Introduction to Queueing Theory

Prof. N. Selvaraju
Department of Mathematics
Indian Institute of Technology Guwahati, India

Lecture - 20
Queues with Bulk Arrivals

Hi and hello, everyone. What we have seen so far are Markovian Queueing Systems, but those systems can be modelled by a birth-death process, and hence we call that as birth-death queuing systems. In such systems, what was happening is that the arrivals occur singly. And the customers are also served one at a time. So, at any given point of time, there could be at most one event of either one arrival or one departure, nothing more than that. So, that makes the state transitions occur only to the nearest neighbours, and hence a birth-death process model, which is a special case of a continuous-time Markova chain, is appropriate. And we could see that it can ah model quite a bit of a large number of systems that exhibit such a behaviour. And the reason we looked at that in the first place is that these are analytically very much tractable in a very easy way and which makes our life easier. And from which you know we could get the insights about the systems very easily. The methodologies can be put into operation in such simple models, which can then be easily extended to more complex models, and that was the idea. And now, what we are going to consider is still Markovian queueing systems only, but these are then modelled by a non birth-death process. So, that is why we call this a general Markovian queueing system does not mean that the previous ones were not; they were also Markovian queueing systems, but those were a specific class of Markovian queueing systems which is what we have considered. But we will now consider models which can be considered or modelled by a general continuous-time Markov chain. So, when we say it is a Markovian queueing system, we are looking for a model, which is the Markovian model, but it is a continuous-time Markov chain. So, it is still in continuous time, but in a state space is discrete; that is the case that we have but by a non birth-death process. So, that models now what they allow? They allow transitions beyond one step with respect to the state. So, it could still be one step, but there will be some transitions that might go beyond. If it gets restricted to only one step, then again, you are going back to your birth-death queueing models. But here is a general one, so they generally allow transitions apart from the one-step that nearest neighbour transitions; they do allow transitions beyond the nearest neighbours case. So, that is, those are the systems, and corresponding models are what then we are going to look at it. Since this still remains, you are still remaining in the Markovian regime; you still have the Chapman Kolmogorov equations hold good backward and forward Kolmogorov equations are true, and the balance equations that we write for generally Markov chain are all still valid. So, what does that makes? That makes the essence of the approach that we have adopted so far in handling a birth-death queueing model that remains the same and is also applicable, and it also helps us to solve such non-birth death Markovian models. What is then the difference? The difference is the complexity might increase, as you might see. That is what is going to happen. From here onwards, without any saying, we are not going to look at a queueing system and look at its transient analysis or time-dependent behaviour where the initial state plays a role, and you have a finite amount of time and which you are looking at the behaviour of the system. Now we are not going to look at that anymore, which we have seen already in the case of birth-death with simple models, and we saw the complexity. But for most of our analysis, the equilibrium analysis would suffice, we will assume, and we look at queueing systems that are only in equilibrium. Now, here onwards, we are not going to look at certain models; we can easily look at the transient behaviour in some way, but we will not do that; we will restrict ourselves to the equilibrium analysis. So, it goes without saying that from this point onwards, in the remaining part of the course, we are going to look at queueing systems that are in equilibrium; that means we are going to do only the study state analysis, ok and we will not worry about the time-dependent behaviour. So, this is what we are going to do next, this general Markovian queueing system, where the appropriate stochastic models that come handy are a general continuous-time Markov chain which is a non birth-death type.

