Math/Mgmt 3502 (Ng/Fall 2007)
Queueing Handout 3
For class on **Dec 3,6,11 2007.**

4.4 Steady-state analysis

The purpose of this section is to analyze the steady-state performances of a queueing system whose steady-state probabilities, p_n of having n customers in the system have already been found.

Recall:
L_s = expected number of customers in the *system.*
L_q = expected number of customers in the *queue.*
W_s = expected waiting time in *system.*
W_q = expected waiting time in *queue.*
The following results pertain to service facility with c parallel servers. Then, given the steady-state probabilities of being in state n, namely, p_n,

$$L_s = \sum_{n=0}^{\infty} np_n$$

$$L_q = \sum_{n=c+1}^{\infty} (n-c)p_n$$

Let λ_{eff} be the *effective* average arrival rate, which is independent of the state n, to distinguish from λ_n. The effective arrival rate is obtain in the following manner:

$$\lambda_{eff} = \sum_{n=0}^{\infty} \lambda_n p_n$$

Here is a relationship between *expected number of customers* and *waiting time.*

$$L_s = \lambda_{eff} W_s$$

$$L_q = \lambda_{eff} W_q$$

Given that μ is the service rate per busy server, the expected service time is $1/\mu$, and so

$$W_s = W_q + \frac{1}{\mu}$$

or

$$L_s = L_q + \frac{\lambda_{eff}}{\mu}$$

To analyze the utility of the service facility, we use:

$$\text{expected number of busy servers, } \bar{c} = L_s - L_q = \frac{\lambda_{eff}}{\mu}$$

and the *percent utilization* of a service facility with c parallel servers is:

$$\text{percent utilization} = \frac{\bar{c}}{c} \times 100$$

Example 4.4.1 Consider a single-server queueing situation in which the arrival and departure rates are constant and given by $\lambda_n = 3$ per hour and $\mu_n = 8$ per hour for all $n \geq 0$. Compute p_n for all $n \geq 0$, λ_{eff}, and the four steady-state measures of performance.

4.5 Special cases of Poisson Queues

This section will concentrate on queues with Poisson arrivals and exponential service times.

4.5A $(M/M/1) : (GD/\infty/\infty)$

This is a single-server model with no limits on capacity of facility or source.

Arrival rates are independent of the number in the system, and hence, $\lambda_n = \lambda$ for all n. And the single server performs service at a constant rate, and hence, $\mu_n = \mu$ for all n. Thus, this model is one of Poisson "birth" and "death" with mean rates λ and μ, respectively.

$$p_n = \left(\frac{\lambda}{\mu}\right)^n p_0, \quad n = 0, 1, 2, \ldots$$

Using

$$\sum_{n=0}^{\infty} p_n = 1$$

,

and provided $\rho < 1$, we obtain

$$p_0 = 1 - \rho \text{ where } \rho \equiv \tfrac{\lambda}{\mu}$$

and hence, for $(M/M/1) : (GD/\infty/\infty)$ system,

$$p_n = (1-\rho)(\rho)^n \quad n = 0, 1, 2, \ldots$$

The basic measures of steady-state performances are:

$$L_s = E\{n\} = \frac{\rho}{1-\rho}$$

$$L_q = L_s - \frac{\lambda}{\mu} = \frac{\rho^2}{1-\rho}$$

$$W_s = \frac{L_s}{\lambda} = \frac{1}{\mu(1-\rho)}$$

$$W_q = \frac{L_q}{\lambda} = \frac{\rho}{\mu(1-\rho)}$$

Example 4.4.2 Turtle-Wax problem. At a car-wash facility, past information gathered indicates that cars arrive for service according to a Poisson distribution with mean 4 per hour. The time for washing, cleaning and waxing each car varies but it is found to follow an exponential distribution with mean 10 minutes per car. The washing facility cannot handle more than one car at a time.

Find all the steady-state measures and performances.

