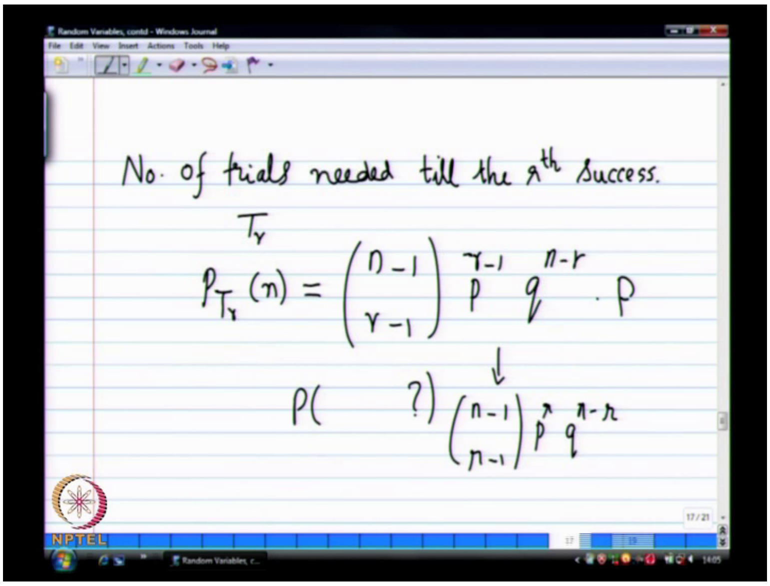
Performance Evaluation of Computer Systems

Prof. Krishna Moorthy Sivalingam
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture No. #08

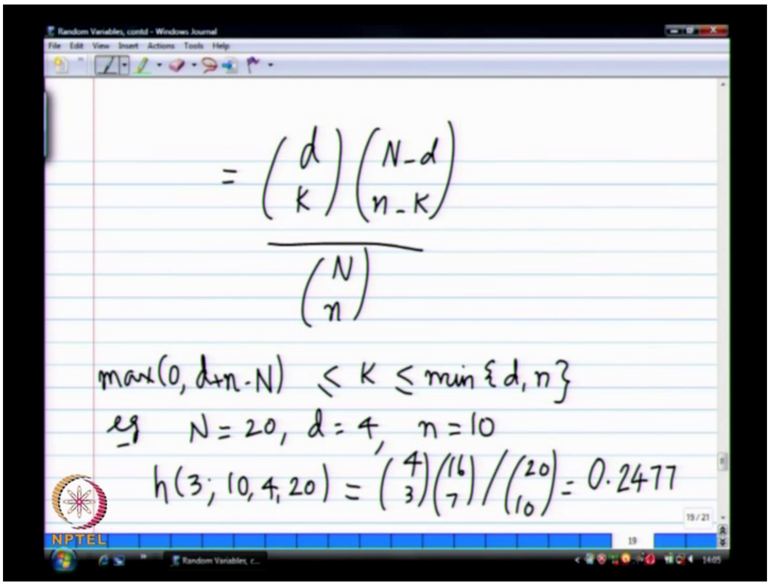
Probability Distributions-III

(Refer Slide Time: 00:18)



There are two pending issues from last class; one is this formula, where this extra p there was missing. So, we need next topic that is all. So, it is total of n trials out of which (no audio from 00:24 to 00:37) so n trials out of which r are successes and n minus r failures. So, that missing p is back is such the p to the power r that is this is basically p power r. So, if we want to rewrite this, it is that is what it, so that way we have r minus 1 failures, sorry r minus 1 successes in the previous n minus 1 failures. So, this was one clarification.

(Refer Slide Time: 01:20)

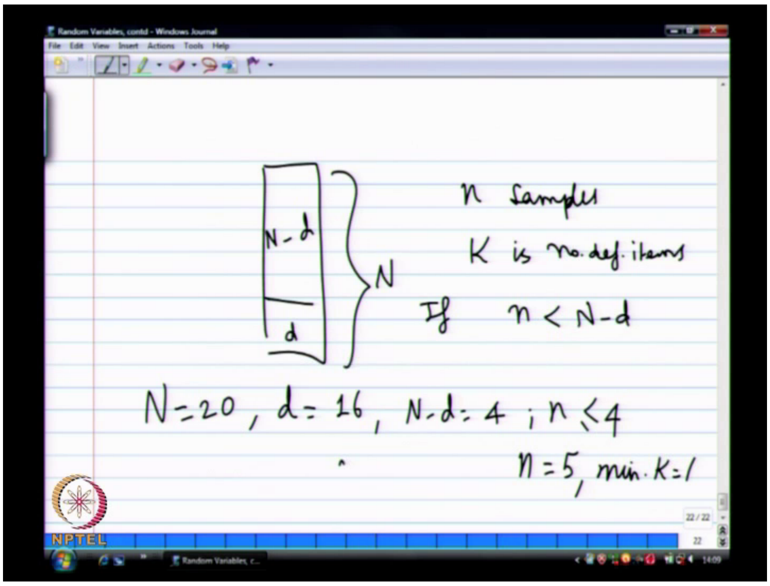


Then the other clarification was to why we need this d plus n minus N. So, we have min of d n that may extends, because you cannot have the maximum number of defective samples or whatever if you are trying to select from that box is bounded by d or if it is you taking d is less than N, and it is got to the N itself. So, only take three samples are a twenty defective items, then you will have it most three defective items, which is what k is representing. On the other side, we have k is you know something max 0 d plus n minus N. So, did I try to explain it, but you are all busy with your to till that in people in to attention, I do not care about it and that finding time. So, what is the explanation for this?

Defective item cannot be Negative.

