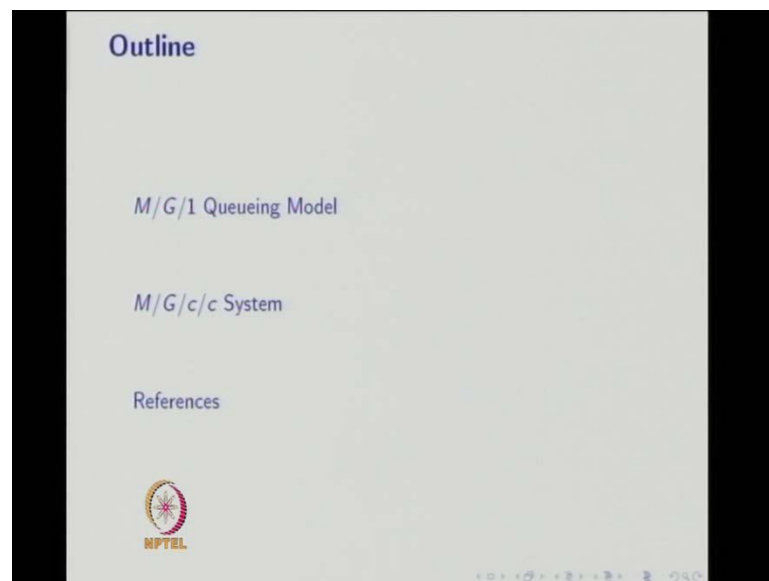


Stochastic Processes
Prof. Dr. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Module - 8
Renewal Processes
Lecture - 4
Non Markovian Queues

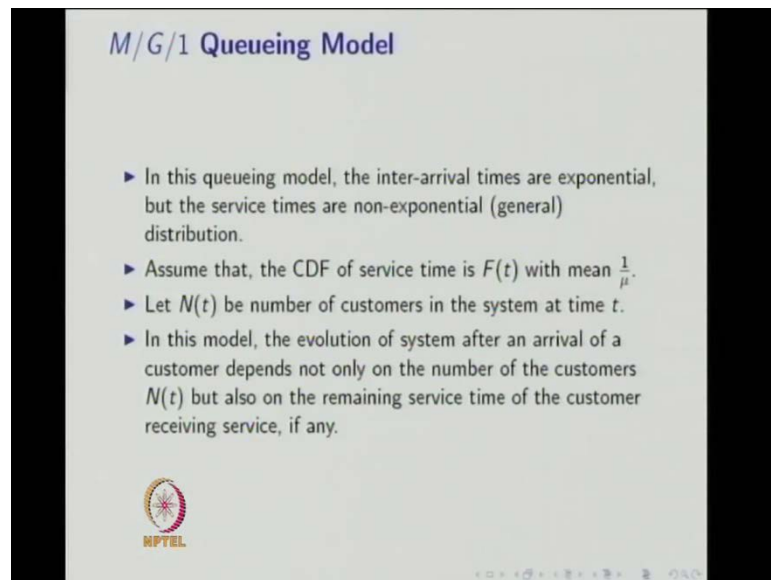
This is stochastic processes, module 8 renewal processes. In the first 3 lectures, we have discussed the renewal processes and its properties. Then later, we have discussed the important limiting theorem on renewal processes. Then, we have discussed the Markov renewal process and Markov regenerative process.

(Refer Slide Time: 01:05)



In between we have discussed the reward renewal process also. In this lecture we are going to cover the non Markovian queues with respect to the service time distribution. In particular, we are going to discuss the M G 1 queue and M G c c system.

(Refer Slide Time: 01:30)



M/G/1 Queueing Model

- ▶ In this queueing model, the inter-arrival times are exponential, but the service times are non-exponential (general) distribution.
- ▶ Assume that, the CDF of service time is $F(t)$ with mean $\frac{1}{\mu}$.
- ▶ Let $N(t)$ be number of customers in the system at time t .
- ▶ In this model, the evolution of system after an arrival of a customer depends not only on the number of the customers $N(t)$ but also on the remaining service time of the customer receiving service, if any.

NPTEL


What is M G 1 queue? In this queueing model the inter arrival times are exponential, but the service times are non exponential, nothing but general distribution. Assume that the CDF of the service time is F of t with the mean 1 by μ . So, suppose you assume that service time is exponential distribution with the mean 1 by μ , then it is a M M 1 queueing model. Let $N(t)$ be the number of customers in the system at time t . In this model, the evaluation of the system after the arrival of customer depends not only on the number of customers N of t , but also the remaining service time of the customers receiving service.

The remaining service is nothing but this is the service times are random variable, the remaining service time is elapsed service or remaining or residual service time. Whenever the service times are exponential distribution, the remaining time or elapsed service time or residual service time all are exponential distributions. Whenever the service times are not a exponential distribution, then the remaining service time will be some other distribution. The evaluation of the system after the service completion depends only on the system of the state. In this queueing model the system changes the state based on two types of transition epochs; one is arrival epochs and other one is service epochs.

(Refer Slide Time: 03:00)

Markov Regenerative Process


- ▶ Whereas the evolution of the system after a service completion depends only on the state of the system.
- ▶ Thus $\{N(t), t \geq 0\}$ is not a semi-Markov process.
- ▶ Suppose that the time origin is taken to be an instant of a departure which left behind exactly j customers.
- ▶ Then, every time a departure occurs leaving behind j customers, the future of $\{N(t)\}$ after such time has exactly the same probability law as the process $\{N(t), t \geq 0\}$ had starting at time 0.
- ▶ One can observe that, $\{N(t), t \geq 0\}$ is a Markov regenerative process (MRGP).



(Refer Slide Time: 03:22)

M/G/1 Queueing Model

- ▶ In this queueing model, the inter-arrival times are exponential, but the service times are non-exponential (general) distribution.
- ▶ Assume that, the CDF of service time is $F(t)$ with mean $\frac{1}{\mu}$.
- ▶ Let $N(t)$ be number of customers in the system at time t .
- ▶ In this model, the evolution of system after an arrival of a customer depends not only on the number of the customers $N(t)$ but also on the remaining service time of the customer receiving service, if any.




First, we are discussing the arrival scenario the evaluation of the system after the arrival of a customer depends not only the number of customers N of t , but also the remaining service time.

(Refer Slide Time: 03:37)

Markov Regenerative Process

- ▶ Whereas the evolution of the system after a service completion depends only on the state of the system.
- ▶ Thus $\{N(t), t \geq 0\}$ is not a semi-Markov process.
- ▶ Suppose that the time origin is taken to be an instant of a departure which left behind exactly j customers.
- ▶ Then, every time a departure occurs leaving behind j customers, the future of $\{N(t)\}$ after such time has exactly the same probability law as the process $\{N(t), t \geq 0\}$ had starting at time 0.
- ▶ One can observe that, $\{N(t), t \geq 0\}$ is a Markov regenerative process (MRGP).




Whereas, the evaluation of the system after service completion depends only on the state of the system. Therefore, N of t is not a semi Markov process because it depends on the remaining service time.

