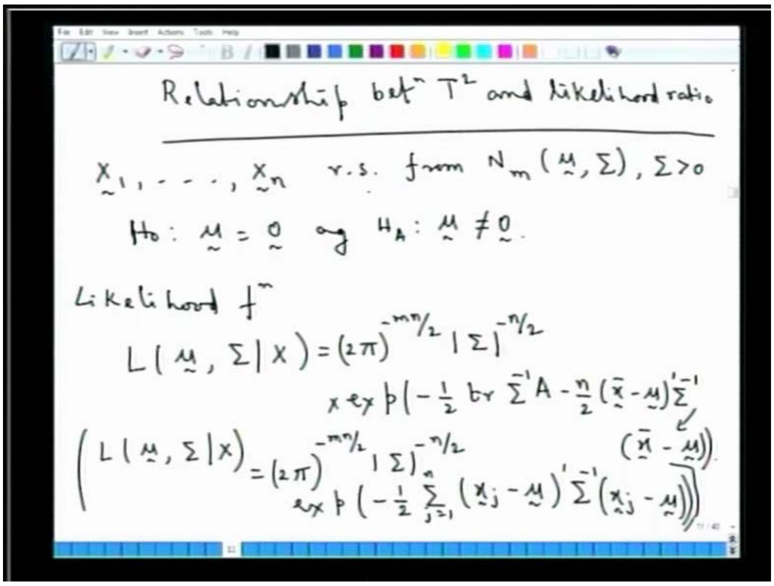
Applied Multivariate Analysis

Prof. Amit Mitra
Prof. Sharmishtha Mitra
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture No. #14

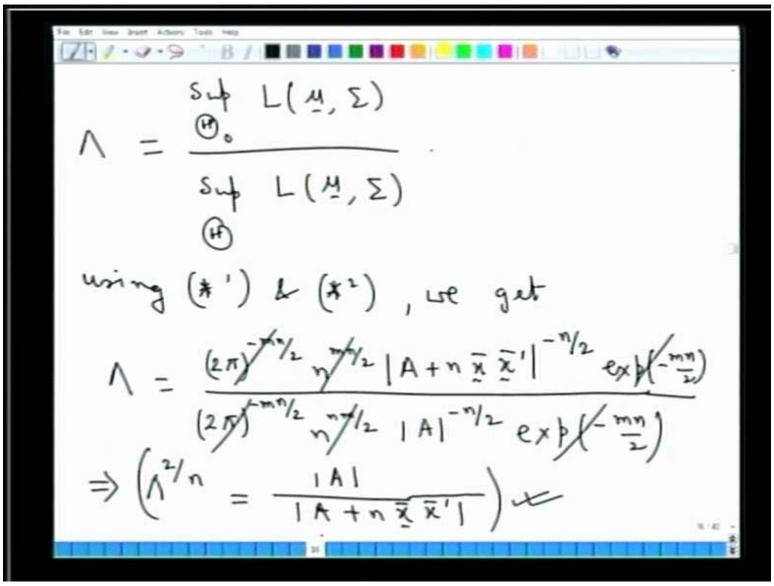
Hotelling's T2 Distribution and Various Confidence Intervals and Regions

(Refer Slide Time: 00:25)



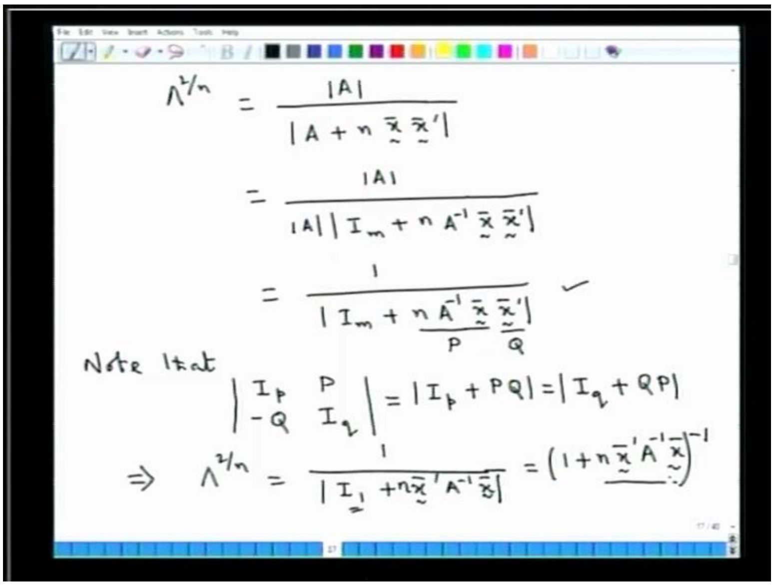
In the last lecture, we were establishing the relationship between hotelling's T square statistic, and the likelihood ratio statistic, and we had considered the that X 1, X 2, X n is a random sample from a multivariate normal distribution with mean vector as mu, and a covariance matrix positive definite sigma, and we were looking at this null hypothesis to be tested that mu equal to a null vector, against an alternate hypothesis that mu is not equal to a null vector.

(Refer Slide Time: 00:46)



And we had derived the likelihood ratio statistic lambda which is given by supremum under script theta naught, where script theta naught was the parameter space under the null hypothesis of the likelihood function L mu sigma; that divided by the supremum of the likelihood function under script theta, where script theta is the unrestricted parameter space not restricting oneself to the null hypothesis space. And then, we had come up to the point that we had shown that this lambda to the 2 by n is equal to determinant of A divided by determinant of A plus n times x bar, x bar transpose.

(Refer Slide Time: 01:25)

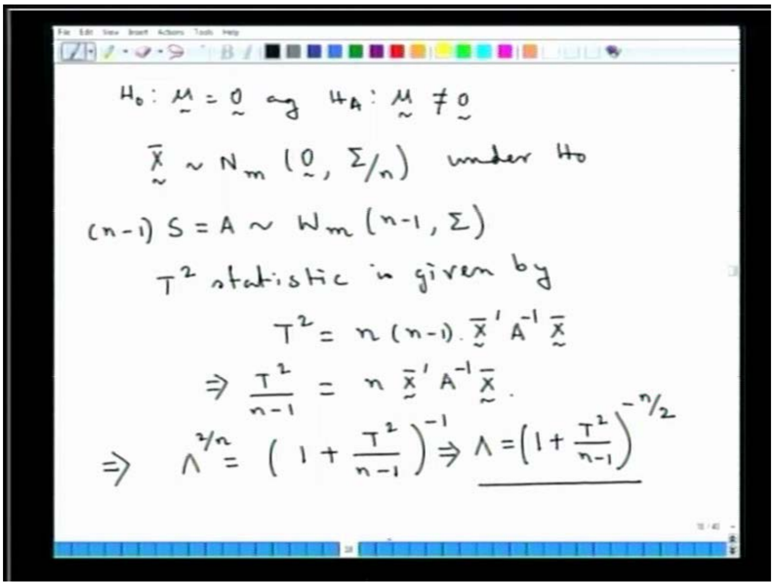


So, we continue from this particular point; we have this lambda to the power 2 by n. That is given by determinant of A, that divided by determinant of A plus n times x bar x bar transpose this is determinant. Now, what we can do here is that we can take determinant of A outside from this denominator to get to the form that this is determinant of A. So, if we take A outside here, what will be having here is I m; that is the dimension there, this plus n times A inverse and x bar x bar transpose it is determinant. So, that this term is just equal to one upon determinant of I m plus n times A inverse x bar x bar transpose.

