信息论课上知识

公理系统零

1. 熵
2. 联合熵
3. 条件熵
4. 相对熵
5. 互信息
6. 链式法则
7. 琴声不等式
8. 马尔科夫链
    I. 概率定义
    II. 数据处理不等式
9. 法诺不等式
10. 文氏图

公理系统一：三条公理、柯西方程、证明

1. $H(X)$是关于 p 的连续函数
2. $H(X) = H(1/n, 1/n \dots 1/n)$ 是关于 n 的单调递增函数
3. 组合原理
4. 用公理系统零验证组合原理

公理系统二：香农辛钦公理系统

1. 六条原理+条件熵的定义 ->如何用公理系统零说明六条公理?

   多个引理：

2. 当 X，Y 独立的时候 $H(X,Y) = H(X) + H(Y)$ ->

   推论：$H(X) = H(X|Y)$

   推论：$H(X1, X2 \dots Xk) = kH(X)$ Xi 独立

3. 如果 X 是有唯一取值的随机变量，那么 $H(X) = 0$
4. 如果 X 是在 A 上一致性分布，Y 是在 B 上的一致性分布，A 包含于 B ,那么 $H(X) <= H(Y)$,当且 $A = B$ 的时候等号成立
5. 单调性原理：$|A| < |B|$, A 包含于 B， X 是在 A 上一致性分布，Y 是在 B 上的一致性分布，那么 $H(X) < H(Y)$
6. 构造 1.4
   I. X 在 A 上取值， $P(X = a) = m_a/n$
   II. $U \sim [n] = \{1,2,3,4, .. n\}$
   III. V_a 是 [n] 上的一个划分，$|V_a| = m_a$
   IV. X' 满足 X' = a 当且仅当 $U \in V_a$
   V. 课下布置任务如何理解 U 和 X 的相互决定关系（如果要证明条件使熵减少会用到）
7. X 在 A 上取值：$H(X) >= 0$
8. 复合使得熵减少 $Y = f(X)$ $H(Y) <= H(X)$
9. 链式法则如何用公理系统证明->可加性原理递推即可，或者数学归纳法
10. $X \sim A$, $H(X) = \log|A|$
11. **Shearer's Lemma：** $H(X1, X2..Xn) \leq \frac{1}{k}\sum_{F \in \mathcal{F}} H(X_F)$ k 是$\forall$ Xi， Xi 至少出现在 F 中的个数

渐进均分性

1. 马尔可夫不等式
2. 切比雪夫不等式
3. 弱大数定律
4. 渐进均分性
5. 典型集

   I. $|log\frac{1}{P(X^n)} - H(X^n)| \leq \varepsilon$

   II. P(X∈ $A^n_{(\varepsilon)}$) > 1-ε when n→ ∞

   III. $|A^n_{(\varepsilon)}| \leq 2^{n(H(X)+\varepsilon)}$

   IV. $|A^n_{(\varepsilon)}| \geq (1-\varepsilon)2^{n(H(X)-\varepsilon)}$ when n→ ∞

   V. $\frac{|A^n_{(\varepsilon)}|}{|X^n|} \to 0$ when n $\to \infty$ & H($A^n_{(\varepsilon)}$) != log|X^n|

数据压缩

1. 信源编码（随机变量 X 的信源编码 C 是 X 到取值空间 D 的一个映射）
2. 信源编码的期望长度 L(C) = $\sum_x p(x)l(c(x))$
3. 编码的类型
   I. 奇异
   II. 非奇异 $x \neq x'$ $then\ c(x) \neq c(x')$
   III. 唯一可解/译 扩展编码也非奇异
   IV. 前缀码 无任何码字是另一个的前缀
4. Kraft 不等式:对所有字母表 D 上的前缀码长 l_i 满足

$$\sum_i^n |D|^{l_{\max} - l_i} \leq |D|^{l_{\max}} \Leftrightarrow \sum_i^n \frac{1}{|D|^{l_i}} \leq 1$$

   反之如果一组长度满足上式，则存在一个相对应的编码方案。

5. 推广的 Kraft 不等式。D 无限依然成立：

   $0.y_1y_2\cdots y_{\mathrm{l}_i} = \sum_{j=1}^{\mathrm{l}_i} y_jD^{-j}$ $\left[0.y_1y_2\cdots y_{\mathrm{l}_i}, 0.y_1y_2\cdots y_{\mathrm{l}_i} + \frac{1}{D^{\mathrm{l}_i}}\right)$

6. 最优码长的界 $L \geq H_D(X)$ 当且对任意 i，$\frac{1}{|D|^{l_i}} = p_i$ 的时候成立（利用相对熵证明）
7. McMillan 不等式：唯一可解码也符合 Kraft 不等式
8. 香农码：最优编码的期望码长被限制在熵与熵+1 之间：即如果设计是最优的话最多多一个单位。

   $\sum_i^n |D|^{-\lceil l_i = -log p_i \rceil} \leq \sum_i^n |D|^{-log p_i} \Rightarrow \exists\, l_i\ s.t.\ log\frac{1}{p_i} \leq l_i \leq log\frac{1}{p_i} + 1$ *满足 Kraft 不等式* $\Rightarrow H_D(X) \leq L^* < H_D(X) + 1$

9. Huffman Code
   Ⅰ. 最优编码
      1. 如果 pj > pk，那么 lj ≤ lk。
      2. 两个最长的代码词长度相同。
      3. 两个最长的代码词只在最后一个位不同，并且对应于两个最不可能的符号。

   II. 哈夫曼编码是最优编码（归纳法）

信道编码

1. 信道通信模型
2. 贝叶斯公式
3. 离散无记忆信道
    I. 无记忆：$p(y_n|x^n, y^{n-1}) = p(y_n|x_n)$
    II. 无反馈：$p(x_n|x^{n-1}, y^{n-1}) = p(x_n|x^{n-1})$
    III. 离散无记忆 $p(y^n|x^n) = \prod_{i=1}^{n} p(y_i|x_i)$
4. 操作性信道容量 C = log(|{通过通道后的可识别输入}|)
5. 信息信道容量 C = 在信道允许的输入概率分布 *p*(*x*) 中，输入 *X* 和输出 *Y* 之间的最大互信息
6. 信容量的计算：
    I. 二进制无噪声信道 （Binary Noiseless Channel）
    II. 非重叠输出的噪声信道 （Noisy Channel with Nonoverlapping Outputs）
    III. 打字机模型（Noisy Typewriter）C=log13
    IV. 二进制对称信道 （Binary Symmetric Channel）C=1−H(p)
    V. 二进制擦除信道（Binary Erasure Channel） C=1−α（α 为擦除概率）
    VI. 对称信道
7. 信道容量的性质
    I. C>=0
    II. C<=log|X| , C<=log|Y|
    III. 是关于 p(x) 的连续函数
8. (M, n) 码
9. 码率
10. 联合典型序列
11. 信道编码定理
12. 信道编码的逆定理