(Refer Slide Time: 04:06)

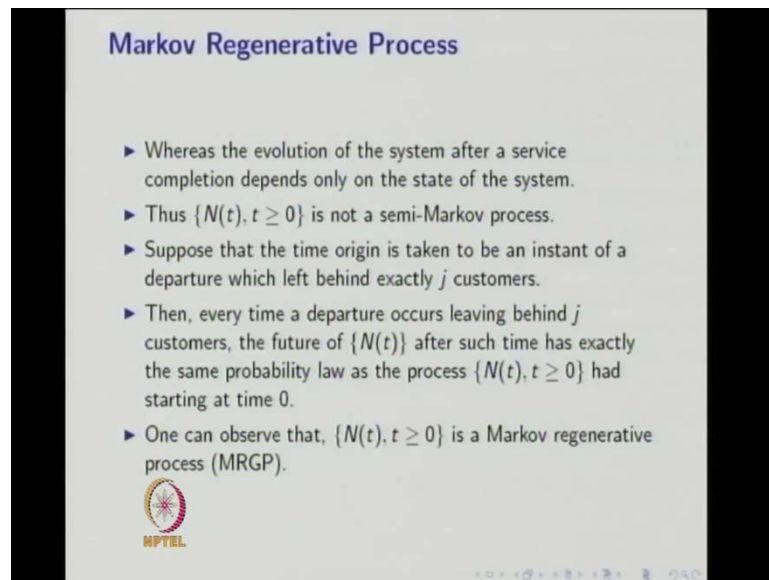
M/G/1 Queueing Model

- ▶ In this queueing model, the inter-arrival times are exponential, but the service times are non-exponential (general) distribution.
- ▶ Assume that, the CDF of service time is $F(t)$ with mean $\frac{1}{\mu}$.
- ▶ Let $N(t)$ be number of customers in the system at time t .
- ▶ In this model, the evolution of system after an arrival of a customer depends not only on the number of the customers $N(t)$ but also on the remaining service time of the customer receiving service, if any.




Since the, since the after the arrival of a customer the remaining service time also play role.

(Refer Slide Time: 04:12)



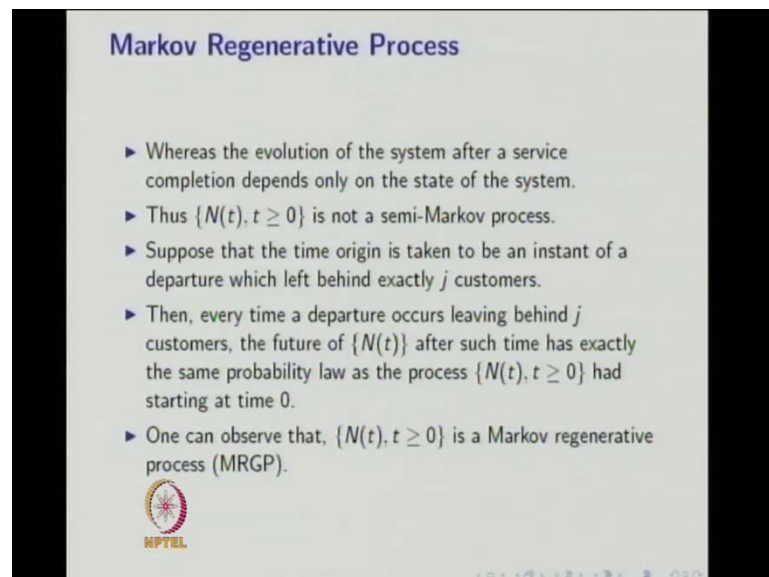
Markov Regenerative Process

- ▶ Whereas the evolution of the system after a service completion depends only on the state of the system.
- ▶ Thus $\{N(t), t \geq 0\}$ is not a semi-Markov process.
- ▶ Suppose that the time origin is taken to be an instant of a departure which left behind exactly j customers.
- ▶ Then, every time a departure occurs leaving behind j customers, the future of $\{N(t)\}$ after such time has exactly the same probability law as the process $\{N(t), t \geq 0\}$ had starting at time 0.
- ▶ One can observe that, $\{N(t), t \geq 0\}$ is a Markov regenerative process (MRGP).




Whereas, after the service completion, it depends only on the state. If you recall the definition of semi Markov process, the memory less property will be satisfied at all time transition instance. The memory less property will be satisfied at all time transition instance. Then only the stochastic process is called as semi Markov process.

(Refer Slide Time: 04:44)



Markov Regenerative Process

- ▶ Whereas the evolution of the system after a service completion depends only on the state of the system.
- ▶ Thus $\{N(t), t \geq 0\}$ is not a semi-Markov process.
- ▶ Suppose that the time origin is taken to be an instant of a departure which left behind exactly j customers.
- ▶ Then, every time a departure occurs leaving behind j customers, the future of $\{N(t)\}$ after such time has exactly the same probability law as the process $\{N(t), t \geq 0\}$ had starting at time 0.
- ▶ One can observe that, $\{N(t), t \geq 0\}$ is a Markov regenerative process (MRGP).



Therefore, here the N of t is not a semi Markov process. Suppose, the time of time origin is taken to be an instant of departure, which left behind exactly j customers, then every time the departure occurs leaving behind j customers, the future of N of t after such time

has exactly the same probability law as the process had started starting at time 0. So, this is so called the probabilistic replica with the Markov property as well as time homogeneity.

If you recall the definition of a Markov region rating process, a stochastic process has the property of probabilistic replica with the few time instance it has the Markov property with the time homogeneity with few states at few states. Therefore, N of t is a Markov region rating process. Please refer the previous lecture for understanding the definition of a Markov region rating process. So, by using the definition of Markov region rating process the N of t will be the Markov region rating process.

(Refer Slide Time: 06:12)

Markov Renewal Process

- ▶ Let n th customer depart at time point t_n .
- ▶ Let X_n be the number of customers in the system just after the departure instant of n th customer.
- ▶ Then $\{(X_n, t_n), n = 0, 1, \dots\}$ is a Markov renewal process, where t_n is the instant when the n^{th} customer departs and $X_n = N(t_n + 0)$.
- ▶ Suppose $Y(t) = X_n, t_n \leq t < t_{n+1}, n = 1, 2, \dots$, then $\{Y(t), t \geq 0\}$ will be a semi-Markov process having embedded discrete time Markov chain (DTMC) $\{X_n, n = 0, 1, \dots\}$.



If you consider the time epochs are departure time epochs. Let N nth customer departure at time instant t_N . Let X_n be the number of customer in the system just after the departure instant of n th customer. That means X_n is nothing but the N of t_n plus 0 just after the departure instant of n th customer that will be treated as the random variable X_n . Then X_n comma t_n will form a Markov renewal process. The t_n are the time instant at the departure epochs not the arrival epochs; only the departure epochs that collection of time points along with x_n form a Markov renewal process, where X_n is n of t_n plus 0.

(Refer Slide Time: 07:30)

Markov Renewal Process

- ▶ Let n th customer depart at time point t_n .
- ▶ Let X_n be the number of customers in the system just after the departure instant of n th customer.
- ▶ Then $\{(X_n, t_n), n = 0, 1, \dots\}$ is a Markov renewal process, where t_n is the instant when the n^{th} customer departs and $X_n = N(t_n + 0)$.
- ▶ Suppose $Y(t) = X_n, t_n \leq t < t_{n+1}, n = 1, 2, \dots$, then $\{Y(t), t \geq 0\}$ will be a semi-Markov process having embedded discrete time Markov chain (DTMC) $\{X_n, n = 0, 1, \dots\}$.



That in words it is the number of customers in the system just after the departure instant of n th customer leaves. We can create a semi Markov process by Y of t is equal to X_n . If that value will be between t_n to t_{n+1} , if it is a same, then it will form a semi Markov process.