The first one that we are looking at in such a scenario is a simple generalization of what you would see in queues where the arrivals or services happen one at a time. What we have now is what we call as bulk queues; what do they mean? That there are many systems in which the arrivals can occur, and or services can be performed in groups which you call in bulk or in batches. For example, when a bus or train arrives or leaves, people arrive or leave in groups right, people may go to a restaurant together and maybe served in batches, lift or elevators or boats handle passengers in batches these types of systems were all that is known as in our language as bulk queues right. Bailey, 1954 introduced the concept of this bulk service in queues, and Gaver introduced the bulk arrival queues. The literature is now quite vast, and you can look at the literature far more; again, we are seeing that it is an elementary introduction kind of thing, only what we are going to see for these kinds of models. So, there are many more; if you want, you can look at the literature. Now, what do we have? We have bulk queues; we are not looking at any generalization of concept; we still want to be within that Poisson exponential or Markovian nature. So, this can be modelled by a non birth-death, but still, a Markov process that is basically a CTMC. Specifically, we consider systems where either the arrivals are in batches or service is provided for batches of customers; these are the two types of bulk queues that we are going to look at it. Of which the first one that we consider is the bulk input or bulk arrival queues which we denote by $M^{[X]}/M/1$. Again this may not be the common notation sometimes; some books use the notation M in the superscript simply X without the square bracket.

So, but will also be that is why I said these all might not be standard notations. So, we might have to be explicitly saying, but at least in our case, like we can be familiar that you know we will be using this kind of notation to denote the bulk arrival or bulk input queues. So, what is the description of the system? So, this has all the underlying assumptions that we had for an $M/M/1$ system.

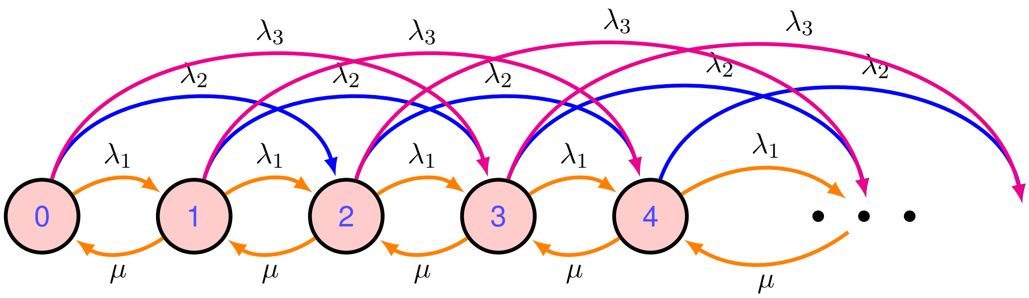
So, the arrivals happen according to a Poisson process with rate lambda service times are IID exponential with parameter mu there is a single server. There is an infinite capacity for queueing first come, first serve, and there is no dependency between the arrival process and service process and whatever the assumptions. These kinds of assumptions that we have had for $M/M/1$ still hold true in this case.

In addition, we said that the arrival stream follows a Poisson process, but now the number of customers in any arriving module, like when you look at an arriving point, the time between two different modules at two different arrivals is still an exponential distribution with parameter lambda. But at each arrival instance, the number of customers who arrive at that point of time is not just a single customer but a batch. That batch size is what we call it by a random variable X, and that has probability mass function $\{c_n\}_{n\geq 1}$, which means that with probability c_1, the batch will be of size 1, with probability c_2, the batch will be of size 2 and so on. This describes the corresponding probability mass function, and this is what is this X that is sitting on the head in a way in the superscript of this M. So, this system is what then we are describing by $M^{[X]}/M/1$, and whenever $c_1 = P\{X = 1\}$ then this will give you give back the original $M/M/1$ system. And the total arrival process now, if you look at it here, is a compound Poisson process right with X denoting batch size.

You think of λ_n being the arrival rate of a Poisson process of batches of size n then, this c_n which is the probability of batch size would be then λ_n/λ, where λ is the composite arrival rate of all batches meaning, that this is $\sum_{n=1}^{\infty} \lambda_n$. So, this gives that proportion of arrivals with that particular rate. So, this is what then would be the probability mass function where this λ_n is the arrival rate of batches of size n. Which means suppose if you think that there is a first instant in which 2 customers came, then there is a second instance in which 3 customers came, then again there is another instance in which 1 customer came, then there is another instance in which there are 2 customers came. So, there is the first instant and the fourth instant where 2. So, between these two, the time if you count like that is what is exponential with this parameter λ_2. Like which each way each one you can differentiate, and each one of them is the arrival rate of the corresponding batch size arrivals. A batch of size 2 arrivals, if you think, then that will be distributed exponentially with parameter $lambda_n$ and for which the probability is c_n.

