Assignment 1

(CS1.502) Information-Theoretic Methods in Computer Science, Spring 2023

Due in class on 27^{th} February (Friday).

INSTRUCTIONS

- Discussions with other students are not discouraged. However, all write-ups must be done individually with your own solutions.
- Any plagiarism when caught will be heavily penalised.
- Be clear and precise in your writing.

Problem 1. (a) Let X and Y be two discrete random variables jointly distributed according to P_{XY}, and $g(X, Y)$ is a function of X and Y. Show that

$$E[g(X,Y)] = \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} g(x,y)P_{XY}(x,y).$$

(2 marks)

(b) Use part (a) to prove the linearity of expectation, i.e., $E[X+Y] = E[X] + E[Y]$.

(1 mark)

Problem 2. Consider two jointly distributed discrete random variables X and Y.

(a) Prove that $H(Y|X) = 0$ if and only if Y is a deterministic function of X, i.e., for all x with $P_X(x) > 0$, there is only one possible value of y such that $P_{XY}(x,y) > 0$.

(2 marks)

(b) Let $g(X)$ be a function of the random variable X. Show that $H(g(X)) \leq H(X)$.

(2 marks)

(c) Show that

$$H(X|Y) \leq H(X) \text{ with equality if and only if } X \text{ and } Y \text{ are independent}$$

using their respective definitions, i.e., without using the non-negativity of mutual information $I(X;Y)$. [*Hint.* Prove that $f : \mathbb{R}_+ \to \mathbb{R}$ defined by $f(t) = t\log t$ is strictly convex and use it.]

(4 marks)

(d) Give an example of a probability distribution P_{XY} for which there exists a $y \in \mathcal{Y}$ such that

$$H(X|Y=y) > H(X).$$

(2 marks)

Problem 3. (a) If $X - Y - Z$ forms a Markov chain, show that $I(X;Y|Z) \leq I(X;Y)$.

(2 marks)

(b) Let X and Y be independent fair binary random variables taking values in $\{0,1\}$, and let $Z = X+Y$. Show that $I(X;Y|Z) > I(X;Y)$.

(2 marks)

(c) If $X - Y - Z$ and $X - Z - Y$ form Markov chains, show that $I(X;Y) = I(X;Z)$.

(1 mark)

Problem 4. Let $X_1, X_2, \ldots, X_n$, and Y are jointly distributed discrete random variables.

(a) Prove the chain rule for conditional entropy, i.e.,

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

[*Hint.* Prove $H(X_1, X_2|Y) = H(X_1|Y) + H(X_2|X_1, Y)$ and use mathematical induction.]

(4 marks)

(b) Prove the chain rule for mutual information, i.e.,

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

(3 marks)