

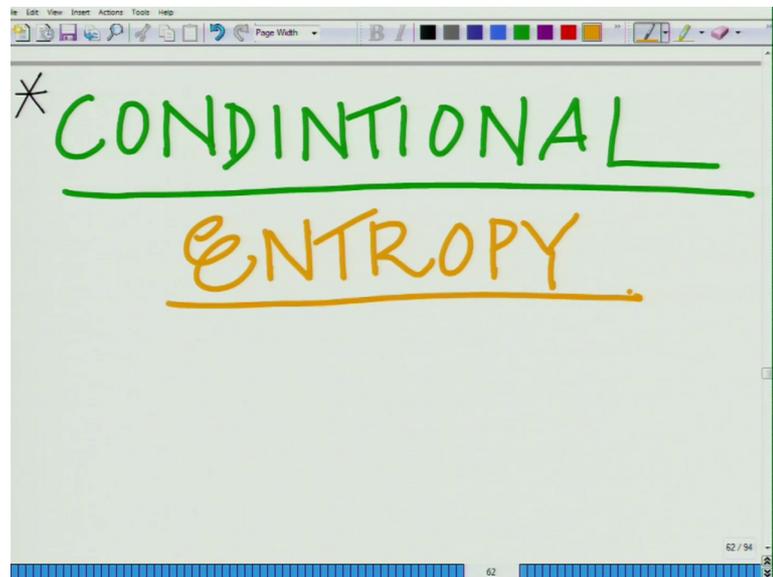
**Principles of Communication Systems - Part II**  
**Prof. Aditya K. Jagannatham**  
**Department of Electrical Engineering**  
**Indian Institute of Technology, Kanpur**

**Lecture – 32**

**Conditional Entropy, Example of Conditional Entropy, Properties of Conditional Entropy**

Hello, welcome to another module in this massive open online course. So far we have seen several concepts in information theory; we have seen the definition of entropy, joint entropy.

(Refer Slide Time: 00:30)



Let us look at yet another related quantity which is the conditional entropy. So, we will start looking at another important quantity in information theory that is the conditional entropy. Entropy as you have already said characterizes it is a measure of information, correct? It is a conditional entropy. So, let us say we have again 2 sources X and Y, we have X and Y then H of the conditional entropy remember, similar H of Y conditioned on X.

(Refer Slide Time: 01:12)

A whiteboard showing the definition of conditional entropy. At the top, 'X, Y' is written with a bracket underneath. To the right, 'Entropy of Y conditioned on X' is written in blue. Below this, the equation is written in blue ink: 
$$H(Y|X) = \sum_{i=0}^{M-1} P(X=s_i) \cdot H(Y|X=s_i)$$

That is entropy of Y conditioned on X equals well summation i equal to 0, i equal to 0 to M minus 1. Probability of X equal to s i times the entropy of Y conditioned on X equal to s i which is equal to well.

(Refer Slide Time: 02:09)

A whiteboard showing the expansion of the conditional entropy equation. The equation is written in blue ink: 
$$= \sum_{i=0}^{M-1} P(X=s_i) \sum_{j=0}^{N-1} P(Y=r_j|X=s_i) \times \log_2 \frac{1}{P(Y=r_j|X=s_i)}$$
 Below the second summation, the expression  $H(Y|X=s_i)$  is written in red ink and underlined with a red line.

Let us look at how these quantities are defined, which is equal to i equal to 0 to M minus 1 probability X equal to s i time summation j equal to 0, correct? J equal to 0 to M minus 1 probability of Y equal to r j conditioned on X equal to s i times log to the base 2 1 over probability, that is this is the entropy of Y given X equal to s i. So, we are using the

conditional probabilities. So, this is the entropy of Y conditioned on X equal to  $s_i$  this is the probability of X equal to  $s_i$ . So, this is basically the entropy of Y condition on the source of X or entropy of Y given X.

