EE411 · Information Theory and Coding

# Assignment 3

刘乐奇 / 12011327

October 12, 2023

## 1 2.15

By the chain rule, we have

$$\begin{aligned} I(X_1;X_2,\ldots,X_n) &= \sum_{i=2}^{n} I(X_1;X_i|X_2,X_3,\ldots,X_{i-1}) \\ &= I(X_1;X_2) + I(X_1;X_3|X_2) + \ldots + I(X_1;X_n|X_2,\ldots,X_{n-1}) \end{aligned}$$

For $X_1 \to X_2 \to \ldots \to X_n$, by the Markov property, $X_i$ and $X_j$ with $j-i>1$ are independent when given $X_p$ with $p \in (i,j)$. Therefore, we have

$$\begin{aligned} I(X_1;X_2,\ldots,X_n) &= I(X_1;X_2) + I(X_1;X_3|X_2) + \ldots + I(X_1;X_n|X_2,\ldots,X_{n-1}) \\ &= I(X_1;X_2) \end{aligned}$$

## 2 2.16

1. By the data processing inequality, we have

$$\begin{aligned} I(X_1;X_3) &\le I(X_1;X_2) \\ &= H(X_2) - H(X_2|X_1) \\ &\le H(X_2) \end{aligned}$$

   Since $H(X_2) \le \log|\mathcal{X}_2| = \log k$, we have $I(X_1;X_3) \le \log k$.

2. From the above, we know that $I(X_1;X_3) \le \log k$. For $k=1$, we have $I(X_1;X_3) \le 0$. Since $I(X_1;X_3) \ge 0$, we have $I(X_1;X_3) = 0$, i.e., $X_1$ and $X_3$ are independent.

## 3 2.32

1. From the table, we can see that $Pr[X=1,Y=a]$ is the greatest among $Pr[X=x,Y=a]$ for $x \in \{1,2,3\}$, and so are the $Pr[X=2,Y=b]$ and $Pr[X=3,Y=c]$. Therefore, the minimum probability of error estimator is like below

$$\hat{X}(Y) = \begin{cases} 1, & y=a \\ 2, & y=b \\ 3, & y=c \end{cases}$$

   and the associated $P_e = Pr[\hat{X} \ne X] = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$.

2. By Fano's inequality, we have

$$P_e \ge \frac{H(X|Y)-1}{\log|\mathcal{X}|-1}$$

   We have

$$\begin{aligned} H(X|Y) &= \sum_{y\in\mathcal{Y}} Pr[Y=y]H(X|Y=y) \\ &= Pr[Y=a]H(X|Y=a) + Pr[Y=b]H(X|Y=b) + Pr[Y=c]H(X|Y=c) \end{aligned}$$

Since

$$H(X|Y=a) = H(X|Y=b) = H(X|Y=c) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{4}\log\frac{1}{4} - \frac{1}{4}\log\frac{1}{4} = \frac{3}{2}\log 2$$

we have

$$H(X|Y) = \frac{3}{2}\log 2(Pr[Y=a] + Pr[Y=b] + Pr[Y=c]) = \frac{3}{2}\log 2$$

Therefore, we have

$$P_e \geq \frac{H(X|Y) - 1}{\log|\mathcal{X}| - 1} = \frac{\frac{3}{2}\log 2 - 1}{\log 2} = \frac{1}{2}$$

So the $P_e$ we calculated before matches well with Fano's inequality.

# 4 2.34

By the data processing inequality, we have

$$I(X_0; X_{n-1}) \geq I(X_0; X_n)$$

Since $I(X_1; X_2) = H(X_1) - H(X_1|X_2)$, we have

$$\begin{aligned} & H(X_0) - H(X_0|X_{n-1}) \geq H(X_0) - H(X_0|X_n) \\ \Longrightarrow \quad & H(X_0|X_{n-1}) \leq H(X_0|X_n) \end{aligned}$$

Therefore, $H(X_0|X_n)$ is non-decreasing with $n$.