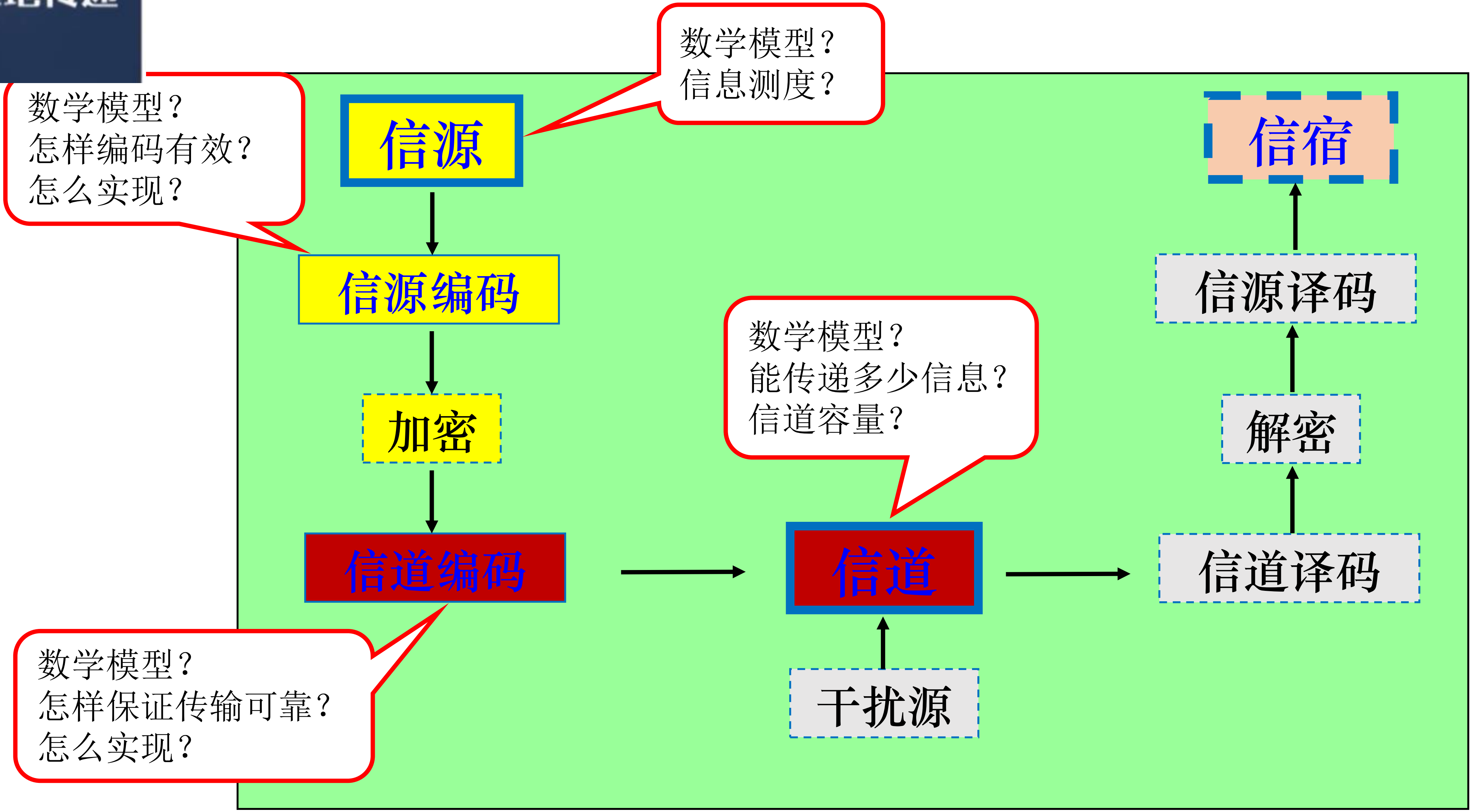
第 4 章　离 散 信 道 容 量

计算机科学与技术学院　孙伟平

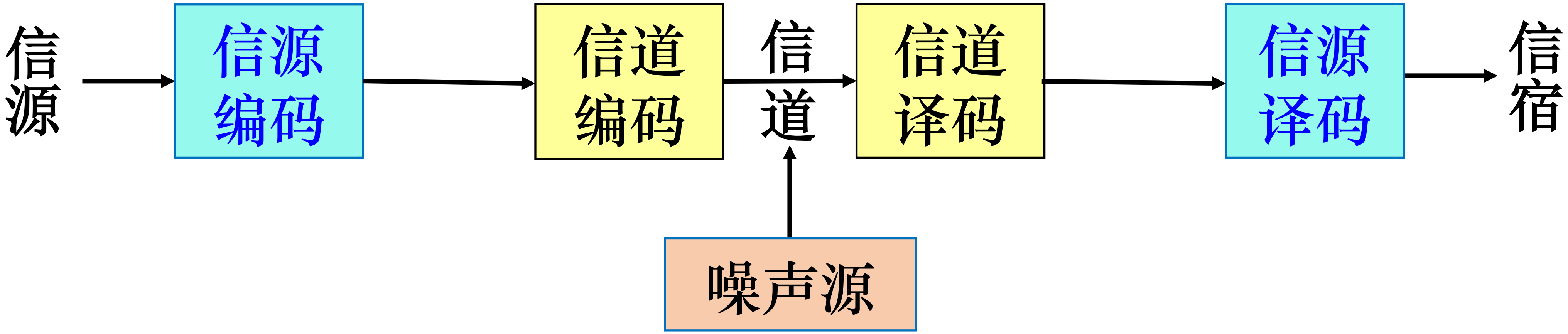
基本内容

在信息可以量度的基础上，研究有效地和可靠地传递信息的科学。



第4章

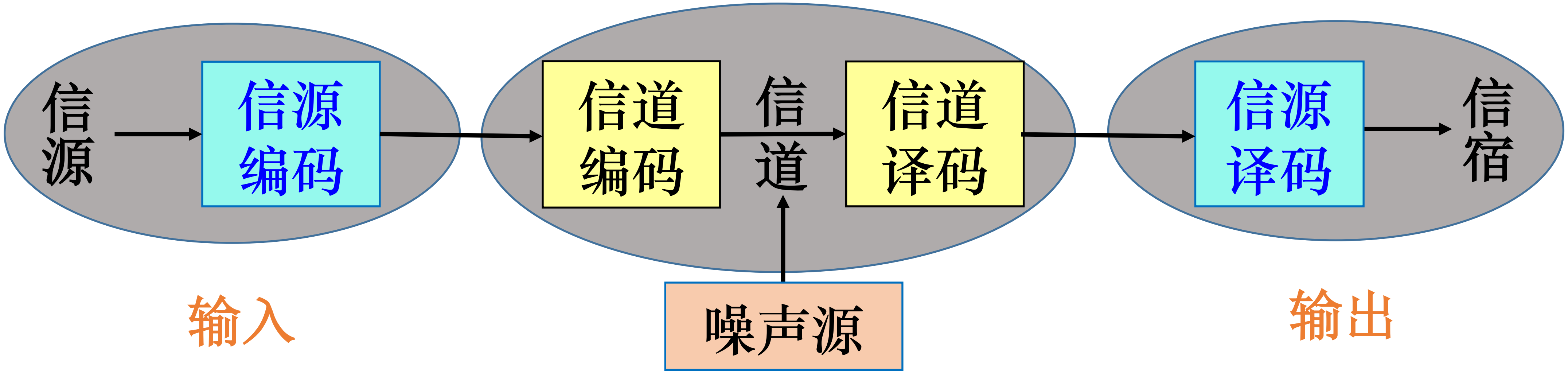
信道



信道是信息传输的通道，是通信系统最重要的组成部分。

研究信道的目的在于研究信道的（最大）传信能力。

为此需要了解信道的特性。



信道在传输信息的同时，会引入各种噪声和随机干扰，使得信号失真，导致接收错误。

由于存在噪声和干扰，信号的输入和输出符号之间是一种统计依赖关系，不再是确定的函数关系。

只要知道了信道的输入信号、输出信号的特性，以及它们之间的统计依赖关系，就可以确定信道的全部特性。

信道的分类

- 按照输入输出信号的幅度和时间特性来分

幅度	时间	信道分类名称
离散	离散	离散信道（数字信道）（讨论编码时有用）
连续	离散	连续信道
连续	连续	模拟信道（波形信道），是实际信道
离散	连续	理论和实用价值较小，不做讨论

- 按照输入输出间的记忆性来分

  无记忆信道 信道在某时刻的输出仅依赖于当前时刻的输入，与先前时刻的输入或输出无关。

