Homework 2

Q1

2.8

With replacement

Let X_i denote the color of ball from i^{th} draw.

$$X_i=\begin{cases}\text{red} & \text{with prob. } \frac{r}{r+w+b}\\ \text{white} & \text{with prob. } \frac{w}{r+w+b}\\ \text{black} & \text{with prob. } \frac{b}{r+w+b}\end{cases} \tag{1}$$

$i=1,2,3,\ldots$

$\therefore H\left(X_i \mid X_{i-1},\ldots,X_1\right)=H\left(X_i\right)$

Without replacement

As we know, the unconditional probability of the color of the ball in kth drawing is red is still $\frac{r}{r+w+b}$, which means $H(X_i)$ is still the same with the replacement case.

However, we know that the conditional entropy is less than the unconditional entropy, i.e.

$$H(X_i) \geq H\left(X_i \mid X_{i-1},\ldots,X_1\right) \tag{2}$$

In this case, the equality is dismissed. Therefore:

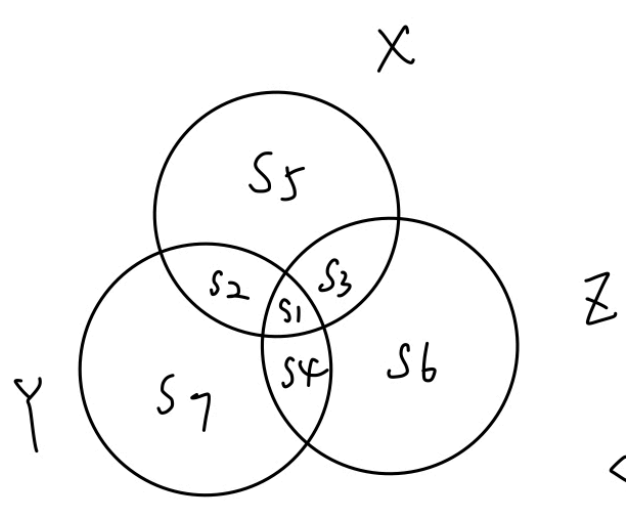
$$\sum_{i=1}^{k} H(X_i) > \sum_{i=1}^{k} H\left(X_i \mid X_{i-1},\ldots,X_1\right) \tag{3}$$

Thus the without replacement one has lower entropy.

2.9

(a)

1. Since entropy is always larger than or equal to zero, first property holds.
2. $\because \rho(X,Y)=H(X|Y)+H(Y|X)=\rho(Y,X)$, $\therefore$ second property holds.
3. To prove third property, we could use information diagram below:



$H(X|Y)+H(Y|X)+H(Y|Z)+H(Z|Y) \geq H(X|Z)+H(Z|X)$

$\Leftrightarrow$

$S_5+S_3+S_7+S_4+S_7+S_2+S_3+S_6$

$\geq S_5+S_2+S_6+S_4$ 成立

4. $\rho(x,y)=0 \Rightarrow H(X|Y)=H(Y|X)=0 \Rightarrow X=f(Y), Y=f^{-1}(X)$
   $\therefore$ there is a one-to-one function mapping from X to Y.

(b)

$$\begin{aligned}\rho(x,y) &= H(X|Y)+H(Y|X)\\ &= H(X,Y)-H(Y)+H(X,Y)-H(X)\\ &= 2H(X,Y)-H(X)-H(Y)\\ &= H(X)+H(Y)-2I(X;Y)\\ &= H(X,Y)-I(X;Y)\end{aligned} \tag{4}$$

2.10

$$
\begin{aligned}
H(X) &= -\sum_{x\in X_1} \alpha p_1(x)\log \alpha p_1(x) - \sum_{x\in X_2}(1-\alpha)p_2(x)\log(1-\alpha)p_2(x) \\
&= \alpha H(X_1) + (1-\alpha)H(X_2) - \alpha\log\alpha - (1-\alpha)\log(1-\alpha)
\end{aligned} \tag{5}
$$