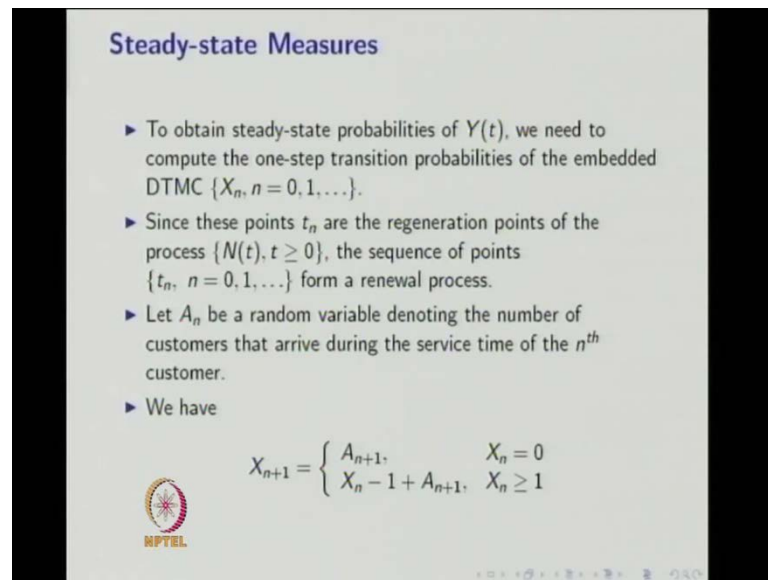
(Refer Slide Time: 08:11)

Markov Renewal Process

- ▶ Let n th customer depart at time point t_n .
- ▶ Let X_n be the number of customers in the system just after the departure instant of n th customer.
- ▶ Then $\{(X_n, t_n), n = 0, 1, \dots\}$ is a Markov renewal process, where t_n is the instant when the n^{th} customer departs and $X_n = N(t_n + 0)$.
- ▶ Suppose $Y(t) = X_n, t_n \leq t < t_{n+1}, n = 1, 2, \dots$, then $\{Y(t), t \geq 0\}$ will be a semi-Markov process having embedded discrete time Markov chain (DTMC) $\{X_n, n = 0, 1, \dots\}$.



(Refer Slide Time: 08:20)



Steady-state Measures

- ▶ To obtain steady-state probabilities of $Y(t)$, we need to compute the one-step transition probabilities of the embedded DTMC $\{X_n, n = 0, 1, \dots\}$.
- ▶ Since these points t_n are the regeneration points of the process $\{N(t), t \geq 0\}$, the sequence of points $\{t_n, n = 0, 1, \dots\}$ form a renewal process.
- ▶ Let A_n be a random variable denoting the number of customers that arrive during the service time of the n^{th} customer.
- ▶ We have

$$X_{n+1} = \begin{cases} A_{n+1}, & X_n = 0 \\ X_n - 1 + A_{n+1}, & X_n \geq 1 \end{cases}$$

NPTEL

That means the system is not moving any other states in between the time duration, then the Y of t will be a semi Markov process. As such N of t is not a semi Markov process. Whereas, N X n comma t_n will form a Markov renewal process and the N of t will be a Markov regenerative process.

With embedded the Y of t will be a semi Markov process, if I make Y of t is equal to X_n , where t lies between t_n to t_{n+1} and having the embedded DTMC X_n . Our interest is to find out the steady state measures. To obtain the steady state probability of Y t , we need the one step transition probabilities of embedded DTMC. Suppose, these time points t_n are the regeneration points, then the t_n 's will form a renewal process. Let A_n be the random variable denoting the number of customers, that arrive during the service time of n^{th} customer. That means X_{n+1} will be A_{n+1} , if X_n was 0. If X_n was greater than or equal to 1.

Then X_{n+1} will be $X_n - 1 + A_{n+1}$. That is nothing but how many customers in the system when the $n+1^{\text{th}}$ customer leaves, that is same as when the n^{th} person leaves how many customers in the system minus 1 is for the $n+1^{\text{th}}$ customer himself. How many customers enter into the system during his service, that is A_{n+1} , whenever the X_n was greater than or equal to 1. If X_n was 0, then the number of customers will be in the system then $n+1^{\text{th}}$ customer leaves will be same.

(Refer Slide Time: 10:08)

Steady-state Measures ...

- ▶ Since service times of all the customers, denoted by B , have the same distribution, the distribution of A_n is same for all n .
- ▶ Denoting, for all n ,

$$\begin{aligned} a_r &= \Pr(A_n = r) \\ &= \int_0^\infty \frac{e^{-\lambda t} (\lambda t)^r}{r!} dB(t), \quad r = 0, 1, \dots \end{aligned}$$

- ▶ Therefore,

$$\begin{aligned} P_{ij} &= \Pr\{X_{n+1} = j / X_n = i\} \\ &= \begin{cases} a_j, & j \geq 0, i = 0 \\ a_{j-i+1}, & i \geq 1, j \geq i-1 \\ 0, & i \geq 1, j < i-1 \end{cases} \end{aligned}$$



As number of customers who enter into the system during the during his service time. Since, the service times of all all the customers have the same distribution the distribution of AS n is same for all n . Therefore, you can go for, for all n , we can go for a r is the r customers enter into the system during the any customer. So, that is probability mass function of a A_n . Now, you can find out the transition probability of system moving from the state i to j with respect to the stochastic process X_n . X_n is embedded Markov chain in the N of t . So, that will be a j , if j is greater than or equal to 0 and i equal to 0, similarly a j minus i plus 1 or 0, according to the values of i and j .

(Refer Slide Time: 11:20)

Steady-state Measures ...

- ▶ Denoting by $P = [P_{ij}]$, we have

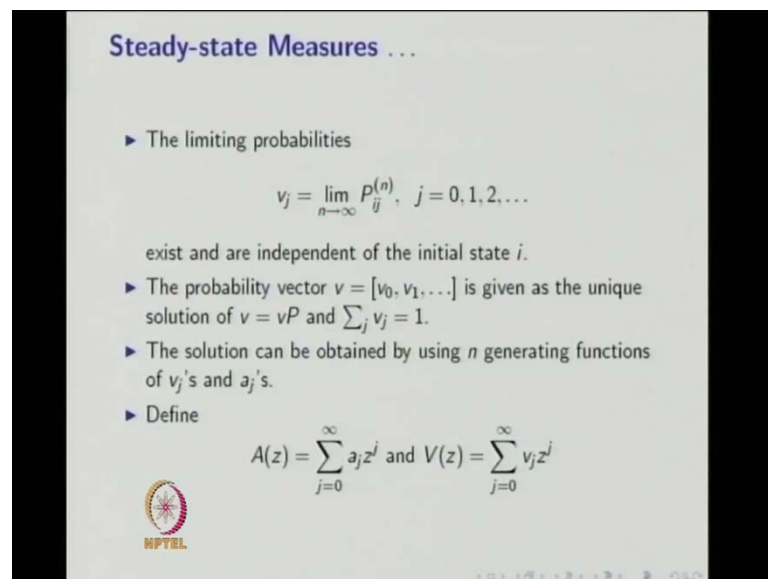
$$P = [P_{ij}] = \begin{bmatrix} a_0 & a_1 & a_2 & \dots & \dots \\ a_0 & a_1 & a_2 & \dots & \dots \\ 0 & a_0 & a_1 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

- ▶ Note that X_n is irreducible Markov chain.
- ▶ When $\rho = \frac{\lambda}{\mu} < 1$, the chain is positive recurrent.
- ▶ Hence the Markov chain is ergodic.



