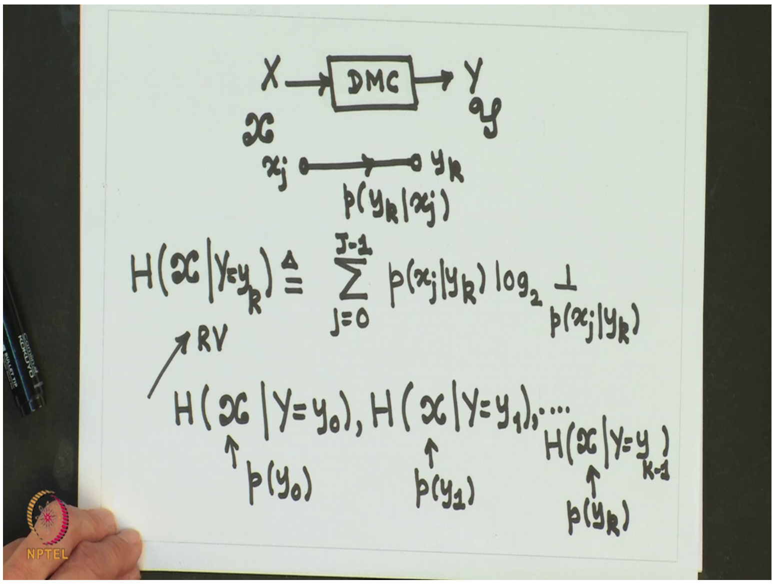
Principles of Digital Communications
Prof. Shabbir N. Merchant
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Lecture - 07
Channel Capacity - I

We are studying information transmission on a communication channel and to start with we have considered a Discrete Memoryless Channel.

(Refer Slide Time: 00:36)



So, a discrete memoryless channel is a statistical model which has an input symbol selected from the alphabet; that is input alphabet. It has an output symbol selected from the output alphabet. And we have transition or conditional probability of an output symbol given an input symbol.

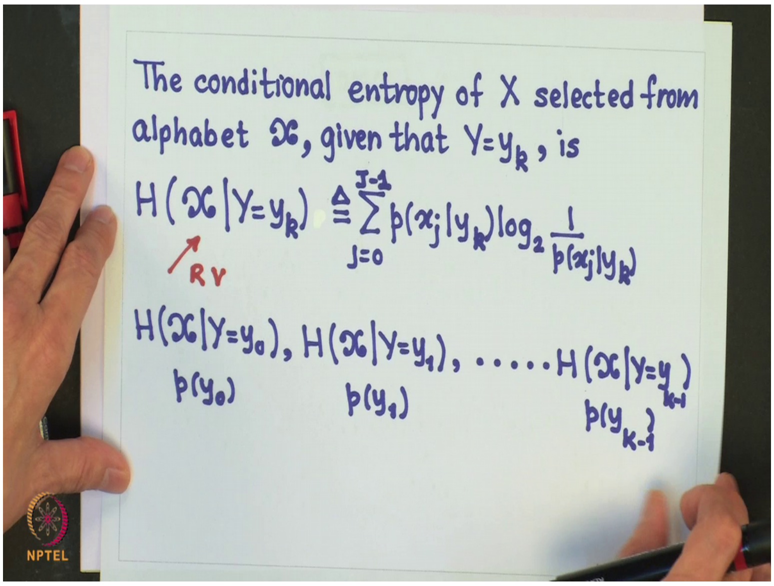
Now, the goal of transmission basically is to learn something about the input symbol or gain information about the input symbol given that we have observed an output symbol. So, similar to the reasoning which we had done earlier for the definition of information measure we can extend it here as follows. The uncertainty which I have about the input symbol given that I have observed the output symbol could be defined as 1 by probability x j given y k and log of this quantity with the base 2.

Now, this is a random variable and we can find out what is the average uncertainty about the input symbol given that we have observed Y is equal to y k that can be achieved by multiplying by conditional probability and summing this for j is equal to 0 to the size of the alphabet input alphabet that is capital J.