We let $f(\alpha) = H(X)$

$$
f'(\alpha) = H(x_1) - H(x_2) - \log\frac{\alpha}{1-\alpha} \tag{6}
$$

Make $f'(\alpha) \geq 0, \alpha \leq \frac{t}{1+t}$, where $t = 2^{H(x_1)-H(x_2)}$.

$$
H(X) = f(\alpha) \leq f(\frac{t}{1+t}) = 2^{H(x_2)+\log(1+t)} = 2^{H(x_2)}\cdot(2^{H(x_1)-H(x_2)}+1) = 2^{H(x_1)} + 2^{H(x_2)} \tag{7}
$$

This means the effective alphabet size of X is smaller than the sum of the effective alphabet size of its two subset-alphabet X_1, X_2.

2.14

(a)

$p(Z=z \mid X=z) = p(Y=z-z \mid X=z)$.

$$
\begin{aligned}
H(Z\mid X) &= \sum_x H(Z\mid X=x) \\
&= -\sum_x p(x)\sum_z p(Z=z\mid X=x)\log p(Z=z\mid X=x) \\
&= -\sum_x p(x)\sum_z p(Y=z-x\mid X=x)\log p(Y=z-x\mid X=x) \\
&= -\sum_x p(x)\sum_y p(Y=y\mid X=x)\log p(Y=y\mid X=x) \\
&= \sum_x p(x)H(Y\mid X=x) \\
&= H(Y\mid X).
\end{aligned} \tag{8}
$$

If X and Y are independent, then $H(Y\mid X) = H(Y)$.

$\because I(X;Z) \geq 0$,

$\therefore H(Z) \geq H(Z\mid X) = H(Y\mid X) = H(Y)$.

Similarly, we can show that $H(X) \leq H(Z)$.

(b)

Let

$$
X = \begin{cases} 1 & p = 1/2 \\ 0 & p = 1/2 \end{cases} \tag{9}
$$

Meanwhile, Y is always equal to -X.

$\therefore H(X) = H(Y) = 1$, but Z is alway equal to 0.

$\therefore H(Z) = 0$.

$\therefore H(X) = H(Y) > H(Z)$

(c)

$$
H(Z) \leq H(Z|X) + H(X) = H(Y|X) + H(X) = H(X,Y) \leq H(Y) + H(X) \tag{10}
$$

If equality holds, then Y and X are independent and (X,Y is also a function of Z.

2.15

By the chain rule of mutual information:

$$
I\left(X_1; X_2,\ldots,X_n\right) = I\left(X_1;X_2\right) + I\left(X_1;X_3\mid X_2\right) + \cdots + I\left(X_1;X_n\mid X_2,\ldots,X_{n-2}\right) \tag{11}
$$

Accorrding to the property of Markov chain, $I\left(X_1;X_k\mid X_2,\ldots,X_{k-2}\right) = 0, k = 2,3,\ldots,n-1,n$

$\therefore I\left(X_1;X_2,\ldots,X_n\right) = I\left(X_1;X_2\right)$

Q2

1. $I(X;Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y)log\frac{p(x,y)}{p(x)p(y)}$

$$
\begin{aligned}
H(X) - H(X|Y) &= \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(y)} - \sum_{x \in X} \sum_{y \in Y} p(x,y) \log p(x) \\
&= I(X;Y)
\end{aligned} \tag{12}
$$

2. $I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) log \frac{p(x,y)}{p(x)p(y)}$

$$
\begin{aligned}
H(X) + H(Y) - H(X,Y) &= -\sum_{x \in X} \sum_{y \in Y} p(x,y) log\, p(x) - \sum_{x \in X} \sum_{y \in Y} p(x,y) log\, p(y) + \sum_{x \in X} \sum_{y \in Y} p(x,y) log\, p(x,y) \\
&= \sum_{x \in X} \sum_{y \in Y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} = I(X;Y)
\end{aligned} \tag{13}
$$

