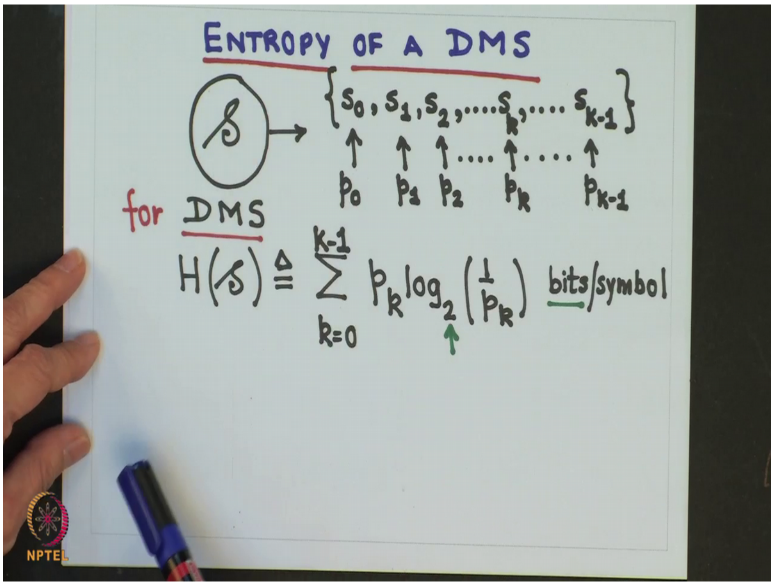
Principles of Digital Communications
Prof. Shabbir N. Merchant
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Lecture – 02
Entropy and it is properties

Hello welcome back. So, in the last module we studied what is the information source? We mathematically define the model for information source. We define the major for information and we also define the concept of entropy for a discrete memory less source. So, the Entropy of a discrete memoryles's source was defined as shown on this slide.

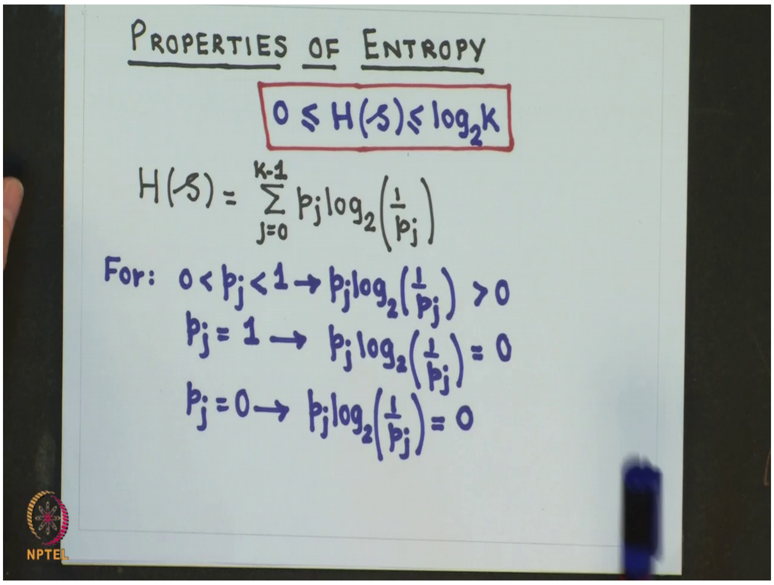
(Refer Slide Time: 00:55)



We have a source and it has source alphabet and each of this elements in the source alphabet is a single or letter, and to each of this symbol we assign the probability of occurrence.

And for discrete memoryless source the entropy turns out to be this expression, which is summation of p k log to the base 2 of 1 by p k and the units for this it is bits per symbol. So, this is the average information, which I get when I observe the output symbol from the source s. Now, today basically we will have a look at the properties of this entropy.

