

## SOLUTION TO EXAM NO. 1

### Problem No. 1: Entropy Calculations

Let  $p(x, y)$  be given by:

X Y	1	2
1	1/8	2/8
2	4/8	1/8

(a) Compute  $H(X)$  and  $H(Y)$ .

$H(X)$  is given by:

$$H(X) = \sum_{x \in \mathcal{X}} p(x) \log \frac{1}{p(x)}$$

And  $H(Y)$  is given by:

$$H(Y) = \sum_{y \in \mathcal{Y}} p(y) \log \frac{1}{p(y)}$$

We need to find  $p(x)$  and  $p(y)$  from  $p(x, y)$  given in the problem statement.

$$\begin{aligned} p(x = 1) &= p(x = 1, y = 1) + p(x = 1, y = 2) \\ &= \frac{1}{8} + \frac{4}{8} = \frac{5}{8} \end{aligned}$$

$$\begin{aligned} p(x = 2) &= p(x = 2, y = 1) + p(x = 2, y = 2) \\ &= \frac{2}{8} + \frac{1}{8} = \frac{3}{8} \end{aligned}$$

$$\begin{aligned}
 p(y = 1) &= p(x = 1, y = 1) + p(x = 2, y = 1) \\
 &= \frac{1}{8} + \frac{2}{8} = \frac{3}{8}
 \end{aligned}$$

$$\begin{aligned}
 p(y = 2) &= p(x = 1, y = 2) + p(x = 2, y = 2) \\
 &= \frac{4}{8} + \frac{1}{8} = \frac{5}{8}
 \end{aligned}$$

Now that we have all the marginal probabilities, we can calculate  $H(X)$  and  $H(Y)$  according to the formulas given above (repeated here).

$$\begin{aligned}
 H(X) &= \sum_{x \in \mathcal{X}} p(x) \log \frac{1}{p(x)} \\
 &= \frac{5}{8} \log \frac{8}{5} + \frac{3}{8} \log \frac{8}{3} = 0.954 \text{ bits}
 \end{aligned}$$

$$\begin{aligned}
 H(Y) &= \sum_{y \in \mathcal{Y}} p(y) \log \frac{1}{p(y)} \\
 &= \frac{3}{8} \log \frac{8}{3} + \frac{5}{8} \log \frac{8}{5} = 0.954 \text{ bits}
 \end{aligned}$$

**(b) Compute  $H(X, Y)$ .**

$H(X, Y)$  is given by:

$$\begin{aligned}
 H(X, Y) &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log \frac{1}{p(x, y)} \\
 &= \frac{1}{8} \log 8 + \frac{4}{8} \log \frac{8}{4} + \frac{2}{8} \log \frac{8}{2} + \frac{1}{8} \log 8 = 1.75 \text{ bits}
 \end{aligned}$$

**(c) Compute  $H(Y/X)$ .**

$$\begin{aligned}
H(Y/X) &= \sum_{x \in \mathcal{X}} p(x)H(Y/X = x) \\
&= \sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y/x) \log \frac{1}{p(y/x)}
\end{aligned}$$

We need  $p(y/x)$  to calculate for  $H(Y/X)$ .

$$\begin{aligned}
p(y = 1/x = 1) &= \frac{p(x = 1, y = 1)}{p(x = 1)} \\
&= \frac{\frac{1}{8}}{\frac{5}{8}} = \frac{1}{5}
\end{aligned}$$

$$\begin{aligned}
p(y = 1/x = 2) &= \frac{p(x = 2, y = 1)}{p(x = 2)} \\
&= \frac{\frac{2}{8}}{\frac{3}{8}} = \frac{2}{3}
\end{aligned}$$

$$\begin{aligned}
p(y = 2/x = 1) &= \frac{p(x = 1, y = 2)}{p(x = 1)} \\
&= \frac{\frac{4}{8}}{\frac{5}{8}} = \frac{4}{5}
\end{aligned}$$