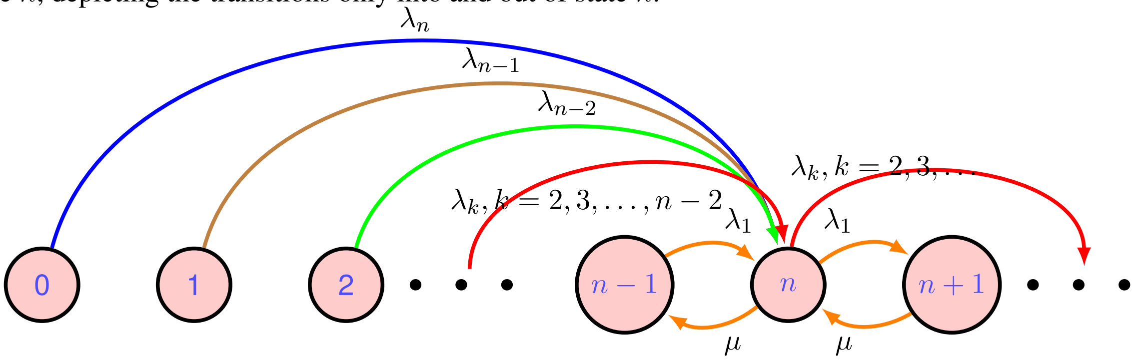
So, this is basically λ_n/λ, so the total arrival process if you look at it. If you are just counting the total number of customers that come during an interval, if you look at it, then that will follow; this is an overlap of this all these Poisson processes, the superposition of all these Poisson processes, which will be a compound Poisson process. Now, you can go back and look at the generalizations that we described for a Poisson process, and at that point of time, we mentioned that this compound Poisson process is what is going to be used in the case of bulk arrival situations. So, it is the total arrival process if I look at it. So, basically is a compound Poisson process which means that in a bulk arrival queue, the arrival process that if you look at the total arrival process, will be a compound Poisson process if you are looking at the number. So, as usual, we are considering N, which is the number of customers in the system in a steady-state, and the underlying process is still Markovian is what we are seeing. Now, this is a sample transition rate diagram which is, of course, it is very difficult to depict everything completely, and it will become quite messy, but at least you will get an idea here.

So, why do we call this sample? Because here this diagram gives the transition rates for the simple case of $M^{[X]}/M/1$ queue, where X takes the value either 1 or 2 or 3 with the corresponding rates $\lambda_1, \lambda_2 and \lambda_3$. So, the sum of these is what is the overall arrival rate of that λ. So, what is the transition that can happen?



For example, when the state is in 0, only then can arrival move to either state 1 or 2 or 3 with the rates $\lambda_1, \lambda_2, \lambda_3$ meaning what. If a batch of size 1 arrives for which the rate is λ_1, it will move to state 1 if a batch of size 2 arrives, which can happen with probability c_2. So, λ time c_2 is what is your λ_2 effectively. So, it can directly move to state 2. So, you see here, from 0, it can move to state 2. So, it goes beyond nearest neighbours whereas, 0 to 1 is still the nearest neighbour transition, but 0 to 2 and similarly if a batch of size 3 arises then it can move from 0 to 3. Since we assume only these 3 batch sizes are possible, it cannot move from 0 to 4. So, the same thing can happen in states 1, 2, and so on. Say, for example, if I look at in state say 2, what can happen, it can move to state 1 by service completion of a single customer. By the way, in this model, you have to remember that the service does not happen for batch; service is given by for 1 customer only at a time. So, that is why with the rate μ, it moves from state 2 to state 1, or with the arrival of batch size 1, it can move to state 3, or with the arrival of batch size 2, which can move to state 4 or with the arrival of batch size 3 it can move to state 5 and so on. So, this is general, which is $1, 2, 3$ with batch sizes of these 3, and this is what would be the rate diagram. Now, if you can imagine, suppose if you allow for all possible things, which means that what will happen all sizes suppose if you allow then 0 to 1, 0 to 2, 0 to 3, 0 to 4, 0 to 5, 0 to 6 everything all arrows will be there. Now, if you look at everywhere somewhere here like, it will be quite messy.