(Refer Slide Time: 03:35)

The image shows a whiteboard with handwritten mathematical equations. At the top, the expression  $H(Y|X=s_i)$  is written in red. Below it, the first equation is:

$$= \sum_{i=0}^{M-1} P_X(s_i) \sum_{j=0}^{N-1} P_Y(r_j|s_i) \cdot \log_2 \frac{1}{P_Y(r_j|s_i)}$$

The second equation is:

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \frac{P_X(s_i) P_Y(r_j|s_i)}{P_X(s_i) P_Y(r_j|s_i)} \cdot \log_2 \left( \frac{1}{P_Y(r_j|s_i)} \right)$$

The whiteboard also shows a toolbar at the top and a status bar at the bottom with the number 63.

Now this can also be written as follows. So, this can let me just simplify the notation little bit this is equal to  $i$  equal to 0  $M$  minus 1 probability of  $s_i$  time summation  $j$  equal to 0 to  $N$  minus 1 probability of  $r_j$  given  $s_i$  times  $\log$  to the base 2  $1$  over probability of  $r_j$  given  $s_i$ .

Now, if you take probability of  $s_i$  inside and multiply this take probability of  $s_i$  into probability of  $r_j$  given  $s_i$  that is  $Y$  equal to  $r_j$  given  $X$  equal to  $s_i$  that is simply probability of  $s_i$  comma  $r_j$ . So, this will be summation  $i$  equal to 0 summation  $j$  equal to 0 to  $N$  minus 1 probability of  $s_i$  into probability  $Y$  equal to  $r_j$  given  $X$  equal to  $s_i$  times  $\log$   $1$  by probability  $r_j$  given  $s_i$ , and of this quantity is basically probability of  $s_i$  comma  $r_j$ . That is  $X$  equal to  $s_i$  and  $Y$  equal to  $r_j$ .

(Refer Slide Time: 05:14)

The image shows a whiteboard with a software interface at the top. The main content is a handwritten equation for conditional entropy. At the top, there are two summation symbols:  $\sum_{i=0}^{M-1}$  and  $\sum_{j=0}^{N-1}$ , followed by the term  $\log_2 \left( \frac{1}{P_r(j|s_i)} \right)$ . Below this, a large box contains the equation: 
$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_r(s_i, j) \log_2 \frac{1}{P_r(j|s_i)}$$
 An arrow points from the text  $H(Y|X)$  Conditional Entropy of Y given X to the boxed equation. The whiteboard also shows a status bar at the bottom with the number 64.

So, this is equal to sum  $i$  equal to 0 to  $M$  minus 1 sum  $j$  equal to 0 to  $N$  minus 1 probability  $s_i$  comma  $r_j$  log to the base 2 1 over probability  $r_j$  given  $s_i$ . This is the quantity conditional entropy, entropy of  $Y$  given  $X$ . This is the conditional entropy of, conditional entropy of this is the entropy of  $Y$  given  $X$ .

So, this is the conditional entropy of so, this is basically the uncertainty in  $Y$ , correct? The uncertainty in  $Y$  given  $X$  having observed  $X$ , or having given the source  $X$ , what is the uncertainty in  $Y$ , what is the uncertainty in  $Y$ . So, this is the in definition of conditional entropy.

(Refer Slide Time: 06:49)

$$H(Y|X) + H(X) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_{r}(s_i, t_j) \log_2 \frac{1}{P_{r}(t_j | s_i)}$$

Now let us look at an interesting property of this quantity conditional entropy. H of Y given X plus H of X I would like to look at what this quantity is this quantity is summation first time I am going to write H of Y given X this is probability of s i comma r j log 2 to the base 1 by probability, I am sorry r j given s i.

(Refer Slide Time: 07:40)

$$\begin{aligned} & \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_{r}(s_i, t_j) \log_2 \frac{1}{P_{r}(t_j | s_i)} \\ & + \sum_{i=0}^{M-1} P_{r}(s_i) \log_2 \frac{1}{P_{r}(s_i)} \end{aligned}$$

$H(Y|X)$   
 $H(X)$

This is basically your H of Y given X plus summation i equal to 0 to M minus 1 probability entropy of X. This we have seen several times before probability of s i times log 2 to the base probability of s i summation over all i.

(Refer Slide Time: 08:24)

The whiteboard shows the following handwritten equations:

$$+ \sum_{i=0}^{M-1} P_r(s_i) \log_2 \frac{1}{P_r(s_i)}$$


---


$$\sum_{j=0}^{N-1} \sum_{i=0}^{M-1} P_r(s_j, s_i) \log_2 \left( \frac{1}{P_r(s_i)} \right)$$

The label  $H(X)$  is written above the second equation.

