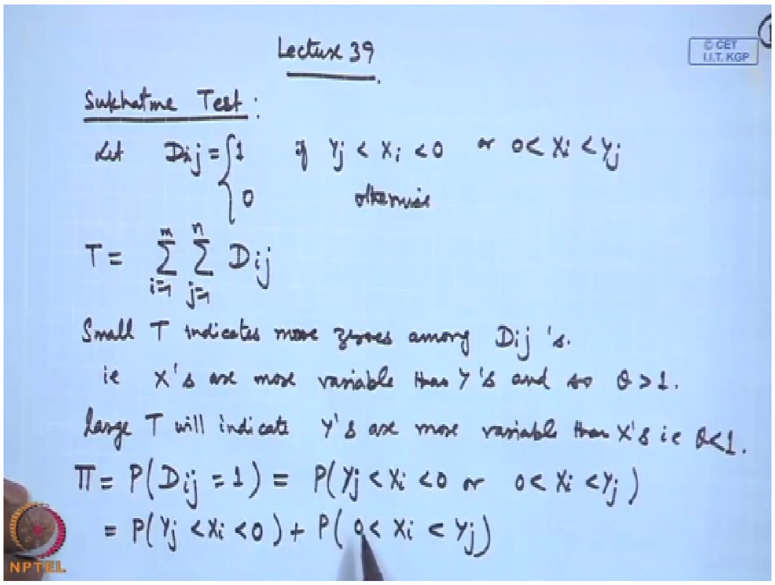
Statistical Methods for Scientists and Engineers
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology – Kharagpur

Lecture–39
Nonparametric Methods – XII

(Refer Slide Time: 00:19)

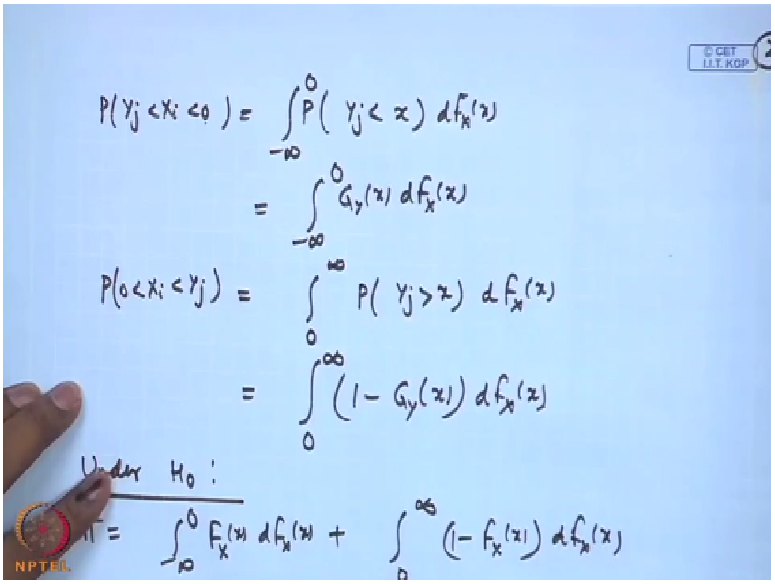


So, we were discussing some tests for the 2 sample scale problem. Let me consider now Sukhatme test. So, for this Sukhatme test let me define $D_{ij}=1$ if $Y_j<X_i<0$ or $0<X_i<Y_j$ and it is = 0 otherwise. You can see that there is a little bit modification here. So, we are on one side of 0. Like if $Y_j<X_i<0$ or if $Y_j>X_i>0$, in both the cases Y_j is farther away from 0 than the X_i. So, that is why you can see that this test is different than the Mann-Whitney or this one.

Because it is not simply based on the order of X_i Y_j but also the positioning from the 0 and we then define the statistics as double summation D_{ij}, $i=1$ to m, $j=1$ to n. So, what it will mean a small t. a small t indicates more 0s among D_{ij}. That means X's are more variable than Y's and so theta will be > 1 and large T will indicate, that is Y's are more variable than X's, that is theta is < 1.

So, you can see that this is a very, very natural kind of definition that has been taken by Sukhatme. However, let us show the calculations for this. Let us consider say pi that is the probability of Dij=1. In terms of that, we will actually derive the mean and variance etc. of this statistic. So, this is = probability of Yj<Xi<0 or 0<Xi<Yj. So, that is equal to probability of Yj<Xi<0 plus probability of 0<Xi<Yj because these are 2 disjointed event, so we can write it as a sum of the probabilities.

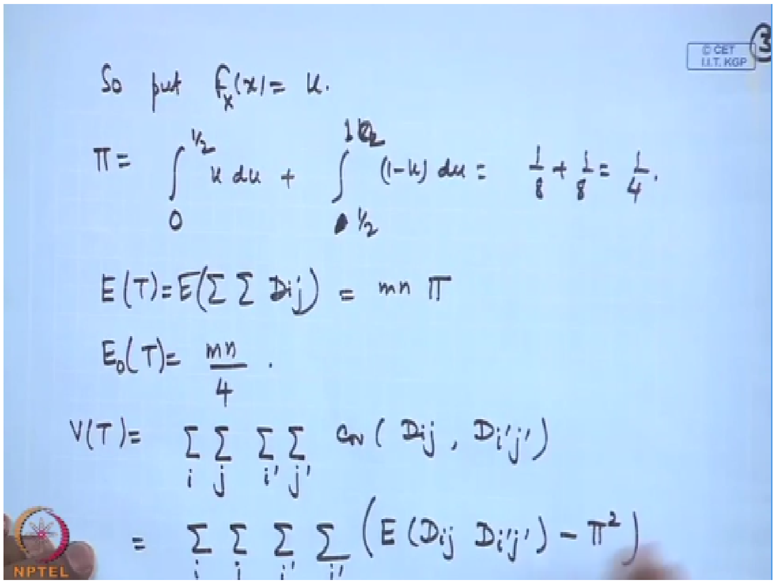
(Refer Slide Time: 03:45)



Now, using the conditioning argument on X, so we can express it like; for example, if I consider Yj<Xi<0. So, this we can consider as the conditioning on Xi, Yj<x*the distribution of x but x is up to 0 only. So, it will be from -infinity to 0, but this can be written as simply Gyx dFf. Similarly, if I consider probability of 0<Xi<Yj, then this is = 0 to infinity probability of Yj>X dFx. Then, this becomes 1- (()) (04:43) of Y, okay.

So, what we do let us consider under H0. So, under H0 pi will become =-infinity to 0. So, this term plus this term, I have written the expressions here Fx dFx+0 to infinity 1-Fx dFx.

