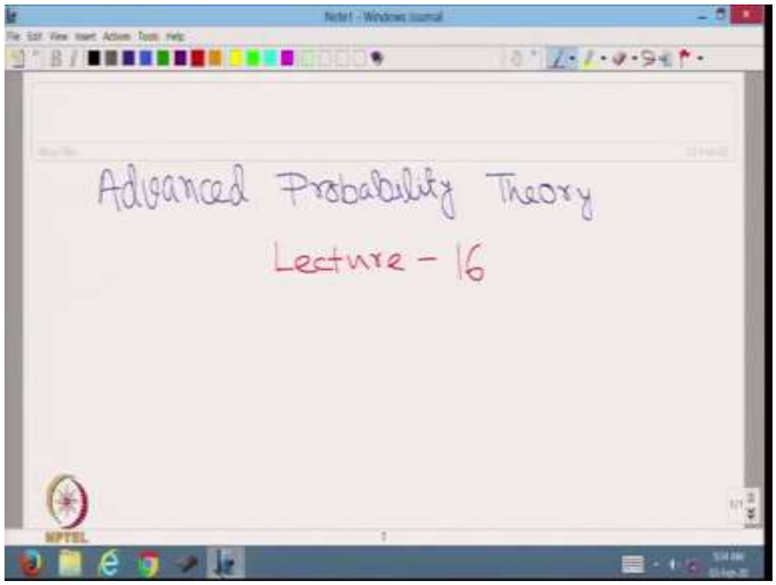
Advanced Probability Theory
Professor Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi
Lecture 16

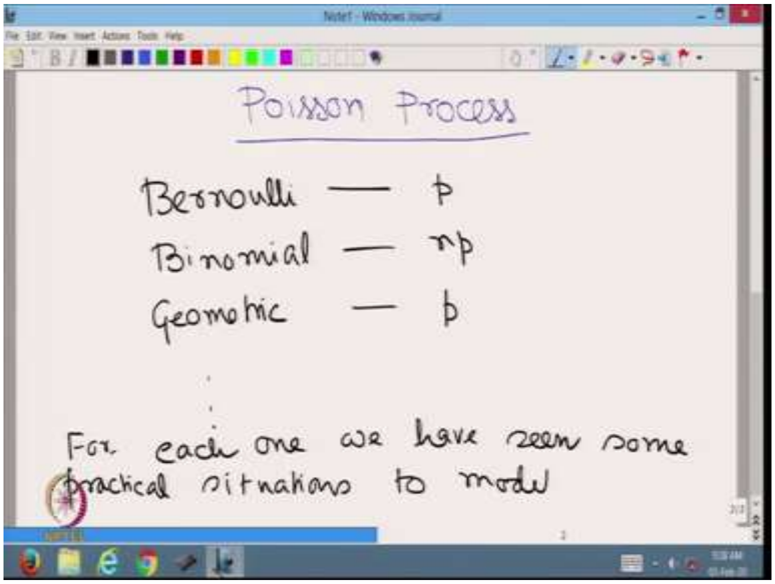
(Refer Slide Time: 00:27)



Welcome students to the MOOC lecture series on Advanced Probability Theory this is lecture number 16. If you notice that over the last 6 weeks, we have studied the basic theory of probability, then random variables, discrete continuous types, and also different moments of different random variables in detail. And each week, we have focused on one particular topic in this week. However, we shall not focus on a particular topic rather, what we shall do, we shall touch upon some basic interesting results that can come from different distributions.

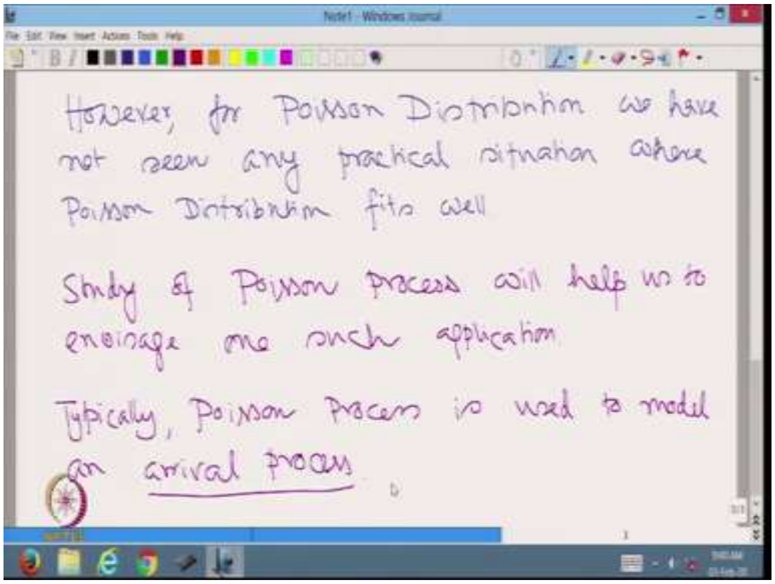
Also we shall study by variate normal distribution, which is a very important probability distribution, when you deal with multivariate data, we shall talk about that later.

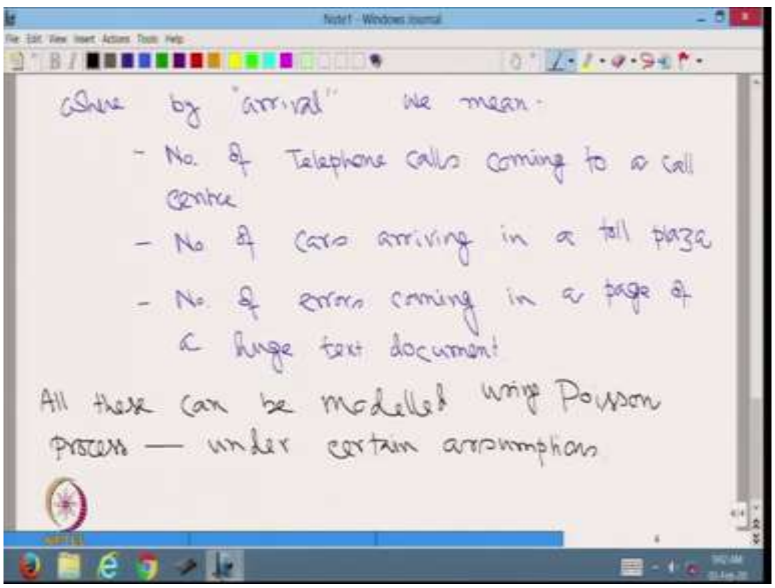
(Refer Slide Time: 01:31)



For todays lecture, the focus is on what is called Poisson process. We have studied many discrete distributions, say binomial, say Bernoulli with parameter p binomial with parameter np geometric with P etc. For each one of them we have seen some practical situations to model them.