So, you can see that, but that is the essence, so you need to catch the essence. Now, once you catch the essence generally in such Markovian models general Markovian models what is done is you look at a typical state and see the transitions into that particular state and out of that state so that that helps us to write down the general balance equations that are what be the general case. General Case: The transition rate diagram only with respect to a general state n, depicting the transitions only into and out of state n.



- For arbitrary batch size distribution, the balance equations are given by

$$\lambda p_0 = \mu p_1$$
$$(\lambda + \mu) p_n = \mu p_{n+1} + \lambda \sum_{k=1}^{n} p_{n-k} c_k, \qquad n \geq 1.$$

  The last term in the second equation says that a total of n in the system can arise from the presence of $n - k$ in the system followed by an arrival of a batch size k.

So, writing down this generic balance equation is the most important step for which you need to understand the system clearly, what is the transitions how the transitions happen from one state to the other state and what are the possibilities you can think about it. So, that is what will give you this equation. Now, once you write down this equation, you see here the same essence of what we did in the birth-death queue in the same way only we are writing down this balance equation, but the thing is that now things are a little bit complex, apart from that there is no other ah issue with respect to this here. So,

$$\lambda p_0 = \mu p_1$$
$$(\lambda + \mu) p_n = \mu p_{n+1} + \lambda \sum_{k=1}^{n} p_{n-k} c_k, \qquad n \geq 1.$$

is what is the balance equation now; once you write this balance equation, then it is the question of now how you will solve it. Again, the methodologies similar approaches can be adopted here. So, this is basically that these two lines describe how the last equation comes up here. So, to solve this equation, we will use a generating function approach.

But whenever you know these batch sizes are small, though theoretically here we said X is a random variable which has support as a set of all positive integers theoretically these are limited and if the batch size is limited. Suppose, when a ferry comes, it cannot be that an infinite amount of passengers comes there; when a bus comes, it cannot be that you know an infinite amount of customers can arrive from the arrival of a bus. So, there is all limited, so if the batch size is small, then the difference equations can also be used, but we will use a generating function approach here to solve this system of equations. Now, for which we define two generating functions

$$C(z) = \sum_{n=1}^{\infty} c_n z_n \quad |z| \leq 1$$
$$P(z) = \sum_{n=0}^{\infty} p_n z_n \quad |z| \leq 1$$

as the generating functions of the batch-size probabilities $\{c_n\}$ and the steady-state probabilities $\{p_n\}$, respectively.

And if you look here, since the batch size is an input to the system, you know what is the batch size because that is part of your model assumptions. So, the $C(z)$ is known, and here this is an input. So, we need to find out from that what is $P(z)$ and from there, you extract to get the p_n's, and once you get p_n's then you can do whatever the usual thing that you want to do with respect to the system state probabilities. Now, let us see how one can do.