4.5B $(M/M/1):(GD/N/\infty)$

Now, we shall consider an embellishment of the previous case in the sense that we now have a *maximum* number of customers allowed in the system, namely, N. In other words, once N customers are in the system, new arrivals either balk or are NOT permitted to join the queue. In terms of the Poisson arrival rates and the exponential service times, we have:

$$\lambda_n = \begin{cases} \lambda & \text{if } n = 0,1,2,\ldots,N-1 \\ 0 & \text{if } n = N, N+1, \ldots \end{cases}$$

$$\mu_n = \mu \text{ for all } n = 0,1,2,\ldots$$

As usual, using

$$\sum_{n=0}^{N} p_n = 1, \Rightarrow \ p_0(1+\rho+\rho^2+\ldots+\rho^N) = 1$$

,

we find that

$$p_0 = \begin{cases} \frac{1-\rho}{1-\rho^{N+1}} & \rho \neq 1 \\ \frac{1}{N+1} & \rho = 1 \end{cases}$$

And p_n are obtained as follows

$$p_n = \begin{cases} \frac{1-\rho}{1-\rho^{N+1}}\rho^n & \rho \neq 1 \\ & \\ \frac{1}{N+1} & \rho = 1 \end{cases} \quad \text{for } n = 0,1,2,\ldots,N$$

Remark: For the above analysis to go through, we do NOT need the assumption that $\rho = \lambda/\mu < 1$. The rest of the steady-state measures are summarized as :

$$\lambda_{eff} = \lambda(1 - p_N)$$

$$L_s = \begin{cases} \frac{\rho[1-(N+1)\rho^N + N\rho^{N+1}]}{(1-\rho)(1-\rho^{N+1})} & \rho \neq 1 \\ & \\ \frac{N}{2} & \rho = 1 \end{cases}$$

$$L_q = L_s - \frac{\lambda_{eff}}{\mu} = L_s - \frac{\lambda(1-p_N)}{\mu}$$

$$W_q = \frac{L_q}{\lambda_{eff}} = \frac{L_q}{\lambda(1-p_N)}$$

$$W_s = W_q + \frac{1}{\mu} = \frac{L_s}{\lambda(1-p_N)}$$

Example 4.4.3.

Consider the **turtle-wax** problem. Suppose that the facility has a total of four parking spaces. If the parking lot is full, then arriving cars will balk or seek service elsewhere.

Use TORA or the above equations to compute the steady-state measures. Also, on the average, how many customers are lost due to the limited parking space?

4.5C $(M/M/c):(GD/\infty/\infty)$

In this model, customers arrive at a Poisson rate of λ and a maximum of c customers may be serviced simultaneously. The service rate per busy server is also a constant, μ. Since there is no capacity on the service facility, $\lambda_{eff} = \lambda$ as before.

$$\lambda_n = \lambda \quad n \geq 0$$

$$\mu_n = \begin{cases} n\mu & n \leq c \\ c\mu & n \geq c \end{cases}$$

$$p_n = \begin{cases} \frac{\lambda^n}{n!\mu^n} p_0 & n \leq c \\ \frac{\lambda^n}{c!c^{n-c}\mu^n} p_0 & n \geq c \end{cases}$$

$$p_0 = \left\{ \sum_{n=0}^{c-1} \frac{\rho^n}{n!} + \frac{\rho^c}{c!(1-\frac{\rho}{c})} \right\}^{-1}$$

$$L_q = \left\{ \frac{\rho^{c+1}}{(c-1)!(c-1)^2} \right\} p_0$$

$$L_s = L_q + \rho$$

$$W_q = \frac{L_q}{\lambda}$$

$$W_s = W_q + \frac{1}{\mu}$$

Some approximations to these calculations, for $\rho \ll 1$:

$$p_0 \cong 1 - \rho \text{ and } L_q \cong \frac{\rho^{c+1}}{c^2}$$

and for ρ/c very close to 1:

$$p_0 \cong \frac{(c-\rho)(c-1)!}{c^c} \text{ and } l_q \cong \frac{\rho}{c-\rho}$$

Example 4.4.4. **Cab company problem**. A very rural town is being serviced by two cab companies. Each of the two cab companies owns two cabs and are known to share the market almost equally. This is evident by the fact that calls arrive at each company's dispatching office at the rate of 10 per hour. The average time per ride is 11.5 minutes. Arrival of calls follow a Poisson distribution, whereas ride times are exponential.

The two companies were recently bought by an investor, Joe. After taking over the two companies, Joe's first action was to try to consolidate the two cab companies into one dispatching office, hoping to either maximize the utility of the service or to provide a faster service to the customers in the community. The question here is whether Joe should consolidate the two companies or not.

Joe noticed that the utilization (ratio of hourly arriving calls to rides) for each company is

$$100\frac{\lambda}{c\mu} = \frac{100 \times 10}{2 \times (60/11.5)} = 95.8\%$$

.

