Massachusetts Institute of Technology
1.200J—Transportation Systems Analysis: Performance and Optimization
Fall 2015 — TA: Wichinpong "Park" Sinchaisri

Recitation 6
Unit 3 — Probabilistic Methodology and Examples of Applications

1 The Poisson Process x Queues

A check-in area at an airport has three parallel, independent and identical servers (all served by real people, as opposed to check-in machines). We shall call these Server 1, Server 2 and Server 3. Passenger arrivals for check-in can be approximated as a Poisson process with a rate of $\lambda = 1.25$ arrivals per minute, while service times at each of the check-in servers has a negative exponential pdf with an expected value of $(1/\mu) = 2$ minutes (i.e., each server has the capacity to serve 30 passengers per hour). If all three servers are busy, arriving passengers join a common queue for the three servers. Every time a server completes service to a passenger, the first passenger in the queue immediately begins service at that server. Of course, when there are fewer than 3 passengers present, one or more servers stay idle.

It is known that at time $t = 0$, all servers are busy and there are no passengers in the queue. Suppose now that Passenger A has just arrived (at $t = 0$) and, finding all three servers busy, becomes the first person in the queue.

(a) What is the expected total time Passenger A will spend in the system, i.e., the time she spends waiting in queue plus the time she spends being served by one of the servers?

(b) What is the probability that two more passengers will join the queue before Passenger A begins service at one of the servers?

(c) What is the probability Passenger A will have her service completed before service is completed to the passenger who occupied Server 2 at the time of passenger As arrival at the system?

(d) What is the probability the system will become empty (i.e., all four customers are served) before the next passenger after Passenger A arrives at the system? (Think of how this can happen.)

(e) Suppose we observe this system during a period when all three servers are continually busy serving passengers. What is the probability that 8 service completions are observed within a 6-minute time interval?

(f) For the same situation as in (g), what is the expected number and the variance of the number of service completions during a 6-minute time interval?

(g) Suppose that Passenger A arrived at the queueing system at exactly 1 p.m. What is the probability that the next customer to arrive at the system will come before 1:02 p.m.? Between 1:02 p.m. and 1:04 p.m.? After 1:04 p.m.?

2 Queueing Theory

Consider a queueing system with three parallel, independent and identical servers. Customers arrive with negative exponential inter-arrival times at an average rate of λ customers per hour. Service times are also negative exponential with an average service rate *per server* equal to μ. The queueing system has infinite queue capacity.

(a) Please draw the state transition diagram of this queueing system.

(b) What relationship between λ and μ must be satisfied for this system to reach steady state?

(c) Assuming that steady state is reached, please write, but do not solve, the set of equations that you would solve to compute the steady state probabilities $P_0, P_1, P_2, P_3, \ldots$

For the remainder of this problem, assume that $\lambda = 10$ per hour and $\mu = 3$ per hour. Assume now (parts d to h) that the queueing system has capacity for only 5 customers, i.e., the three receiving service plus a maximum of two people waiting. Any prospective customers who show up at the queueing system when two customers are waiting in queue will be denied access. They just go away and are lost forever, as far as this queueing system is concerned.

(d) Find all the state probabilities $P_0, P_1, P_2, P_3, \ldots$ for all the possible states of the system.

(e) Does this system reach steady state? Please explain in a sentence or two.

(f) What fraction of prospective customers will be served by this queueing system? In other words, what is the probability that a random prospective customer will be able to obtain access to this queueing system?

(g) Suppose a customer has just joined the system and is the second person in the queue of those waiting to enter service. What is the expected total time in the system of this customer (i.e., the sum of his expected waiting time to enter service and his expected service time)?

(h) What is the value of W, the expected total time spent in the system by those customers who actually join the system?

3 Solution 1

(a) The time she spends waiting in queue is the time until the first service completion from any of three servers. Because the three check-in servers constitute a Poisson Process with a combined rate of $3\mu = 3(0.5) = 1.5$ passengers per minute, that is, the time until the first service completion from any of three servers has a negative exponential pdf with parameter 3μ, it follows that $\mathbf{E}$[time spent in the queue] $= \frac{1}{1.5} = 0.67$ minutes.

