Name: ________________________

OR 647: Queueing Theory
Mid-Term Solutions
March 16, 2016

1. [18 points] Arrivals to a queueing system follow a Poisson process with rate 18 per hour. Service times are exponential with rate 20 per hour. Each server is paid $40 per hour while working and $10 per hour while idle. The company determines that it loses $2 per customer per hour spent waiting in the queue, due to "ill-will" or lost customer loyalty.
   a. What is the hourly cost to the company if 1 server is employed?
   b. What is the hourly cost to the company if 2 servers are employed?
   c. If the company adds a 3rd server, would the cost go up or down compared with 2 servers? (Give a one- or two-sentence explanation without working the numbers out fully).

Part a: The hourly cost due to "ill will" is $\$2 \cdot L_q$.
The hourly cost per server is $\$40\cdot\rho + \$10(1-\rho)$, since ρ is the fraction of time a server is busy.

Thus, the total hourly cost is $2L_q + c(30\rho + 10)$, where c is the number of servers.

For $c = 1$, we have $\rho = 18/20 = .9$. And,

$$L_q = \frac{\rho^2}{1-\rho} = \frac{.81}{.1} = 8.1$$

The total hourly cost is:

$$2L_q + c(30\rho + 10) = 16.2 + (30\cdot 0.9 + 10) \boxed{= \$53.2}$$

Part b: For $c = 2$, we have $r = 18/20 = 0.9$ and $\rho = 18/40 = .45$.

$$p_0 = \left[\frac{r^c}{c!(1-\rho)} + \sum_{n=0}^{c-1}\frac{r^n}{n!}\right]^{-1} = \left[\frac{.9^2}{2(.55)} + 1 + .9\right]^{-1} \approx .37931$$

$$L_q = p_0 \frac{r^c}{c!}\frac{\rho}{(1-\rho)^2} \approx .37931\,\frac{.9^2(.45)}{2(.55)^2} \approx .22853$$

The total hourly cost is:

$$2L_q + c(30\rho + 10) = .45705 + 2(30\cdot 0.45 + 10) \boxed{= \$47.457}$$

Note: You can also compute the server cost as $\$10p_0 + \$50p_1 + \$80(1 - p_0 - p_1)$.

Part c: It will go up. The total work is conserved so the total non-idle costs ($40 per hour) will be the same no matter how many servers. Adding a 3rd server will basically add an additional cost of $10 per hour for idle time. The potential benefit comes from a reduction in the cost associated with L_q. But $2L_q$ is about 0.9 .45705, so the maximum possible reduction in this cost is about $0.90, significantly below $10.

2. [18 points] Gamma Airlines has a large call center to handle customer service operations. Calls arrive to the center according to a Poisson process with rate 500 per hour. Call times are exponentially distributed with a mean of 12 minutes. The center has 115 representatives. It is observed that the fraction of callers who do not immediately connect with a representative is .095. Gamma Airlines makes the following promise to its customers: If a caller does not immediately connect with a representative, Gamma Airlines gives the caller a $5 travel

voucher that can be applied to future purchases. Approximately 70% of customers eventually use the travel voucher, while 30% never use them. Representatives are paid $30 per hour.

a. What is the approximate hourly cost to Gamma Airlines? State any assumptions you make.
b. If the arrival rate increases to 600 calls per hour, what number of representatives would be needed to achieve the same level of service (in terms of fraction of callers who do not immediately connect)?
c. With the original assumptions (500 calls per hour), if the number of representatives is reduced to 110, what would be the approximate fraction of callers who do not immediately connect? (Use the square-root law to answer the question.)

<u>Part a</u>

Hourly cost to pay representatives = $30·115 = $3450
Hourly cost for travel vouchers = 500 · .095 · .7 · $5 = $166.25
Total hourly cost = $3,616.25

A key assumption: The $5 voucher does not induce someone to buy a ticket they would not have bought otherwise. Otherwise, the voucher would not be a cost to the airlines but a benefit.