Now, you can make one step transition probability matrix, that is nothing but the P , the row sums are going to be 1. Note that X_n is irreducible Markov chain because each state is reachable by every other states. If we make ρ , that is nothing but λ/μ is less than 1, λ is the parameter for the inter arrival time which is exponential distribution. So, whenever ρ is less than 1, the chain is a positive recurrent. ρ is less than 1 is required because only then the mean recurrence time will be finite. Hence, the Markov chain will be a positive recurrent.

(Refer Slide Time: 12:22)




Steady-state Measures ...

- ▶ The limiting probabilities

$$v_j = \lim_{n \rightarrow \infty} P_{ij}^{(n)}, \quad j = 0, 1, 2, \dots$$
 exist and are independent of the initial state i .
- ▶ The probability vector $v = [v_0, v_1, \dots]$ is given as the unique solution of $v = vP$ and $\sum_j v_j = 1$.
- ▶ The solution can be obtained by using n generating functions of v_j 's and a_j 's.
- ▶ Define

$$A(z) = \sum_{j=0}^{\infty} a_j z^j \quad \text{and} \quad V(z) = \sum_{j=0}^{\infty} v_j z^j$$



Therefore, it is ergodic. Irreducible positive recurrent gives the Markov chain is ergodic. Therefore, the Markov chain is the embedded Markov chain is ergodic.

So, we can find out the limiting probabilities that will exist as it is independent of initial state i . You can find out the limiting probabilities by solving v is equal to vP and summation of v_j is equal to 1. The solution can be obtained by using the probability generating function of v 's and a 's v 's are probability vector steady state probability vector. a 's are the number of customers arrived during the service of any customer.


(Refer Slide Time: 13:25)

Steady-state Measures ...

► Now,

$$\begin{aligned}
 A(z) &= \sum_{j=0}^{\infty} a_j z^j \\
 &= \sum_{j=0}^{\infty} z^j \left(\int_0^{\infty} \frac{e^{-\lambda t} (\lambda t)^j}{j!} dF(t) \right) \\
 &= \int_0^{\infty} e^{-\lambda t} e^{\lambda z t} f(t) dt \\
 &= \int_0^{\infty} e^{-t(\lambda - \lambda z)} f(t) dt \\
 &= f^*(\lambda - \lambda z)
 \end{aligned}$$

Here $f^*(s)$ is the Laplace transform of $f(t)$.



So, $A(z)$ is nothing but the probability generating function for a_j 's and a_j 's are probabilities of z is nothing but probability generating function for the random variable for the probabilities a_j 's. Now, you can find out a_j of z in terms of the Laplace transform of the function F of t . Because A of z is nothing but the summation of a_j into z power j and you can replace a_j by the integration. Now, integration and summation can be interchanged and F of t is nothing but F of t is nothing but the distribution of service time.

(Refer Slide Time: 14:00)


Steady-state Measures ...

► Since service times of all the customers, denoted by B , have the same distribution, the distribution of A_n is same for all n .

► Denoting, for all n ,

$$\begin{aligned}
 a_r &= Pr(A_n = r) \\
 &= \int_0^{\infty} \frac{e^{-\lambda t} (\lambda t)^r}{r!} dF(t), \quad r = 0, 1, \dots
 \end{aligned}$$

► Therefore,

$$\begin{aligned}
 P_{ij} &= Pr\{X_{n+1} = j | X_n = i\} \\
 &= \begin{cases} a_j, & j \geq 0, i = 0 \\ a_{j-i+1}, & i \geq 1, j \geq i-1 \\ 0, & i \geq 1, j < i-1 \end{cases}
 \end{aligned}$$


F of t is nothing but the distribution of service time, therefore the small f of t is a probability density function of the service time. So, the f star is nothing but the Laplace transform of F of t.


(Refer Slide Time: 14:20)

Steady-state Measures ...

► Now,

$$\begin{aligned}
 A(z) &= \sum_{j=0}^{\infty} a_j z^j \\
 &= \sum_{j=0}^{\infty} z^j \left(\int_0^{\infty} \frac{e^{-\lambda t} (\lambda t)^j}{j!} dF(t) \right) \\
 &= \int_0^{\infty} e^{-\lambda t} e^{\lambda z t} f(t) dt \\
 &= \int_0^{\infty} e^{-t(\lambda - \lambda z)} f(t) dt \\
 &= f^*(\lambda - \lambda z)
 \end{aligned}$$

Here $f^*(s)$ is the Laplace transform of $f(t)$.




So, the probability generating function of a of j's is nothing but the Laplace transform of probability density function of a service time distribution.

(Refer Slide Time: 14:40)

Steady-state Measures ...

► Expected number of arrivals during service time is given by

$$\begin{aligned}
 &= A'(1) = \frac{d}{dz} A(z) \Big|_{z=1} \\
 &= \frac{d}{dz} f^*(\lambda - \lambda z) \Big|_{z=1} \\
 &= \frac{df^*}{ds} \cdot \frac{ds}{dz} \Big|_{z=1} \\
 &= \frac{df^*}{ds} \Big|_{s=0} \cdot \frac{d(\lambda - \lambda z)}{dz} \Big|_{z=1} \\
 &= -\lambda \frac{d}{ds} \left(\int_0^{\infty} e^{-st} f(t) dt \right) \Big|_{s=0} \\
 &= -\lambda \left(\int_0^{\infty} -t f(t) dt \right) \\
 &= \frac{\lambda}{\mu} = \rho
 \end{aligned}$$



Now, we can find the expected number of arrivals during the service time that is nothing but since we got the probability generating function differentiate with respect to z and

substitute z is equal to 1 will be the average number of arrivals during the any customers service time. So, you can simplify that by using the Chain rule.

(Refer Slide Time: 15:14)

Steady-state Measures ...

► Expected arrival is given by

$$\begin{aligned}
 E(\text{Arrival}) &= A'(1) = \frac{d}{dz} A(z) \Big|_{z=1} \\
 &= \frac{d}{dz} f^*(\lambda - \lambda z) \Big|_{z=1} \\
 &= \frac{df^*}{ds} \cdot \frac{ds}{dz} \Big|_{z=1} \\
 &= \frac{df^*}{ds} \Big|_{s=0} \cdot \frac{d(\lambda - \lambda z)}{dz} \Big|_{z=1} \\
 &= -\lambda \frac{d}{ds} \left(\int_0^\infty e^{-st} f(t) dt \right) \Big|_{s=0} \\
 &= -\lambda \left(\int_0^\infty -t f(t) dt \right) \\
 &= \frac{\lambda}{\mu} = \rho
 \end{aligned}$$


So, that will be nothing but λ by μ where λ by μ is nothing but the ρ . So, the expected number of arrivals is ρ . ρ is nothing but λ by μ .


(Refer Slide Time: 15:25)

Steady-state Measures ...