So, this will be the average uncertainty about the input symbol x given that we have observed the output symbol to be Y is equal to y k and this we define as the conditional entropy of the input given that I observed output symbol to be y k. So, if you look at this itself is a random variable and it takes different values as follows. We are presuming that the output alphabet is of size k and each of this occurs with the probability here. This will be the probability of y 0 this will occur with the probability y 1 and this will occur with the probability y k.

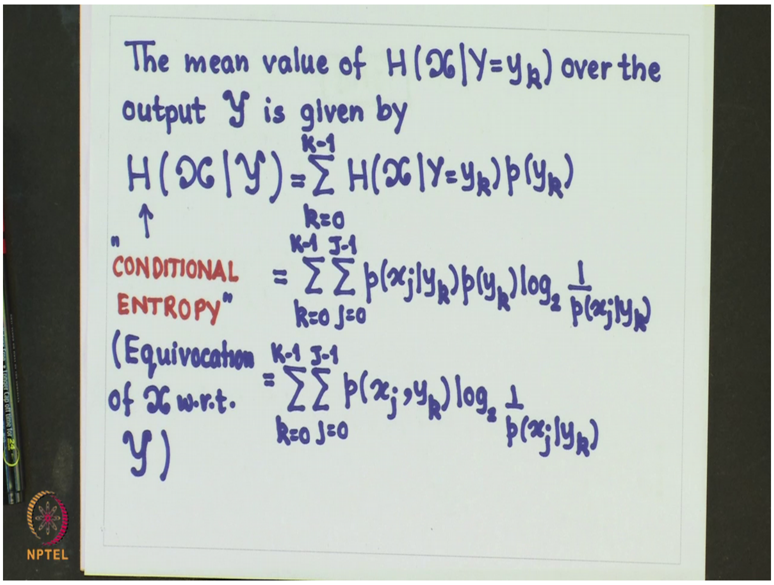
So, now, we can compute the average value of this random variable and if we do that we will get the mean value of this random variable over the output alphabet Y, correct.

(Refer Slide Time: 05:19)



So, the conditional entropy of X selected from alphabet x, given that Y is equal to y k is written as it this way define and then these are being random variable I can sum it over the output alphabet and what I will get is as follows.

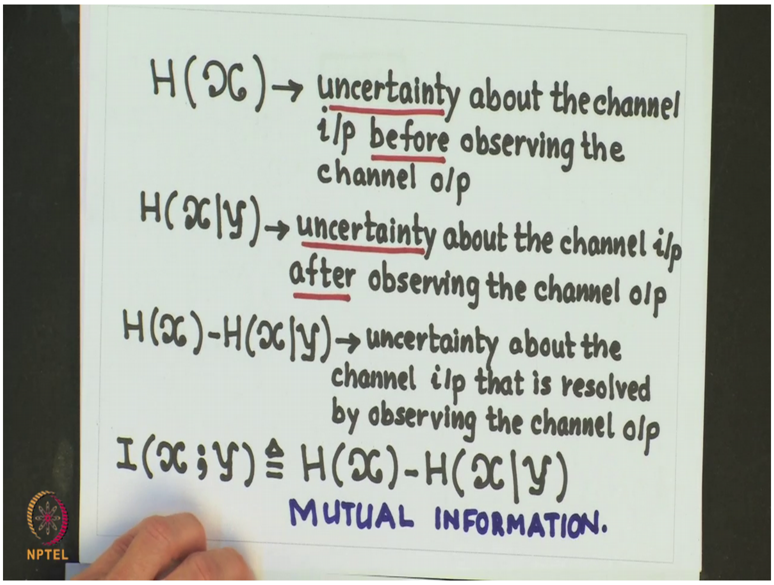
(Refer Slide Time: 05:30)



So, the earlier quantity I take it I multiplied by the probability of the output symbol, correct and I will get the average value of this, correct and this is basically known as conditional entropy of x given y and this we can write it as shown here in this equation simple and this can be simplified using the base rule. This multiplied by this conditional probability multiplied by the marginal probability will give us the joint probability distribution of x j and y k.

