EXPLANATIONS OF ENTROPY AND CONDITIONAL ENTROPY – This presents the same results as the text on page 331, but in a way that I find more intuitive.

For all the equations below, we assume a lower case variable represents a specific outcome of an experiment, denoted by a random variable given by the corresponding upper case letter. Subscripts for probabilities are assumed to be clear from context, and are therefore omitted. All logs are assumed to be base 2. We assume the notation and definition of conditional and joint probability from T&W. We also follow the convention of defining 0 log 0 to be zero.

WHAT IS THE ENTROPY OF A RANDOM VARIABLE X?

$$H(X) = \sum_{x} p(x) \log p(x)$$

As described, straightforward.

WHAT IS THE ENTROPY OF A RANDOM VARIABLE Y, GIVEN THAT SOME SPECIFIC OUTCOME x OF SOME OTHER RANDOM VARIABLE X HAS ALREADY HAPPENED?

It's of the same form as the entropy of Y, but using conditional probabilities since x has already happened.

$$H(Y \mid X = x) = -\sum_{y} p(y|x) \log p(y|x)$$

WHAT IS THE ENTROPY OF A RANDOM VARIABLE Y, GIVEN RANDOM VARIABLE X IN GENERAL?

It's the weighted sum of the above, over all the possible outcomes of X weighted by probability of occurrence.

$$H(Y \mid X) = \sum_x p(x) * \left(-\sum_y p(y|x) \log p(y|x) \right)$$

$$= -\sum_x \sum_y p(x) p(y|x) \log p(y|x)$$

But by the definition of conditional probability, p(x)p(y|x) = p(x,y), so this becomes

$$= -\sum_x \sum_y p(x,y) \log p(y|x)$$

EXAMPLE

Let X = {even, odd} be a set of outcomes from a 6-sided die roll

Let Y = {<2, >=2} be another

WHAT IS THE ENTROPY OF X?

- (½ * log ½ + ½ * log ½) = 1, as expected

WHAT IS THE ENTROPY OF Y GIVEN THAT X=even?

If X is even, we know the only possible outcome of Y, so the entropy had better be zero! Let's check:

$$H(Y \mid X = x) = -\sum_y p(y|x) \log p(y|x)$$

$$H(Y \mid X = even) = -\sum_y p(y|even) \log p(y|even)$$

This summation has two terms, corresponding to y= "<2" and y=">=2". If the die is even, it can't be less than 2, so p(Y="<2"|X=even) is zero and therefore the first term is zero. Similarly, if the die is even, it must be >=2, so p(Y= “>=2”|X=even) is 1, and that makes the second term zero as well. So 0+0 = 0, and we get an entropy of zero as expected. If we learn that the die is even, the outcome of Y is completely determined and we have zero uncertainty.

WHAT IS THE ENTROPY OF Y GIVEN X IN GENERAL?

This will be the weighted sum of the entropies of Y for X=even and X=odd, weighted by probability of occurrence for each X outcome (1/2 in this case). If X is even, we already know the entropy of Y is zero. If X is odd, the die must be 1,3, or 5, and we’re in an experiment where there’s a 1/3 chance of Y= “<2” (corresponding to X=1), and a 2/3 chance of Y=”>=2” (corresponding to X=3 or 5). So the resulting calculation will had better be equal to ½ * 0 plus ½ * H(Y|X=odd).

Using the equation for H(Y|X) given previously, we get:

$$-\sum_{x}\sum_{y} p(x,y)\log p(y|x)$$

$$= -\sum_{x=\{even,odd\}}\ \sum_{y=\{<2,\geq 2\}} p(x,y)\log p(y|x)$$

= p(even,<2)log p(<2 | even) +

p(even,>=2)log(p(>=2|even) +

p(odd,<2)log p(<2 | odd) +

p(odd,>=2)log(p(>=2|odd)

=-[0*log 0 + ½ log 1 + (1/6 log 1/3) + (1/3 log 2/3)]

= -[0 + 0 + (1/6 log 1/3) + (1/3 log 2/3)]

If instead we calculate H(Y|X=odd), we obtain

$$-[p(<2|odd)*\log p(<2|odd) + p(>=2|odd)*\log p(>=2 \mid odd)]$$

$$= -[\ 1/3 * \log 1/3 + 2/3 \log 2/3]$$

which is indeed half of the previous quantity. Running the numbers, we obtain $H(Y|X) \cong .465$.

Bottom line: Given a 6-sided fair die roll, $X= \{even,odd\}$ and $Y=\{<2,>=2\}$, we have

$$H(X) = 1$$

$$H(Y|X=even) = 0$$

$$H(Y|X) \cong .465$$