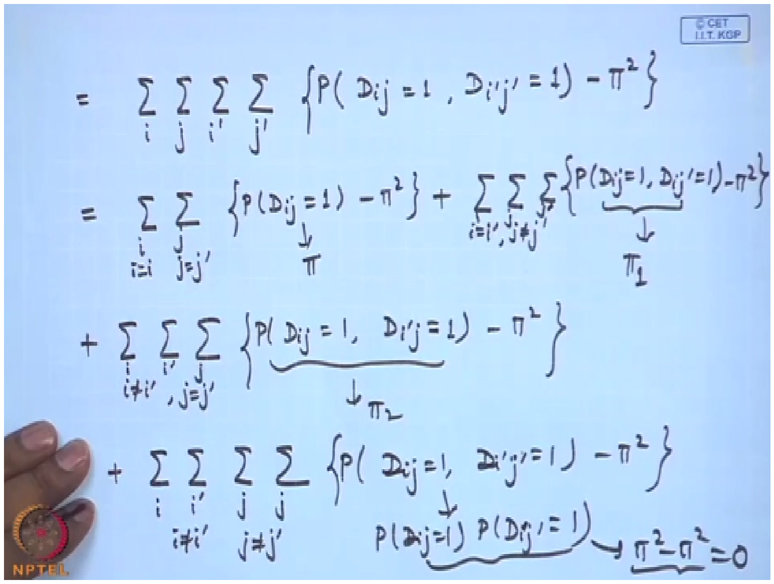
(Refer Slide Time: 05:30)



So, let us put say Fx=U. Then, pi= at -infinity this is 0, at 0 this will become 1/2. So, this is becoming udu+0 to 1/2. Well actually second one will become 1/2 to 1 and the this is 1-u du. So, both are actually 1/8+1/8=1/4. So, under the null hypothesis, the probability that Dij=1 is actually becoming = 1/4. So, if I consider the expectation of T under the null hypothesis; of course this is equal to double summation Dij, so that is = mn pi, but under the null hypothesis, this is simply becoming mn/4.

So, you can see that actually the symmetry will come around this value. Let us look at similarly the variance term here. So, variance T is =, we can write the general term covariance Dij Di prime j prime where the sum is over all i's and j's here. So, this is then = expectation of Dij Di prime j prime-pi square, i j i prime j prime. So, this will be one only when Dij and Di prime j prime both are one. In all other cases, this will be = 0. So, this is simply equal to the probability of Dij=1 and Di prime j prime=1.

(Refer Slide Time: 08:03)



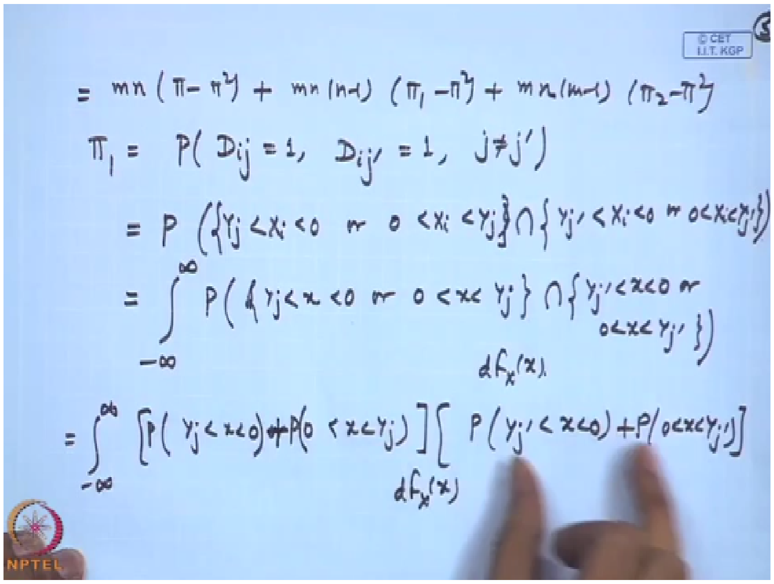
So, we can express it as double summation, quarterpole summation probability of Dij=1, Di prime j prime=1-pi square. Another thing that we observe since this Dij is based on Xi and Yj, therefore Di prime j prime will be based on Xi prime Yj prime. Since, the random samples are taken, therefore these will be totally independent. So, this can be written separately here. So, let me express it in a full form here.

There will be all cases; i can be = i prime, j can be = j prime and so on. So, let us consider all the cases here. So, one case is when i=i prime, j=j prime, then this will become simply double summation. So, this term then can be written as if i=i prime, j=j prime, then this term will become probability of Dij=1-pi square. Now, let us consider other case> One case will be when i=i prime, j != j prime, so in that case this will become triple summation, that is i j j prime. So, this is then = probability of Dij=1, Dij prime=1-pi square.

Now, this we have to calculate separately, so let me give a notation for this. This will become pi1. So, this is again pi here, this is pi1. Then, there will be another case. The other case will be when i != i prime but j=j prime. So, this is i i prime j. So, this is probability of Dij=1, probability of Dij prime j=1-pi square. This one let us name it as pi2 and then there will be choice when all of them are different, i.e., i i prime j j prime, here i != i prime, j != j prime.

This is probability Dij=1, Di prime j prime=1-pi square. In this case, this is actually = Dij=1, Di prime j prime=1, why because Xi Xi prime Yj Yj prime they are all independent. So, Dij will become independent of Dij prime Di prime j prime. So, then this is nothing but pi square, so pi square-pi square, that is it becoming = 0. So, this term vanishes. We are left with the this, this and this term.

(Refer Slide Time: 11:53)

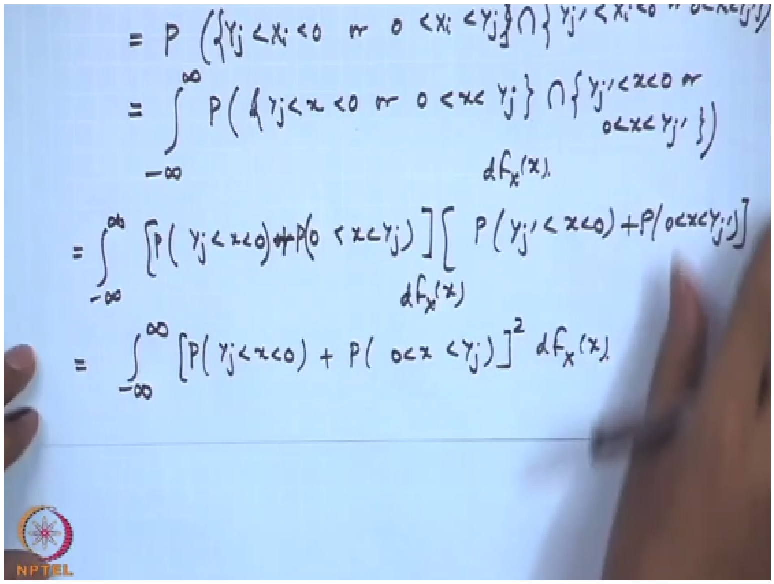


So, in terms of the notations pi, pi square pi1, pi2, etc. we can express it as, so then this we write as mn pi-pi square+mn*n-1 pi1-pi square+mn*m-1 pi2-pi square. Let us look at the counting of these terms. Here we are taking over all i j. So, there will be mn terms. In the second one, here I am taking i = i prime but j != j prime. So, these are n*n-1 and i's are m, so m*n*n-1 pi1-pi square. Then, in the third one j = j prime, so that is n terms and then i != i prime, that is m*n-1 term.

So, it becomes mn*m-1 pi2-pi square and the last one which is actually m*n-1 n*m-1 but actually this is becoming 0 because this is pi square-pi square. So, we are left with this much only. Now, let us consider the expressions for these quantities under general and null hypothesis. So, pi1 let us look at for example, so that is equal Dij=1, Dij prime=1 where j != j prime. So, that is = probability of Yj<Xi<0 or 0<Xi<Yj, that is Dij=1 and we will be taking intersection with the event Yj prime<Xi<0 or 0<Xi<Yj prime.