► Now,

$$v_j = v_0 a_j + \sum_{i=0}^{j-1} v_i a_{j-i-1}, \quad j = 0, 1, 2, \dots$$

► Multiplying by z^j on both the sides and taking the sum, we get

$$\begin{aligned}
 \sum_{j=0}^{\infty} V_j z^j &= \sum_{j=0}^{\infty} v_0 a_j z^j + \sum_{j=0}^{\infty} \left(\sum_{i=1}^{j+1} v_i a_{j-i+1} \right) z^j \\
 V(z) &= v_0 \sum_{j=0}^{\infty} a_j z^j + \sum_{i=1}^{\infty} \left(\sum_{j=i-1}^{\infty} v_i a_{j-i+1} z^j \right) \\
 &= v_0 A(z) + \frac{1}{z} [V(z) - v_0] A(z) \\
 V(z) - \frac{1}{z} A(z) V(z) &= v_0 A(z) - \frac{1}{z} v_0 A(z) \\
 V(z) &= \frac{v_0 A(z)(z-1)}{z - A(z)}
 \end{aligned}$$


Our interest is to find out the steady state probability probabilities V_j 's. So, the V_j 's you can relate in terms of by expanding V is equal to V_p you will get V_j is equal to V naught a j 's plus summation form. So, multiply z^j in both sides and taking the

summation, you will get after doing, after simplification we will get $V(z)$ in terms of $A(z)$ with $V(0)$. $V(0)$ is no, $V(0)$ is the steady state probability of no customer.

(Refer Slide Time: 16:15)

Pollaczek-Khinchin (P-K) Formula

► We know that $V(1) = A(1) = 1$,

$$V(1) = \lim_{z \rightarrow 1} v_0 \left(\frac{A'(z)(z-1) + A(z)}{1 - A'(z)} \right)$$

$$1 = \frac{v_0 A(1)}{1 - A'(1)} = \frac{v_0}{1 - A'(1)}$$


provided $A'(1)$ is finite and is less than 1.

► Taking $\rho = A'(1)$ we get, $v_0 = 1 - \rho$.

► Hence,

$$V(z) = \frac{(1 - \rho)(z - 1)f^*(\lambda(1 - z))}{z - f^*(\lambda(1 - z))}.$$

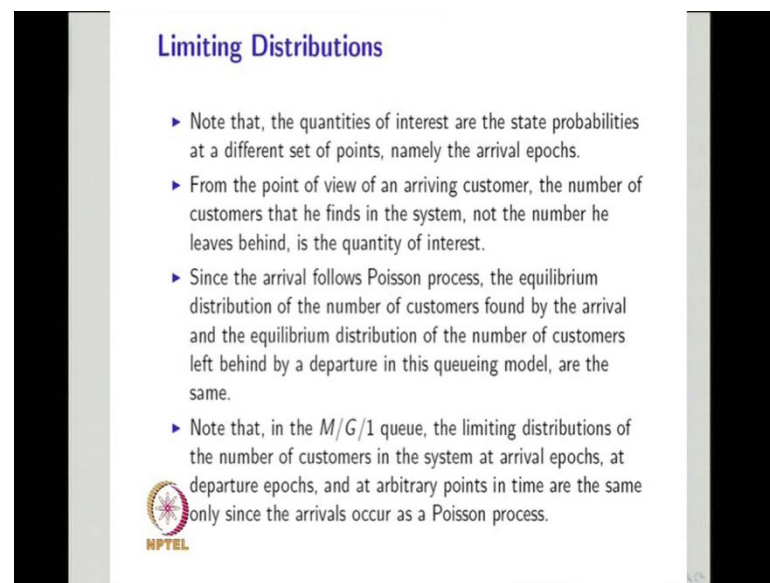
► This is known as **Pollaczek-Khinchin (P-K) formula**.



Sorry $V(z)$ in terms of $A(z)$ and $V(0)$. You know that since $V(z)$ and $A(z)$ are probability generating functions. So, v_0 and $A(1)$ are equal to be 1. So, using that relation, you can find out what is a value of $V(0)$. So, $V(0)$ is nothing but $1 - A'(1)$ and $A'(1)$ is nothing but the expected arrivals. Just now we got expected arrival is equal to ρ , therefore $V(0)$ is equal to $1 - \rho$.


So, here we got $V(z)$ in terms of $A(z)$ with $V(0)$ and already we have the relation $A(z)$ is a Laplace transform of probability density function of service time distribution. Just now we got $V(0)$ in terms of $1 - \rho$, $V(0)$ is equal to $1 - \rho$. Therefore, the probability generating function of the steady state probabilities will be in terms of Laplace transform of probability density function of service time distribution with ρ . This equation is known as Pollaczek Khinchin or PK formula. So, this is the formula used to find out the steady state probabilities because if you get the coefficient of z that is steady state probabilities.

(Refer Slide Time: 17:45)



Limiting Distributions

- ▶ Note that, the quantities of interest are the state probabilities at a different set of points, namely the arrival epochs.
- ▶ From the point of view of an arriving customer, the number of customers that he finds in the system, not the number he leaves behind, is the quantity of interest.
- ▶ Since the arrival follows Poisson process, the equilibrium distribution of the number of customers found by the arrival and the equilibrium distribution of the number of customers left behind by a departure in this queueing model, are the same.
- ▶ Note that, in the $M/G/1$ queue, the limiting distributions of the number of customers in the system at arrival epochs, at departure epochs, and at arbitrary points in time are the same only since the arrivals occur as a Poisson process.

 NPTEL


Note that the quantities of interest are the state probabilities at a different set of points, namely the arrival epochs from the past of view of arriving customer. The number of customers that he finds in the system not the number, he leaves behind is the quantity of interest. Since, the arrival follows Poisson process, the equilibrium distribution of the number of customers find found by the arrival. The equilibrium distribution of the number of customers left behind departure in this queuing model are the same.

Note that the the $M/G/1$ the limiting distribution of the number of customers in the system at arrival epochs and the departure epochs and at the arbitrary time points are the same only, since the arrival occurs as a Poisson process. So, we found the limiting probabilities at the departure epochs, but since the arrival follows the Poisson process, the limiting distribution of number of customers in the system at the arrival epochs. The departure epochs and at the arbitrary time points all are same.

(Refer Slide Time: 19:07)

P-K Mean Value Formula

- ▶ Average number of customers in the system in steady-state is given by
$$L_s = \left. \frac{dV(z)}{dz} \right|_{z=1}$$
- More generally,
$$L_s = \rho + \frac{\lambda^2(E(B^2))}{2(1-\rho)}.$$
- ▶ This is known as the P-K mean value formula.
- ▶ Here, $E(B^2)$ is the second order moment about the origin for the service time.
- ▶ This result holds true for all scheduling disciplines in which the server is busy if the queue is non-empty.
- ▶ The expected system size L_s can be computed without $V(z)$ also.



This is the standard result, which we shall be using without proof. Now, you can find out the average measures average number of customers in the system in steady state is given by differentiate the probability generating function. Substitute z is equal to 1 will be the average number of customers. If you do the simplification, you will get ρ times ρ plus λ square expectation of the service time distribution whole square expectation of B square divided by 2 times 1 minus ρ .

This equation is known as Pk mean formula. Previous one was the PK formula because that gives the steady state probabilities. Whereas, this gives the average measures, therefore this is called the PK mean formula. Here expectation of B square is the second order moment about the origin for service time. This result holds true for all schedule link discipline in which the server is busy if the queue is non empty.

(Refer Slide Time: 20:35)

Steady-state Measures ...

►


$$L_s = E(X_{n+1}) = E(X_n)$$

$$L_s = L_s - E(U(X_n)) + E(Arrival)$$

where

$$U(X_n) = \begin{cases} 1, & X_n > 0 \\ 0, & X_n = 0 \end{cases}$$

►

$$\begin{aligned} E(U(X_n)) &= E(Arrival) \\ &= \int_0^\infty E(Arrival|s=t)dB(t) \\ &= \int_0^\infty \lambda t b(t) dt \\ &= \lambda E(B) = \frac{\lambda}{\mu} = \rho \end{aligned}$$


The expected system size L_s can be completed without V_z also. Because if you know ρ as well as if you know the expectation of B^2 you can find out the expected system size. Now, we are deriving the L_s in a different way not via the PK formula. So, the derivation is as follows. So, the L_s can be written in terms of expectation of X_{n+1} and you can find out expectation of U of X_n using this you will get is equal to ρ .