Now, we note that we have a following result that, if we are looking at determinant of I say of P dimension I q minus Q A matrix and A matrix P here. Now, the dimension of this P is P by Q, and the dimension here is Q rows, and P columns. So, this determinant can be written as I p this plus P times Q that using the result of this partition the matrixes determinant; this is also I q plus Q times P. Now, in this expression here, if we take, if we use this same result; here, we will take this term as say P matrix; this actually is a vector, this is an m by 1 vector, and this is a Q, which also is a vector.

So, we will be able to write using this result here, this lambda to the power 2 by n that is equal to 1 divided by determinant of I 1. So, it is basically a scalar quantity, it is just equal to 1 that plus Q times P. So, we can interchange this, and we can write this as now, this note that this n is a scalar quantity. So, what will be having here is x bar n times x bar transpose A inverse and x bar determinant of that now, this is a scalar quantity this also is a scalar quantity. So, this is nothing, but 1 plus n x bar transpose A inverse x bar to the power minus 1.

(Refer Slide Time: 04:44)



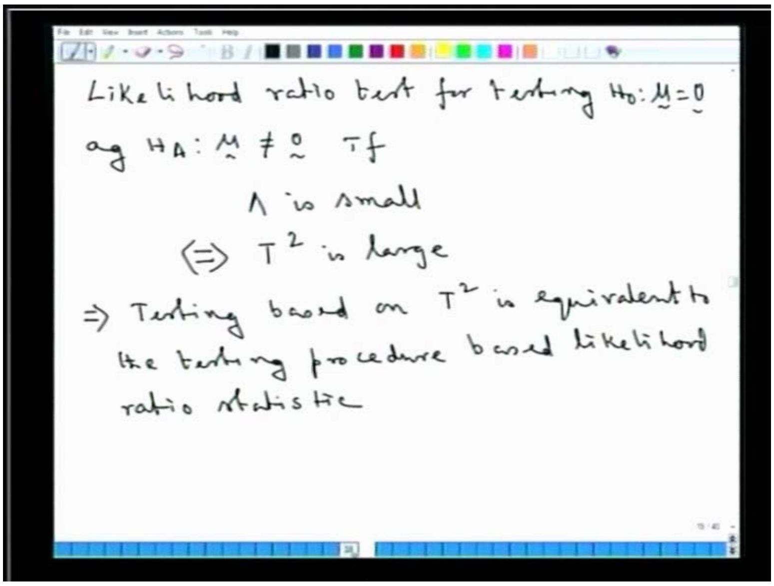
Now, we will see what is the relationship of this with a T square statistic? Now, the corresponding T square statistic, if we have H naught, mu equal to a null vector; this is to be tested against our alternate hypothesis that this mu is not equal to a null vector. We will have the distribution of x bar to be a multivariate normal with a mean vector as a null vector, and a covariance matrix as sigma by n this under H naught, and what more is n minus 1 times S with a divisor n minus 1. Let us, write that to be equal to A, which is actually equal to A, this would follow a wishart distribution on m dimension with degrees of freedom as n minus 1, and an associated variance, covariance matrix as sigma.

So, what would be the T square statistic from here so, if we have this setup; then the T square statistic is given by T square; which is, if we recall the result that T square, how T square is formed from a wishart distribution, and a multivariate normal distribution; this would be given by n times n; n is the degrees of freedom of the associated wishart which is n minus 1 times this x bar transpose and then this wishart matrices inverse x bar. So, this statistic is going to be given by this, and hence, this would imply that this T square by n minus 1 this is equal to n times x bar transpose A inverse x bar.

So, T square by n minus 1 in case of testing from a multivariate normal distribution with mean vector equal to a null vector is going to be given by this, and thus, if we have this lambda to the power 2 by n to be given by this actually as, we can see is same as, what we have obtained out here; so, this would imply that this lambda. Let me, take that 2 by n to the other side; and let me, first keep it in the form, that it was this would be given by 1 plus T square the observed value in case of small x bar, and the associated A matrix. So, this is T square by n minus 1 whole to the power minus 1.

This would further imply that the likelihood ratio statistic is going to be given by 1 plus T square by n minus 1 whole raise to the power minus n by 2. So, if this is the relationship between the likelihood ratio statistic, and the hotelling’s T square statistic. We can say the following, that the likelihood ratio test would reject the null hypothesis likelihood ratio test for testing this H naught may be equal to a null vector against the alternate hypothesis H A, that mu is not equal to the null vector. If lambda is small, that is, how a likelihood ratio test goes along, because it is looking at the ratio of the two supremums: one in the denominator is supremum over the null space, and the one in the denominator is a supremum over the entire parameter; space not restricting oneself to the null space.

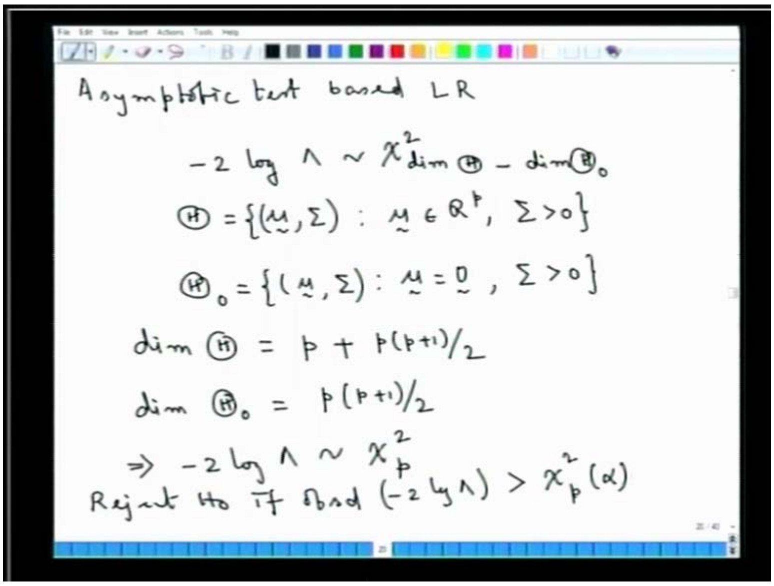
(Refer Slide Time: 07:49)



So, the likelihood ratio test would reject if lambda is small; this is equivalent, because we have this relationship; that this lambda is 1 upon 1 plus T square to the power n by 2, and hence lambda being small is equivalent to saying that; T square is large, and that is, what is the desired thing. So, this is equivalent to saying that, if T square is large. So, this would imply that the testing based on T square is equivalent to the testing procedure based on the likelihood ratio, likelihood ratio statistic.

So, the two actually are equivalent this gives us another justification actually of the T square, a hotelling's T square statistic; that, if we are looking at using the hotelling's T square statistic for testing such null hypothesis as, what is framed that testing procedure is equivalent to a likelihood ratio test principle. Now if this lambda, what we have derived is this particular quantity? We can derive the distribution of this lambda or we can use an asymptotic distribution; another distribution of lambda is going to be based on the distribution of T square T square remember another null hypothesis has got a F distribution on m n minus m plus 1 degrees of freedom, and hence the distribution of this lambda can theoretically be derived, when we have the relationship between this lambda, and T square.

