Statistical Methods for Scientists and Engineers
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology – Kharagpur

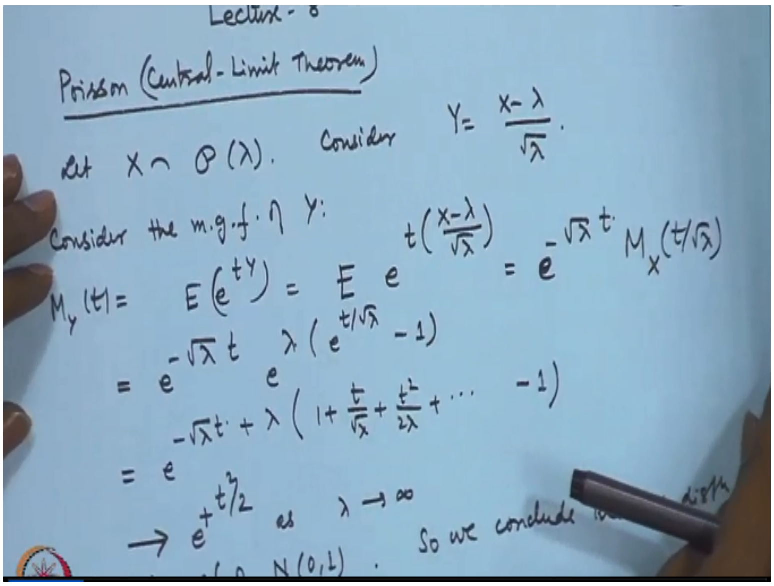
Lecture - 08
Sampling Distributions

In the previous lecture I have introduced normal distribution. Now I did not actually give you how that normal distribution arises. So first of all let us look at the historical development of the normal distribution and then why it has come to be placed as one of the most important distribution in the theory of statistics. Historically if we look at the origins then probably the mathematician.

Gauss was the first one who derived the density function of the normal distribution when he was studying the planetary observations and he derived it as the distribution of the errors so that is why it is also called error distribution and also the function which we use e to the power –z square/2 this is also called error function, but then one of the important results which is called central limit theorem.

So initially it was obtained as a limiting form of binomial distribution or the Poisson distribution, but then later on it was found as a general limiting distribution. So let me state some of these main developments.

(Refer Slide Time: 01:44)



So the first one is called you can say Poisson it is actually called de Moivre–Laplace central limit theorem, but I will just for short I will call it Poisson central limit theorem. It is basically named after the de Moivre–Laplace. The result is that if I consider X to be a Poisson distribution with parameter lambda. Let us consider say Y=x-lambda/root lambda. That means this is actually the standardized form.

The reason is that in Poisson distribution mean and variance both are same=lambda. Let us look at the moment generating function. Consider the m. g. f of y so Myt that is= expectation of e to the power ty that is= expectation of e to the power t x-lambda/root lambda. So this is=e to the power-root lambda*t then this is nothing, but moment generating function of x at the point t/root lambda.

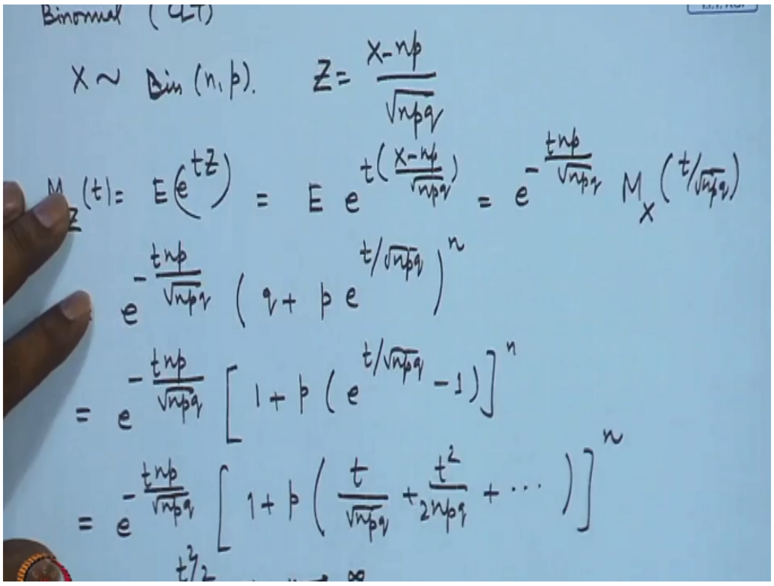
Moment generating function of the Poisson distribution is known to us so that we write for t/root lambda so that is= e to the power-root lambda t e to the power lambda *e to the power t/root lambda-1. So we can do some simplification here it is e to the power –root lambda t+ lambda. Consider the expansion of this 1+t/root lambda +t square/2 lambda and so on-1. So easily you can see this 1 cancels out then next term root lambda t also gets cancelled out.

And if I take the limit as lambda tends to infinity then all the term here will get cancelled out except the third term here. So this will converge to e to the power-t square/2 as lambda tends to infinity rather +t square/2. So this is the moment generating function of normal 0, 1. So we conclude that the distribution of y converges to normal 0, 1 as lambda tends to infinity. Now this is one of the first manifestations of normal distribution or you can say origins.

Because what we are seeing we are looking at actually the rate of arrival in a Poisson process is not it that is we are looking at how many occurrences in an interval of length t. So that is a discrete random variable, but as lambda becomes large that means the rate is more that means there will be more and more number of occurrences. So this can be approximated by a continuous distribution.

So here what we are saying is after standardization it is becoming x-lambda/root lambda that is normal 0, 1 that means roughly we are saying x has a normal distribution with mean, lambda and variance lambda.

(Refer Slide Time: 05:48)



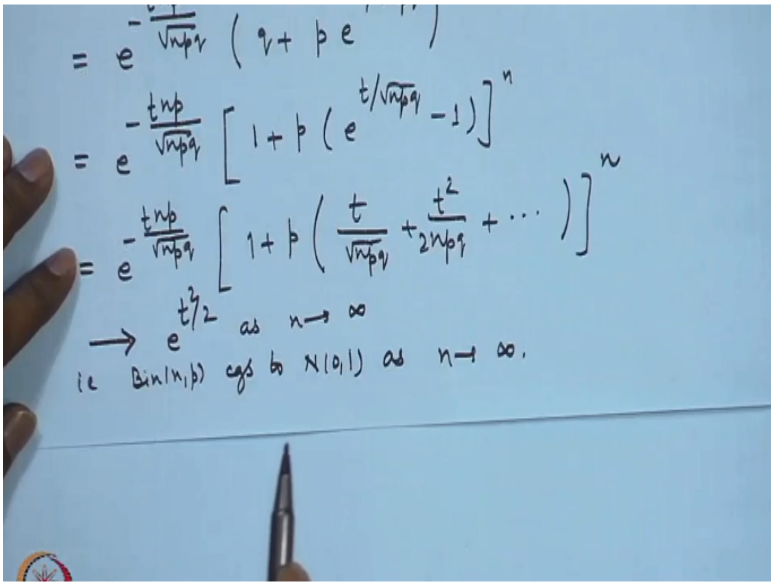
Let us look at another one which I call say binomial limit theorem or which is again attributed to de Moivre–Laplace. So let us consider say x follows binomial n, p distribution. And once again let us consider say Z=X-np/root npq. Let us again consider the moment generating function of z that is=expectation of e to the power t x-np/root npq. So that is=e to the power-tnp/root npq and the remaining part will be moment generating function of x at the point t/root npq.

