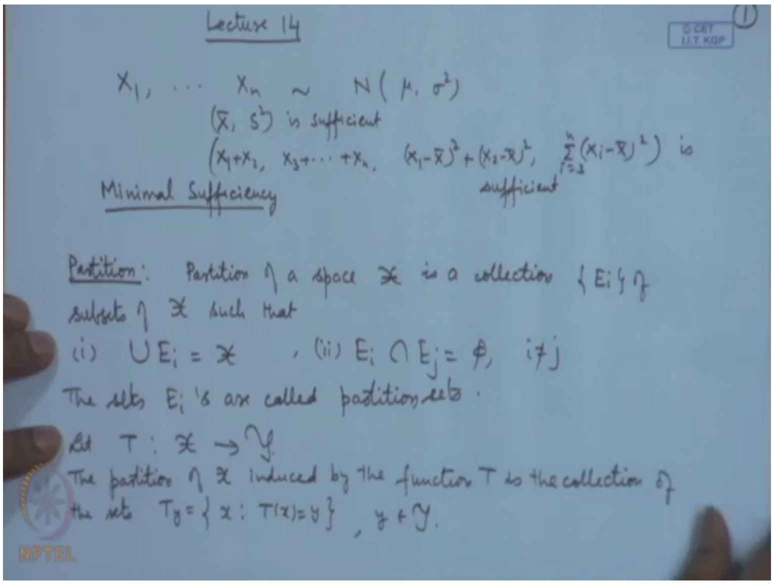
**Statistical Inference**

**Prof. Somesh Kumar**
**Department of Mathematics**
**Indian Institute of Technology, Kharagpur**

**Lecture – 27**
**Minimal Sufficiency, Completeness – I**

We have considered the concept of Sufficiency and I related to the Fisher's information major. We showed that if a statistic is sufficient then the information contained in the sufficient statistic is the same as the information contained in the whole sample. However, we are also seen that for a given problem there can be various sufficient statistics.

(Refer Slide Time: 00:53)



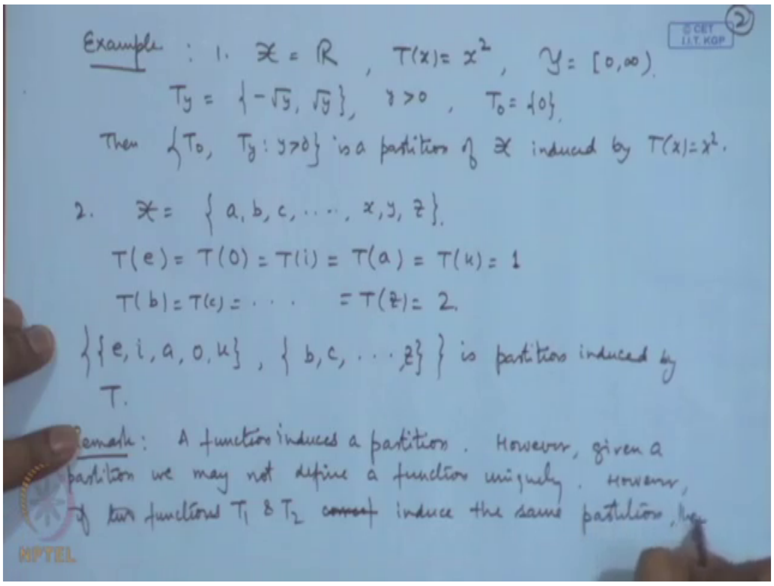

Suppose I consider a random sample X 1 X 2 X n, say X 1 X 2 X n from say normal mu sigma square then by using factorization theorem we showed that X bar and S square is sufficient. But, using the same argument we may also say that X 1 plus X 2 X 3 plus and so on X n and similarly we can say X 1 minus X bar square plus X 2 minus X bar square and so on. Say X i minus X bar whole square i is equal to 3 to n; this is also sufficient and like that we can write several sets of sufficient statistics.

So now, which one to use? Then naturally it should occur that the one which leads to the maximum reduction of the data should be used that leads to the concept of minimal sufficiency. However, this concept I will introduce through the concept of partition. So, let me introduce the concept of partition.