  有记忆信道 信道在某时刻的输出与其它时刻的输入或输出有关。

- 按照输入输出间是否有确定关系来分

  无干扰信道 信道输入与输出之间具有确定的相互关系，即传输过程中不会出现差错。

  有干扰信道 与无干扰信道相对。

- 按照统计特性的平稳性来分

平稳信道 统计特性不随时间改变而改变，也称为恒参信道，时不变信道。

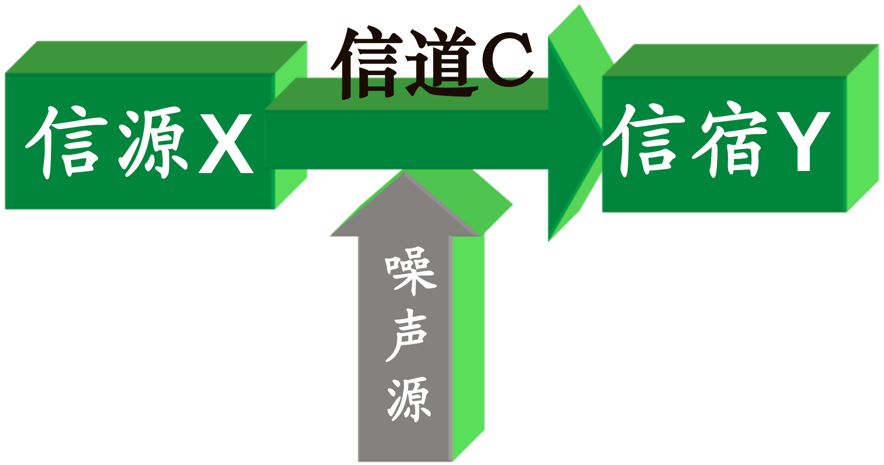
非平稳信道 统计特性随时间改变而改变，也称为随参信道，时变信道。

- 按照输入和输出的个数来分

单用户信道 单输入单输出的单向通信。

多用户信道 双向通信或三个或更多用户之间相互通信的情况。

4.1.1 单符号离散信道的数学模型



由于噪声和干扰的存在，信道的输入和输出之间存在统计依赖关系。研究信道就要研究输入输出信号特性和它们的统计依赖关系，在此基础上研究信息在信道中的流通。

离散信源X的数学模型为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} a_1, & a_2, & \cdots, & a_i, & \cdots, & a_n \\ p(a_1), & p(a_2), & \cdots, & p(a_i), & \cdots, & p(a_n) \end{Bmatrix}$$