(Refer Slide Time: 02:42)

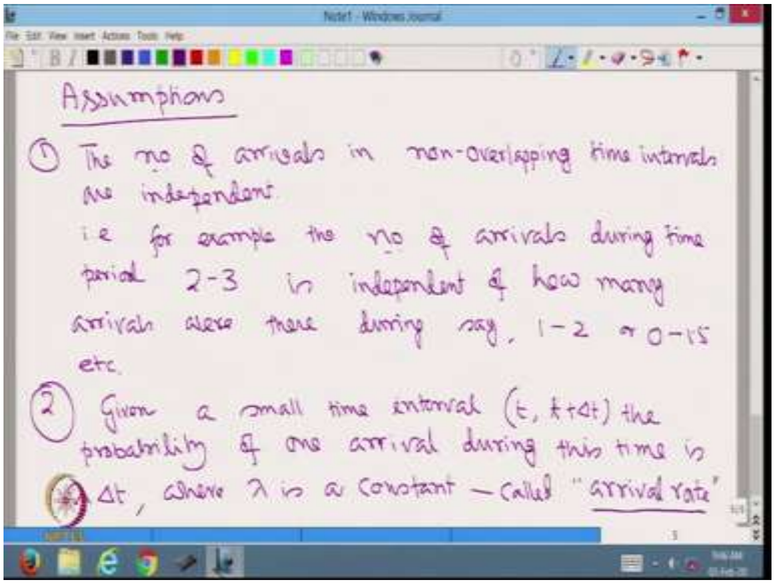




However, for Poisson distribution we have not yet seen any practical situation where Poisson distribution fits well. Study of Poisson process will help us to envisage one such application. Typically Poisson process is used to model an arrival process, this is very important that a very specific type of practical problems can be modeled with Poisson distribution or Poisson process where by arrival we mean many situations say number of telephone calls coming to a call center.

It may mean number of cars arriving in a toll plaza it may mean number of errors coming in a page of a huge text document all these can be modelled using Poisson process under certain assumptions, okay. So, let us first see what are the basic assumptions.

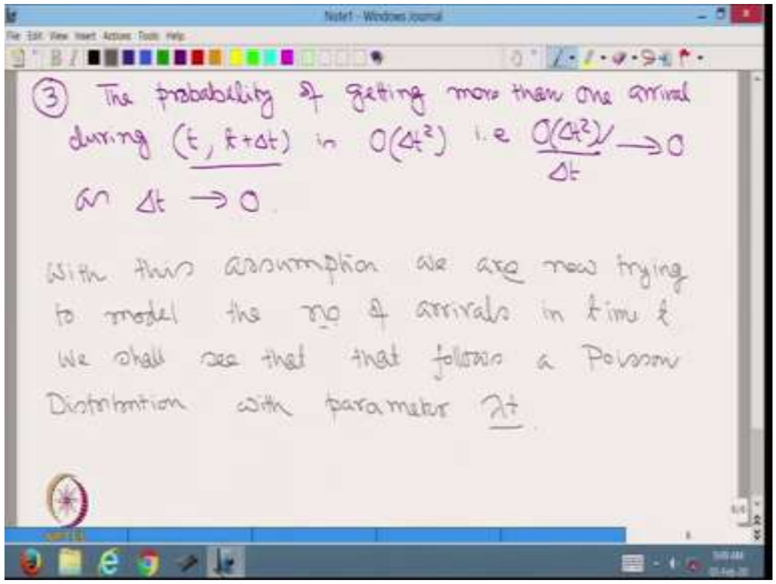
(Refer Slide Time: 06:28)



One, the number of arrivals in non-overlapping time intervals are independent that is, say for example, the number of arrivals during time period 2 to 3 is independent of how many arrivals were there during say 1 to 2 or 0 to 1.5 etc.

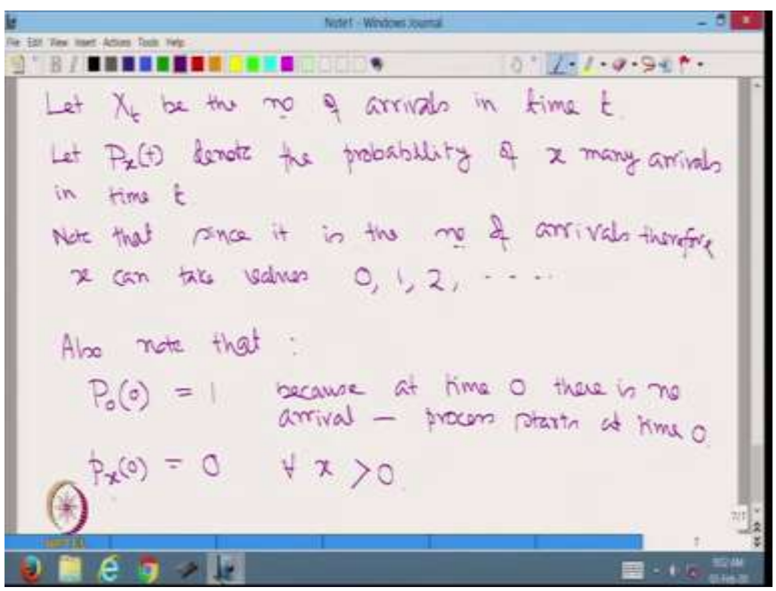
Second assumption is that given a small time interval t to t plus delta t the probability of one arrival during this time is lambda times delta t, where lambda is a constant and we call it arrival rate.

(Refer Slide Time: 09:20)



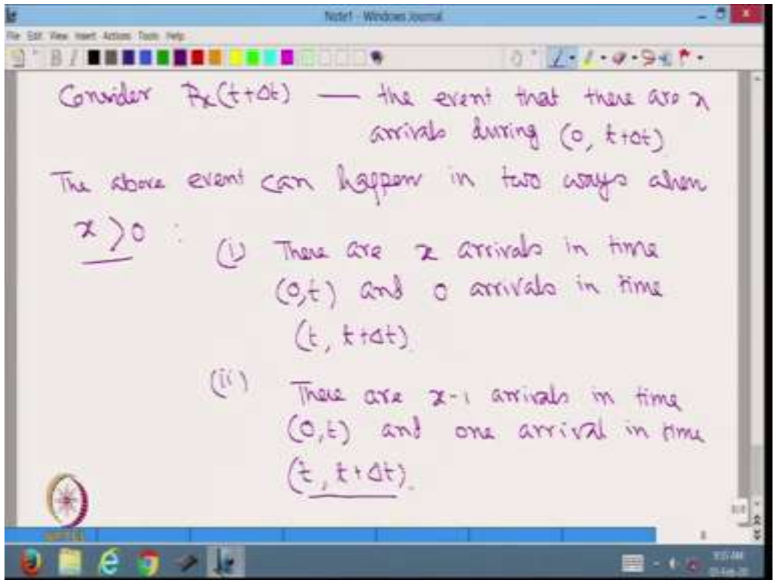
Assumption 3, the probability of getting more than one arrival during t to t plus delta t, a small interval is order of delta t square that is order of delta t square upon delta t will go to 0, as delta t goes to 0 that means that this is the very-very small quantity even in comparison with delta t with this assumption. We are now trying to model the number of arrivals in time t we shall see that, that follows a Poisson distribution with parameter lambda t.