Moment generating function of the binomial is q+ p e to the power t whole to the power n. So from there we conclude that it is e to the power-tnp/root npq q + p e to the power t/root npq whole to the power n. So this we can write as e to the power-tnp/root npq and this term we write as q I write as 1-p. So I can write as 1+p e to the power t root npq-1 whole to the power n.

This one I can expand so this is becoming e to the power –tnp/root npq 1+p. So this will become 1+ something so that 1 will cancel out and you will get t/root npq+ t square/npq+2npq and so on. That means higher power of t will come and higher power of n will come here. So if I take the limit as n tends to infinity. See this term will give me n time here so that will get cancelled out.

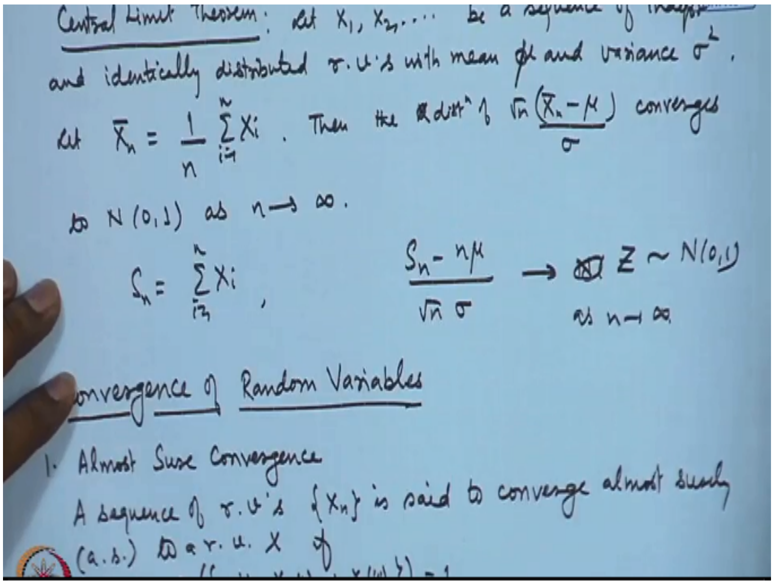
And then the remaining term will give me again t square because this npq term with n will get cancelled out. So as this tends to it will converge to e to the power t square/2 as n tends to infinity.

(Refer Slide Time: 08:54)



So we have the second central limit theorem that is binomial distribution converges to normal 0, 1 as n tends to infinity. We have earlier seen that as n tends to infinity and p tends to 0 such that np tends to lambda then binomial converges to Poisson, but if we simply have a condition that n tends to infinity then actually the binomial distribution can be approximated by a normal distribution.

(Refer Slide Time: 09:29)



Now from here actually we have the more general central limit theorem. Let x1, x2 etcetera be a sequence of independent and identically distributed random variables with mean mu and variance sigma square. Let us consider say xn bar to be the mean of the first n variables here. Then the distribution of xn bar-mu/sigma * root n converges to normal 0, 1 as n tends to infinity.

In fact this is one of the basic central limit theorem. Here I have made the assumptions that the random variables are independent and identically distributed so with mean mu and variance sigma square. Actually what is the significance of this result? What I am saying here is that if I consider the mean of the observations then that will be normal no matter what is the original distribution.

An alternative version of this can be written in terms of the summation also like if I consider Sn= sum of the first n observations then the central limit theorem will be Sn-n mu/root n sigma then this will converge to z which follows normal 0, 1 as n tends to infinity that means either we consider the sample sum or the sample mean the limiting distribution will be normal no matter what is the original distribution.

Of course we have to have the existence of the mean and the variance. Later on the generalization of this results have been done to the sequence of random variables which may be non identically distributed that means you may have here mu I and here you may have sigma I square, but then you can consider the suitable version here by replacing by the mean of the first n mean and here similarly convergences.

So similar versions do exist later on and of course the condition will be slightly more stringent rather than we consider variance we consider something more than the variance that is more than the second moment should exist and then even more further generalizations are there where the concept of independent has also been relaxed. However, in this course I will not be mentioning the full statement of this central limit theorem.

Those who are interested may look at some of the books on limit distribution or the advanced probability theory. For example, the book by (()) (13:05) Kingsman and Tailor etcetera where all this results are mentioned. Now this is the result which actually places normal distribution in the center of theory of statistics because no matter what original distribution you are starting with but if you consider the mean of the n random variables.

Then that is having a limiting distributing which is normal. Now what is the significance of this? The significance is that in most of the practical problems for example you can see here the problem of measurements for example how the Gauss arrived at it because he was considering the measurements of the astronomical distances and many other planetary observations he was considering.

So that mean in place of one observation you will take several times the observations to account for the error and then you will take the average of those observations rather than taking individual observation you consider the average of the measurements taken several times. So Sn becomes large this convergence is to the normal distribution. So this is one of the practical aspects also.

For example, if you consider the performance of a student in an examination. Now in examination different questions will be there because the question paper consists of several questions for example it may have 30 questions or it may have 50 questions. So the score of the student will be actually the total performance over all the questions that means the assumption of n being large can applied and if we assume that his ability in answering the questions will be similar.

Then the marks or the score of the student can be considered to be normally distributed. Similar thing happens almost in various areas of human life. For example, if you consider human abilities or the height of a person say the distance a person can travel in an hour and so on. Many of these things have been found to follow normal distribution. Related to this central limit theorem there are some other simplistic concepts also which we call in general laws of large numbers.

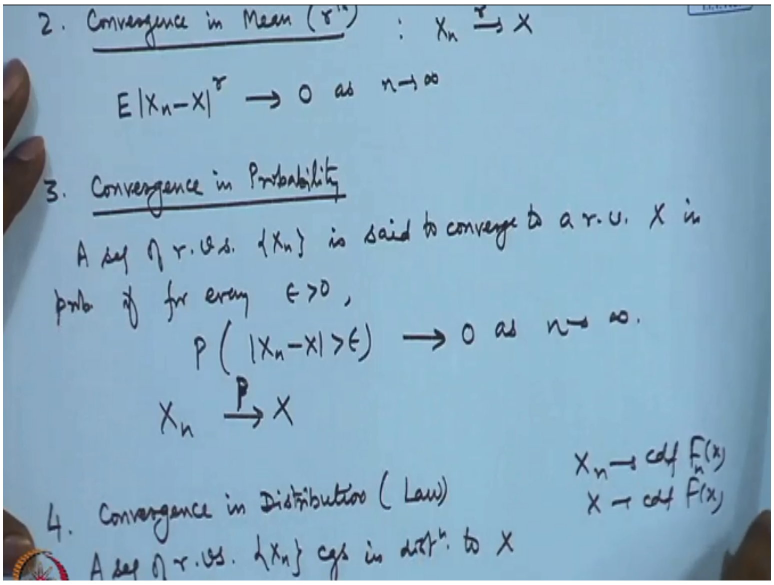
Let me just mention the simplest version as we have mentioned here this for the independent and identically distributed random variables. Before that briefly I will just mention because here we are talking about the convergence. So this convergence is clear in what sense it is. I am shown m g f converges which means that if I am considering the cdf then cdf of this quantity will converge to the cdf of normal 0, 1.

And similarly in the binomial if I am considering the cdf of x-np/root npq then that will converge to the cdf of normal 0,1 and similarly for the Poisson. So that means I am talking about something like convergence in distribution. Now likewise I can introduce some more convergences. Here in a brief form I will introduce those convergences and based on that I will talk about the laws of large numbers.