- Multiplying each equation by z^n and summing over n, we have

$$\lambda \sum_{n=0}^{\infty} p_n z^n + \mu \sum_{n=1}^{\infty} p_n z^n = \frac{\mu}{z} \sum_{n=1}^{\infty} p_n z^n + \lambda \sum_{n=1}^{\infty} \sum_{k=1}^{n} p_{n-k} c_k z^n$$

- Now, observe that $\sum_{k=1}^{n} p_{n-k} c_k$ is the probability mass function for the sum of the steady-state system size and batch size. And, the PGF of this sum is the product of the respective PGFs.

$$\text{i.e.,} \quad \sum_{n=1}^{\infty} \sum_{k=1}^{n} p_{n-k} c_k z^n = \sum_{k=1}^{\infty} c_k z^k \sum_{n=k}^{\infty} p_{n-k} z^{n-k} = C(z)P(z).$$

- Substituting in the above,

$$\lambda P(z) + \mu[P(z) - p_0] = \frac{\mu}{z}[P(z) - p_0] + \lambda C(z)P(z)$$
$$\implies P(z) = \frac{\mu(1-z)p_0}{\mu(1-z) - \lambda z[1 - C(z)]}, \qquad |z| \leq 1.$$

  Again here we have one quantity to be determined still, which is p_0, and as we have said earlier, this p_0 will be then determined using this condition that is satisfied by this generating function which is $P(1) = 1$. Now, let us see how I can do this.

- Rewriting the $P(z)$ above,

$$P(z) = \frac{p_0}{1 - (\lambda/\mu) z \overline{C}(z)}, \qquad \text{where } \overline{C}(z) = \frac{1 - C(z)}{1 - z}.$$

  ■ Observe that $\overline{C}(z)$ is the generating function of the complementary batch-size probabilities $\overline{C}_n = P\{X > n\} = 1 - P\{X \leq n\}$

So, this $C(z)$ is basically the generating function corresponding to this $\bar{C}_n$'s. So, the $\bar{C}(z) = \sum_{n=0} \bar{C}_n z^n$. So, you can simply work it out if you want to see this; otherwise, it is not a very difficult one. And the $C(z)$ and the complementary batch size probability, the complementary CDF itself plays a critical role in many different applications. Say, for example, in reliability theory, where you look for the survival of some instrument or machine or anything beyond a certain time point. So, you are looking at the $P\{X > n\}$ or in actuaries; for example, you look at the survival probabilities, that is why this is also called a survival function in those areas. So, this $\frac{1-C(z)}{1-z}$ is nothing but $\bar{C}(z)$ where $C(z)$ is the probability generating function of the PMF of the batch size. Whereas you take the CDF of that batch size and take its complement for that function; if you find the generating function, that is what you will get as $\bar{C}(z)$. Now, if you want, you can work it out, but you can find it elsewhere in your text also; you will find it.

- Then $1 = P(1) = \dfrac{p_0}{1-(\lambda/\mu)\overline{C}(1)}$ and $\overline{C}(1)$ can be found from $\overline{C}(z)$, using L'Hopital rule (once).
  We then get $\overline{C}(1) = \sum_{n=0}^{\infty} P(X > n) = E[X]$. Therefore,
  $$p_0 = 1 - (\lambda/\mu)E[X] = 1 - \rho, \quad \left(\rho = \frac{\lambda E[X]}{\mu}\right)$$
  and hence
  $$P(z) = \frac{1-\rho}{1-(\lambda/\mu)z\overline{C}(z)}, \qquad |z| \leq 1.$$
  We can then find p_n's from the $P(z)$ above.

  Now, again you note here $p_0 = 1 - \rho$, so this will be strictly greater than 0 if $\rho < 1$.

  ⧫ So, $\rho < 1$ is the necessary and sufficient condition for the steady state to exist in the usual manner as usually you expect.

  So, this is what is the condition under which this steady-state would exist in the given in the case of here in the $M^{[X]}/M/1$ model where my ρ is now $\frac{\lambda E[X]}{\mu}$. Because this is the total arrival rate, $\lambda E[X]$ is the total arrival rate; that is what you will get here. So, this is once we have $\frac{1-\rho}{1-(\lambda/\mu)z\bar{C}(z)}$ now if we want, we can invert and if it is possible, not just if you want, but it should be possible, and it is possible for many reasonable assumptions if you make with respect to batch size then it is possible to invert it. .