<u>Part b</u>

In part a, r = 500 / 5 = 100. The number of extra servers is Δ = 15. By the square-root staffing law, to maintain the same level of service, Δ should increase by a factor $\sqrt{600/500} \approx 1.09545$, which would give $\Delta = 1.09545 \cdot 15 = 16.43$. Since the new offered load is 600/5 = 120, the new number of servers should be 120 + 16.43 = 136.43 (or round to 136 or 137).

<u>Part c</u>

$$r + \beta\sqrt{r} = c$$
$$100 + \beta\sqrt{100} = 110$$
$$100 + 10\beta = 110$$
$$\beta = 1$$

The probability of non-zero wait is about: $\dfrac{\phi(\beta)}{\phi(\beta)+\beta\Phi(\beta)}$.

From the table, this is about: $\boxed{\dfrac{.241971}{.241971 + .841345} \approx .2234}$

3. [12 points] A telecommunication line can handle a maximum of 5 calls at one time. Caller demand follows a Poisson process with a rate of 100 calls per hour. Call lengths are exponential with a mean of 6 minutes. When the system is full, arriving callers are blocked and do not return. Determine the fraction of customers blocked.

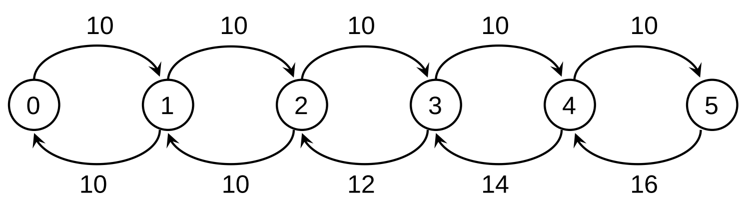
This is a M/M/c/c queue with r = 100 / 10 = 10 (μ = 10 per hour for 6 min service times) and c = 5. Using the Erlang-B formula,

$$p_c = \frac{10^5/5!}{1+10+10^2/2!+10^3/3!+10^4/4!+10^5/5!}$$
$$= \frac{10^5}{120+120\cdot 10+60\cdot 10^2+20\cdot 10^3+5\cdot 10^4+10^5}$$

$$\boxed{= \frac{100000}{177320} \approx .564}$$

4. [19 points] A single server queue has Poisson arrivals with rate 10 per hour and exponential service times with a mean of 6 minutes. The maximum number of customers in the system is 5 (i.e., 4 in the queue and 1 in service; customers arriving when the system is full are lost). Further, customers in line may become impatient and leave the system. More specifically, any customer who is 2^{nd} or 3^{rd} or 4^{th} in line will renege with rate 2 per hour. That is, the renege time for any such customer is exponential with mean ½ hr. These customers will not renege once they are 1^{st} in line.
    a. Draw the rate transition diagram for the CTMC associated with this system.
    b. Suppose you know the stationary probabilities p_n. Give W_q in terms of these probabilities (where W_q is defined as the average time in queue among all customers who join the system, including those who renege).
    c. Find the fraction of time the server is busy (i.e., a numerical value).
    d. If the arrival rate were very large (instead of 10), what would be the approximate average time waiting in queue among customers who eventually receive service?

Part a



Number in System

Note: When there are, for example, 5 in the system, there are 4 customers in line, 3 of whom may renege, so the reneging rate is 6.

Part b

$$L_q = p_2 + 2p_3 + 3p_4 + 4p_5$$

By Little's law:

$$\boxed{W_q = \frac{p_2 + 2p_3 + 3p_4 + 4p_5}{\lambda(1 - p_5)}}$$

Because some arrivals are blocked to the system, "λ" is $\lambda(1 - p_5)$ in Little's law. Note:

Part c

$$1 = p_0 + p_1 + p_2 + p_3 + p_4 + p_5 =$$

$$p_0[1+1+1+(10/12)+(10\cdot 10)/(12\cdot 14)+(10\cdot 10\cdot 10)/(12\cdot 14\cdot 16)] = 1$$

$$p_0 \approx .208$$

The fraction of time the server is busy is about $\boxed{p_0 \approx .792}$

Part d: If the arrival rate is very large, then a customer who receives service is very likely to arrive at the end of the line, in which case the expected wait in line is 1/16 + 1/14 + 1/12 + 1/10 ≈ .317 hrs.