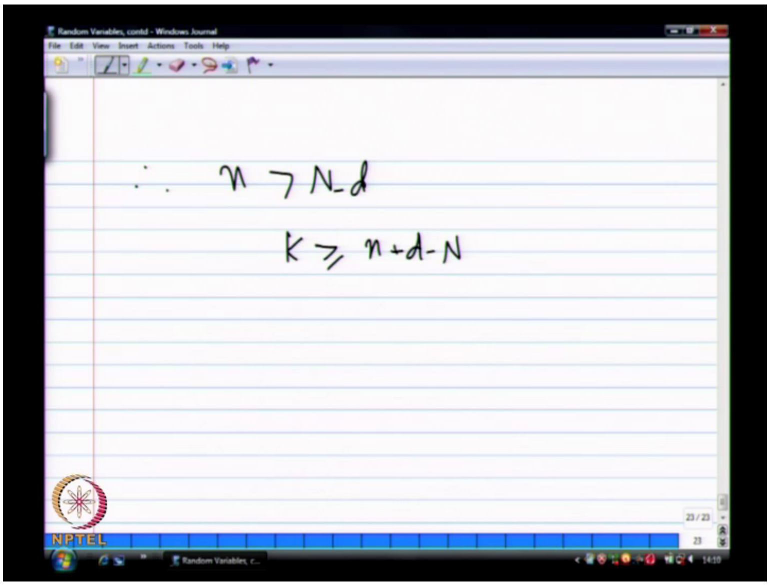
So, the number of non defective items cannot negative they were. So, that is going back in.

(Refer Slide Time: 02:24)



So, if I have a system, where I have total of N items and I have total of d defective items. So, there are N minus d non defective items and if look, see my condition here. So, this condition that we want is accurately so when picking n samples, so K is a number of defective items. And the lowest possible value for K would be, if let us see. So, if f n is less than I will go an example; so what is my example here? So, N equals 20 d equals 16. So, if I have to a 16 defective items in a total set of 20. So, the non defective items is 4 and if I, let say choose if something… So, if n is less than or equal to 4, what will be the minimum value, least value of K we pick only four items, then what can happen, what will be least number of defective item be 0. I have the worst case, if it gets no defective item at all, but if n equals 5 then I could have I have to pick all. I will have at least one defective that will end of choosing whether only four non defective items therefore, there are in the 5 experiments, they could pick the first four terms, you could pick the non defective items the last I have to going till defective one therefore, min value of k. So, minimum value of K is 1. So, therefore, whereas, go to the next page.

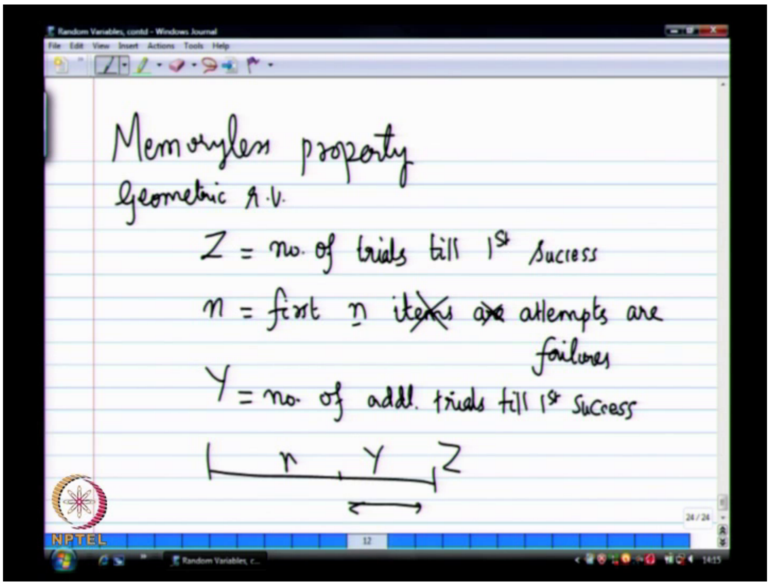
(Refer Slide Time: 05:03)



So, if n is greater than N minus d, is where you have this condition. So, then your K will be so your, K has to be greater than or equal to n minus n plus n minus n plus d minus N make sense. So, that is where that, that is why k has to be either 0 or something else max of just on. So, if will not, you can go back note it out that is a. So, that is where we stopped off last time, with all those to different cables; any questions on the discrete types? When you know, usually looks like a lot of you know math up in the sky, but just get you back something about some of these you will have to go through some definitions in a past you have simply gone into Markov chains without giving much of the background, much I thought giving give you little bit of trigger. So, lot of the step is there in our (( )) book too all this variables are there in the chapter, they mention anything is not much explanations simply a list of like variable, but they thought some more explanations is needed. So, that is about finishes all the discrete random variables that we have covered so for. So, any questions on these? And most of this we encounter on databases. So, depending on what we use later on will choose one or more of this type of variable. Questions?

So, now, let us look at this one property called the memory less property memory less property, we will we look at discrete variables as well as continuous variables or random variables.

(Refer Slide Time: 06:52)



So, which variables are random variable in the discrete set has the memory less property geometric. You will agreed or first time hearing that, how many of you are getting a first time that the geometric is memory less geometric is memory less? So, it will it not for it flow. So, only those did the math code. So, did in these which code you have probability code? So, waiting for them to get a probability, have you done Markov process? Not? So, when you get to Markov process, in something you until then this where ever. So, let us look at this geometric variable and why that has this memory less property. So, the idea of geometric is keep tossing until you find the first success. So, that the probability of number of trials needed is what is described by this geometric distribution, given your parameter P that is a single parameter.

So now, let us say that will use some variables Z is a number of trials till the first success thus the Z and so, if you trying to transmit the packet. So, this is the number of trials to the first success. Let us say, already try transmitting this packet some n number of times. So, you try n times and I have to lot of editing in audio. So, n times you try and then it means failures. So, the first n terms are failures. See the total number of trials and n is the first n attempts, which are all that the failures. So, I use this variables Y, this is the kind to find number of additional trials given. So, the random variable Y is the number of additional trials that I have to do before.

Now you want to see, what the distribution of this Y is be given the fact that already the first n times have been failures, will let make a difference to my probability minus till the clear for success. So, this is Z is total number of trials have done Y and I am looking for lambda and n which is failure. So, why is there representation for? So, after conducting n experiments and looking at the distribution for Y and see whether it has a same distribution as a started with initially.

So, after enclose by I will get.

The first success. So, I am trying to find out whether distribution.