So,  $H(Y|X)$  plus  $H(X)$ . Now look at this, this entropy of  $X$  I can also write as,  $H(X)$  I can also write this as remember probability of  $s_i$  is nothing but probability of  $s_i, j$  summed over all  $j$ . So, I can also write this as well,  $j$  equal to 0 to  $N$  minus 1. I can also introduce a summation over all  $j$   $i$  equal to 0 to  $M$  minus 1 probability of  $r_j, s_i$  over probability of  $s_i$ . And therefore,  $H(Y|X)$  plus  $H(X)$  you can see ah,

(Refer Slide Time: 09:03)

The whiteboard shows the following handwritten equations:

$$\Rightarrow H(Y|X) + H(X)$$

$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_r(s_i, s_j) \log_2 \frac{1}{P_r(s_j, s_i)}$$

$$+ \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_r(s_i, s_j) \log_2 \frac{1}{P_r(s_i)}$$

This was this is basically summation  $i$  equal to 0 to  $M$  minus 1 summation  $j$  equal to 0 to  $N$  minus 1, probability  $s_i, r_j$  log 2 to the log to the base 2 1 over probability  $r_j, s_i$



(Refer Slide Time: 11:06)

The image shows a whiteboard with a digital interface at the top. The main content is a mathematical derivation. At the top, there is a double summation formula: 
$$= \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} P_r(s_i, t_j) \log_2 \left( \frac{1}{P_r(s_i, t_j)} \right)$$
 Below this, the expression  $H(X, Y)$  is written in purple. Three orange arrows point from this label to the terms in the boxed equations below: 'uncertainty in Y given X' points to  $H(Y|X)$ , 'uncertainty in X' points to  $H(X)$ , and 'joint uncertainty X and Y' points to  $H(X, Y)$ . The first boxed equation is  $H(Y|X) + H(X) = H(X, Y)$ . The second boxed equation is  $H(X|Y) + H(Y) = H(X, Y)$ . The whiteboard interface includes a toolbar at the top and a status bar at the bottom showing '67 / 94'.

Now, you can observe this is summation  $i$  equal to 0 to  $M$  minus 1 summation  $j$  equal to 0 to  $N$  minus 1  $r$   $j$  equal to 0 to  $N$  minus 1 probability  $s_i$  comma  $r_j$  log 2 to the base 1 over probability  $s_i$  comma  $r_j$ , and if you can recollect this is nothing but the joint entropy of  $X$  comma  $Y$ .

So, the very interesting relation that we have here is that  $H(Y|X)$  conditional entropy of  $Y$  conditional entropy  $X$   $H(Y|X)$  plus  $H(X)$  is nothing but the joint entropy. This is the conditional entropy  $H(Y|X)$  plus  $H(X)$  equals  $H(X, Y)$ . Similarly you can also show naturally, is nothing special about  $X$  and  $Y$ . So, if I reverse the roles of  $X$  and  $Y$   $H(X|Y)$  plus  $H(Y)$  is also equal to  $H(X, Y)$ , this is another interesting property of this is basically another way to express again. So, I have  $H(Y|X)$  plus  $H(X)$  is equal to the joint entropy  $H(X, Y)$ , which is the same as  $H(X|Y)$  plus  $H(Y)$  and. So, we have also seen  $H(Y|X)$  is the conditional entropy you can also see this is basically nothing but the uncertainty remember entropy is the uncertainty, entropy is the uncertainty.

So, conditional entropy is the uncertainty remaining in  $Y$ . One can say so, you have conditional entropy is nothing but the uncertainty remember is the information conveyed by a source or is a uncertainty in a source. So, we can say the conditional entropy is the uncertainty remaining in  $Y$  having observed  $X$ .