So, let us look at the physical interpretations of this entropies and conditional entropy. To be particular, this is also known as equivocation of x with respect to y.

(Refer Slide Time: 06:47)

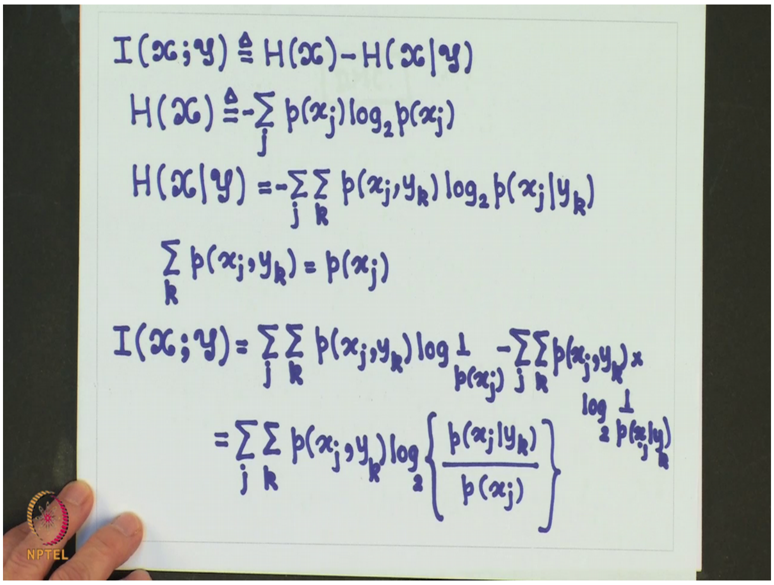


So, we had input to the channel the alphabet x and if we calculate the entropy of that that is nothing, but uncertainty about the channel input before observing the channel output. Next, we calculated this quantity which is known as conditional entropy. What does that mean? Physically it means this is the average uncertainty about the channel input after observing the channel output.

So, now if we take a difference between the two, this minus this, correct, that should be the uncertainty about the channel input that is resolved by observing the channel output, correct. So, this looks very intuitively satisfying this quantity which is the difference between the entropy of the input alphabet and the conditional entropy of the input alphabet given the output alphabet is defined as I x semicolon y like this and this is known as mutual information.

Let us rewrite this mutual information in a different manner.

(Refer Slide Time: 08:42)

