

Stochastic Processes - 1
Dr. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology – Delhi

Lecture – 65
M/M/c Queueing Model

This is a lecture 05, application of a continuous time Markov chain in simple Markovian Queueing models. The first lecture we have discussed the definition of a stochastic process in particular continuous time Markov chain. Then we have considered the Kolmogorov differential equation, Chapman Kolmogorov equation, the transient solutions for the CTMC. The second lecture we have discussed the special case of continuous time Markov chain.

That is a birth death process we have discussed in lecture 02. Lecture 03, the special case of birth death process, which are very important stochastic process that is Poisson process is discussed in the lecture 03. In the lecture 04, we have discussed the M/M/1 queueing model that is a very special and important queueing and then the underlying stochastic process for the M/M/1 queueing model that is a birth death process with the birth rates are λ and death rates are μ .

(Refer Slide Time: 01:57)

Contents

Applications of CTMC:

- M/M/c Queueing Model
- M/M/1/N Queueing Model
- M/M/c/K Queueing Model
- M/M/c/c loss system
- M/M/ ∞
- Finite Source Markovian Queueing Model

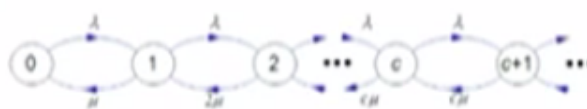
In the fourth lecture, we have discussed only the M/M/1 queueing model. In these lecture we are going to consider the other simple Markovian Queueing models as an application of a continuous time Markov chain. So in these lecture, I am planning to discuss other than

M/M/1 queueing model. I am going to discuss a simple Markovian Queueing models starting with the M/M/c infinity queueing model.

Then the finite capacity model Markovian setup M/M/1/N Queueing model. Then I am going to discuss the multi-server finite capacity model; that is M/M/c/K Queueing model. After that I am going to discuss the loss system; that is M/M/c/c model. For an infinite server model that is M/M infinity also I am going to discuss. At the end, I am going to discuss the Finite Source Markovian Queueing model also.

(Refer Slide Time: 02:59)

M/M/c Queueing Model



- Arrival follows Poisson process with rate λ
- Service times follow exponential distribution with parameter μ
- c servers
- Arriving customer finds n customers in system
 - $n < c$: it is routed to any idle server
 - $n \geq c$: it joins the waiting queue - all servers are busy

Whereas the other 5 models, the population is infinite source, so the last one is the Finite Source Markovian Queueing model also I am going to discuss as the application of continuous time Markov chain. The first model is a multi server infinite capacity Markovian Queueing model. The letter M denote a; the inter arrival time is the exponentially distributed its parameter lambda.

The service time by the each server that is exponentially distributed with the parameter mu and all we have a more than one servers; suppose you can consider as a c, where c is a positive integer and all the servers are identical and each server is doing the service which is exponentially distributed with the parameter mu, which is independent of the all other servers and the service time is independent with the inter arrival time also.

With these assumptions, if you make a random variable x of t is the number of customers in the system at any time t that is a stochastic process. Since the possible values of number of

customers in the system at any time t that is going to be 0, 1, 2 and so on. Therefore, it is a discrete state and you are observing the queueing system at any time t therefore it is a continuous time.

So, discrete state continuous time stochastic process and if you observe the system is keep moving into the different states because of a either arrival or the service completion from the any one of the c servers. So suppose, there are no customer in the system and the system moves from the state 0 to 1 by one arrival. So the inter arrival time is exponentially distributed therefore the rate in which the system is moving from the state 0 to 1 is λ .

Like that you can visualise the rates for a system moving from 1 to 2, 2 to 3 and so on. Whereas, whenever the system size is 1, 2 and so till c , since we have a c number of a servers in the system; whoever entering into the system they will start getting the service immediately. Suppose the system goes from the state 1 to 0, that means the customer enter into the system and he gets the service immediately.

And the service time is exponentially distributed with the parameter μ . Therefore, whenever the service is completed the system goes form the state 1 to 0, therefore the rate is μ . Whereas from 2 to 1, there are 2 customers in the system and both are under service at any time if any one of the servers; if any one of the servers complete the service, then the system moves from 2 to 1.

So the service completion will be minimum of the service time of the both the servers. Since each server is doing the service are exponentially distributed with the parameter μ ; therefore, the minimum of a 2 exponential and both are independent also. Therefore, that is also going to be an exponentially distributed with the sum of parameters so it is going to be parameter will be $\mu + \mu$ that is 2μ .

So this system moves from the state 2 to 1 will be; the rate will be 2μ . Like that it will be keep going till the state from c to $c - 1$, that means we have c servers. Therefore, whenever the systems size is also less than or equal to c ; that means all the customers are under service. Now we will discuss the rate in which the system is moving from the state c plus 1 to c .

The system state is a $c + 1$, that means when the number of customers in the system that is $c + 1$. We have c servers; therefore, one customer will be waiting for the service, waiting in the queue. Therefore, this system is moving from $c + 1$ to c that is nothing but one of the server completed the service out of c servers. Therefore, the rate will be the service time; completion service time will be exponentially distribution with the parameter $c\mu$ not $c + 1\mu$.