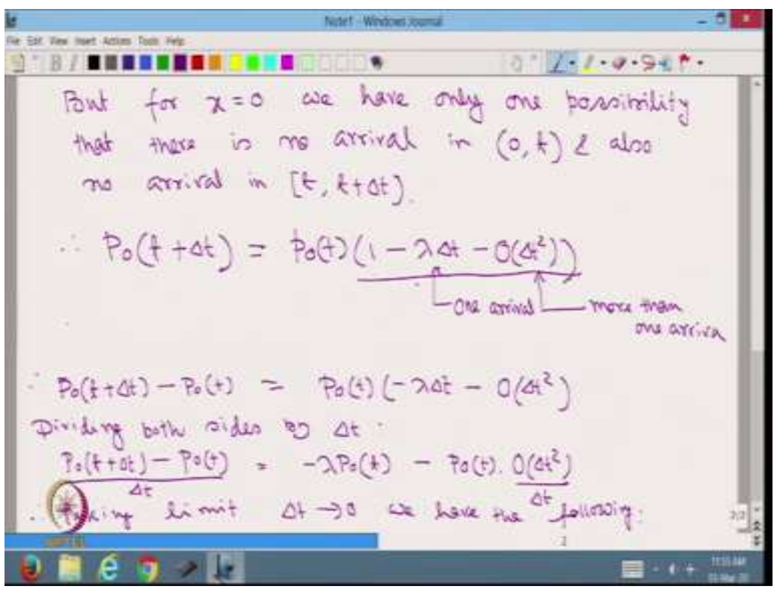
(Refer Slide Time: 11:54)



So, let us start, let xt be the number of arrivals in time t, let Pxt denote the probability of x many arrivals in time t. Note that since it is the number of arrivals therefore x can take values 0, 1, 2, etc. Also note that P 0 0 is equal to 1 because at time 0 there is no arrival that is process starts at time 0 and px 0 is to 0 for all x greater than 0, because the same reason that at times 0 there is no arrival.

(Refer Slide Time: 14:15)





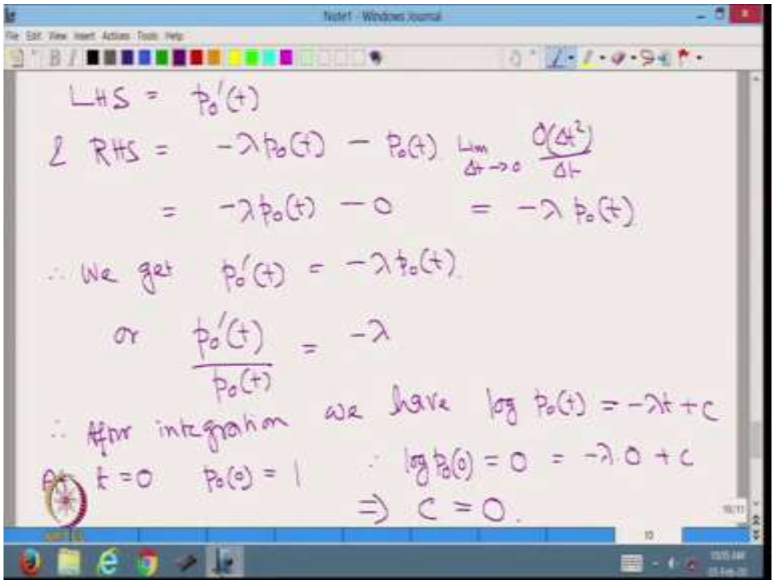
So consider Pxt plus delta t that is the event, there are x arrivals during 0 to t plus delta t, the above event can happen in two ways when x is greater than 0, what are they? The first one is that there are x arrivals in time 0 to t and 0 arrivals in time t to t plus delta t. Other way is that there are x minus 1 arrivals in time 0 to t and one arrival in time t to t plus delta t.

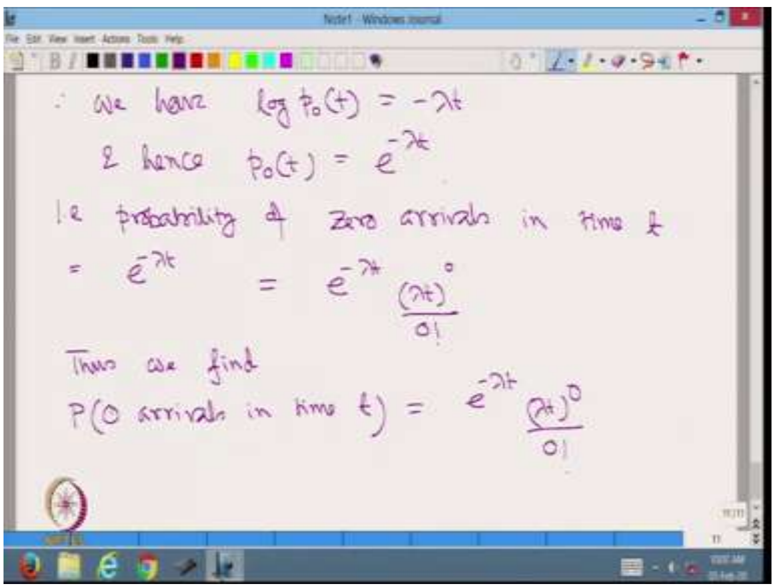
And we have seen that or we have assumed and we have assumed that there cannot be more than arrival more than one arrival in this period. Therefore, these are the only two possibilities when x is greater than 0.

But for x is equal to 0 we have only one possibility that there is no arrival in 0 to t and also no arrival in t to t plus delta t. Therefore, P0 t plus delta t is equal to P0 t into 1 minus lambda times delta t minus order of delta t square because in time interval t to delta t, this is the probability of one arrival. This is the probability of more than one arrival.

Therefore, one minus this quantity gives us the probability of zero arrivals during this period. Therefore, P knot t plus delta t minus P knot is equal to P knot t multiplied by minus lambda times delta t minus Big O of delta t square. Dividing both sides by delta t P knot t plus delta t minus P knot t divided by delta t is equal to minus lambda P knot t minus P knot t into Big O of delta t square divided by delta t. Therefore, taking limit delta t is going to 0 we have the following.