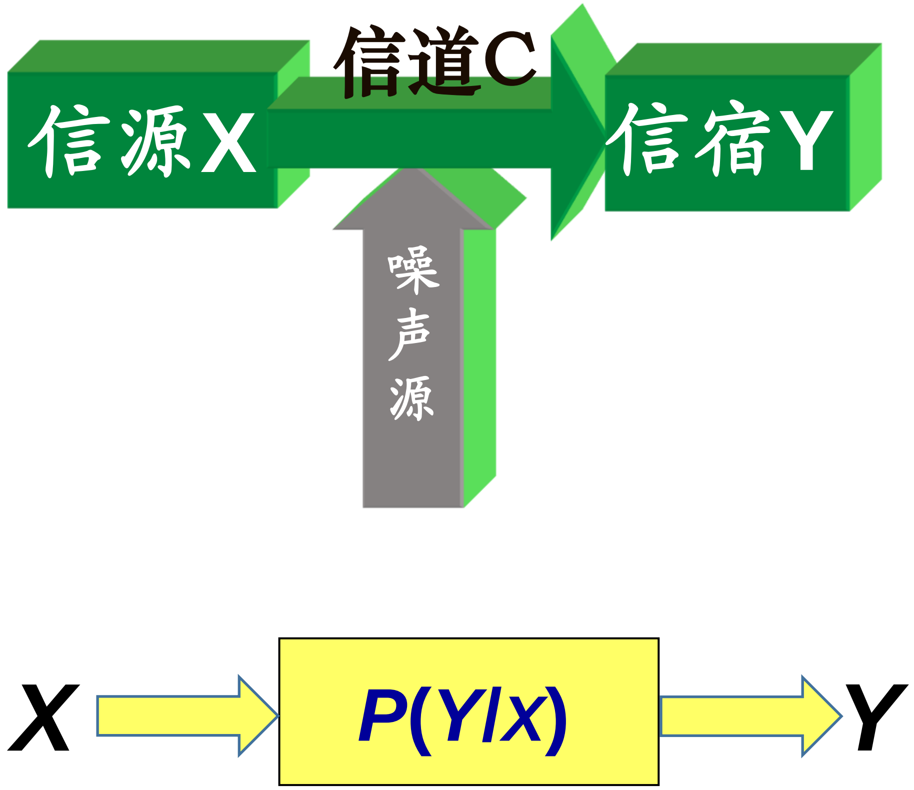
$$0 \le p(a_i) \le 1, \sum_{i=1}^{n} p(a_i) = 1$$

离散信宿Y的数学模型为

$$\begin{bmatrix} Y \\ P(Y) \end{bmatrix} = \begin{Bmatrix} b_1, & b_2, & \cdots, & b_i, & \cdots, & b_m \\ p(b_1), & p(b_2), & \cdots, & p(b_i), & \cdots, & p(b_m) \end{Bmatrix}$$

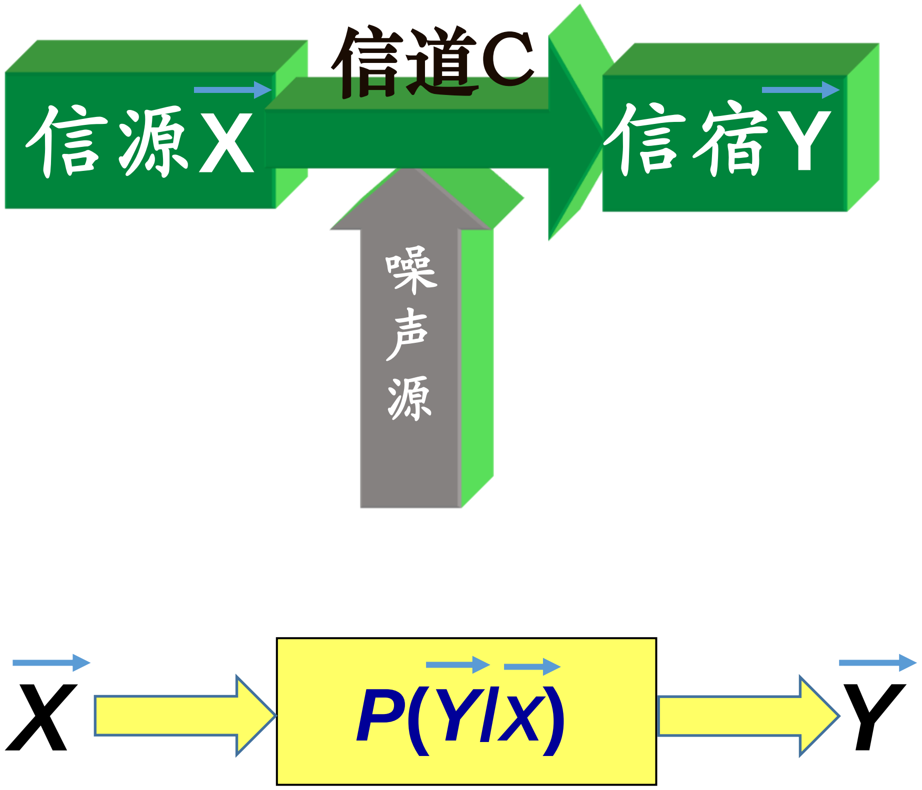
$$0 \le p(b_i) \le 1, \sum_{i=1}^{m} p(b_i) = 1$$