So the concept of convergence of random variables. Although I will not go into deep in this concept here. I will only mention those who are interested may read the advance text on the probability theory as I mentioned just sometime before. In particular, there are 4 types of convergences. The first one is called Almost Sure Convergence. So a sequence of random variables.

Of course we assume that the probability space will be the same for all of them. So Xn is said to converge almost surely that is I will write in short to a. s to a random variable X if probability of the set such that Xn omega converges to X omega is=1. This is called Almost Sure Convergence.

(Refer Slide Time: 17:58)

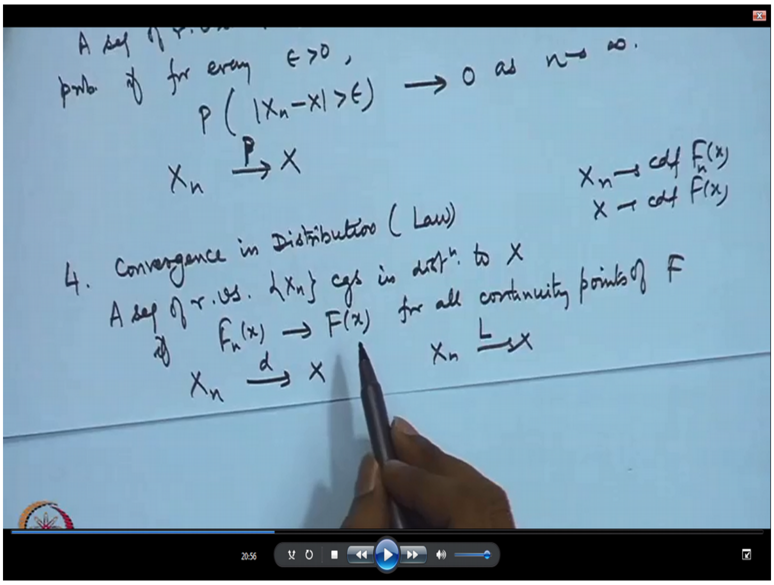


Then you have Convergence in Mean. So actually I consider rth mean. If we have expectation of modules Xn-x to the power r converging to 0. So we say a sequence of random variables Xn converges to X in rth mean if expectation of modules Xn-X to the power r goes to 0 as n tends to infinity. So we say here Xn converges to X in rth mean. The notation for almost sure convergence is we write Xn converges to X almost surely.

So similarly we have convergence in rth mean then we have convergence in probability. So once again a sequence of random variables Xn is said to converge to a random variable X in probability if for every epsilon>0 probability of modules Xn-X > epsilon this goes to 0 as n tends to infinity. We actually write in notational terms as Xn converges to X in probability and sometimes capital P and sometimes small p is used and the convergence in distribution which we actually used in the central limit theorem.

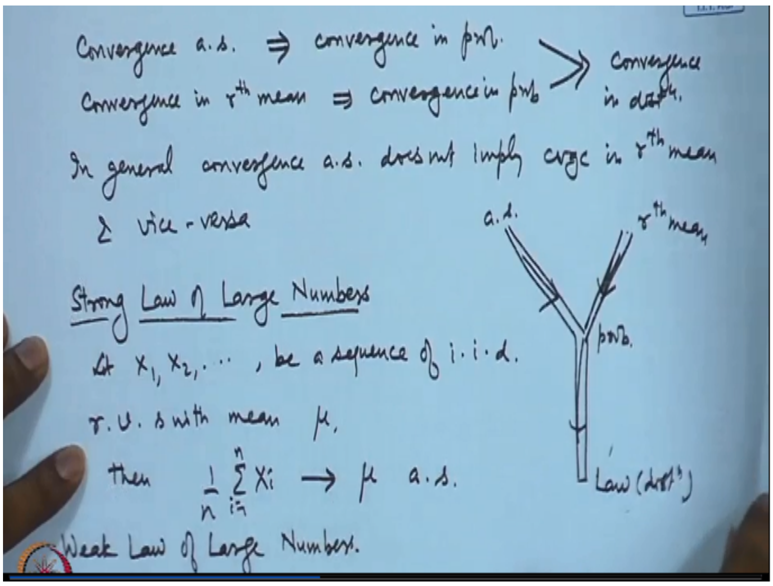
But let me formally write it convergence in distribution also it is called convergence in law. So let us consider say Xn has cdf Fn and x as cdf say Fx. So we say a sequence of random variables Xn converges in distribution to X.

(Refer Slide Time: 20:25)



If Fnx converges to Fx for all continuity points of F that means this x is a point at which F is continuous and the notational is Xn converges to X in distribution or sometimes we say Xn convergence to Xn law. The first thing is that one should ask that what is the relation between these various types of convergences. So without going into proves and other things I will mention this thing.

(Refer Slide Time: 21:05)



Convergence almost surely implies convergence in probability convergence in mean implies convergence in probability and of course convergence in probability implies convergence in distribution. Of course neither of convergence almost surely or rth mean imply each other without any conditions. In general convergence almost surely does not imply convergence in rth mean and vice versa.

Convergence in probably does not imply convergence almost surely convergence in probability does not imply convergence on rth mean. Convergence in distribution does not imply convergence in probability. So actually we can describe this relation that means the flow of convergence in the form of a funnel. So you have convergence to almost surely convergence in rth mean.

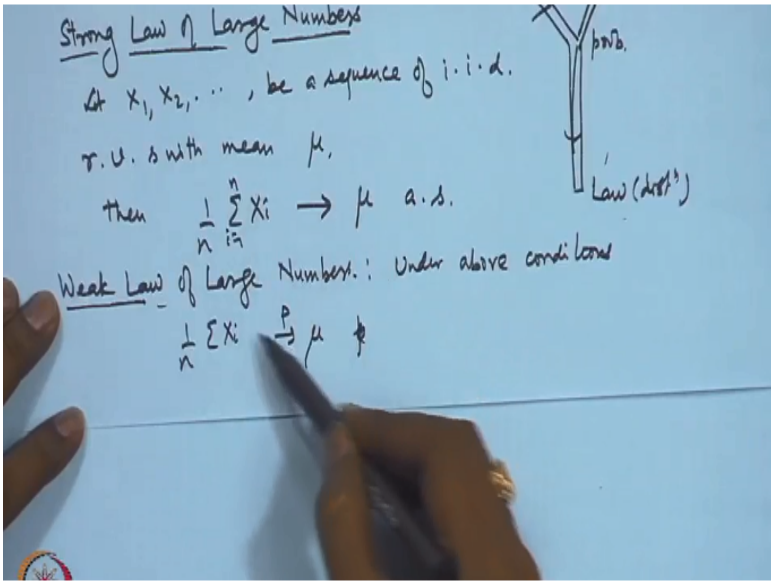
So suppose I pour a liquid in a funnel then the liquid will flow down so this is convergence in probability and this is convergence in law or distribution. So convergence is almost surely implies convergence in probability. Convergence in probability implies convergence in law. Convergence in rth mean implies convergence in probability, but neither of this imply each other.

Of course under certain conditions convergence almost surely will imply convergence in rth mean and vice versa. Similarly, if I impose some condition on the convergence in probability it will imply convergence in rth mean or it will imply convergence in almost surely and similarly if I put some condition in the random variable then convergence in law may also imply convergence in probability.

Now the purpose of giving this one is to tell about laws of large number like you have the central limit theorem we have strong law of large numbers. So let x1, x2 and so on be a sequence of independent and identically distributed random variables with mean mu. Then 1/n sigma Xi=1= 1 to n. This converges to mu almost surely that means in the long run the mean of observations is converging to its actual unknown or original mean.

