

Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi


Lecture – 87

(Refer Slide Time: 00:01)

Average Number in the System

$$E(N) = \text{Average number of customers in the system in steady-state}$$
$$= \sum_{n=0}^{\infty} n p^n (1-p) = p(1-p) \sum_{n=0}^{\infty} n p^{n-1}$$
$$= p(1-p) \frac{1}{(1-p)^2}$$
$$E(N) = \frac{p}{1-p}$$

Also, $\text{Var}(N) = \frac{p}{(1-p)^2}$




Other than stationary distribution, one can find out the average measures also in the system. So, suppose you make a E of N that is nothing, but the average number of customers in the system in steady state. Since we know the probability distribution substitute type p i n's here, therefore, n times p i n summation over n that is going to be the average number of customers in the system.

If you do little simplification, you will get ρ divided by 1 minus ρ where, ρ is less than 1 . So, this is the average number of customers in the system. And also, one can get variance of the number of customers in the system also. For that, you have to find out the E of N square then using that formula, you can get the variance of n also. So, here we are getting a mean and variance of number of customers in the system in steady state.

(Refer Slide Time: 01:03)

Average Number in the Queue

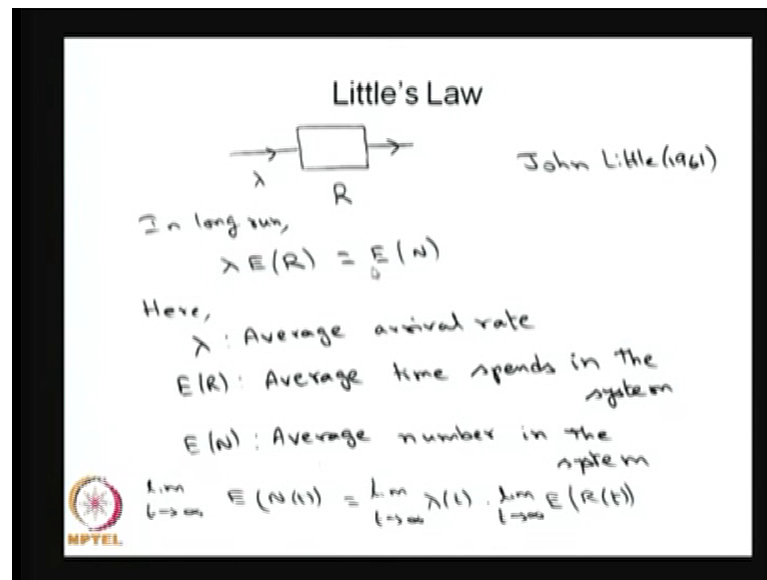
$$E(Q) = \text{Average number of customers in the queue in steady-state}$$
$$= \sum_{n=1}^{\infty} n \pi_{n+1} = \sum_{n=1}^{\infty} n \rho^{n+1} (1-\rho)$$
$$= \rho^2 (1-\rho) \cdot \frac{1}{(1-\rho)^2}$$
$$= \frac{\rho^2}{(1-\rho)}$$


Also one can find average number in the queue. So, the letter q is a random variable and here, we are finding the expectation of q , that is, average number of customers in the queue; that means, a before getting the service or how many customers in the system. We have only one server in the system.

And whenever the service is going on and all other arriving customers will be queued; that means, when n plus 1 customers in the system; n people are in the queue. Therefore, summation n times π_{n+1} .

Do the simplification. You will get average number of customers in the queue. Also substitute the π_{n+1} from the one I have discussed in the stationary distribution.

(Refer Slide Time: 01:59)



Here, I am going to relate the average measures using the Little's law. This is proven by John Little, 1961. This is valid for any system in which arrival comes into the pattern with the arrival rate λ and R is a time spent in the system and leave the system after the service or whatever, the things are over.

Then, in a longer run, one can say the arrival rate multiplied by the average time spent in the system that is same as average number in the system. So, this relation is a valid for whatever be the underlying distribution, whatever be the underlying distribution of the service, underlying distribution of the arrival.

What it says if you have a system in which the arrival rate is a mean? Arrival rate is λ and the mean time spent in the system is a expectation of R . Then, that product will give average number in the system.

Since indirectly it says, whenever the system has a long run in a stable system, the expectation of a average number of customers during the interval 0 to t as a t tends to infinity that is going to be have a limit expectation of n . And the arrival rate λ of t , that also has the mean arrival rate as a t tends to infinity, that is also going to be a sum having element constant λ .

And similarly, the average spent by the customers in the system at any time t and if you make a t tends to infinity, that expectation quantity also has a limit. Therefore, you will

have a lambda times a expectation of R is same as expectation of n. Now, using this Little's law, I am going to find out the measures for the mm 1 queue model.