Now, let us look at some of the performance measures. Of course, many things can be done, but of course, some of the performance measures are what we will see as the basic ones.

- We first find L from $P(z)$ as
  $$L = P'(1) = (1-\rho)(\lambda/\mu)\frac{\overline{C}(1) + \overline{C'}(1)}{(1-(\lambda/\mu)\overline{C}(1))^2} = \frac{(\lambda/\mu)(E(X) + \overline{C'}(1))}{1-\rho}.$$
  Now, $\overline{C'}(1)$ can be found from $\overline{C}(z)$, using L'Hopital rule (twice) to $\overline{C'}(z) = [1 - C(z) - (1-z)C'(z)]/(1-z)^2$ as $\overline{C'}(1) = C''(1)/2 = E[X(X-1)/2]$.

Therefore,

$$L = \frac{(\lambda/\mu)(E[X]+E[X^2])}{2(1-\rho)} = \frac{\rho + (\lambda/\mu)E[X^2]}{2(1-\rho)} = \frac{\rho}{1-\rho}\left[\frac{E(X)+E(X^2)}{2E(X)}\right], \quad \left(\rho = \frac{\lambda E[X]}{\mu}\right)$$

And here, you will see in this form that $\rho/1-\rho$ is what normally you would find for the $M/M/1$ queue, and this is multiplied by some $\frac{E(X)+E(X^2)}{2E(X)}$ factor which is based upon the first two moments of the batch size probabilities. Similarly, like this is a mean number in the system, say if you want to compute the variance in the system you can then express, you will get an expression which we will not do here, but those of you who are interested you might do, if you want to know the variance of the system size which will have a similar expression, but that will now involve the first three moments of the batch size probabilities right so that is how one would get it. Now, once we get this L, then by using Little's law and the the the box results that we have exhibited between W, W_q, L, L_q. So, you can employ Little's law to get the other quantities.

- But now you have to remember that the arrival rate of customers is $\lambda E(X)$ and hence the other performance measures using Little's law are given by

$$W = \frac{L}{\lambda E[X]}, \quad W_q = W - \frac{1}{\mu}, \text{ and } L_q = \lambda E[X] W_q = L - \rho.$$

But if you want waiting time distribution, but that is a little bit complicated, then you have to look at now the two quantities; if we are looking at this, just a hint, there is nothing that we are going to do with that, but I will just highlight the point when for some of you might be interested to look further. The waiting time total waiting time suppose if you look at in the queue, this will be two components one is the arrivals are happening in batches. So, the number of customers who arrived up to the previous batch that will be one time, they have to be served all. Then within his batch, then, how his service time is going to come when depending upon that, there is another time. So, the sum of these two times will give you the total waiting time, and accordingly, the analysis has to be carried out, which is even carried out in the general service time distribution set up; you people can look into that. So, we will not worry about that; we are looking at the average performance measure with respect to these four. Now, with this in mind like, let us look at some of the simple examples.

Example. *[M/M/1 Queue]*
If $c_1 = P\{X=1\} = 1$ and $c_n = 0$ for $n > 1$, then we obtain the corresponding results for an $M/M/1$ queue.

Example. *[Constant Batch Size]*

- If, for some $K > 1$, $c_K = P\{X = K\} = 1$ and $c_n = 0$ for $n \neq K$, then

$$L = \frac{\rho + K\rho}{2(1-\rho)} = \frac{K+1}{2}\frac{\rho}{1-\rho}, \qquad (\rho = \lambda K/\mu)$$

which is equal to the M/M/1 mean system size multiplied by $(K+1)/2$. Also,

$$L_q = L - \rho = \frac{2\rho^2 + (K-1)\rho}{2(1-\rho)}.$$

The inversion of $P(z)$ to get p_n's is reasonably easy for small values of K.

But if it is for large values, of course, then things will be difficult, but for small values, you can still do it; it is not going to be that complex. That is what you might see.