3.

$$
\begin{aligned}
I(X;Y|Z) &= \sum_{x \in X} \sum_{y \in Y} \sum_{z \in Z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} \\
&= \sum_{x \in X} \sum_{y \in Y} \sum_{z \in Z} p(x,y,z) \log p(x,y|z) - \sum_{x \in X} \sum_{y \in Y} \sum_{z \in Z} p(x,y,z) \log p(y|z) - \sum_{x \in X} \sum_{y \in Y} \sum_{z \in Z} p(x,y,z) \log p(x|z) \\
&= H(X|Z) + H(Y|Z) - H(X,Y|Z)
\end{aligned} \tag{1
$$

Q3

1. **Sufficiency**:
   $\because p(x) = q(x)$
   $\therefore D(p||q) = -\sum_{x \in X} p \log \frac{q}{p} = 0$

   Necessity:

$$
-D(p||q) = \sum_{x \in X} p \log \frac{q}{p} \leq \log \sum_{x \in X} p \frac{q}{p} = \log_{x \in X} \sum q \leq \log 1 = 0 \tag{15}
$$

   By Jensen Inequality's condition of equality, $\frac{q(x)}{p(x)} = Const, \forall x \in X$.

   $\therefore \log Const = 0, Const = 1$

   $\therefore p(x) = q(x)$

2. **Sufficiency**:

   $\because X, Y$ are independent

   $\therefore p(x,y) = p(x)p(y)$

   $\therefore I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} = 0$

   Necessity:

   $\therefore -I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) log \frac{p(x)p(y)}{p(x,y)} \leq 0$

   Similar to equation(4), by Jensen Inequality's condition of equality, we have $p(x)p(y) = p(x,y)$, which means X and Y are independent.

3. According to the conclusion reached in Q3.1, we have

$$
D(p(y|x)||q(y|x)) = 0 \Longleftrightarrow p(y|x) = q(y|x), \text{for all } y \text{ and } x \text{ such that } p(x) > 0 \tag{16}
$$

   Here $p(x) > 0$ is to gurantee $q(y|x) \neq 0$, which is enough for obtaining the conclusion. Note $q(x) = p(x), \forall x \in X$ is unnecessary here.

4. $I(X;Y|Z) = \sum_{x \in X} \sum_{y \in Y} \sum_{z \in Z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)}$

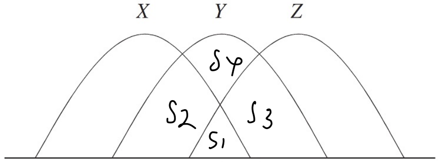
   According to the conclusion in Q3.2, we have $I(X;Y|Z) = 0$ if and only if X and Y are conditionally independent given Z.

Q4

$$
p(z|x,y) = p(z|y) \Longleftrightarrow \frac{p(x,y,z)}{p(x,y)} = \frac{p(y,z)}{p(y)} \Longleftrightarrow \frac{p(x,y,z)}{p(y,z)} = \frac{p(x,y)}{p(y)} \Longleftrightarrow p(x|y,z) = p(x|y) \tag{17}
$$

Q5

1.



$X \perp Y \mid Z \Rightarrow S_2 = 0$

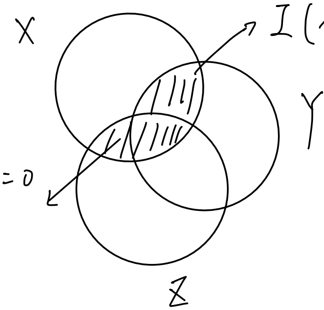
$X \perp Z \Rightarrow S_1 = 0$

$\therefore S_1 + S_2 = 0$

$\therefore X \perp Y$

2.

