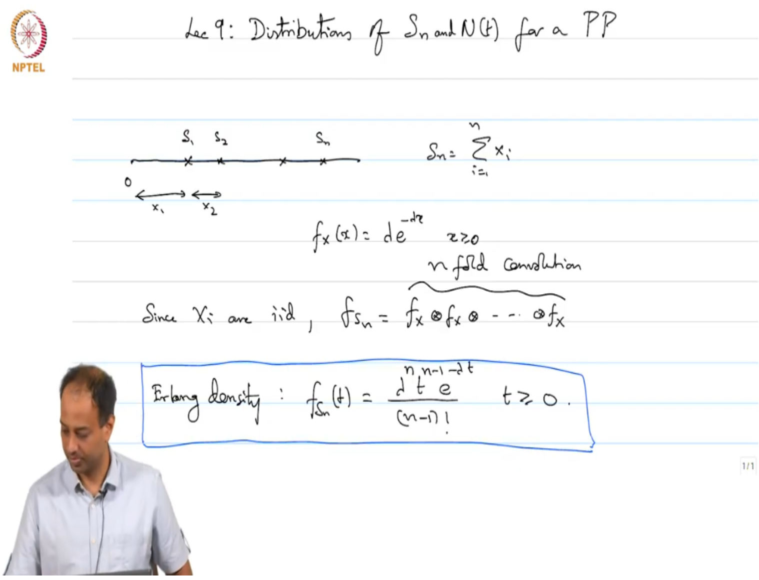
Stochastic Modeling and the Theory of Queues
Prof. Krishna Jagannathan
Department of Electrical Engineering
Indian Institute of Technology - Madras

Module - 2
Lecture - 9
Distribution of Arrival Epoch Sn and N(t) for a Poisson Process

(Refer Slide Time: 00:16)



Good morning. Today, we will discuss the distribution of the arrival epochs S_n and the number of arrivals until time t, $N(t)$ of a Poisson process. So, recall that we defined a Poisson process as a counting process where these interarrival times $X_1, X_2, . . .$, et cetera are IID exponentials, and these are of course the arrival epochs $S_1, S_2 ... S_n$. We know that $S_n = \sum_{i=1}^{n} X_i$.

So, the distributions of X_i's are independent identically distributed exponentials with parameter λ. So, what is the distribution of S_n? Multiply in Laplace domain. Of course, see, the basic thing is that these are independent random variables, so, you can, the density of the sum is given by a convolution of the density of each of these. And of course, the density of each of these is the same, which is an exponential.

So, if you take; so, $f_X(.)$; for each of these X_i's, $f_X(.) = \lambda e^{-\lambda x}$, $x \geq 0$. And, since these X_i's are independent, we can write; since X_i are IID, we can write $f_{S_n} = f_X \otimes f_X \otimes ... f_X$. It is an n fold convolution. This is an n fold. So, you can sit and convolve the exponential distribution n times. Or if you know something about Laplace transforms, you can take Laplace transform, multiply it.

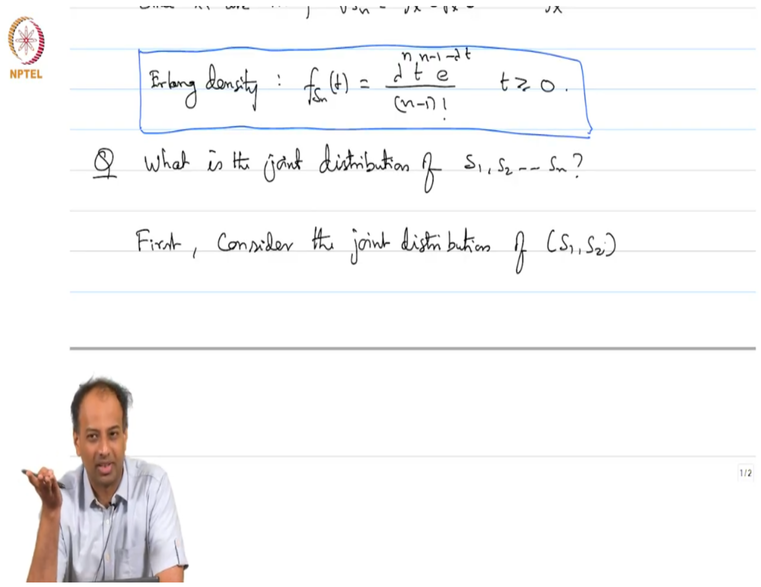
So, you take the n^{th} power of the exponentials Laplace transform and then invert back. So, what you get when you do this is something known as the Erlang density. So, you get $f_{S_n}(.)$ when you do all this n fold convolution, you get,