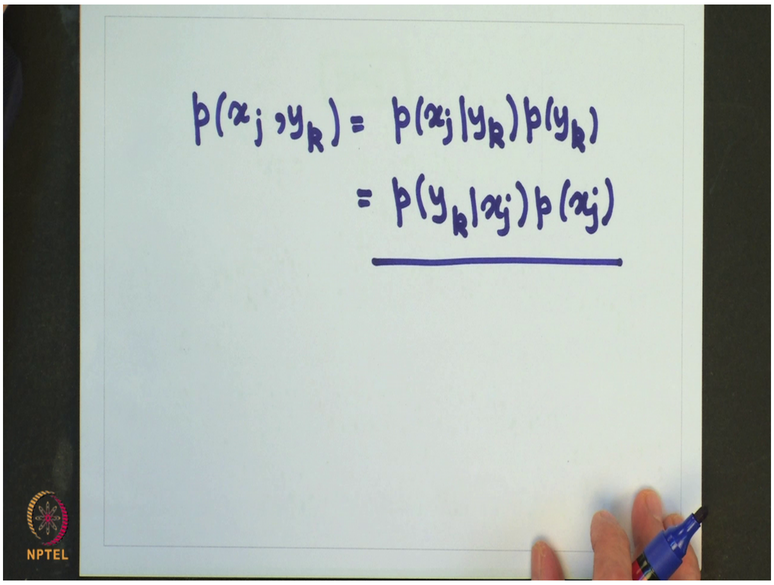


So, let us look at this slide. I have by definition of a mutual information between x and y as follows this is by definition. Now, if I we know that the entropy of the source is given by this expression. Please note that, I have put a minus sign because I have taken the denominator p x j into the numerator and the conditional entropy which we have just defined. We showed that that quantity is equal to this quantity. Again, please note that, I have taken the denominator term into the numerator and that is why we get the minus sign and just for the sake of simplicity I am avoiding writing this summation from j equal to 0 to j equal to capital J minus 1 and similarly, here the k is equal to k equal to 0, summation up to capital k minus 1, ok. It is understood when I write this.

So, and if we use this relationship from the probability that if I take the joint probability and I sum it over k, then I will get the marginal probability for the symbol x j using this inequality we can write H x as this form. So, instead of p H z summation p x j over k, I have substituted that basically with joint distribution summit over k will give me p x j and then this quantity will give me the entropy the source and this quantity is basically here, I repeat it, this is here multiplied by this term, ok, fine.

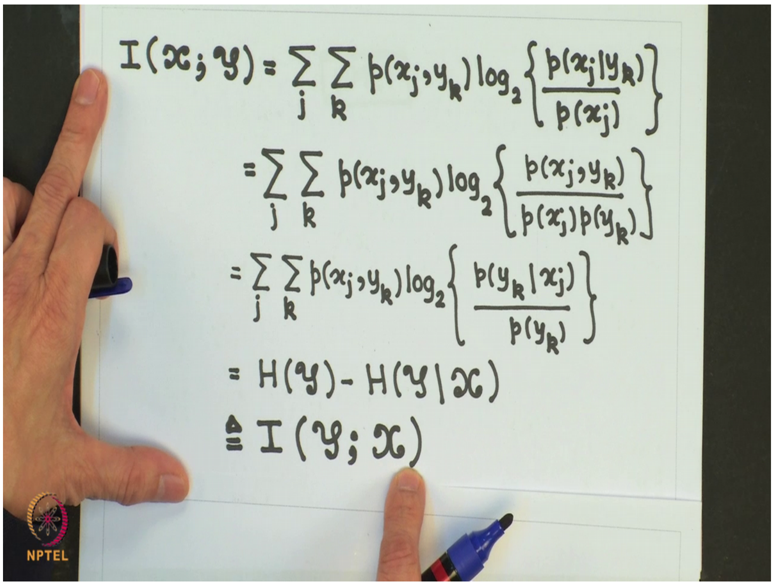
So, if we just simplify this what we get is, this term, correct. It is very simple to understand that this and this will give H x and this with this will give you conditional entropies, ok fine. Given this relationship and now, we use the Bayes' rule.

(Refer Slide Time: 11:22)



Bayes' rule says that the joint probability is equal to this expression or is also equal to this expression. Based on this Bayes' rule I can rewrite my mutual information as follows.

(Refer Slide Time: 12:05)