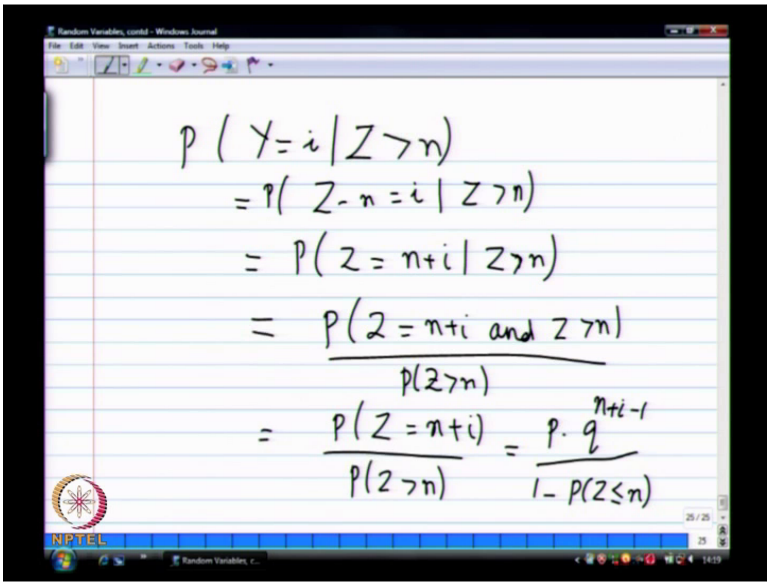
For this variable Y.

What is the distribution for that, given that we know that first n is set 0 is get to that Y is independent of that.

(( ))

The distribution for Y is the same as when you started at the initial distribution, that initial distribution for Z we know that, this p into q time i minus 1 or whatever the variable is it have using a. So, if it is p into q Z minus 1.

(Refer Slide Time: 11:07)



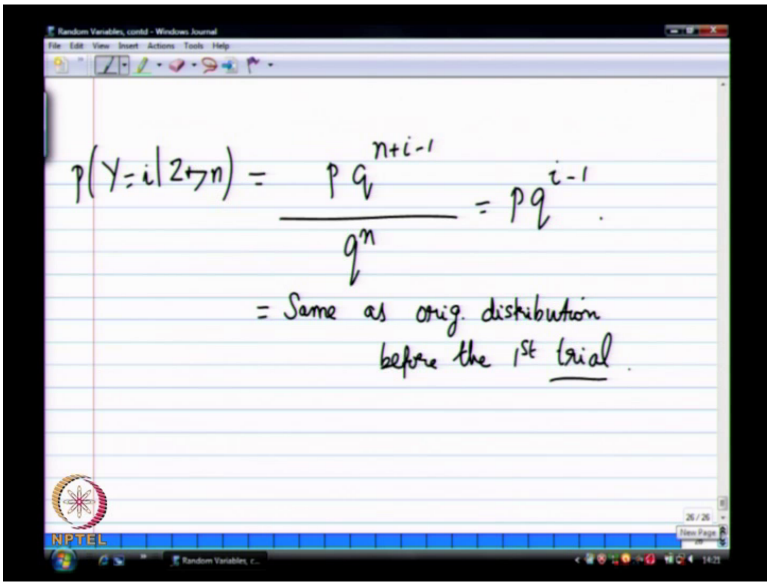
So that we will do as a conditional probability; so I want to find out what is the distribution for Y equals i given that Z is…

(No audio from 11:23 to 11:32)

So, what is the distribution for Y, given that the discrete variable, which we just will little bit time deriving, because everything that we use later on depends upon this business of memory less. So, when I go to Markov chains whether discrete or continuous, we end of having this memory less property. So, that is why we have try to this. I given that, so how do you precede now as a recollect your probability skills? So, how do, why what is the probability distribution here; what is probability distribution of a given b are they conditional on b of a and b divided by probability of d, this is goes back your probability days.

So, if you do not remember that suddenly this is coming a new, then we should go back I am look at that first definition. So, this is basically, what is the probability that Z equals n i and Z is greater than n divided by Z that is looking like to interesting on that can I do this. So, Z equals for any value of i greater than 1, it the second part of the conditional redundant. So, this second part, I can simply ignore and just say, what is the probability that Z equals see, Y equals 1. So, if Z equals n plus one order, naturally greater than n. So, therefore, the second part of the conditions is subsumed by first part of the, for all values of i greater than 1 or gets an equal to one. So, what is the probability that p equal Z equals n plus i in all know, it is p, we use q to did not 1 minus p. So, this is simply q n plus i minus one divided by. So, this is the probability that 1 minus this.

(Refer Slide Time: 14:35)



So, this becomes q. So, P Z greater than one simply Z, all the n trials where failures . So, what is the probability of all the n trials being failures q to the n. So, it simply put that is q to the n. So, remember the definition, what is the probability that Y equals i given that the Z is greater than n is simply a… So, you can see that is no dependent on n on draw, this is the original distribution, we started with we simply said what is that probability of taking four trials it simply p into q 3 and even after 10 trials hundred trials does not matter the probability does not change this is what is called as a memory less property of the geometric variable. So, this is the same as the…

(No audio from 16:00 to 16:27)

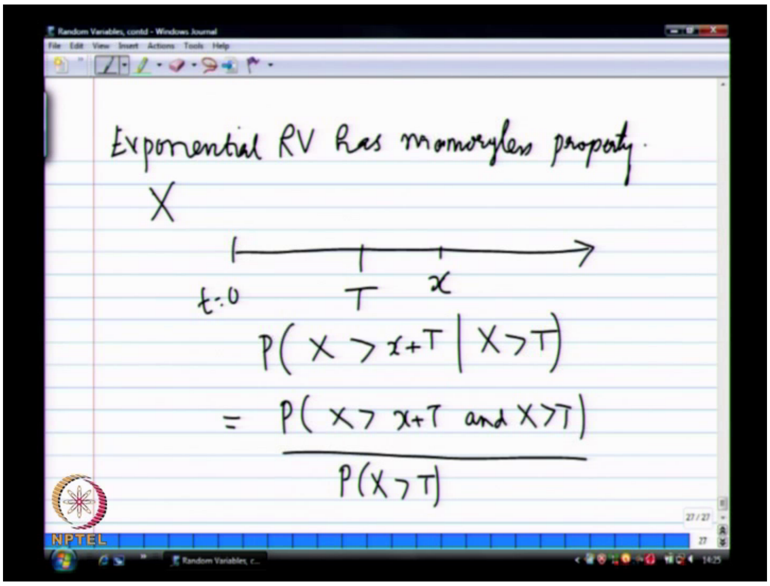
Question?

(No audio from 16:30 to 16:48)

So, then let us we have done with the discrete random variable so for. Questions? We will go to the exponential now. So, the same property in the case of continuous random variable is shown up is shows up in the exponential random variable.