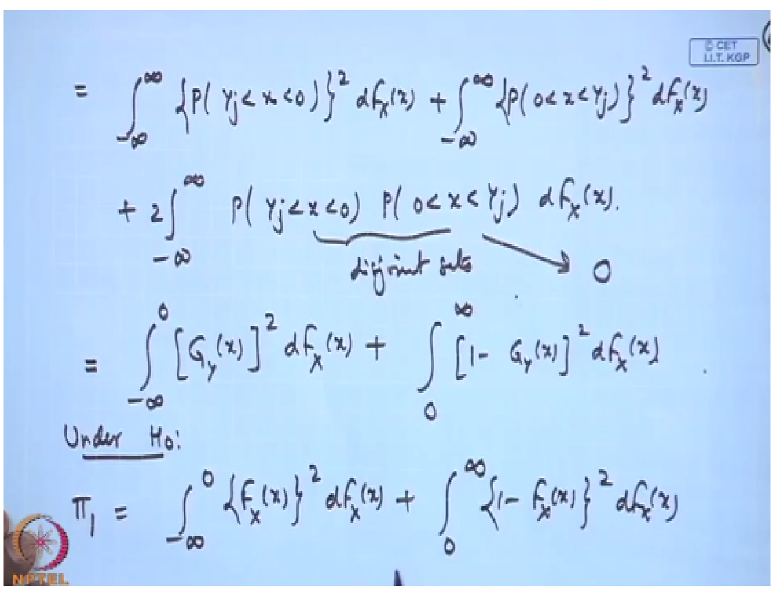
So, here you can notice here Xi is fixed here, so we can do the conditioning on that. So, this becomes probability of Yj<X<0 or 0<X<Yj intersection Yj prime<X<0 or 0<X<Yj prime dFx. So, that is = Yj and Yj prime are independent. Therefore, these 2 probabilities can be written as a product here, Yj<X<0 or 0<X<Yj. In fact, you can write it as sum here into probability of Yj prime<X<0+probability of 0<X<Yj prime dfx. Since, Yj and Yj prime have the same distribution, therefore this quantity will be same as this.

(Refer Slide Time: 16:00)



So, we can write it as the square term here.

(Refer Slide Time: 16:04)

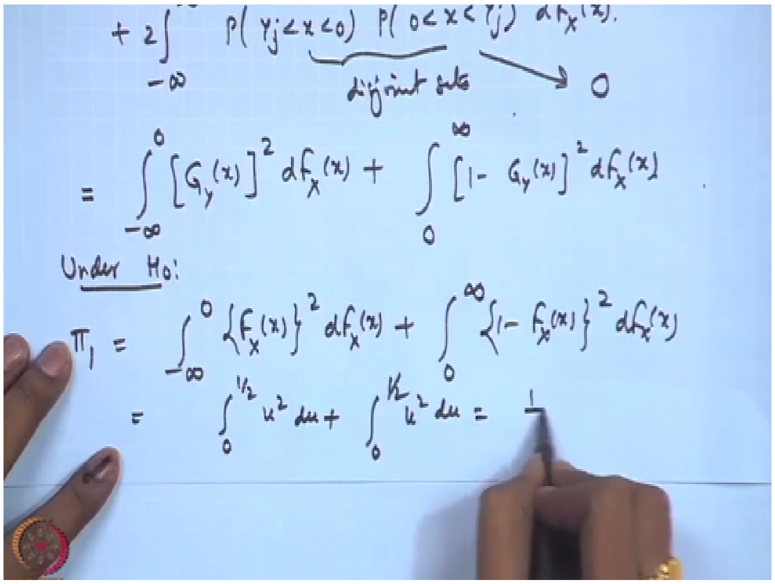


So, that is = probability of Yj<X<0 square dfx-infinity to infinity probability of 0<X<Yj square dfx+twice-infinity to infinity probability of Yj<X<0*probability 0<X<Yj dfx. Now, you look at

this one. See, this is saying Yj<X< and of course X<0. This is Yj>X. Now, under the same distribution of f, that means same distribution of X we are having 2 disjoint sets here. So, these 2 are disjoint sets. If these are 2 disjoint sets, so therefore this would be = 0.

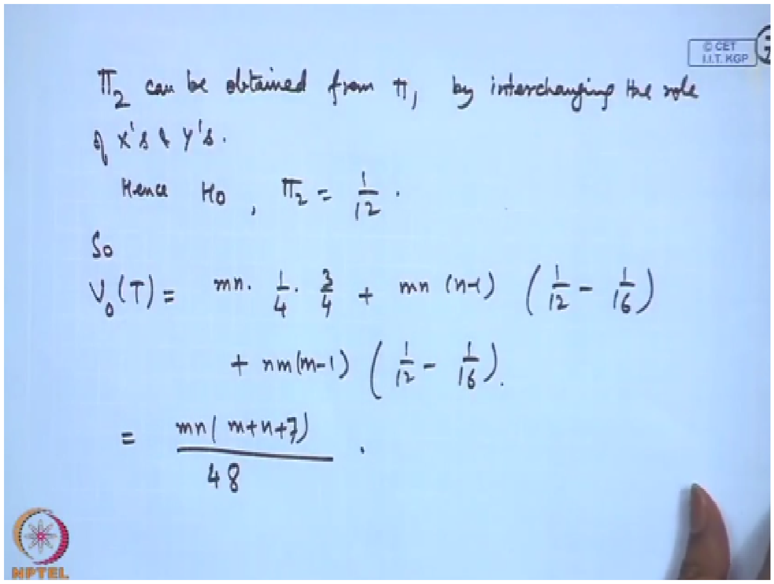
Because these 2 events cannot occur together, like if I have to put integral, then for this one it is –infinity to 0, for this one it has to be 0 to infinity. So, both of them cannot occur simultaneously. So, this term will become simply = 0. So, this one now it is = -infinity to 0. This is the CDF of Y square*dfx and the second one is then 0 to infinity 1-CDF of X dfx. So, this is the general expression now we have obtained for pi1. Now, under the special case when F and G are same, then this become = -infinity to 0 Fx square dfx+0 to infinity 1-Fx square dfx.

(Refer Slide Time: 18:34)



So, when you put Fx = u, then this can be written as say 0 to 1/2 u square du+0 to ½, 1-Fx you can put here. So, u square du. So, this is then becoming = 1/12. This will become 1/3 here, so u cube/3. So, when you put 2 here 1/24+1/24 is = 1. In a similar way, if you look at the expression for pi2. In pi2, what is happening, the roles of Xi's and Yj's you can interchange. So, I will not write the expression for that now in full detail.