(Refer Slide Time: 20:05)



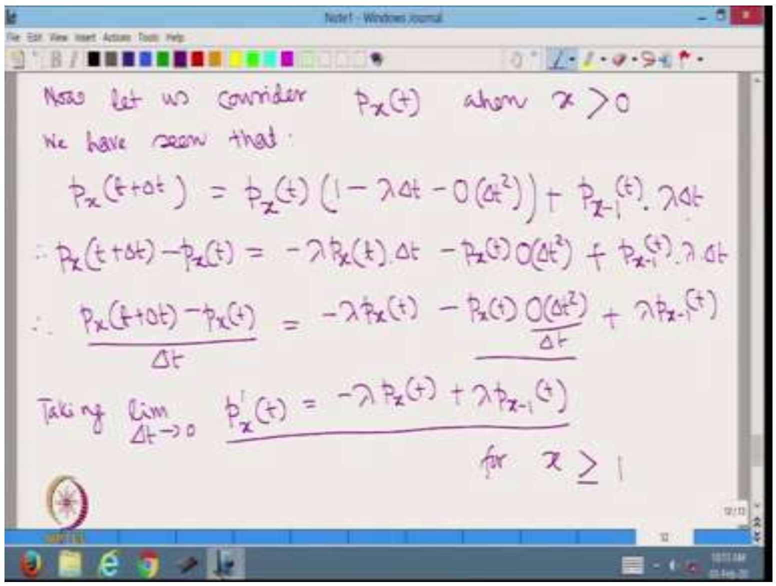


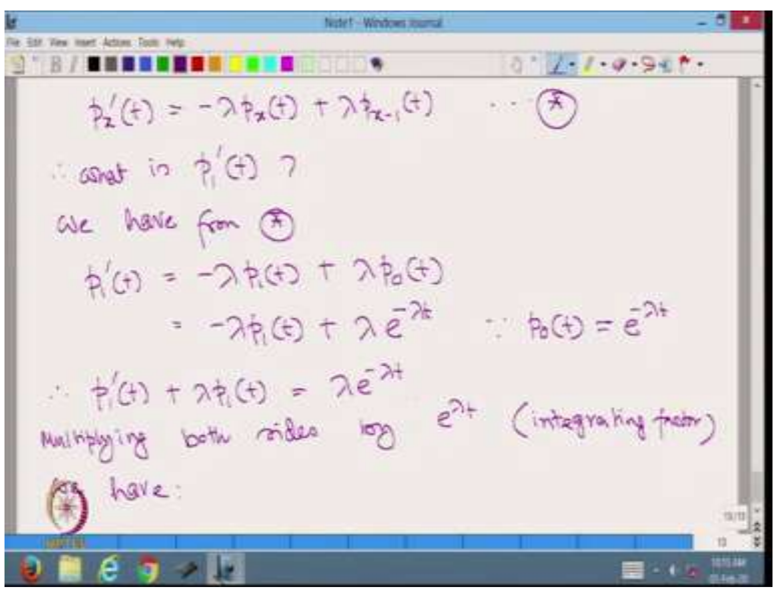
LHS is equal to P0 prime t and RHS is equal to minus lambda p0 t minus p0 t into limit delta t going to 0 Big O of delta t square upon delta t. Therefore, this is equal to minus lambda p0 t minus 0 is equal to minus lambda p0 t. Therefore, we get P0 prime t is equal to minus lambda p0 t. Or p0 prime t upon p0 t is equal to minus lambda.

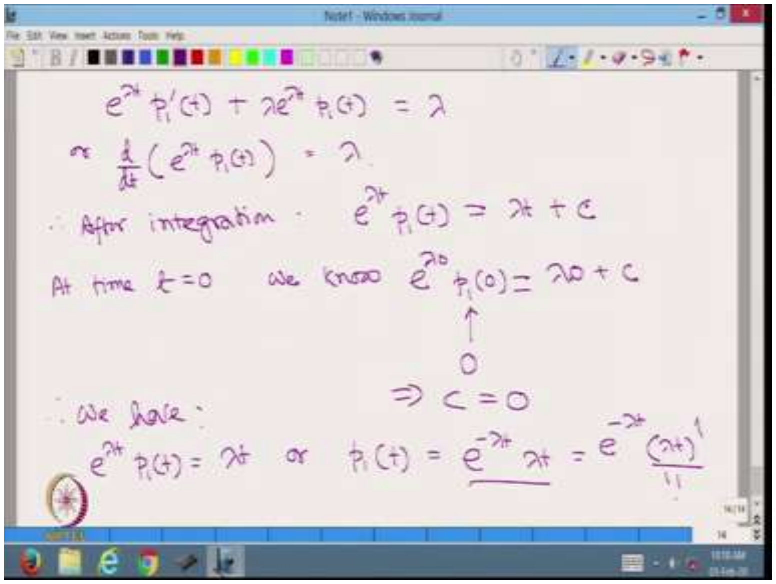
Therefore, after integration we have log of p0 t is equal to minus lambda t plus C at t is equal to 0, P 0 0 is equal to 1. Therefore, log of p 0 0 is equal to 0 is equal to minus lambda times 0 plus C implies C is equal to 0. Therefore, we have log of p0 t is equal to minus lambda t and hence p0 t is equal to e to the power minus lambda t that is probability of 0 arrivals in time t is equal to e to the power minus lambda t which is equal to e to the power minus lambda t, lambda t to the power 0 upon factorial 0.

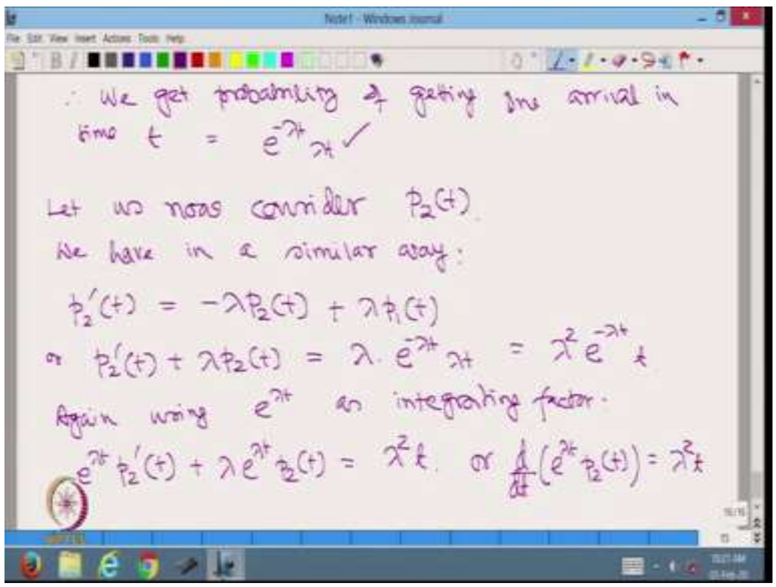
Thus, we find probability zero arrivals in time t is equal to e to the power minus lambda t, lambda t to the power 0 upon factorial 0.

(Refer Slide Time: 23:44)









Now let us consider P xt when x is greater than 0. We have seen that P xt plus delta t is equal to is probability of x arrivals in time t multiplied by 0 arrivals in time t to t plus delta t plus probability of x minus 1 arrivals in time t, multiplied by one arrival in time t to t plus delta t. This we have already seen.

Therefore, P x t plus delta t minus p x t is equal to minus lambda p x t into delta t minus p x t order of Big O of delta t square plus P x minus 1 t into lambda delta t. Therefore, p x t plus delta t minus p x t upon delta t is equal to minus lambda p x t minus p x t Big O of delta t square upon delta t plus lambda times p x minus 1 t taking limit delta t going to 0, we have p prime x t is equal to minus lambda P x t this goes to 0 plus lambda P x minus 1 t.