(Refer Slide Time: 13:15)

The image shows a whiteboard with a mathematical derivation. At the top, the equation is written as:

$$= \sum_{i=0} \sum_{j=0} Pr(x_i, y_j) \log_2 \frac{1}{Pr(y_j | x_i)}$$

Below the equation, there are handwritten annotations. A purple arrow points from the text  $H(Y|X)$  to the inner sum over  $j$ . Another purple arrow points from the text  $H(Y|X) = \text{uncertainty remaining in } Y \text{ having observed } X$  to the entire equation. The text  $H(Y|X)$  is written in purple, and the explanatory text is in orange.

So, this is  $H(Y|X)$  equals uncertainty remaining, the uncertainty remaining in  $Y$  having observed  $X$ . And  $H(X|Y)$  and of course, this is a joint information in  $X$  and  $Y$  this information. So, uncertainty so, let us look at it this way let us write it over here. Uncertainty remaining in  $Y$  given  $X$ ,  $H(X)$  is uncertainty in  $X$  this is joint information or joint uncertainty in  $X$  and  $Y$ , joint uncertainty in  $X$  and  $Y$ . So, uncertainty in  $X$  plus the uncertainty remaining in  $Y$  given  $X$ , that is a joint information of  $X$  and  $Y$  joint entropy of  $X$  comma  $Y$ . So, we have this interesting property.

Let us look at a simple example to understand this better all right. So, let us look at a simple example, let us go back to our previous example that we had where we had this table of probability  $X$ .

(Refer Slide Time: 15:16)

Example:

$$P_r(Y=r_j | X=s_i) = \frac{P_r(s_i, r_j)}{P_r(s_i)} = \frac{\frac{1}{16}}{\frac{1}{4}} = \frac{1}{4}$$

		$r_0$	$r_1$	$r_2$	$r_3$	
$s_0$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$	$P_{i,3} = P(X=s_0, Y=r_3) = \frac{1}{32}$
$s_1$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$	$= \frac{1}{32}$
$s_2$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	
$s_3$	$\frac{1}{4}$	$\frac{1}{4}$	0	0	0	

$P(Y|X)$

So, this is what you have, this is your  $s_0, s_1, s_2, s_3$ . This is your  $r_0, r_1, r_2, r_3$  the joint probabilities are  $\frac{1}{8}, \frac{1}{16}, \frac{1}{32}, \frac{1}{32}, \frac{1}{16}, \frac{1}{8}, \frac{1}{32}, \frac{1}{32}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{4}$ . Now each entry  $p_{ij}$  remember equals  $p(X=s_i, Y=r_j)$ . For instance this denotes  $p_{03}$  equals  $p(X=s_0, Y=r_3) = \frac{1}{32}$ . And from this we have computed the marginal probabilities of each  $s_i$  probability of  $X=s_i$  is simply the sum of the probabilities in the corresponding row. That is in sum the probabilities of  $s_0$  corresponding to each  $r_j$ , that is  $p(s_0, r_j)$  is summation over  $j$  gives the probability of  $s_0$  that is, we know we know that by the total probability route.

So, that is we have computed this before. So, probability of each  $s_i$  all  $s_i$  are equiprobable, probabilities  $\frac{1}{4}$  is probability of  $r_0$  is  $\frac{1}{8}$ ,  $\frac{1}{4}$  star  $\frac{1}{8}$   $\frac{1}{4}$  star  $\frac{1}{8}$   $\frac{1}{4}$  star  $\frac{1}{8}$   $\frac{1}{4}$  star  $\frac{1}{8}$ .

(Refer Slide Time: 17:37)

The image shows a whiteboard with handwritten mathematical expressions. At the top left, there is a green fraction  $\frac{1}{4}$  with a blue '4' written below it. To the right, there is a horizontal line with four vertical tick marks below it, and the expression  $P(Y|X)$  written below the line. Below this, the expression  $P(Y|X=s_0)$  is written. To the right of this, there is a red arrow pointing to the expression  $P(Y=s_j|X=s_0)$ , with a red  $P_{ij}$  written above it. Below this, the expression  $= \frac{P_r(Y=s_j, X=s_0)}{P_r(X=s_0)}$  is written in green. At the bottom right of the whiteboard, there is a small number '69'.

