Information Theory HW3

許博翔

November 2, 2023

Problem 1.

(a) Since N_0 is deterministic from $X_1, X_2, \ldots, X_{N_0}$, N_1 is deterministic from $X_1, X_2, \ldots, X_{N_1}$, there is $I(N_0; X_1, \ldots, X_{N_0}) = H(N_0) = \frac{1}{3}\log 3 + \frac{2}{3}(\log 3 - 1) = \log 3 - \frac{2}{3}$, $I(N_1; X_1, \ldots, X_{N_1}) = H(N_1) = \sum_{i=1}^{\infty} \frac{1}{2^i} i = \sum_{i=1}^{\infty} \frac{1}{2^i} \sum_{j=1}^{i} 1 = \sum_{j=1}^{\infty} \sum_{i=j}^{\infty} \frac{1}{2^i} = \sum_{j=1}^{\infty} \frac{1}{2^{j-1}} = 2$.

(b) Let's assume $n \geq 2$ (because for $n = 1$ there is nothing to be computed).

Claim: $X_1, X_2, \ldots, X_{n-1}$ are mutually independent Ber$(\frac{1}{2})$.

Proof. $\forall x \in [0,1]^{n-1}$, there is exactly one $x^* \in [0,1]^n$ (which is $(x_1, x_2, \ldots, x_{n-1}, x_1 \oplus \cdots \oplus x_{n-1})$) s.t. $2 \mid \sum_{i=1}^{n} x_i^*$ and $\forall 1 \leq i \leq n-1,\ x_i^* = x_i$.

$\therefore \Pr((X_1, \ldots, X_{n-1}) = x) = \Pr((X_1, \ldots, X_n) = x^*) = 2^{-(n-1)}$.

$\Rightarrow (X_1, \ldots, X_{n-1})$ is an uniform distribution on $[0,1]^{n-1}$, which means $X_1, X_2, \ldots, X_{n-1}$ are mutually independent Ber$(\frac{1}{2})$. ∎

Similarly, for any distinct $i_1, i_2, \ldots, i_{n-1}$, $X_{i_1}, \ldots, X_{i_{n-1}}$ are mutually independent.

Let $1 \leq i \leq n-1$,

$$I(X_i; X_{i+1} | X_1, \ldots, X_{i-1}) = H(X_i | X_1, \ldots, X_{i-1}) - H(X_i | X_1, \ldots, X_{i-1}, X_{i+1})$$

$$\overset{X_1,\ldots,X_i \text{ are mutually independent}}{=} H(X_i) - H(X_i | X_1, \ldots, X_{i-1}, X_{i+1})$$

$$\overset{X_1,\ldots,X_{i+1} \text{ are mutually independent if } i<n-1}{=}$$

$$\begin{cases} H(X_i) - H(X_i) = 0, \text{ if } i < n-1 \\ H(X_i) - H(X_1 \oplus \cdots \oplus X_{i-1} \oplus X_{i+1} | X_1, \ldots, X_{i-1}, X_{i+1}) = H(X_i) - 0 = 1, \text{ if } i = n-1 \end{cases}.$$

Problem 2.

(a) $I(X_4; X_1) = I(X_4; X_2) - I(X_4; X_2|X_1)$

$I(X_3; X_1) = I(X_3; X_2) - I(X_3; X_2|X_1) = I(X_3; X_2) - I(X_3; X_2|X_1, X_4) - I(X_4; X_2|X_1)$

$\Rightarrow I(X_1; X_3) + I(X_2; X_4) = I(X_3; X_2) - I(X_3; X_2|X_1, X_4) - I(X_4; X_2|X_1) + I(X_2; X_4) = I(X_2; X_3) + I(X_1; X_4) - I(X_3; X_2|X_1, X_4) \leq I(X_2; X_3) + I(X_1; X_4)$.

(b) It's equivalent to two Markov's chains: $X_1 - X_2 - X_3, X_1 - X_2 - X_4$.

$I(X_3; X_1) = I(X_3; X_2) - I(X_3; X_2|X_1)$

$I(X_4; X_1) = I(X_4; X_2) - I(X_4; X_2|X_1)$

$I(X_1; X_2) + I(X_3; X_4) \geq I(X_1; X_2) + I(X_4; X_1) - I(X_4; X_1|X_3) \geq I(X_2; X_1) + I(X_4; X_1) - I(X_2; X_1|X_3) = I(X_1; X_3) + I(X_1; X_4)$.

Problem 3.

(a) Let $X_i \in \mathcal{X}^{(i)}$.

$I(X;Y) = H(X) - H(X|Y) \overset{I \text{ is deterministic from } X}{=} H(X,I) - H(X|Y) = H(X|I) + H(I) - H(X|Y) \overset{I \text{ is deterministic from } Y}{=} H(X|I) + H(I) - H(X|Y,I) = I(X;Y|I) + H(I)$.