So, this is the result that we have for x greater than equal to 1. Therefore, what we have? We have p x prime t is equal to minus lambda p x t plus lambda p x minus 1 t, therefore, what is p1 prime t? This is we have say let us call it star from star p1 prime t is equal to minus lambda p1 t plus lambda p0 t is equal to minus lambda p1 t plus lambda times e to the power minus lambda t. Since, p0 t we have already obtained is equal to e to the power minus lambda t.

Therefore, p1 prime t plus lambda times p1 t is equal to lambda e to the power minus lambda t multiplying both sides by e to the power lambda t, which is the integrating factor we have e to the power lambda t into p1 prime t plus lambda e to the power lambda t p1 t is equal to lambda or d dt of e to the power lambda t p1 t is equal to lambda.

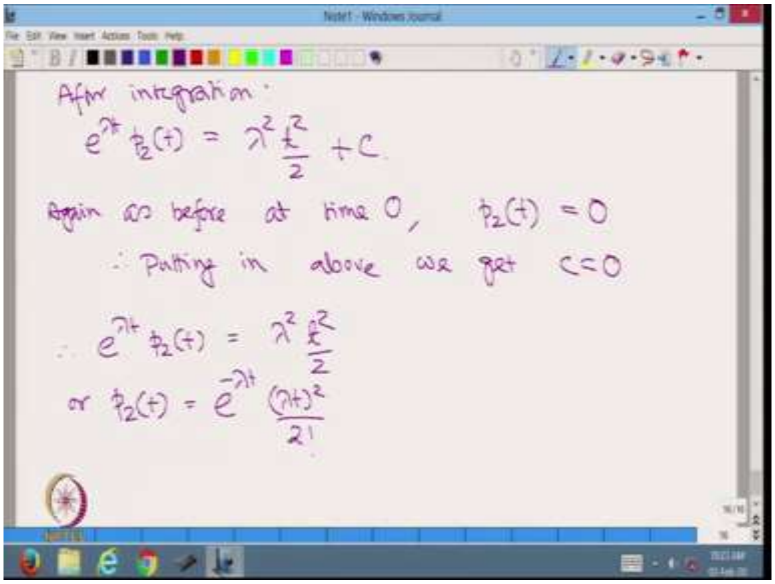
Therefore, after integration what we get is e to the power lambda t into p1 t is equal to lambda t plus C at time t is equal to 0. We know e to the power of lambda 0 into p1 at 0 is equal to lambda 0 plus C. Now this is 0, we have already seen at that at time 0, there would not be any arrival implies c is equal to 0.

Therefore, we have e to the power lambda t p1 t is equal to lambda t or p1 t is equal to e to the power minus lambda t lambda t is equal to e to the power minus lambda t lambda t to the power one upon factorial 1.Therefore, we get probability of getting one arrival in time t is equal to e to the power minus lambda t into lambda t.

Let us now consider P2 t, we have in a similar way p2 prime t is equal to minus lambda p2 t plus lambda p1 t or P2 prime t plus lambda p2 t is equal to lambda into e to the power minus lambda t lambda t. This is the value of p1 t is equal to lambda square e to the power minus lambda t into t.

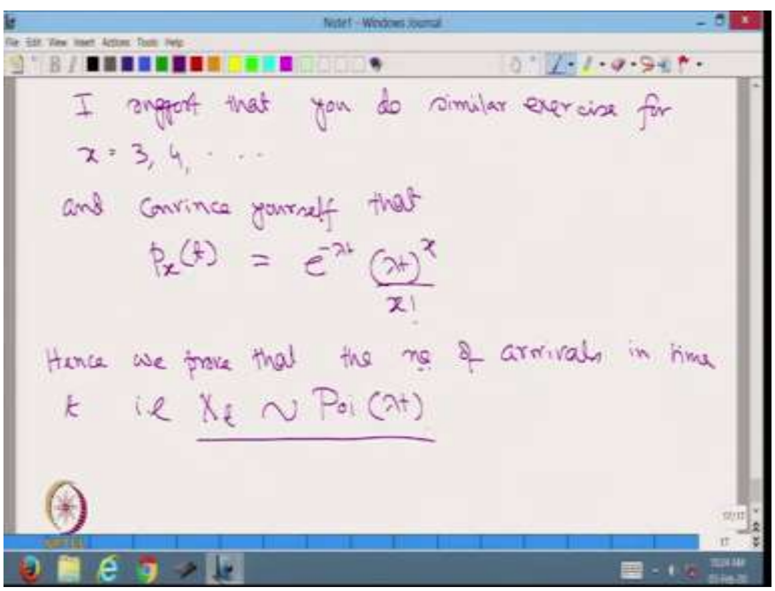
Again using e to the power lambda t as integrating factor we have e to the power lambda t into p2 prime t plus lambda e to the power lambda t P2 t is equal to lambda square t or d dt of e to the power lambda t p2 t is equal to lambda square t.

(Refer Slide Time: 33:04)



After integration what we have e to the power lambda t P 2 t is equal to lambda square t square by 2 plus c again as before at time 0 p2 t is equal to 0. Therefore, putting in above we get c is equal to 0. Therefore, e to the power lambda t p2 t is equal to lambda square t square by 2 or P2 t is equal to e to the power minus lambda t lambda t square upon factorial 2.

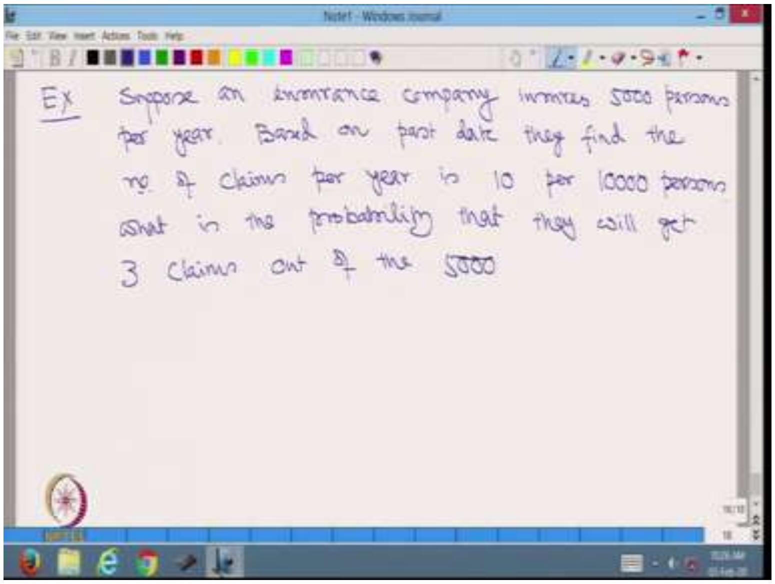
(Refer Slide Time: 34:20)



I suggest that you do similar exercise for x is equal to 3, 4, etc. and convince yourself that probability of x many arrivals in time t is equal to e to the power minus lambda t lambda t to the power x upon factorial x. Hence, we proved that the number of arrivals in time t that is x t the random variable is distributed as well so with lambda t.

