Homework 1

Due: 13:10, 09/26, 2024 (in class)

Homework Policy: (READ BEFORE YOU START TO WORK)

- Copying from other students' solution is not allowed. If caught, all involved students get 0 point on that particular homework. Caught twice, you will be asked to drop the course.
- Collaboration is welcome. You can work together with **at most one partner** on the homework problems which you find difficult. However, you should write down your own solution, not just copying from your partner's.
- Your partner should be the same for the entire homework.
- Put your collaborator's name beside the problems that you collaborate on.
- When citing known results from the assigned references, be as clear as possible.

1. (Phase transition in error probability) [20]

For a lossless source coding problem, let $\epsilon^*(n,k)$ denote the smallest possible error probability that any (n,k) source code can ever achieve, that is,

$$\epsilon^*(n,k) := \min \left\{ \epsilon \,\middle|\, \text{there exists an } (n,k,\epsilon) \text{ source code} \right\}.$$

For a discrete memoryless source $\{S_i \,|\, i \in \mathbb{N}\}$, S_i's being i.i.d. copies of a discrete random variable S with entropy $\mathrm{H}(S)$, prove the following statements.

a) If $\mathrm{R} > \mathrm{H}(S)$, then
$$\lim_{n\to\infty} \epsilon^*(n, \lfloor n\mathrm{R} \rfloor) = 0 \qquad [10]$$

b) If $\mathrm{R} < \mathrm{H}(S)$, then
$$\lim_{n\to\infty} \epsilon^*(n, \lfloor n\mathrm{R} \rfloor) = 1. \qquad [10]$$

Note that these are alternative ways to state the achievability part and the converse part of the lossless source coding theorem in the lecture, respectively.

2. (Entropy calculation) [16]

a) Let $X_1, X_2, \ldots, X_n$ be n discrete random variables with disjoint alphabets $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ respectively. Let J be a random index, independent of everything else, taking values in $\{1, 2, \ldots, n\}$ with

$$\mathsf{Pr}\{J = j\} = p_j,\ \ j = 1, 2, \ldots, n.$$

Find $\mathrm{H}(X_J)$ in terms of $\mathrm{H}(X_1), \mathrm{H}(X_2), \ldots, \mathrm{H}(X_n)$ and $p_1, p_2, \ldots, p_n$. [8]

b) A biased coin (Head with probability p and Tail with probability $(1-p)$) is flipped until the first Head occurs. Let N denote the number of flips required.

Find $\mathrm{H}(N)$ if it exists, or show that it does not exist. [8]

3. (Mixing Increases Entropy) [14]

Consider a probability vector $\boldsymbol{p} = (p_1, ..., p_i, ..., p_j, ..., p_d)$.

a) For $i \neq j \in \{1, 2, ..., d\}$, an $\{i, j\}$-mixing of $\boldsymbol{p}$, called $\boldsymbol{p}_{\{i,j\}}$, is another probability vector where both the i-th and the j-th coordinates are replaced by $\frac{p_i + p_j}{2}$.

Show that

$$\mathrm{H}(\boldsymbol{p}) \leq \mathrm{H}\left(\boldsymbol{p}_{\{i,j\}}\right).$$ [8]

b) For $\mathcal{I} \subseteq \{1, 2, \ldots, d\}$, an $\mathcal{I}$-mixing of $\boldsymbol{p}$, $\boldsymbol{p}_{\mathcal{I}}$, is another probability vector where for all the $i \in \mathcal{I}$, the i-th coordinate p_i is replaced by $\frac{\sum_{i \in \mathcal{I}} p_i}{|\mathcal{I}|}$.

Show that

$$\mathrm{H}(\boldsymbol{p}) \leq \mathrm{H}(\boldsymbol{p}_{\mathcal{I}}).$$ [6]