CS258 Information Theory Homework 2

Zhou Litao 518030910407 F1803016

March 14, 2020

Exercise 1 Show that $D(p\|q) = 0$ if and only if $p(x) = q(x)$.

Proof.

$$\begin{aligned}
-D(p(x)\|q(x)) &= \sum_{x\in\mathcal{X}} p(x)\log\frac{q(x)}{p(x)} \\
&\leq \log(\sum p(x)\frac{q(x)}{p(x)}) \qquad\qquad \text{(By Concavity of log(x))} \\
&= \log(\sum q(x)) \leq \log 1 = 0
\end{aligned} \tag{1}$$

The first equality holds if and only if

$$\frac{q(x)}{p(x)} = k, \text{ for every } x \in \mathcal{X} \text{ such that } p(x) > 0$$

The second equality holds if and only if there exists no x such that $p(x) = 0$ while $q(x) > 0$. Since we have $\sum p(x) = 1$, we know that

$$\frac{q(x_1)}{p(x_1)} = \frac{q(x_2)}{p(x_2)} = \cdots = \frac{\sum q(x)}{\sum p(x)}.$$

By the second condition we know that $\sum q(x) = \sum p(x) = 1$. Hence $p(x) = q(x)$ for all $x \in \mathcal{X}$.

□

Exercise 2 Show that $I(X;Y) \geq 0$, with equality if and only if X and Y are independent.

Proof.

$$\begin{aligned}
-I(X;Y) &= \sum_{(x,y)\in\mathcal{X}\star\mathcal{Y}} p(x,y)\log\frac{p(x)p(y)}{p(x,y)} \\
&\leq \log\left(\sum_{(x,y)\in\mathcal{X}\star\mathcal{Y}} p(x,y)\frac{p(x)p(y)}{p(x,y)}\right) \\
&= \log\sum_{(x,y)\in\mathcal{X}\star\mathcal{Y}} p(x)p(y) = \log\left(\sum_{x\in\mathcal{X}} p(x)\sum_{y\in\mathcal{Y}} p(y)\right) = \log 1 = 0
\end{aligned} \tag{2}$$

The equality holds if and only if

$$\frac{p(x,y)}{p(x)p(y)} = k, \text{ for every } (x,y) \in \mathcal{X}\star\mathcal{Y} \text{ such that } p(x,y) > 0$$

Since $\sum p(x,y) = 1$, we know that $k = 1$. That is to say, $p(x,y) = p(x)p(y)$ for every possible x, y. X and Y are independent.

□

Exercise 3 Show that $D(p(y|x)\|q(y|x)) \geq 0$ with equality if and only if $p(y|x) = q(y|x)$ for all x and y such that $p(x) > 0$.

Proof.

$$\begin{aligned} -D(p(y|x)\|q(y|x)) &= \sum_{x\in\mathcal{X}} p(x) \sum_{y\in\mathcal{Y}} p(y|x) \log \frac{q(y|x)}{p(y|x)} \\ &\leq \sum_{x\in\mathcal{X}} p(x) \log \sum_{y\in\mathcal{Y}} q(y|x) \\ &\leq \sum_{x\in\mathcal{X}} p(x) \log 1 = 0 \end{aligned} \tag{3}$$

The first equality holds if and only if

$$\frac{q(y|x)}{p(y|x)} = k, \text{ for every } y \in \mathcal{Y} \text{ such that } p(y|x) > 0 \text{ given } p(x) > 0 \text{ with } x \in \mathcal{X}$$

The second equality holds if and only if given $p(x) > 0$, there exists no y such that $p(y|x) = 0$ while $q(y|x) > 0$. Since we have $\sum p(y|x) = 1$, we know that

$$\frac{q(y_1|x)}{p(y_1|x)} = \frac{q(y_2|x)}{p(y_2|x)} = \cdots = \frac{\sum q(y|x)}{\sum p(y|x)}.$$

It follows from the second condition that $\sum q(y|x) = \sum p(y|x) = 1$. Hence the equality holds if and only if $p(y|x) = q(y|x)$ for all $p(x) > 0$.

□

Exercise 4 Show that $I(X;Y|Z) \geq 0$ with equality if and only if X and Y are conditionally independent given Z.

Proof.

$$\begin{aligned} -I(X;Y|Z) &= \sum_{(x,y,z)\in\mathcal{X}\star\mathcal{Y}\star\mathcal{Z}} p(x,y,z) \log \frac{p(x|z)p(y|z)}{p(x,y|z)} \\ &\leq \sum_{z\in\mathcal{Z}} p(z) \log \sum_{(x,y)\in\mathcal{X}\star\mathcal{Y}} p(x,y|z) \frac{p(x|z)p(y|z)}{p(x,y|z)} \\ &= \sum_{z\in\mathcal{Z}} p(z) \log \left(\sum_{x\in\mathcal{X}} p(x|z) \sum_{y\in\mathcal{Y}} p(y|z) \right) \\ &= \sum_{z\in\mathcal{Z}} p(z) \log 1 = 0 \end{aligned} \tag{4}$$

The equality holds if and only if

$$\frac{p(x,y|z)}{p(x|z)p(y|z)} = k, \text{ for every } (x,y) \in \mathcal{X} \star \mathcal{Y} \text{ such that } p(x,y) > 0 \text{ given } p(z) > 0 \text{ with } z \in \mathcal{Z}$$

Since $\sum p(x,y|z) = 1$, we know that $k = 1$. That is to say, $p(x,y|z) = p(x|z)p(y|z)$ for every possible x, y given $p(z) > 0$. Therefore the equality holds if and only if X and Y are independent given Z.

□

Exercise 5 Let $u(x) = \frac{1}{|\mathcal{X}|}$ be the uniform probability mass fuction over X, and let $p(x)$ be the probability mass function for X, Then

$$0 \leq D(p\|u) = \log|\mathcal{X}| - H(X)$$

Proof. From Exercise 1 we know $D(p\|u) \geq 0$. By definition of mutual entropy we have

$$\begin{aligned} D(p\|u) &= \sum_{x\in\mathcal{X}} p(x) \log \frac{p(x)}{u(x)} = \sum_{x\in\mathcal{X}} p(x) \log |\mathcal{X}| p(x) \\ &= \log |\mathcal{X}| \sum_{x\in\mathcal{X}} p(x) - \sum_{x\in\mathcal{X}} p(x) \log p(x) = \log |\mathcal{X}| - H(X) \end{aligned} \tag{5}$$

□

Exercise 6 (Conditioning reduces entropy) Show that

$$H(X|Y) \leq H(X)$$

with equality if and only if X and Y are independent.

Proof. We know that $I(X;Y) = H(X) - H(X|Y)$. From Exercise 2 we know that $I(X;Y) \geq 0$. It follows that $H(X|Y) \leq H(X)$ with equality if and only if X and Y are independent. □