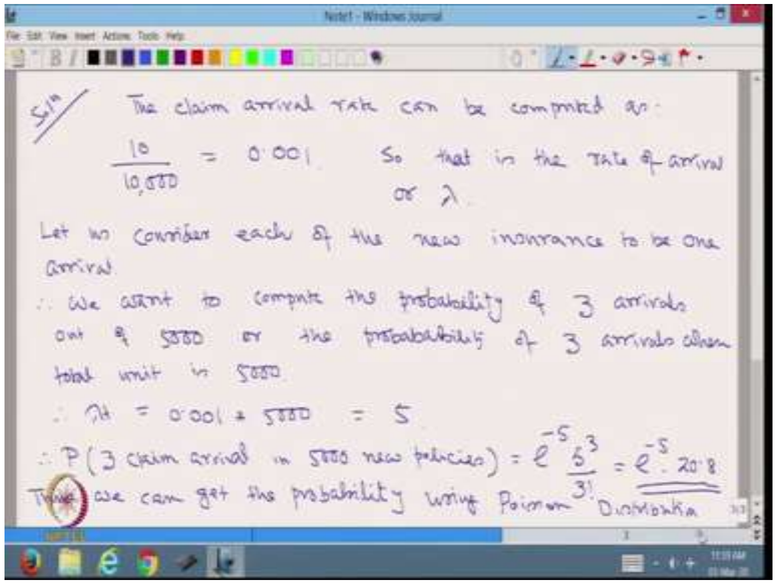
So, very interesting result that we can model the number of arrivals in time t using a Poisson distribution under of course certain assumptions as we have mentioned earlier.

(Refer Slide Time: 35:56)



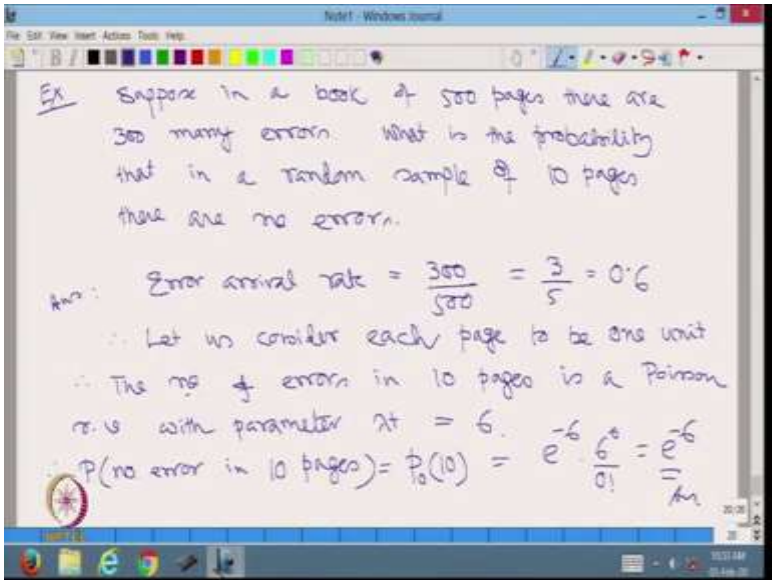
Example, suppose an insurance company insures 5000 persons per year based on past data they find the number of claims per year is 10 per 10000 persons. What is the probability that they will get three claims out of the 5000?

(Refer Slide Time: 37:33)



Solution the claim arrival rate can be computed as 10 upon 10000 is equal to 0.001 so, that is the rate of arrival or lambda. Let us consider each of the new insurance to be one arrival. Therefore, we want to compute the probability of 3 arrivals out of 5000 or the probability of three arrivals when total unit is 5000. Therefore, lambda t is equal to 0.001 multiplied by 5000 is equal to 5, therefore, probability 3 claim arrival in 5000 new policies is equal to e to the power minus 5, 5 to the power 3 upon factorial 3 is equal to e to the power minus 5 into something like say 20.8. Hence, thus we can get the probability using Poisson distribution.

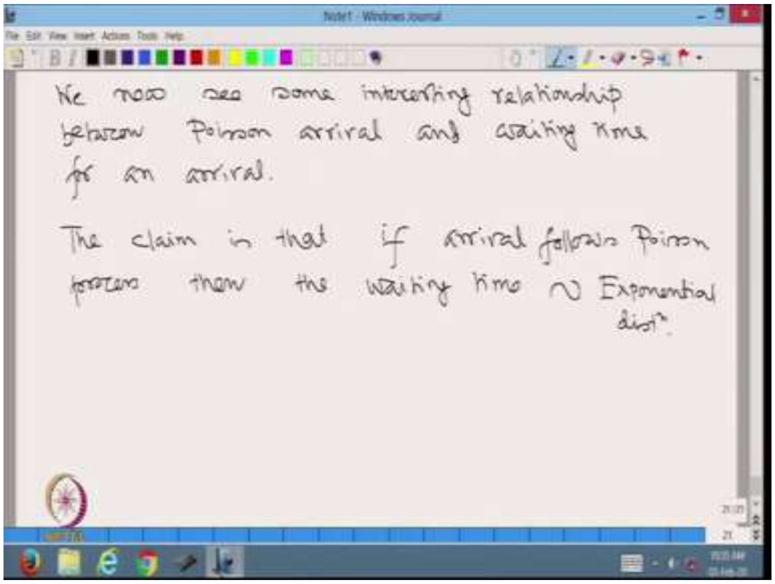
(Refer Slide Time: 40:57)



Another example, example suppose in a book of 500 pages there are 300 many errors, what is the probability that in a random sample of 10 pages there are no errors? Answer, we can see that errors arrival rate is equal to 300 upon 500 is equal to 3 by 5 is equal to 0.6. Therefore, let us consider each page to be 1 unit, therefore the number of errors in 10 pages is a Poisson random variable with parameter lambda t is equal to 6.

