

## 单项选择题

1. 以下关于联合熵的命题 C 恒为真

- A.  $H(X_1, \dots, X_n) = H(X_1) + H(X_1, \dots, X_{n-1}|X_1)$
- B.  $H(X_1, \dots, X_n) < H(X_1) + H(X_1, \dots, X_{n-1}|X_1)$
- C.  $H(X_1, \dots, X_n) > H(X_1) + H(X_1, \dots, X_{n-1}|X_1)$
- D.  $H(X_1, \dots, X_n) \neq H(X_1) + H(X_1, \dots, X_{n-1}|X_1)$

2.  $F$ 是一个多对一的函数, 则以下为真的是 C

- A.  $H(F(X)) = H(X)$
- B.  $H(F(X)) > H(X)$
- C.  $H(F(X)) < H(X)$
- D. 都有可能

3. 随机变量 $X$ 和 $Y$ 独立, 有相同的概率分布,  $H(X)$ 的对数底数为 $\alpha$ , 则以下为真的是 B

- A.  $P(X=Y) = \alpha^{-H(X)}$
- B.  $P(X=Y) \leq \alpha^{-H(X)}$
- C.  $P(X=Y) \geq \alpha^{-H(X)}$
- D. 都有可能

4. 对不同的 $i$ ,  $(X_i, Y_i)$ 之间是概率独立的离散型联合随机变量, 概率分布为 $P(X, Y), P(X)$ 和 $P(Y)$ 分别为各个 $X_i$ 和 $Y_i$ 的概率分布,  $i = 1, 2, \dots, n$ . 根据大数定律, 当 $n$ 趋于无穷大时, 随机变量 $\frac{1}{n} \log \frac{P(X_1, \dots, X_n)P(Y_1, \dots, Y_n)}{P(X_1, Y_1, \dots, X_n, Y_n)}$ 的极限是 C

- A.  $H(X|Y) - H(Y|X)$
- B.  $H(X|Y) + H(Y|X)$
- C.  $I(X; Y)$
- D.  $-I(X; Y)$

5. 一个二元对称无记忆离散信道的容量为0.8比特, 信道编码采用二进制形式, 每个原始数据分组为12位, 要使译码的差错概率能够任意地接近于0, 信道码字的长度最短不能低于 C

- A. 8位
- B. 10位
- C. 16位
- D. 20位

## 分析题

1.  $X, Y$ 是随机变量, 互信息量 $I(X; Y)$ 在条件概率 $P_{y|x}$ 固定的情况下,  $X$ 的概率分布 $P_x$ 是凸函数还是凹函数?

解: 凹函数.

$$P_y = \sum_x P_x \cdot P_{y|x}, \quad \frac{\partial P_y}{\partial P_x} = P_{y|x}$$

互信息量

$$\begin{aligned} I(X; Y) &= \sum_{x,y} P(x, y) \log P_{y|x} - \sum_y P_y \log P_y \\ &= \sum_{x,y} P_x \cdot P_{y|x} \log P_{y|x} - \sum_{x,y} P_x \cdot P_{y|x} \log P_y \end{aligned}$$

其Hessian矩阵 $H$ 中每个元素为二阶导数 $\frac{\partial^2 I}{\partial P_x \partial P_{x'}}$ .

一阶导数:  $\frac{\partial I}{\partial P_x} = \sum_y P_{y|x} \log P_{y|x} - \sum_y P_{y|x} \log P_y - \sum_{x',y} P_{x'} \cdot P_{y|x'} \frac{P_{y|x}}{P_y}$

二阶导数:  $\frac{\partial^2 I}{\partial P_x \partial P_{x'}} = - \sum_y P_{y|x'} \frac{P_{y|x}}{P_y}$

对于任意向量 $\mathbf{u} = [u_1, u_2, \dots, u_n]^T$ ,  $\mathbf{u}^T H \mathbf{u}$ 的值为

$$\sum_{x,x'} u_x u_{x'} \frac{\partial^2 I}{\partial P_x \partial P_{x'}} = - \sum_y \frac{1}{P_y} \left( \sum_x P_{y|x} u_x \right)^2 \leq 0$$

所以其Hessian矩阵为半正定矩阵,  $I(X;Y)$ 是关于变量 $P_x$ 的凹函数.

## 2. 互信息量 $I(X;Y)$ 总是非负的吗?

**解:**  $I(X;Y) \geq 0$ 恒成立.

互信息量可以写作

$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

$\because p(x,y) \geq p(x) \cdot p(y)$  当且仅当 $X, Y$ 概率独立时取等号.