Note: This assumes the customer does not renege. If you include customers who renege, then you have to consider a weighted sum.

5. [14 points] A graduate program in systems engineering has full time students and part time students. The number of *full-time* students who join the program each year is a Poisson random variable with a mean of 30. The number of *part-time* students who join the program each year is a Poisson random variable with a mean of 20. 30% of full-time students graduate in 1.5 years and 70% graduate in 2 years. 50% of part-time students graduate in 3 years and 50% graduate in 4 years.
   a. What is the average number of students enrolled in the program?
   b. On average, at a given moment in time, what fraction of enrolled students are full-time and what fraction are part-time?

Part a: The average arrival rate of full-time students to the "system" (i.e., the graduate program) is λ = 30 per year. The mean time spent in the system for a full-time student is W = .3(1.5) + .7(2) = .45+1.4 = 1.85 years. Thus, the average number of full-time students in the system is L = 30(1.85) = 55.5.

The average arrival rate of part-time students to the system is λ = 20 per year. The mean time spent in the system is .5(3) + .5(4) = 3.5 years. Thus, the average number of full-time students in the system is L = 20(3.5) = 70.

The total number of students enrolled in the program is 55.5 + 70 = 125.5

Part b: The fraction of students who are full time is 55.5 / (70 + 55.5) ≈ .442. The fraction of students who are part time is 70 / (70 + 55.5) ≈ .558. Note that there are more part-time students in the program even though they arrive at a slower rate.

6. [19 points] A queueing system has 2 queues, a left queue and a right queue. Both queues are served by a single common server, who uses a "serve the longest queue first" discipline. This means that, following a service completion, the server chooses the next customer from the front of the longer of the two lines. As a tie-breaking rule, if both lines have the same length, the server takes a customer from the left queue. (If both queues are empty, the server takes the next arriving customer from either stream.)
   a. Determine $W_{q\mathrm{L}}$ (the average waiting time in queue among customers in the left queue), $W_{q\mathrm{R}}$ (the average waiting time in queue among customers in the right queue) and W_q (the average waiting time in queue among all customers).
   b. Find L_q. State the time horizon you are using to measure L_q.

Left Queue			Right Queue		
Customer	Arrival Time	Service Time	Customer	Arrival Time	Service Time
1	0	5	A	3	8
2	6	4	B	7	2
3	12	6	C	8	3
4	16	3	D	12	5
5	20	2	E	14	1
6	25	7	F	23	4

Part a
The spreadsheet solution is given below:

						Queue state on departure	
Customer	Arrival Time	Service Time	Start Service	End Service	Tm in Queue	# Left	# Right
1	0	5	0	5	0	0	1
2	6	4	18	22	12	3	2
3	12	6	22	28	10	3	3
4	16	3	28	31	12	2	3
5	20	2	36	38	16		
6	25	7	39	46	14		
				Wq1	10.66666667		

						Queue state on departure	
Customer	Arrival Time	Service Time	Start Service	End Service	Tm in Queue	# Left	# Right
A	3	8	5	13	2	2	3
B	7	2	13	15	6	2	3
C	8	3	15	18	7	2	2
D	12	5	31	36	19	2	2
E	14	1	38	39	24		
F	23	4	46	50	23		
				Wq2	13.5		

Thus, Wq1 = 10.67
Wq2 = 13.5
Wq is the average of the two (since each group has the same number of customers) = (10.67 + 13.5) / 2 = 12.0833.

Part b:
The system empties at time 50, so we can compute Lq via Lq = λ Wq = (12/50) Wq = 2.9