$$f_{S_n}(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}, \ t \geq 0$$

So, you can easily calculate the density of the n^{th} arrival epoch. So, you put $n = 1$, you get back the; so, for $n = 1$, what happens? S_1, right? $S_1 = X_1$. So, you should get back what? The usual exponential which you do. You can just put $n = 1$ and check that it works out. Now, this is just the marginal distribution of S_n. So, in order to specify the process; see, what specifies a process is the joint distribution of these X_i's or S_i's.

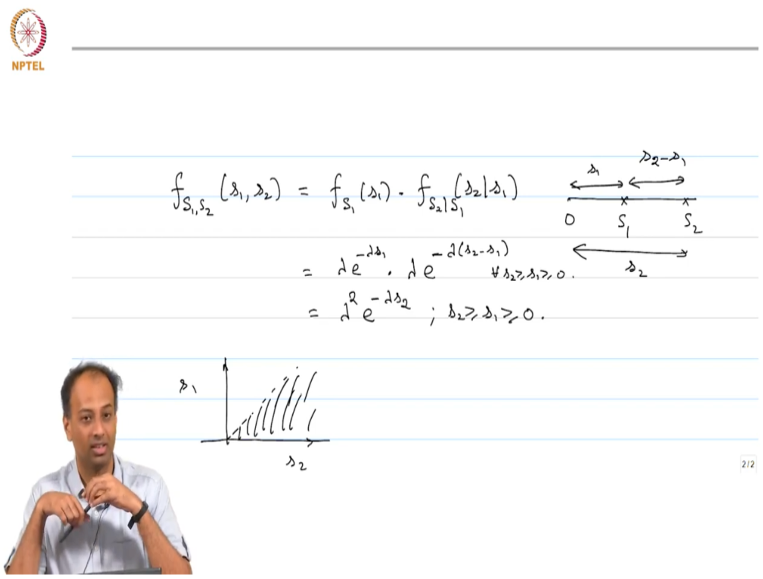
The joint distribution of X_is, of course, we know; they are all independent exponentials. But the joint distribution of S_i's, we have to calculate. So, what we have calculated through this convolution formula is only the marginal. So, if I ask you what is the joint distribution of $S_1, S_2, \ldots S_n$ for any n, then what happens? It is a little more non-trivial. See what I mean? So, question:

(Refer Slide Time: 04:40)



What is the joint distribution of S_1, S_2,... S_n? This is the question. Now, let us do this for $n = 2$. So, S_1 and S_2 are the first two arrival epochs. Now, S_1 and S_2 are of course dependent. X_1 and X_2 are independent, but S_1 and S_2 are dependent. So, they have some joint distribution, which is not just the product or anything. Now, of course, S_2 has to be bigger than or equal to S_1; so, clearly dependent.

(Refer Slide Time: 05:51)



So, what is $f_{S_1, S_2}(s_{1,} s_2)$? You can write this as, using the definition of conditional densities, you can write this as $f_{S_1}(s_1) \cdot f_{S_2|S_1}(s_2|s_1)$. This you know, right? you know the definition of conditional density; that is what I am using. The reason I am doing this is because, once I condition on S_1, the further time to S_2 is well-known. What is it? It is exponential. So, I know that. So, I want to exploit that property.