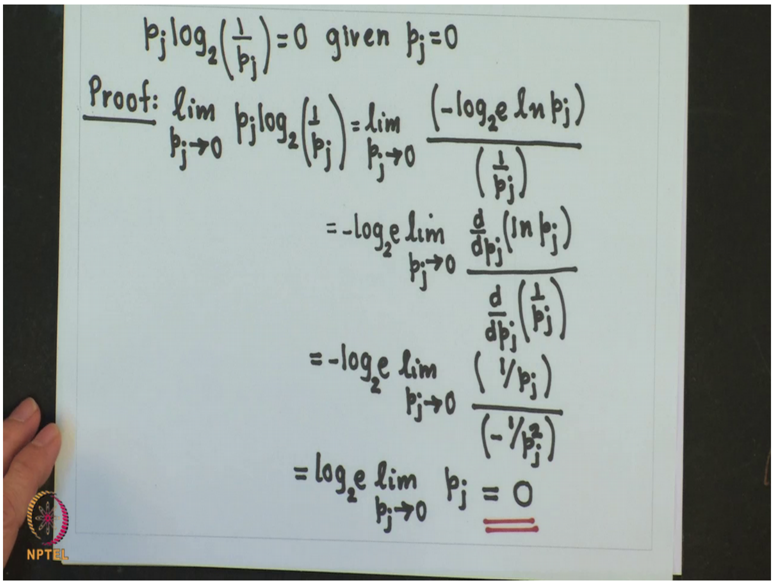
(Refer Slide Time: 02:00)



So, we will study this property of the entropy, which says that entropy is always greater than equal to 0 and it is bounded by log to the base 2 k, where k denotes the cardinality of the set alphabet or the size of the alphabet or you could say the number of symbols or letters, which are there in the source alphabet.

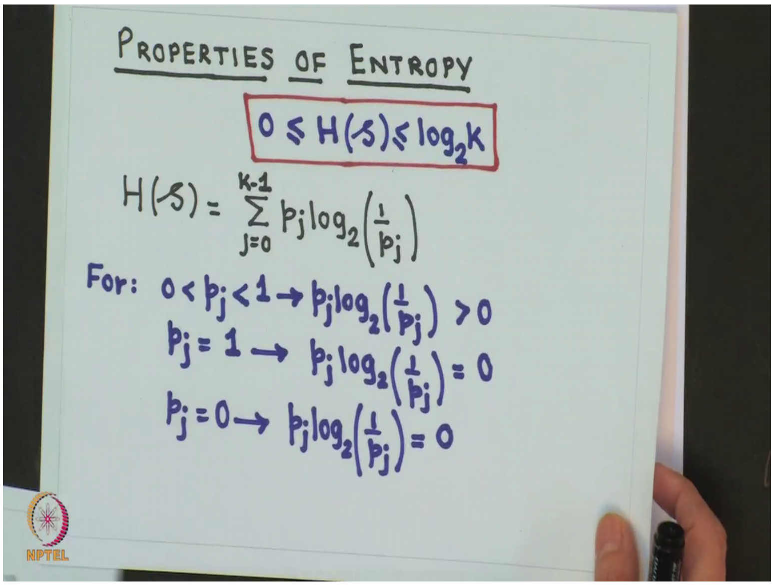
Now, we know that the expression for entropy is given here. Now, let us look at each of the term in this expression. So, if you look at this p j log 2 1 by p j, then it is not very difficult to see that for p j greater than 0 and less than 1, this quantity is always larger than 0. For p j equal to 1 these is equal to 0 and for p j equal to 0 this quantity again is equal to 0. Now, for p j equal to 0 to show this quantity is equal to 0, we can use simple the hospital rule and evaluate that.

(Refer Slide Time: 03:33)



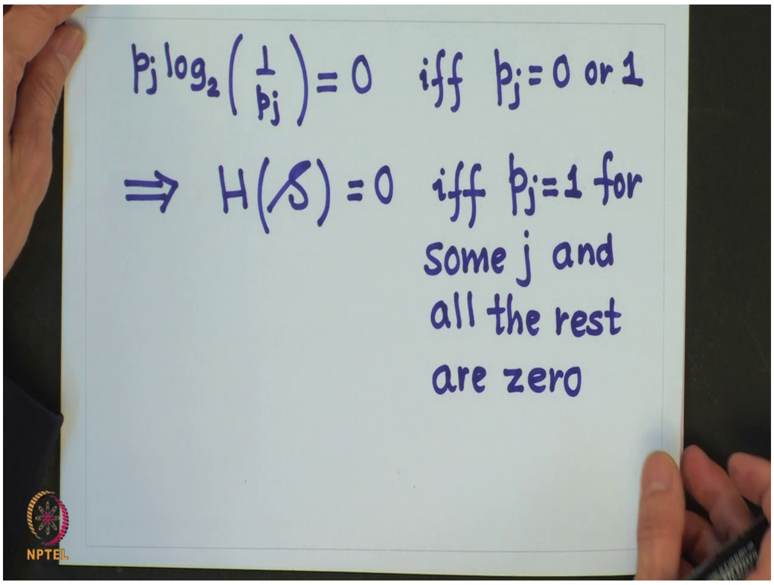
So, just I will flash that the slide to show how to do it. So, you take the limit p j tending to 0 of this quantity and you continue it very clear from this slide how to proceed? And you can show that when p j tends to 0 this quantity is equal to 0.