$X \in \{a_1, a_2, \cdots, a_n\}$ $Y \in \{b_1, b_2, \cdots, b_m\}$



单符号离散信道的数学模型： 三要素 $\{X \quad P(Y/X) \quad Y\}$

$X_i \in \{a_1, a_2, \cdots, a_n\}$ $Y_j \in \{b_1, b_2, \cdots, b_m\}$



多符号离散信道的数学模型： $\{\vec{X} \quad P(\vec{Y}/\vec{X}) \quad \vec{Y}\}$

信道转移/传递概率矩阵

$$p(b_j / a_i) = P(Y = b_j \mid X = a_i) \quad (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, m)$$

$$P_{n\times m} = \begin{bmatrix} p(b_1/a_1) & p(b_2/a_1) & \cdots & p(b_m/a_1) \\ p(b_1/a_2) & p(b_2/a_2) & \cdots & p(b_m/a_2) \\ \vdots & \vdots & \vdots & \vdots \\ p(b_1/a_n) & p(b_2/a_n) & \cdots & p(b_m/a_n) \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nm} \end{bmatrix}$$

根据条件概率的归一性，$\sum_{j=1}^{m} p(b_j / a_i) = 1, \quad i = 1, 2, \ldots, n$

例 二元对称信道（BSC)输入符号集为A={0,1}，输出符号集为B={0,1}

$$
\begin{array}{ccc}
0 & \xrightarrow{\bar{p}} & 0 \\
 & \overset{p}{\searrow} \; \overset{p}{\nearrow} & \\
1 & \xrightarrow[\bar{p}]{} & 1
\end{array}
$$

$$\bar{p} = 1 - p$$

信道转移概率矩阵为：

$$P = \begin{bmatrix} p(0|0) & p(1|0) \\ p(0|1) & p(1|1) \end{bmatrix} = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$

常用概率及其关系

先验概率 $p(a_i)$

传递/转移概率 $p(b_j / a_i)$

后验概率 $p(a_i / b_j)$

联合概率 $p(a_i b_j) = p(a_i) p(b_j / a_i) = p(b_j) p(a_i / b_j)$

贝叶斯公式 $$p(a_i / b_j) = \frac{p(a_i b_j)}{p(b_j)} = \frac{p(a_i) p(b_j / a_i)}{\sum_{i=1}^{n} p(a_i) p(b_j / a_i)}$$

输出符号概率--全概公式 $$p(b_j) = \sum_{i=1}^{n} p(a_i b_j) = \sum_{i=1}^{n} p(a_i) p(b_j / a_i)$$

$$\leftrightarrow P_Y = P_X P_{Y/X}$$

离散信源X的数学模型为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} a_1, & a_2, & , & a_i, & , & a_n \\ p(a_1), & p(a_2), & , & p(a_i), & , & p(a_n) \end{Bmatrix}$$

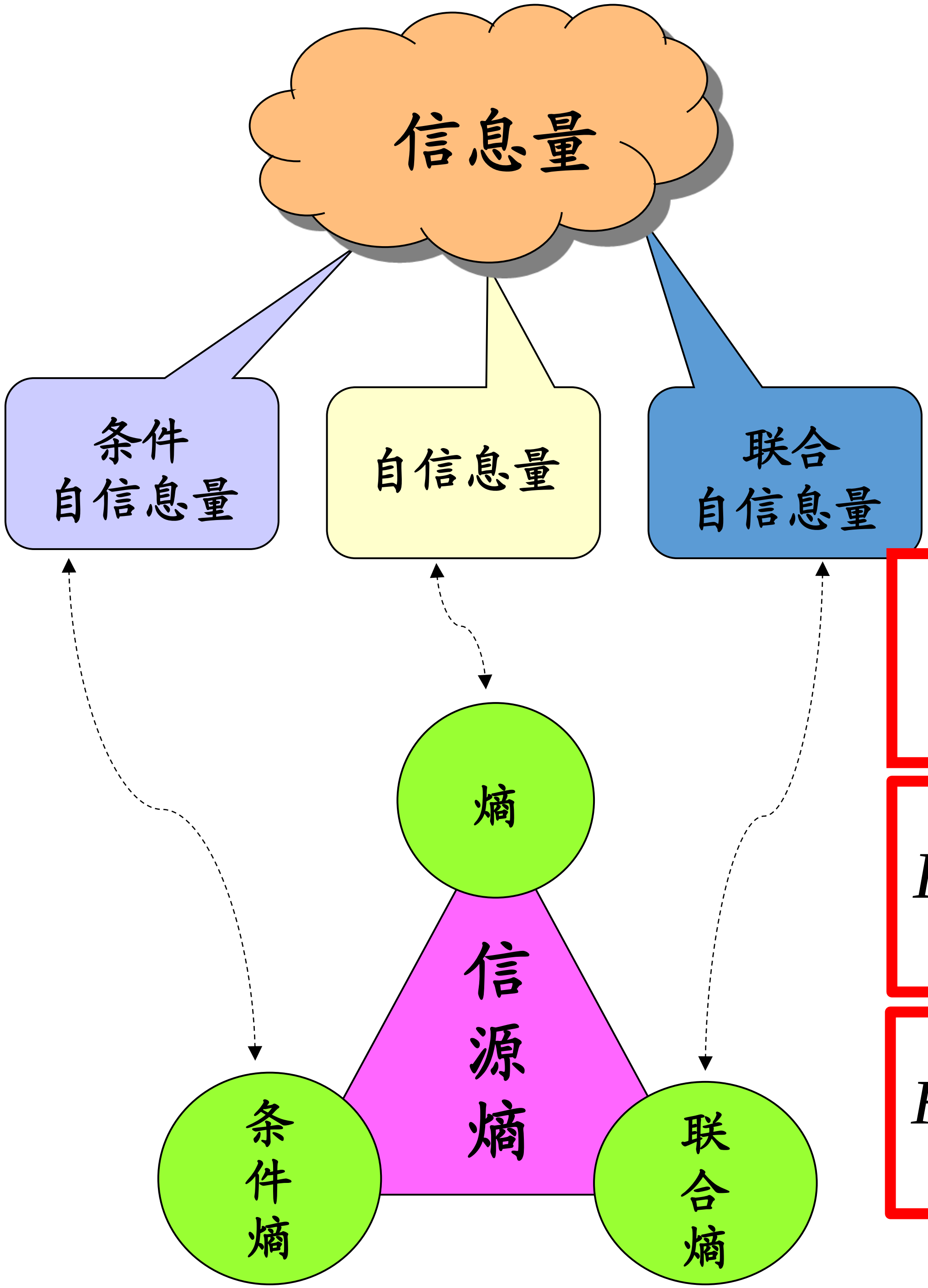
$$0 \le p(a_i) \le 1, \sum_{i=1}^{n} p(a_i) = 1$$