So, you are looking at something like this. So, this is 0; that is S_1; that is S_2. So, given that this guy is realised as s_1, what is the distribution of that? So, given $S_1 = s_1$, this width is just an exponential. So, if you want this to be s_2, then this width should be just $s_2 - s_1$; that is what we are going to use. It is very simple. So, this is, of course, $S_1 = X_1$. So, this is what? The first term is $\lambda e^{-\lambda s_1}$.

So, you are conditioning now on S_1, and you want $S_2 = s_2$, which means that the realisation of X_2 should be $s_2 - s_1$. So, this is just the density of the X_2, evaluated at $s_2 - s_1$. And this is true for all $s_2 \geq s_1 \geq 0$. I am just using the fact that, condition on S_1; $s_2 - s_1$ is an exponential; $s_2 - s_1$ *is* X_2. That is all that I am using.

So, what does that work out to be? That works out to be $\lambda^2 e^{-\lambda s_2}$, because this cancels with $e^{-\lambda s_1}$. And this is true for $s_2 \geq s_1 \geq 0$. So, the joint density is,

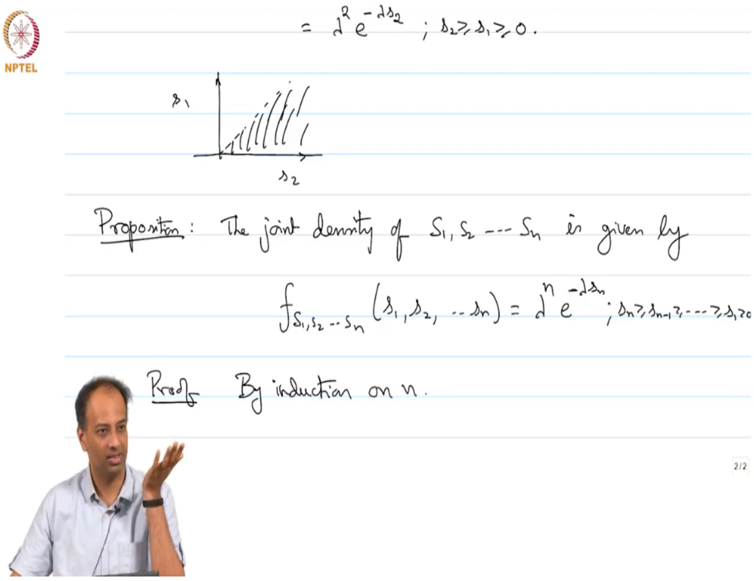
$$f_{S_1, S_2}(s_{1,} s_2) = \lambda^2 e^{-\lambda s_2}$$

Now, where is s_1? See, this $f_{S_1, S_2}(* \ ,*)$should be a function of $s_1$and s_2, right? But it is only a function of s_2, which means it is? It is independent.

It is constant. It is constant in s_1, except that, it means, the s_1 shows up in the constraint. So, $s_1 \leq s_2$; it clearly has to be, right? So, if you look at the two-dimensional plane, let us say this is s_2, that is s_1. The density is non-zero, only in the range $s_2 \geq s_1 \geq 0$. So, it is non-zero only here. Is that clear to everyone?

And as a function of s_2, there is a dependency $e^{-\lambda s_2}$, but as a function of s_1 which is, I have drawn in the vertical axis, it is constant. So, if you pick any s_2 and you change s_1, the density remains constant, the density is coming out of the plane of the board, if you like. So, it is decaying in s_2 but constant in s_1, and it is defined in this region which is below the 45-degree line. Now, you can do the same trick for $s_1, s_2, \ldots s_n$. So, we can state this now.