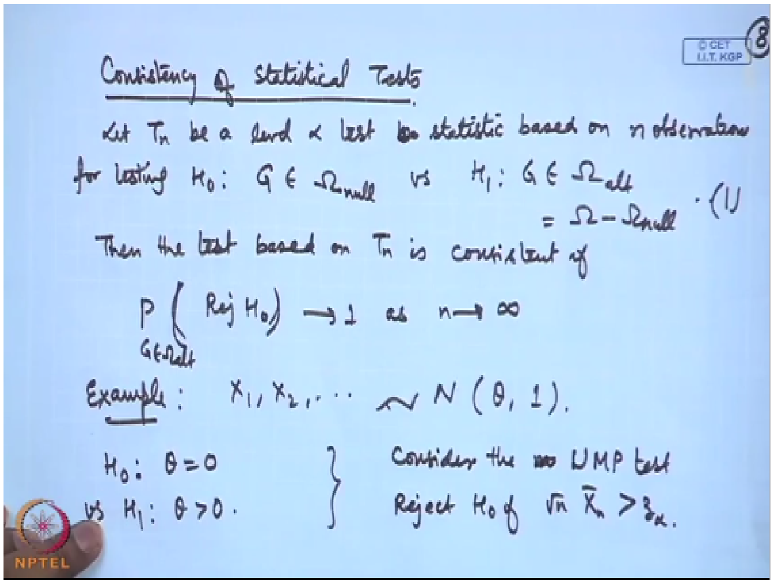
(Refer Slide Time: 19:32)



Pi2 can be obtained from pi1 by interchanging the role of X's and Y's. So, under H0, pi2 will then become = 1/12. So, variance of T under H0, that is a mn 1/4 3/4, that is pi-pi square+mn*n-1 1/12-1/16+nm*nm-1 1/12-1/16. Of course, you can simplify this. It becomes mn*m+n+7/48. So, the null distribution for Sukhatme test statistic has been obtained here. The expectation T is mn/4 and the variance of T under the null hypothesis is obtained.

So, this Sukhatme test statistic can also be used for testing the 2 sample scale problem. As I mentioned here that the small t indicates that theta is > one, a large T will indicate theta < 1. So, this can be used and also we have obtained the null distribution of that. As I mentioned now, I am discussing the large sample property of the tests for the nonparametric situations.

(Refer Slide Time: 21:33)

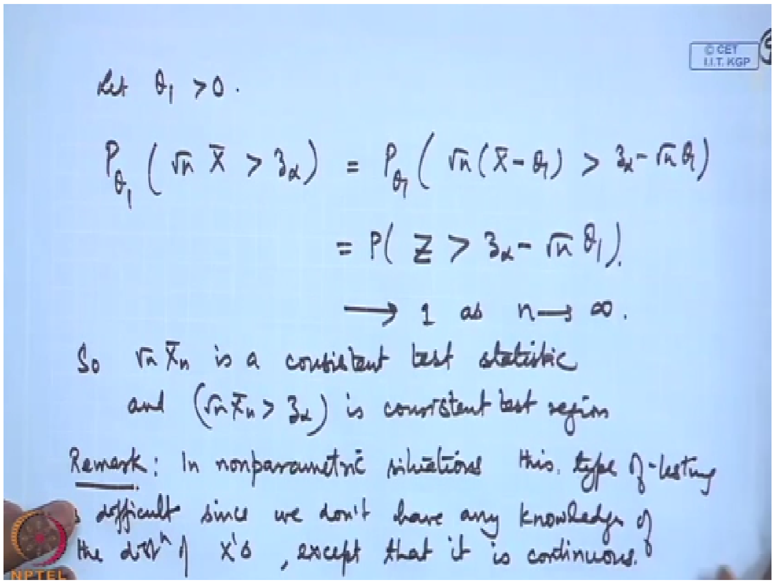


So, this property is called the consistency of statistical tests. So, you can actually think of the consistency property of the estimator. In the point estimation, how do we define the consistency. We consider the probability that the estimator approaches the true value of the parameter converges to 1. In the case of testing, we can consider the power function. If the function approaches 1, that means the power becomes large and large as the sample size increases, then we can consider it as a consistent test.

So, it is similar to the consistency of the estimator in the sense that here the power will increase here. So, let me define here. Let Tn be a level alpha test statistic based on n observations. For testing, H0 say G belongs to omega null versus H1 G belongs to omega alternative which is actually = omega-omega null, okay. This is my testing problem here. Then, the test based on Tn is consistent if probability say G belongs to omega alternative rejecting H0. This goes to 1 as n tends to infinity.

So, it is same the power of the test going to 1. Let me consider a simple application of this. Let us consider say observations from a normal distribution with mean theta and variance unity and we are considering the standard test for the hypothesis testing problem, theta=0 against theta>0. So, consider the most powerful, that is we call it UMP test here, uniformly most powerful test that is reject H0 if the root n Xn bar>Z alpha, that is a level alpha test here.

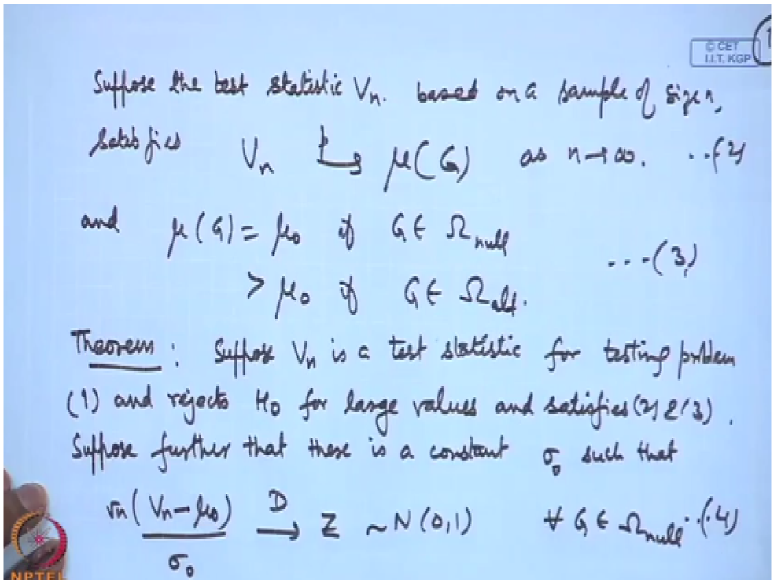
(Refer Slide Time: 24:47)



So, we consider at a point theta 1. So, let us take say theta 1>0, what is the power at this point, root n X bar>Z alpha, that is = probability of root n X bar-theta 1>Z alpha-root n theta 1. When theta=theta 1, this will have the standard normal distribution. So, this is probability of Z>Z alpha-root n theta 1. Now, as n tends to infinity what happens here, here theta 1 is positive, therefore this value will go to -infinity.

So, Z > –infinity this will go to 1 as n tends to infinity. So, root n Xn bar is a consistent test statistic and this test is actually consistent that is root n Xn bar>Z alpha, this is consistent test region, that is consistent critical region here, okay. In the nonparametric situation, directly specifying this kind of thing is difficult here, because we do not have the knowledge of the probability distribution here. So, we cannot write down this kind of statement. So, we define in a different way. In nonparametric situations, this type of testing is difficult since we do not have any knowledge of the distribution of X's, except that it is continuous.