(No audio from 17:04 to 17:33)

(Refer Slide Time: 17:07)



So, how do you go about to things and I have proved it already the other one…

(No audio from 17:38 to 17:59)

It is formal T of S is greater than X, sorry given.

Probability of X greater than T given X greater than S.

Let us try that. So, here what we are having this, either something happening in the time access, usually we look at that is time inter arrival time or something like that. So, that say, have just waiting for the bus to arrive. So, you as a started a time t equals 0 and you waited for some time will call this T; obviously so, you have waited for such a long time bus then you want to know how much longer will have to, what is the distribution for the remaining waiting time, normally what you expect, if you know the buses coming every half hour vary for twenty minutes? You know that the probability of bus coming next 10 minutes is fairly high.

So, but here, we will try to look at… So, given that so, X is my random variable and use as little X. See, use X in T use with X in T. So, what is the probability that the arrival will be before this time X. So, Z what is the probability that X is greater than x plus T, given that X is greater than T. The probability will one and so this probability X greater that will also give me the 111 distribution. So, if I solve it and if is no there is known distribution then I will random. So, how do we go about this? So, we do like before, what is the probability that X is greater than x plus T and X is greater that T divided by probability that X is greater than T.

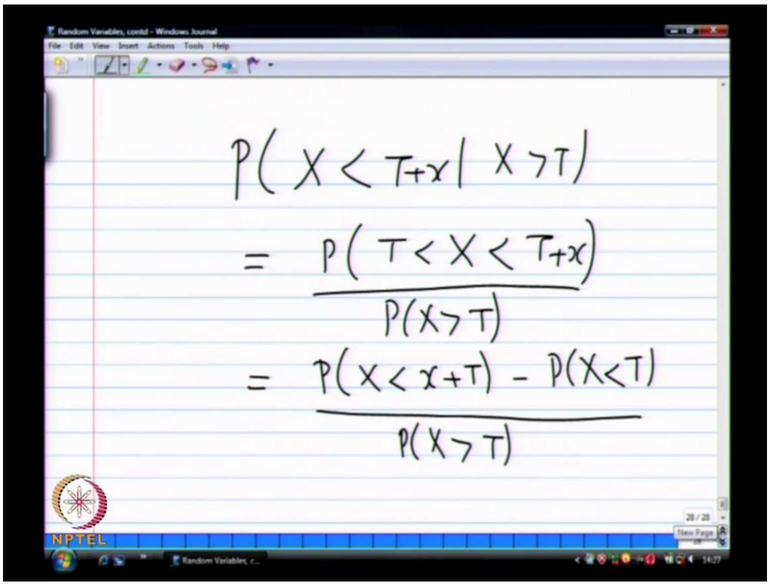
Sir, our variable X is that is the event is the normal arrival of bus.

In the arrival of the bus.

And then you have taken X is greater than T.

So, which means that, what is I have say the you get distribution for this. I can get the we can also derive at the other way if you want.

(Refer Slide Time: 20:52)



So, let us, as it should be normal in a bus, we look at arrival of first time. So, what is the probability that the bus will come before T plus x, given that I waited already for T one X and just will redefining this original thing ? So, this makes sense. So, this I can write as X is… So, X remembers that again conditional. So, this one to be two X, two should be greater that T and X be less than x plus T. So, therefore, this is my condition.

(No audio from 21:36 to 22:14)

(( ))

This can that…

T and time.

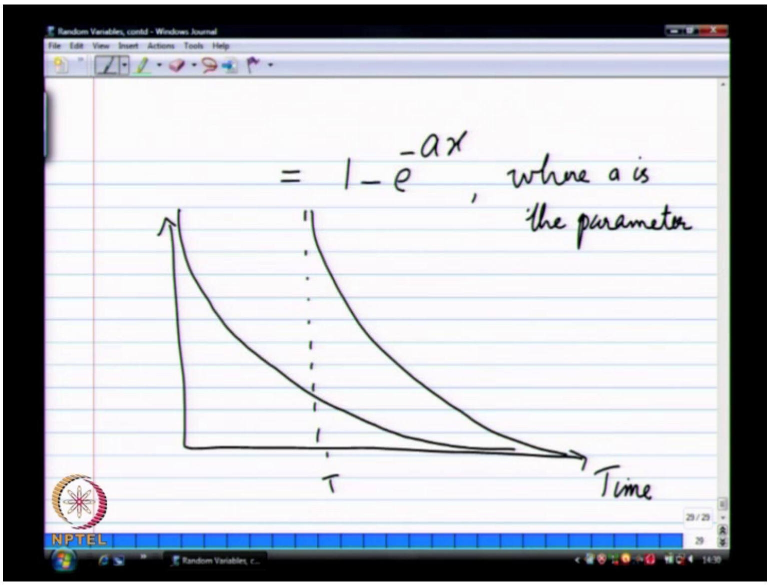
The time diagram if you from r one r two pitch like that. Small X, is the duration between the couple T, no X is…

Correct duration (( ))

(No audio from 22:40 to 22:56)

This is T.

(Refer Slide Time: 23:13)



We used a right?

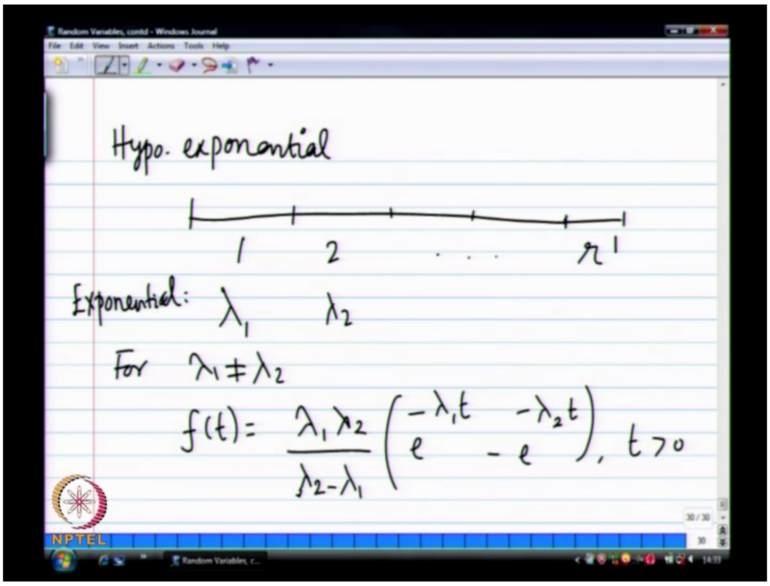
(No audio from 23:00 to 23:39)

So, we look at the total the distribution of this variable x. So, this is the same as the c d of that, they started with one minus e power a x or lambda x, that is series of x before we started c d f of x even after time T one its. So, that is why therefore, fact that you waited for this much time for must not make difference to the probability as such. See, if I want to look at this pictorially, this is your time access. So, the distribution was something like this, when I started this distribution for the waiting time for the bus, now I waited for T time units. So, let say shifted my access and you in still going to be. So, it simply time shifted by T, we can write this mathematically also, but the tell T that whatever probability X of X minus T is the same as X. So, this is again these are the two set of classic memory less property and this only applicable for these two distribution.

