Assignment 2 - Information Theory and Inference

Let X be a random variable representing a month of the year,

$$X \in \mathcal{A}_X = \text{'January', 'February', 'March', 'April', 'June', 'July', \ldots}$$

$$\ldots, \text{'August', 'September', 'October', 'November', 'December'}\}$$

Assume all months are equiprobable,

$$p_X(x) = 1/12, \quad \forall x \in \mathcal{A}_X$$

Let then Y be a random variable representing the first letter of the month,

$$Y \in \mathcal{A}_Y = \text{'J', 'F', 'M', 'A', 'S', 'O', 'N', 'D'}$$

and Z be a random variable representing the last letter of the month,

$$Z \in \mathcal{A}_Z = \text{'Y', 'H', 'L', 'E', 'T', 'R'}$$

Compute the mutual information:

- $I(X:Y)$
- $I(X:Z)$
- $I(Y:Z)$

and

$I(X:YZ)$

- $I(X:Y) = H[X] + H[Y] - H[X,Y]$
  $= H[Y] - H[Y|X]$

When $Y = y$ is the 1st letter of $X = x$, we have $p(Y|X) = 1$ and $= 0$ for the other cases. Then $H[Y|X] = 0$

$$H[Y] = -\sum_y p_Y(y) \log_2 p_Y(y)$$

$$= -\frac{3}{12}\log\frac{3}{12} - \frac{1}{12}\log\frac{1}{12} - \frac{2}{12}\log\frac{2}{12}$$

$$-\frac{2}{12}\log\frac{2}{12} - \frac{1}{12}\log\frac{1}{12} - \frac{1}{12}\log\frac{1}{12}$$

$$-\frac{1}{12}\log 1/12 - \frac{1}{12}\log\frac{1}{12}$$

$$\simeq 2{,}9 \text{ [BIT]} = I(X:Y)$$

- $I(X:Z) = H[X] + H[Z] - H[X,Z]$

$$= H[Z] - H[Z|X]$$

As explained previously but for $Z=z$ when $X=x$ we have that $H[Z|X]=0$

$$H[Z] = -\sum_z p_Z(z) \log_2 p_Z(z)$$

$$= -\frac{4}{12}\log\frac{4}{12} - \frac{1}{12}\log 1/12 - \frac{1}{12}\log 1/12$$

$$-\frac{1}{12}\log\frac{1}{12} - \frac{1}{12}\log\frac{1}{12} - \frac{4}{12}\log\frac{4}{12}.$$

$$\simeq 2{,}3\ [BIT] = I(X:Z)$$

- $I(Y:Z) = H[Y] + H[Z] - H[Z,Y]$

$$(= H[Y] - H[Y|Z])$$

We already have $H[Y]$ and $H[Z]$ from the previous calculations.

For $H[Y,Z]$, we are looking for the months with the same initial and last letters i.e. $Y =$ "J" and $Z =$ "Y" (*)

This happens only for January and July, so we have $p_{Y,Z}(y,z) = \frac{2}{12}$ for (*) and $= 1/12$ otherwise.

$$H[Y,Z] = -\sum_{y,z} p_{Y,Z}(y,z)\log_2 p_{Y,Z}(y,z)$$

$$= -\frac{2}{12}\log\frac{2}{12} - 10\,\frac{1}{12}\log\frac{1}{12} \simeq 3{,}4\ [BIT]$$

$$I(Y:Z) = H[Y] + H[Z] - H[Y,Z]$$

$$= 2{,}9 + 2{,}3 - 3{,}4$$

$$= 1{,}8 \text{ [BIT]}$$

- $I(X:Y,Z) = H[Y,Z] - H[Y,Z|X]$

If we know the exact month, then we know its first and last letters.

$\Rightarrow H[Y,Z|X] = 0$

$\Rightarrow I(X:Y,Z) = H[Y,Z]$ $\rightarrow$ we computed previously

$\Leftrightarrow$

$$I(X:Y,Z) = 3{,}4 \text{ [BIT]}$$