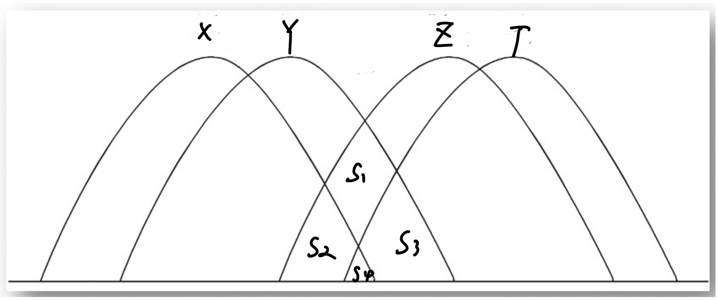
$I(X;Y \mid Z) = 0$



$I(X;Z) = 0$

$\therefore X \perp Y$

3.



$Y \perp Z \mid T \Rightarrow$

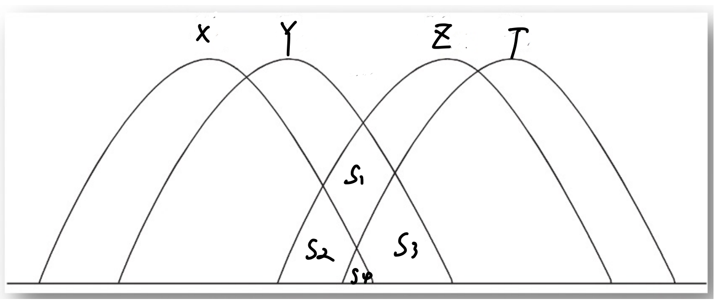
$S_1 + S_2 = 0, \; S_1 = S_2 = 0$

$I(Y; Z \mid (X, T))$

$= S_1 = 0$

$\therefore Y \perp Z \mid (X, T)$

4. ①



$I(X;T) + I(Y;Z)$

$= 2S_4 + S_1 + S_2 + S_3$

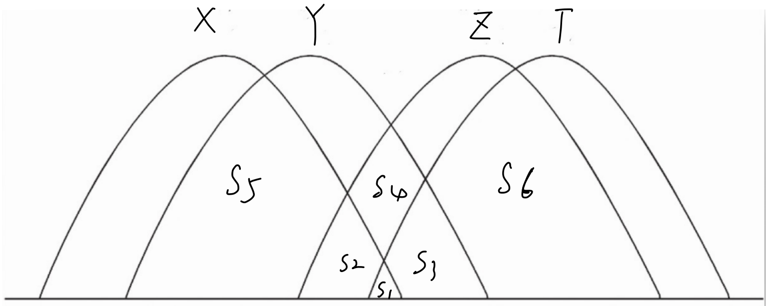
$I(X;Z) + I(Y;T)$

$= S_2 + S_4 + S_3 + S_4$

$\because S_1 \geq 0$

$\therefore$ 不等式成立

②

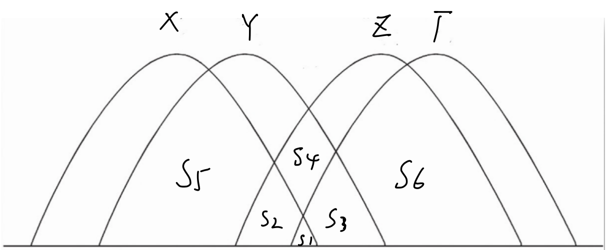


$I(X;T) + I(Y;Z) = 2S_1 + S_2 + S_3 + S_4$

$I(X;Y) + I(Z;T) = S_5 + S_2 + S_1 + S_6 + S_3 + S_1$

S_4 与 $S_5 + S_6$ 无法比较大小，故不等式不恒成立

③



$I(X;Y) + I(Z;T) = S_5 + S_2 + S_1 + S_6 + S_3 + S_1$

$I(X;Z) + I(Y;T) = S_2 + S_1 + S_1 + S_3$

∴ 不等式成立.