The time she spends being served has a negative exponential pdf with parameter $\mu = 0.5$, and thus $\mathbf{E}$[time spent being served] $= \frac{1}{0.5} = 2$ minutes.

Therefore, the expected total time Passenger A will spend in the system is given by $\mathbf{E}$[time spent in the queue] + $\mathbf{E}$[time spent being served] = 2.67 minutes.

(b) Let X_p be the time until the next passenger arrival, and X_s be the time until service completion from any of the three servers. The passenger arrival process is a Poisson process with a rate of $\lambda = 1.25$ arrivals per minute, and the service completion process is a (combined) Poisson process with a (combined) rate of $3\mu = 1.5$ passengers per minute. Therefore, the probability that the next event will be a passenger arrival is given by

$$\mathbf{P}(X_p < X_s) = \frac{\lambda}{\lambda + 3\mu} = 0.4545.$$

Similarly, the probability that the next event will be a service completion is given by

$$\mathbf{P}(X_p > X_s) = \frac{3\mu}{\lambda + 3\mu} = 0.5455.$$

Thus, we have $\mathbf{P}$(**exactly** two more passengers will join the queue before Passenger A begins her service) = $\mathbf{P}$(the next three events are two passenger arrivals followed by a service completion) = $\mathbf{P}$(1st event is a passenger arrival)$\cdot\mathbf{P}$(2nd event is a passenger arrival)$\cdot\mathbf{P}$(3rd event is a service completion) = (0.4545)(0.4545)(0.5455) = 0.1127.

Also, we have $\mathbf{P}$(**at least** two more passengers will join the queue before Passenger A begins her service) = $\mathbf{P}$(the next two events are passenger arrivals) = $\mathbf{P}$(1st event is a passenger arrival)$\cdot\mathbf{P}$(2nd event is a passenger arrival) = (0.4545)(0.4545) = 0.2066.

(c) Passenger A will have her service completed before service is completed to the passenger who occupied Server 2 at the time of passenger A's arrival if (i) either Server 1 or Server 3 completes the service before Server 2, and (ii) Passenger A has her service completed before the passenger being served at Server 2. Note that the process of service completions by Server 1 and Server 3 is a Poisson process with a (combined) rate of $2\mu = 1$ passenger per minute.

Thus, we have $\mathbf{P}$(Passenger A will have her service completed before service is completed to the passenger who occupied Server 2 at the time of passenger A's arrival) = $\mathbf{P}$(Server 1 or Server 3 completes the service before Server 2)$\cdot\mathbf{P}$(Passenger A has her service completed before the passenger being served at Server 2) $= \left(\frac{2\mu}{2\mu+\mu}\right)\left(\frac{\mu}{\mu+\mu}\right) = 1/3$.

(d) The system will become empty before the next passenger after Passenger A arrives at the system if each of the next four events is a service completion, rather than a passenger arrival.

- **P**(1st event is a service completion) is given by $\frac{3\mu}{3\mu+\lambda}$. Note that the two competing processes are the service completion process by any of the **three** servers and the passenger arrival process.
- **P**(2nd event is a service completion | 1st event is a service completion) is again $\frac{3\mu}{3\mu+\lambda}$ because all **three** servers are still busy.
- Once there are two service completions and no passenger arrival, only **two** servers will be busy. Thus, **P**(3rd event is a service completion | 1st and 2nd events are service completions) is given by $\frac{2\mu}{2\mu+\lambda}$.
- Similarly, we have **P**(4th event is a service completion | the first three events are service completions) is given by $\frac{\mu}{\mu+\lambda}$.