(Refer Slide Time: 21:05)


Steady-state Measures ...

► Now,

$$\begin{aligned} X_{n+1} &= X_n - U(X_n) + A_n \\ X_{n+1}^2 &= X_n^2 + U^2(X_n) + A_n^2 \\ &\quad - 2X_n U(X_n) - 2A_n U(X_n) + 2A_n X_n \end{aligned}$$

where $U^2(X_n) = U(X_n)$ and $X_n U(X_n) = X_n$.

► Now, taking expectations on both the sides, we get

$$\begin{aligned} E(X_{n+1}^2) &= E(X_n^2) + E(U^2(X_n)) + E(A_n^2) \\ &\quad - 2E(X_n U(X_n)) - 2E(A_n U(X_n)) + 2E(A_n X_n) \end{aligned}$$


After you do the simplification, you can get the expectation of X_{n+1} square also in terms of expectation of X_n square as well as expectation of A_n square.

(Refer Slide Time: 21:19)

P-K Mean Formula


► Using $E(X_{n+1}^2) = E(X_n^2)$ and $\text{Var}(\text{Arrival}) = \rho + \lambda^2 \sigma_B^2$ where σ_B^2 is the variance of the service time distribution, we get

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

which is the P-K mean formula.

► Using this, other measures such as L_q , T_s and T_q can be obtained as follows

$$L_q = L_s - \lambda E(B); \quad E(B) = \frac{1}{\mu}$$

$$T_s = \frac{L_s}{\lambda}; \quad T_q = \frac{L_q}{\lambda}$$


So, once you know the expectation of X_{n+1} square and variance of arrival where variance of arrival means the variance of number of arrivals during a service time. You can get the expected number of customers in the system in steady state. So, this is called the PK mean formula without using the PK formula. Once you know the L_s that is average number of customers in the system, you can find out the L_q , that is average number of customers in the queue. T_s is nothing but total time spent in the system and the T_q is, nothing but the sorry, T_s is nothing but the average time spent in the system and the T_q is average time spent in the queue. Using Little's formula, you can find the all other measures.


(Refer Slide Time: 22:17)

Special Case

- For the case where the service time is constant, $\text{Var}(B) = 0$, then the P-K formula for $M/D/1$ queue reduces to

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

where $\rho = \lambda/\mu$ and $1/\mu$ is the constant service time.



As a special case, if the variance is 0, variance of service time is 0 that means it is a M D 1 queue.

(Refer Slide Time: 22:48)


Steady-state Measures ...

- Alternatively, the steady state probabilities can be obtained in the following way. When the embedded DTMC is irreducible and ergodic, the limiting probability $v_j = \lim_{n \rightarrow \infty} P_{ij}^{(n)}$, $j = 0, 1, 2, \dots$ are given as unique solutions of

$$v_j = a_j v_0 + \sum_{i=1}^{j+1} a_{j-i+1} v_i, \quad j = 0, 1, 2, \dots \quad (1)$$

- Suppose the mean sojourn time in state j , $\frac{1}{\mu_j}$ (finite) and $\{X_n, n = 0, 1, \dots\}$ is irreducible and ergodic, then

$$P_j = \frac{\frac{1}{\mu_j} v_j}{\sum_k \frac{1}{\mu_k} v_k}, \quad j = 0, 1, 2, \dots$$


 v_j are obtained by solving (1).

Then you can get the average number of customers in steady state will be rho plus rho square divided by 2 times 1 minus rho. Here the rho is nothing but lambda by mu where 1 by mu is a constant service time. Alternatively you can find out the steady state probabilities by solving V is equal to Vp .

(Refer Slide Time: 23:31)

Special Case

- ▶ Here also, as a special case, when the service time follows exponential distribution with mean $\frac{1}{\mu}$ can be discussed.
- ▶ Assume that $\lambda < \mu$. The one step transition probability matrix P becomes

$$P = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots \\ \frac{\mu}{\lambda+\mu} & 0 & \frac{\lambda}{\lambda+\mu} & \dots & \dots \\ 0 & \frac{\mu}{\lambda+\mu} & 0 & \frac{\lambda}{\lambda+\mu} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$


You can use the mean sojourn time that is 1 divided by μ_j for the state j . Then you can find out the steady state probabilities P_j 's nothing but average sojourn time multiplied by the steady state probabilities of embedded Markov chain, the way we have done it in the semi Markov process.

(Refer Slide Time: 23:47)

Special Case ...


- ▶ Then the solution of $V = VP$ is

$$v_j = \left(\frac{\lambda+\mu}{\lambda}\right) \left(\frac{\lambda}{\mu}\right)^j v_0, \quad j = 1, 2, \dots$$

- ▶ Hence, for the $M/M/1$ queueing system, the limiting probabilities are given by

$$\begin{aligned} \pi_j &= \frac{\left(\frac{\lambda+\mu}{\lambda}\right) \left(\frac{\lambda}{\mu}\right)^j \left(\frac{1}{\mu}\right)}{1 + \sum_{k=1}^{\infty} \left(\frac{1}{\mu}\right) \left(\frac{\lambda+\mu}{\lambda}\right) \left(\frac{\lambda}{\mu}\right)^k}, \quad j = 0, 1, \dots \\ &= \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^j, \quad j = 0, 1, \dots \end{aligned}$$

- ▶ The expected number of customers in the system is given by the Pollazek-Khintchine formula.



As a special case, you can verify when the service time follows exponential distribution with mean $1/\mu$ you can get the 1 step transition probability matrix, in this form for the embedded Markov chain. Then you can solve $V = VP$ with the summation


of V_i is equal to 1, we will get V_j 's. Then if you substitute the V_j 's as well as substitute the V_j 's, then you can get the steady state probabilities for the M M 1 queue, that is same as one minus rho times rho power j.

(Refer Slide Time: 24:22)

Steady-state Measures ...

- ▶ The formula states that mean queueing delay is given by:
$$T_q = \frac{\rho(1 + \mu^2 \sigma_B^2)}{2\mu(1 - \rho)}$$
- ▶ Hence, the average time spent in the system is:
$$T_s = T_q + \frac{1}{\mu} = \frac{1 + \mu^2 \sigma_B^2}{2(\mu - \lambda)} + \frac{1}{\mu}$$
- ▶ By Little's formula, we know: $L_s = \lambda T_s$.
- ▶ Hence
$$L_s = \rho + \frac{\rho^2 + \lambda^2 \sigma_B^2}{2(1 - \rho)}$$

where $1/\mu$ is the mean and σ_B^2 is the variance of the service time distribution and $\rho = \frac{\lambda}{\mu}$.




For any M G 1 queue, you can find out the average time spent in the queue. That is a mean queueing delay. Once you know the mean queueing delay, if you add the average time spent in the, you can get the average time spent in the system by adding average service time. That is one by mu, so the T_s will be T_q plus $1/\mu$ that will be the average time spent in the system. By Little's formula, you can get L_s . L_s is equal to lambda time T_s . You know the T_s , so from that you can get the lambda s, which is same as what you got it in the PK formula, PK mean formula.