Similarly, if I consider that is called weak law of large numbers.

(Refer Slide Time: 24:43)



That is under the conditions given above 1/n sigma Xi converges to mu in probability. So the names weak and strong law are simply related to the stronger convergence here and the weaker convergence here, but both are true and actually the generalization of these results are there for example on i .i. d. random variables or the random variable which are independent, but not identically distributed or dependent and so on.

So I have stated it in simplest form. Now what is the practical meaning of this one. Now the practical meaning you can see here that as we consider observations repeatedly then what we are saying is that the average performance or average measure or average yield or average height etcetera it will converge to the true value of the mean. So now these are the useful adds in the sampling because when we do the sampling I will be just be coming through that concept in it little time.

We are considering the sample mean there. So what we are concluding here is the sample mean is almost becoming equal to the population mean Sn becomes large. So that is what allows us to use statistics in practical sphere. Let me come to those concepts now. Let me just to wind up this particular section let me mention few things. We have considered certain continuous distributions.

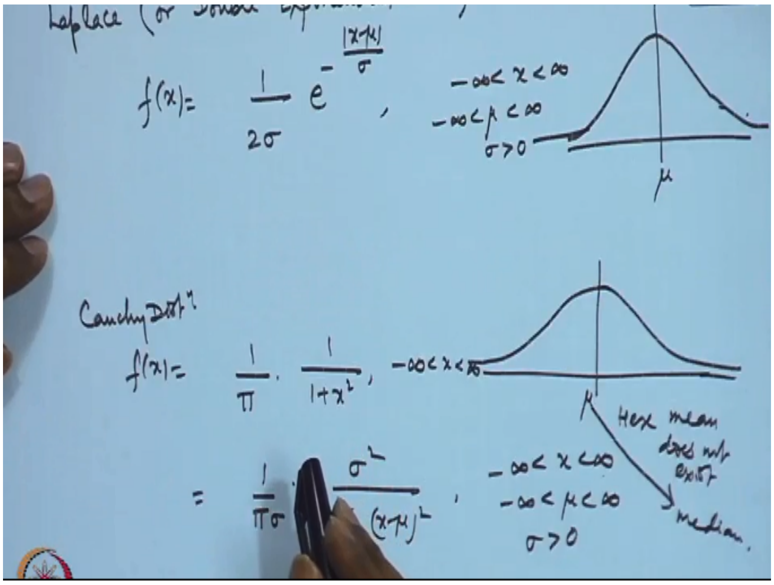
Initially I started with the distributions which are arising as the waiting time of the occurrences. Now waiting time of the occurrences has one important interpretation that is they can be considered distributions which are representing life of system, life of components for example you are considering mechanical system, electrical system, electronic system or any type of organism.

If we are considering the failure first failure or rth failure etcetera, then those distributions can be modeled by exponential gamma distribution etcetera then we also considered in terms of failure rate and then we could look at the distributions which are like Weibull distribution or extreme value of distributions etcetera then I will consider one of the simpler one which is called the uniform distribution and then I have introduced the normal distribution.

I have established now that it is one of the most important distribution in the theory of statistics because of the law of averages that we are saying that if I consider the average of the observation then that is having approximately normal distributions We have also seen the laws of large numbers that is not to say that there are not other important distributions. There are very large number of continuous distribution that one can think of.

For example, in the normal distribution I am considering the tails to go rapidly to 0 because e to the power x-mu square/sigma square when we are considering. So as X goes to +infinity or –infinity the shape of the curve that means it goes to 0 it goes to 0 very rapidly, but there may be distributions where you may not require that. For example, in place of square you may have only linear.

(Refer Slide Time: 28:30)



Then that gives you double exponential distribution or which is also called Laplace Distribution. Laplace are double exponential distribution. So I will just write down the density ½ sigma e to the power x-mu/sigma. So here you can see that tails will be flatter than that of a normal distribution, but of course here mean is again mu median is mu and the peak is also at mu that is the mode is also mu that is called double exponential distribution.

The name double exponential because in the usual exponential distribution you have only one side. Now I have both the sides here so that is why the name double exponential is also coming. You can think of even flatter versions that means in place of exponential function you have only quadric or something like that that means even flatter tails may be there. So for example you have Cauchy distribution.

So let me write the simplest form 1/1+x square or a general version of this could be 1/1+x-mu whole square. So let me put sigma square here and then you may have 1/pi and 1/ sigma will be coming here that is –infinity <x < infinity –infinity< mu < infinity and sigma positive. Now this is applicable to system or you can say where the convergence to +infinity or –infinity is quite slow.

For example, you may consider the decay of radioactivity of say nuclear fallout. So as you know that it is very prolong process and the similar thing in various chemical degradation etcetera you can see the time taken to complete the process may be too large. In fact, Cauchy distribution I gave as an example earlier in fact here the mean itself does not exist this is symmetric. Here mean does not exist. So that means higher order moments will also not exist.

So median is mu, mu is median. Similarly, we have distributions such as beta distribution log normal distribution. There are quite a large number of basically there is a family of distribution that can be described using various functions. So I will stop this discussion here let us move to another concept that is of sampling. That means we move to the use of probability theory for making inferences.

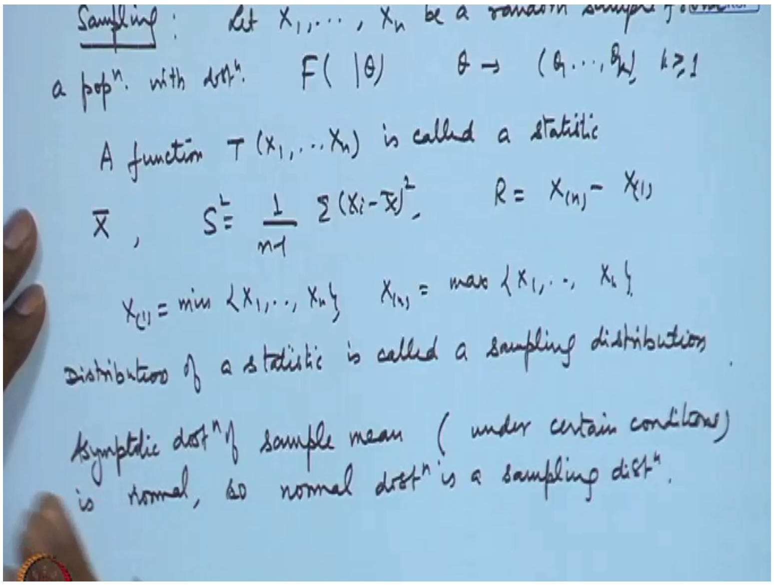
So to start with extremely simply problem. Suppose we want to estimate the average expenditure on the say medical by the people of for example by the people of a state or by the people of a country. Now what one has to do for this study that means one thing is that you take the data from each household of the country, but this not a very useful situation because in a similar way one may be looking at expenditure on say education.

One may be looking at the expenditure on entertainment, one may be looking at the say expenditure on travel and so on. Now if one does a complete enumeration of the population for each of this thing then it is going to be horrendous task and practical studies cannot be done because for example if you are having a large geographical area a country or a state then you will not be able to conduct it in a very reasonable point of time or in a very reasonable timeframe.

And also the resources that will be required will be huge. So what one suggest is that one can use sample. Now the theory of sampling I will be covering at other point of time right now I am introducing from the point of view of distribution that we look at the distribution that arise in the sampling. So suppose we have taken a sample. So let us consider now why the sampling is justified.