Now what we have to do is we have to compute, let us say the joint probability. So, we want to compute the probability of well Y given X. Let us first start by computing probability of Y given X equal to  $s_0$ . For that we will need the probability of Y equals  $r_j$  given X equal to  $s_0$ . Now if you look at this for instance probability of Y equal to  $s_0$ , given X equal to  $s_0$  equals probability given X Y equal to, sorry  $r_1$ . Let us say Y equal to  $r_1$  Y equal to  $r_1$  equals probability of  $s_0$  comma  $r_1$  divided by probability of  $s_0$  comma  $r_1$ , for instance that is  $1$  by  $16$  divided by,  $1$  by  $16$  divided by probability of  $s_0$  probability of  $s_0$  is  $1$  by  $4$  equals  $1$  by  $4$ .

So, probability of remember we are using we are using the simple principle that probability of a given b is probability of a intersection b divided by probability of b. We already know the for various joint probabilities, for instance if we want to compute probability of  $s_0$   $r_j$  given X Y equal to  $r_j$  given X equal to  $s_0$  that is simply equals probability of Y equals  $r_j$  comma X equal to  $s_0$ , given the probability that X equals  $s_0$ . We already know these quantities, this is nothing but  $p_{0j}$  that is probability of X equal to  $r_0$  comma Y equal to  $r_j$ . And probability of X  $s_0$  we already know this probability these are the marginal probabilities.

(Refer Slide Time: 19:59)

The whiteboard shows the following calculations:

$$P(Y=5|X=8_0) = \frac{P(Y=5, X=8_0)}{P(X=8_0)}$$
$$P(Y=r_0|X=8_0) = \frac{1/8}{1/4} = \frac{1}{2}$$
$$P(Y=r_1|X=8_0) = \frac{1/16}{1/4} = \frac{1}{4}$$
$$P(Y=r_2|X=8_0) = \frac{1}{8} = P(Y=r_3|X=8_0)$$

So, we see for instance probability of Y equal to r 0 given X equal to r 0 that is the probability of r 0, s 0 that is 1 by 8 divide the probability X equal to s 0 that is 1 by 4. So, this is 1 by 8 divided by 1 by 4 equals well half similarly probability Y equals r 1 given X equals s 0. We have already calculated that that is basically 1 by 16 divided by 1 by 4 equals, 1 by 4 probability Y equals r 2 given X equal to s 0 equals, well 1 by 32 divided by 1 by 4 1 by 8 which is also equal to the probability that Y equal to r 3 given X equal to s 0.

(Refer Slide Time: 21:10)

The whiteboard shows the following calculation for conditional entropy:

$$H(Y|X=8_0) = H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\right)$$
$$= \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 + \left(\frac{1}{8} \log_2 8\right) \times 2$$
$$= 1.75 \text{ bits}$$

So, the entropy given, entropy of Y given X equal to s 0. Corresponds to entropy of a source with the 4 probabilities, probability Y equal to r 0 given X equal to s 0 that is half Y equal to r 1 given X equal to s 0 that is 1 by 4 and the remaining probabilities that is 1 by 8 1 by 8, which is basically again, we confirm the entropy of the source this is basically half log to the base 2 1 over half that is 2 1 by 4 log to the base 2 4 plus 1 by 8 log to the base 2 8 and find there are 2 symbols of probability 1 by 8. So, into multiplied by 2. So, this is 1.75 bits. Similarly one can calculate H of Y.

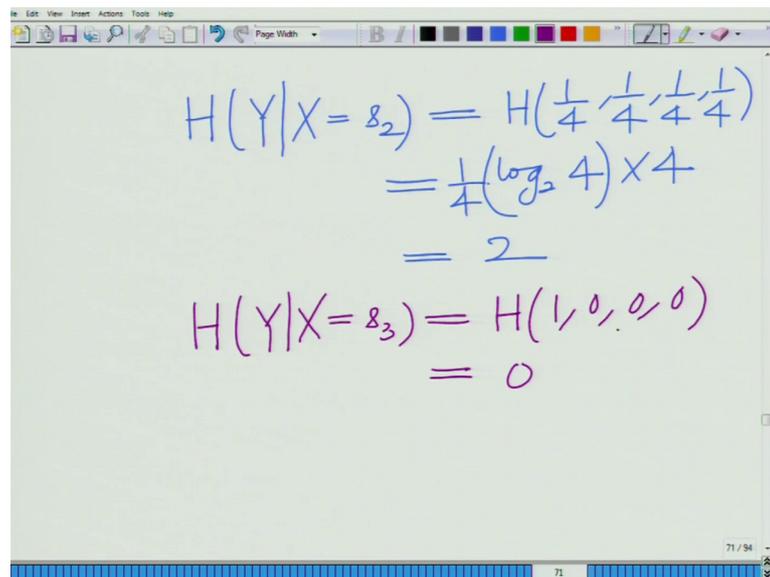
(Refer Slide Time: 22:09)