(Refer Slide Time: 04:00)



So, what it means basically that all the terms out here are either 0 or greater than 0. So, the summation will be; obviously, greater than equal to 0 ok.

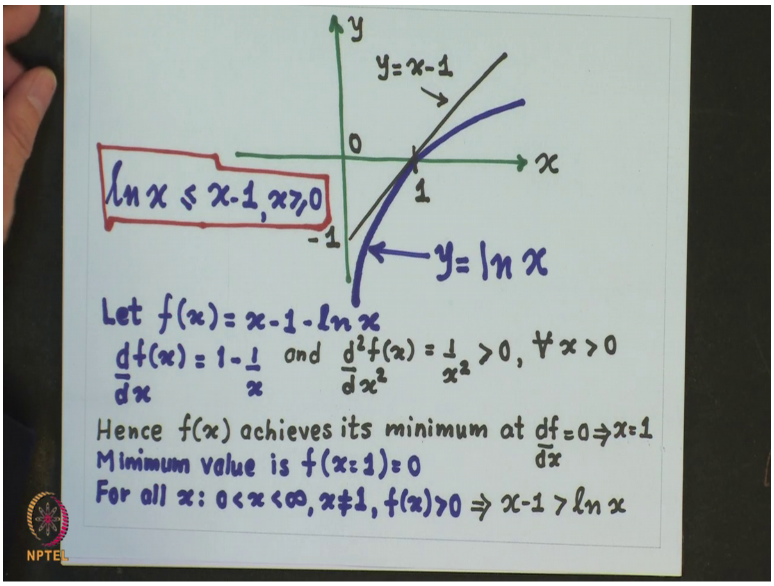
(Refer Slide Time: 04:18)



So, p j log 2 1 by p j is equal to 0, if and only if p j is equal to 0 or 1. So, this implies that entropy of the source is equal to 0, if and only if p j is equal to 1 for some j and all the rest are 0.

So, H s is equal to 0 will happen when this condition is satisfied; that means, there is no uncertainty of the output of the source ok. Now, let us try to get the upper bound ok.

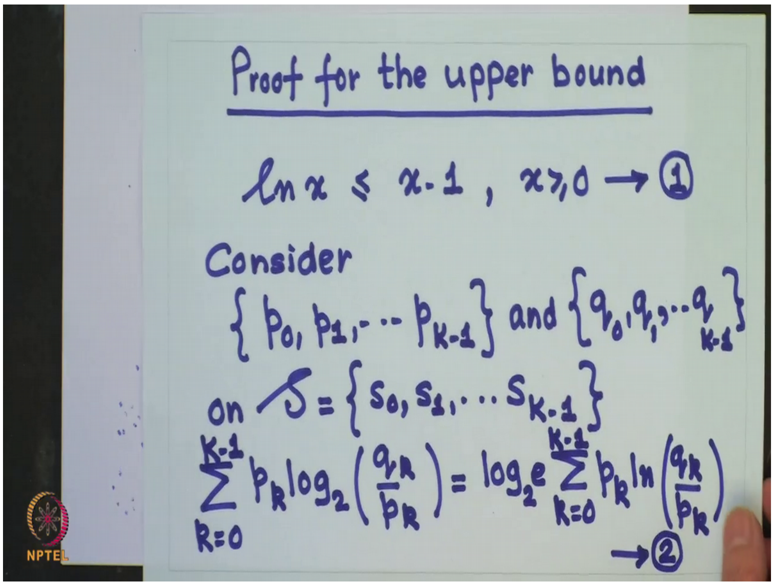
(Refer Slide Time: 05:36)



To do that basically will use this relationship between the log x and x minus y. Now, I have plotted the equation y is equal to x minus 1 and y is equal to log x. So, from this graph it is very clear and we can see that log x is always less than equal to x minus 1, for x greater than equal to 0 and they will be equal when x is equal to 1. The mathematical formal proof is also given here I can form a function as shown here you take the first derivative and it is second derivative.

