Exercise 2.1

Proof for $\mathcal{X} = \{0, 1, \cdots k-1\}$, we set $\{P_1, P_2 \cdots P_k\}$ be the pmf for $\mathcal{X}$

And following the $f(x) = -t\log t$ is concave function in $[0, 1]$

~~H(x)~~ And there is a constraint $\sum_{i=1}^{k} P_i = 1$

so, $H(x) = \sum_{x\in\mathcal{X}} P_x \log \frac{1}{P_x} = E[\log \frac{1}{P_x}]$

By using Jensen's inequality, we can rewrite as

$$H(x) = E[\log \frac{1}{P_x}] \le \log E[\frac{1}{P_x}]$$

$$= \log \sum_{x\in X} P_x \cdot \frac{1}{P_x}$$

$$= \log |\mathcal{X}|$$

As desired, we prove that is upper bound of entropy.

Exercise 2.2

a). Firstly, we know $D(P_x \| V_x) = \sum_{x\in X} P_x(x) \log \frac{P_x(x)}{V_x(x)} = \sum_{x\in X} P_x(x) \log \frac{P_x(x)}{\frac{1}{|X|}}$

$$\cancel{= \sum_{x\in X} P_x(x) \log \frac{V_x(x)}{P_x(x)}}$$

$$H(x) = \sum_{x\in X} P_x(x) \log \frac{1}{P_x(x)} = -\sum_{x\in X} P_x(x) \log P_x(x)$$

$$= -\sum_{x\in X} P_x(x) \log \left[\frac{1}{|X|} \cdot \frac{P_x(x)}{\frac{1}{|X|}}\right]$$

$$= -\sum_{x\in X} P_x(x) \log \frac{1}{|X|} - \sum_{x\in X} P_x(x) \log \frac{P_x(x)}{\frac{1}{|X|}}$$

$$= -\log \frac{1}{|X|} - D(P_x \| V_x)$$

$$= \log |X| - D(P_x \| V_x)$$

b) Firstly, we know $H(X|Y) = \sum_{x\in X, Y\in Y} P(x,y) \log \frac{1}{P(X|Y)}$

$$= -\sum_{x\in X, Y\in Y} P(x,y) \log \left[\frac{1}{|X|} \cdot \frac{P(X|Y)}{\frac{1}{|X|}}\right]$$

$$= -\sum_{x\in X, Y\in Y} P(x,y) \log \frac{1}{|X|} - \sum_{x\in X, Y\in Y} P(x,y) \log \frac{P(X|Y)}{\frac{1}{|X|}}$$

$$= -\log \frac{1}{|X|} - \sum_{x\in X, Y\in Y} P(x,y) \log \frac{P(x,y)}{P_Y \cdot \frac{1}{|X|}}$$

$$= \log |X| - D(P_{XY} \| V_X \times P_Y)$$

扫描全能王 创建

c)

From (a) and (b), we can get

$$I(X:Y) = H(X) - H(X|Y)$$
$$= \log|X| - D(P_X \| U_X) - [\log|X| - D(P_{XY} \| U_X \times P_Y)]$$
$$= D(P_{XY} \| U_X \times P_Y) - D(P_X \| U_X)$$
$$= \sum_{x\in X, y\in Y} P(x,y) \log \frac{P(x,y)}{P_Y \frac{1}{|X|}} - \sum_{x\in X} P_X(x) \log \frac{P_X(x)}{\frac{1}{|X|}}$$
$$= \sum_{x\in X, y\in Y} P(x,y) \log \frac{P(x,y)}{P_Y \cdot \frac{1}{|X|}} - \sum_{x\in X, y\in Y} P(x,y) \log \frac{P_X(x)}{\frac{1}{|X|}}$$
$$= \sum_{x\in X, y\in Y} P(x,y) \cdot \log \frac{P(x,y)}{P_X \cdot P_Y}$$
$$= D(P_{XY} \| P_X \times P_Y)$$

Exercise 2.3

By setting the alphabet size at most 2 bits

a). $H(X|YZ)=0 \Rightarrow H(X|YZ) = \sum_Y P_Y \sum_Z P_Z H(X|Y=y, Z=z) = 0.$

$\Rightarrow P_{X|YZ} = 0$ or $P_{X|YZ} = 1$

$H(X|Y) = H(X|Z) = 1 \Rightarrow \sum_Y P_Y H(X|Y=y) = \sum_Z P_Z H(X|Z=z) = 0$

$\Rightarrow P_{X|Y} = \frac{1}{2}$ and $P_{X|Z} = \frac{1}{2}$

For example

	X	Y	Z
①	0	0	0
②	1	0	1
③	1	1	0
④	0	1	1

b). $I(X:Y|Z) = 1 \Rightarrow H(Y|Z) - H(Y|XZ) = 1 \Rightarrow H(Y|Z) = 1$ and $H(Y|XZ) = 0$