(No audio from 24:56 to 25:11)

So, in generally if we find the random variable that you are study in measuring continuous random variable let has as memory less property then we can state that is exponential because only one that has a property. So, that is we test of an memory less property, now we look at two three set of you know derivates of this exponential and that point will stop with allow an random variable customs.

(Refer Slide Time: 25:45)

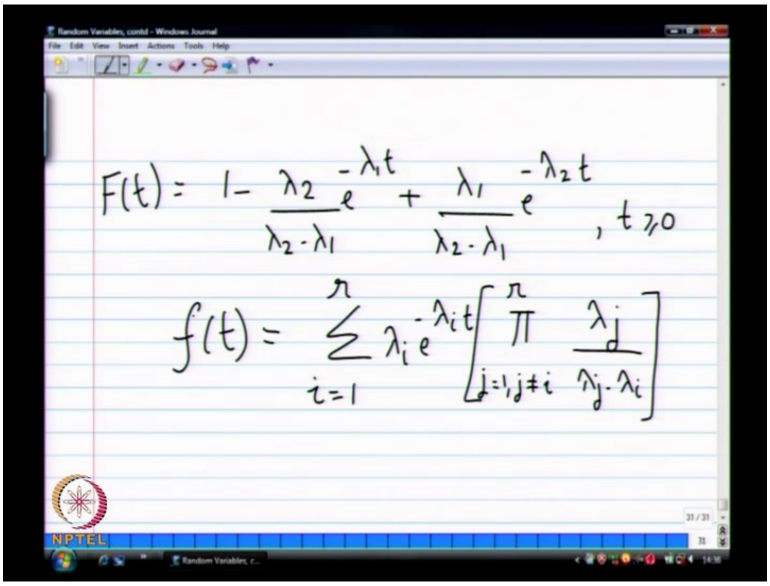


So, the next look, a hypo exponential we that comes often that a little with networks of queues chain of queues. So, rest of this I am going to use lambda 1, lambda 2 and so on. That is the what is there in this the Trivedi's book and just for whatever a call with for a lambda, do not get confuse what is this you have process. So, you have your, going through your entire system is going through several faces, let say this is the pipeline and you going through this faces and each face you spend the certain amount of time and the time that you spend is that is simply not. So, this is face one, face two and so on. The total of r phases the time that you spend in each phases is lambda 1 lambda 2 and so on. So, this is the exponential the theory, that gives.

(No audio from 27:03 to 27:20)

So, before I give the general. So, if I have only two phases. So, what is the distribution for a f of t? So, for lambda 1 not equal to lambda 2. So, whether you have concatenation of phases with each of them, you know that there is the particular this exponential distribution you know the parameter that you can you can you know the distribution of the entire parameter of the entire combine process. So, this is your f of t and just write down f t, for the sake of completeness, we not we not, we may not use the reach of this formulas in future.

(Refer Slide Time: 28:35)



But just for complete the shake end, because the notes when my go on the web or YouTube you have this.

(No audio from 28:40 to 28:50)

So, this for two, if for a two stage process many times you have two stage process, if I want to generalized n stage process then it is.

(No audio from 28:58 to 29:46)

Can you define here?

What does the random variable capture here?

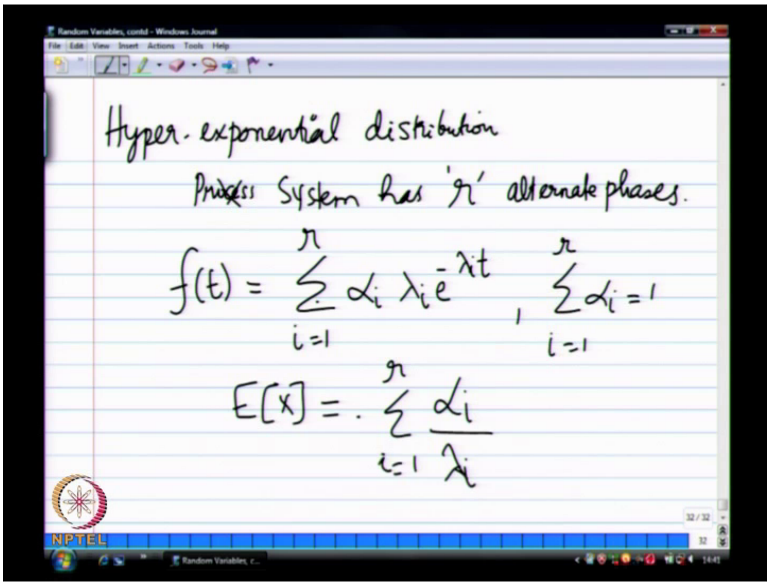
So, this is what it is capturing is, if I have like an assembly. So, let see you are a spending time in two queues and the service time in each queue is exponential. So, your spending time let us say, you have where, you are going to the bank, you first, you have to stand and teller, I want to vary out, wait for the service, you just go the teller one thus, not stand you get waiting simply service time at the first teller for whatever. Your process is doing is the exponential with this parameter lambda 1, then you shunt do another person, who will again make you, we has to service you and then that is also exponential term with service lambda two. So, then we have two separate process, it one follows the other, then we have this hypo exponential the overall. So, the overall waiting time; so T is your total waiting time in the system. So, that is one variant of exponential, then another variant.

