Information Theory

Problem Set 06 - Dependent Random Variables

Luís Felipe Ramos Ferreira

lframos_ferreira@outlook.com

1. (a) $H(X, Y)$ is the joint entropy of X and Y. It means how much information, on average, each of the joint outcomes carries. ON onther words, is the expected value of information of the joint outcomes from ensembles X and Y.

$$H(X, Y) = \sum_{xy \in \mathcal{A}_x \mathcal{A}_y} P(x, y) log \frac{1}{P(x, y)}$$

(b) $H(X|Y)$ is the conditional entropy of X given Y. It represents the average information infomration content of X given each $y \in \mathcal{A}_y$.

$$H(X|Y) = \sum_{xy \in \mathcal{A}_x \mathcal{A}_y} P(x, y) log \frac{1}{P(x|y)}$$

(c) $I(X, Y)$ is the mutual information between X and Y. It is the average reduction of uncertainty/ gain of information about X when learning the values of Y, or vice-versa.

$$I(X; Y) = H(X) - H(X|Y)$$

(d) $I(X : Y|Z)$ is the conditional mutual information between X and Y given Z. It is the mututal information between X and Y given that the ensemble Z is known.

$$I(X; Y|Z) = H(X|Z) - H(X|Y, Z)$$

2. (a) a

(b) b

(c) c

3. $$H(X, Y) = \sum_{xy \in \mathcal{A}_x \mathcal{A}_y} P(x, y) log \frac{1}{P(x, y)} =$$

4. 4
5. 5
6. 6
7. 7
8. 8