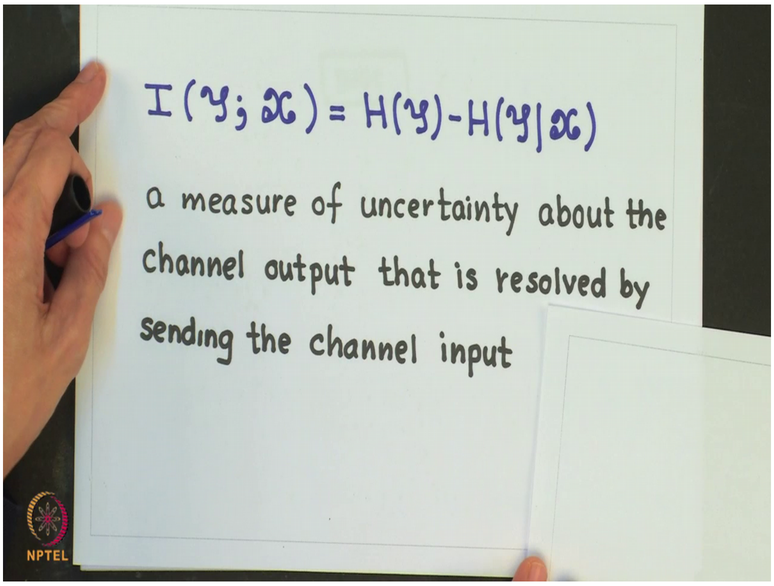


So, this is my definition for mutual information now what I do is basically I will multiply the numerator and denominator by probability of y k and here also I multiplied by probability of y k. If I do that basically what I get here after multiplying this term by probability of y k I get the joint probability of x j y k.

So, I can again using the Bayes' rule I can rewrite this term out here as probability of y k, x j multiplied by probability of x j. So, probability of x in the numerator will get cancelled with the probability of x j in the denominator and what I will be left with is probability of y k given x j upon probability of y k. Now, this by definition is equal to entropy of alphabet y minus conditional entropy of alphabet y given alphabet x, correct. So, this will give me this term and this term will give me this term ok, fine. So, this is equal to neutral information between y and x.

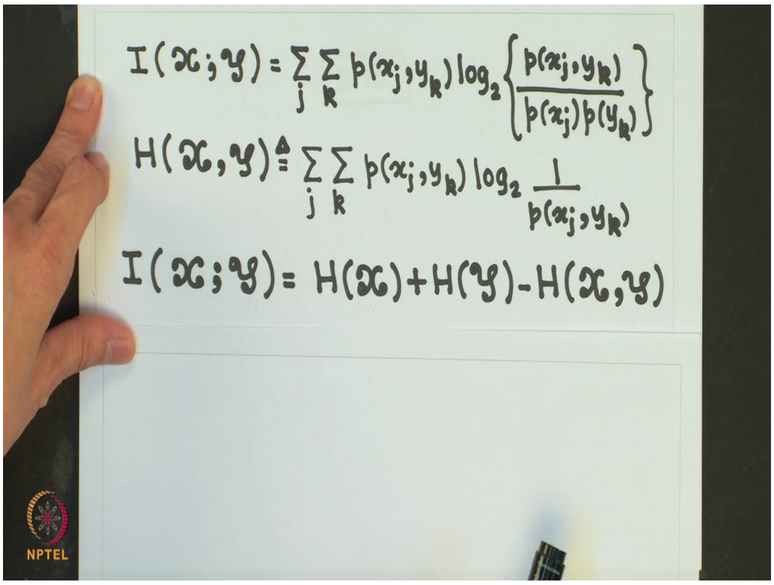
So, what I have shown you basically that mutual information between x and y is same as the mutual information between y and x. So, symmetry condition is satisfied by mutual information definition.

(Refer Slide Time: 14:10)



So, what is this, actually how do you interpret this? I can interpret this mutual information between y and x as a measure of uncertainty about the channel output that is resolved by sending the channel input, correct. So, this is the physical interpretation which we could give.

(Refer Slide Time: 14:44)

