TUTORIAL VI

1 Channel capacity

Definition 1.1 (Information capacity)**.** *The information capacity of a channel $W_{Y|X}$ is given by $C(W_{Y|X}) = max_{P_X} I(X;Y)$, where the joint distribution of X, Y is defined by $P_{XY}(x, y) = P_X(x)P_{Y|X}(y|x)$.*

1. For a discrete channel $W_{Y|X}$ with input alphabet $\mathcal{X}$, output alphabet $\mathcal{Y}$, let $C(W)$ denote the channel capacity of W. Show that
   (a) $C(W) \geq 0$.
   (b) $C(W) \leq \log_2 |\mathcal{X}|$.
   (c) $C(W) \leq \log_2 |\mathcal{Y}|$.
   (d) $I(X;Y)$ is a continuous concave function of $p(x)$.
2. Given a channel $W_{Y|X}$ and channel capacity $C(W) = \max_{p(x)} I(X;Y)$. Suppose you apply a preprocessing step to the output by forming $\tilde{Y} = g(Y)$.
   (a) Does it strictly improve the channel capacity?
   (b) Under what conditions does the capacity not strictly decrease?

2 Binary Erasure Channel

A binary erasure channel with input alphabet $\{0, 1\}$ and output alphabet $\{0, 1, \mathrm{E}\}$ is defined by the following transition probabilities.

$$p_{Y|X}(0|0) = p_{Y|X}(1|1) = 1 - \alpha, \qquad p_{Y|X}(\mathrm{E}|0) = p_{Y|X}(\mathrm{E}|1) = \alpha$$

Essentially, a fraction α of the input bits are erased (represented by the symbol E).

1. Determine the capacity of the channel.
2. If there is (noiseless) feedback on whether the input bit is received or erased, how do you achieve a rate equal to the capacity (you can send the same message several times) ?
3. Suppose that there is no feedback and we use the following coding scheme: encode 0 as 000 and 1 as 111. Decode 000, E00, 0E0, 00E, EE0, E0E, 0EE to 0 and similarly decode 111, E11, 1E1, 11E, EE1, E1E, 1EE to 1. In case EEE is received, then choose one of $0, 1$ at random. What is the probability of error for the code?

3 Fun with Fano

1. Consider the two following pairs of correlated random variables:
   1. X is uniform on $\{0, 1\}^n$, Y equals the first $n/2$ bits of X.
   2. With probability $\alpha \in [0; 1]$, X is uniform on $\{0; 1\}^n$ and $Y = X$; and with probability $1 - \alpha$, X is uniform on $\{0; 1\}^n$ and Y is the all 0s string.

   Suppose we observe Y and estimate $\hat{X} = g(Y)$. What is the minimum possible value of $\mathbf{P}(\hat{X} \neq X)$ in the above two examples ? What lower bound does Fano's inequality give in the two examples ?
2. For two vectors $u, v \in \{0, 1\}^n$, we denote by $\Delta(u, v)$ the following set: $\Delta(u, v) = \text{Card}(\{j \in \{1, \ldots, n\} : u_j \neq v_j\})$. Suppose X and Y are two correlated random variables taking values in $\{0; 1\}^n$.
   For $i \in \{0, \ldots, n\}$, we define $\theta_i = \mathbf{P}(\Delta(X, Y) = i)$. Prove that

$$H(X|Y) \leq \sum_{i=0}^{n} \theta_i \log_2 \left(\binom{n}{i} \frac{1}{\theta_i} \right)$$

   (Hint: Define the random variable $\Delta(X, Y)$ and mimic steps from the proof of Fano's inequality)

4 Expurgation

Let C be a M-code with error probability $P_{err}(C) = \delta$.

1. Show that you can build a $\lfloor M/2 \rfloor$-code with maximal error probability $\leq 2\delta$.