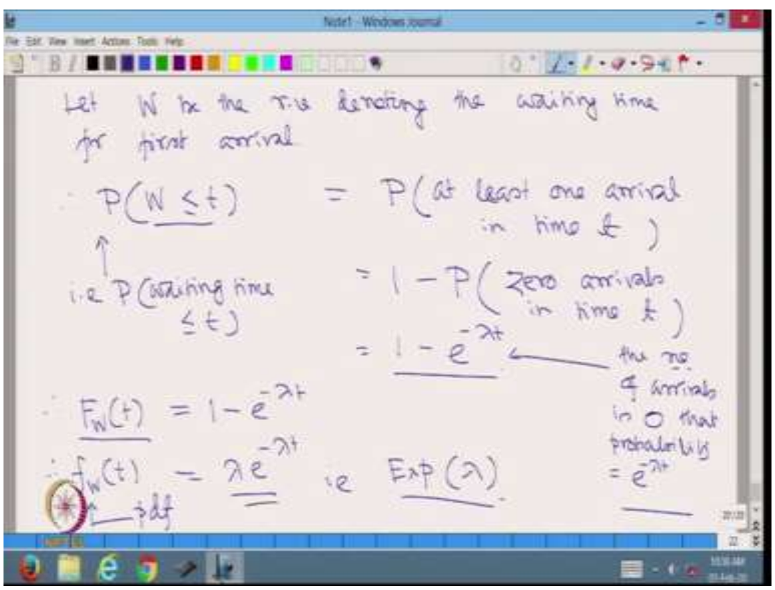
Therefore, probability no errors in 10 pages is equal to P0 of 10 is equal to e to the power minus 6, 6 to the power 0 upon factorial 0 is equal to e to the power minus 6 that is the answer.

(Refer Slide Time: 43:57)



We now see some interesting relationship between Poisson arrival and waiting time for an arrival, the claim is that if arrival follows Poisson process then the waiting time is exponential distribution.

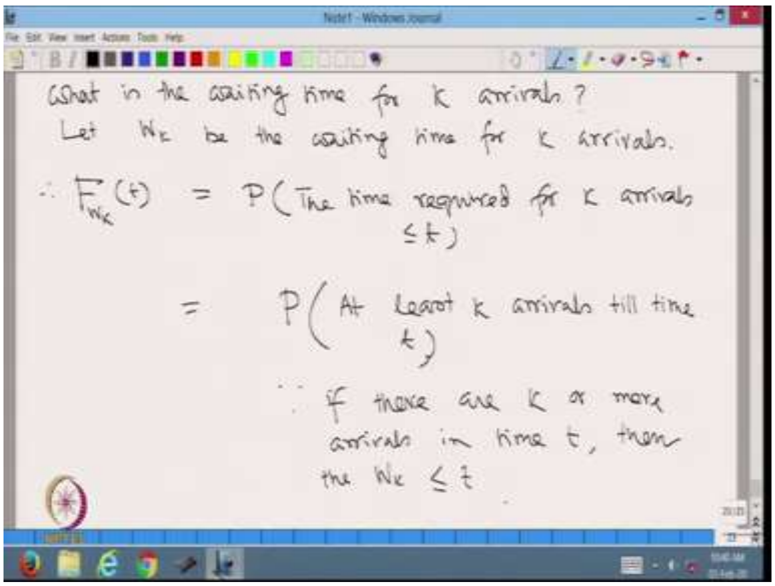
(Refer Slide Time: 45:25)

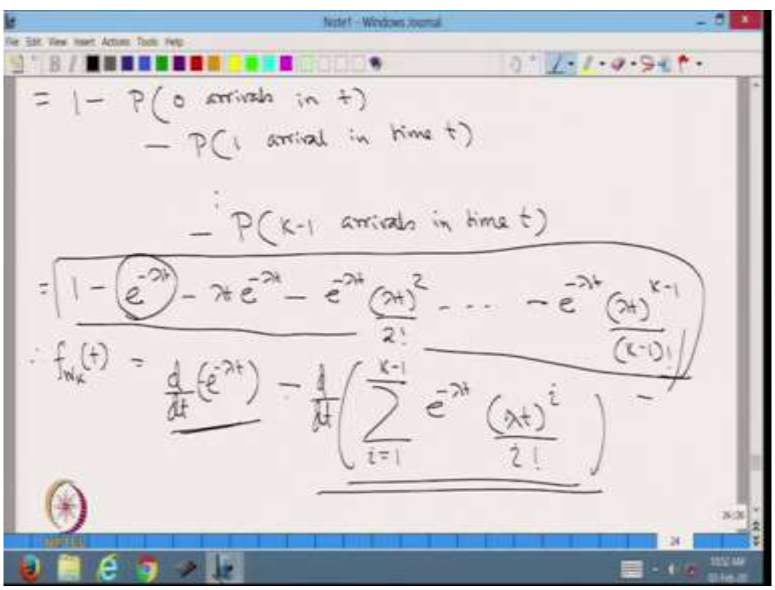


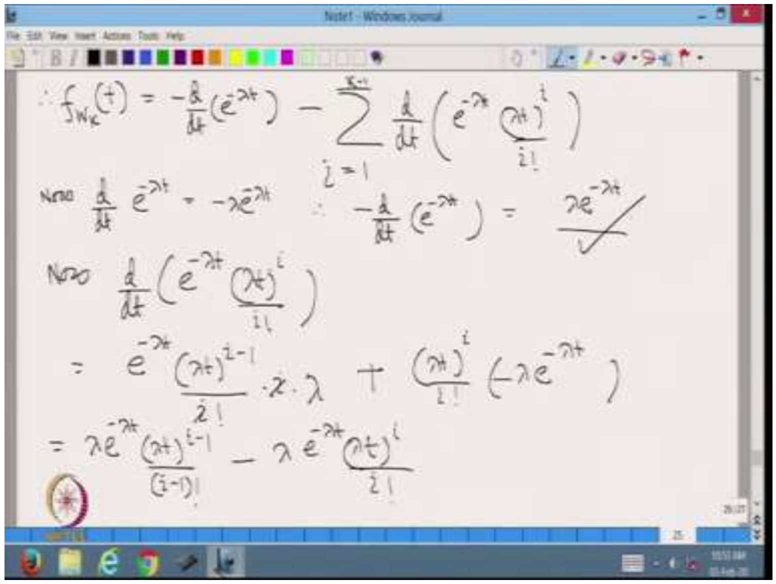
Let w be the random variable denoting the waiting time for first arrival, therefore, probability w is less than equal to t that is probability waiting time less than equal to t is equal to probability of at least one arrival in time t, because if there is one arrival before time t, then waiting time has to be less than equal to t is equal to 1 minus probability of 0 arrivals in time t is equal to 1 minus e to the power minus lambda t. This is because the number of arrivals is 0 that probability is e to the power minus lambda t that we have already seen.

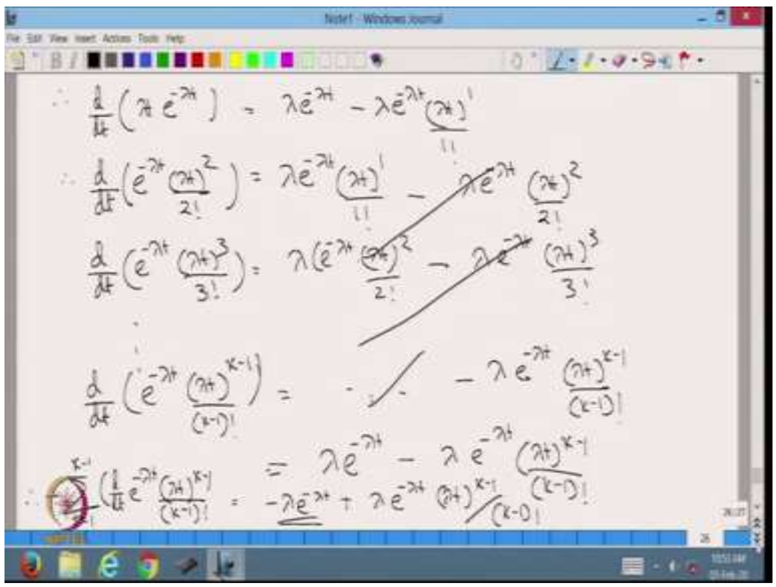
Therefore, if w of t the cumulative distribution function is equal to 1 minus e to the power minus lambda t. Therefore f w of t that is the PDF is equal to lambda e to the power minus lambda t that is exponential with lambda. So, this is a very interesting relationship between exponential and Poisson distribution.