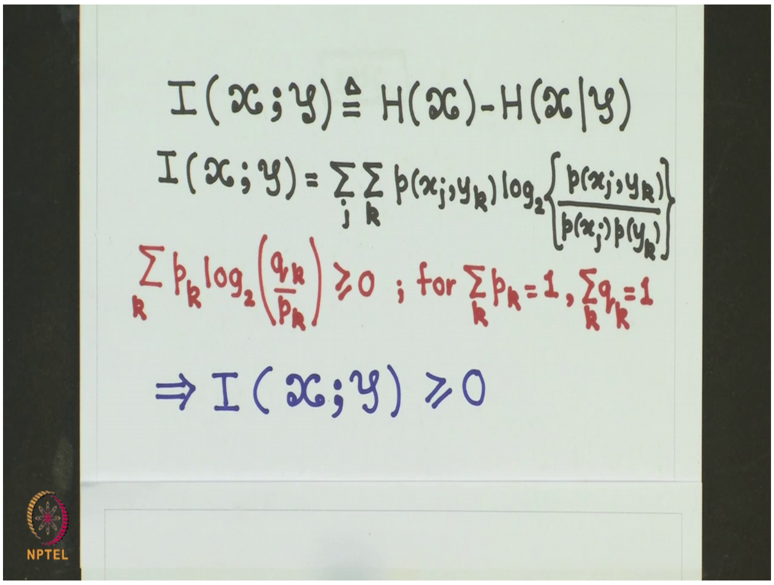


Now, let us look at another interesting property of the mutual information. Again, I go back to the mutual information definition and one of the form in which mutual information can be written is we just saw this is given by this expression. So, if you look at this expression I can rewrite this in a different form. If I define joint entropy of two alphabets x and y as follows so, if you look at basically the joint entropy of alphabet x and y is given by this expression, correct. It is intuitively satisfying based on the reasoning which we did for the definition of information major, correct.

So, what this means basically that I can rewrite this expression as entropy of x, this and you take with this term, this you take with this term you get entropy of y and this basically if you take with this term I will get minus is a joint entropy.

Another interesting property of mutual information follows.

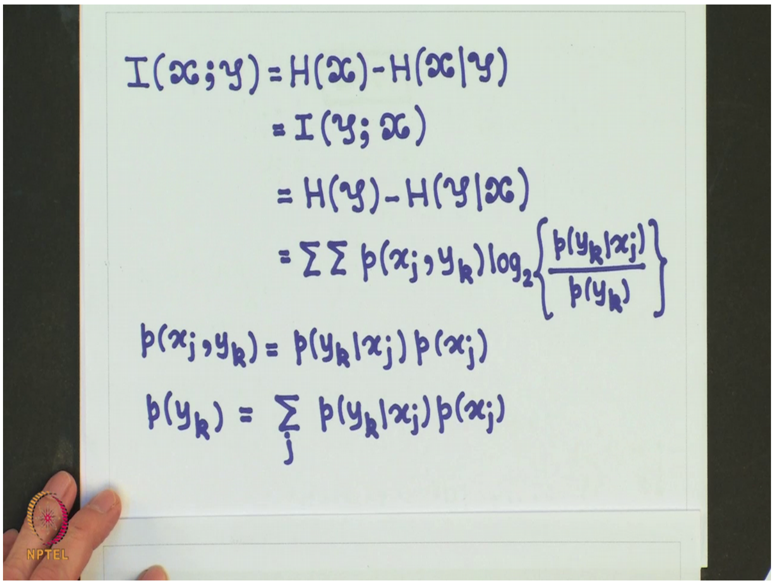
(Refer Slide Time: 16:00)



Again, I know I have defined my mutual information to be of this form, correct. Now, if you see that basically I can rewrite this as this term, correct. We have seen this earlier. Now, if you look at this form out here this form is very similar to the inequality which we derived earlier in this course we have seen that summation p k log to the base 2 of q k by p k is always greater than equal to 0 if p k and q k satisfy this constraint.

So, in this framework if you just take the similarity between this equation and this equation we see that the mutual information is always greater than or equal to 0. So, what it implies basically or observing the output my uncertainty about the input can become lower, but it can never increase and that is why intuitively also it is satisfying that this mutual information will be greater than or equal to 0. It will be equal to 0 when the joint probability x j, y k is equal to probability of x j multiplied by probability of y k; that means, x j and y k are independent, and that also looks intuitively satisfied ok.

