NUS

National University of Singapore

Name : LUO ZIJIAN

Matric.No: A0224725H

MUSNET: E0572844

Subject: Information Theory

Assignment: Homework NINE

Date: October 31st

Prof: Marco Tomamichel

Exercise 9.1

For z-channel, we can get $p(Y|X)$ with this following matrix

$$Q = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad x, y \in \{0, 1\}$$

And we suppose $Pr(X=1) = P$,

$$H(Y|X) = Pr(X=0) \cdot 0 + Pr(X=1) \cdot 1 = P$$

$$H(Y) = H[Pr(X=1)] = H\left[\frac{P}{2}\right]$$

So, we can $I(X;Y) = H(Y) - H(Y|X) = H\left(\frac{P}{2}\right) - P$

In conclusion: this channel mutual information $H\left(\frac{P}{2}\right) - P$

Exercise 9.2

For empirical distribution, we know $f_x(x) := n^{-1} \cdot \sum_{i=1}^{n} \delta_{x_i, x}$

And then x^n are n i.i.d. copies of X, $p_{X^n} = p_X^{\otimes n}$

$$p_{X^n}(x) = \prod_{i=1}^{n} P(x_i)$$

$$= 2^{\sum_{i=1}^{n} \log P(x_i)}$$

$$= 2^{\sum_{i=1}^{n} \cdot P(x_i|x^n) \log P(x_i)}$$

$$= 2^{-n\left[\frac{\sum_j x_i}{n} \cdot \log \frac{1}{P(x_i)}\right]} \quad \text{(we use the } f_x(x) \text{ definition)}$$

$$= 2^{-n\left[\frac{\sum \delta x_i}{n} \log \frac{1}{P(x)} \cdot \frac{f_x(x)}{f_x(x)}\right]}$$

$$= 2^{-n\left[H[f_x(x)] + D(f_x(x) \| P(x))\right]}$$

$$= 2^{-n(H(f_x) + D(f_x \| P_x))}$$

It is desired, the probability that x^n being any sequence $x \in X^n$ depends only on its type and P_X

Exercise 9.5

Following the instructions from lecture notes and the paper of Vincent Poor, for this $(2^{nR}, n)$ code, we let encoder and decoder to use random mappings. which means $P[M \neq \hat{M}] \leq \varepsilon$

And then use the Proposition 6.9

$$|M| \leq \max_{P_X \in \mathcal{P}(\mathcal{X})} \min_{Q_Y \in \mathcal{P}(\mathcal{Y})} \frac{1}{\beta_\varepsilon^*(P_{XY} \| P_X \times Q_Y)}$$

where $P_{XY}(x,y) = P_X(x) \cdot W_{Y|X}(y|x)$ is the joint distribution of channel input and output and $\beta_\varepsilon^*(P_{XY} \| P_X \times Q_Y)$ is the minimal error of the second kind.

And we know the encoder generate a codeword $x^n(m, w)$, and the decoder generate an estimate $\hat{m}(y^n, w)$

Therefore, we construct these sets $A_x = \{y \in \mathcal{Y} : (x,y) \notin A\}$ for all $x \in \mathcal{X}$

Given a fixed channel output y, its probability of (assigns $\hat{M} = m$)

$$P(\hat{M} = m \mid Y = y) = \frac{I\{y \in A_{e(m)}\}}{\sum_{m'} I\{y \in A_{e(m')}\}} = \frac{I\{y \in A_{e(m)}\}}{I\{y \in A_{e(m)} + \sum_{m' \neq m} I\{y \in A_{e(m')}\}}$$

Then to analysis the error for this code

$$P(M \neq \hat{M} \mid M = m, E = e) = 1 - \sum_{y \in \mathcal{Y}} W(y|e(m)) \cdot P[\hat{M} = m \mid Y = y]$$

$$= \sum_{y \in \mathcal{Y}} W(y|e(m)) \left(1 - \frac{I\{y \in A_{e(m)}\}}{I\{y \in A_{e(m)} + \sum_{m' \neq m} I\{y \in A_{e(m')}\}}\right)$$

$$\leq \sum_{x \in \mathcal{X}, y \in \mathcal{Y}} P_X(x) W(y|x) \left(I\{y \notin A_x\} + (|M|-1) \sum_{x' \in \mathcal{X}} P_X(x') I\{y \in A_x\}\right)$$

And summarizing this, we can get

$$P(M \neq \hat{M}) \leq \varepsilon + |M| \beta_\varepsilon^*(P_{XY} \| P_X \times P_Y)$$

FALCON

And then use the part of (Theorem 6.8 in lecture notes)

$$2^{[nR]} \leq \varepsilon \cdot \frac{1}{\beta_\varepsilon^*(P_{X^nY^n} \| P_{X^n} \times P_{Y^n})}$$

And take logrithm, and choose P_X as the maximiser in the definition of mutual information $I(X:Y) = I(W)$

$$R \leq \lim_{n\to\infty} \frac{1}{n} \log \frac{1}{\beta_\varepsilon^*(P_{XY}^n \| P_X^n \times P_Y^n)} = D(P_{XY} \| P_X \times P_Y) = I(X:Y) = C(W)$$

As desired, we prove that this randomization does not increase the capacity of the DMC