And you can show that log x will be always equal less than or equal to x minus 1 correct. So, x minus 1 is always greater than log x correct this is a fundamental relationship, which will exploit in our module on information theory ok. So, based on this let us derive the upper bound for this.

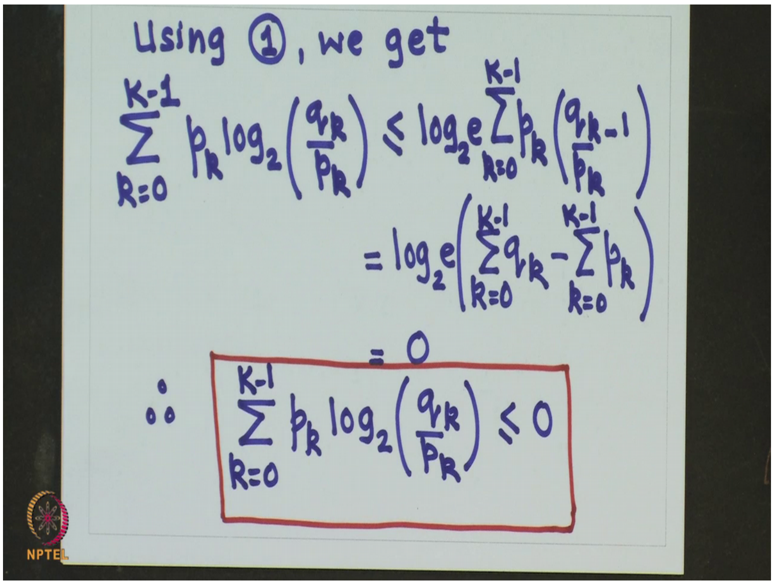
(Refer Slide Time: 06:49)



So, proof for the upper bound I will use the relationship log x is less than equal to x minus 1 for x greater than equal to 0.

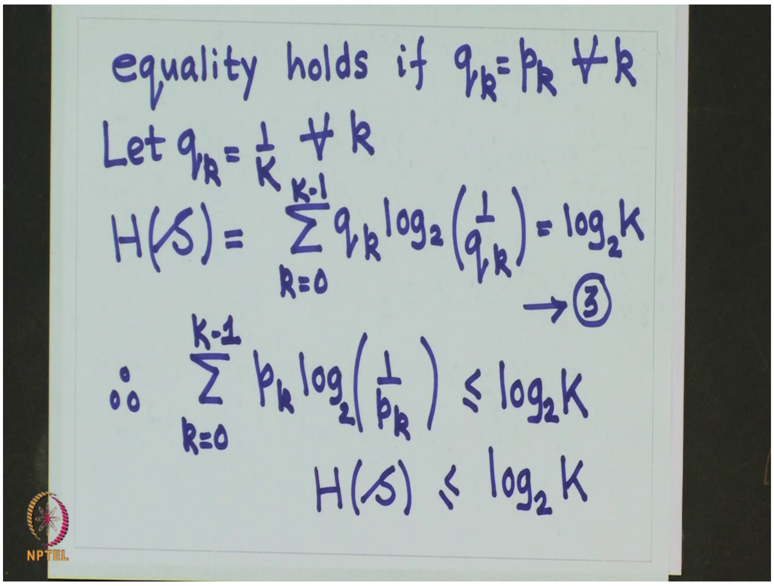
Now, consider 2 probability distributions in terms of p s and terms of q, both these distributions are defined on the source alphabet s. Now, we can write this expression p k log to the base 2 q k p k k is equal to 0 to k minus 1, as log to the base 2 of e and summation of p k log this log is to the natural base q k by p k let me call this as equation 2.

(Refer Slide Time: 09:00)



So, now using 1, we get the following relationship this is equal to less than less than or equal to ok. Using 1 this is equal to I can write it and remember these are the probability distributions q k s and p k s. So, the summations both are 1 and this is equal to 0 therefore, I get this relationship. This is a fundamental inequality, which is used very often in the literature on information theory.