(Refer Slide Time: 11:21)



Proposition: Once you understand the case $n = 2$, the case for general n is easy. The joint density is ,

$$f_{S_1, S_2, \ldots, S_n}(s_1, s_2, \ldots s_n) = \lambda^n e^{-\lambda s_n}, \; s_n \geq s_{n-1} \geq \ldots \geq s_1$$

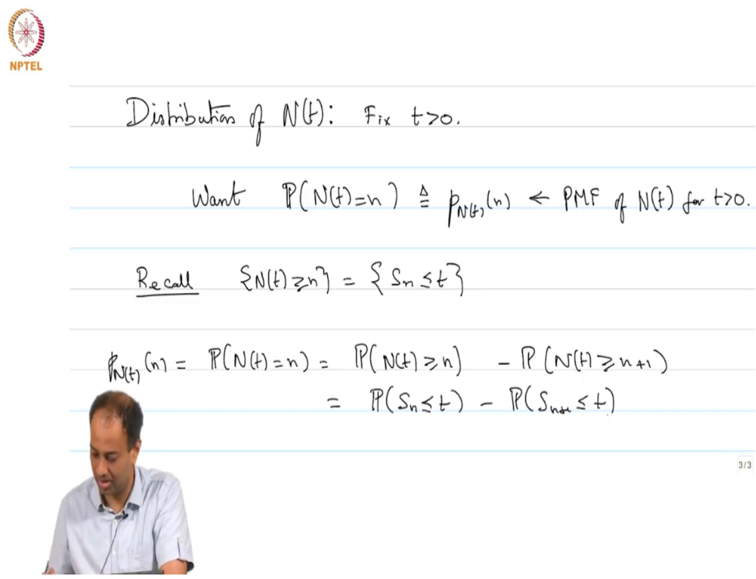
So, again, this joint density $f_{S_1, S_2, \ldots, S_n}(s_1, s_2, \ldots s_n)$ is explicitly a function only of s_n, but the other random variables $s_1, s_2, \ldots s_{n-1}$ occur in the constraints.

So, in an n-dimensional space; of course, all of these are non-negative random variables; so, you can; anyway looking at only the non-negative orthant. And in the non-negative orthant, it exists only in a part of the orthant where the n^{th} coordinate is bigger than $(n-1)^{th}$ coordinate is bigger than the $(n-2)^{th}$coordinate and so on; similar to the picture I drew here, this picture right here, except in the n-dimensions.

So, this is the joint density. How do you prove this? Yes. Proof is by induction. In particular, for the case $n = 2$, you already proved. Make $n = 2$ the base case. You already proved, from first principles. Then you make an induction hypothesis saying that the k^{th} joint density is this, $f_{S_1, S_2, .., S_k}(.)$. Then you look at that joint density of $f_{S_1, S_2, .., S_{k+1}}(.)$. Then you write that in terms of the joint density of $f_{S_1, S_2, .., S_k}(.) f_{S_{k+1} | S_1, S_2, .., S_k}(.)$.

Of course, you have an induction hypothesis for the first term, which is the joint density of the first k. And then, of course, the conditional density, given the first k, is simply another exponential. Then, the same trick works, induction will do the job. I think you can complete this easily.

(Refer Slide Time: 14:24)



The Distribution of $N(t)$. So, you are fixing some t; $N(t)$, of course, is a random variable. You fix $t = 10$, $t = 100$, whatever you want and you are looking at the total number of arrivals until the time t; is of course a random variable; and you are looking at its distribution. It is a non-negative integer value random variable. So, you are talking about the; so, for non-negative random variable, integer value random variable, you have to talk about the PMF.

So, you want to get the probability mass function of $N(t)$. So, you want this; want $P(N(t) = n)$. So, we can denote this by $p_{N(t)}(n)$. This is a PMF, probability mass function of $N(t)$, for a fixed $t > 0$. Now, recall that there is this equivalence. Recall that the event $\{N(t) \geq n\} = \{S_n \leq t\}$. We know the distribution of S_n is Erlang.