Therefore, the probability that the system will become empty before the next passenger after Passenger A arrives at the system is given by

$$\left(\frac{3\mu}{3\mu+\lambda}\right)\left(\frac{3\mu}{3\mu+\lambda}\right)\left(\frac{2\mu}{2\mu+\lambda}\right)\left(\frac{\mu}{\mu+\lambda}\right) = 0.0378.$$

(e) During a period when all three servers are continually busy serving passengers, the process of service completions by any of the three servers is a Poisson process with a (combined) rate of $3\mu = 1.5$ passengers per minute. Therefore, the probability that 8 service completions observed within a 6-minute time interval is given by

$$\frac{(1.5\cdot 6)^8 e^{-1.5\cdot 6}}{8!} = 0.1318.$$

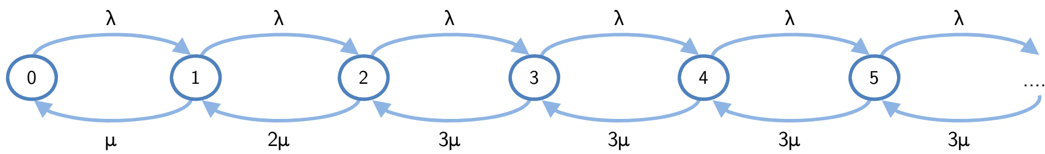
(f) The number of service completions during a 6-minute time interval has a Poisson distribution with parameter $(3\mu)(6) = 9$, and thus its expected value and variance are 9.

(g)

- **P**(the next customer will come before 1:02 pm) = **P**(at least one passenger arrival during a 2-minute interval) = 1 - **P**(no passenger arrival during a 2-minute interval) $= 1 - e^{-2\lambda} = 1 - e^{-2.5} = 0.918$.
- **P**(the next customer will come between 1:02 pm and 1:04 pm) = **P**(no passenger arrival during the first 2-minute interval) $\cdot$ **P**(at least one passenger arrival during the next 2-minute interval) $= e^{-2\lambda}(1 - e^{-2\lambda}) = e^{-2.5} - e^{-5} = 0.0753$.
- **P**(the next customer will come after 1:04 pm) = **P**(no passenger arrival during a 4-minute interval) $= e^{-4\lambda} = e^{-5} = 0.0067$.

3.1 Solution 2

(a) The state transition diagram of this queueing system:



(b) This is an M/M/3 system, and the steady state condition is given by $\frac{\lambda}{3\mu} < 1$. Loosely speaking, the steady state can be reached if the arrival rate is smaller than the service rate at the full capacity.

(c) From the state transition diagram given in part (a), we obtain the following set of balance equations.

$$
\begin{aligned}
\lambda P_0 &= \mu P_1 && \Rightarrow P_1 = \left(\frac{\lambda}{\mu}\right) P_0 \\
\lambda P_1 &= 2\mu P_2 && \Rightarrow P_2 = \frac{1}{2}\left(\frac{\lambda}{\mu}\right)^2 P_0 \\
\lambda P_2 &= 3\mu P_3 && \Rightarrow P_3 = \frac{1}{6}\left(\frac{\lambda}{\mu}\right)^3 P_0 \\
\lambda P_3 &= 3\mu P_4 && \Rightarrow P_4 = \frac{1}{6\cdot 3}\left(\frac{\lambda}{\mu}\right)^4 P_0 \\
\lambda P_4 &= 3\mu P_5 && \Rightarrow P_5 = \frac{1}{6\cdot 3^2}\left(\frac{\lambda}{\mu}\right)^5 P_0 \\
&\vdots && \quad \vdots \\
\lambda P_{k-1} &= 3\mu P_k && \Rightarrow P_k = \frac{1}{6\cdot 3^{k-3}}\left(\frac{\lambda}{\mu}\right)^k P_0 \qquad (k > 3)
\end{aligned}
$$

In addition, we also need $\sum_{n=0}^{\infty} P_n = 1$.

(d) The state transition diagram for this capacitated queueing system is simply the truncated version of the diagram in part (a). From part (c), we have

$$P_1 = \left(\frac{10}{3}\right) P_0,$$
$$P_2 = \frac{1}{2}\left(\frac{10}{3}\right)^2 P_0,$$
$$P_3 = \frac{1}{6}\left(\frac{10}{3}\right)^3 P_0,$$
$$P_4 = \frac{1}{6\cdot 3}\left(\frac{10}{3}\right)^4 P_0,$$
$$P_5 = \frac{1}{6\cdot 3^2}\left(\frac{10}{3}\right)^5 P_0,$$