(No audio from 30:36 to 30:53)

So, you have series what it could be the other thing parallel. So, you have process that can take one or say r faces and, but each of in have only one of those in r phases it is in parallel. So, you go to the bank, either you go to the DD counter, I go to the cash withdrawal counter, you go to some you know passbook update counter, you only do one of these three things and each of those you have a known for distribution that is exponential with the parameter lambda 1, lambda 2 and so on. So, in that case what is the average waiting time? So, this is this is appear not just in the bank, but also even look at CPU process, the process comes in some times and goes to, I was some time goes to computation, some time goes to whatever those have. So, just so the s to data s data to files and so on.

So, depend each of those will take different time, but you want to look at the overall distribution for the process in general, because you may be you can try to proof properties if you know the overall distribution. So, that is what that is what this is useful. So, that is your hyper exponential.

(Refer Slide Time: 32:42)



(No audio from 32:00 to 32:19)

So, the so, there is a process or that you are here the process, not there stochastic process, simply that is a random variable you are going through some system. I will say system as the say r alternate phases and each time you are tossing this kind of whatever it is, you are going to go through only one those through phases. So, what is the overall distribution in you have the each of the distribution thus the exponential. So, this that is where the f of t is known; so this is again simple. So, there is lambda i, e power lambda i it is simply another. So, you know the probability of going to each one of those phases, which is alpha i. So, the probability, point three we go to one phase, point four go to somewhat phase and so on. And therefore, this is simply the summation of onwards, where your alpha i sums up to, that is why I want to get that mat lab.

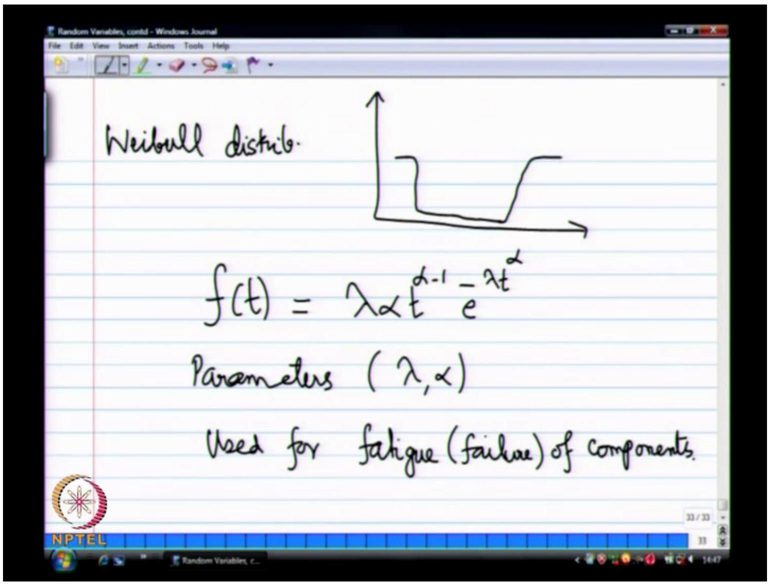
So, you can look, plot some of these things such you want, because sometimes, you find the certain distributions fit hyper better or hypo better or hyper better than simply purely exponential and the exponential still the easiest for getting close form, but for more complicated systems, you may find that the hyper better fits, because you can actually the describe the system that particular way. Then as it is very similar, we had to accept this very alpha i parameter much on and if you are looking. And what is the expected value? Alpha parameter is the probability of going to one of those three phases. So, a customer entering will go through one of those three tellers or counters with some probability of that is some predetermined. So, that is part of your input, input tells input is what alpha and what is lambda.

So, what would be the expected value, for this is the expected value for the exponential variable one over lambda. So, you can just guess what is this can be simply alpha by we can prove this with you know taking a whole class, but… So, that is the expected value and there is again if you look at the E of X square that is also similar with that, because in the case of exponential what is the E of X square two by lambda square. So, this is again of to alpha by lambda i square that i have leave.

(No audio from 35:36 to 35:47)

So, that is the hyper exponential. This two, three more I have done of this variable.

(Refer Slide Time: 36:28)



So, then sum then we have this is called is Weibull distribution or an what I did there, but it something which is not intended. So, when people, where one of things from our performance side of networks and so on. We normally look at exponential, many other things for the service time, the time spent in the service or even if it is the CPU system, you look at as different processors or look at the time taken by each process how long thus the CPU, I have born process thus how long the CPU born process, as you taken then you can, you try to set of characterize them as exponential and so on. So, there is another field, which uses the extensively these variables is reliability theory, where you want to look at probable and reliability the variable for the how long this particular lifetime of a product. So, you want characterize the life time of product written you started of basing the life time is exponential. So, if you are buying a cell phone today, how do we characterize the life time of that, if you want to, I say plot the life time of this at the probability of failure of this component the module it look like. So, it every instant plot the fail the failure or likelihood that this.

Amplify that the n number of see how many time it will fate, you can just you know number of times with have failed divided by f n s see small (( )).

So, you have sat to characterize this as a distribution, you have looked at several devices for long time this is by if by most of the electronic items, if you look at it, this is a guess that, very likely it will fail in the first few days also; that means, somewhere in the quality testing something withdraw. So, the probability of failure that draw it will be little bit high initially and then it will dropdown to 0 for long time and then once it gets with three years, four years then the probability of failure, now starts in a keeping out the end of the something like that, that is your and this your life time this is your. So, if you are able to go past, if you are survive, if your particular products survives for the first two three months without any failure, then I will simply survive for a plot time for something goes on and thus again the generalization. So, we try to. So, therefore this is group that tries to look at life time of product reliability also.

So, what is the reliability, try to the life time at every instant of time will this product fail or not and it is useful in machinery and in aircraft also in aircraft, as million of components and each component as a failure probability. So, you want to compute the overall probability of. So, when going in a design to planes, it has make sure that at least theoretically that can give you some the reliability information that this entire arc of function will, without failures for this extended periods of time. So, this Weibull distribution comes from that side of theory, where people are looking at the life time or of particular product and I find that exponentially is not really capturing the life time, because you know the exponential simply goes on very short period. So, this you add one more parameter. So, called shape parameter that will help you fine.

So, again how do you determine the shape of the parameter that comes from observation of the product about yours, you build the product you look keep on measuring, it is life time four years you see what happen. So, now, what we do. So, I have. So, where I had lambda, I am scaling that by lambda alpha and then I have my e power instead of lambda T which I had before, I am alpha is shape parameter alpha is a shape parameter. So, alpha that will have to figure out. So, your system is specified by two parameter; one is lambda, one is alpha. So, if you are going to say this product is follows a Weibull distribution, you will say alpha equals two or alpha equals three or something like that and then is more term is rather T alpha is input.