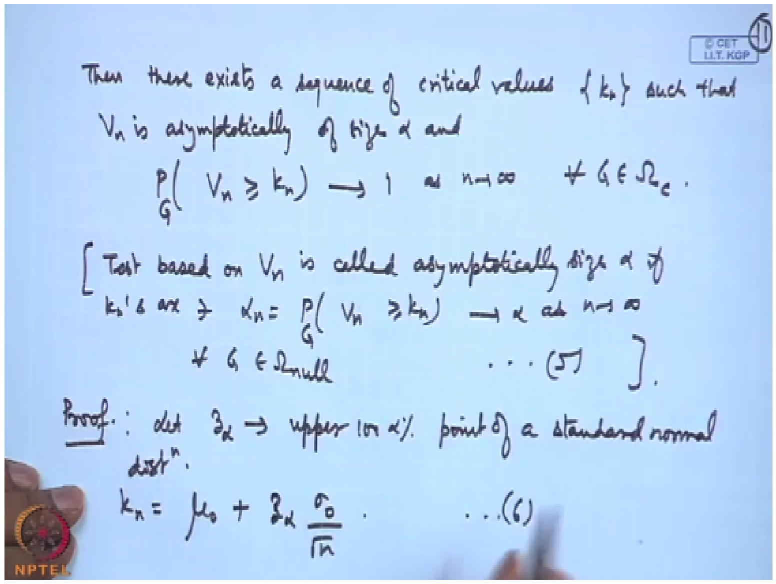
(Refer Slide Time: 27:30)



Let us consider suppose the test is statistic Vn based on a sample of size n satisfies Vn converges to mu of G in probability as n tends to infinity and the function mu satisfies mu G=mu0 if G belongs to omega null and it is > mu0 if G belongs to omega alternative. Let me give this numbering here. So, we have the following result then regarding the consistency here. Suppose Vn is a test statistic for the situation 1, 1 is the hypothesis testing problem, G belongs to omega null against H1, G belongs to omega alternative.

So, suppose Vn is a test statistic for testing problem 1 and rejects H0 for large values and satisfies 2 and 3, that means it is consistent that is convergence and probability to mu G function and this mu G itself satisfies that under the null hypothesis, it is equal to some fixed value mu0 and under the alternative hypothesis it is > mu0. So, basically we are trying to put it in the framework of a parametric testing problem here. Suppose further that there is a constant sigma0 such that root n Vn-mu0/sigma0 converges in distribution to standard normal distribution for all that is under the null hypothesis.

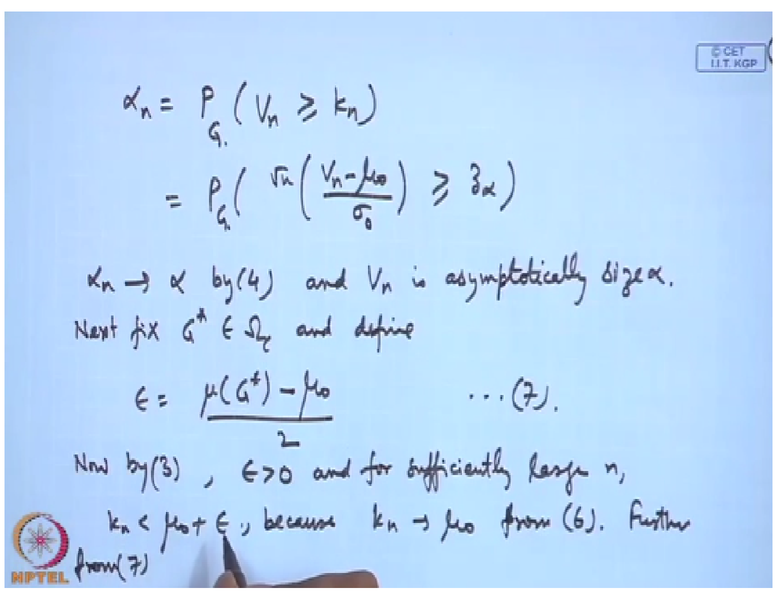
(Refer Slide Time: 30:32)



Then, there exists a sequence of critical values Kn such that Vn is asymptotically of size alpha and probability of Vn >=K under G is going to 1 as n tends to infinity for all G in the alternative, okay. Asymptotically size alpha let me define here, test based on Vn is called asymptotically size alpha if Kn's are such that alpha n is = probability Vn >=Kn goes to alpha as n tends to infinity for all G belonging to omega null.

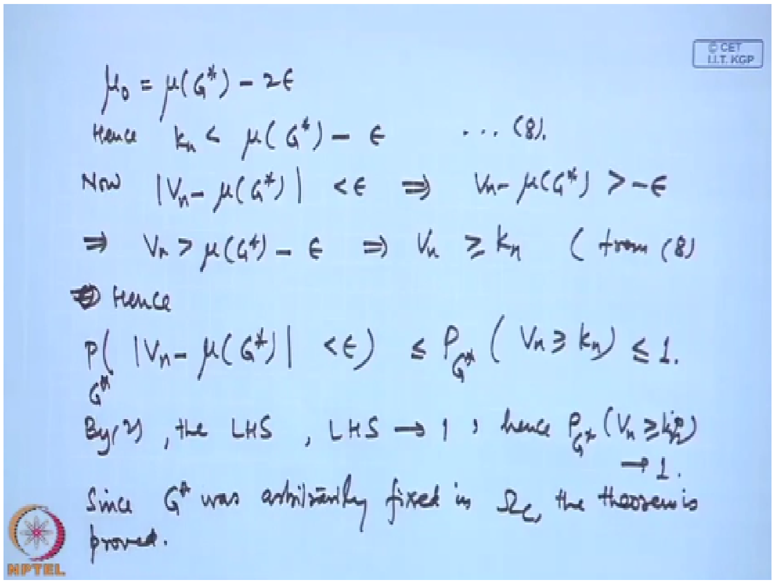
Let me prove this. So, let Z alpha with the upper hundred alpha percent point of the standard normal distribution. So, let us define say Kn=mu0+Z alpha sigma0/root n.

(Refer Slide Time: 32:55)



So, let us consider say alpha n=probability of Vn >=Kn, so that is = probability of root n Vn-mu0/sigma0 >=Z alpha. So, alpha n goes to alpha/4, we have assumed here the asymptotic normality here and Vn is asymptotically size alpha. Now, we fix here G* belonging to omega C and define epsilon=mu G*-mu0/2. So, by 3, epsilon will be > 0 and for sufficiently large normal, Kn<mu0+epsilon since Kn goes to mu0 from 6.

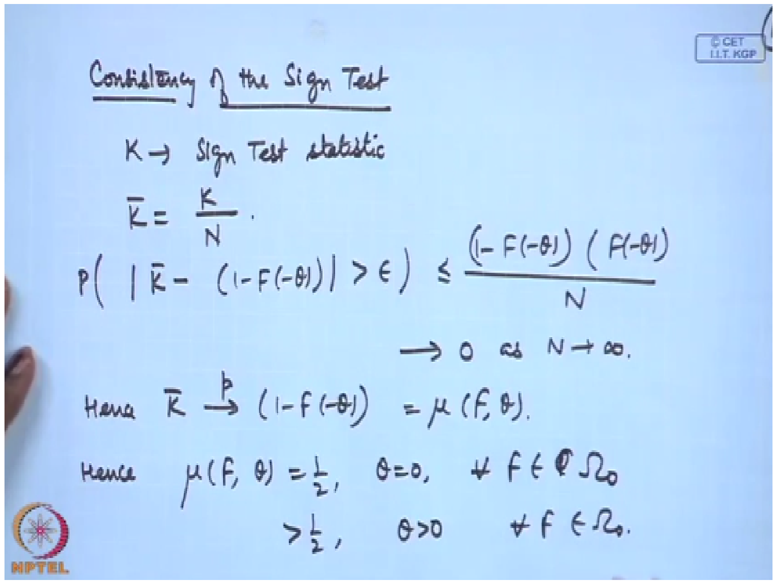
(Refer Slide Time: 34:46)



From 7, we will have mu0=mu of G*-2 epsilon, hence Kn<mu of G*-epsilon. So, if we consider now modulus of Vn-mu G* < epsilon, this will imply that Vn-mu G* is > -epsilon which implies that Vn is > mu G*-epsilon which implies that Vn is >=Kn from 8 because of this condition here. So, this implies that probability of modulus Vn-mu of G*<epsilon for the distribution G*, it is <= probability of Vn >= Kn under the distribution G* that is <= 1.