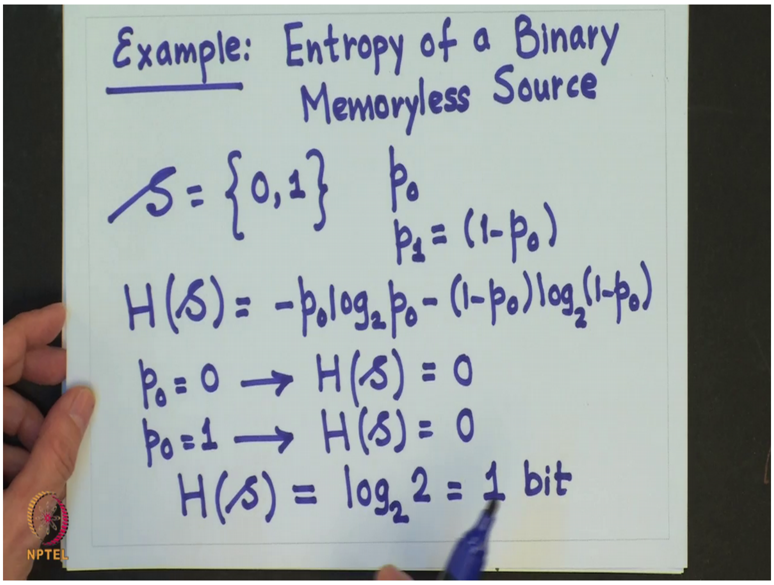
(Refer Slide Time: 11:12)



And now we see that in this expression equality holds, if q k is equal to p k for all k. So, let q k equal to 1 by capital K for all k. If we do that than for this probability distribution over the source alphabet, we will get the entropy of the source equal to log to the base 2 equal to k correct.

So, what it means, that this relationship is always valid provided the conditions on q k and p k are satisfied. So, using this relationship and using this result, we can conclude. So, I have proved this is nothing, but the entropy of the source is less than equal to log 2 K fine. So, let us take an example of a binary memoryless source and calculate this entropy.

(Refer Slide Time: 13:32)

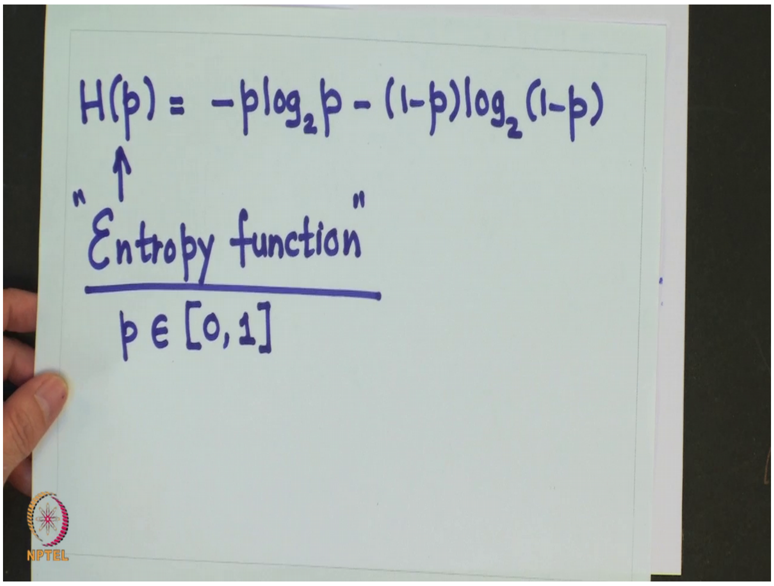


So, Entropy of a Binary Memoryless Source. So, in this case the source alphabet will be 2 binits 0 and 1 let us assume the probability is p 0 and p 1 will be; obviously, equal to 1 minus p 0. So, the entropy for this source would be equal to minus p 0 log to the base 2 p 0 minus 1 minus p 0 log to the base 2 1 minus p 0.

So, we know that for this case for p 0 equal to 0, you can show that entropy is equal to 0 and for p 0 equal to 1, the entropy is again equal to 0. And based on the earlier result, which we derived the maximum value of this entropy for a binary memoryless source would be equal to log 2 to the base 2 is equal to 1 bit 1 bit per symbol correct. So, the binary memoryless source attains the maximum value equal to 1 bit.