So, a partition of a space say x is a collection say E i of subsets of x such that union of E i is equal to x and E i intersection E j is equal to phi for i not equal to j. That means, the it is mutually exclusive sets and the union is equal to the full that means, mutually exclusive and exhaustive subsets of a space that is called a partition. The sets E i they are called partition sets.

Let T be a function from say x into another space y then the partition of x induced by the function T is the collection of the sets say T y which is defined as the set of all those values x such that T x is equal to y. That means, corresponding to distinct values of y look at the inverse image set and each of this set when you consider then that forms a partition. Let me explain through an example here.

(Refer Slide Time: 04:11)

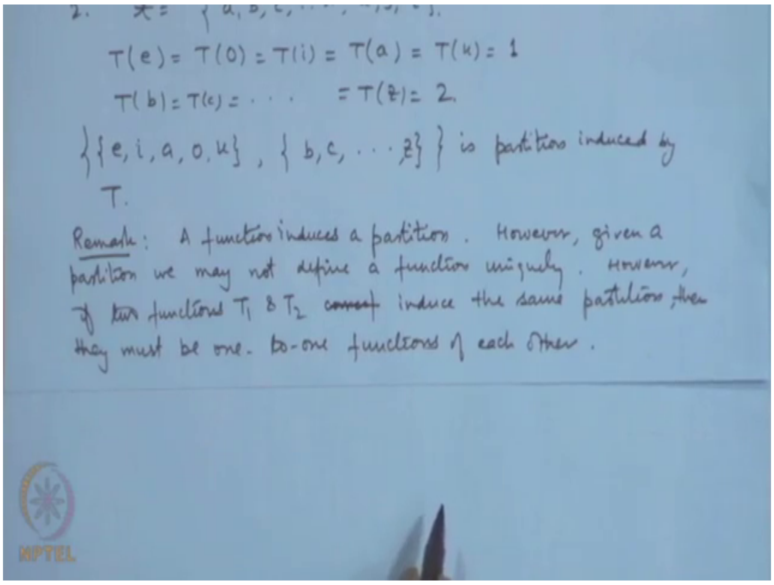



Say I consider x to be the set of real numbers and let us considered the function T x 2 be say x square. Then what is y then? y is the 0 to infinity. Then if you consider say T y then it will be equal to minus root y root y for all y greater than 0 and T 0 is of course, 0. Then T 0 and T y for y greater than 0 this is a partition of x induced by T x.

Let me give another example. Suppose, I considered say x to be the set of say alphabets; a, b, c, d and so on up to x, y, z then I define say T of say e T of o T of i and say T of a u say is equal to for example, I write 1 and T of say remaining things remaining characters like b c and so on T of z is equal to say 2. Then, if I considered the inverse images of 1 and 2 respectively, then, what you will get? e i a o u and another set will be remaining b c and so on then this is the partition induced by this function T.

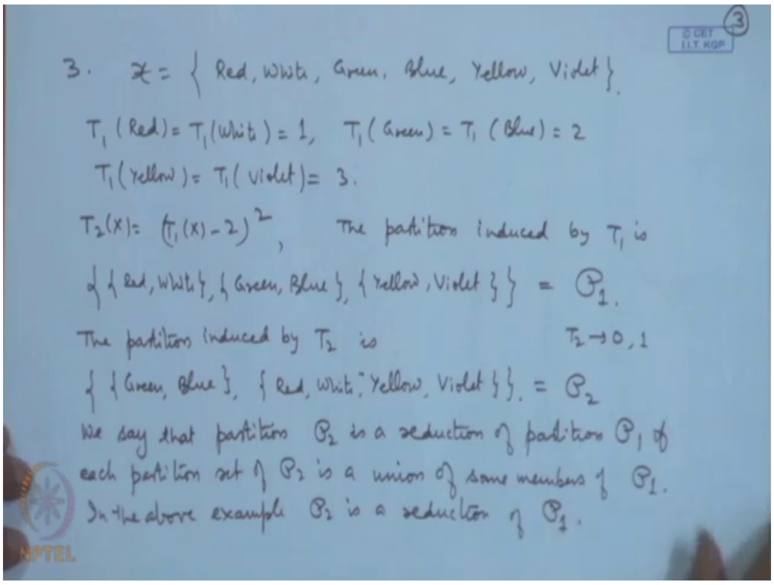
Now we can say that a function induces a partition, but given a partition you may not be necessarily able to define a function uniquely. However, given a partition we may not define a function uniquely.

