1.  $P(0, y) = (1-p)p^y$  for  $y \in \mathbb{N}$ . For  $x \in \mathbb{N}_+$ ,  $P(x, y) = (1-p)p^{y-x+1}$  for  $y \in \{x-1, x, x+1, \dots\}$ .
2.  $m = \frac{p}{1-p}$ .
3.  $\Phi(t) = \frac{1-p}{1-pt}$  for  $|t| < \frac{1}{p}$ .
4.  $q = 1$  if  $0 < p \leq \frac{1}{2}$  and  $q = \frac{1-p}{p}$  if  $\frac{1}{2} < p < 1$ .
5. The chain is transient if  $p > \frac{1}{2}$ , null recurrent if  $p = \frac{1}{2}$ , and positive recurrent if  $p < \frac{1}{2}$ .
6.  $\mu_0 = \frac{1-p}{1-2p}$  for  $p < \frac{1}{2}$ .

Graphs of  $t \mapsto \Phi(t)$  and  $t \mapsto t$  when  $p = \frac{1}{3}$

 Graphs

Graphs of  $t \mapsto \Phi(t)$  and  $t \mapsto t$  when  $p = \frac{2}{3}$

 Graphs

Curiously, the parameter  $q$  and the classification of the chain are the same in the last two models.

Consider the queuing chain whose arrival distribution is the Poisson distribution with parameter  $m \in (0, \infty)$ . Thus  $f(n) = e^{-m} m^n / n!$  for  $n \in \mathbb{N}$ . Find each of the following:

1. The transition matrix  $P$
2. The mean  $m$  of the arrival distribution.
3. The generating function  $\Phi$  of the arrival distribution.
4. The approximate value of  $q$  when  $m = 2$  and when  $m = 3$ .
5. Classify the chain as transient, null recurrent, or positive recurrent.
6. In the positive recurrent case, find  $\mu_0$ , the mean return time to 0.

Answer

1.  $P(0, y) = e^{-m} m^y / y!$  for  $y \in \mathbb{N}$ . For  $x \in \mathbb{N}_+$ ,  $P(x, y) = e^{-m} m^{y-x+1} / (y-x+1)!$  for  $y \in \{x-1, x, x+1, \dots\}$ .
2. The parameter  $m$  is the mean of the Poisson distribution, so the notation is consistent.
3.  $\Phi(t) = e^{m(t-1)}$  for  $t \in \mathbb{R}$ .
4.  $q = 1$  if  $0 < m \leq 1$ . If  $m > 1$  then  $q$  is the solution in  $(0, 1)$  of the equation  $e^{m(q-1)} = q$  which can be expressed in terms of a special function known as the *Lambert W function*:

$$q = -\frac{1}{m} W(-me^{-m}) \quad (16.12.13)$$

For  $m = 2$ ,  $q \approx 0.20319$  For  $m = 3$ ,  $q \approx 0.059520$

5. The chain is transient if  $m > 1$ , null recurrent if  $m = 1$ , and positive recurrent if  $m < 1$ .
6.  $\mu_0 = \frac{1}{1-m}$  for  $m < 1$ .

Graphs of  $t \mapsto \Phi(t)$  and  $t \mapsto t$  when  $m = \frac{1}{2}$

 Graphs

Graphs of  $t \mapsto \Phi(t)$  and  $t \mapsto t$  when  $m = 2$

 Graphs

---

This page titled [16.12: Discrete-Time Queuing 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.