(Refer Slide Time: 04:15)

$R(W)$: time spend (waiting time) by any customer
Using Little's formula,

$$E(R) = \frac{E(N)}{\lambda}$$
$$= \frac{\rho}{\lambda(1-\rho)}$$
$$E(R) = \frac{1}{\mu - \lambda}$$

Now,

$$E(W) = E(R) - \frac{1}{\mu}$$
$$= \frac{\rho}{\mu - \lambda}$$



So, suppose R denotes the time spent in the system by the customer and W denotes the waiting time by any customer in the system, then I can use the Little's formula; the Little's law in the previous one.

So, if I know the mean arrival rate and if I know mean number of customers in the system in a longer run, using these, I can find out the average time spent in the system. If I know to use the to use using Little's law, if I know the average number of customers in the system longer run and if I know the arrival rate, then I can find out the average time spent in the system in a longer run.

Similarly, I can once I know the average time spent in the system, if I subtract the average time of my own service, then that is going to be the average a time waiting in the queue. So, this is a average a time waiting in the queue that is same as an average a time spent in the system minus my own average service time.

The mm 1 queue model, the service time is a exponentially distributed with the parameter mu. Therefore, the average is 1 by mu. So, the difference will give the average time. Waiting in the queue by any customer not only the average measures for the mm 1 queue, one can find out the actual distribution for or as well as W also.

Because, this is a very simplest Markovian queuing model whereas, for all other models, it is little complicated. But still, one can get it. So, this is the easy model in which you one can find out the distribution of the time spent in the system as well as the time or as well as the waiting time by a customer in the queue.

(Refer Slide Time: 06:17)

Distribution of Waiting Time

$$W = \begin{cases} 0, & n=0 \\ S_1 + S_2 + \dots + S_n, & n=1, 2, 3, \dots \end{cases}$$

$$P[W \leq t] = \begin{cases} 0, & t < 0 \\ 1-p, & t = 0 \\ ?, & 0 < t < \infty \end{cases}$$

$W/N=n \sim \text{Gamma}(n, \mu)$

For $t > 0$

$$P[W \leq t] = \sum_{n=1}^{\infty} \int_0^t \frac{\mu^n x^{n-1} e^{-\mu x}}{(n-1)!} dx (1-p)^n$$

First, let us go for finding out the distribution of waiting time. Waiting time means, if no one in the system, when you arrive, then your waiting time is 0, you are immediately going to get the service. So, the service time is your time spent in the system, ok.

Usually, the time spent in the system is the time of your service plus the time of the waiting time. So, here I am finding the only the distribution of waiting time first. So, whenever the system is a 0, your waiting time is 0 whenever no customer in the system, the waiting time is 0.

Whenever more than or equal to one customer in the system, then the waiting time is same as the remaining a service time for the customer who is under service plus the customers in the queue before you join in the queue.

So, those people service time addition plus the residual or the remaining a service time of the customer who the first customer who is under service. So, this total time is the waiting time. Whenever the system is non empty, whenever the system is empty then the waiting time is 0.

Therefore, the W is a random variable. Either it takes a value 0 or it takes a value greater than 0 based on the times time of a service of a previous n people in the ahead of you. Therefore, W is a mixed random variable, which has the probability mass function at 0 as well as a density function between the interval 0 to infinity.

So, let me try for finding out the cdf of this random variable. So, the cdf is going to be 0 till 0 at the 0. It has a cdf 1 minus rho because and the waiting time is 0 that is, the equivalent of a no one in the system.

So, in the long run, no one in the system that is a π naught and the π naught probability is a π not probability is a 1 minus rho.

(Refer Slide Time: 08:49)

Take $\rho = \frac{\lambda}{\mu}$

Then,

$$\pi_0 = 1 - \rho$$

$$\pi_n = (1 - \rho) \rho^n \quad ; \quad \rho < 1 \text{ (stable system)}$$

$n = 1, 2, \dots$

ρ : offered load (traffic intensity)
 ρ : server utilization

NPTEL

The system is the empty in a longer run; that is, 1 minus rho. Therefore, the cdf at a 0 that is same as a 1 minus rho of that is π naught between the interval 0 to infinity one. We have to find out the distribution of W whenever n customers before you before you join in the system.

That conditional distribution the distribution of a W given the number of customers in the system is n . That distribution is nothing, but the service time of a the n customers. The first customer remaining service time the service time of the first customer is exponential distribution the residual or remaining a service time of the first customer, that is also exponential distribution because of memory less property.

So, this is exponential distribution. This is second customer service times that is exponential distribution. And similarly, for the n th customer also, the service time is exponentially distributed and the way we made the assumption all the service times are independent and each one is exponentially distributed with the parameter μ . Therefore, this is a sum of n independent exponentially distributed random variables.

