EE8993 Information Theory

Exam no. 1

Vinod K. Valsalam

MS Computational Engineering

Problem no. 1 : Entropy Calculations

Let p(x,y) be given by:

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

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

$$p(X = 1) = \frac{1}{8} + \frac{4}{8} = \frac{5}{8}$$

$$p(X = 2) = \frac{2}{8} + \frac{1}{8} = \frac{3}{8}$$

$$H(X) = \sum_x p(x) log \frac{1}{p(x)}$$

$$= \frac{5}{8}\log\frac{8}{5} + \frac{3}{8}\log\frac{8}{3}$$

$$= 0.95443 \text{ bits}$$

$$p(Y = 1) = \frac{1}{8} + \frac{2}{8} = \frac{3}{8}$$

$$p(Y = 2) = \frac{4}{8} + \frac{1}{8} = \frac{5}{8}$$

$$H(Y) = \sum_y p(y) log \frac{1}{p(y)}$$

$$= \frac{3}{8}\log\frac{8}{3} + \frac{5}{8}\log\frac{8}{5}$$

$$= 0.95443 \text{ bits}$$

(b) Compute H(X,Y)

$$H(Y|X) = \sum_{x,y} p(x,y) log \frac{1}{p(y|x)}$$

$$p(y|x) = \frac{p(x,y)}{p(x)}$$

$$H(Y|X) = \frac{1}{8}\log\frac{1}{1/5} + \frac{2}{8}\log\frac{1}{2/3} + \frac{4}{8}\log\frac{1}{4/5} + \frac{1}{8}\log\frac{1}{1/3}$$

$$= 0.79556 \text{ bits}$$

$$H(X,Y) = H(X) + H(Y|X)$$

$$= 0.95443 + 0.79556$$

$$= 1.74999 \text{ bits}$$

(c) Compute H(Y/X)

$$H(Y|X) = 0.79556 \text{ bits}$$

(Already computed in part (b))

(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.

$$H(X,Y) = H(X) + H(Y|X)$$

Entropy is non–negative. Therefore,

$$H(X) \geq 0$$

Hence,

$$H(X,Y) \geq H(Y|X)$$

For the given problem, H(X,Y) and H(Y|X) are numerically different because H(X) is non–zero. If H(X) = 0, then H(X,Y) and H(Y|X) will be equal.

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 Info Theory), Mr. Woods obviously would like to choose a format that maximizes his chances of winning. It is 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.

There are only two possible outcomes for any match : Tiger Woods wins or loses. Therefore the the maximum value of entropy is 1 bit, which is attained when Tiger Woods has an equal chance of winning or losing. As the number of games played increases, the probability of Tiger Woods winning also increases. As the probability increases from 0 ($n = 1$) to 1 ($n = \infty$), entropy increases from 0, reaches the maximum when probability is 0.5, and then decreases to 0.

Tiger Woods should aim to keep the entropy low, but in his favor i.e., he should choose a 'n' such that the entropy is low and is decreasing with respect to n. The person who wins the maximum number of matches will be the best gofer in the world. We can easily easily see that H(A) = 1 when n = 2 and H(A) < 1 for n > 2. If 3 matches are played, H(A) is not low enough to ensure that Tiger Woods will win the series, because the odds are against him. If n = 5, matches 3, 4 and 5 are clearly in favor of Tiger Woods and he can be confident of winning the series. To enhance his chances of becoming the best golfer in the world, Tiger Woods can decide to play 7 matches, thereby decreasing H(A) further to his advantage.

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

As explained in the previous section, H(A) is less than 1 and decreasing for n > 2. This means that for n > 2, Tiger Woods has a better chance of winning the match than Jack Nicklaus. If only one match is played to decide the winner of the series, betting on Jack Nicklaus is sure to win. For n = 3, although H(A) < 1 in favor of Tiger Woods, the odds of winning the series is slightly in favor of Jack Nicklaus because he is sure to win the first match. It is not a good idea to bet for this value of n because both have a good chance of winning the series. For n = 5 or n = 7, the odds are heavily in favor of Tiger Woods. Therefore, betting on Tiger Woods is a good betting option.