(No audio from 40:25 to 40:40)

So, alpha is an input. Alpha is a input. So, the parameters for this are lambda or alpha and this is again generalization of your exponential, if your alpha equals one, this becomes your regular exponential only for alpha greater than one, is its is describing something else. So, when we get that system then, I will try to plot for alpha equals one two and so on. Show how the lifecycle, how to describe the life cycle of a particular process; so this again just something. So, this is use to describe fatigue, if you look at PC to PC or so many thing you have memory CPU and so on. Then vary of an the first thing go us you can or drive in a simply has to replace that which is on a stock system, but the middle of afraid have you goes than look for not for hard drive replace that. So, based on this distribution they will plan the redundancy in the system how to increase reliability and so on.

Sir.

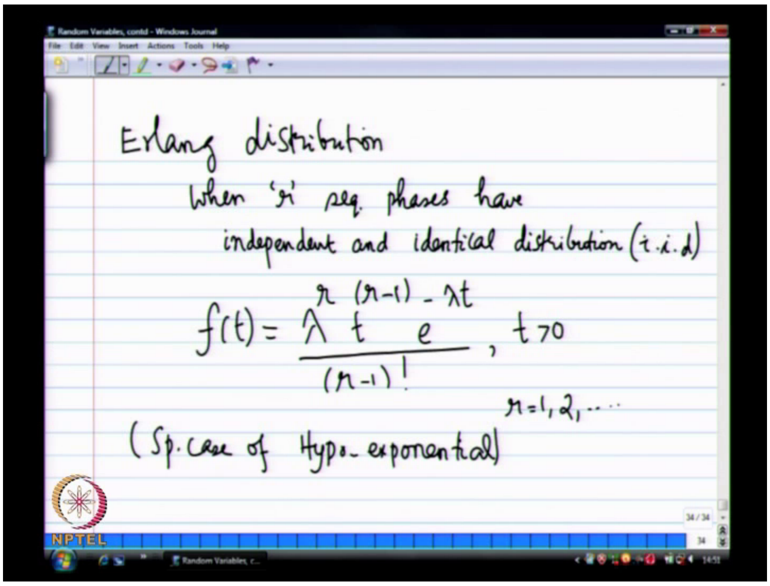
What are the possible values for alpha in space.

It will be greater than one, but to figure out a the value for one that is usually something that will have to your estimate.

Over time greater than one.

What is alpha is less than one alpha is less than one, that may T are alpha minus one is there. Let me check up on that to you see, if I mean theoretically, you get some number, but whether products are there with the particular alpha. So, then we have will cover those variance.

(Refer Slide Time: 42:43)



So, any other variance of anything else try to exponential something called Erlang distribution. Erlang computes in various on the name Erlang of your site. The Erlang's formula and so on. But this is Erlang distribution, so this Erlang distribution is a specialization of your hypo exponential in hypo exponentially is sequential in each of them having different lambda. I you have to they have are identical and I d independent identical distributed. So, when are sequential phases?

((No audio from 43:26 to 43:40))

So, we back to the exponential, that is the definition, that is why the sequential phases is coming in distribution this is your. So, called i i. So, their time spent each of the phases is independent of the other phases, it does not dependent than on the definition, where independent than and an identical. So, with the same parameter lambda then the time total time spend then the system as distribution as follows then this is the…

((No audio from 44:10 to 44:47))

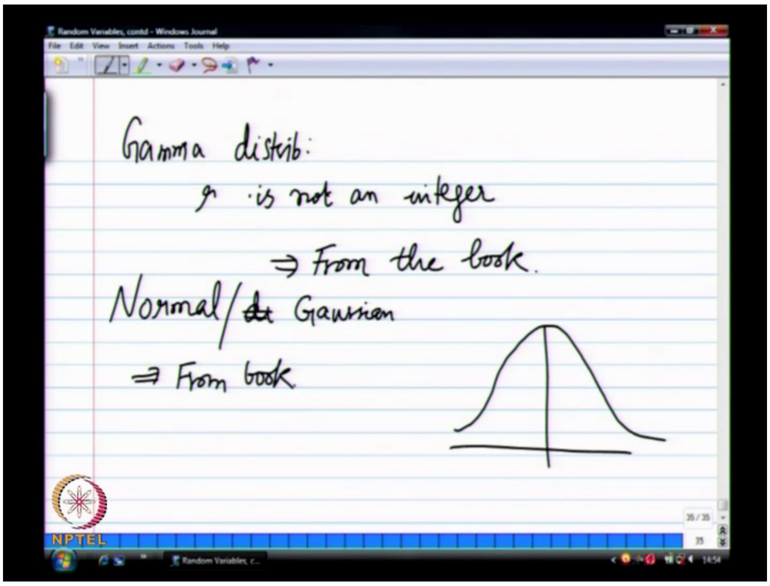
Sir hyper or hypo.

Hypo; hyper is alternate, this is sequential.

This is the same, let I we let last in arrival for sense in queuing theory they were also identical and independent.

When you looking at arrival pass it us to each of the links, then we assume that it is I T. So, in this case the time spent in each of the phases is considered to be I T with this parameter rate parameter lambda and. So, this r f it is an integer it is this Erlang, some cases I have no, I have to look at some better examples for that become relate to having get than that search is, if you a have non integral number of phases, if you have this r is non integral is not integer it could be any real fractional value or real value than you get the gamma distribution, which is little bit difficult to grasp because why would has to become one point two phases. Should I have one phases two phases and so on. But what is my number of sequential phases one point two, I have to find an example for that, but for…

(Refer Slide Time: 46:33)



