Notes and Comments on Chapter 2

In a queueing system, it is natural to think of people or parts "getting in line" – for example, at the bank, the cafeteria, the airport security check, the post office. Items in a factory may join a line for inspection. However, a group of broken machines, waiting for the service of a repairperson, don't physically "get in line." But, conceptually, the queue and its discipline (see page 21) exist just the same.

Let's look a bit more at Kendall's notation (often "Kendall notation"). An excellent summary of it is at http://en.wikipedia.org/wiki/Kendall's_notation.

1. The first symbol represents the arrival pattern. "*M*" means exponential distribution (exponential time between arrivals). Why not "*E*"? "*E*" is for "Erlang" distribution (an important distribution in its own right, which is the sum of *n* identically distributed and independent exponential distributions), and the "*M*" stands for "Markovian" or "memoryless." What is meant by "memoryless"? It means that experience so far means nothing for prediction. For example, if interarrival times are truly memoryless, the statement "Someone's sure to come soon – we've had no arrivals for a long time" is false. More formally, p(next arrival within 10 minutes) = p(next arrival within 10 minutes | last arrival was 50 minutes ago). (Remember that this is conditional probability notation; read the vertical bar as "given that."
2. The second symbol specifies the service time distribution. It is often reasonable to argue that interarrival times are exponential (hence memoryless), but rarely reasonable to argue that service times are. If service times are truly memoryless, a statement like "Her service is sure to be done soon – the clerk has spent 30 minutes with her already!" are false.
3. The third symbol specifies the number of servers (e.g., parallel machines, open wickets at post office, number of repair people).
4. The fourth symbol specifies the number of places in the system (places to wait + number of servers). If a 3-barber shop has 4 chairs to wait in, the number of places is 7. If a factory work cell has four machines and the buffer can hold six parts which are waiting, the number of places is 10.
5. The fifth symbol specifies the number of potential customers, usually assumed infinite (e.g., if we are modeling a bank). If there are 10 machines in total, and 2 mechanics to service them, the fourth symbol is 2 and the fifth is 10.
6. The sixth symbol is the queue discipline (e.g., FCFS or FIFO, the default, is "first-come, first-served").

The easiest type of queue to analyze mathematically is $M/M/1/\infty/\infty/\text{FCFS}$. In view of the defaults, this notation would be abbreviated to $M/M/1$.

On page 25, notice the formula for Lq (length of queue, pages 22-23). Of course, it's nice for Lq to be small. So $(1-\rho)$, in the denominator, needs to be large. But ρ is the fraction of time the server(s) is/are utilized. $0 \leq \rho \leq 1$. $\rho = \lambda / \mu$ (page 24). It's nice for ρ to be large (get value for the cost of the server by keeping the server busy!). But these objectives conflict! What happens to Lq if $\rho = 1$? These comments apply also to the formula on page 26 for Wq, the average waiting time (page 22), which is also "small is better." Among operations research specialists, in systems with high variability, $\rho = 0.8$ (80%) is often considered the "sweet spot."

Reassurance: *Do not memorize formulas. Do not memorize Kendall's notation.*

Not far beyond *M*/*M*/1, mathematical analysis breaks down. As just one example, suppose a bank proposes the rule "open another teller window when six or more customers are waiting" (e.g., change the third parameter of Kendall's notation). How much will that help? Would saying "five" or "seven" instead of "six" be better? For such questions we need simulation. What if the arrival rate changes? For example, at the bank, the interarrival times may be exponential (hence the number of in nearby offices rush to arrivals Poisson) all day, but with a smaller mean time from 11am to 1pm (people working the bank during lunch hour). For these and many other issues, discrete-event simulation is required.

Briefly review the Poisson distribution (was it covered in your statistics class?).

Memorize this sentence: The interarrival times are exponentially distributed if and only if the number of arrivals is Poisson distributed.

You will soon do a homework based on chapter 2.

Questions welcome as always!