EE411 Information Theory and Coding

Homework #1

Site Fan
fanst2021@mail.sustech.edu.cn

Exercise 1

Entropy of functions. Let X be a random variable taking on a finite number of values.What is the (general) inequality relationship of $H(X)$ and $H(Y)$ if

(a) $Y = 2^X$?

(b) $Y = \cos X$?

Solution

(a) $Y = 2^X$. If X takes on a finite number of values, then 2^X will also take on a finite number of values. The transformation from X to 2^X is a bijection, i.e., each X maps uniquely to a value of 2^X and vice versa. Therefore $H(X) = H(Y)$.

(b) $Y = \cos X$. The mapping $X \to \cos X$ is not one-to-one. Multiple values of X can map to the same value of $\cos X$. Hence the uncertainty in the original distribution is reduced when transformed to $\cos X$. Therefore $H(X) \geq H(Y)$.

Exercise 2

Minimum entropy. What is the minimum value of $H(p_1, ..., p_n) = H(\mathtt{p})$ as $\mathtt{p}$ ranges over the set of n -dimensional probability vectors?

Find all $\mathtt{p}$'s which achieve this minimum.

Solution

Minimize $H(\mathtt{p}) = -\sum\limits_{i=1}^{n} p_i \log p_i$, where $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$

The entropy is minimized when there exists $p_i = 1$ such that for any $j \neq i$, $p_j = 0$.

The minimum $H(\mathtt{p}) = 0$, which is achieved when the probability mass concentrates on a single element, i.e., there is only one $p_i = 1$ while others $p_j = 0$.

Exercise 3

Entropy of functions of a random variable. Let X be a discrete random variable. Show that the entropy of a function of X is less than or equal to the entropy of X by justifying the following steps:

$$H(X, g(X)) \stackrel{(a)}{=} H(X) + H(g(X)|X) \tag{2.166}$$

$$\stackrel{(b)}{=} H(X); \tag{2.167}$$

$$H(X, g(X)) \stackrel{(c)}{=} H(g(X)) + H(X|g(X)) \tag{2.168}$$

$$\stackrel{(d)}{\geq} H(g(X)). \tag{2.169}$$

Thus $H(g(X)) \leq H(X)$.

Solution

Using the Chain Rule, we have

$$H(X, g(X)) \overset{(a)}{=} H(X) + H(g(X)|X)$$

$$H(X, g(X)) \overset{(c)}{=} H(g(X)) + H(X|g(X))$$

Since x to $g(x)$ is an **onto** mapping, for any $x \in \mathcal{X}$, $p(g(x)|x) = 1$, therefore $H(g(X)|X) = 0$, $H(X, g(X)) \overset{(b)}{=} H(X)$.

On the other hand, x to $g(x)$ may be not **one-to-one** (e.g. $X \to \cos X$), $p(x|g(x)) \leq 1$, therefore $H(X|g(X)) \geq 0$, $H(X, g(X)) \overset{(d)}{\geq} H(g(X))$.

From (a) to (d), we have $H(g(X)) \leq H(X)$.

Exercise 4

Zero conditional entropy. Show that if $H(Y|X) = 0$, then Y is a function of X, i.e., for all x with $p(x) > 0$, there is only one possible value of y with $p(x, y) > 0$.

Solution

For conditional entropy $H(Y|X) = -\sum_x p(x) \sum_y p(y|x) \log p(y|x) = 0$, we can imply that for any $p(x) > 0$, the inner term $\sum_y p(y|x) \log p(y|x) = 0 (\star)$.

And for a fixed x, $\sum_y p(y|x) = 1$, for any $0 < p(y|x) < 1$, $p(y|x) \log p(y|x) < 0$, meaning that to satisfy $(\star)$, there could only be one y such that $p(y|x) = 1$ and others be 0, to let $\sum_y p(y|x) \log p(y|x) = 0$.

To put it another way, for every x such that $p(x) > 0$, there is only one possible value of y with $p(y|x) = 1$, which means that Y is a function of X.