$\therefore \frac{p(x,y)}{p(x) \cdot p(y)} \geq 1$ , 即 $\log \frac{p(x,y)}{p(x)p(y)} \geq 0$

$\therefore I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \geq 0$

## 3. 假设一个卫星转发系统的模型表达为一个Markov链 $X \rightarrow Y \rightarrow Z$ (其中 $X \rightarrow Y$ 和 $Y \rightarrow Z$ 分别为所谓的上行链路和下行链路), 你可以得出 $I(X;Z)$ , $I(X;Y)$ 和 $I(Y;Z)$ 的值之间有怎样的关系? 陈述并予以证明.

**解:** 数据处理不等式:  $I(X;Y) \geq I(X;Z)$ ,  $I(Y;Z) \geq I(X;Z)$ .

$\because X \rightarrow Y \rightarrow Z$ 形成了Markov链,  $\therefore p(x,y,z) = p(x)p(y|x)p(z|y)$ ,

$\therefore p(x,z|y) = p(x,y,z)/p(y) = p(x|y)p(z|y)$  在给定 $Y$ 的条件下,  $X$ 与 $Z$ 互相独立

$\therefore I(X;Z|Y) = 0$

根据互信息量的链式法则:

$$\begin{aligned} I(X;Y,Z) &= I(X;Z) + I(X;Y|Z) \\ &= I(X;Y) + I(X;Z|Y) \\ &= I(X;Y) \end{aligned}$$

由于互信息量始终非负 $I(X;Y|Z) \geq 0$

$$I(X;Y) = I(X;Z) + I(X;Y|Z) \geq I(X;Z)$$

Markov链具有对称性, 故 $I(Y;Z) \geq I(X;Z)$ 的证明与 $I(X;Y) \geq I(X;Z)$ 的证明同理.

4. 有人说, 通过大幅度提升上行链路  $X \rightarrow Y$  的容量便有可能显著提升以上卫星信道的容量, 该论断正确吗?

**解:** 不正确.

通过上题的证明,  $I(X; Y) \geq I(X; Z)$  和  $I(Y; Z) \geq I(X; Z)$  同时成立,  $I(X; Z)$  上限不仅取决于  $I(X; Y)$ , 也取决于  $I(Y; Z)$ . 因此仅仅依靠大幅提高  $I(X; Y)$  的最大值, 即上行链路的信道容量, 而不对下行链路进行优化, 该信道的容量将无法突破下行链路的信道容量.

## 证明题

1. 随机变量  $X_1, X_2, \dots, X_n$  彼此概率独立具有相同的概率分布  $P(X)$ ,  $Y_1, Y_2, \dots, Y_n$  彼此概率独立具有相同的概率  $P(Y)$ ,  $H$  表达熵, 对任何正整数  $n$  和正数  $\varepsilon$  定义集合

$$A_\varepsilon^{(n)} = \{(x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n) : \left| -\frac{1}{n} \log p(x_1, x_2, \dots, x_n) - H[X] \right| < \varepsilon, \\ \left| -\frac{1}{n} \log p(x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n) - H[X, Y] \right| < \varepsilon\}$$

和集合

$$B_\varepsilon^{(n)} = \left\{ (x_1, x_2, \dots, x_n) : \left| \frac{1}{n} (x_1 + x_2 + \dots + x_n) \right| < \varepsilon \right\}$$

证明:  $n$  趋近于无穷大时有极限

$$P \left[ (x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n) \in A_\varepsilon^{(n)} \right] \rightarrow 1$$

**解:**

弱大数定律:

$$-\frac{1}{n} \log p(X^n) \rightarrow -E[\log p(X)] = H(X) \quad \text{依概率}$$

因此, 给定  $\varepsilon > 0$ , 存在  $n_1$ , 使得对于任意  $n > n_1$ ,

$$\Pr \left( \left| -\frac{1}{n} \log p(X^n) - H(X) \right| \geq \varepsilon \right) < \frac{\varepsilon}{2} \quad (1)$$

同理, 存在  $n_2$ , 使得对于任意  $n > n_2$ ,

$$\Pr \left( \left| -\frac{1}{n} \log p(X^n, Y^n) - H(X, Y) \right| \geq \varepsilon \right) < \frac{\varepsilon}{2} \quad (2)$$

那么对于任意  $n > \max(n_1, n_2)$ , 将不等式(1), (2)的集合之并的概率也小于  $\varepsilon$ . 对于充分大的  $n$ , 集合  $A_\varepsilon^{(n)}$  的概率大于  $1 - \varepsilon$ , 则

$$\lim_{n \rightarrow \infty} P \left[ (x_n, y_n) \in A_\varepsilon^{(n)} \right] = 1$$