and

$$1 = P_0 + P_1 + P_2 + P_3 + P_4 + P_5.$$

We have

$$P_0 = \frac{1}{\left(1+\left(\frac{10}{3}\right)+\frac{1}{2}\left(\frac{10}{3}\right)^2+\frac{1}{6}\left(\frac{10}{3}\right)^3+\frac{1}{6\cdot 3}\left(\frac{10}{3}\right)^4+\frac{1}{6\cdot 3^2}\left(\frac{10}{3}\right)^5\right)}$$
$$= 0.0327,$$
$$P_1 = 0.1091,$$
$$P_2 = 0.1819,$$
$$P_3 = 0.2021,$$
$$P_4 = 0.2246,$$
$$P_5 = 0.2495.$$

(e) Although the arrival rate of 10 per hour is larger than the maximum service rate of $3 \cdot 3 = 9$ per hour, this system reaches steady state because there is an upper limit on how long the queue can get. In this case, a large fraction of prospective customers will be turned away.

(f) The probability that a random prospective customer will be able to obtian access to this queueing system is equal to the probability that he/she does not find a full system, which is given by

$$1 - P_5 = 0.7505.$$

(g) First of all, note that during the time period that the customer waits in the queue, all three servers are busy, and the service completion process is a (combined) Poisson process with a (combined) rate of 3μ per hour. Because he has to wait for exactly two service completions before he can enter service, the expected waiting time in the queue is given by $2 \cdot \frac{1}{3\mu}$ hours. Because the customer will be served by only one server, the expected service time is given by $\frac{1}{\mu}$ hours. Therefore, the expected total time in the system of this customer equals to

$$2 \cdot \frac{1}{3\mu} + \frac{1}{\mu} = \frac{2}{9} + \frac{1}{3} = \frac{5}{9} \text{ hours,}$$

which is about 33.33 minutes.

(h) From part (f), we have that the probability that a random prospective customer will join the system is 0.7505. Given that a customer actually joins the system, the (conditional) probility that he/she will find another n customers in the system is then given by $\frac{P_n}{0.7505}$, where $n \in \{0, 1, 2, 3, 4\}$. We also have that

- $\mathbf{E}$[waiting time for a customer who finds zero, one, or two customers in the system] = 0,
- $\mathbf{E}$[waiting time for a customer who finds three customers in the system] = $\mathbf{E}$[time until the first service completion from any of the three servers] = $\frac{1}{3\mu}$, and
- $\mathbf{E}$[waiting time for a customer who finds four customers in the system] = $\mathbf{E}$[time until the second service completion from any of the three servers] = $2 \cdot \frac{1}{3\mu}$, as discussed in part (g).

Therefore, $\mathbf{E}$[total time spent in the system by those who actually join the system]

$$\begin{aligned}
&= \mathbf{E}[\text{service time}] + \mathbf{E}[\text{waiting time}] \\
&= \frac{1}{\mu} + \sum_{n=0}^{4} \mathbf{P}(\text{a customer finds } n \text{ existing customers in the system}) \\
&\quad \cdot \mathbf{E}[\text{waiting time for a customer who finds } n \text{ existing customers in the system}] \\
&= \frac{1}{\mu} + \frac{0.0327}{0.7505}(0) + \frac{0.1091}{0.7505}(0) + \frac{0.1819}{0.7505}(0) + \frac{0.2021}{0.7505}\left(\frac{1}{3\mu}\right) + \frac{0.2246}{0.7505}\left(\frac{2}{3\mu}\right) \\
&= 0.4297.
\end{aligned}$$

Alternatively, we can solve for the expected total time spent in the system by those who actually join the system using Little's law. The expected number of customers in the system L is given by

$$\sum_{n=0}^{5} nP_n = 0(0.0327) + 1(0.1091) + 2(0.1819) + 3(0.2021) + 4(0.2246) + 5(0.2495) = 3.22$$

The actual arrival rate of this capacitated system λ' is equal to $\lambda(1 - P_5) = 7.51$ customers per hour. By Little's law, it follows that

$$W = \frac{L}{\lambda'} \approx 0.43 \text{ hours}$$