(Refer Slide Time: 10:37)

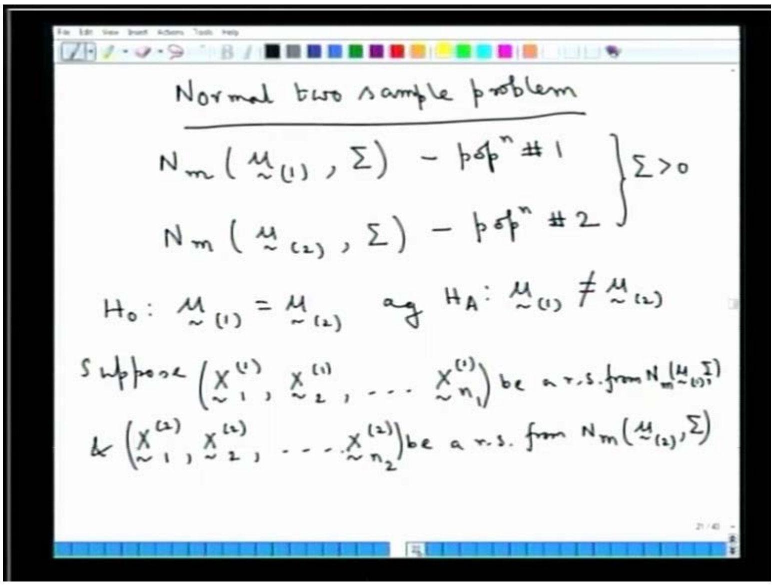


For large sample asymptotic test based on this likelihood principle can be obtained an asymptotic test based on L R, likelihood ratio is going to be given or rather would be based on the following fact, that this minus 2 log lambda, this would follow a central chi square on the degrees of freedom, which is the difference of the dimension of script theta, and the dimension of script theta naught. So, this minus 2 log lambda, while lambda is given by this likelihood ratio, which is this particular quantity; one can look at what is this degrees of freedom; now, what is the dimension of script theta, and dimension of script theta naught. Let me, write the two spaces that it would be easy to see, what is the dimension of the two spaces.

So, this is mu sigma space, where this mu belongs to R to the power P, and this sigma is positive definite; it is a symmetric matrix with P into p plus 1 by 2 distinct elements, and script theta naught is the null space, which is the set of all mu sigmas; such that, mu under the restriction of the null spaces, the null vector, and this sigma is still positive definite. So, the dimension of script theta would be the P dimensions; that are present in mu, this plus the number of free parameters in sigma, which is p into p plus 1 by 2, and the dimension of script theta naught, this is equal to p into p plus 1 by 2. The distinct element corresponding to the sigma matrix and mu is specified and nothing comes actually from there. So, this would imply that minus 2 log lambda would follow asymptotically a chi square random variate on just P degrees of freedom. Now reject, we will reject the null hypothesis, reject H naught, if observed value of this test statistic that is minus 2 log lambda. The observed value of that is greater than chi square P alpha, when this is the upper alpha percent cut off point of a central chi square P degrees of freedom.

So, this is the asymptotic test, that is based on the likelihood ratio principle, which of course we have shown, that it is equivalent to the hotelling's the test that is going to be based on the hotelling's T square statistic.

(Refer Slide Time: 13:34)



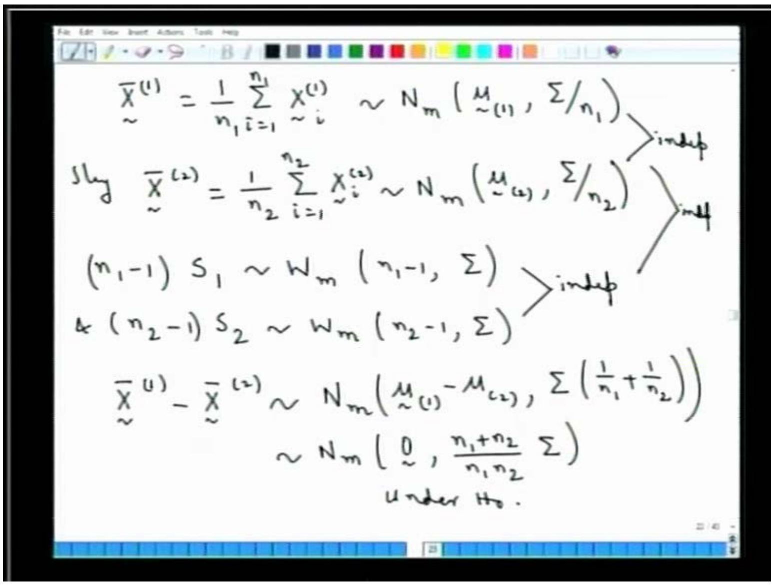
Now, let us look at this a normal two sample problem, and how to use the hotelling's T square statistic in order to perform testing. In this situation, normal two sample problem: what is the problem here? We have two multivariate normal populations, say N p 1 will be using N m. So, let me stick to that particular notation that, multivariate normal with a mean vector as mu 1, and a covariance matrix as sigma.

So, this is population number 1, this is population number one and the second population is multivariate normal m with a different mean vector, and the same covariance matrix sigma. So, this is what is characterizing the second population, and sigma of course is positive definite, we have this common assumption that sigma is positive definite for the two populations. So, we have two multivariate normal populations different in the mean vector and the covariance matrix are same. Now, in this particular situation the point of interest is to test the following null hypothesis, that this mu 1 vector is equal to mu 2 vector; this is to be tested against the alternate hypothesis that this mu 1 vector is not equal to this mu 2 vector.

It is a standard testing procedure, if we look at the corresponding univariate counter part then the clearly the univariate counter part also, we have done the testing of course in univariate counter part, when we have two univariate normal populations with same variance, and different mean component; then the testing of that of course is obtained using the T statistic the students T statistic, and this we are going to show that, the testing of this particular problem in the multivariate to sample to multivariate normal population can also be obtained using a hotelling's T square statistic.

Now, suppose we have two sets of random samples; now suppose X 1 2, X 2 1, X 1 n 1, be a random sample from the first population; which is n m mu one sigma, and we have a second set of sample from the second population X 2 1, X 2 2, and X say n 2 2, be a random sample from the second multivariate normal population. That is, it is a random sample from mu 2 sigma. So, we will be using this two sets of random samples; this is the set of random samples n 1 size from the first population; this is the set of random sample from the second population of size n 2.

(Refer Slide Time: 17:07)



Now, if we have this particular setup then the following can be easily said that, suppose x one bar is one upon n summation i equal to, I am sorry this is 1 upon n 1. So, I am looking at the first population, i equal to 1 to up to n 1 of this X 1 i. So, this is the sample mean random vector from the first population; this has got a distribution multivariate normal m with a mean vector as mu 1 and a covariance matrix as sigma by n 1.

Now, similarly we have the second population giving us, this mean vector, sample mean vector based on the n 2 random samples from the second population; this is 1 upon n 2 summation i equal to 1 to up to n to X i random samples from the second population. So, these follow a multivariate normal similarly, with a mean vector as mu 2, and a covariance matrix as sigma by n 2. Now, these two are going to be independent, why because this X 1 bar is based on X i 1 terms n 1 of them, and this is going to be based on the other set of random samples from the second population.