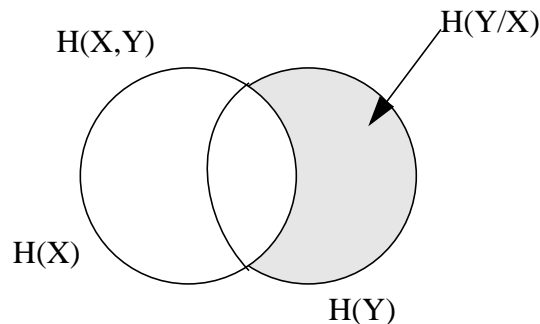
$$\begin{aligned}
p(y = 2/x = 2) &= \frac{p(x = 2, y = 2)}{p(x = 2)} \\
&= \frac{\frac{1}{8}}{\frac{3}{8}} = \frac{1}{3}
\end{aligned}$$

$$\begin{aligned}
H(Y/X) &= \sum_{x \in \mathcal{X}} p(x)H(Y/X = x) \\
&= \sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y/x) \log \frac{1}{p(y/x)} \\
&= \frac{5}{8} \left( \frac{1}{5} \log 5 + \frac{4}{5} \log \frac{5}{4} \right) + \frac{3}{8} \left( \frac{2}{3} \log \frac{3}{2} + \frac{1}{3} \log 3 \right) = 0.796 \text{ bits}
\end{aligned}$$

**(d) Justify any differences between  $H(X,Y)$  and  $H(Y/X)$ . Note that this doesn't require that the answers to parts (a)-(c) are correct.**

By definition,  $H(X, Y) = H(X) + H(Y/X)$

Also look at the Venn diagram of the above definition given below:



$H(X, Y)$  is the entropy of  $X$  and  $Y$  which is the sum of the entropy of  $X$  and  $Y$  if  $X$  and  $Y$  are independent. If  $X$  and  $Y$  are dependent as in the Venn diagram, then  $H(X, Y)$  is the sum of the entropy of  $X$  and the entropy of  $Y$  given that  $X$  is known.

## Problem No. 2: Professional Golf

**Jack Nicklaus and Tiger Woods decide to play a series of matches to decide who is the best golfer in the world (of course, everyone knows it is Arnold Palmer). Tiger Woods, being a Stanford graduate, recalls the Information Theory course he took from Prof. Cover and decides to put what he learned to good use.**

**In discussing the format of the match with Jack Nicklaus (who didn't take Information Theory), Mr. Woods obviously would like to choose a format that maximizes his chances of winning. It is a given that the longer they play, the more Jack Nicklaus will tire (hence, the better chance Tiger Woods has to win). Let's suppose the probability of Tiger Woods winning a given match is  $P(A) = 1 - 1/n$ , where  $A$  is the event that Tiger Woods wins and  $n$  is the match number. Obviously, they won't play indefinitely, so let's assume  $1 \leq n \leq 8$ .**

**(a) Using concepts discussed in this course, determine what strategy Tiger Woods should use in deciding the number and format of matches to be played.**

We want Tiger Woods to win. Tiger Wood is more likely to win if  $p(A) > 0.5$ . So if he plans to play for one game, he will lose  $p(A)=0$ . If he plans to play for two games, Nicklaus will win the first game, and they both have equal chances of winning the second game so Nicklaus is still in better shape than Woods. If Woods plans to play for three games, Nicklaus still wins the first game, they both have equal chance of winning the second game, but Woods will have a better chance of winning the third game. So Woods should plan to play at least three games to have a chance of beating Nicklaus.

**(b) If I want to bet on this match, obviously I would like to maximize my chances of winning. Using the concepts discussed in this course, develop and prove an optimal betting strategy. In other words, under what conditions should I bet?**

If they play for one game, you should bet a lot of money on Nicklaus since we know for sure he will win. If they play for two games and if you are poor, don't bet on this match since either has equal chance of winning. If they play for three games, bet on Woods since he has better chance to win the match. If they play more and more games per match, bet on Woods.