$\Rightarrow P_{Y|Z} = \frac{1}{2}$ and $[P_{Y|XZ} = 0$ or $P_{Y|XZ} = 1]$

$I(X:Y) = 0 \Rightarrow H(Y) - H(Y|X) = 0$

For example

	X	Y	Z
①	0	0	1
②	0	1	0
③	1	0	0
④	1	1	1

扫描全能王 创建

c). $I(X:Y)=1 \Rightarrow H(Y)-H(Y|X)=1 \Rightarrow H(Y)=1$ and $H(Y|X)=0$

~~$I(X:Y|Z) = H(Y|Z)$~~ $\Rightarrow Y$ is uniform distributed

$\Rightarrow P_{Y|X}=1$ or 0

And $I(X:Y|Z)=H(Y|Z)-H(Y|XZ)=0$

for example:

	X	Y	Z
①	0	0	0
②	0	1	1
③	1	0	1
④	1	1	0

d) $I(X:Y)=I(X:Z)=1 \Rightarrow H(Y)-H(Y|X)=1$ and $H(Z)-H(Z|X)=1$

But for $I(Y:Z)=0 \Rightarrow$ $P_{Y|X}=0$ or 1 / $P_{Z|X}=0$ or 1

so, for example:

	X	Y	Z
①	0	1	1
②	1	0	0

Exercise 2.5

For the first part, if we know $X_1, X_2 \cdots X_n$ are mutually independent

so, $I(X_1,\cdots X_n : Y_1 \cdots Y_n) = H(X_1,X_2\cdots X_n) - H(X_1,X_2\cdots X_n|Y_1,Y_2\cdots Y_n)$

$= H(X_1)+H(X_2|X_1)+\cdots H(X_n|X_{n-1}) - H(X_1,X_2\cdots X_n|Y_1,Y_2\cdots Y_n)$

Because $X_1,\cdots X_n$ are mutual independent, we can rewrite above equation

$= H(X_1)+H(X_2)+\cdots H(X_n) - H(X_1,X_2\cdots X_n|Y_1,Y_2\cdots Y_n)$

$= H(X_1)+H(X_2)+\cdots H(X_n) - [H(X_1|Y_1Y_2\cdots Y_n) + H(X_2|X_1,Y_1\cdots Y_n) + \cdots + H(X_n|X_1\cdots X_{n-1},Y_1\cdots Y_n)]$

~~Owing $H(X_1|Y_1) \le H(X_1|Y_1$~~

$= \sum_{i=1}^{n} H(X_i) - [H(X_1|Y_1Y_2\cdots Y_n) + \cdots + H(X_n|Y_1\cdots Y_n)]$

Owing to $H(X_1|Y_1 Y_2\cdots Y_n) \le H(X_1|Y_1)$, $H(X_2|Y_1\cdots Y_n) \le H(X_2|Y_2)$

$\cdots, H(X_n|Y_1\cdots Y_n) \le H(X_n|Y_n)$

so for above ~~equa~~ statement, we can rewrite

$\ge \sum_{i=1}^{n} H(X_i) - [H(X_1|Y_1) + \cdots H(X_n|Y_n)]$

$= \sum_{i=1}^{n} H(X_i) - \sum_{i=1}^{n} H(X_i|Y_i)$

$= \sum_{i=1}^{n} I(X_i:Y_i)$

扫描全能王 创建

And for the second part, if we know X_i is ~~to~~ conditional independent of all the remaining random variables, given Y_i)

so, $I(X_1 \cdots X_n : Y_1 \cdots Y_n) = H(X_1, X_2 \cdots X_n) - H(X_1, X_2 \cdots X_n \mid Y_1, Y_2 \cdots Y_n)$

$= H(X_1) + H(X_2 \mid X_1) + \cancel{H(\ )} + \cdots H(X_n \mid X_{n-1})$

$- [H(X_1 \mid Y_1 Y_2 \cdots Y_n) + H(X_2 \mid X_1, Y_1 \cdots Y_n) + \cdots + H(X_n \mid X_1 \cdots X_{n-1}, Y_1 \cdots Y_n)$

$= H(X_1) + H(X_2 \mid X_1) + \cdots H(X_n \mid X_{n-1})$

$- [H(X_1 \mid Y_1) + H(X_2 \mid Y_2) + \cdots H(X_n \mid Y_n)]$

Owing to $\cancel{H(X_2) \le}$ $H(X_2 \mid X_1) \le H(X_2)$

$\vdots$

$H(X_n \mid X_{n-1}) \le H(X_n)$

Therefore, above equation, we can rewrite

$\le H(X_1) + H(X_2) \cdots H(X_n) - [H(X_1 \mid Y_1) + H(X_2 \mid Y_2) + \cdots H(X_n \mid Y_n)]$

$= \sum_{i=1}^{n} H(X_i) - \sum_{i=1}^{n} H(X_i \mid Y_i)$

$= \sum_{i=1}^{n} I(X_i : Y_i)$

In a word, the proof is finished.

扫描全能王 创建