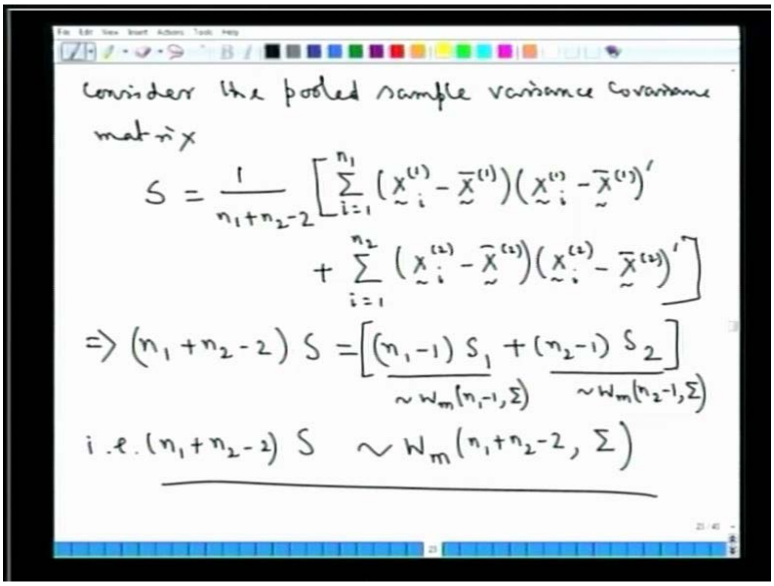
Now, these being set of random samples this, and this. So, this set of random sample is independent of this set of random samples, and within the random samples sets actually this X 1 1 would be independent with each of them. So, this set is also, set of independent random vectors, and this set also is independent random vectors, and the two are mutually independent, because they are random samples from two different multivariate normal populations. So, we have first of all this particular result. Now, corresponding to the sample variance, covariance matrix; what we can say is that n 1 minus 1 times S 1, where S 1 is a sample variance, covariance matrix based on the set of n 1 random samples from the first population. This would follow a wishart distribution on degrees of freedom as n 1 minus 1, and a covariance matrix as sigma, and similarly based on the second set of random samples n 2 minus 1 times S 2; this would also follow, a wishart distribution wishart m n 2 minus 1, and sigma.

Now, once again these two are independent, and what is more important is to realize that this X 1 bar is independent of S 1; that is the result that, we have proved for multivariate normal distribution X 2 bar is independent of this S 2 quantity. So, that what we can further say that all these four statistic. They are mutually independent, first of all from the set of random samples of size n 1; this X 1 bar, and S 1 are independent from the set of random samples of size S 2 from the second population. These X 2 bar, and S 2 X 2 bar, and S 2 are independent, and they are mutually independent, because they are random samples from two different multivariate normal populations.

So, this is what we are? These are the building blocks for getting into the [hotelling's] T square statistic for this particular setup. Now let us, look at the distribution of this X 1 bar minus X 2 bar, what is going to be the distribution of this the distribution of this; we have X 1 bar to be this multivariate normal X 2 bar to be this multivariate normal distribution these two are independent. So, one can very easily show that, this also is following a multivariate normal distribution m dimensional with mu 1 minus mu 2; as it is mean vector, and the covariance matrix would just be the sum of the two covariance matrices, because X 1 bar, and X 2 bar, are independent.

So, what is it that is equal to sigma times 1 upon n 1 plus 1 upon n 2. So, this is one upon n 1 plus 1 upon n 2. Now, this would follow under the null hypothesis that mu 1 is equal to mu 2. This would follow a multivariate normal with a mean vector as null vector, and a covariance matrix, as n 1 plus n 2 by n 1 n 2 times sigma. This is the distribution under the null hypothesis H naught being true. So, this as far as combing the two mean sample mean vectors.

(Refer Slide Time: 22:08)



Now, what we can-we say further is that consider the pooled sample variance, covariance matrix, consider the pooled sample variance covariance matrix that would be given by S, which is 1 upon n 1 plus n 2 minus 2. This is actually same as that type; what we usually do with univariate theory also. So, that would be the pooling here 1 upon n 1 X i 1 minus X 1 bar. So, this is for the part that is coming from the first set of random samples. So, it is X i 1 minus the mean coming from all these X i 1 quantities that multiplied by X i 1 vectors, minus X bar 1 vector transpose, this plus the second set of random samples i equal to 1 to up to n 2 X i 2. So, these are based on the second set of random samples deviation taken from the second the random mean vector. The sample mean vector is based on the second set of random samples; this is X 1 X i 2 minus X bar 2; it is transpose.

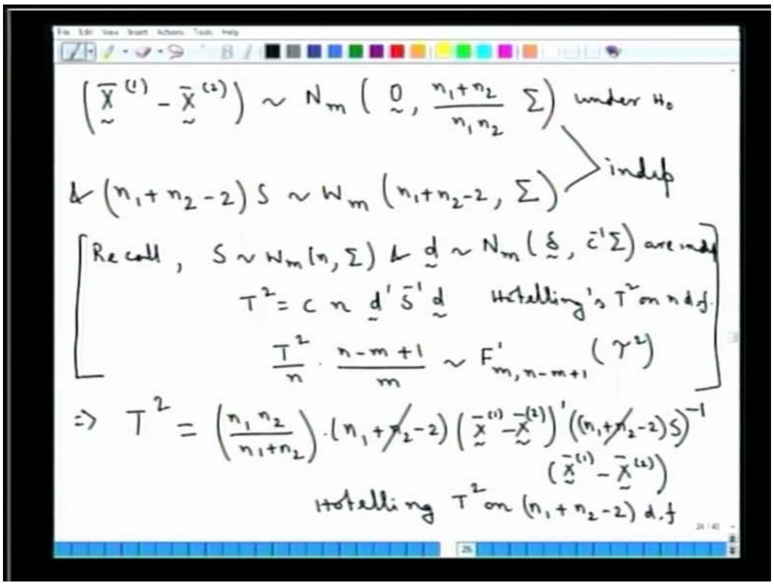
So, this is the pooled sample variance, covariance matrix, and what is this term equal to this term is nothing, but n 1 minus 1 S 1. So, that we can say that this is n 1 plus n 2 minus 2 times S, this is equal to the first term; which is n 1 minus 1 times S 1 plus n 2 minus 1 times S 2. So, we have these two the sum of these two giving us n 1 plus n 2 minus 2 times the pooled sample variance, covariance matrix. Now, what can we say about the distribution of this note that this quantity here follows, a wishart distribution wishart m n 1 minus 1 degrees of freedom, and an associated variance, covariance matrix as sigma.

And Similarly, the second part here; this follows, a wishart distribution with the same dimensionality as the degrees of freedom changes; it is n 2 minus 1, and the same variance, covariance matrix as sigma. Now these two are independent; that is what we had said in the previous slide here, that this these two statistic, they are going to be independent, because they are based on two different random samples, and hence this is just the sum of two independent wishart distributions; now by the additive property of the wishart distribution keeping the dimensionality, and the associated sigma the associated variance, covariance matrix to be same.

We will have this particular sum of two independent wishart distributions with the same dimension, and same sigma, this will have a wishart distribution m, and the degrees of freedom pooling up. So, its n 1 plus n 2 minus 2 degrees of freedom, and the same variance, covariance matrix as sigma; that is, what we now have is n 1 plus n 2 minus 2 times this pooled sample variance, covariance matrix to have this particular distribution.

Now, we we now have the building blocks for getting into the [hotelling's] T square statistic; one building block is here, and the second building block is basically here, right because we have one part here a multivariate normal distribution, and we have another wishart distribution. Now critically, this multivariate normal distribution, and this this multivariate normal distribution and this wishart distribution, they are going to be independent.

(Refer Slide Time: 26:25)



Let me, write it in a compact form here, that this X 1 bar minus X 2 bar this is a random vector. This follows a multivariate normal distribution m with a null vector, as it is mean vector under the null hypothesis, and this n 1 plus n 2 divided by n 1 n 2 times sigma under H naught. We are looking at it under H naught, because any way we will have to look at the distribution of the test statistic which is going to be hotelling's T square under the null hypothesis; otherwise, this remains mu 1 minus mu 2; this, and we also have n 1 plus n 2 minus 2 times S; this to follow a wishart distribution, wishart m the use of freedom as n 1 plus n 2 minus 2, and an associated variance, covariance matrix as sigma, and the two are independent.