As a special case when the service time is exponential distribution with the mean $1/\mu$, you can get the average time spent in the average number of customers in the system will be rho divided by $1 - \rho$. This is same as the average number of customers in the M M 1 queue. In this derivation, we assume first come first serve scheduling to simplify the analysis.

(Refer Slide Time: 25:14)

Special Case

- ▶ When service time is exponentially distributed with mean $\frac{1}{\mu}$, then
$$L_s = \rho + \frac{2\rho^2}{2(1-\rho)} = \frac{\rho}{1-\rho}$$
- ▶ In this derivation, we assume FCFS scheduling to simplify the analysis.
- ▶ However, the above formulae are valid for any scheduling discipline in which the server is busy if the queue is non-empty, no customer departs the queue before completing service, and the order of service is not dependent on the knowledge about service times.
- ▶ Hence, using Little's formula, the average time spent in the system is given by $\lambda T_s = L_s$.




But the above formula are valid for any scheduling discipline in which the server is busy, if the queue is non empty. Then no customer departs the queue before completing the service and order of service not dependent on the knowledge about the service time. If these conditions are satisfied, then for any scheduling discipline, you can use the above formula of the average number of customers in the system.

(Refer Slide Time: 26:23)

Example

- ▶ People are entering cricket stadium at New Delhi to watch a cricket match.
- ▶ There is only one ticket line to purchase tickets. Each ticket purchase takes an average of 20 seconds.
- ▶ The average arrival rate is 2 persons per minute.
- ▶ Find the average length of queue and average waiting time in queue assuming $M/G/1$ queueing with service time follows uniform distribution between 15 and 25 seconds.
- ▶ Departure rate: $\mu = 20$ seconds/person or 3 persons/minute.
- ▶ Arrival rate: $\lambda = 2$ persons/minute
- ▶ $\rho = 2/3$
- ▶ Given $B \sim U(15, 25)$, $E(B) = 1/3$ minute

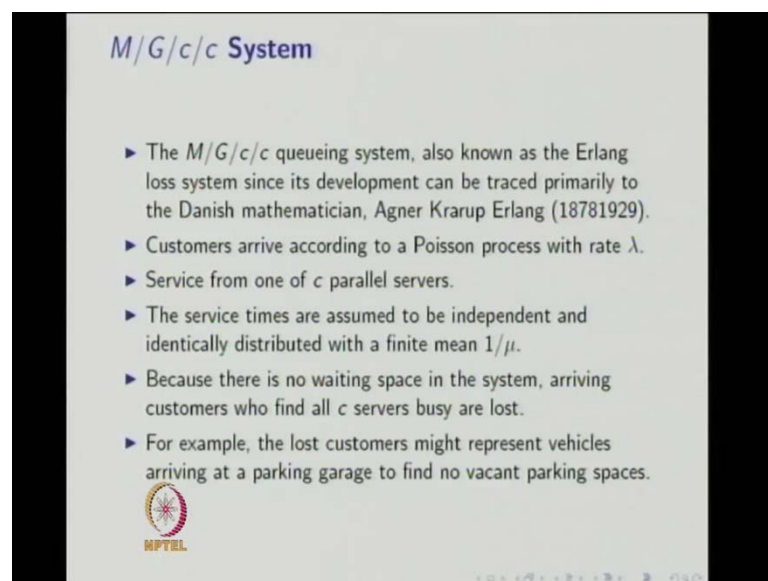


So, using Little's formula, you can find out the average time spent in the system $\lambda T_s = L_s$. As a simple example, consider the people entering cricket stadium

at New Delhi to watch the cricket match. There is only one ticket line to purchase tickets. Each ticket purchase takes average of 20 seconds; the average arrival rate is 2 persons per minute. So, the question is find the average length of queue as well as average waiting time in queue.


Assuming the queuing model is a $M/G/1$ queue. With the service follows uniform distribution between 15 to 25 five seconds. With this given information, you can get the departure rate, arrival rate because it is 20 seconds per person. Therefore, the rate will be 3 persons per minute and arrival rate is 2 persons per minute. Therefore, you can get lambda that is 2 by 3. So, is irreducible positive recurrent Markov chain, therefore the steady state probabilities exists and given the steady service time distribution is uniform distribution between the interval 15 to 25, you can get the measures of steady state probabilities as well as all the average measures.

(Refer Slide Time: 27:58)



M/G/c/c System

- ▶ The $M/G/c/c$ queueing system, also known as the Erlang loss system since its development can be traced primarily to the Danish mathematician, Agner Krarup Erlang (1878-1929).
- ▶ Customers arrive according to a Poisson process with rate λ .
- ▶ Service from one of c parallel servers.
- ▶ The service times are assumed to be independent and identically distributed with a finite mean $1/\mu$.
- ▶ Because there is no waiting space in the system, arriving customers who find all c servers busy are lost.
- ▶ For example, the lost customers might represent vehicles arriving at a parking garage to find no vacant parking spaces.

 NPTEL

Now, we move into the second non Markovian queuing model not queuing model. Second non Markovian system, that is $M/G/c/c$ system, because in this model there is no queuing. The $M/G/c/c$ queueing system is also known as Erlang loss system, since its development can be traced primarily to the Danish mathematician Erlang. Customers arrive according to the Poisson process with the rate lambda service from of c parallel servers.

The service times are assumed to be independent and identically distributed with the finite mean $1/\mu$, because there is no waiting space in the system arriving customers who find all c servers busy are lost. Therefore, this system is called a loss system not a queueing system. For example, lost customers might represent vehicles arriving at a parking garage to find no vacant parking spaces.

(Refer Slide Time: 29:22)

Steady-state Distribution

- ▶ The offered load per server, often termed the traffic intensity, is given by

$$\rho = \frac{\lambda}{c\mu} = \frac{a}{c}$$
- ▶ Let $N(t)$ be the number of occupied servers at time t .
- ▶ $\{N(t), t \geq 0\}$ be a stochastic process with state space $S = \{0, 1, 2, \dots, c\}$.
- ▶ Define

$$\pi_j = \lim_{t \rightarrow \infty} P(N(t) = j), \quad j = 0, 1, 2, \dots$$
- ▶ The steady-state distribution is given by

$$\pi_j = \frac{a^j / j!}{\sum_{i=0}^c a^i / i!}, \quad j = 0, 1, 2, \dots, c$$
- ▶ This distribution is truncated Poisson distribution with parameter a .

Our interest is to find out the steady state distribution, the offered load per server often termed the traffic intensity is given by $\lambda / (c\mu)$. If you denote a is equal to λ / μ , then the ρ is nothing but a / c . Let $N(t)$ be the number of occupied servers at time t . So, $N(t)$ is a continuous time discrete state stochastic process with state space S . Define the limiting probabilities $\lim_{t \rightarrow \infty} P(N(t) = j)$. That $N(t)$ is equal to j that is π_j are the steady state distribution.

You can find easily that is nothing but $a^j / j!$ divided by $\sum_{i=0}^c a^i / i!$. So, this is a steady state distribution of $M/G/c/c$ loss system. So, this distribution is a truncated Poisson distribution with the parameter small a , where a is nothing but λ / μ . Once we know the limiting distribution, you can find the other measures the first measure.