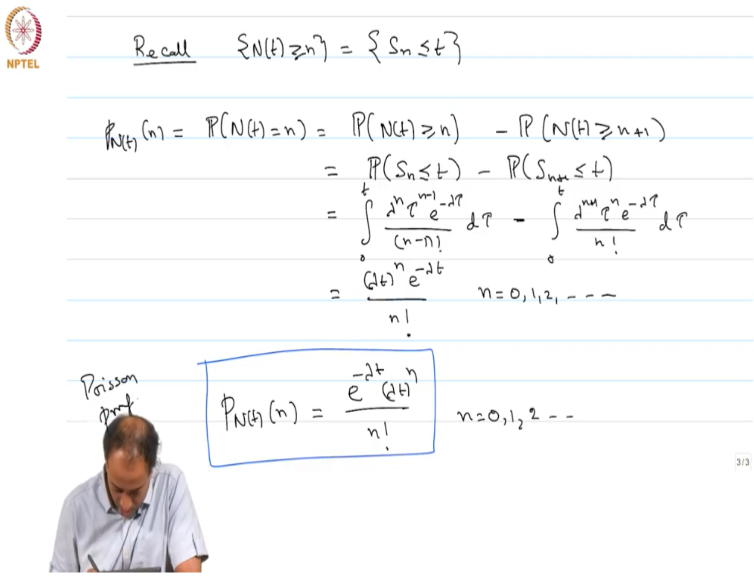
So, $P(S_n \leq t\}$, we can easily write out; it is simply the integral of the Erlang density. This is nothing but the Erlang CDF, if you take probability of this. So, you can look at the $P(N(t) \geq n\}$ is easy to get. That is really all there is to it. Since you know the distribution of S_n, you can calculate the distribution of $N(t)$ using this equivalence.

It is a mostly mechanical exercise, but you can do the following for example. So, you can write; perhaps you can; this is, well, one way to do it. $P(N(t) = n)$ is what you want. $P(N(t) = n) = p_{N(t)}(n)$; is this correct? Maybe not. No, this is not correct. So, I want to write; $P(N(t) \geq n)$. If I write $n - 1$here, I will be okay, right? So, maybe I should write it like this,

$$P(N(t) = n) = P(N(t) \geq n) - P(N(t) \geq n + 1)$$

I think you will agree if I write; Somewhere, anything wrong here? No, I changed; I mean, I was off by 1, now I think I am okay. This is correct. So, but now, this is of course equal to $P(S_n \leq t) - P(S_{n+1} \leq t)$; Why? This is from the equivalence.

(Refer Slide Time: 18:58)



So, this, you can write as,

$$P(N(t) = n) = \int_0^t \frac{\lambda^n \tau^{n-1} e^{-\lambda\tau}}{(n-1)!} d\tau - \int_0^t \frac{\lambda^{n+1} \tau^n e^{-\lambda\tau}}{n!} d\tau$$

What have I done? So, I want to look at $P(S_n \leq t)$, which is the CDF of S_n, which is the running integral from 0 to t of the Erlang density. See, I am writing τ here, because I want the variable of integration to be different from what the limit is; that is why I put τ, if you are wondering why. And similarly, I have done the same thing with n replaced with $n - 1$. So, now you can fight it out. There is nothing more to it.

You will do this integral integration by patch, whatever; you people do this faster than I can do, right? So, finally, you fight it out and you will get a nice answer. You will get, this is equal to; this answer I know,

$$P(N(t) = n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}, n = 0, 1, 2... \text{ (because } 0! = 1\text{)}$$

Actually, see, we have to, you will directly get this for $n = 1$ onwards; $n = 0$, you have to do it separately, because this formula would not hold, because you will get $(-1)!$ and all that; does not make sense.

So, for $n = 0$, $P(N(t) = 0) = P(X_1 > t)$. That we already know. $\{N(t) = 0\} = \{X_1 > t\}$; $P(X_1 > t) = e^{-\lambda t}$. So, nevertheless you get that for $N(t) = 0$. For $N(t) = 1$ onwards, this calculation is valid. So, this calculation here is valid for $N(t) = 1$, 2... onwards. For $N(t) = 0$, you have to do it separately, like I just spoke out. But nevertheless, the formula here will be valid. If you take $n = 0$ factorial, you will get it. So, this is worth putting in a box.

$$p_{N(t)}(n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}, n = 0, 1, 2...$$