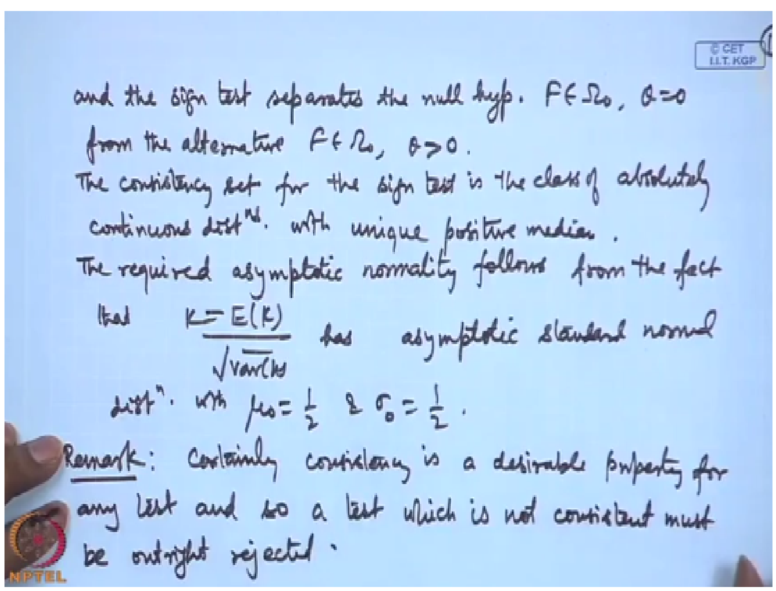
Now by the equation number 2 that we have taken here, that is Vn goes to mu G as n tends to infinitely, so therefore the left-hand side converges to 1, hence probability of Vn >=Kn goes to 1. Since G* was arbitrarily fixed in omega C, the theorem is proved. We will prove the consistency of some standard test here.

(Refer Slide Time: 37:00)



Let us consider say consistency of the sign test here. The sign test that we had introduced firstly for testing whether the median is = 0 or > 0 or < 0. So, let us consider here K is the sign test statistic. So, let us define say K bar=K/N, so probability of K bar-1-F of –theta > epsilon. This is <= by (()) (37:53) in equality 1-F of –theta*F of –theta/N and this goes to 0 as N tends to infinity. So, K bar converges to 1-F-theta in probability, that is mu of F theta. Mu of F theta=1/2 and it is > 1/2 for theta>0.

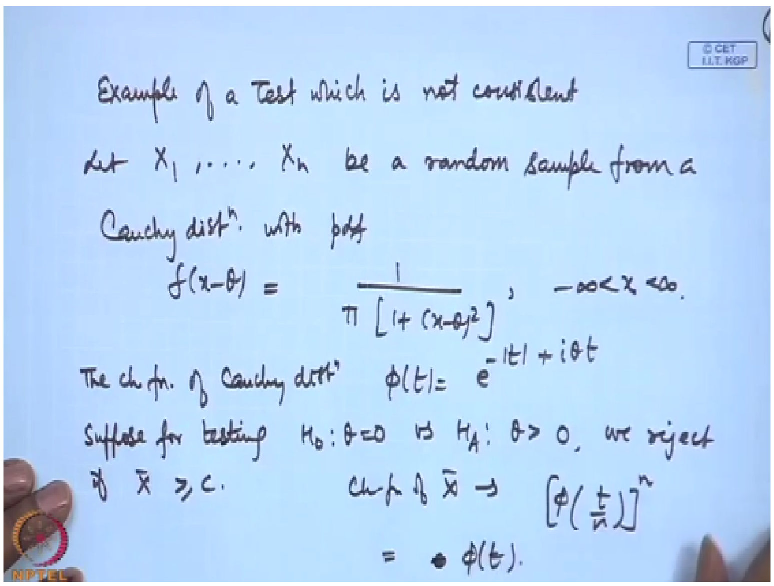
(Refer Slide Time: 38:42)



The sign test separates the null hypothesis F belonging to omega0, theta=0 from the alternative F belonging to omega0 theta>0. The consistency set for the sign test is the class of absolutely continuous distributions with unique positive median. The required asymptotic normality will follow from the fact that K-expectation K/square root of variance K has asymptotic standard normal distribution with mu0=1/2 and sigma0=1/2.

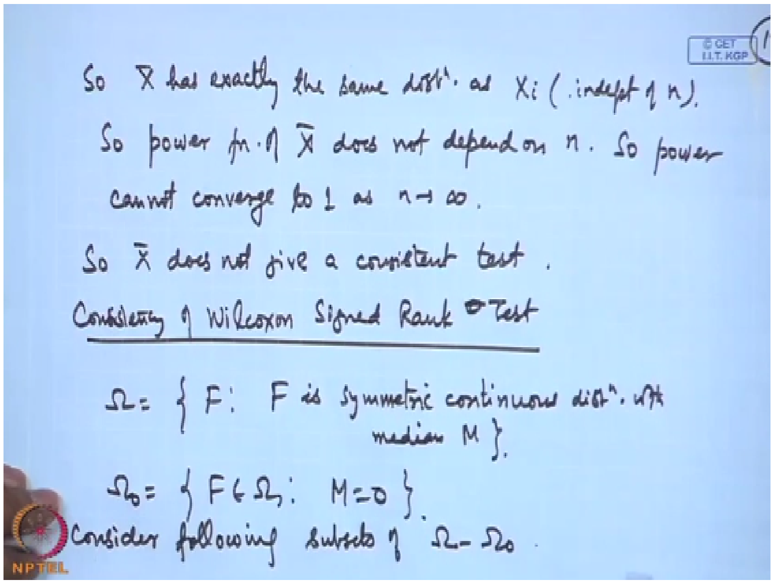
Any reasonable test should be actually consistent, therefore actually consistency does not provide a criteria for distinguishing among tests; however, if a test is not consistent, then certainly it is a defective test. So, that means basically all the go test must be consistent test. So, let me just give it as a remark here. Certainly, consistency is a desirable property for any test; and so, a test which is not consistent must be outright rejected.

(Refer Slide Time: 41:52)



Let me consider a test which is not consistent. So, let us consider say X1, X2,…Xn be a random sample from a Cauchy distribution, that is with pdf suppose I am considering the location form 1/pi 1+x-theta square. Actually, we know the characteristic function. The characteristic function of Cauchy distribution, that is phi t=e to the power –modulus t+i*theta t. So, suppose for testing H0, theta is = 0 against alternative theta > 0, we reject if X bar is >=C. So, if we consider this characteristic function of X bar, then it is same as phi of t/n to the power normal that is equal to same thing basically, because of this form it will turn out to be phi t itself.