4.5D $(M/M/c):(GD/N/\infty)$, $c \leq N$

In this model, customers arrive at a Poisson rate of λ and a maximum of c customers may be serviced simultaneously. The service rate per busy server is also a constant, μ. Since there is a capacity on the service facility, the maximum queue size must be $N - c$.

$$\lambda_n = \begin{cases} \lambda & 0 \leq n < N \\ 0 & n \geq N \end{cases}$$

$$\nu_n = \begin{cases} n\mu & 0 \leq n < c \\ c\mu & c \leq n \leq N \end{cases}$$

Letting $\rho \equiv \frac{\lambda}{\mu}$, and using the usual expression for p_n and p_0, we have:

$$p_n = \begin{cases} \frac{\rho^n}{n!} p_0 & 0 \leq n \leq c \\ \\ \frac{\rho^n}{c! c^{n-c}} p_0 & c \leq n \leq N \end{cases}$$

where

$$p_0 = \begin{cases} \left\{ \sum_{n=0}^{c-1} \frac{\rho^n}{n!} + \frac{\rho^c [1-(\rho/c)^{N-c+1}]}{c!(1-\rho/c)} \right\}^{-1} & \rho/c \neq 1 \\ \left\{ \sum_{n=0}^{c-1} \frac{\rho^n}{n!} + \frac{\rho^c}{c!}(N-c+1) \right\}^{-1} & \rho/c = 1 \end{cases}$$

$$L_q = \begin{cases} p_0 \frac{\rho^{c+1}}{(c-1)!(c-\rho)^2} \left\{ 1 - (\rho/c)^{N-c} - (N-c)(\rho/c)^{N-c}(1-(\rho/c)) \right\} & \rho/c \neq 1 \\ p_0 \frac{\rho^c (N-c)(N-c+1)}{2c!} & \rho/c = 1 \end{cases}$$

$$L_s = L_q + (c - \tilde{c}) = L_q + \frac{\lambda_{eff}}{\mu}$$

where

$$\tilde{c} = \text{expected number of idle servers} = \sum_{n=0}^{c} (c-n) p_n$$

$$\lambda_{eff} = \lambda(1 - p_N) = \mu(c - \tilde{c})$$

Example 4.4.5. Refer to the **Cab company problem**.

Assume that our friend Joey had made the decision to consolidate the two cab companies. One obvious solution to the excessive waiting time for the customers in the **cab company problem** is to buy more cabs. However Joey does not have good enough of a credit background to make this a feasible choice. Hence, to alleviate this problem, Joey had instructed the dispatching office to turn off customers in a polite way once the waiting list reaches 16 customers.

Now, Joey wants to study the effect of his decision on the waiting time, in particular, what is the expected waiting time for the customers W_q.

Did Joey make a good decision?

4.5E $(M/M/\infty):(GD/\infty/\infty)$, Self-service model

In this model, the number of servers is infinite because the customer is the server.

Note: This does **not** apply to service facilities with finite number of equipment, notwithstanding the fact that customers service themselves.

As in section **4.3**, we have

$$\lambda_n = \lambda \quad \text{for all } n \geq 0$$

$$\mu_n = n\mu \quad \text{for all } n \geq 0$$

$$p_n = \frac{\lambda^n}{n!\mu^n} p_0 = \frac{\rho^n}{n!} p_0$$

It follows that

$$p_0 = e^{-\rho}$$

Hence, in closed form,

$$p_n = \frac{e^{-\rho}\rho^n}{n!}, \quad n = 0, 1, 2, \ldots$$

$$L_s = E\{n\} = \rho$$

$$W_s = \frac{1}{\mu}$$

$$L_q = W_q = 0$$

(Why is $W_q = 0$ and why is $W_s = 1/\mu$?)

The following table gives some results of steady-state measures for several c values to $(M/M/c):(GD/\infty/\infty)$ model.