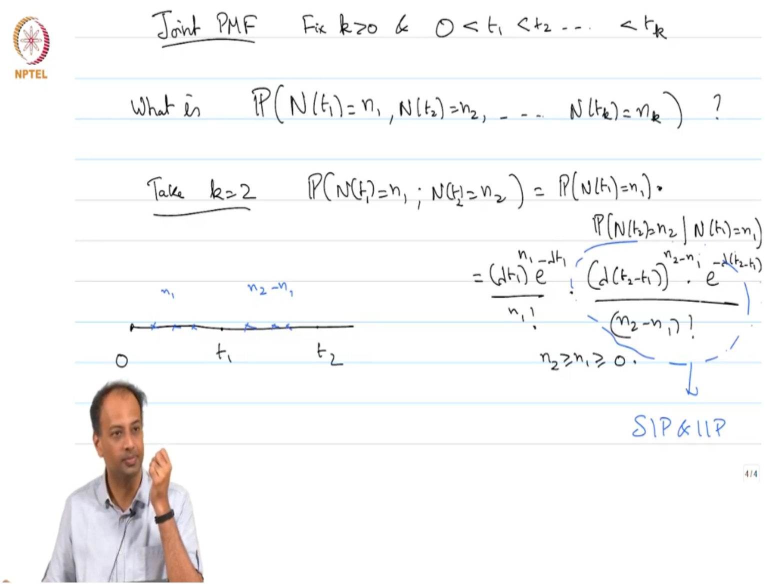
This is called the Poisson PMF. This is true for; so, n is equal to 0, 1, 2, et cetera. This is called the Poisson PMF. It is Poisson PMF with parameter λt. So, for any t, $N(t)$ is Poisson distributed with parameter λt.

So, we know that; what we did is, we know S_n is Erlang distributed. So, using the equivalence between S_n and $N(t)$ which we already derived, we just got it, we just did some algebraic manipulations and got it. So, nothing very greatly involved here; some big integrals involved. Of course, the PMF of $N(t)$ alone is not satisfactory. What do you actually want?

See, this $N(t)$ is a sequence of random variables indexed by t. So, you want to characterise all finite order joint distributions of $N(t_1)$, $N(t_2)$ et cetera, just like $S_1 . . . S_n$. Here, for any given t_1, t_2, ... t_k, you want the joint PMF of $N(t_1)$,$N(t_2)$,...$N(t_k)$. So, that is the next thing we will do. **"Professor - student conversation starts"** Yes? Sorry. I; so, in this? No. I mean; see, this, whatever I have written down here, this expression out here is valid for $n \geq 1$.

So, all of this is true for; this guy is true for $n \geq 1$. So, I should really write; so, I can only write $n = 1, 2,..$, because this expression is valid only for $n \geq 1$. I am sneaking the 0 in by saying that you can make a separate argument for $N(t) = 0$. That is all. **"Professor - student conversation ends"**

(Refer Slide Time: 25:24)



Now, joint distribution; joint PMF. So, fix some $k > 0$ and $0 < t_1 < t_2 <\ldots< t_k$. So, you want the joint PMF of $N(t_1)$,$N(t_2)$, et cetera. So, what is $P(\,N(t_1) = n_1,\, N(t_2) = n_2, \ldots ,\, N(t_k) = n_k)$? Again, it is easy to do, I mean, you will take this $k = 2$. So, what is the joint PMF $P(\,N(t_1) = n_1,\, N(t_2) = n_2)$? So, you first take $k = 2$, which is the simplest case. We want $P(\,N(t_1) = n_1,\, N(t_2) = n_2)$.

This can be written as $P(N(t_1) = n_1)\,.\,P(\,N(t_2) = n_2 \,|\, N(t_1) = n_1)$; this is by this conditioning. Got it? So, it just comes down to; see, $P(N(t_1) = n_1)$ I already know. $P(N_1(t) = n_1) = \frac{(\lambda t_1)^{n_1} e^{-\lambda t_1}}{n_1!}$. That is just the first term, which I already know.

Then, you are looking at what is $P(\,N(t_2) = n_2 \,|\, N(t_1) = n_1)$? Now, you have to use some property of the Poisson process. So, if there are; so, there have been n_1 arrivals till t_1, you want another $n_2 - n_1$ arrivals to come, in an interval of width $t_2 - t_1$. See, by the stationary increment property, the number of arrivals in any interval is only a function of width, which is $t_2 - t_1$ in this case.