(Refer Slide Time: 48:05)









Now, let us generalize this, what is the waiting time for k arrivals? Let Wk be the waiting time for k arrivals therefore, if Wk of t that is probability the time required for k arrivals is less than equal to t is equal to probability of at least k arrivals till time t. Since, if there are k or more arrivals in time t, then the waiting time Wk is less than equal to t is equal to 1 minus probability 0 arrivals in time t minus probability 1 arrival in time t.

Up to minus probability of k minus 1 arrivals in time t is equal to 1 minus e to the power minus lambda t minus lambda t e to the power minus lambda t minus e to the power minus lambda t, lambda t whole square upon factorial 2 up to e to the power minus lambda t lambda t to the power k minus 1 upon factorial k minus 1.

Therefore, f w k t this we get by differentiating this term is equal to d dt of minus e to the power minus lambda t minus ddt of sigma i is equal to 1 to k minus 1 e to the power minus lambda t into lambda t to the power i upon factorial i. So, I have separated out this from the rest, because it has only one term involving t others have two terms involving t.

Therefore, f w k t is equal to minus d dt of e to the power minus lambda t minus sigma i is equal to one to k minus 1 d dt of e to the power minus lambda t lambda t to the power i upon factorial i, because, with the derivative of the sum is equal to sum of derivatives. Now, d dt of e to the power minus lambda t is equal to minus lambda e to the power minus lambda t therefore, minus d dt of e to the power minus lambda t is equal to lambda e to the power minus lambda t.

Now, d dt of e to the power minus lambda t, lambda t to the power i upon factorial i is equal to e to the power minus lambda t, lambda t to the power i minus 1 upon i factorial then i will come into picture because i into lambda t to the power i minus 1 multiplied by lambda plus lambda t to the power i upon factorial i into minus lambda e to the power minus lambda t is equal to lambda e to the power minus lambda t, lambda t to the power i minus one upon i minus one factorial.

Because, this i cancels with 1 i, minus lambda e to the power minus lambda t lambda t to the power i upon factorial i therefore, d dt of lambda t e to the power minus lambda t is equal to lambda e to the power minus lambda t minus lambda e to the power minus lambda t into lambda t to the power one upon factorial 1. Therefore, d dt of e to the power minus lambda t lambda t to the power 2 upon factorial 2 is equal to lambda e to the minus lambda t into lambda t to the power 1 upon factorial 1 minus lambda e to the power minus lambda t lambda t to the power 2 upon factorial 2.

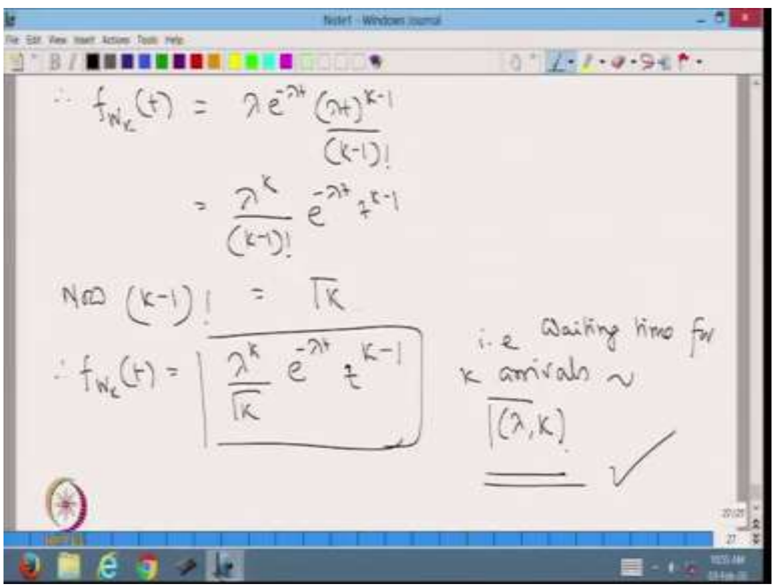
In a similar way d dt of e to the power minus lambda 2, lambda t lambda t cube upon factorial 3 is equal to lambda e to the power minus lambda t into lambda t square upon factorial 2 minus lambda e to the power minus lambda t lambda t 2 cube upon factorial 3 like that if we go then the last term is d dt of e to the power minus lambda t, lambda t to the power k minus 1 upon k minus 1 factorial is equal to.

Let us write the last term, this is going to be lambda e to the power minus lambda t lambda t to the power k minus 1 upon k minus 1 factorial. Now, we note that, this cancels with this, this will cancel with this finally this term will be cancelled. Therefore, the whole result is coming out to be lambda e to the power minus lambda t minus lambda e to the power minus lambda t, lambda t to the power k minus 1 upon factorial k minus 1.

Now, let us look at this entire term we have computed this derivative but it will have a minus sign and this d dt of minus e to the power minus lambda t that we have already calculated to be lambda e to the power minus lambda t. Therefore, minus of summation i is equal to 1 to k minus 1 d dt of e to the power minus lambda t lambda t to the power k minus 1 upon k minus 1 factorial is equal to minus lambda e to the power minus lambda t plus lambda e to the power minus lambda t, lambda t to the power k minus 1 upon k minus 1 factorial.

Now, this term will cancel with this because this is lambda e to the power minus lambda t and this is minus lambda e to the power minus lambda t.

(Refer Slide Time: 58:06)



Therefore, if w k t is equal to what we get is equal to lambda e to the power minus lambda t lambda t to the power k minus 1 upon factorial k minus 1 is equal to lambda to the power k upon k minus 1 factorial e to the power minus lambda t t to the power k minus 1. Now k minus 1 factorial is equal to gamma k this we have seen earlier.

Therefore, f w k t is equal to lambda to the power k upon gamma k e to the power minus lambda t, t to the power k minus 1 that is waiting time for k arrivals distributed as gamma with lambda k.

This is not very surprising, since sum of exponential with same parameter lambda follows gamma distribution. Therefore, this was something we would expect and we have got that result.

Okay friends, I stopped here today. In the next class, I shall start with conditional expectation and variance and also we shall see some important inequalities like Chebyshev's Inequality and Markov Inequality. Okay then thank you so much.