数学模型 → 信息测度

自信息（三种）

信源熵（三种）

定义，物理意义，
计算公式，性质



$$I(a_i) = -\log p(a_i)$$

$$I(a_i b_j) = -\log p(a_i b_j)$$

$$I(a_i / b_j) = -\log p(a_i / b_j)$$

$$H(X) = E[I(a_i)] = -\sum_{i=1}^{n} p(a_i)\log_2 p(a_i)$$

$$H(XY) = E[I(a_i b_j)] = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i b_j)\log p(a_i b_j)$$

$$H(X/Y) = E[I(a_i / b_j)] = -\sum_{j=1}^{m}\sum_{i=1}^{n} p(a_i b_j)\log p(a_i / b_j)$$

如果X表示信源的随机变量，Y表示信宿的随机变量，则：

$H(X/Y)$ 表示信宿在收到Y后，信源X仍然存在的不确定性。

这是信道传输失真所导致的。

称$H(X/Y)$为**信道疑义度**，也称为**损失熵**。

称$H(Y/X)$为**噪声熵**。

$H(X/Y) \le H(X)$ 即收到输出后对于输入的不确定度一般会减少。

$$I(a_i b_j) = I(a_i) + I(b_j/a_i)$$
$$= I(b_j) + I(a_i/b_j)$$

$$H(XY) = H(X) + H(Y/X)$$
$$H(XY) = H(Y) + H(X/Y)$$

$$H(XY) \le H(X) + H(Y)$$

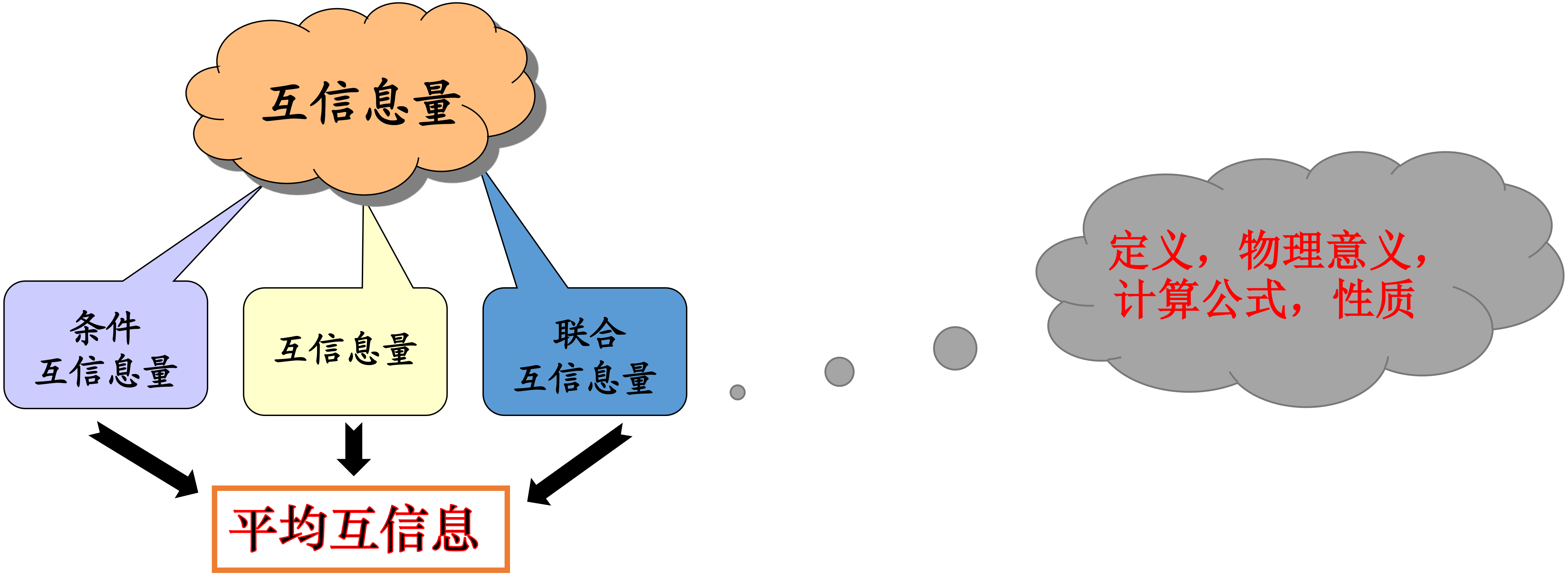
当X和Y相互独立时：$H(X/Y) = H(X)$

$$H(Y/X) = H(Y)$$
$$H(XY) = H(X) + H(Y)$$

单符号离散信道的数学模型：$\{X \quad P(Y/X) \quad Y\}$

数学模型 → 信息测度

- 互信息
- 平均互信息-交叉熵



4.1.2 互信息量及其性质

1 互信息量的定义

定义b_j对a_i的互信息量为

$$I(a_i;b_j)=I(a_i)-I(a_i/b_j)=\log\frac{p(a_i/b_j)}{p(a_i)},\quad (i=1,2,\cdots,n;\ j=1,2,\cdots,m)$$