(Refer Slide Time: 07:45)




However, if 2 functions say T 1 and T 2 induced the same partition then they must be one to one functions of each other. Let me take another example.

(Refer Slide Time: 08:09)



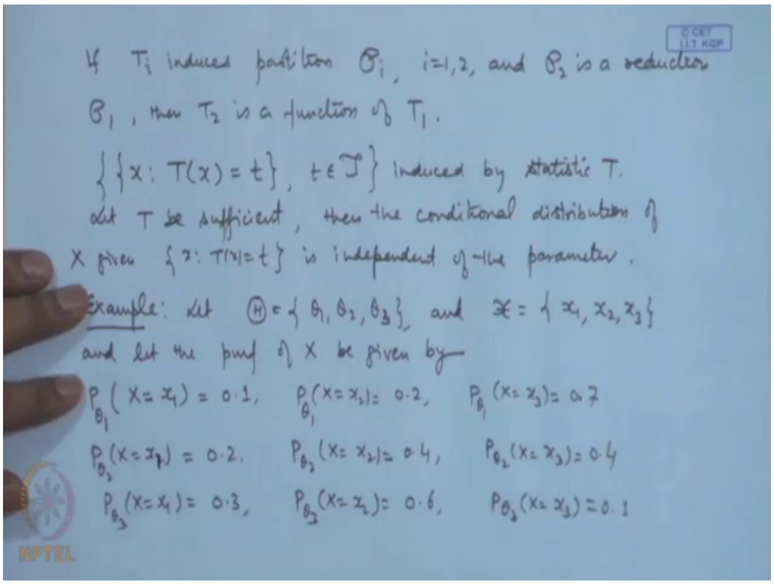

Let us consider say x as the colors red, white, green, blue, yellow and say violet. I define a function T 1 of say red and T 1 of white, I assign the value say 1. Say T 1 of green and T 1 of blue is equal to say 2. T 1 of yellow and T 1 of violet supposed I define it to be 3 and I define say T 2 X is equal to T 1 X minus 2 whole square. Now, if you define this then corresponding to T 1 value is equal to 1, T 2 will become 1 and corresponding to T 1 is equal to 3, also T 2 will become 1 corresponding to the value T 1 is equal to 2 T 2 will be 0.

So, the partition induced by the partition induced by T 1 is red white in one set, green and blue in another set and yellow and violet in another set. Because, the value 1 corresponds to red and white, the value 2 corresponds to green and blue the value 3 corresponds to yellow and violet. So, let me call this partition P1. Let us also find out the partition induced by T 2. Now, let us see. T 2 is taking values T 2 can take value 0 that is 1 T 1 is equal to 2 and when T 1 is equal to 1 or 3, T 2 is taking value 1. So, the value 0 is corresponding to green and blue.

So, the partition which is induced by T 2; one set consists of green and blue and 1 is correspondent to then red, white, yellow and violet; red, white, yellow and violet. Let me call this partition P2. So, we say that partition P2 is a reduction of partition P1 if each partition set of P2 is a union of some members of P1.

Now, here you observe the partition sets in P2 this set is the same as the set and the second set here is a union of 2 sets of P1. So, in the above example P2 is a reduction of P2 P1. Now from here we can make out that if the partition induced by T 2 is a reduction of the partition induced by T 1 then T 2 is a function of T 1.