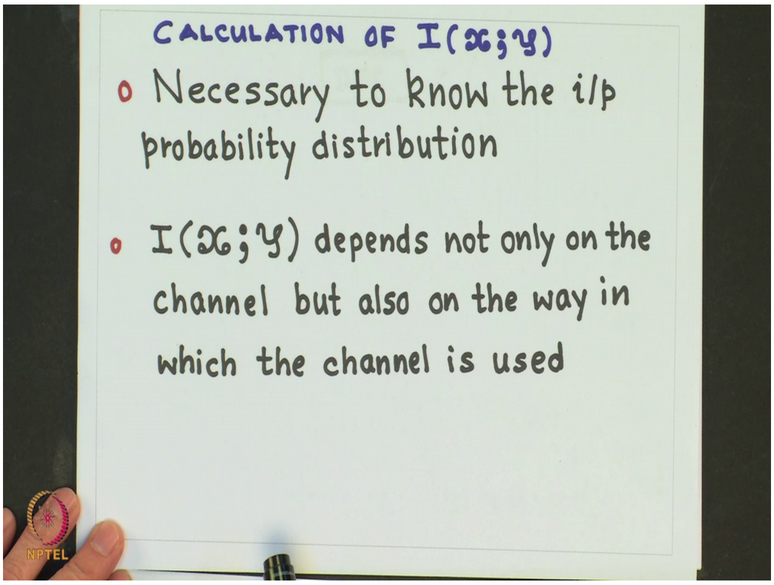
(Refer Slide Time: 17:50)



Now, if you look at the definition of mutual information which I repeat here and this is also equal to mutual information between y and x and this I can rewrite it as this. All this we have proved in the earlier slides. Now, what happens the advantage of writing mutual information between x and y in this form is that it is more convenient from the point of view of calculation of mutual information because the channel is specified by the conditional probability of the output symbol given the input symbol and so, if we see that basically this equation becomes this expression based on this.

And, now, look at how do I compute joint probability x j, y k. This can be computed by using the Bayes' rule as follows and then if you look at the marginal probability p y k can be computed from the conditional probability of p y k given x j and the probability of x j, correct. You also see that for the calculation of mutual information it is not only dependent on the channel in the form of this probability of y k given x j, but it is also dependent on the way in which we use the channel; that means, this probability x j s, correct.

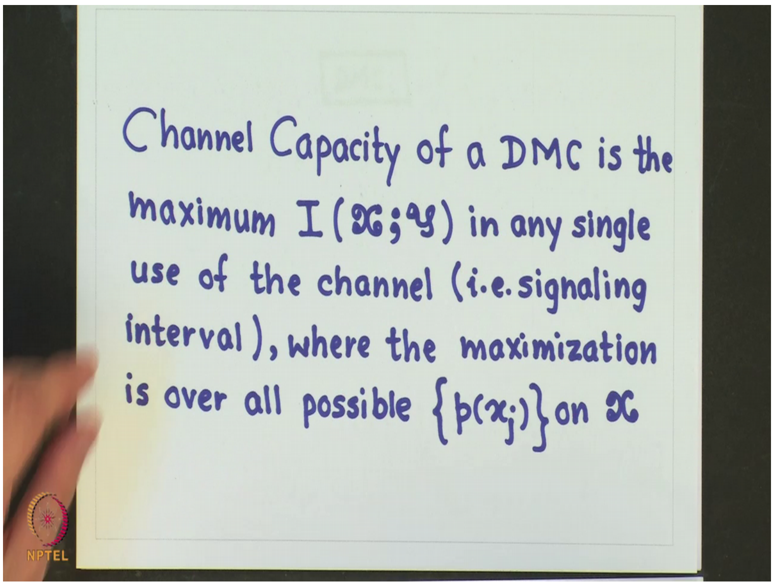
(Refer Slide Time: 19:43)



Note these two points that mutual information requires. The input probability distribution and the value of this mutual information is not only dependent on the channel, but on the way the channel is used.

