Homework 1

Q1:

$$\lim_{x \to 0} x \log x = \lim_{x \to 0} \frac{\log x}{1/x} = \lim_{x \to 0} \frac{1/x}{-1/x^2} = 0 \quad \text{(L'Hopital's rule)} \tag{1}$$

Q2:

$$\begin{aligned} H'(p) &= -\log p - 1 + \log(1-p) + 1 = \log(1-p) - \log p \\ H''(p) &= \frac{1}{p-1} - \frac{1}{p} = -\frac{1}{(1-p)\cdot p} < 0, p \in [0,1] \end{aligned} \tag{2}$$

$\therefore H(p)$ is concave in $p \in [0,1]$.

Q3:

$$\begin{aligned} \min f(x,y,z) = x\log x + y \log y + z \log z \\ s.t \;\; x + y + z = 1 \\ x, y, z \geq 0 \end{aligned} \tag{3}$$

Consider its Lagrangian:

$$g(x,y,z;\lambda) = f(x,y,z) - \lambda(x+y+z) \tag{4}$$

(a).

$$\begin{aligned} \frac{\partial g}{\partial x} &= \log x + 1 - \lambda \\ \frac{\partial g}{\partial y} &= \log y + 1 - \lambda \\ \frac{\partial g}{\partial z} &= \log z + 1 - \lambda \end{aligned} \tag{5}$$

$$\nabla g = (\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}) \tag{6}$$

(b).

Let $\nabla g = 0$, we get $x = y = z = e^{\lambda - 1}$. Thus, $\min f$ is attained only if $x = y = z$.

E2.1

(a)

$p(X = x) = \frac{1}{2^x}$,

$$\begin{aligned} H(X) &= -\sum_{n=1}^{\infty} p(n) log(p(n)) \\ &= \log 2 \sum_{n=1}^{\infty} \frac{n}{2^n} \\ &= 2 \log 2 \\ &= 2 \text{ bits} \end{aligned} \tag{7}$$

(b)

We design the sequence to be: "Is X = 1? If not, is X = 2? if not, is X = 3,? if not, is X = 4? if not,...."

The expected number of the questions is up to the expected number of flips required. We have:

$$E(X) = \sum_{n=1}^{\infty} \frac{n}{2^n} = 2 \tag{8}$$

Here $H(X) = E(X) = 2$.

E2.2

(a)

$\because Y = 2^X$ is an injective function

$\therefore$

$$H(X)=-\sum_{X}^{\infty}p(x)log(p(x))=-\sum_{Y}^{\infty}p(y)log(p(y))=H(Y) \tag{9}$$

(b)

Let denote $y=f(x)$. Then

$$p(y)=\sum_{y=f(x)}p(x) \tag{10}$$

Obviously, $p(x)\le p(y)$.

$$\begin{aligned} H(X) &= -\sum_x p(x)\log p(x) \\ &= -\sum_y \sum_{y=f(x)} p(x)\log p(x) \\ &\ge -\sum_y \sum_{y=f(x)} p(x)\log p(y) \\ &= -\sum_y p(y)\log p(y) \\ &= H(Y) \end{aligned} \tag{11}$$

Equality is reached only when $p(y)=\sum_{y=f(x)}p(x)=p(x)$, which means $y=f(x)$ is injective.

$\therefore H(X)\ge H(Y)$, with equality if $Y=CosX$ is injective on its domain.

E2.5

$\because H(Y|X)+H(X)=H(X,Y), H(Y|X)=0$

$$H(X,Y)=-\sum_x\sum_y p(x,y)\log p(x,y) \tag{12}$$

$$H(X)=-\sum_x\sum_y p(x,y)\log p(x) \tag{13}$$

Since $H(X)-H(X,Y)=0$

$$H(X)-H(X,Y)=\sum_x\sum_y p(x,y)\log p(y|x) \tag{14}$$

$\because p(x,y)\log p(y|x)\le 0$

$\therefore$ the equality is reached only when $p(y|x)=1$ always holds, $(p(x,y)\neq 0)$

Which means Y is determined by X.

$\therefore Y$ is a function of X.

E2.6

According to the chain rule, we have:

$$I(X;Y)=I(X;Y\mid Z)+I(Z;Y)-I(Z;Y\mid X) \tag{15}$$

Then we could just focus on $\Delta=I(Z;Y)-I(Z;Y\mid X)$.

(a)

if Z and Y are independent, then $I(X;Y)\le I(X;Y\mid Z)$

Ex.

> *X, Z and Y could all be independent Bernoulli(1/2)*

(b)

if Z and Y are independent given the condtion X, then $I(X;Y)\ge I(X;Y\mid Z)$

Ex.

> *We could design a random distribution like below:*
>
> *p(Y=0|X=0) = 1 / 4, p(Y=1|X=0) = 3 / 4,*
>
> *p(Z=0|X=0) = 1 / 3, p(Z=1|X=0) = 2 / 3,*
>
> *p(Y=0|X=1) = 1 / 5, p(Y=1|X=1) = 4/ 5,*

> $p(Z=0|X=1) = 1/2$, $p(Z=1|X=1) = 1/2$,
>
> $p(X=0) = 1/6$, $p(X=1) = 5/6$

E2.7

(a)

For n coins, there are 2n+1 situations:

1. All of equal weight
2. One of them is ligher
3. One of them is heavier

Meanwhile, each weighing could have three outcomes:

1. Equally weigh
2. Left > right
3. Right > left

Thus, we can at most distinguish 3^k different pattern. Let

$$3^k \geq 2n+1 \tag{16}$$

We derive $n \leq (3^k - 1)/2$

(b)

By applying single error Hamming code, we could solve this problem like this: let denote the coin index in tenary system as below:

$$\begin{array}{c|cccccccccccc|c}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & n \\
\hline
i=1,3^0 & 1 & -1 & 0 & 1 & -1 & 0 & -1 & -1 & 0 & 1 & 1 & 0 & \Sigma_1 = 0 \\
i=2,3^1 & 0 & 1 & 1 & 1 & -1 & -1 & 1 & 0 & 0 & 0 & -1 & -1 & \Sigma_2 = 0 \\
i=3,3^2 & 0 & 0 & 0 & 0 & 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 & \Sigma_3 = 0
\end{array} \tag{17}$$

Then our strategy is: for weighting time i, we place all these coins(index denoted by n) by following rules,

- If $n_i = -1$, left pan
- If $n_i = 0$, aside
- If $n_i = 1$, right pan

The result for each weighting follows rules below:

- equal: 0
- left is heavier: -1
- right is heavier: 1

1. If the weighting result is (r1, r2, r3), which is in the alphabet, then the corresponding heavier number will be $3^0 r_1 + 3^1 r_2 + 3^2 r_3$,.
2. If the weighing result is the opposite of a certain coding of coin n, then it means coin n is lighter.