(Refer Slide Time: 12:25)

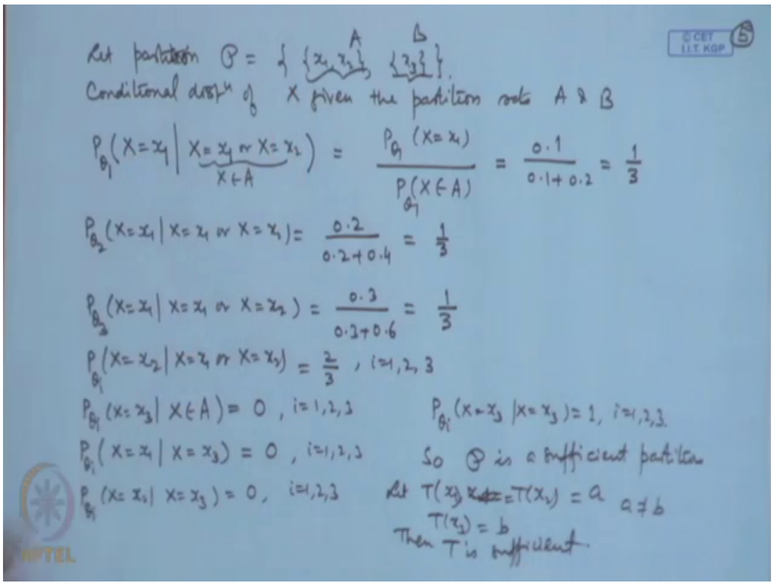



So, we can make this statement that if T i induces partition say P i for i is equal to 1, 2 and P2 is a reduction of P1 then T 2 is a function of T 1. So, now when we talk about a sufficient statistics T then the corresponding partition sets will be T x is equal to T where T values is over the set of values of T. This is the partition sets induced by statistic T. Let T be sufficient.

Now, the definition of sufficiency says that the conditional distribution of the data that is X 1 X 2 X n given T must be independent of the parameter. So, we can say that the conditional distribution the conditional distribution of X given x such that T x is equal T is independent of the party of the parameter. So, let me take one example. Let the parameter is space consists of say 3 points; theta 1, theta 2, theta 3 and let us consider say the variable space consisting of 3 points x1, x2, x3 and let the probability mass function of X be given by probability X is equal to x 1.

Now, when theta is equal to theta 1 that is equal to 0.1 probability X is equal to x 2 is equal to say 0.2, probability X is equal to x 3 is equal to say 0.7. So, when the parameter value is theta 1 this is a probability distribution. When theta 2 is the parameter value the probability distribution is given by say 0.2; 0.4 and say 0.4. And when theta 3 is the 2 parameter value, suppose the probability distribution is given by 0.3, 0.6 and 0.1. Let me define a partition here. Let us consider partition x1, x2 and x3.

(Refer Slide Time: 15:59)




Now, considered the conditional distribution of X given the partition sets. Here we have 2 sets. So, 1 set is let me call it A, another set I call B. So, what is the probability of X is equal to x 1 given say X is equal to x 1 or X is equal to x 2 that is given that x belongs to A. Now, here it will be dependent upon that means, you have to calculate this probability under different parametric configurations.

Let me consider say theta is equal to theta 1. When I take probability of X is equal to x 1 given x belonging to A then we apply the conditional probability formula. this will be equal to probability of X is equal to x 1 divided by probability x belonging to A that is when X equal to x 1 or X equal to x 2. Now, when theta is equal to theta 1 probability of X equal to x 1 is 0.1 and probability of X equal to x 2 is equal to 0.2. So, this value turns out to be 0.1 divided by 0.1 plus 0.2 that is equal to 1 by 3.