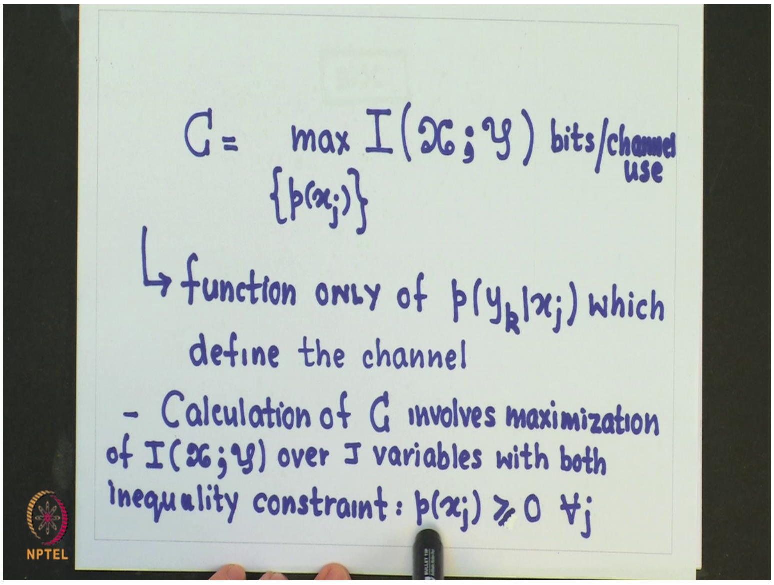
So, now we define what is known as a channel capacity as follows.

(Refer Slide Time: 20:07)



Channel capacity of a discrete memoryless source is the maximum value of the mutual information in any single use of the channel; that is, the signaling interval where the maximization is over all possible, input symbol distribution on the alphabet x.

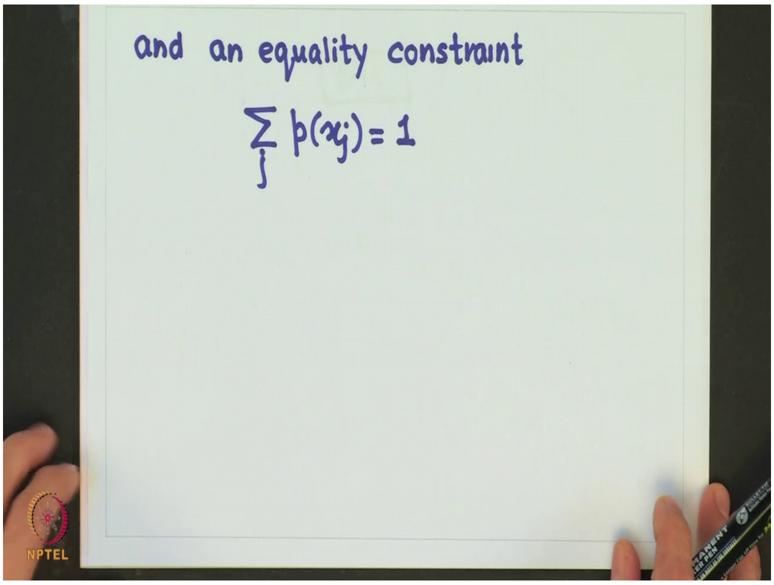
(Refer Slide Time: 20:43)



So, I can say that my channel capacity is equal to maximization of the mutual information and this maximization is over the probability distribution of the input symbols in the input alphabet and the units is basically bits per channel use.

And since this is function only of now, this is function, this once I define the channel capacity this is function only of transition or conditional probability which basically define the channel, correct. So, the calculation of this capacity involves maximization over j variables because that is the size of your input alphabet and it has to satisfy two constraint: one is the inequality constraint, which is all the probabilities should be, obviously, greater than or equal to 0 for all j and the other condition is the equality constraint.

(Refer Slide Time: 21:50)



Which says that, sum of the probabilities of the symbols in the alphabet should be equal to 1, having defined the channel capacity.

Now, we will try to evaluate the channel capacity for an important practical communication channel and that is binary symmetric channel. And we will do this in the next class.

Thank you.