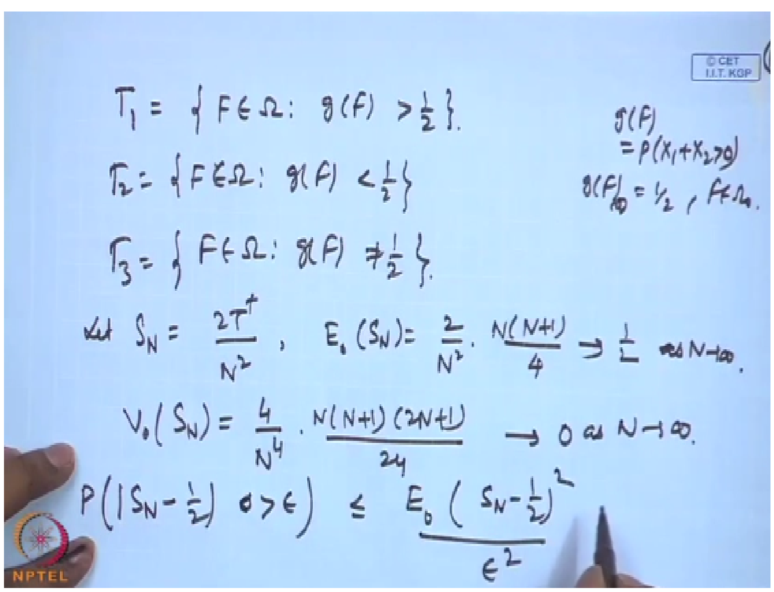
(Refer Slide Time: 43:53)



So what we are concluding here is that X bar has exactly the same distribution as Xi, that is it is independent of n. So, power function of X bar does not depend upon n. So, power cannot converge to 1 as n tends to infinity. So, X bar does not give a consistent test. So, this is an example of a bad test. I have proved the consistency of the sign test. Let us also consider the consistency of the Wilcoxon test.

So, omega is the class of symmetric continuous distribution with median given by m and omega0 is the class where we say that median=0. So, we consider the following subsets for alternatives, following subsets of omega-omega0. So, these are for defining the alternative hypothesis here.

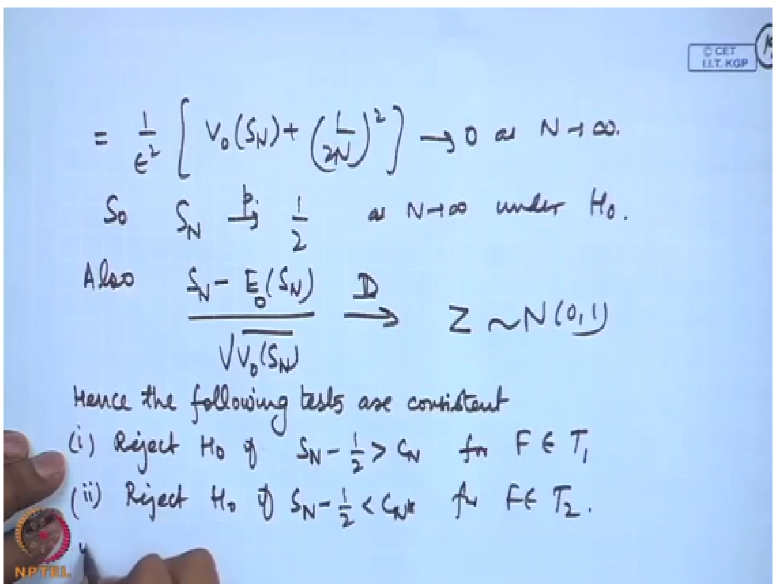
(Refer Slide Time: 46:05)



Let me call it gamma one, that is F belongs to omega where Gf>1/2. So, basically here Gf is probability of X1+X2>0. Under null hypothesis, this is = 1/2. Omega 2 is where Gf<1/2 and gamma 3 f belonging to omega where Gf != 1/2. Let us consider here Sn, that is 2T+/Normal square. So, if I consider expectation of Sn that is = 2/N square*N*N+1/4. Naturally, this will converge to 1/2 as N tends to infinity.

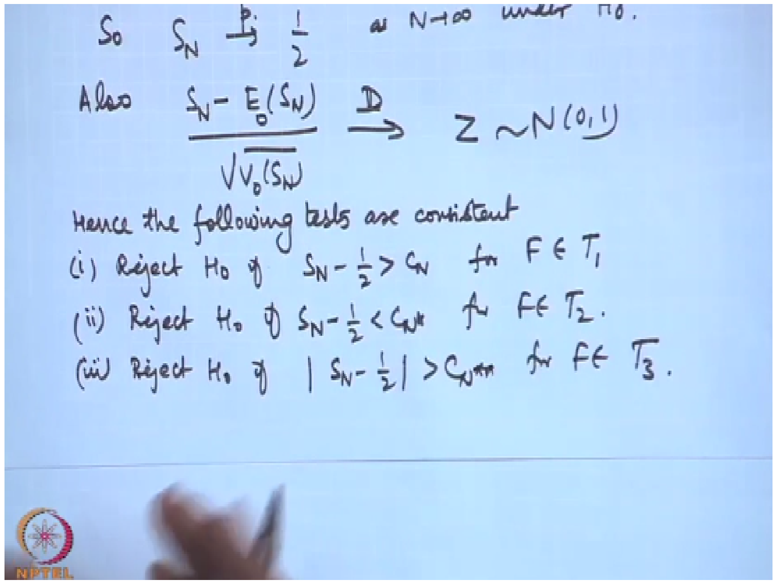
Similarly, if I consider variance of Sn, then that is 4/n to the power 4 variance of T+ that is N*N+1*2N+1/24 that goes to 0 as N tends to infinity, this goes to 1/2. So, if I consider probability of Sn-1/2 modulus being > epsilon, then by (()) (47:43) inequality it is <= expectation of Sn-1/2 square/epsilon square and this we simply split into 2 parts.

(Refer Slide Time: 47:53)



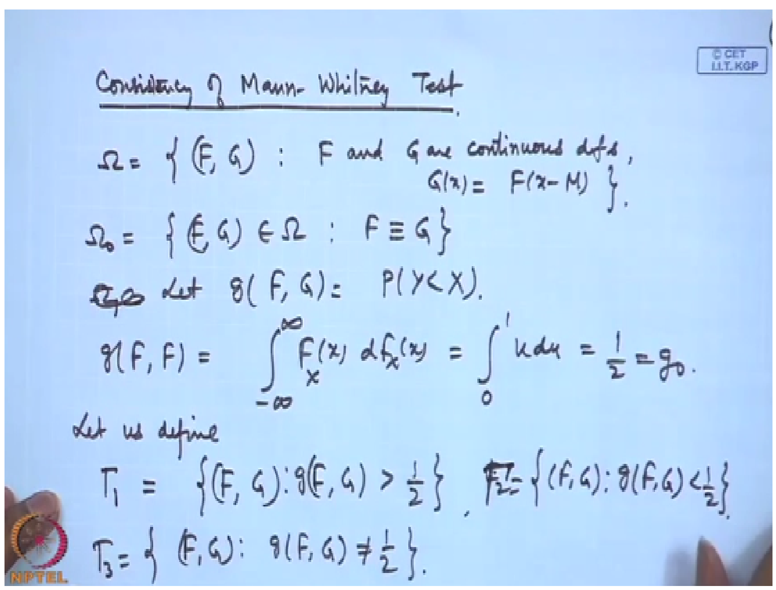
That is = 1/epsilon square variance of Sn+1/2n whole square. So, this goes to 0 as N tends to infinity. So, what we have proved here that Sn converges to 1/2 in probability as N tends to infinity under H0 and the asymptotic distribution of Sn is also normal, hence the following tests will be consistent, that is reject H0 if Sn-1/2>Cn for F belonging to gamma 1, reject H0 if Sn-1/2<some Cn star for F belonging to gamma 2.

