Homework 1

EE 668 — Information Theory

Name: Ali Zafari **Student Number:** 800350381 **Fall 2022**

Problem 1.1

a. The alphabet for random variable X is:

$$\mathcal{X} = \{1, 2, 3, ...\}$$

and its probability mass function (the coin is **fair**):

$$P(X = x) = (\frac{1}{2})^{x-1} \times \frac{1}{2} = (\frac{1}{2})^x$$

Entropy is as follows:

$$\begin{aligned} H(X) &= -\sum_x P(x) \log P(x) = -\sum_{x=1}^{+\infty} (\frac{1}{2})^x \log (\frac{1}{2})^x \\ &= -\sum_{x=1}^{+\infty} x(\frac{1}{2})^x \\ &= \frac{\frac{1}{2}}{(1-\frac{1}{2})^2} \\ &= 2 \end{aligned}$$

b. With a fair coin ($p = \frac{1}{2}$), the best questions to be asked are the questions which have the highest information for guessing the value of X. First question should be "Is X equal to 1?". There is a chance of 50% that we find out X is zero. If not, then the next reasonable question would be "Is X equal to 2?". Even by having a No answer for this question, we are confident that we have removed uncertainty by 75% about the value of X. By doing so, (on average) only 25% of times we will be required to ask 3 or more questions to find the value of X. To calculate the average number of questions need to be asked to find the exact value of X in this approach, we calculate the expected value of the number of questions as $\sum_{x=1}^{+\infty} x(\frac{1}{2})^x = 2$. This value is equal to the entropy of the random variable.

Problem 1.2

a. First we need to calculate the marginal distribution for both X and Y.

X	0	1
$P(X)$	$\frac{1}{2}$	$\frac{1}{2}$

Y	0	1
$P(Y)$	$\frac{5}{12}$	$\frac{7}{12}$

The entropies can then be calculated as:

$$H(X) = -\sum_x p(x) \log p(x) = 1$$
$$H(Y) = -\sum_y p(y) \log p(y) = 0.97987$$

b. Conditional probabilities:

$P(X\|Y)$	$Y=0$	$Y=1$
$X=0$	$\frac{3}{5}$	$\frac{3}{7}$
$X=1$	$\frac{2}{5}$	$\frac{4}{7}$

$P(Y\|X)$	$Y=0$	$Y=1$
$X=0$	$\frac{1}{2}$	$\frac{1}{2}$
$X=1$	$\frac{1}{3}$	$\frac{2}{3}$

The conditional entropies are the expected value of the conditional probability over the joint entropy:

$$H(X|Y) = -\sum_y \sum_x p(x,y) \log p(x|y) = 0.97928$$
$$H(Y|X) = -\sum_x \sum_y p(x,y) \log p(y|x) = 0.95915$$

c. Venn diagram.

$$H(X,Y) = -\sum_x \sum_y p(x,y) \log p(x,y) = 1.95915$$

d.

$$H(Y) - H(Y|X) = 0.97987 - 0.95915 = 0.02072$$

e.

$$I(X;Y) = H(X) - H(X|Y) = 1 - 0.97928 = 0.02072$$

f. Venn diagram:

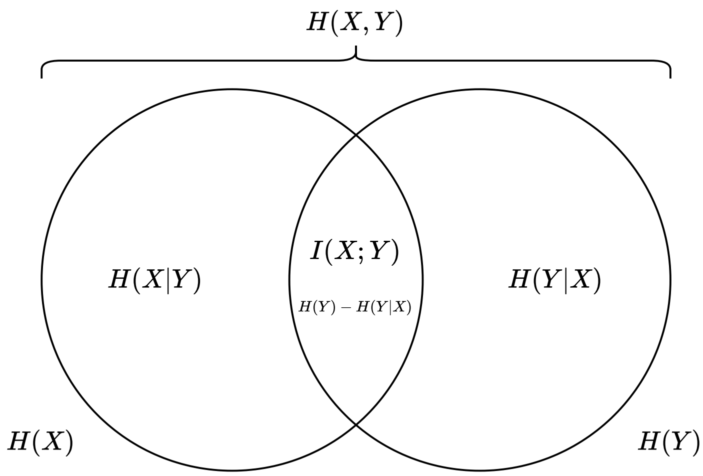

Figure 1: Venn Diagram of Information Measures of X and Y

Problem 1.3

$$H(p) = -\sum_x p(x) \log p(x) = 1.5$$
$$H(q) = -\sum_x q(x) \log q(x) = 1.5850$$

$$D(p||q) = \sum_x p(x) \log \frac{p(x)}{q(x)} = 0.0849$$
$$D(q||p) = \sum_x q(x) \log \frac{q(x)}{p(x)} = 0.0817$$

Last two equations verify that $D(p||q) \neq D(q||p)$.

Problem 1.4

Taylor series expansion of an infinitely differentiable function $f(x)$ at point $x = a$ is

$$\begin{aligned} f(x) &= \sum_{n=0}^{+\infty} \frac{f^{(n)}}{n!}(x-a)^n \\ &= f(a) + \frac{f^{(1)}(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \dots \end{aligned}$$

We now consider Taylor expansion of $\ln(x)$ at $x = 1$:

$$
\begin{aligned}
\ln(x)\Big|_{x=a} &= \ln(a) + \frac{1}{1!}\frac{1}{x}(x-a) + \frac{1}{2!}(-\frac{1}{x^2})(x-a)^2 + \dots \\
\ln(x)\Big|_{x=1} &= 0 + x - 1 - \frac{1}{2}(x-1)^2 + \dots
\end{aligned}
$$

Since the term $-\frac{1}{2}(x-1)^2$ is always negative we can write:

$$\ln(x) \leq x - 1$$

As $x > 0$ we can use a change of variable as $y = \frac{1}{x}$ (noting that y has a range of $y > 0$ same as x):

$$
\begin{aligned}
\ln(\frac{1}{y}) &\leq \frac{1}{y} - 1 \\
-\ln(y) &\leq \frac{1}{y} - 1 \\
\ln(y) &\geq 1 - \frac{1}{y}
\end{aligned}
$$