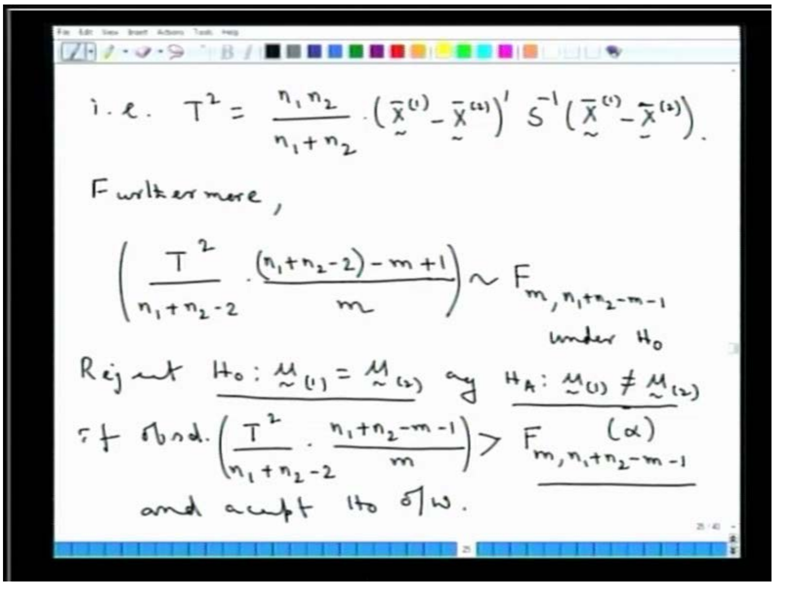
So, we can use; let me, write once again what we are now going to recall we are going to recall, that if S has a wishart m n sigma, and d; a random vector this is; how we had defined a hotelling's T square statistic? We have d; a random vector, which is having a multivariate normal with a mean vector as delta, and the covariance matrix as c inverse sigma, where c is a scalar quantity? If we have these two random vector, and random matrix are independent. Then this T square statistic was given by c times n d prime S inverse d. This is, how we had defined hotelling's T square. So, this is hotelling's T square on n degrees of freedom; the degrees of freedom is associated with this n; here, which is the degrees of freedom of the associated wishart distribution. So, if we have that further more what we will be having is T square by n in to n minus m plus 1; that divided by m, this to follow a non-central f distribution on m n minus m plus 1 degrees of freedom, and a non-centrality parameter equal to tau square; which is going to be given by c times delta prime sigma inverse delta.

So, this is what is the definition, and the distribution of a hotelling’s T square, which can be framed from a wishart distribution, and a multivariate normal distribution that being independent. So, we are going to use this definition, and the distribution of the hotelling’s T square statistic, in order to frame the hotelling’s T square statistic out of this particular problem. So, this would imply that the T square for the given problem would be given by what is c here; c is the inverse of this particular quantity. So, that is n 1 n 2 this divided by n 1 plus n 2. So, that is how a c term, and then what is n in our problem; n is the degrees of freedom associated with the wishart distribution, which is in our case n 1 plus n 2 minus 2.

So, we have taken care of these two terms; what is d prime? The d prime is the prime of this particular term here, because that is the multivariate normal distribution out there. So, we will have this X 1 bar minus X 2 bar transpose, and then we have S inverse S inverse S is, what S is the associated wishart distribution. So, we will have that as n 1 plus n 2 minus 2 this multiplied by this S; the pooled sample variance, covariance matrix inverse of that, and that multiplied by this d. So, this d is once again this X 1 bar minus x 2 bar.

So, this is now our hotelling’s T square; this is the hotelling’s T square for the given problem, for the given two sample normal problem hotelling’s T square on how many degrees of freedom? The degrees of freedom would be associated with the degrees of freedom of the underline wishart distribution. So, this is a hotelling’s T square on n 1 plus n 2 minus 2 degrees of freedom. So, what happens here? This term cancels out with this term, and whatever remains is the hotelling’s T square statistic.

(Refer Slide Time: 31:20)



That is, this T square is nothing, but this n 1 n 2, this divided by n 1 plus n 2, this times X 1 bar minus X 2 bar; it is transpose S inverse, where S is the pooled sample variance, covariance matrix X 1 minus, this X 2. Now, further more we will have to find the distribution of this T square or some constant times; this T square function in order to find the distribution of that, what we are going to use is this result for the general hotelling's T square distribution. So, that we will have this T square; now T square divided by what here divided by n? What is n; n is the degrees of freedom associated with the wishart distribution.

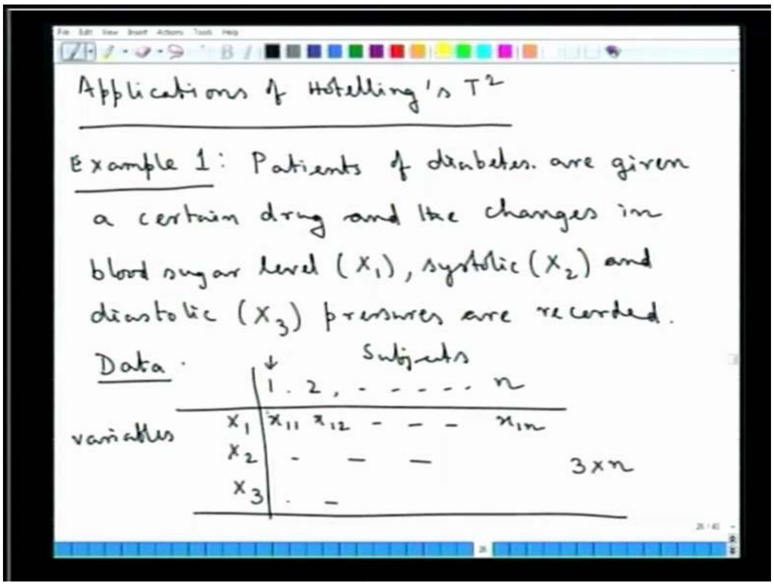
So, we have n 1 plus n 2 minus 2, which is now playing the role of n; then we have the constant here which was n minus m plus 1. So, that this is now our n minus m; m is the dimension. So, m remains as, it is this plus 1 this divided by m. So, this is equal to this will, this is going to follow a F distribution on what degrees of freedom; degrees of freedom is first is m, and the second is this is n 1 plus n 2 minus 2 minus m plus 1. So, this is equal to n 1 plus n 2 minus m minus 1.

And what is the non-centrality parameter? The non-centrality parameter of this F distribution if at all that would be given from this particular mean vector here mu. So, that is going to be this null vector prime, the inverse of this into the null vector which is equal to 0. So, this is a central F distribution on these degrees of freedom. So, if we have the hotelling's T square, which is from the two sample normal problem to have this particular distribution; we will reject the null hypothesis, we know it is distribution. This is under the null hypothesis this is under the null hypothesis.

So, if it is not under the null hypothesis, then what we will be having is this to be a non central F distribution, and the non-centrality parameter in such a situation would be given by mu one minus mu two prime, and then inverse of this particular quantity, that is n 1 plus n 2 by n 1 n 2 into sigma, whole inverse, that multiplied by mu 1 minus mu 2 that would be the distribution or rather that would be the non-centrality parameter of the F distribution; if this is not considered under the null hypothesis. Now, using this, what we can do is to look at the rejection regions. So, reject H naught, which is mu 1 equal to mu 2 against the alternate hypothesis H A, which is mu 1 not equal to mu 2, if observed value of this T square divided by n 1 plus n 2 minus 2 this multiplied by n 1 plus n 2 minus m minus 1 this divided by m is greater than the upper alpha percent tabulated value of the central F distribution on n 1 plus n 2 minus m minus 1 degrees of freedom.

