Assignment 4

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

Due on 23rd April (Sunday).

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. For some $n \in \mathbb{N}$, let $X = (X_1, X_2, \ldots, X_{8n})$ be a random variable uniformly distributed on $\mathcal{X} = \{0,1\}^{8n}$. Note that each element in $\mathcal{X}$ can be equivalently treated as an integer in $[0 : 2^{8n} - 1]$ in its binary representation. Define two random variables Y and Z jointly distributed with X in the following manner:

$$Y = \begin{cases} X, & \text{if } X \bmod 8 = 0 \\ 1, & \text{otherwise} \end{cases},$$

and

$$Z = (X_1, X_2, \ldots, X_{n+1}).$$

(a) For any two random variables A and B jointly distributed according to P_{AB} on $\mathcal{A} \times \mathcal{B}$, let $\mathrm{P_c}(A|B)$ denote the maximum expected probability of correctly guessing A after observing B, i.e., $\mathrm{P_c}(A|B) = \sum_{b \in \mathcal{B}} P_B(b) \max_{a \in \mathcal{A}} P_{A|B}(a|b)$. Show that $\mathrm{P_c}(X|Z) = \frac{1}{2^{7n-1}} < \frac{1}{8} \leq \mathrm{P_c}(X|Y)$.

(3 marks)

(b) Show that $I(X;Y) = n + \frac{7}{8}\log\frac{8}{7} < n + 1 = I(X;Z)$.

(5 marks)

(c) Using parts (a) and (b), deduce that mutual information is not very suitable for modelling the information leakage in the current scenario.

(2 marks)

Problem 2. Consider a joint probability distribution P_{XY} on a finite set $\mathcal{X} \times \mathcal{Y}$. Let $\mathrm{P_c}(X)$ denote the maximum expected probability of correctly guessing X, i.e., $\mathrm{P_c}(X) = \max_{x \in \mathcal{X}} P_X(x)$, and $\mathrm{P_c}(X|Y)$ denote the maximum expected probability of correctly guessing X after observing Y, i.e, $\mathrm{P_c}(X|Y) = \sum_{y \in Y} P_Y(y) \max_{x \in \mathcal{X}} P_{X|Y}(x|y)$.

(a) Show that $\mathrm{P_c}(X) \leq \mathrm{P_c}(X|Y)$.

(2 marks)

(b) The quantity $L(X \to Y) \triangleq \log \frac{1}{\mathrm{P_c}(X)} - \log \frac{1}{\mathrm{P_c}(X|Y)}$ can be viewed as the amount of information about X that is leaked through Y. Compute $L(X \to Y)$ when P_X is the uniform distribution on $\mathcal{X}$ and Y is a deterministic function of X.

(2 marks)

Problem 3. Given a joint distribution P_{XY} on a finite set $\mathcal{X} \times \mathcal{Y}$, the maximal leakage from X to Y is given by $\mathcal{L}(X \to Y) = \log \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} P_{Y|X}(y|x)$.

(a) Show that $\mathcal{L}(X \to Y) \leq \min\{\log |\mathcal{X}|, \log |\mathcal{Y}|\}$. Also, argue that $\mathcal{L}(X \to Y) = \log |\mathcal{X}|$ if and only if X is a deterministic function of Y.

(5 marks)

(b) Prove that $2^{\mathcal{L}(X \to Y)}$ is convex in $P_{Y|X}$ for a fixed support P_X.

(1 mark)