物理含义：

是事件b_j出现前后关于输入a_i存在的不确定度的减少量。

是事件b_j出现以后信宿获得的关于输入a_i的信息量。

1 互信息量的定义

定义b_j对a_i的互信息量为

$$I(a_i;b_j)=I(a_i)-I(a_i/b_j)=\log\frac{p(a_i/b_j)}{p(a_i)},\quad (i=1,2,\cdots,n;\ j=1,2,\cdots,m)$$

同理，定义a_i对b_j的互信息量为

$$I(b_j;a_i)=I(b_j)-I(b_j/a_i)=\log\frac{p(b_j/a_i)}{p(b_j)},(i=1,2,\cdots,n;\ j=1,2,\cdots,m)$$

上述两个互信息量有何关系?

互信息量的**计算公式**：

$$I(a_i;b_j)=I(a_i)-I(a_i/b_j)=\log\frac{p(a_i/b_j)}{p(a_i)}$$

从信宿看

$$I(a_i;b_j)=I(b_j)-I(b_j/a_i)=\log\frac{p(b_j/a_i)}{p(b_j)}$$

从信源看

$$I(a_i;b_j)=I(a_i)+I(b_j)-I(a_ib_j)=\log\frac{p(a_ib_j)}{p(a_i)p(b_j)}$$

从通信系统整体看

$$I(a_ib_j)=I(a_i)+I(b_j/a_i)=I(b_j)+I(a_i/b_j)$$

互信息量的引出，使信息流通问题进入了定量分析的范畴，为信息流通的定量测量打下了基础。

例[4.1.1] 继续讨论例2.2.1题，即某地二月份天气构成的信源为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} a_1(\text{晴}), & a_2(\text{阴}), & a_3(\text{雨}), & a_4(\text{雪}) \\ \frac{1}{2}, & \frac{1}{4}, & \frac{1}{8}, & \frac{1}{8} \end{Bmatrix}$$

$$I(a_1) = 1\,bit \quad I(a_2) = 2bit$$

$$I(a_3) = 3bit \quad I(a_4) = 3bit$$

一天有人告诉你：“今天不是晴天。”把这句话作为收到的消息b_1，收到b_1后，各种天气出现的概率变成后验概率了。其中

$$p(a_1/b_1) = 0, \quad p(a_2/b_1) = \frac{1}{2}, \quad p(a_3/b_1) = \frac{1}{4}, \quad p(a_4/b_1) = \frac{1}{4}$$

例[4.1.1] 继续讨论例2.2.1题，即某地二月份天气构成的信源为

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} a_1(\text{晴}), & a_2(\text{阴}), & a_3(\text{雨}), & a_4(\text{雪}) \\ \frac{1}{2}, & \frac{1}{4}, & \frac{1}{8}, & \frac{1}{8} \end{Bmatrix}$$

$$p(a_1/b_1)=0,\quad p(a_2/b_1)=\frac{1}{\ },\quad p(a_3/b_1)=\frac{1}{\ },\quad p(a_4/b_1)=\frac{1}{\ }$$

这表明从 b_1 分别得到了 a_2、a_3、a_4 各1比特的信息量，也可以理解为消息 b_1 使 a_2、a_3、a_4 的不确定度各减少了1比特。

$$I(a_1;b_1)=\log_2\frac{p(a_1/b_1)}{p(a_1)}=\log_2 0=-\infty$$

$$I(a_1)=1\,bit \quad I(a_2)=2bit$$

$$I(a_3)=3bit \quad I(a_4)=3bit$$

$$I(a_2;b_1)=\log_2\frac{p(a_2/b_1)}{p(a_2)}=\log_2\frac{1/2}{1/4}=1(bit)$$

$$I(a_3;b_1)=I(a_4;b_1)=1(bit)$$

2 互信息量的性质 $I(a_i;b_j)=I(a_i)-I(a_i/b_j)$

1. 对称性 $I(a_i;b_j)=I(b_j;a_i)$
2. 当X和Y相互独立时，互信息为0。
3. 互信息量可正可负可为0。
4. 互信息量不可能大于其中任意一个事件的自信息量。

$$I(a_i;b_j)\le I(a_i) \qquad I(a_i;b_j)\le I(b_j)$$

某事件的自信息是发生其他事件能提供的关于该事件的最大信息量。

3 联合互信息量

三个变量的联合互信息量

$$I(a_i;b_jc_k)=I(a_i)-I(a_i/b_jc_k)=\log\frac{p(a_i/b_jc_k)}{p(a_i)}$$

4 条件互信息量

$$I(a_i;b_j/c_k)= I(a_i/c_k)- I(a_i/b_jc_k)=\log\frac{p(a_i/b_jc_k)}{p(a_i/c_k)}$$

联合互信息与互信息、条件互信息的关系

$$\boxed{I(a_i;b_jc_k)=I(a_i;c_k)+I(a_i;b_j/c_k)}$$

$$\boxed{I(a_i;c_kb_j)=I(a_i;b_j)+I(a_i;c_k/b_j)}$$

$$\begin{aligned}I(a_ib_j)&=I(a_i)+I(b_j/a_i)\\&=I(b_j)+I(a_i/b_j)\end{aligned}$$