(Refer Slide Time: 49:15)



Thirdly reject H0 if modulus of Sn-1/2>say Cn double star for F belonging to gamma 3. All of these 3 tests statistics will be consistent.

(Refer Slide Time: 49:40)

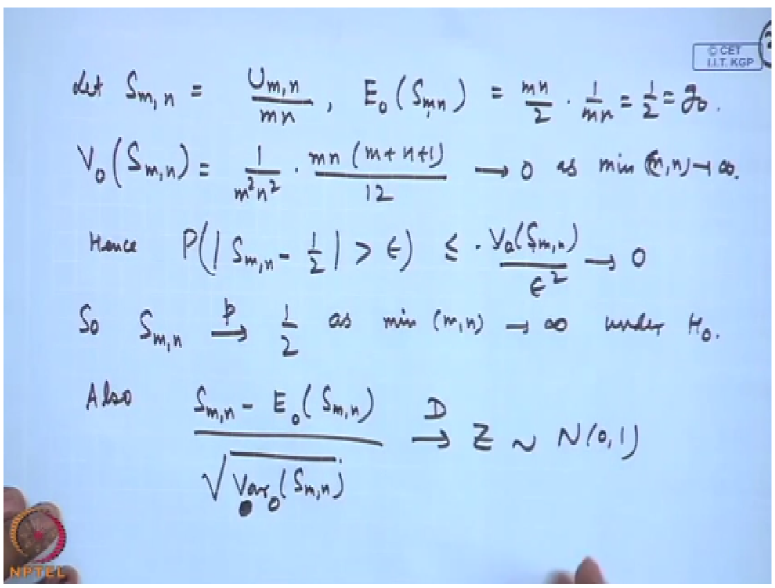


Since, this Mann-Whitney test was simply a variation here from the Wilcoxon, let us prove the consistency of Mann-Whitney also. I would just like to explain once again here the consistency of the Wilcoxon signed rank statistic. In the null hypothesis, we are saying the median is 0. So, here we are saying on either side that the median is > 0, < 0 or != 0.

So, to prove this what we considered is that consistency under asymptotic normality, then for the one-sided alternative, that is theta > 0 when we are having the right-hand side as the rejection region, then this is a consistent test. For m<0 when we have the alternative, the left hand rejection region is consistent and the 2-sided rejection region will be consistent when we have the 2-sided alternative hypothesis here.

Now, let us consider consistency of Mann-Whitney test statistic. So, we define omega=the class of all 2 sample problems. So, F and G are continuous distribution functions and Gx=F of x-M. So, omega0 is the case when M=0 that means we are considering F, G belonging to omega such that F and G are the same. Let us define G of F, G that is = probability of Y<X. SO, G of F, F that is = F of X*dfx that is = integral du 0 to 1 that is = 1/2 that we call it is = G0. So, we define the alternative hypothesis sets as F, G such that G of F, G is > 1/2, < 1/2, gamma 2 or gamma 3.

(Refer Slide Time: 52:51)

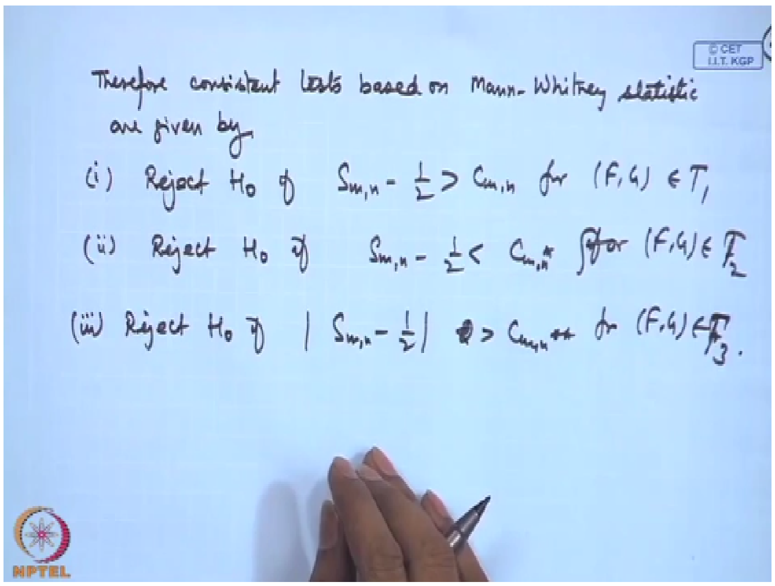


Let us consider here Smn=Umn/mn, this Umn was the Mann-Whitney statistics, so we are considering scaling by mn here. So, expectation of Smn=mn/2*1/mn=1/2=G0. Let us consider variance of Smn=1/m square n square mn*m+n+1/12, this goes to 0 as minimum of mn goes to infinity.

Because one of the mn's will cancel out and the term will become 1/m+1/n+1/mn, so if minimum of mn goes to 0, then both of the terms will go to 0 and if I consider then Smn-1/2 probability of this > epsilon, then this is <=, well again we can show that this is <= variance of Smn/epsilon square, this goes to 0. So, we are concluding that Smn goes to 1/2 in probability as minimum of mn goes to infinity under H0.

Also the asymptotic distribution is established under H0, this goes to Z following normal 0, 1. Since these 2 properties are satisfied, we conclude that the Mann-Whitney test statistics will be consistent provided we define it in the following question. We have the 3 alternative hypothesis, one is when we are considering gFG>1/2, so we consider the right-handed rejection region, here we consider the left-handed rejection region, here we consider the 2-sided rejection region here.

(Refer Slide Time: 55:30)



So, let me define it here, therefore consistent tests based on Mann-Whitney statistic are given by reject H0 if Smn-1/2>Cmn for F, G belonging to gamma 1, reject H0 if Smn-1/2 is < Cmn star for F, G belonging to gamma 2 and thirdly, reject H0 if modulus of Smn-1/2 is > Cmn double star for F, G belonging to gamma 3. So, all of these test functions are actually consistent tests here.

Here we have considered 2 types of 2 sample problems. In one of the 2 sample problems, we are shifting by a location and in another one we are shifting by a scale. So, we want to know whether the shifting is actually significant or not, that means like if we are shifting by the location then we are saying whether that shifting is in the positive direction or it is in the negative direction.

Similarly, in the scale, we are considering > 1 or < 1, that means whether we are introducing more variability or we are considering less variability. One may also think of general 2 sample problem in which we do not talk about the location scale, rather we consider whether the 2 distributions are the same or not. It is something like we consider in the one sample problem that we test whether the given distribution function is of a given form. So, we have for example a chi-square test for goodness of it.

We also introduced the Kolmogorov Smirnov test for single sample problem. So, in a similar way, if we consider a more general form of the hypothesis for a 2 sample problem, that means we simply say whether the 2 distributions are the same or not, then you can consider it as a goodness of a problem and we can consider a Kolmogorov Smirnov sample tests for this. So, in the lecture I will be actually discussing about the Kolmogorov Smirnov test and we will discuss the concept of efficiency of the tests also. So, in the next lecture, I will take up this part here.