Now, in information theory in literature we come across 1 function, which is very closely related to the expression given by the entropy of a binary memoryless source. In the sense if I assume p 0 to be a variable, then I can write this expression as minus p log to the base 2 p minus 1 minus p log to the base 2 of 1 minus p.

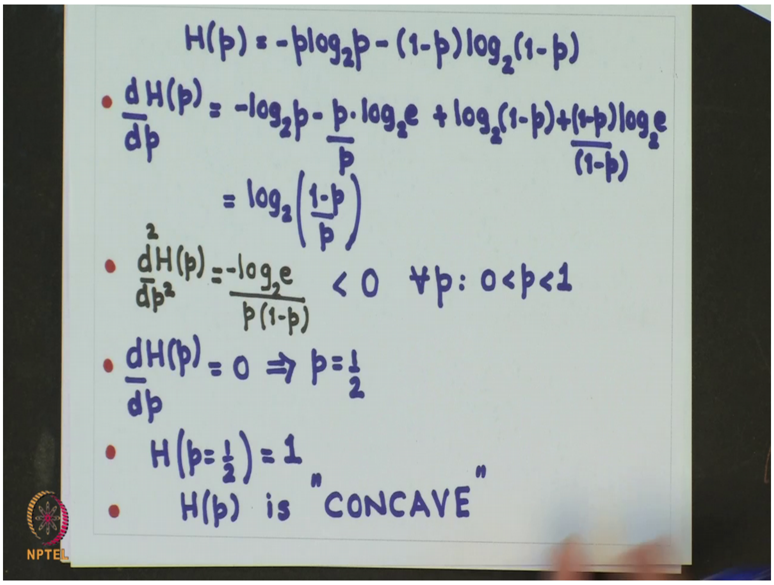
(Refer Slide Time: 16:11)



Now, when I do this then this is known as a “entropy function”. And, this is used very often in information theory fine. Now, so, if you take this H p what are the characteristics of this in entropy function, it is easy to see that for p between 0 and 1 H p will be minimum when p is equal to either 0 or 1, and the maximum will be p is equal to half and the maximum value of this entropy function would be equal to 1.

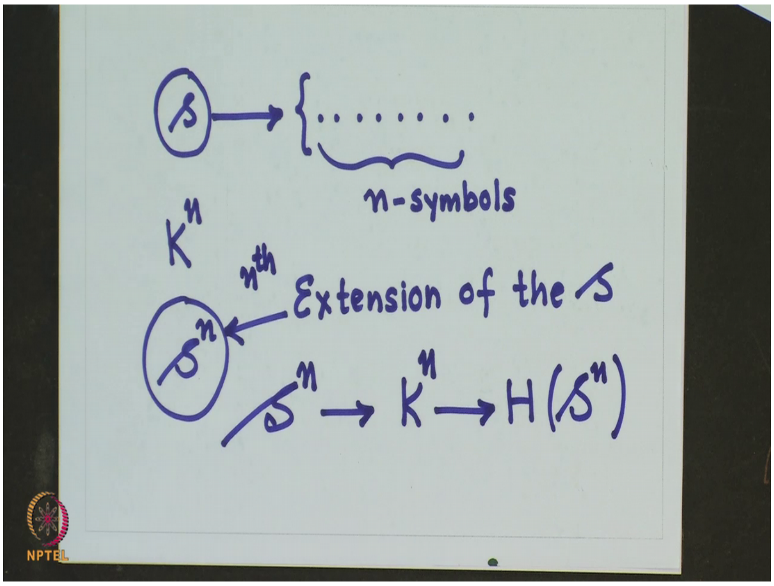
Now, we can again mathematically prove this correct the procedure is very simple, what you could do is basically write this function entropy function?

(Refer Slide Time: 17:39)



Take it is first derivative you will get this, you take a second derivative you will get this second derivative is always less than 0, for p between 0 to 1. So, what it implies that if I equate this first derivative equal to 0, the p equal to half point gives me the maximum point correct. And the maximum value is equal to 1 and based on this second derivative being less than 0, I can also conclude that my entropy function is concave ok.