(b) The capacity $= \max\limits_{P_I} I(X;Y) = \max\limits_{P_I} E_{(X,Y)\sim P_{X,Y}}(\log \frac{P_{Y|X}(Y|X)}{P_Y(Y)}) = \max\limits_{P_I} \sum\limits_{i=1}^{l} P_I(i)(I(X_i;Y_i) - \log P_I(i)) = \max\limits_{P_I} \left(\sum\limits_{i=1}^{l} P_I(i)C^{(i)} + H(I) \right)$.

(c) Consider the distribution: $P_J(i) = \dfrac{2^{C^{(i)}}}{\sum_{j=1}^{l} 2^{C^{(j)}}}$.

$\sum\limits_{i=1}^{l} P_I(i)C^{(i)} + H(I) = \sum\limits_{i=1}^{l} P_I(i) \log \dfrac{2^{C^{(i)}}}{\sum_{j=1}^{l} 2^{C^{(j)}} P_I(i)} + \sum\limits_{i=1}^{l} P_I(i) \log \sum\limits_{j=1}^{l} 2^{C^{(j)}} = \sum\limits_{i=1}^{l} P_I(i) \log \dfrac{P_J(i)}{P_I(i)} + \log \sum\limits_{j=1}^{l} 2^{C^{(j)}} = -D(P_I \| P_J) + \log \sum\limits_{j=1}^{l} 2^{C^{(j)}} \geq \log \sum\limits_{j=1}^{l} 2^{C^{(j)}}$,

with equality $\iff D(P_I \| P_J) = 0 \iff P_I = P_J$.

$\therefore$ the capacity $= \log \sum\limits_{j=1}^{l} 2^{C^{(j)}}$, and the distribution P_I is $P_I(i) = \dfrac{2^{C^{(i)}}}{\sum_{j=1}^{l} 2^{C^{(j)}}}$.

Problem 4.

(a) Suppose that $X \sim \text{Ber}(q)$.

$\Rightarrow P_Y(0) = 1 - q + pq = 1 - \frac{1}{2}q, P_Y(1) = q(1-p) = \frac{1}{2}q$.

$I(X;Y) = H(X) + H(Y) - H(X,Y) = -q\log q - (1-q)\log(1-q) - \frac{1}{2}q\log(\frac{1}{2}q) - (1-\frac{1}{2}q)\log(1-\frac{1}{2}q) + (1-q)\log(1-q) + 2\cdot\frac{1}{2}q\log(\frac{1}{2}q) = -q\log q + \frac{1}{2}q\log(\frac{1}{2}q) - (1-\frac{1}{2}q)\log(1-\frac{1}{2}q) = -q - \frac{1}{2}q\log(\frac{1}{2}q) - (1-\frac{1}{2}q)\log(1-\frac{1}{2}q)$.

Let $\dfrac{dI(X;Y)}{dq} = -1 - \frac{1}{2}\log(\frac{1}{2}q) - \frac{1}{2}\log e + \frac{1}{2}\log(1-\frac{1}{2}q) + \frac{1}{2}\log e = -1 + \frac{1}{2}\log\dfrac{1-\frac{1}{2}q}{\frac{1}{2}q} = 0$.

$\Rightarrow \log\dfrac{1-\frac{1}{2}q}{\frac{1}{2}q} = 2$.

$\Rightarrow \dfrac{1-\frac{1}{2}q}{\frac{1}{2}q} = 4$.

$\Rightarrow q = \dfrac{2}{5}$.

$\therefore I(X;Y) \leq -\frac{2}{5} - \frac{1}{5}\log\frac{1}{5} - \frac{4}{5}\log\frac{4}{5} = -\frac{2}{5} - \frac{8}{5} + \log 5 = \log 5 - 2$, with equality iff $P_X^* = \mathrm{Ber}(\frac{2}{5}), P_Y^* = \mathrm{Ber}(\frac{1}{5})$.

(b) Since the equality in (a) is an if and only if condition, so the input distribution is unique.

(c) $D(P_{Y|X}(\cdot|0)\|P_Y^*(\cdot)) = P_{Y|X}(0|0)\log\frac{P_{Y|X}(0|0)}{P_Y^*(0)} = 1\log\frac{1}{1-\frac{1}{2}q} = -\log(1-\frac{1}{2}q) = \log 5 - 2$.

$D(P_{Y|X}(\cdot|1)\|P_Y^*(\cdot)) = P_{Y|X}(0|1)\log\frac{P_{Y|X}(0|1)}{P_Y^*(0)} + P_{Y|X}(1|1)\log\frac{P_{Y|X}(1|1)}{P_Y^*(1)} = \frac{1}{2}\log\frac{1}{2(1-\frac{1}{2}q)} + \frac{1}{2}\log\frac{1}{2(\frac{1}{2}q)} = -\frac{1}{2}(\log(2-q) + \log q) = -\frac{1}{2}(3 - \log 5 + 1 - \log 5) = \log 5 - 2$.