Now that is because of the laws of large number and the central limit theorem because in the long run what we are saying is the sample mean acts as the population mean. The distribution of the sample mean after certain normalization converges to a normal distribution and so on. So these are the properties which allows us to use the sampling.

(Refer Slide Time: 34:00)



So let me briefly go to the sampling. So let x1, x2, xn be a random sample from a population with distribution some distribution it will have say capital F and you may have some parameter there F theta the theta maybe vector or scalar theta maybe theta 1, theta 2, theta k where k could be greater than or=1. As you have already seen examples like binomial distribution you have 2 parameters n and p.

In Poisson distribution, you have parameter lambda which is one parameter. In gamma distribution you have parameter r and lambda, in exponential distribution you have parameter lambda and so on. So when I say this is random sample from this basically I am saying each of x1, x2, Xn will have independent and identically distribution F. Now I consider a function say T of x1, x2, Xn this is called a statistic.

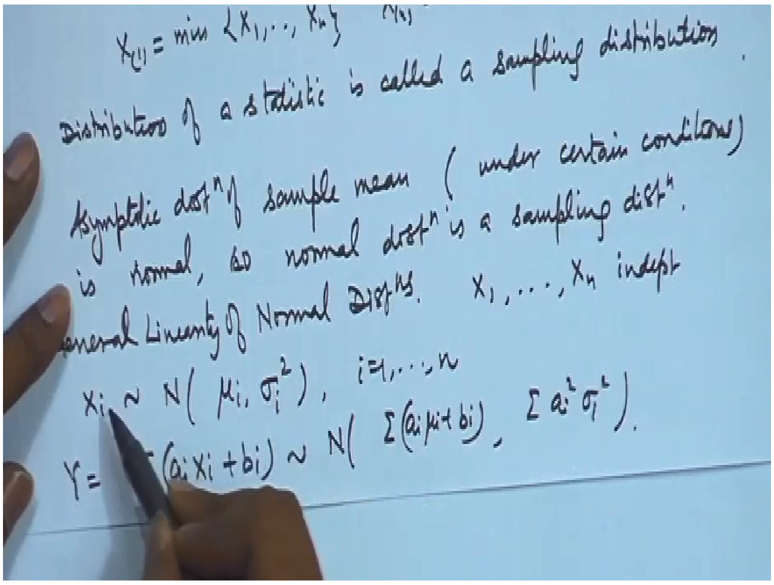
For example, X bar say S square suppose I consider 1/n-1 sigma Xi- X bar whole square. Suppose I consider say range that is maximum-the minimum and so on where this x1 is the minimum of the observations Xn is the maximum of the observations and so on. So these are all example of statistics. Now distribution of a statistic is called a sampling distribution. Now by the central limit theorem.

We can say that normal distribution itself a sampling distribution because I am obtaining it as a limiting distribution or asymptotic distribution of the sample mean. So asymptotic distribution of sample mean of course under certain conditions is normal. So normal distribution is a sampling distribution. Let us also consider say for example I gave you the linearity property of the normal distribution

I also discussed the additive properties of some distributions. For example, if you add certain random variables which are geometric then the sum will become negative binomial where the probability of p of success in individual trial is considered to be constant. We looked at the sum of exponential then that is gamma and so on. A similar property is true for the normal distributions also.

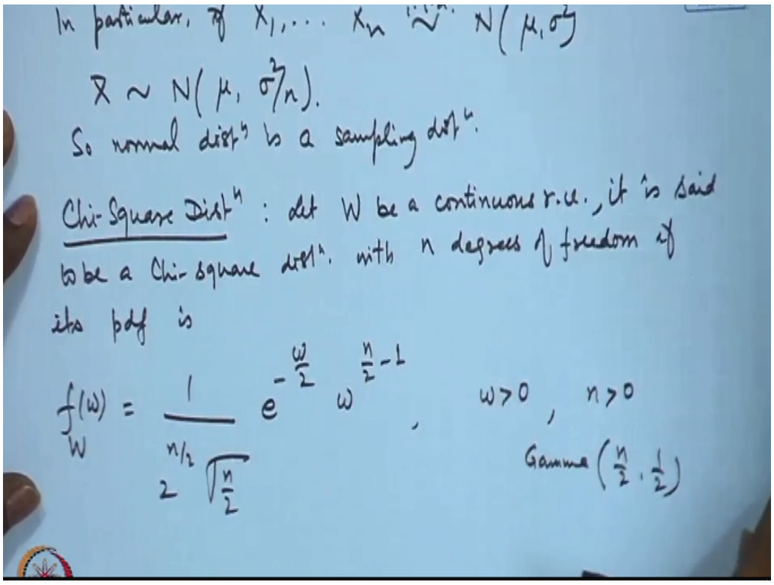
Here this asymptotic distribution is normal, but if original distributions are normal then the sum is also normal.

(Refer Slide Time: 38:08)



So I state the general linearity property of normal distributions. Let us consider say x1, x2, Xn independent where Xi follows normal mu i sigma i square or i=1 to n. Then if I consider say Y=sigma ai Xi+ bi that is a general linear combination of x1x2 Xn then that is following normal with ai mu i+ bi sigma ai square sigma i square.

(Refer Slide Time: 39:25)



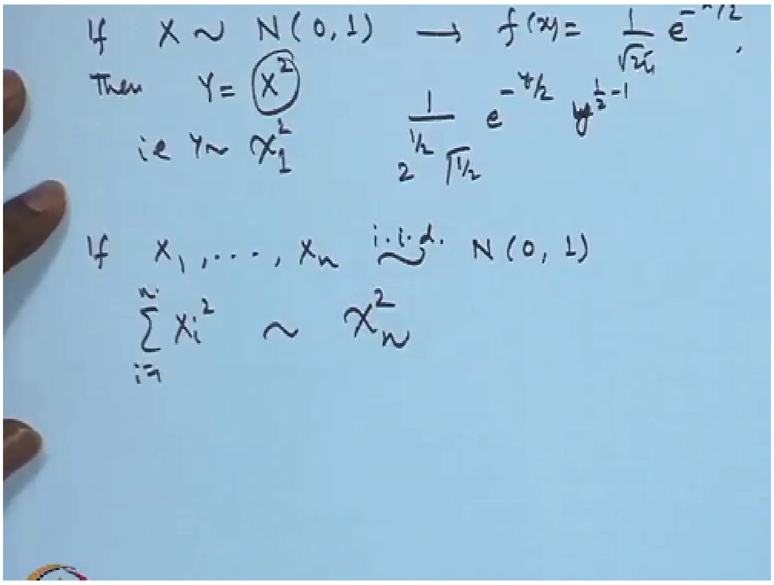
So in particular if I am taking the mean here then that will also have mean which is the mean of these mu1, mu2, mu1 and if I take the variances then this will become sum of the variances and divided by n square. In particular, if I take if x1, x2, xn follow normal mu sigma square that means if they are independent and identically distributed then X bar will follow normal with mean mu and variance sigma square/n.

So normal distribution itself is a sampling distribution in the finite sense also. Here it is asymptotically a sampling distribution, but here it is a fixed sample size also it is a sampling distribution. Now let me introduce some other sampling distributions which arise in the study of distributions of various statistics. So let us consider first which is known as Chi-square distribution.