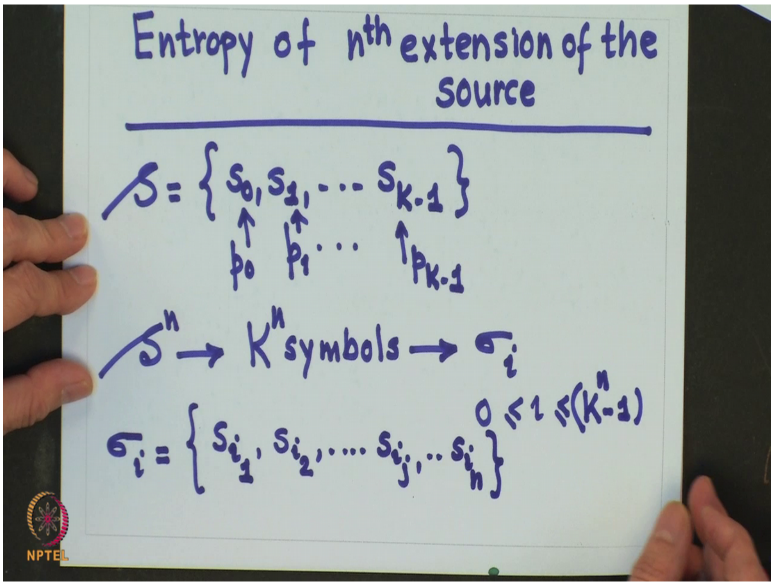
(Refer Slide Time: 18:30)



Now, given an information source with the source alphabet and the probabilities assigned to these symbols of this alphabet. If you look at the sequence which we get from here, I can group some of the symbols into 1 larger symbol correct. So, what I mean to say is that suppose the output from this source, if I group them or form a subsequence of n symbols. Then this group or subsequence can take how many values it will be K raised to n where K is the cardinality of the set alphabet associated with the source s.

So, now I can assume that this new symbols, which I get k raised to n is being emitted from another source and that source I indicate by what is known as s raised to n for, and this is known as extension of the original source s and since I have grouped in terms of n symbol, this is known as n-th extension of the source. So, I will have new source represented as this, it will have k raised to n symbols and I am interested in calculating the entropy of this source extended source.

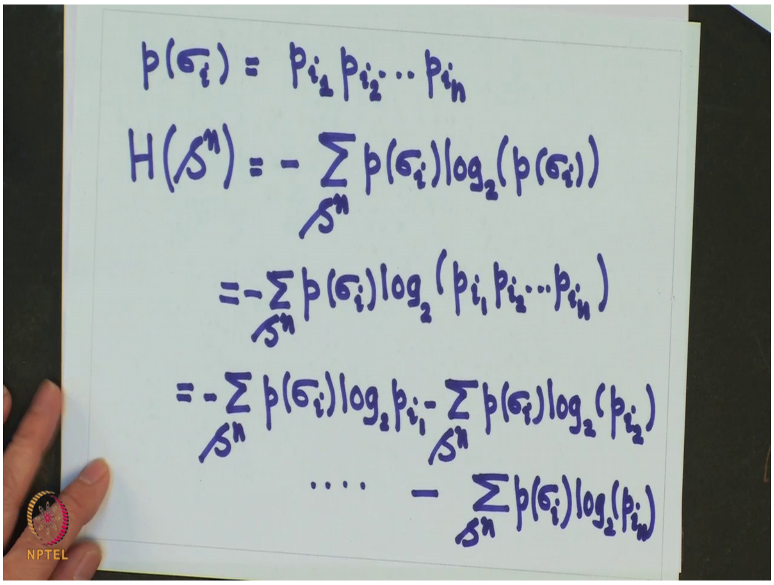
(Refer Slide Time: 20:47)



So, let us quickly see how to do this? So, Entropy of n-th extension of the source; so I have my original source with the alphabet given as follows, to each of this there is a probability assigned. Now, this extended source s n will have K raised to n symbols I denote this symbols by sigma I where I will take value from 0 to K raise to n minus 1.

Now, each of this sigma I that is the new symbol in this n-th extension could be writ10 as follows it will be S i 1. So, S i 1 can take any symbol from this alphabet then you will have S i 2. This basically can take again any symbol from this source alphabet and so on. So, each of these new symbols will have n symbols from the original source alphabet ok. So now, we will assume that the source is a discrete memoryless source.