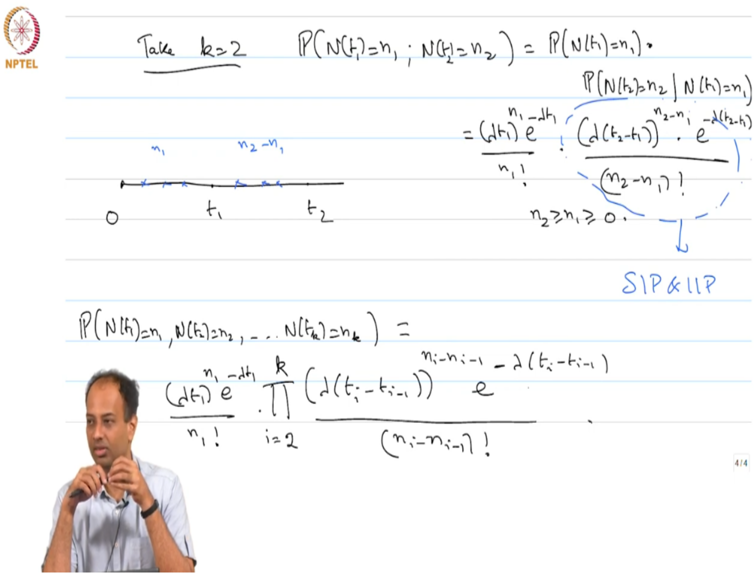
Also, given that there are n_1 arrivals till time t_1, the number of arrivals in $(t_1, t_2]$ is independent of the number of arrivals in $(0, t_1]$. Why? IIP, independent increment property. So, you have to use both SIP and IIP to calculate this guy. So, let me write this,

$$P(N(t_2) = n_2 \,|\, N(t_1) = n_1) = \frac{\lambda(t_2 - t_1)^{(n_2 - n_1)} e^{-\lambda(t_2 - t_1)}}{(n_2 - n_1)!}$$

You want $n_2 - n_1$ arrivals in that interval $t_2 - t_1$ times $e^{-\lambda(t_2 - t_1)}$ over $(n_2 - n_1)!$. This is of course true for $n_2 \geq n_1 \geq 0$. And this, for this particular term, to get this term, I have used SIP and IIP. So, to just give you a picture; this is 0, this is t_1, that is t_2. So, you had some n_1 arrivals here. So, given that you had n_1 arrivals in $(0, t_1]$, you want to have another further $n_2 - n_1$ arrivals here.

Of course, by the stationary increment property, the number of arrivals in $(t_1, t_2]$has the same distribution as the number of arrivals in $(0, t_2 - t_1]$. And what is more; given that there are n_1 arrivals in $(0, t_1]$, the number of arrivals in $(t_1, t_2]$is independent of this number of arrivals that have already taken place. So, I am using both these things together to write this. So, similarly for k I can write down; the same trick works.

(Refer Slide Time: 32:00)



I hope I do not make any mistakes. Let me write this down.

$$P(N(t_1) = n_1, N(t_2) = n_2, \ldots, N(t_k) = n_k)$$

$$= \frac{(\lambda t_1)^{n_1} e^{-\lambda t_1}}{n_1!} \prod_{i=2}^{k} \frac{(\lambda(t_i - t_{i-1}))^{n_i - n_{i-1}} e^{-\lambda(t_i - t_{i-1})}}{(n_i - n_{i-1})!}$$

Is that correct? Because you can do this repeated conditioning. In particular, you can use induction; you can use the previous result as a base case; make this induction hypothesis for k and prove it for $k + 1$. This expression is basically, you do this multiple times; whatever I did before for 2, you do this multiple times.

And I hope there are no off by 1 errors here; looks correct. So, that is the joint distribution of these. So, given any k and t_1, t_2, ... t_k, you can calculate the joint distribution of $N(t_1)$ through $N(t_k)$ using IIP and SIP.