2. 请估计集合  $A_\varepsilon^{(n)}$  的大小的上界, 并给出你的证明 (除大数定律外, 其他环节要求给出推导)

解:  $2^{n(H(X,Y)+\varepsilon)}$

$$\begin{aligned} 1 &= \sum p(x^n, y^n) \\ &\geq \sum_{A_\varepsilon^{(n)}} p(x^n, y^n) \\ &\geq |A_\varepsilon^{(n)}| 2^{-n(H(X,Y)+\varepsilon)} \end{aligned}$$

因此

$$|A_\varepsilon^{(n)}| \leq 2^{n(H(X,Y)+\varepsilon)}$$

3. 随机变量  $U_1, U_2, \dots, U_n$  彼此概率独立且每一个与  $X$  有相同的概率分布  $P(X)$ ,  $V_1, V_2, \dots, V_n$  彼此概率独立且每一个与  $Y$  有相同的概率分布  $P(Y)$ , 此外每个  $u_i$  和  $v_i$  也概率独立.  $(x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n)$  的集合  $A_\varepsilon^{(n)}$  如上, 请确定

$$P \left[ (u_1, u_2, \dots, u_n; v_1, v_2, \dots, v_n) \in A_\varepsilon^{(n)} \right]$$

的上界并给出证明.

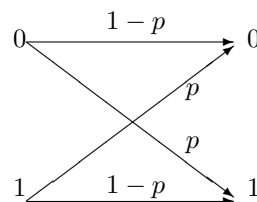
解:  $2^{-n(I(X;Y)-3\varepsilon)}$

$$\begin{aligned} \Pr((U^n, V^n) \in A_\varepsilon^{(n)}) &= \sum_{(x^n, y^n) \in A_\varepsilon^{(n)}} p(x^n) p(y^n) \\ &\leq 2^{n(H(X,Y)+\varepsilon)} 2^{-n(H(X)-\varepsilon)} 2^{-n(H(Y)-\varepsilon)} \\ &= 2^{-n(I(X;Y)-3\varepsilon)} \end{aligned}$$

## 计算题

1. 设二元对称离散无记忆信息的传输差错概率为  $p$ , 记为  $BSC(p)$ , 请计算其容量  $C$

解:  $1 - H(p)$



$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(Y) - \sum p(x) H(Y|X=x) \\ &= H(Y) - \sum p(x) H(p) \\ &\leq 1 - H(p) \quad \text{当且仅当输入分布是均匀分布时取等号} \end{aligned}$$

这里的 $H(p) = -p \log p - (1-p) \log(1-p)$ .

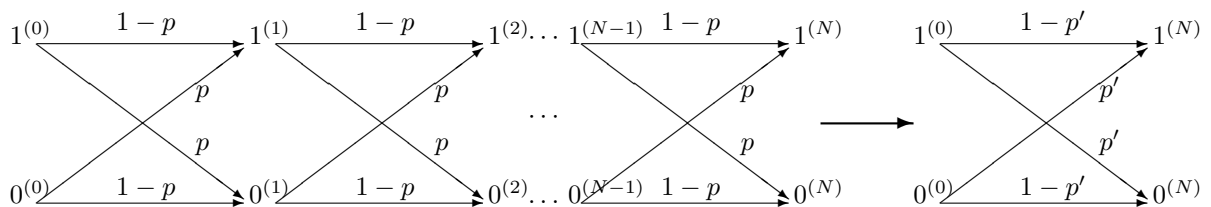
$$C = \max_{p(x)} I(X; Y) = 1 - H(p)$$

2. 将 $N$ 个 $BSC(p)$ 信道串联, 结果得到一个等效的BSC信道. 计算其信道容量 $C$ (用 $N$ 和 $p$ 表示)

**解:**  $1 - H(\frac{1}{2}(1 - (1 - 2p)^N))$

$$X_0 \rightarrow \boxed{\text{BSC}} \rightarrow X_1 \rightarrow \dots \rightarrow X_{N-1} \rightarrow \boxed{\text{BSC}} \rightarrow X_N$$

每个信道的原始差错概率为 $p$ , 那么整个串联信道的差错概率 $p'$ 即:



当 $x = p, y = 1 - p$ 时,  $(x + y)^N$ 的二项展开的奇次项之和:

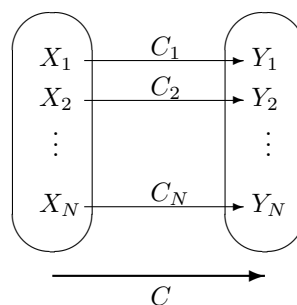
$$p' = \frac{1}{2}(x + y)^N - \frac{1}{2}(y - x)^N = \frac{1}{2}(1 - (1 - 2p)^N)$$