Now will just…

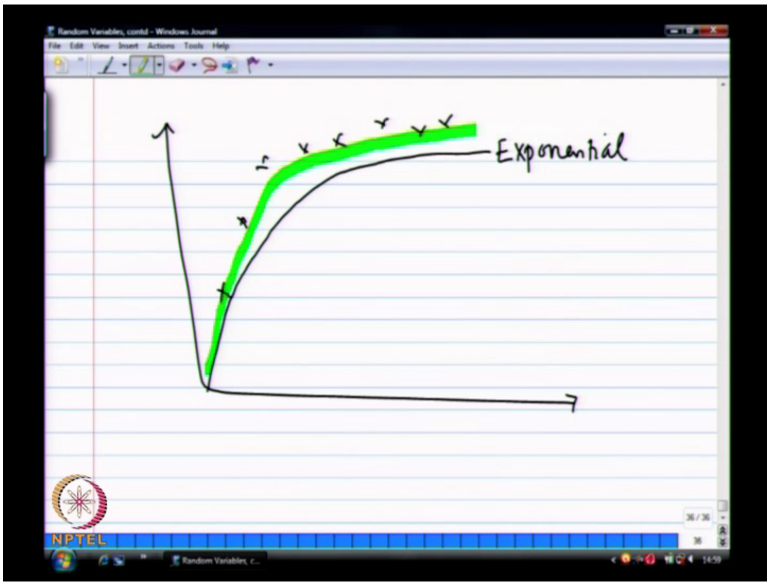
I little take it look at the book ever two uniforms had and…

Then one more or normal distribution which is also, I am also let you look at the book.

So, normal or Gaussian is one that we are know then, we want that we hope that every time I grow my grading we have this perfect bulk have for doing the grading will also, I will ask you to look at the book. So, what are the parameters for the normal mew and sigma the mean and the standard deviation based on that we can and then that is useful and several different ways, many times we assume that things are normal for us normal distribution and based on that we do lot of computations, all that is the round of the sum of random variable that is.

So, what will do is since, some of these variables we do not encounter in all the kind of performance evaluation in. I have an assignment later on, where you will find some of some research articles that use as one at these one of these things and then will go and look for, how these are used otherwise now it is just a bunch of distributions. There were return on board and when ever in see them, but we see them and really of examples and how people have use this to model use normally start to model with some simpler distribution any find that your performance result are not actually matching . So, here is one set of classic example.

(Refer Slide Time: 49:28)



So, you know you have. So, instead that reverse that we to do, what we these we have, we can measure certain things like life time of a product look at the cumulative distribution for something and then you try to fit it back these one of the known distributions because than you want to use the distributions to do your mathematical modeling and so on to do plan for that future I want. So, let us that you know you have some data, that looks like this for something that you measure your measuring f be cumulative your c d f, what you measuring this is the measured values and then you summarize to summarize that this would be some exponential with some lambda, whatever one point 5 you think one point 5 is the I know the correct parameter for you than you turn up like you that is like look like this. So, this is your best fit thinking at is exponential. So, this is for some input parameter let say this is the time that each process run each packets take in the system and so on.

So, you haves the measured value and you this predicted value then you run your theoretical analysis on using this as your input parameter saying that my input is now exponential with this, then you compute delay throughput delay, whatever it is an if find that there is a big gap between, what you think is computed as a delay and what actually is reported by the simulation. So, then you know that the three exponential was the over simplification than you can maybe I should try to do this Weibull or may be some hypo exponential may be the actually two separate processes that is going on. So, in your even sink is this mobile phone also a two phases; one is the normal phase, one is you know when it the where entire start showing up. Even you measure you have the same thing normally everything is working, good one see where enter start showing then we going to the second part of the system which will be another, where the lifetime will be totally different your parameter for what that will be different from the first part of it.

Then you find some other hypo hyper whatever it is and then you find that making use color here no this is highlighted, then you find something else that is better in terms of a distribution itself your parameter estimation is, but just you have in the distribution you have to change your alpha value. So, that is where some of these you know these other distributions come in handy, when if variable to better predict behave you should try to think it as something. Let us simplify, the not you as book as started with we are assuming that in the exponential call that is means that suppose itself equation, that is also it could be something that does not of it anything at all. So, we just use this as a convenience for some analysis analytical purposes something like this, you could try to fit it with that, but the thing it to that assumption can you lets your trying to assume that you are doing some queuing analysis . So, you know if the results for M G one and you are trying to say there I can have, but for you are something that you are assuming at the arrival process is Poisson any where, if it is not Poisson then if you want that to get some addition information.

Then you might see and I have to use something besides Poisson, as the arrival processor where as you can fit it, but these not curve fitting alone, you trying to find out you trying to describe the distribution for the process and based on that you want to do some analysis. For example, you want predict variant and these know the variance for these going to be different such variance, such absorb and based on variance you want to do measurement in term, let you doing some say resource allocation say you want to allocate buffers based on the variance, in this time then you find that your original variance calculation is not adequate, it because your find that you allocated certain number of buffers and these buffers keep overflow in because your assumption was incorrect, you want to look at a more stronger assumption, you will not get the perfect answers, because sometimes there is possible that there is in they does not fit any distribution at all, but you want to do the best guess for planning purposes for performance purposes and so on.

It the standard assumption that was used, even now in the industries at how long thus they average cell phone call type be. In those days, the measure, measured the finally, set this is exponential with period, with average of 5 minutes that is than these over seventies, when eighties, when phone calls very expensive that is the normal you will still see exponential call with 5 second 5 five minutes of holding time, today it is not that the case at all. You find that that call time may be is not exponentially something else and 5 second 5 minutes is also there multiple values for that know business users it have less time errors as general thought as would have more times you want and that case you are the way you provision the system the number of channels we allocate the system will depend upon that. So, that is where these things will be useful for predicting what we have resources to allocate than that is another side of the picture, not just performance alone and also if you are going to be setting performance goals than I have again you will have some understanding of parts. So, if you want to set the delay to some three hundred milliseconds and so on. You should know that it is the fusible delay, that the system can actually tell you.

For that, you need to correct proper station of the correct. So, in there is the networking field at least we have few models, we have the old Poisson model and now exponentially varying packets lies we did that when we did MM one, but whose it packet their exponentially length distribution, but it is not bi model many time its bi model. So, what we do? Then we need have to some of it by model into this and then use that to get your better understanding this and then it say that, when self similar it is not Poisson is memory less know it is not memory less, the distribution itself similar, which means depends upon the previous distribution also the previous packet lengths too.

And then they also looked at for video for example, this is more of a more of MPP or MM PP Markov model, Poisson process is not a simple straight Poisson. All these things people propose they validate and then they ultimately, you never really know how the data files are going to be, what is the typical distribution for that number of packets in a particular video frame; that is still can have you know. So, you end up with traces. So, you try to do a combination (( )). So, that is the end of random variables.