It is a , we have only c servers therefore the minimum of a exponentially distributed with the parameters μ and so on with the c exponentially distributed random variables, therefore that is going to be exponentially distribution with the parameter $\mu + \mu + \dots$, whereas $c\mu$ is therefore it is going to be $c\mu$. Like that the rate will be; the death rate will be $c\mu$ after $c + 1$ onwards.

Whereas from 0 to c , it will be μ to 2μ , 3μ and so on till $c\mu$ after that it will be a $c\mu$ from the state from c , $c + 1$ to c , $c + 2$ to $c + 1$ and so on and if you see the state transition diagram, you can observe that it is a birth death process. So before that, let me explain; what is the M/M/c infinity means, whenever c customers or c servers or any one of the c servers are available, then the customers will get the service immediately.

(Refer Slide Time: 09:40)

- Birth-death process with state-dependent death rates

$$\mu_n = \begin{cases} n\mu, & 1 \leq n \leq c \\ c\mu, & n \geq c \end{cases}$$

If all the c servers are busy, then the customers have to wait till any one of the c servers are going to be completing their service, so that is the way the system works. Therefore, you will have the system size, the system size the underlined stochastic process is going to be a birth

death process as a special case of a continuous time Markov chain. Because the transitions are only the neighbours transition with the forwards rates, that is lambda.

And backwards rates are the death rates are going to be mu to 2 mu and so on. Therefore, this is a special case of our continuous time Markov chain, the underlined stochastic process for the M/M/c infinity model that is the birth death process. The birth rates are lambda whereas the death rates depend on the n the mu n is the function of n. Therefore, it is called state dependent death rates.

It might not be the function n times mu, it can be a function of n, then we can use the word state dependent. So here it is a linear function. So state dependent death rates and the death rates are n times mu, whenever n is lies between 1 to c and the mu n is going to be c times mu for n is greater than or equal to c that you can observe it from the state transition diagram also.

(Refer Slide Time: 11:10)

M/M/c Queuing Model

• Steady-state or equilibrium solution when $\frac{\lambda}{c\mu} < 1$

$$p_n = \begin{cases} \frac{\lambda^n}{n! \mu^n} p_0 & 1 \leq n \leq c \\ \frac{\lambda^n}{c^{n-c} c! \mu^n} p_0 & n > c \end{cases}$$

Using normalizing constant

$$\sum_{n=0}^{\infty} p_n = 1 \Rightarrow p_0 = \left[\sum_{n=0}^{c-1} \frac{1}{n!} \left(\frac{\lambda}{\mu}\right)^n + \frac{1}{c!} \left(\frac{\lambda}{\mu}\right)^c \left(\frac{c\mu}{c\mu - \lambda}\right) \right]^{-1}$$

The death rates are going to be c mu here also c mu and so on. Therefore, this is the birth death process with the state dependent death rates. Now our interest is to find out the steady state or equilibrium solution. Since it is an infinite capacity model, if you observed the birth death process with the infinite state space, then you need a condition, so that the steady state probability is exist.

So whenever lambda by c mu is less than 1, whenever lambda by c mu is less than 1, you can find out the limiting probabilities, so sometimes I use the letter pn, sometimes I use the word

p_n , both are one and the same. So you find out the steady state probability by solving a p_n is equal to; p_0 is equal to 0 and the summation of a p_n is equal to 1 and if you recall the birth death processes steady state probabilities the p_0 has the 1 divided by the series.

Whenever the denominator series converges, then you will get the p_n 's. So either I use the p_n 's or p_n 's both are one and the same. So here summation of a p_n is equal to 1 and p and if you make a vector p , p times q is equal to 0, if you solve that equation and the denominator of a p_0 that expression that is going to be converges only if λ divided by $c \mu$ is less than 1. So therefore whenever this condition is there, the queueing system is stable also.

If you put c is equal to 1, you will get the $M/M/1$ q . So using the normalising condition you are getting the p_0 and p_0 is 1 divided by this, so this is a series, so this series is going to be converges only if this condition is satisfied. So by solving that equations, you are getting p and c in terms of p_0 and using normalising constant you are getting a p_0 . Therefore, this is the steady state also known as the equilibrium solution for the $M/M/c$ infinity model.

(Refer Slide Time: 13:40)

The probability that an arriving customer has to wait on arrival

$$= \sum_{n=c}^{\infty} P_n = \frac{P_c}{1-\rho} \quad \left(\text{reg. } c \left(c, \frac{\lambda}{\mu} \right) \right)$$

This is known as Erlang's C formula.

$c \left(c, \frac{\lambda}{\mu} \right)$: fraction of time all the c servers are busy.

Let N_q denote the number of customers in the queue.

$$P[N_q = j, W > 0] = P[N = c+j] = \rho^j P_c$$

$$= \rho^j (1-\rho) c \left(c, \frac{\lambda}{\mu} \right)$$

So here we are using the birth death process with the birth rates are λ and the death rates are given in these form and use the same logic of the stationary distribution for the birth death process. Using that, we are getting the steady state or equilibrium solution for the $M/M/c$ model.