$$\begin{aligned}
 &= \frac{1}{2} \log_2 2 + \frac{1}{4} \log_2 4 \\
 &\quad + \left( \frac{1}{8} \log_2 8 \right) \times 2 \\
 &= 1.75 \text{ bits} \\
 H(Y|X=s_1) &= H\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{8}, 0\right) \\
 &= 1.75
 \end{aligned}$$

So, we have so, using the conditional probabilities of the various symbols Y equal to r 0, r 1, r 2, r 3. Conditioned on the fact that X equal to 0 we are able to calculate the conditional entropy of Y given X equal to s 0.

Next the conditional entropy of Y given X equal to s 1. This is equal to the entropy corresponding to the probabilities well 1 by 4 again the same 1 by 4, 1 by 2, 1 by 8 comma 0 which is equal to 1.75, H of Y given X equal to s 2.

(Refer Slide Time: 22:50)



The image shows a whiteboard with handwritten mathematical derivations. The first derivation, written in blue ink, calculates the conditional entropy  $H(Y|X=s_2)$  as  $H(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ , which simplifies to  $\frac{1}{4}(\log_2 4) \times 4 = 2$ . The second derivation, written in purple ink, calculates  $H(Y|X=s_3)$  as  $H(1, 0, 0, 0)$ , which simplifies to 0. The whiteboard interface includes a menu bar at the top with options like 'Edit', 'View', 'Insert', 'Actions', and 'Tools', and a status bar at the bottom showing '71 / 94'.

$$H(Y|X=s_2) = H\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$$
$$= \frac{1}{4}(\log_2 4) \times 4$$
$$= 2$$
$$H(Y|X=s_3) = H(1, 0, 0, 0)$$
$$= 0$$

This is equal to H of, this probabilities will be equal 1 by 4, 1 by 4, 1 by 4, 1 by 4 because, if you look at given X equal to s 2 all the symbols r j the joint probability of all symbols is 1 by 16. So, each probability that is Y equal to that Y equal to r j given X equal to s 2 will be 1 by 16 divided by 1 by 4 that is 1 by 4.

So, the conditional probabilities will all be 1 by 4 in this case and the joint the entropy will be well, 1 by 4 log to the base 2 4 into there are 4 symbols with probability 1 by 4 So, this is basically 2. And finally, H of Y given X equal to s 3 this will be equal to H of 1, 0, 0, 0 because you can see in this case the joint probability of all symbols other than the symbol r 0, r 0. So, there conditional probability of the 0, conditional probability of X equal to Y equal to r 1, r 0 given X equal to X 3 will be 1 by 4 divided by 1 by 4 that is equal to 1 ok.

So, that is an entropy of a source with probabilities 1, 0, 0 and this is basically certain which means given X equal to s 3 the only symbol Y can take will be r, r 0, because the rest all have 0 probabilities. So, there is no uncertainty which means this conditional entropy given X equal to s 3 will naturally be 0. And that you can get also from this calculation corresponding to probabilities 1, 0, 0, 0.

(Refer Slide Time: 24:42).

The image shows a whiteboard with handwritten mathematical equations. The top equation is  $H(Y|X=s_3) = H(1, 0, 0) = 0$ . Below it is the general formula  $H(Y|X) = \sum_{i=0}^{M-1} P_r(X=s_i) H(Y|X=s_i)$ . The bottom equation shows the calculation:  $\frac{1}{4} \times 1.75 + \frac{1}{4} \times 1.75 + \frac{1}{4} \times 2 + \frac{1}{4} \times 0$ . The whiteboard interface includes a toolbar at the top and a status bar at the bottom showing '72 / 94'.