So, this is the upper alpha percent cut off point of a central F distribution on m n 1 plus n 2 minus m minus 1 degrees of freedom, and thus this testing of this problem of testing, this null hypothesis against this alternate hypothesis; here, would be achieved, and we will reject the null hypothesis, if the observed value of this quantity the test statistic exceeds, the tabulated value, and we accept H naught; otherwise so, that completes the proof actually or rather the derivation of the test statistic, and formulation of this particular testing problem. In case of two sample multivariate normal population, we have seen how this testing problem can actually be framed, and testing can be carried out for one sample multivariate or rather one population multivariate normal problem, and two population multivariate normal problems.

(Refer Slide Time: 36:22)



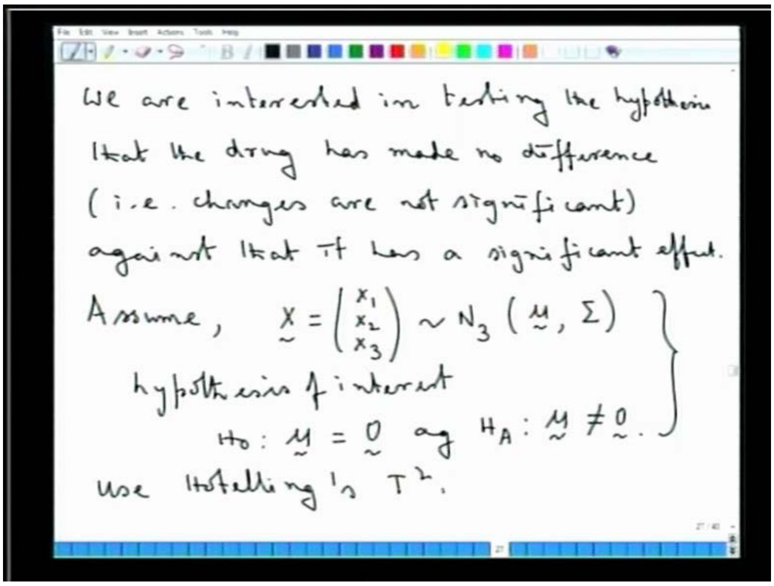
Let us, now look at some practical data in order to see, how this particular concept of hotelling's T square can be applied for real life data. So, applications of hotelling's T square is, what we are now going to see from somewhere life data examples; these are standard examples, hotelling's T square; we will look at two examples as, I said that these are standard examples; now, the problem is the following we have patients of diabetes are given a certain drug, certain drug and the changes in blood sugar level. Let, that be denoted by variable X 1 systolic. Let us, denote that by X 2 variable, and diastolic diastolic say X 3 variable pressures are recorded.

So, the problem is the following; it is a practical data example; that there are patients say n of them having diabetes are given a particular drug, and then in order to charge whether the drug is effective the following three variables are recorded; the recording on the three variables are taken the three variables are blood sugar level changes; changes in the blood sugar level changes in the systolic pressure, and the changes in the diastolic pressures are recorded. Now, what is the structure of the data the data looks like the following; suppose, I have got these as subjects, that is patients, we have n such patients. So, these are n patients; the data are recorded for these n patients; all the three variables there are three variables X 1, X 2, and X 3.

So, corresponding to these three variables corresponding to the subject; one we record that patients change in the blood sugar level, and record it out here, that patients change in the systolic pressure here, and the change in the diastolic pressure here. So, we will able to complete this particular data completely. So, this now is a three by n data matrix which holds all the recording for the n patients on these three variables of interest.

Now, the problem is framed or rather the question; that is posed is whether the drug which was administered on these n patients looks as, if it is an effective drug or not.

(Refer Slide Time: 40:00)



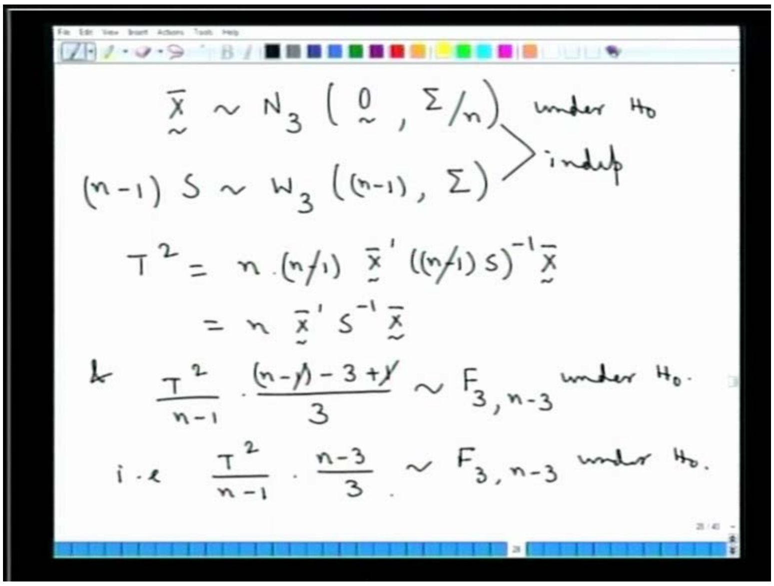
So, the question in hand is to answer is the following: we are interested in testing the hypothesis that the drug which is administered; which is under study has made no difference no difference with respect to the three variables; that is the changes are not significant that is the changes are not significant significantly different from 0 or not significant. This is to be tested against that it has made a change or it has it has a significant effect it has a significant effect.

So, this basically is a problem; practical problem; now, in order to test this particular or rather carry forward this particular problem, and then use this particular data in order to answer the question of interest that, whether the drug at all is effective or not that is going to be tested on the based based on on the basis of the three variables; that we have taken recordings on we assume the following: we assume that this X 3 random vector which is comprising of this X 1 variable, X 2 variable, and this X 3 variable. This follows a three dimensional normal distribution with a mean vector as mu, and a covariance matrix as sigma.

So, this is the mean vector corresponding to this random vector out here. So, the first component would not correspond to the expected change in X 1 variable in that multivariate normal population. Now, in terms of this assumption that we are made on this random vector, if we are interested in testing the hypothesis; that the drug had made no difference, that is the changes are not significant. These are the variables denoting the changes in the respective variables. So, the hypothesis of interest hypothesis of interest is thus the testing of H naught what is that it is testing for this change vector in the population V is a null vector.

This is to be tested against the alternate hypothesis, that this mu vector is not equal to a null vector. Now the problem looks in attractable form now. So, if we have made this particular assumption, and translated this particular hypothesis in terms of the mean vector of this multivariate normal population we can use a hotelling's T square statistic in order to test this practical problem use hotelling's T square.