该串联信道可等效为差错概率为 $p'$ 的 $BSC(p')$ 信道:

$$C = 1 - H(p') = 1 - H(\frac{1}{2}(1 - (1 - 2p)^N))$$

3. 将 $N$ 个 $BSC(p)$ 信道串联且这 $N$ 个信道相互独立 (无串扰), 结果得到一个输入和输出为 $N$ 维的二进制向量的矢量信道, 并请计算其容量 $C$ (用 $N$ 和 $p$ 表示)

**解:**  $N(1 - H(p))$



信道间相互概率独立

$$p(y_1, \dots, y_N | x_1, \dots, x_N) = \prod_{i=1}^N p(y_i | x_1, \dots, x_N)$$

$$p(y_i | x_1, \dots, x_N) = \frac{p(x_1, \dots, x_N, y_i)}{p(x_1, \dots, x_N)} = \frac{p(x_i, y_i) p_{n \neq i}(x_1, \dots, x_n, \dots, x_N)}{p(x_i) p_{n \neq i}(x_1, \dots, x_n, \dots, x_N)} = p(y_i | x_i)$$

联合条件熵满足

$$H(Y_1, \dots, Y_N | X_1, \dots, X_N) = - \sum_{x, y} p(x_1, \dots, x_N, y_1, \dots, y_N) \log p(y_1, \dots, y_N | x_1, \dots, x_N)$$

$$= - \sum_{i=1}^N \sum_{x, y} p(x_i, y_i) \log p(y_i | x_i)$$

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

整个信道的互信息量

$$I(X_1, \dots, X_N | Y_1, \dots, Y_N) = H(Y_1, \dots, Y_N) - H(Y_1, \dots, Y_N | X_1, \dots, X_N)$$

$$= H(Y_1, \dots, Y_N) - \sum_{i=1}^N H(Y_i | X_1, \dots, X_N)$$

$$\leq \sum_{i=1}^N H(Y_i) - \sum_{i=1}^N H(Y_i | X_i) \quad \text{当且仅当所有的Y概率独立时取等号}$$

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

$$= \sum_{i=1}^N I(X_i, Y_i)$$

信道容量

$$C = \max_{p(x_1, \dots, x_N)} I(X_1, \dots, X_N | Y_1, \dots, Y_N)$$

$$\leq \max_{p(x_1, \dots, x_N)} \sum_{i=1}^N I(X_i, Y_i)$$

$$\leq \sum_{i=1}^N \max_{p(x_1, \dots, x_N)} I(X_i, Y_i)$$

$$= \sum_{i=1}^N C_i$$

$$= N(1 - H(p))$$

4. 将两个容量分别为 $C_1$ 和 $C_2$ 的BSC信道依概率组合.即码字 $X$ 的以概率 $\alpha$ 在信道1传输,或以概率 $1 - \alpha$ 在信道2传输,但两种传输不同时发生。请计算这种组合的最大容量 $C$  (用 $C_1$ 和 $C_2$ 表示)

解:  $\log(2^{C_1} + 2^{C_2})$

$$X = \begin{cases} X_1 & \text{概率为}\alpha \\ X_2 & \text{概率为}1-\alpha \end{cases}$$

令

$$\theta(X) = \begin{cases} 1 & \text{当} X = X_1 \\ 2 & \text{当} X = X_2 \end{cases}$$

$\theta$ 是 $Y$ 的函数, 即 $X \rightarrow Y \rightarrow \theta$ 是Markov链

$$I(X; Y, \theta) = I(X; \theta) + I(X; Y|\theta) = I(X; Y) + I(X; \theta|Y) = I(X; Y)$$

$$\begin{aligned} I(X; Y) &= I(X; \theta) + I(X; Y|\theta) \\ &= H(\alpha) + \alpha I(X_1; Y_1) + (1 - \alpha) I(X_2; Y_2) \end{aligned}$$

信道容量取决于关于 $\alpha$ 的函数

$$f(\alpha) = H(\alpha) + \alpha C_1 + (1 - \alpha) C_2$$

对函数求极值优化

$$\frac{df}{d\alpha} = -\log \alpha + \log(1 - \alpha) + C_1 - C_2 = 0$$

当 $\alpha = 2^{C_1} / (2^{C_1} + 2^{C_2})$ 时, 取到极大值

$$C = \max_{\alpha} f(\alpha) = \log(2^{C_1} + 2^{C_2})$$