Chapter 2

Xuan

2020 年 9 月 10 日

1 离散信源熵与互信息

1.1 信源熵

$H(X) = \sum_i p(x_i) \log_2 \frac{1}{p(x_i)} = -\sum_i p(x_i) \log_2 p(x_i)$

性质：

1. $H(X) \geq 0$
2. $H(X) \leq \log_2 |X|$

当某一符号 x_i 的概率为 p_i 为零时，$p_i \log_2 p_i$ 在熵公式中无意义，为此规定这时的 $p_i \log_2 p_i$ 为零。当信源 X 中只含一个符号 x 时，必定有 $p(x) = 1$，此时信源熵 $H(x)$ 为零，是确定信源。

1.1.1 证明 $H(x) \leq \log_2 |X|$:

方法一： 求偏导，略

方法二： 利用相对熵

假设 q 服从均匀分布，即 $q_i = \frac{1}{|X|}$

$$\begin{aligned} D(p||q) &= \sum_i p_i \log_2 \frac{p_i}{q_i} && (1) \\ &= \sum_i p_i \log_2 p_i - \sum_i p_i \log_2 q_i && (2) \\ &= -H(x) - \sum_i p_i \log_2 q_i \geq 0 && (3) \\ &= -H(x) + \sum_i p_i \log_2 |X| \geq 0 && (4) \end{aligned}$$

所以我们可以得到

$$H(x) \leq \log_2 |X|$$

1.2 相对熵

1.2.1 证明 $D(p||q) \geq 0$

下凸函数的性质 $pf(x_1) + (i-p)f(x_2) \geq f(px_1 + (1-p)x_2)$

Jensen's Inequality: 对于下凸函数而言，

$$Ef(X) \geq f(EX)$$

上凸函数反之。

1. $|X| = 2$ 时，由下凸函数性质可证
2. 数学归纳法

Assume $|X| = k-1$, Jensen's Inequality holds

Prove $|X| = k$, Inequality holds as well.

Assume $\sum_{i=1}^{k-1} p(x_i) = q = 1 - p(x_k)$

$$
\begin{aligned}
Ef(X) &= p(x-k)f(x_k) + \sum_{i=1}^{k-1} p(x_i)f(x_i) && (5)\\
&= p(x_k)f(x_k) + q\sum_{i=1}^{k-1} \frac{p(x_i)}{q} f(x_i) && (6)\\
&= p(x_k)f(x_k) + (1-p(x_k))f(\sum_{i=1}^{k-1} \frac{p(x_i)}{1-p(x_k)} x_i) && (7)\\
&\geq f(p(x_k)x_k + \sum_{i=1}^{k-1} p(x_i)x_i) && (8)\\
&= f(EX) && (9)
\end{aligned}
$$