**Problem No. 3: The Browsing Dog**

**A dog walks on the integers, possibly reversing direction at each step with probability  $p = 0.1$ . Let  $X_0 = 0$ . The first step is equally likely to be positive or negative. A typical walk might look like this:**

$$(X_1, X_2, \dots, X_n) = (0, -1, -2, -3, -4, -3, -2, -1, 0, 1, \dots)$$

**(a) Prove the chain rule for entropy.**

Let  $X_1, X_2, \dots, X_n$  be drawn according to  $p(x_1, x_2, \dots, x_n)$ . Then

$$H(X_1, X_2, \dots, X_n) = \sum_{i=1}^n H(X_i / X_{i-1}, \dots, X_1)$$

Proof: By repeated application of the two-variable expansion rule for entropies, we have

$$H(X_1, X_2) = H(X_1) + H(X_2 / X_1)$$

$$H(X_1, X_2, X_3) = H(X_1) + H(X_2 / X_1) + H(X_3 / X_2, X_1)$$

...

$$H(X_1, X_2, \dots, X_n) = H(X_1) + H(X_2/X_1) + \dots + H(X_n/X_{n-1}, \dots, X_1)$$

$$H(X_1, X_2, \dots, X_n) = \sum_{i=1}^n H(X_i/X_{i-1}, \dots, X_1)$$

(b) Find  $H(X_1, X_2, \dots, X_n)$ .

$$H(X_0, X_1, \dots, X_n) = H(X_0) + H(X_1/X_0) + \sum_{i=2}^n H(X_i/X_{i-1})$$

Since  $H(X_0) = 0$  We know  $X_0 = 0$

$H(X_1/X_0) = 1$  The dog can take one of two directions

$$\sum_{i=2}^n H(X_i/X_{i-1}) = (n-1)H(X_2/X_1)$$

$$H(X_0, X_1, \dots, X_n) = 0 + 1 + (n-1)H(0.1, 0.9)$$

(c) Find the entropy rate of the browsing dog.

Entropy rate is given by:

$$H(X) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n)$$

$$H(X_0, X_1, \dots, X_n) = 0 + 1 + (n-1)H(0.1, 0.9)$$

$$H(X) = \lim_{n \rightarrow \infty} \frac{1}{n+1} H(X_0, X_1, \dots, X_n)$$

$$H(X) = \lim_{n \rightarrow \infty} \frac{1 + (n-1)H(0.1, 0.9)}{n+1}$$

$$\begin{aligned}
H(X) &= \lim_{n \rightarrow \infty} \left[ \frac{1}{n+1} + \frac{nH(0.1, 0.9)}{n+1} + \frac{H(0.1, 0.9)}{n+1} \right] \\
&= H(0.1, 0.9) \\
&= 0.469 \text{ bits}
\end{aligned}$$

**(d) What is the expected number of steps the dog takes before reversing direction?**

Let  $n$  be the number of steps taken before reversing.

$$E[n] = \sum g(n)p(n)$$

$g(n)$  is a function of  $n$ ,  $g(n) = n$ .

$$p(n) = [p(a)]^{n-1} [p(b)]^1$$

where  $a$ =same direction and  $b$ =change direction

$$p(n) = 0.9^{n-1} 0.1^1$$

$$\begin{aligned}
E[n] &= \sum_{i=1}^n n 0.9^{n-1} 0.1^1 \\
&= \frac{0.1 \sum_{i=1}^n n 0.9^{n-1}}{0.9}
\end{aligned}$$

$$\text{Since } \sum_{i=1}^n nr^n = \frac{r}{(1-r)^2}$$

$$\begin{aligned}
E[n] &= \frac{0.1(0.9)}{0.9(1-0.9)^2} \\
&= 10
\end{aligned}$$

The dog takes 10 steps before reversing. It reverses on the 11th step.