(Refer Slide Time: 31:02)

Erlang B Formula

- ▶ Because Poisson arrivals see time averages (PASTA), the long-run proportion of arriving customers who see c servers busy is precisely, denoted by Erlang B formula $B(c, a)$, is given by
$$B(c, a) = \frac{a^c / c!}{\sum_{i=0}^c a^i / i!}$$
- ▶ When a and c are large, it can be difficult to compute due to the presence of factorials and potentially very large powers. Use the following recursive formula
$$B(k, a) = \frac{aB(k-1, a)}{k + aB(k-1, a)}, \quad k = 1, 2, \dots, c$$
where $B(0, a) = 1$.
- ▶ The Erlang B formula is a fundamental result for telephone traffic engineering problems and can be used to select the appropriate number of trunks (servers) needed to ensure a small proportion of lost calls (customers).

Is Erlang B formula because Poisson arrivals see time averages, that is pasta. You can find the long run proportion of the arriving customers, who see c servers busy. That is denoted by Erlang B formula as the function of c and a , where c is the number of servers in the system and a is λ/μ . That is nothing but the loss probability that is $a^c / c!$ divided by summation i is equal to 0 to c $a^i / i!$ when a and c are large. It can be difficult to compute due to the presence of factorials.

(Refer Slide Time: 33:44)

Erlang B Formula ...

- ▶ For a fixed, the blocking probability $B(c, a)$ monotonically decreases to zero as c increases, and for c fixed, $B(c, a)$ monotonically increases to unity as a increases.
- ▶ When the service times are i.i.d. exponential random variables with mean $1/\mu$, the system is an $M/M/c/c$ loss system. In this case,
$$B(c, a) = \frac{(\lambda/\mu)^c / c!}{\sum_{i=0}^c (\lambda/\mu)^i / i!}$$

Due to the presence of factorials and potentially very large powers a power c , when c is very large, then a power c as well as this factorials giving trouble. So, we can use recursive formula, we can use the recursive formula to compute the Erlang B formula, that is in terms of that is $B(k, a)$. In terms of $B(k-1, a)$ with the initial condition $B(0, a) = 1$.

So, that means to find the value of $B(1, a)$, you use $B(0, a)$. Then to find $B(2, a)$, use of, use the value of $B(1, a)$ and so on, so finally, you can get $B(c, a)$. In this recursive formula, we are avoiding the factorial as well as the large powers. The Erlang B formula is a fundamental result for telephone traffic engineering problems and can be used to select the appropriate number of servers need to ensure a small portion of lost customers. So, this is the way using the Erlang process, we can find out or we can select appropriate number of servers for the $M/G/c/c$ losses too.

For a fixed a where a is λ/μ the Erlang B formula or the blocking probability or lost probability monotonically decreases to 0 as c increases. Whereas, for a fixed c , where c is a number of servers, the blocking probability monotonically increases to unity as λ/μ increases. As a special case, when the service times are i.i.d random variables each having exponential random distribution with mean $1/\mu$ the system becomes $M/M/c/c$ loss system.

(Refer Slide Time: 35:00)

Erlang C Formula

- ▶ Related to the Erlang B formula is the Erlang C formula (or Erlang delay formula) for the $M/M/c$ system (or Erlang delay system), which includes an infinite-capacity queue to accommodate arriving customers who find all c servers busy.
- ▶ For this model, P_c is interpreted as the long-run proportion of customers who experience a delay before their service begins.
- ▶ The model assumes that the customers are willing to wait as long as needed to receive service.
- ▶ The Erlang C formula is given by

$$C(c, a) = \frac{\frac{a^c}{c!(1-\rho)}}{\sum_{i=0}^{c-1} \frac{a^i}{i!} + \frac{a^c}{c!(1-\rho)}}$$

- ▶ Note that above result does not hold for arbitrary service time distributions, and it requires that the traffic intensity ρ does not exceed unity.

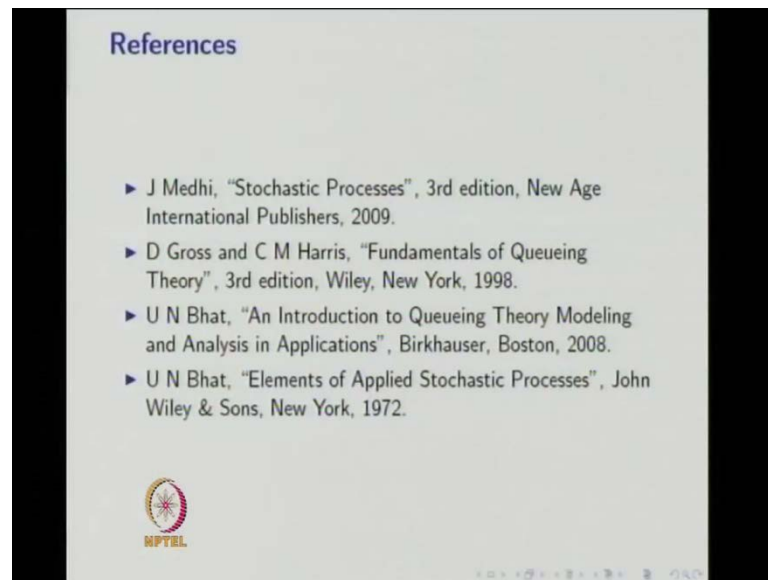
So, you can have Erlang B formula for the $M/M/c/c$ system also. With the function of service time of i.i.d random variables, each having exponential distribution with mean $1/\mu$ related to the Erlang B formula. We are going to discuss the other one called Erlang C formula. This is for the $M/M/c$ system not for $M/G/c/c$ system. This is for the $M/M/c$ sorry, $M/M/c$ system, which includes an infinite capacity queue to accommodate arriving customers who find all C servers are busy. That means this is the queuing system, queuing and delay system. We have seen servers and infinite capacity queue to accommodate arriving customers who find all C servers are busy.

So, corresponding to $M/M/c$ queuing and delay system, we have the formula called Erlang C formula or Erlang delay formula. In this model the P_c is interpreted as long run proportion of customers, who experience a delay before the service begins. The model assumes the customers are willing to wait as long as needed to receive service. So, the Erlang C formula is nothing but the blocking probability for the $M/M/c$ queuing and delay system. That is in terms of that is written in this form a^c divided by $c!$, factorial multiplied by $1 - \rho$ divided by this summation form.

Since, it is the queuing and delay system, you need additional condition to have the Erlang C formula, the additional condition is it requires the traffic intensity ρ does not exceed 1. That means as long as ρ is less than 1, the system is stable the corresponding $M/M/c$ queuing and delay system will be stable. Hence, the steady state probabilities exists and once the steady state probabilities exists, you can find the loss probability and that loss probability same as Erlang C formula.

So, to have a Erlang C formula, it requires the traffic intensity ρ has to be less than 1. Note that the above result does not hold for arbitrary service time distribution. So, this Erlang C formula is valid only for service times are exponential distributed not for arbitrary service time distribution, whereas, Erlang B formula is valid both for $M/G/c/c$ loss system and $M/M/c/c$ loss system. Erlang C formula is valid only for $M/M/c$ queuing and delay system, with the restriction ρ has to be less than 1. Whereas, Erlang B formula, the value of a is λ/μ need not be less than 1, because that is a finite capacity and loss system.

(Refer Slide Time: 38:25)



In this, we have discussed non Markovian queues in particular M G 1 queuing system, M G c c loss system, M M c c loss system, Erlang B formula for M G c c loss system as well as M M c c loss system and finally, we have discussed Erlang C formula for the M M c queuing and delay system. With these, lecture 3 lecture 4 is completed and here is the reference for lecture 4.