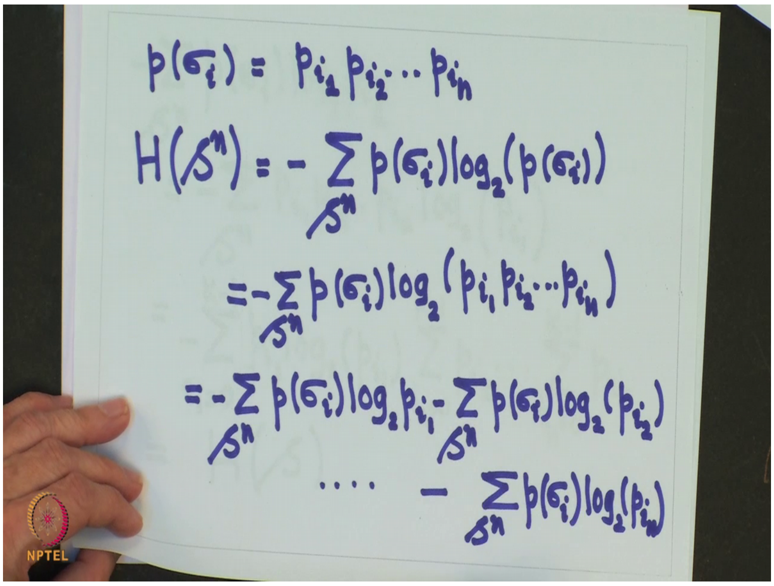
(Refer Slide Time: 23:00)



So, if we do this then the probability of sigma i is equal to probability of i 1 multiplied by probability i 2 up to probability i n.

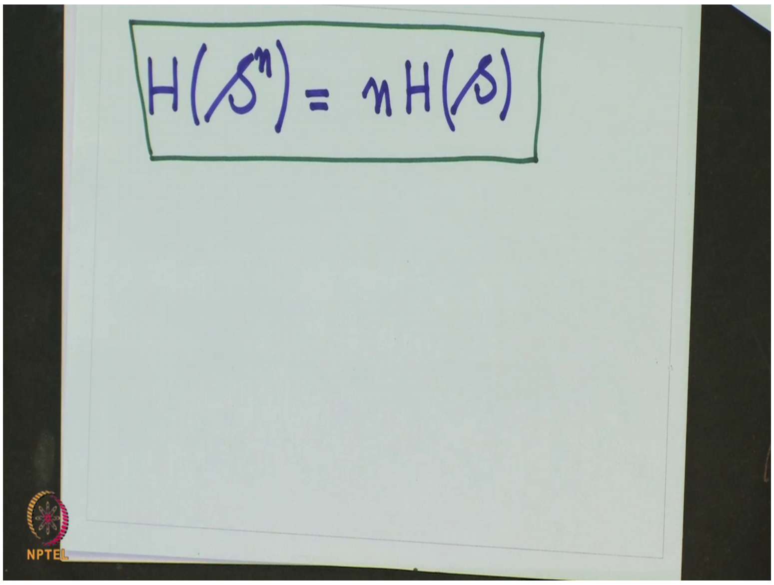
So, let us calculate the entropy of the source, that would be equal to minus summation p sigma i, log to the base 2 of probability sigma i and this should be sum over all the symbols from the extended source. So, this can be rewritten as, which can be rewritten again fine. Now, let us look at 1 other term in this expression, let us take the first term without loss of generality.

(Refer Slide Time: 25:00)



So, if I take the first term this I can write it as and this I can rewrite it as fine. Now, you see basically that is all these summations are 1. So, and this term out here sorry this is minus sign out here. So, this term is your entropy of the source. So, going back to this expression each of this basically is equal to H s and we have n terms like this.

(Refer Slide Time: 26:46)



So, what I get is that entropy of my n-th extension of my original source turns out to be n times the entropy of my original source.

So, in this class we learned about the properties of the entropy we showed that entropy is always greater than or equal to 0 it is value. And, it is bounded by the log to the base 2 k where k denotes the cardinality of the set alphabet. We also examined as an example the entropy of the binary memoryless source. And, we derived the entropy of an extent n-th extension of a source and showed that it is equal to n times and showed that, it is equal to n times the entropy of the original source.