Example. *[Geometric Batch Size]*
Let $X \sim Geo(\alpha)$, i.e., $c_n = (1-\alpha)\alpha^{n-1},\ n = 1, 2, \ldots, 0 < \alpha < 1$.

Then $\rho = \lambda E[X]/\mu = (\lambda/\mu)/(1-\alpha)$ and $C(z) = (1-\alpha)\sum_{n=1}^{\infty} \alpha^{n-1} z^n = \dfrac{z(1-\alpha)}{1-\alpha z}$.

$$\begin{aligned}
\implies P(z) &= \frac{(1-\rho)(1-z)}{1-z-(\lambda/\mu)z[1-C(z)]} = \frac{(1-\rho)(1-z)}{1-z-(\lambda/\mu)z[1-z\frac{(1-\alpha)}{(1-\alpha z)}]} \\
&= \frac{(1-\rho)(1-\alpha z)}{1-z[\alpha+(\lambda/\mu)]} \\
&= (1-\rho)\left(\frac{1}{1-z[\alpha+(\lambda/\mu)]} - \frac{\alpha z}{1-z[\alpha+(\lambda/\mu)]}\right).
\end{aligned}$$

Utilizing the formula for the sum of geometric series, we get

$$\begin{aligned}
p_n &= (1-\rho)\{[\alpha+(\lambda/\mu)]^n - \alpha[\alpha+(\lambda/\mu)]^{n-1}\} \\
&= (1-\rho)(\lambda/\mu)[\alpha+(\lambda/\mu)]^{n-1}, \qquad n > 0
\end{aligned}$$

Now, let us take care you know a typical example which is exactly as presented in the textbook.

Example. *[A machine-line production system]*

- Consider a multistage machine-line process that produces an assembly in quantity.
- After the first stage, many items are found to have one or more defects, which must be repaired before they enter the second stage and this job of making the adjustment is done by one worker.
- It is observed that the number of defects is 1 or 2 (most of the times).
- The interarrival times for units with one defect $\sim Exp(\lambda_1)$ and with two defects $\sim Exp(\lambda_2)$, where $\lambda_1 = 1/h$ and $\lambda_2 = 2/h$.
- The worker's service time distribution is $Exp(\mu)$, where $1/\mu = 10$ minutes.
- Subsequently, it was observed that the rate of defects have increased, and it was decided to put an additional worker who who will handle exclusively the 2-defect items, while the original worker will handle only 1-defect items.
- When to add the additional worker will be based on a cost analysis, especially based on L.

Example. *[contd...]*

- We will find the average number of units in the system, L, under the assumption of only two possible batch sizes, from the given data.

$$\lambda = \lambda_1 + \lambda_2 = 3, \quad \mu = 6, \quad c_1 = \frac{\lambda_1}{\lambda} = \frac{1}{3}, \quad c_2 = \frac{\lambda_2}{\lambda} = \frac{2}{3}$$
$$E[X] = \frac{1}{3} + 2 \times \frac{2}{3} = \frac{5}{3}, \qquad E[X^2] = \frac{1}{3} + 2^2 \times \frac{2}{3} = 3$$
$$\rho = \lambda \frac{E[X]}{\mu} = \frac{5}{6}, \qquad L = \frac{\rho + (\lambda/\mu)E[X^2]}{2(1-\rho)} = \frac{\frac{5}{6} + \frac{3}{2}}{2(1-\frac{5}{6})} = 7.$$

- *Extra:* If needed, we can get p_n's from $P(z) = \dfrac{1}{6 - 3z - 2z^2}$ as

$$p_n = (0.116)(0.880)^n + (0.050)(-0.379)^n, n \geq 0.$$

Example. *[contd...]*
Cost Analysis:
▶ C_1 = The cost per unit time per waiting repair.
▶ C_2 = The cost of a worker per unit time.