Let W be a continuous random variable. It is said to have chi square distribution with n degrees of freedom. So the parameter of chi square distribution is actually called degrees of freedom. If its pdf is ½ to the power n/2 gamma n/2 e to the power-w/2 w to the power n/2-1. Here of course n is positive. See if you look at it carefully this is actually nothing, but a gamma distribution. This is gamma distribution with actually n/2 and ½.

R=n/2 and lambda=1/2. So this is actually not a new distribution, but I am introducing it as a separate name chi square distribution because I will show it as a sampling distribution. Notationally, we write it as w follows chi square n. Let us introduce it as a sampling distribution.

(Refer Slide Time: 42:10)

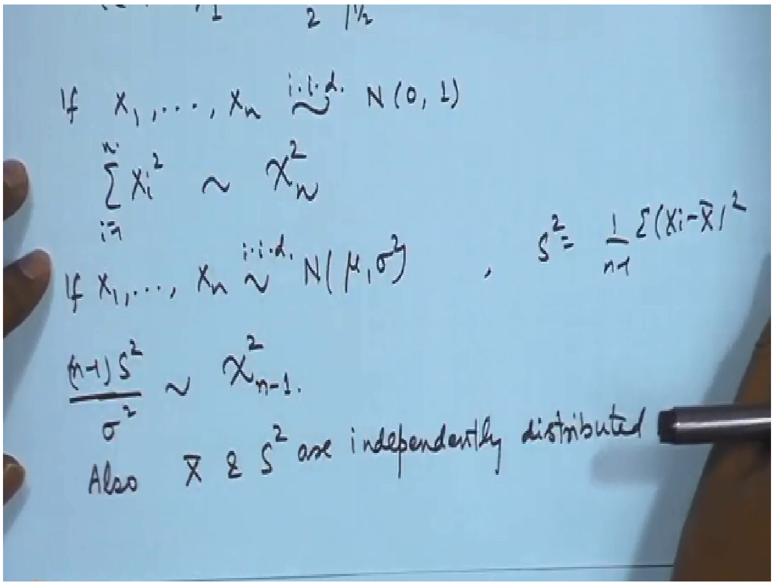


If say X follows normal 0, 1 then if I consider say Y=x square and we can derive a distribution very easily in fact let me just demonstrate it here what is the density function of this that is 1/root 2 pi e to the power –x square/2. So if I consider x square this is a (()) (42:34) transformation so I will get it as ½ to the power 1/2 gamma 1/2 that is 1/root 2 pi e to the power –y/2 Y to the power 1/2-1 which is nothing.

But that is y follows chi square distribution on one degree of freedom which is again gamma and in gamma we know that if lambda is common then additive property is followed. So if I consider if x1, x2, xn are independent and identically distributed normal 0,1 random variables then sigma Xi square that will follow chi square on n degrees of freedom. Now that is one derivation of the chi square distribution as a sampling distribution.

It is arising as the distribution of the sum of squares of n observations from a standard normal distribution, but we also can derive it from a general normal distribution.

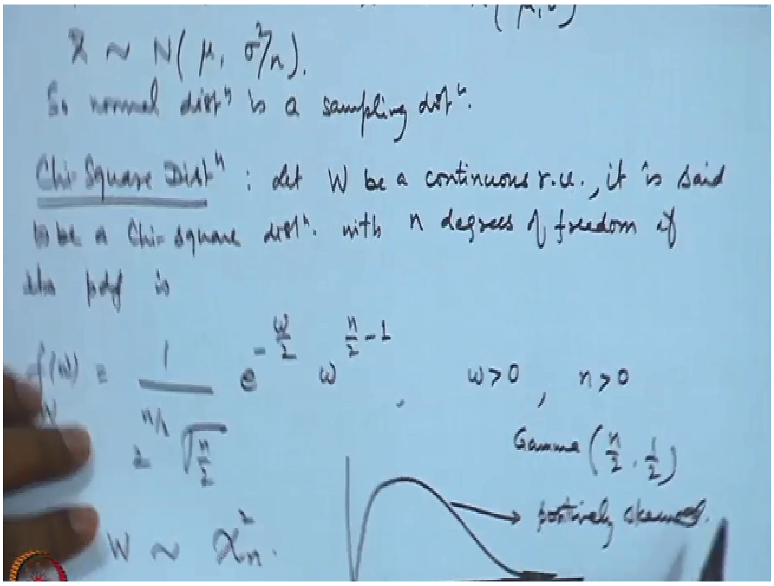
(Refer Slide Time: 43:43)



If x11, x2, xn follow normal mu sigma square. I have introduced S square that is S square was 1/n-1 sigma Xi-X bar square. Then n-1 S square/sigma square that will follow chi square distribution on n-1 degrees of freedom. So this shows it has a sampling distribution of the sample variance also X bar and S square independently distributed. For more details about the derivation of this and this results etcetera you may look at the (()) (44:38) lecture on probability and statistics.

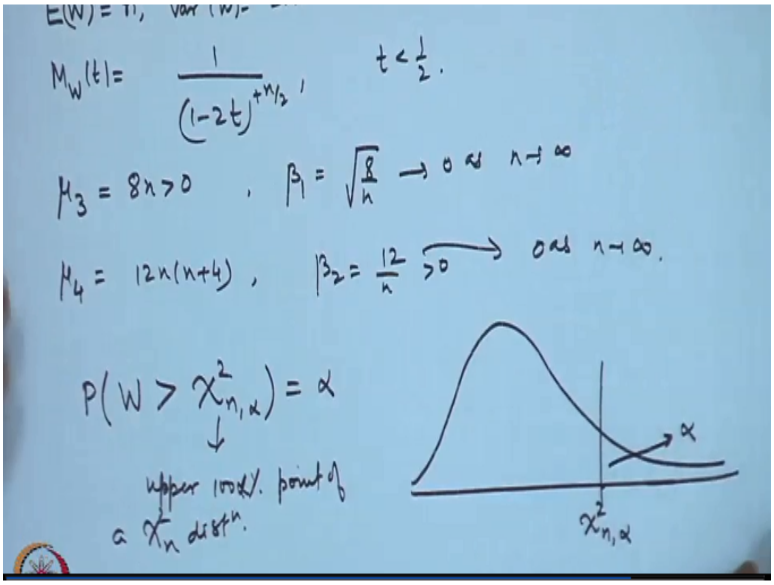
And also you can look at the books which I have mentioned in the references. So I will not get into too much details of each of this distribution. Let us just look at the properties of this thing.

(Refer Slide Time: 44:56)



If we look at the form of the density function this function is actually if we plot it of course it will depend upon n, but it is positive so it will be usually positively skewed. In fact, let us look at the coefficient etcetera.

(Refer Slide Time: 45:21)



Let me write expectation of chi square is= to the degrees of freedom, variance of chi square=twice the degrees of freedom. It is a moment generating function is=1/1-12 t to the power+ n/2 for t<1/2. If we look at the measures of skewness and Kurtosis it turns out that the third central moment is=8n which is of course positive, but if I look at say measure of skewness then you can see it is=root 8/n.

So of course this goes to 0 as n becomes large. That means it will converge to symmetry if n is large which is okay because I am obtaining chi square as the distribution of a sum here. If you look at the distribution of the sum, then by central limit theorem as n becomes large. The distribution of sigma Xi square-n/root 2n that will converge to normal 0, 1. You can actually write here suppose I am calling it as a U then u-n/root 2 n that will converge to normal 0,1 as n tends to infinity.

So therefore you can also look at the measure of kurtosis mu 4 is 12 n*n+4. So beta 2= 12/n which is positive, but this goes to 0 as N tends to infinity. Then regarding the calculation of the probabilities like in the normal probabilities all the probabilities we were able to calculate through the standard normal probability curve which are tabulated. Now for the chi square distribution if you see you are considering again if you look at the probabilities related to this.