And therefore, finally, we have H of Y given X equals summation well i equal to 0 to M minus 1 M equal to 4 in this case probability X equal to s i into entropy H of Y given X equal to s i which is equal to well, that is equal to probability of X equal to s 0 is 1 by 4 into entropy X given Y given X equal to s 0 that is 1.75. Plus probability of X equal to s 1 is 1 by 4 into entropy Y given X equal to s 1 that is 1.75, plus 1 by 4 probability of X equal to s 2 entropy Y given X equal to s 2 which is equal to 2 plus 1 by 4 it is probability X equal to s 3 into entropy Y given X equal to s 3 is 0.

(Refer Slide Time: 25:48)

The image shows a whiteboard with handwritten mathematical equations. The top equation is  $\frac{1}{4} \times 1.75 + \frac{1}{4} \times 1.75 + \frac{1}{4} \times 2 + \frac{1}{4} \times 0$ . The bottom equation shows the final result:  $= \frac{22}{16} = \frac{11}{8} = 1.375$  bits. The whiteboard interface includes a toolbar at the top and a status bar at the bottom showing '72 / 94'.

And if you calculate this you will get the answer this is equal to 22 by 16 equals 11 by 8 which is equal to well, this is equal 22 by I think this equal to 1.375 because this is equal to 1.375.

(Refer Slide Time: 26:14)

The image shows a whiteboard with handwritten mathematical derivations. At the top, the fraction  $\frac{22}{16}$  is simplified to  $\frac{11}{8}$ , which is then converted to the decimal 1.375, labeled as 'bits' and identified as  $H(Y|X)$ . Below this, the entropy of X is calculated as  $H(X) = 2 \text{ bits} = H(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ . The joint entropy is then calculated as  $H(X) + H(Y|X) = 2 + 1.375 = 3.375$ , which is also labeled as  $H(X,Y)$ . A note 'Joint Entropy of X,Y' is written in yellow. At the bottom, the equation  $H(Y) + H(X|Y) = H(X,Y)$  is written in yellow, with an arrow pointing to it and the word 'check' written below.

And now if you remember, H of X equal to 2 we have derived this is basically this quantity is H of Y given X, all the source symbols have all the symbols of X of uniform probability 1 over 4. So, H of X equal to 2 we have calculated this H of X equal to 2 this is equal to H of because all the marginal probabilities are equal to 1 by 4, and you can check that H of X plus H of Y given X is equal to 2 plus 1.375 equals 3.375 which is indeed equal to H of X given Y. We had also computed H of X given Y which is now the joint entropy of X comma Y.

Similarly, you can also check that H of well, Y plus H of X given Y should also be equal to H of X comma Y, you can also check this. So, you can check you can check this. So, basically that is what we have we have seen the definition of conditional entropy. So, in this module what we have seen is we have explored yet another concept that is the conditional entropy. The conditional entropy of Y given X and X given Y which is the conditional entropy of Y given X is the uncertainty remaining in Y having observed X.

(Refer Slide Time: 27:59)

The image shows a digital whiteboard with the equation  $H(X) + H(Y|X) = H(X,Y)$  written in blue ink and enclosed in a red rectangular box. Below the equation, there are three handwritten annotations in green ink: 'information in X' with an arrow pointing to  $H(X)$ , 'conditional Entropy information in Y given X' with an arrow pointing to  $H(Y|X)$ , and 'Joint information in X, Y.' with an arrow pointing to  $H(X,Y)$ . The whiteboard interface includes a toolbar at the top with various drawing tools and a status bar at the bottom showing '73 / 94'.

So, we have  $H$  of  $X$  plus  $H$  of, the important property that we have seen  $H$  of  $H$  of  $Y$  given  $X$  equals  $H$  of  $Y$ . This is conditional entropy or information in  $Y$  uncertainty in  $Y$  given  $X$ . This is the information in  $X$ . This is the joint information in  $X$  comma  $Y$ ; this is the joint information in  $X$  comma  $Y$  all right. So, we have seen conditional entropy, it is properties and an example to illustrate how to compute the conditional entropy all right.

So, we will stop here and continue with other aspects in the subsequent modules.

Thank you very much.