- When there is only one worker, the system is an $M^{[X]}/M/1$ system with the assumption of only two possible batch sizes (1 and 2) and L is computed as above.
  Then the expected cost per unit time of the single-server system is $C = C_1 L + C_2$

- If a second worker is added, then there is an additional C_2 cost, and we now have two queues: ▶ The singlet line is $M/M/1$ and $L_1 = \dfrac{\frac{\lambda_1}{\mu}}{1 - \frac{\lambda_1}{\mu}}$

  ▶ The doublet line is $M^{[X]}/M/1$ with fixed batch size of 2 and $L_2 = \dfrac{3\frac{\lambda_2}{\mu}}{1 - 2\frac{\lambda_2}{\mu}}$.

  ▶ Hence, the average number of units in the system is

$$L = L_1 + L_2 = \frac{\frac{\lambda_1}{\mu}}{1 - \frac{\lambda_1}{\mu}} + \frac{3\frac{\lambda_2}{\mu}}{1 - 2\frac{\lambda_2}{\mu}}.$$

Example. *[contd...]*

- Therefore, with the second worker, the new expected cost is

$$C^* = C_1 \left(\frac{\frac{\lambda_1}{\mu}}{1 - \frac{\lambda_1}{\mu}} + \frac{3\frac{\lambda_2}{\mu}}{1 - 2\frac{\lambda_2}{\mu}} \right) + 2C_2.$$

- Decision is based on the comparative magnitude of C and C^*. An additional worker is added whenever $C^* < C$ or

$$C_1 \left(\frac{\frac{\lambda_1}{\mu}}{1 - \frac{\lambda_1}{\mu}} + \frac{3\frac{\lambda_2}{\mu}}{1 - 2\frac{\lambda_2}{\mu}} \right) + 2C_2 < C_1 L + C_2$$

$$\Rightarrow C_1\left(\frac{\frac{\lambda_1}{\mu}}{1-\frac{\lambda_1}{\mu}}+\frac{3\frac{\lambda_2}{\mu}}{1-2\frac{\lambda_2}{\mu}}\right)+C_2<C_1L=C_1\left(\frac{\rho+(\lambda/\mu)E[X^2]}{2(1-\rho)}\right)=C_1\left(\frac{\frac{\lambda_1}{\mu}+3\frac{\lambda_2}{\mu}}{1-\frac{\lambda_1}{\mu}-2\frac{\lambda_2}{\mu}}\right)$$

Example. *[contd...]*

- *That is*

$$C_2<C_1\left(\frac{\frac{\lambda_1}{\mu}+3\frac{\lambda_2}{\mu}}{1-\frac{\lambda_1}{\mu}-2\frac{\lambda_2}{\mu}}-\frac{\frac{\lambda_1}{\mu}}{1-\frac{\lambda_1}{\mu}}-\frac{3\frac{\lambda_2}{\mu}}{1-2\frac{\lambda_2}{\mu}}\right)$$

- Using the values of the parameters, we arrive at a decision criterion of

$$C_2<\frac{19C_1}{5}.$$

So, using the parameter values then, we can obtain the decision criteria to be $C_2<\frac{19C_1}{5}$. Now, an additional worker is involved or is added provided the C_1, C_2 satisfies $C_2<\frac{19C_1}{5}$; otherwise, it might be a good idea from the cost perspective to continue with the single worker. If the cost of the worker is less than $19/5$ times the cost per unit of waiting that you are associating. So, depending on these two costs, then the decision can be taken on whether to add a second worker or not. Now, the same can be reversed similar analysis can be reversed if you are given C_1, C_2. Suppose if you want to determine the rate at which the service needs to be expanded or anything that you want to decide with respect to that; one can also do a similar analysis along these lines.

So, this is one simple example of how one can use this kind of model applicable to any model that you have done so far. So, when you are basically comparing this kind of thing that you have here. So, this is about the bulk arrival models that we have seen today. So, this is what we are going to do. So, for more, of course, you can look at literature, but we are not going to deal with anything beyond at this level for the bulk arrival models here, so we will see in the next lecture.
Thank you bye.