(Refer Slide Time: 43:26)



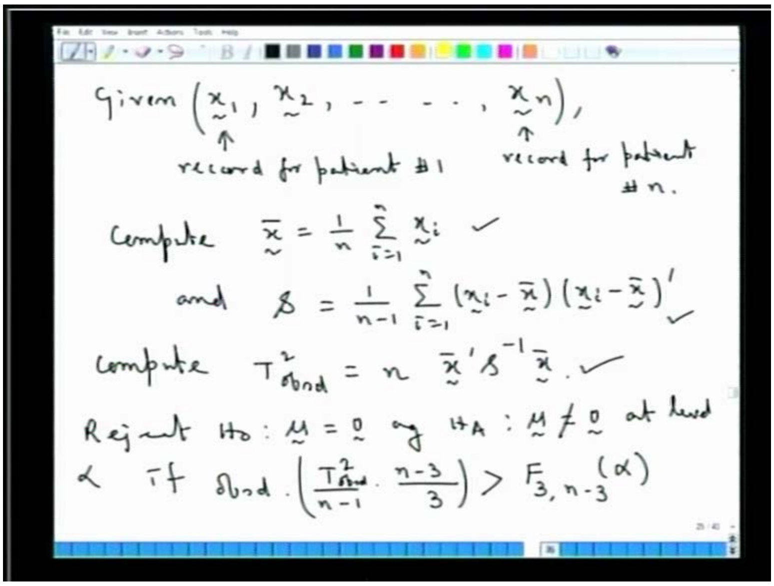
In what way what are the things that we will be requiring we will be requiring as the constituent elements as this x bar to have a three dimensional normal distribution with a mean vector as null vector under the null hypothesis and sigma by n to be it is covariance matrix. So, this is under the null hypothesis; that the mu vector, the mean vector mu is equal to a null vector; otherwise, this would be just a general mu vector, and what more we have a random sample of size n. So, we will have n minus 1 times S; S is the sample variance covariance matrix with a divisor n minus 1. So, this will have a wishart distribution; now, three dimension with a degrees of freedom as the number of samples that has been taken n minus 1, and the covariance matrix as sigma, which is of course, unknown the two are going to be independent from the multivariate normal sampling distribution theory.

So, this is what we have already so, using these two we can frame a hotelling's T square statistic, which is going to be given by n times degrees of freedom into this X bar transpose; the mean vector is a null vector, and this is an n minus 1 times S whole inverse of that that multiplied by this X bar vector. So, what we have we have these two cancelling out, and this is our good old hotelling's T square statistic; that is n times X bar S inverse X bar.

Now, further more we will have this T square; this T square divided by it is degrees of freedom; degrees of freedom is n minus 1, then this is n minus n; n here is 3 this plus 1 that divided by 3. This would follow a central F distribution under H naught 1 degrees of freedom as 3, and what is this equal to this is n minus 1 degrees of freedom minus m, that is 3 plus 1. So, this one cancels out plus 1, and minus 1. So, we will have this as n minus 3 this under H naught.

So, we have the null distribution of T square this divided by n minus 1 into this n minus 3 by 3; this follows central f distribution 3 n minus 3 under our H naught. So, if we have this particular term here from the data the data is this. So, we will have the data as this say X 1 1, X 2, X 1 2, say X the way that we write it it is no unique way of representing this indexes. So, these are basically the small x is are the data that we have in our hand.

(Refer Slide Time: 46:33)

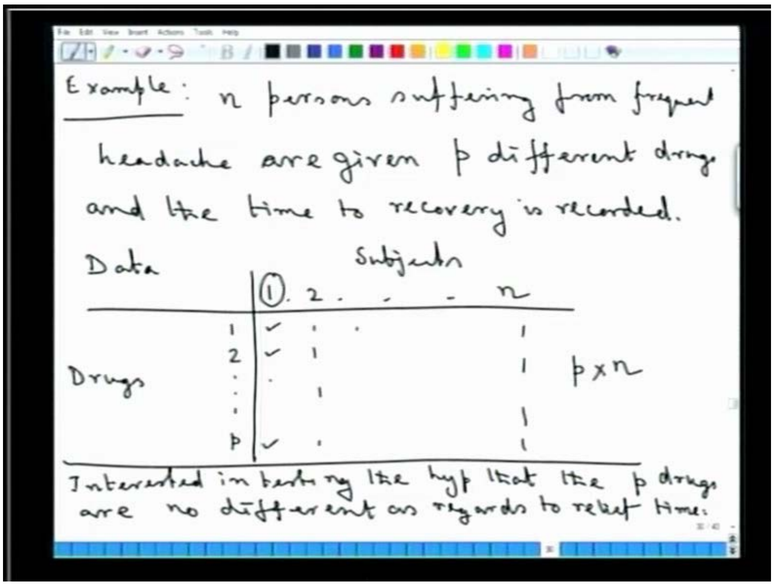


So, given the data given the data which is X 1? This is three dimensional vector. So, this is the recordings or the record for patient number one; then we have this X 2 vector these are the actual recordings. So, this is X n this is the three dimensional record for patient number n. So, these are the data. So, given this particular set of information; we compute X bar; the sample mean, and well this is going to be given by 1 upon n summation i equal to 1 to n of these X i vectors, and the S matrix with a divisor n minus 1. So, that that is given by i equal to 1 to n X i minus X bar X i minus X bar; it is transpose the random matrix, and this is actually the observed sample variance, covariance matrix.

So, given these two one can easily compute this T square observed T square observed would be n times X bar transpose, this S inverse, and this X bar, and we will reject once this is computed we have the testing reject H naught; that is that our new vector is a null vector reject H naught equal to a null vector against the alternate hypothesis H A; that this mean vector is not equal to a null vector at level alpha say, if we have observed value of this T square divided by what was it? It was T square by n minus 1 n minus 3 by 3.

Now, once we have computed this is T square observed. So, if this quantity; now, we will reject the null hypothesis if of course, T square is large, because that is what is going to give us the more deviation from the null hypothesis point, and hence we will reject the null hypothesis; if this observed value of this quantity here exceeds the tabulated value of this F distribution on three n minus 3 degrees of freedom at alpha point. So, this is the upper alpha percent cut off point of a central F distribution on three n minus three degrees of freedom. So, that is how we obtain from the given data this testing problem.

(Refer Slide Time: 49:35)

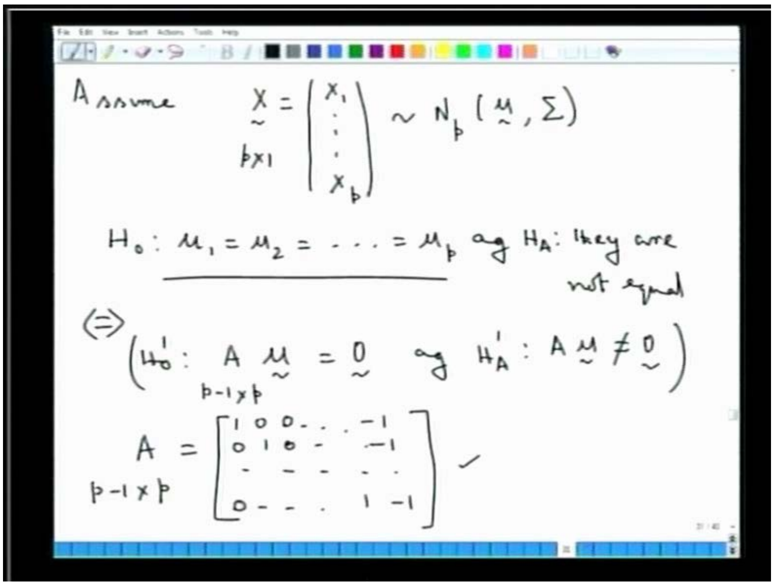


Now, let us look at example number two: this also is a standard type of example that one comes across in real life examples. So, that this is this problem that n persons suffering from frequent headaches; headaches are given p different drugs; this setup is a bit different p different drugs; and the time to recovery; recovery is recorded. So, the problem is the following: that, we have n persons suffering from headache for the given problem, it may be for some other disease also; the same approach can be applied there given p different drugs, and then the time to recovery corresponding to each of these persons for each of these p drugs are now recorded.