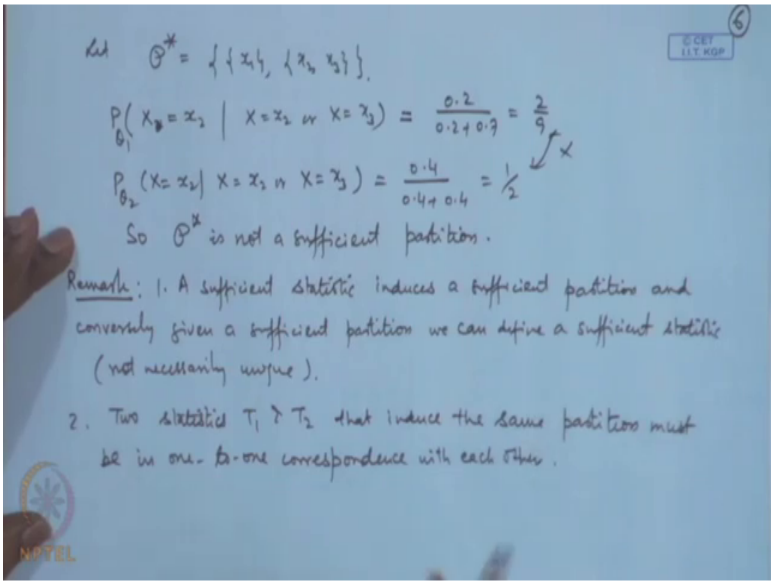
Now, if I consider say theta 2 X is equal to x 1 given X equal to x 1 or X equal to x 2; again you notice here when theta is equal to theta 2 the corresponding probabilities are 0.2 and 0.4. So, this value will turn out to be 0.2 divided by 0.2 plus 0.4 that is once again 1 by 3. And in a similar way if I calculate when theta is equal to theta 3 this is turning out to be 0.3 by 0.3 plus 0.6 that is again 1 by 3. So, what we have found that the conditional probability of X is equal to x 1 given x belongs to A is free from theta. However, we need to check other configurations also; that means, what is the probability of X is equal to x 2 given X equal to x 1 or X equal to x 2.

Once again when theta is equal to theta 1 this value will become 0.2 divided by 0.1 plus 0.2. When theta is equal to theta 2 it will become equal to 0.4 divided by 0.2 plus 0.4. When theta is equal to theta 3 this probability will be equal to 0.6 divided by 0.3 plus 0.6 that means, it is equal to 2 by 3 in all the cases. That means, it is free from the value of theta. Similarly, if I consider X is equal to x 3 given X belonging to A then this is going to be 0 because when you take the numerator probability of X equal to x 3 intersection X belongs to A that is going to be 0.

So, this is equal to i is equal to 1, 2, 3. Then similarly, if I consider probability of X is equal to x 1 given x belonging to B that is X is equal to x 3 then that is also 0 and this will be true for all i is equal to 1, 2, 3. If I consider probability of X is equal to x 2 given X equal to x 3 that is also going to be 0 for i is equal to 1, 2, 3. If I consider theta i X is equal to x 3 given X equal to x 3 that is going to be 1 for I is equal to 1 2 3. So, we can say that this is a sufficient partition.

So, if I consider a function T which is assigning one value to say T x 1 and T x 2 as some value and T x 3 as another value let me call it value a and this value as b where a is different from b then T must be sufficient. That means, a function which partitions the variable space into 2 parts, one part is consisting of x 1 and x 2 and another one is consisting of x 3 then that will be sufficient. Just to tell that this is not a unique way of looking at it for example: in plays of x 1, x 2, x 3, where I have clubbed x 1 and x 2, suppose I had clubbed in a different way.

(Refer Slide Time: 21:21)




For example, if I had clubbed let me consider another partition, let me call it P star. Suppose I had put x 1 and say x 2, x 3. Let us see what is the probability of say x 2 given X is equal to x 2 or X is equal to x 3. Now, when theta is equal to theta 1 this probability will be equal to now x 2 given, x 2 is equal to x 3. So, this will become equal to 0.2 divided by 0.2 plus 0.7 0.2 divided by 0.2 plus 0.7 that is 2 by 9.

Now, if I consider theta 2 same probability then for theta 2 these values are 0.4 and 0.4. So, this value will be turning out to be half, this is not same as this. So, P star is not a sufficient partition. Consequently any statistic which will induce this partition will not be sufficient. Let me a sufficient statistic induces a sufficient partition and conversely given a sufficient partition we can define a sufficient statistics. Of course, this is not necessarily unique.

And also 2 statistics T 1 and T 2 that induce the same partition must be in one to one correspondence with each other.