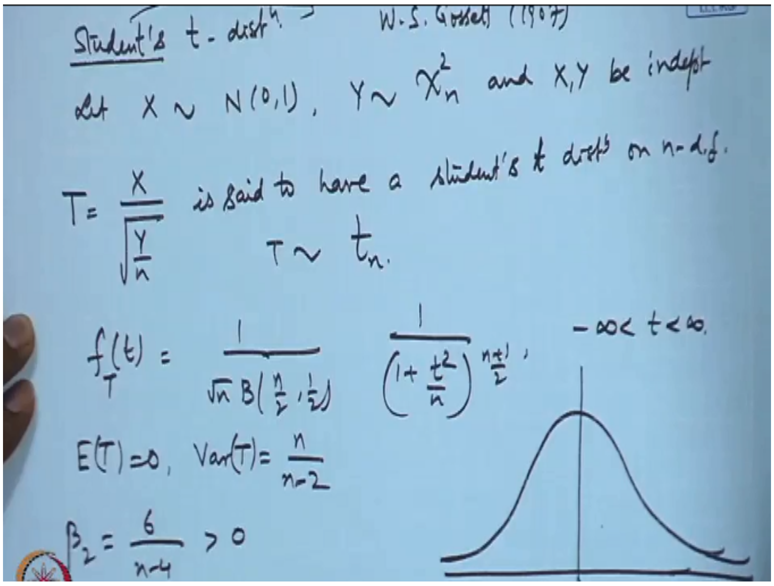
This will give you a incomplete gamma function. Now for various values of n the tables of chi square cdf are tabulated, but that will be too complicated. So to consolidate or you can say to make it in a compact form what is tabulated is of this form. If this probability is alpha, then this point is called chi square n alpha. That means probability of W > chi square n alpha=alpha.

So for different values of n and alpha this percentile points of chi square distribution this is called upper 100 alpha percent point of a chi square n distribution. So the tables of this are given in almost all the statistical books and tables this is tabulated. So for example you can see if I have n=10 and alpha= 0. 1 then the value is given to be 4.865 and so on. Here actually let me see this is 0.05 is actually 1-alpha so that will be= 0.95 so the point is for example 3.94.

If I take 0.9 then it is 4.865 and so on. So the tables of chi square and alpha are given for different values of n and alpha. And as I have mentioned that Sn becomes very large it is not required because then the distribution of U-n/root 2 and can be approximated by normal distribution so those tables are not given. Generally, in the book they tabulate up to n=30 or sometimes up to 60 or something like that.

So we have shown chi square as a sampling distribution in sampling from a normal distribution. Let us look at some further sampling distributions.

(Refer Slide Time: 50:00)

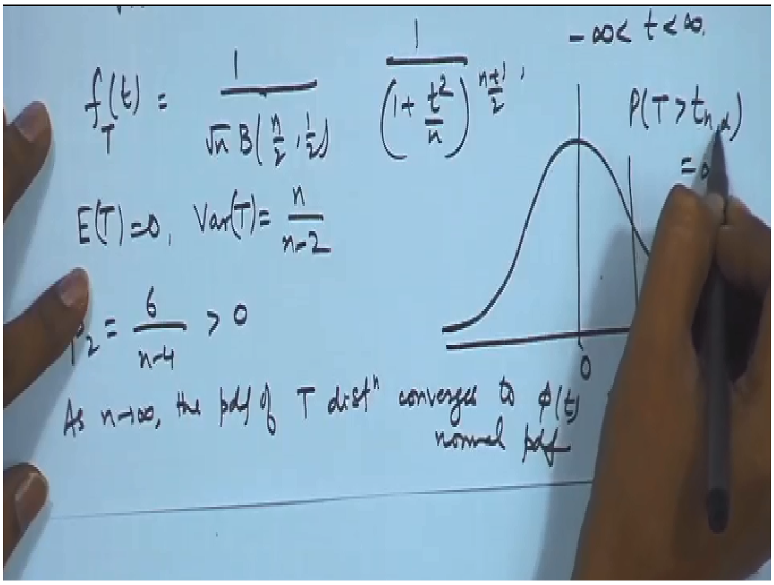


The next one we call Student's t distribution. Let X follow normal 0,1 and Y follow chi square n and Xy be independent. Then if I define t=x/root y/n then this is said to have a Student's t distribution on n degrees of freedom and we write here t follows t on n degrees of freedom. This name is student actually because of statistician W. S. Gosset who gave it in 1907, but he (()) (50:59) published under the student that is why it is called a Student's t distribution.

One can easily derive the density function of t that is 1/root let us give the exact form here that is 1/root n beta n/2 1/2 1/1+ t square/n to the power n+1/2. As you can see this is also symmetric distribution around 0 and it has some important things for example expectation will be 0. The variance of T is n/n-2 of course you can see that this will converge to 1 as n tends to infinity.

And if we look at mu 4 and say beta 2 here beta 2 for this distribution is actually=6/n-4 which is positive. Actually this distribution is closely resembling a normal distribution.

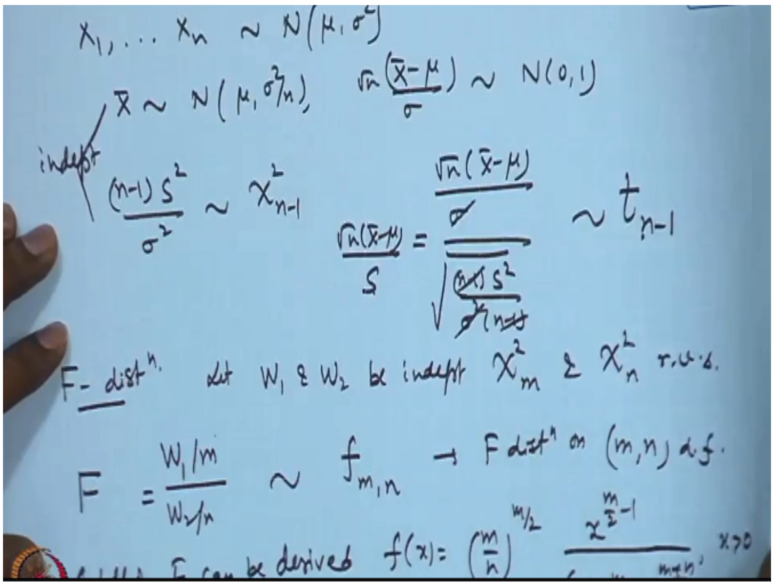
(Refer Slide Time: 52:23)



And in fact you can prove as n tends to infinity the pdf of T distribution converges to phi t that is the standard normal. Now usually for n > or=30 the approximation is quite good and that is why generally the tables of t distribution will be tabulated up to n=30 only. And because of the symmetry if we consider the point here t n alpha that is this probability is=alpha probability of t> tn alpha=alpha.

So for different values of n and alpha t and alpha values are tabulated. So this is called the upper 100 alpha% point of the t distribution. Now to look at it as more as a sampling distribution.