Now, the data structure is of the following form in this problem. So, these are what we have as the subjects, which are these patients 1 2 up to n, and then what we have here is the drugs. So, we have p drugs being administered. So, we have all these p drugs on this side. Now, corresponding to this person number 1; we will administer all these p drugs to the same person, and then the time to recovery from a particular headache by administering drug number 1 is put here, the time to recovery of the same patient; when drug two is administered at some other point of time is recorded, and like this, we have all the recordings taken for the first patient, and likewise for the second patient, and the third, and all the patients are recorded.

So, we have this data in the form of a p by n matrix; now what can be the point of interest in such a problem, the point of interest in such a problem for any practical purposes would be to know whether these p drugs that, when actually administered or rather p drugs under consideration whether they are same with respect to time to recovery or not. So, we would be interested in the following problem let me write it. So, we may be interested in testing the hypothesis that the p different drugs the p different drugs are no different as regards to relief time. So, that is the setup of the problem we have p drugs, and we are trying to see whether the drugs differ among themselves or there is no significant difference at all between the p drugs that were administered for this particular illness,

(Refer Slide Time: 53:26)



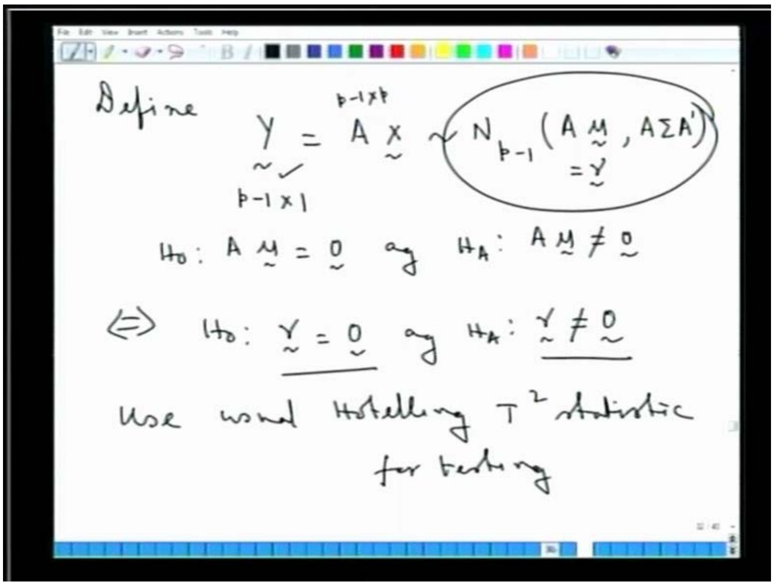
Now, how to frame this particular problem? We would require some assumptions in order to test this particular problem; we will assume that this X, which is the time to recovery for the p different drugs these are p random variables here. So, X 1 is the random variable which denotes the time to recovery for drug number one x two denotes the random variable; denoting the time to recovery for the second drug and so on, X p denotes the time to recovery for the p th drug.

So, we will assume that this p dimensional random vector this has got a multivariate normal distribution n p mu sigma, where each of these elements here mu 1, mu 2, mu p; they are the mean time to recovery for the p drugs. Now, what is the null hypothesis in terms of this multivariate normal distribution assumption the null hypothesis would be mu 1 equal to mu 2 equal to mu p; this is to be tested against the alternate hypothesis H A, that they are not equal that is the alternate hypothesis is not H naught.

Now, in terms of what we have framed here, I will be able to write this in terms of an alternate hypothesis that, I will say for the moment, it is a mu equal to a null vector against the alternate hypothesis. Let me, write first as H naught prime, and H A prime against the alternate hypothesis that A mu is not equal to a null vector; what is A p by p minus 1 by p matrix of constants; which is going to be given by the following: that the first entry here is 1, 0; here minus 1 here. So, that when I multiply this A by mu? What we will be at the first element is mu 1 minus mu p; that is equal to 0. So, that we will be able to say from this first line that mu 1 is equal to mu p, the second would be mu 2 equal to mu p, and likewise mu 1 equal to mu 2 mu 1 equal to mu p mu 2 equal to mu p and so on. The last, that is p minus 1th row of this matrix of constant this will be minus 1, and the previous element would be 1.

So, this will lead us to, if we look at, what is a mu equal to null? This would give us mu 1 equal to mu p, mu 2 equal to mu p, mu p minus 1 equal to mu p. So, all of them basically are equal. So, mu 1 equal to mu 2 equal to mu p; this hypothesis is translated in terms of this hypothesis. So, with this as, we will have this particular problem being carried forward by defining a new set of random variables; a random vector Y is defined as A times X; now this is our p minus 1 by p matrix of constant, and hence this Y vector is, now A p minus 1 dimensional random vector.

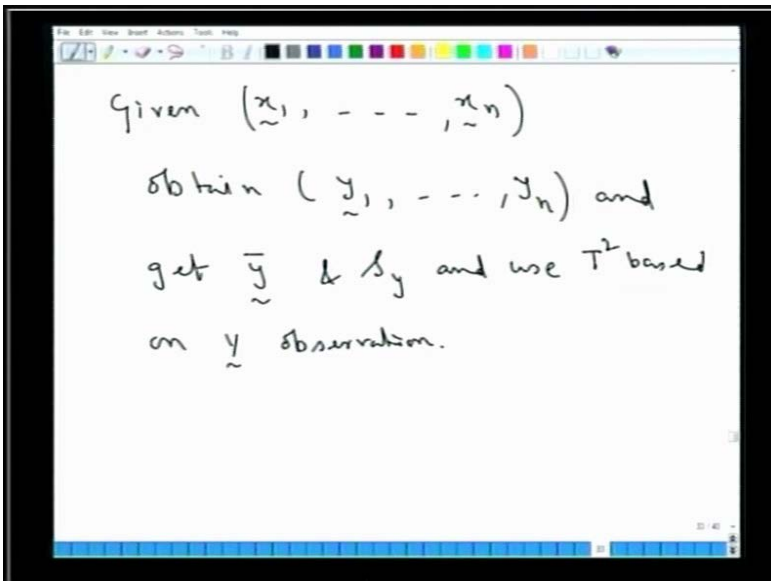
(Refer Slide Time: 56:36)



Now, this is going to have a multivariate normal distribution n p minus 1, and with the mean as A mu, and a covariance matrix, as S sigma A prime. Now, if that is there let us, denote this by equal to a vector gamma; now this H naught that A mu equal to null vector to be tested against the alternate hypothesis H A; that A mu is not equal to a null vector is thus equivalent to this hypothesis. In terms of the y random variables, Y random vector, this gamma is equal to a null vector against the alternate hypothesis; that gamma is not equal to a null vector.

So, we have cleverly translated the given problem in terms of this random p minus one dimensional random vector Y, which is now having this as it is multivariate normal distribution, and we are going to test that the mean of that particular multivariate normal distribution is a null vector, against this particular null that gamma is not equal to a null vector, and the testing for that would be carried forward using exactly the previous approach, which we had in the previous setup, that X follows a multivariate normal with a mean vector mu, and some variance, covariance matrix. We If we are interested in testing, the mean vector to be a null vector against that it is not then use we can use the usual hotelling's T square statistic for testing.

(Refer Slide Time: 58:54)



So, from the given problem what we had was this in terms of x, we will translate this particular problem in terms of y that is possible, because given this set x 1, x 2, x n, given this x one, x two, x n, that is what was the original random sample obtained this y 1, y 2, y n, which is going to have a dimensionless; now y 1 y 2, y n, and based on this y 1, y 2, y n; we can get this y bar vector, and this S Y matrix. The sample variance, covariance matrix that is based on y 1, y 2, y n, and use T square based on y observations, and that completes the testing for this particular problem; in the next lecture, what we are going to see is to look at some important properties of hotelling's T square statistic. Thank you.