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 :*

$(X_0, X_1, \ldots, X_n) = (0, -1, -2, -3, -4, -3, -2, -1, 0, 1, \ldots)$.

(a) Prove the chain rule for entropy.

If $X_1, X_2, \ldots, X_n$ are random variables drawn from p(x), then

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

Proof :

$$\begin{aligned}
H(X_1, X_2, \ldots, X_n) &= -E_{p(x_1,\ldots,x_n)} \log p(x_1, \ldots, x_n) \\
&= -E_{p(x_1,\ldots,x_n)} \log \prod_{i=1}^{n} p(x_i \,|x_{i-1}, \ldots, x_1) \\
&= -E_{p(x_1,\ldots,x_n)} \left[\sum_{i=1}^{n} \log p(x_i \,|x_{i-1}, \ldots, x_1) \right] \\
&= \sum_{i=1}^{n} -E_{p(x_1,\ldots,x_n)} \log p(x_i \,|x_{i-1}, \ldots, x_1) \\
&= \sum_{i=1}^{n} -E_{p(x_1,\ldots,x_i)} \log p(x_i \,|x_{i-1}, \ldots, x_1) \\
&= \sum_{i=1}^{n} H(X_i \,|X_{i-1}, \ldots, X_1)
\end{aligned}$$

(b) Find $H(X_0, X_1, \ldots, X_n)$.

$H(X_0) = 0$ ($X_0 = 0$ is given)

$H(X_1) = H(0.5)$ (First step is equally likely to be positive or negative)

$H(X_2 \,|X_1, X_0) = H(0.1)$ (There are two possibilities : The dog reverses direction with probability 0.1 and continues in the same direction with probability 0.9)

$$\begin{aligned}
H(X_0, X_1, \ldots, X_n) &= \sum_{i=0}^{n} H(X_i \,|X_{i-1}, \ldots, X_1) \\
&= \sum_{i=0}^{n} H(X_i \,|X_{i-1}, X_{i-2}) \qquad \text{(Markov process of order 2)} \\
&= H(X_0) + H(X_1 \,|X_0) + \sum_{i=2}^{n} H(X_i \,|X_{i-1}, X_{i-2}) \\
&= H(X_0) + H(X_1 \,|X_0) + (n-1) H(X_i \,|X_{i-1}, X_{i-2}) \qquad \text{(Process is stationary)} \\
&= 0 + H(0.5) + (n-1) H(0.1)
\end{aligned}$$

$$= 0 + 1 + (n-1)\left(0.1 log \frac{1}{0.1} + 0.9 log \frac{1}{0.9}\right)$$

$$= 1 + (n-1)0.46899$$

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

Entropy rate, $H(\mathcal{X}) = \lim_{n\to\infty} H(X_n \,|X_{n-1}, \ldots, X_1)$

$$= H(X_2 \,|X_1, X_0)$$

$$= H(0.1)$$

$$= 0.46899$$

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

Probability that the dog reverses direction at the k^{th} step $= (0.9)^{k-1}(0.1)$

To get the expectation of a variable, we take the product of the probability and the value of the variable and sum it over all possible values of the variable. Therefore, the expected number of steps the dog takes before it reverses direction can be calculated as

$$E[steps] = \sum_{k=1}^{\infty} (0.9)^{k-1} 0.1k + 1$$

The one is added because one step is always taken before changing direction.

$$E[steps] = \frac{0.1}{0.9} \sum_{k=1}^{\infty} (0.9)^k k + 1$$

$$= \frac{0.1}{0.9} \frac{0.9}{(1-0.9)^2} + 1$$

$$= 11 \text{ steps}$$