4.1.3 平均互信息量及其性质

1 平均互信息量的定义

Y对X的**平均互信息量**，也称平均交互信息量、交互熵。

$$\boxed{I(X;Y)} = E[I(a_i;b_j)] = \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i b_j) I(a_i;b_j)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i b_j) \log \frac{p(a_i/b_j)}{p(a_i)}$$

$$I(X;Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i b_j)\left[\log p(a_i/b_j) - \log p(a_i)\right] = H(X) - H(X/Y)$$

X的平均不确定性

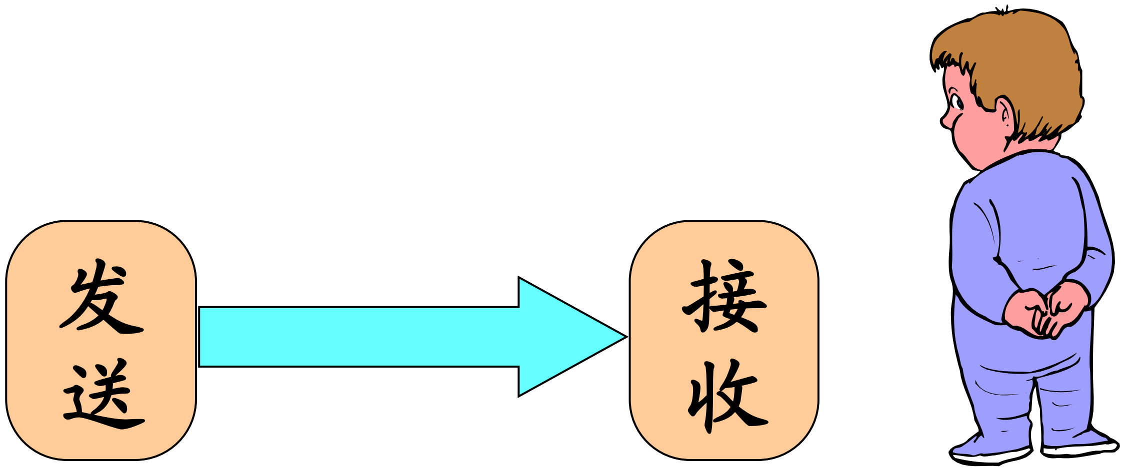
已知Y后，关于X的平均不确定性

$$I(X;Y)=H(X)-H(X/Y)$$

收到Y前后，关于X的平均不确定性减少的量

=

由Y获得的关于X的平均信息量



同理，X对Y的平均互信息量

$$I(Y;X)=E[I(b_j;a_i)]=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\log\frac{p(b_j/a_i)}{p(b_j)}$$

$$I(Y;X)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\left[\log p(b_j/a_i)-\log p(b_j)\right]=H(Y)-H(Y/X)$$