(Refer Slide Time: 53:30)



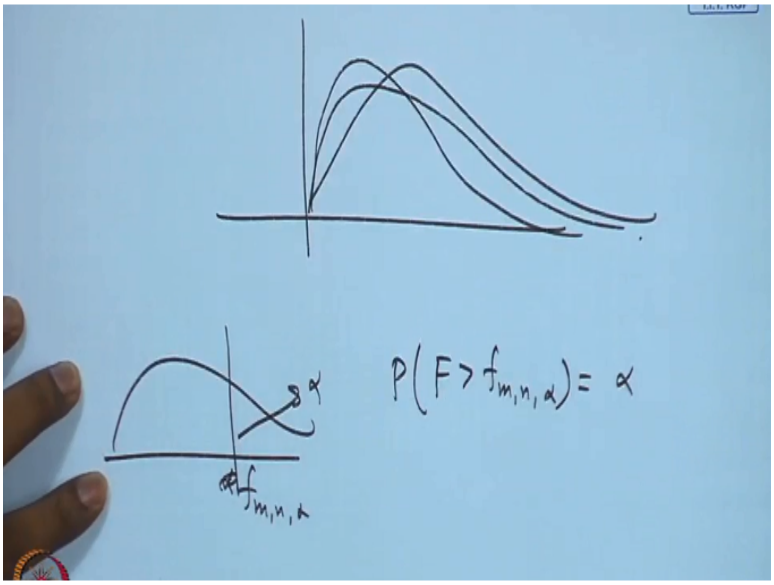
If I consider say x1, x2, xn follow normal mu sigma square and if I consider say X bar then x bar follows normal mu sigma square/n. Therefore, if I consider x bar-mu/sigma *root n that will follow normal 0,1. At the same time if I look at n-1 S square/sigma square that follows chi square on n-1 degrees of freedom and these 2 are independent as I mentioned earlier. So if I consider root and X bar-mu/sigma/ root of n-1 S square/sigma square* degrees of freedom here. Then this will follow t distribution on n-1 degrees of freedom.

But if you simplify this then you get this as root n X bar-mu/S. So here you differentiate this is root n X bar-mu/sigma that is normal 0,1. Root n x bar –mu/S is t distribution on n-1 degrees of freedom. So this is a sampling distribution. To the end, I will define one more distribution that is called F distribution. Let say W1, W2 be independent chi square say m and chi square n random variables.

Then let us define W1/m/w2/n let me call it F. Then this is said to follow F distribution on m n degrees of freedom that is F distribution on m n degrees of freedom. One is for the numerator chi square variable and one degree of freedom for the denominator chi square variable. One can again write down the density function pdf of this F can be written as so that is let us say fx that is= m/n to the power m/2 X to the power m/2-1, 1+m, nx to the power m+ n/2 where x is positive.

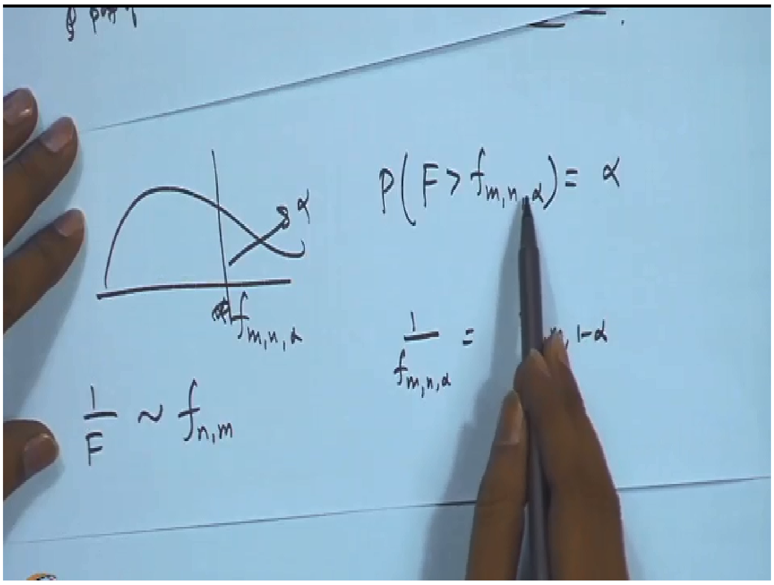
So this is the pdf of this. One can derive it using the usual distribution theory and you can easily see that this is a skewed distribution.

(Refer Slide Time: 56:10)



You will have of course it will vary with the value of m and n but various forms will be skewed here. These are the things and just to give you if I consider say one of them and I consider this alpha this probability was alpha and this is called f, m, n alpha that is the probability of F>f, m, n alpha this is=alpha.

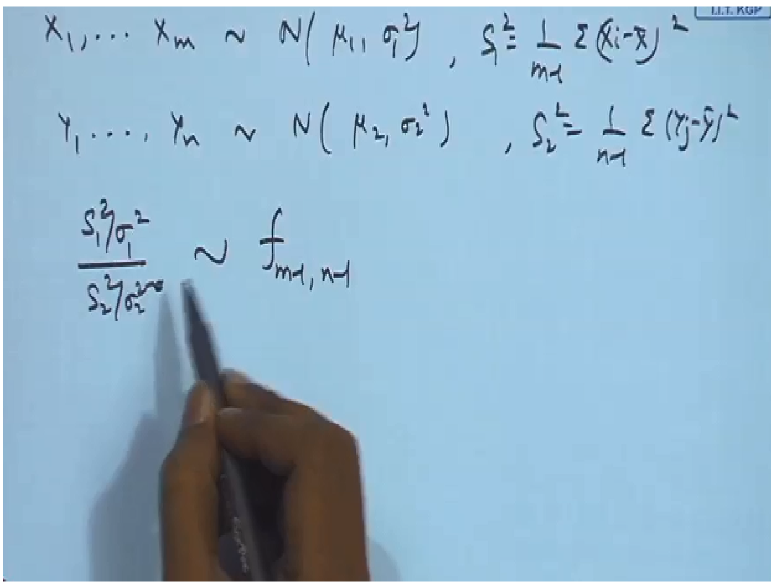
(Refer Slide Time: 56:55)



Another thing that you can notice here is that I consider 1/f here I/f is also f that will be f on nm degrees of freedom. Therefore, if I find out the relation 1/f mn alpha that is=fnm 1-alpha this relation is there. So for different values of m and n and alpha the values of f, m, n alpha has been tabulated. So one can look at, but since this is 3 dimensional things. Therefore, only for selected values of alpha you can find the tables of the percentile points of f distribution.

So I have given important distributions as the sampling distribution.

(Refer Slide Time: 57:44)



Just to end here let me give you here suppose I consider a random sample from normal mu 1 sigma 1 square and I define S1 square as 1/m-1 sigma Xi-X bar square. Similarly, I consider another random sample from normal mu2 sigma 2 square whereas 2 square is 1/n-1 sigma y j-y bar square then if I consider S1 square/ sigma 1 square/S2 square/sigma 2 square that will have f distribution on m-1, n-1 degrees of freedom. So this is also a sampling distribution.

I have derived normal distribution, chi square distribution, t and f distribution as sampling distributions when we are sampling from the normal populations, but normal distribution itself is a sampling distribution in a more general form, more general sense because it is also a sampling distribution of the sample mean from any population with of course finite variance provided the random variables are i. i. d. So you consider a sample mean.

And the conditions have been relaxed also that means that identical thing can be relaxed or independent thing can be relaxed and therefore in a more general sense in a normal distribution is a sampling distribution. These sampling distributions are very useful when we will do the inference that means we will consider confidence intervals for the parameters of mean and variances when we will consider the testing of hypothesis for the mean and variances etcetera.

So in the next module of this course, we will be covering various aspects of this. You can look at the problem sets on this module of probability and distribution theory which is available on the website so that will be very useful to look at the problem for this. With this, I complete this section.