	$\rho = .1$				$\rho = 9$			
Measure	$M/M/10$	$M/M/20$	$M/M/50$	$M/M/\infty$	$M/M/10$	$M/M/20$	$M/M/50$	$M/M/\infty$
W_s	.1	.1	.1	.1	1.6	1.0	1.0	1.0
W_q	2.5×10^{-19}	1.8×10^{-41}	0	0	.668	.0001	6×10^{-23}	0
L_s	.1	.1	.1	.1	15.02	9.0	9.0	9.0
L_q	2.5×10^{-19}	1.8×10^{-41}	0	0	6.02	.0092	5.6×10^{-22}	0
p_0	.90484	.90484	.90484	.90484	.00007	.00012	.00012	.00012

From these results, speculate a relationship between the results of the $(M/M/\infty):(GD/\infty/\infty)$ model and that of the $(M/M/c):(GD/\infty/\infty)$ model.

4.5F $(M/M/R):(GD/K/K)$, $R<K$ - machine servicing model.

This is a queueing model which assumes that there are R repairmen available for servicing a total of K machines.

Usually in this type of model, an arrival corresponds to a machine breaking down, while a departure corresponds to a broken-down machine just got fixed. (Why is this model one of finite calling source?)

Let λ be the rate of breakdown *per* machine. As in the same analysis in section 6.3, we obtain

$$\lambda_n = \begin{cases} (K-n)\lambda & 0 \le n \le K \\ 0 & n \ge K \end{cases}$$

$$\mu_n = \begin{cases} n\mu & 0 \le n \le R \\ R\mu & R \le n \le K \\ 0 & n > K \end{cases}$$

For $R > 1$,

$$L_q = \sum_{n=R+1}^{K} (n-R)p_n$$

$$L_s = L_q + (R - \bar{R}) = L_q + \frac{\lambda_{eff}}{\mu}$$

where

$$\bar{R} = \text{expected number of idle repairmen} = \sum_{n=0}^{R}(R-n)p_n$$

$$\lambda_{eff} = \mu(R-\bar{R}) = \lambda(K-L_s)$$

Example 4.4.6. **Maytag Repairmen problem.**

Since the consolidated cab company made customers wait that long, it went bankrupt and our friend Joey decides to go into the laundromat business. He borrowed money from his big daddy and bought a laundry shop which has a total of 22 Maytag laundry machines. On the average, an operative machine breaks down every 2 hours. It takes 12 minutes on the average, for the lonely Maytag repairman to repair a broken down machine.

Joey is interested in determining the number of Maytag repairmen needed to keep his shop running "reasonably" smooth.

(**Note**: This is a pretty subjective question in the sense that you have to figure out what Joey meant by "reasonably" smooth.)

4.6 Non-Poisson Queueing models.

In general, non-Poisson queueing models do not have tractable analytical results. Hence, the usual way to attack these models is via simulation techniques. (If you want more info, take an OR course that covers simulation in every sense of that term $\ddot{\smile}$).

However, there is one particular non-Poisson queueing model that have closed form analytical results, and that is $(M/G/1) : (GD/\infty/\infty)$ queueing model. In this model, the arrival distribution is still Poisson with rate λ. However, the service time is described by a general probability distribution with mean, $E\{t\}$ and variance, $var\{t\}$. And the only analytical closed form results are the four basic steady-state measures, L_s, L_q, W_s, and W_q.

The steady-state probabilities, p_n are **not** easily obtained in closed form. (WHY?)

Let λ be the (Poisson) arrival rate at the single-server facility. Let $E\{t\}$ and $var\{t\}$ be the respective mean and variance of the service-time distribution. (**Note:** that the service rate μ is $1/E\{t\}$, and $\lambda_{eff} = \lambda$). Then, we have the following results, provided $\lambda E\{t\} < 1$:

$$L_s = \lambda E\{t\} + \frac{\lambda^2(E^2\{t\} + var\{t\})}{2(1 - \lambda E\{t\})}$$

$$W_s = \frac{L_s}{\lambda}$$

$$L_q = L_s - \lambda E\{t\}$$

$$W_q = \frac{L_q}{\lambda}$$

Here are the simplified results of L_s for a special case of the service-time distribution.

When the service-time is approximately constant, $var\{t\} = 0$, and hence, we have:

$$L_s = \rho + \frac{\rho^2}{2(1 - \rho)}$$

Example 4.4.7. Return to our **Turtle-wax** problem where cars arrive for service at a Poisson rate of $\lambda = 4$ per hour.

Suppose that in the **Turtle-wax** problem, the washing is done by automatic machines, so that the service-time may be considered the same and constant for all cars. The washing machine cycle takes 10 minutes exactly.

Analyze this new model, and make some comparisons with the results obtained in the $(M/M/1) : (GD/\infty/\infty)$ model before.