Therefore, the sum of n exponentials that is going to be a gamma distribution with the parameters n and μ , there are many ways of finding out the distribution. But, here, I am just explaining through the distribution concept. This is some of the n independent exponential distribution. Therefore, you can conclude it is gamma distribution with the parameter n and μ .

Once you know the conditional distribution, our interest is to find out the unconditional one; that means, a for t is greater than 0 and cdf at the point t . That is nothing, but what is the conditional density, probability density. And what is the probability of n ? Customers in the system that multiplication with the possible n will give the cdf between the interval 0 to t . So, I have a density function of a gamma distribution probability density function with the parameters n and μ .

And, this is a probability density function multiplied and integration between 0 to t . That will give the cdf and the condition unconditional multiplied by probability of n customers in the system. That will be the summation that will give the unconditional. Therefore, the cdf is going to be summation n is equal to 1 to infinity. Integration 0 to t of, the probability density function of a gamma distribution multiplied by n customer in the system if you do the simplification, you will get a $1 - e^{-\mu t}$.

(Refer Slide Time: 12:04)


Distribution of Waiting Time

For $t \geq 0$
 $P[W \leq t] = 1 - e^{-\mu(1-\rho)t}$

Hence,
$$P[W \leq t] = \begin{cases} 0, & t < 0 \\ 1 - \rho, & t = 0 \\ 1 - e^{-\mu(1-\rho)t}, & 0 < t < \infty \end{cases}$$

Hence,
 $P[W = 0] = 1 - \rho$

and $f_w(t) = \rho(\mu - \lambda) e^{-(\mu - \lambda)t}, t > 0$



Therefore, you can substitute here oh here I made a mistake. So, here it is multiplied by rho. So, 1 minus rho times 1 minus rho times.

So, 1 minus rho times e power minus rho e power minus mu times 1 minus rho that is going to be the. So, once you are getting the cdf, you can conclude this is a mixed random variable with the probability mass at a 0 is 1 minus rho and the density function.

Between the interval 0 to infinity, that is a rho times mu minus lambda times e power minus mu minus lambda times t; that is, the probability density function for a distribution of waiting time.

(Refer Slide Time: 13:26)

Distribution of Response Time


$$R = S + S_1 + S_2 + \dots + S_n$$

$$P[R \leq t] = \begin{cases} 0 & , t \leq 0 \\ ? & , 0 < t < \infty \end{cases}$$

$$R/n \sim \text{gamma}(n+1, \mu)$$

For $t > 0$

$$P[R \leq t] = \sum_{n=0}^{\infty} \int_0^t \frac{\mu^n x^{n-1} e^{-\mu x}}{n!} dx (1-\rho)^n$$

$$= 1 - e^{-\mu(1-\rho)t}$$


Similarly, one can get the distribution of a response time also or the total time spent in the system. The total time spend in the system that is nothing, but that is a random variable and the residual service time of the first customer who is in the system plus all the remaining n customers in the system in the queue plus your own service time.

Therefore, here, this is not a mixed random variable. This is a continuous random variable because your service time is a continuous random variable which is exponentially distributed.

Therefore, the R is going to be sum of your own service plus the remaining service of the first person in the system if and so on, till the n th customer who is in the queue.


Therefore, this is the cdf of the random variable R . Here also, one can argue when n customer in the system before him who enter into the system, that is a sum of a exponential independent random variable and so on. Therefore, this is going to be a gamma distribution with the parameters n plus 1 μ .

And for t greater than 0, find out the cdf using the first conditional then, unconditional multiplied by 1 minus ρ times ρ power n summation over n is equal to 0 to infinity. Because, there is a possibility no one in the system or 1 customer, 2 customer and so on. Therefore, the running index 0 to infinity. Do the simplification, we will get to 1 minus e power minus ρ times this. Therefore, you can substitute here.

(Refer Slide Time: 15:13).

Distribution of Response Time

Hence,

$$P[R \leq t] = \begin{cases} 0 & , t \leq 0 \\ 1 - e^{-\mu(1-\rho)t} & , 0 < t < \infty \end{cases}$$
$$R \sim \text{EXP}(\mu(1-\rho))$$
$$\therefore E(R) = \frac{1}{\mu(1-\rho)} = \frac{1}{\mu - \lambda}$$


And, if you see, the cdf is the same as the cdf of exponential distribution with the parameter, that is, μ times 1 minus ρ . Therefore, you can conclude the total time spend in the system is exponentially distributed with the parameters μ times 1 minus ρ . If you find out the average time, that is going to be 1 divided by the parameter. That is this.

The same thing you got it in the average response time from the Little's formula using a. Once you know the value of λ and the expected number in the system, using Little's law you got expectation of a time spent in the system, that is same result. So, here we are getting first finding the distribution of time spent in the system or response time, then we are finding the average time.