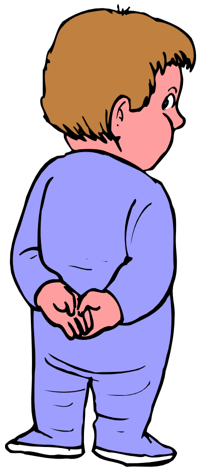
$$I(Y;X)=I(X;Y)$$

Y的平均不确定性

已知X后，关于Y的平均不确定性

$$I(X;Y)=H(Y)-H(Y/X)$$

发出X前后，关于Y的平均不确定性减少的量 = 由X获得的关于Y的平均信息量



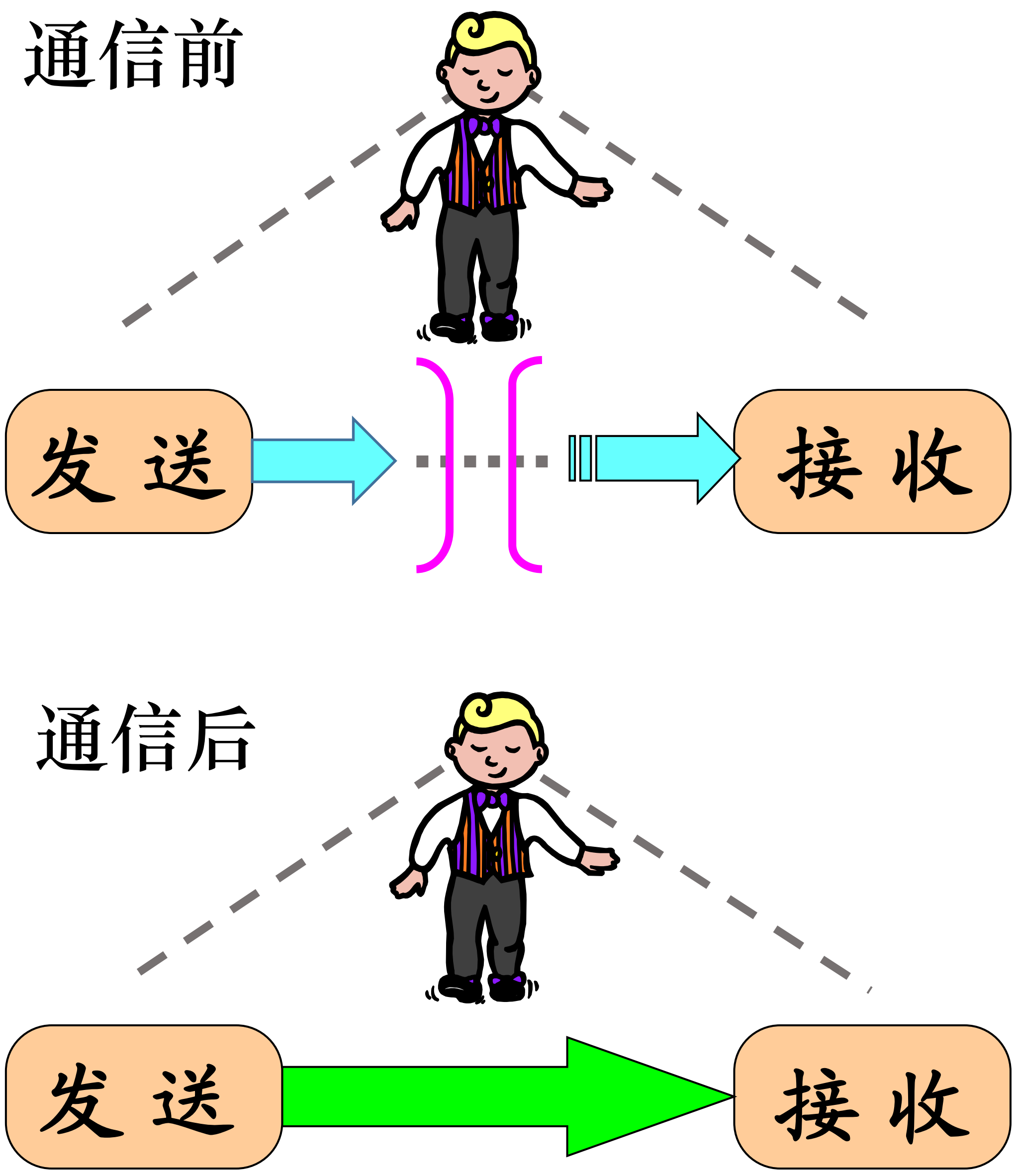
发送 → 接收

$$p(a_i b_j) = p(b_j)p(a_i/b_j)$$
$$= p(a_i)p(b_j/a_i)$$

$$I(X;Y)$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i b_j)\log\frac{p(a_i b_j)}{p(a_i)p(b_j)}$$
$$= H(X) + H(Y) - H(XY)$$

通信前　　通信后

通信前后，整个系统不确定度减少的量，即信道中流通信息量的整体测度。



平均互信息量的计算公式：

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\left[\log p(a_i/b_j)-\log p(a_i)\right]=H(X)-H(X/Y)$$

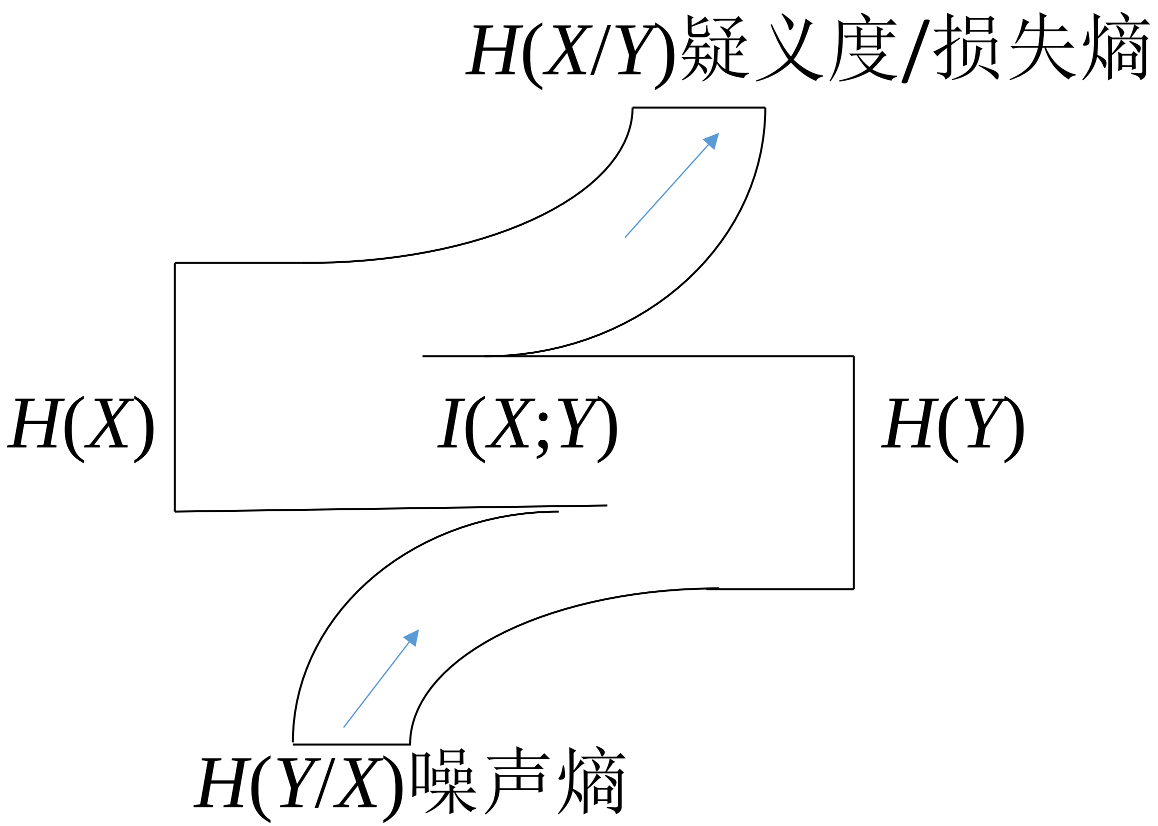
疑义度/损失熵

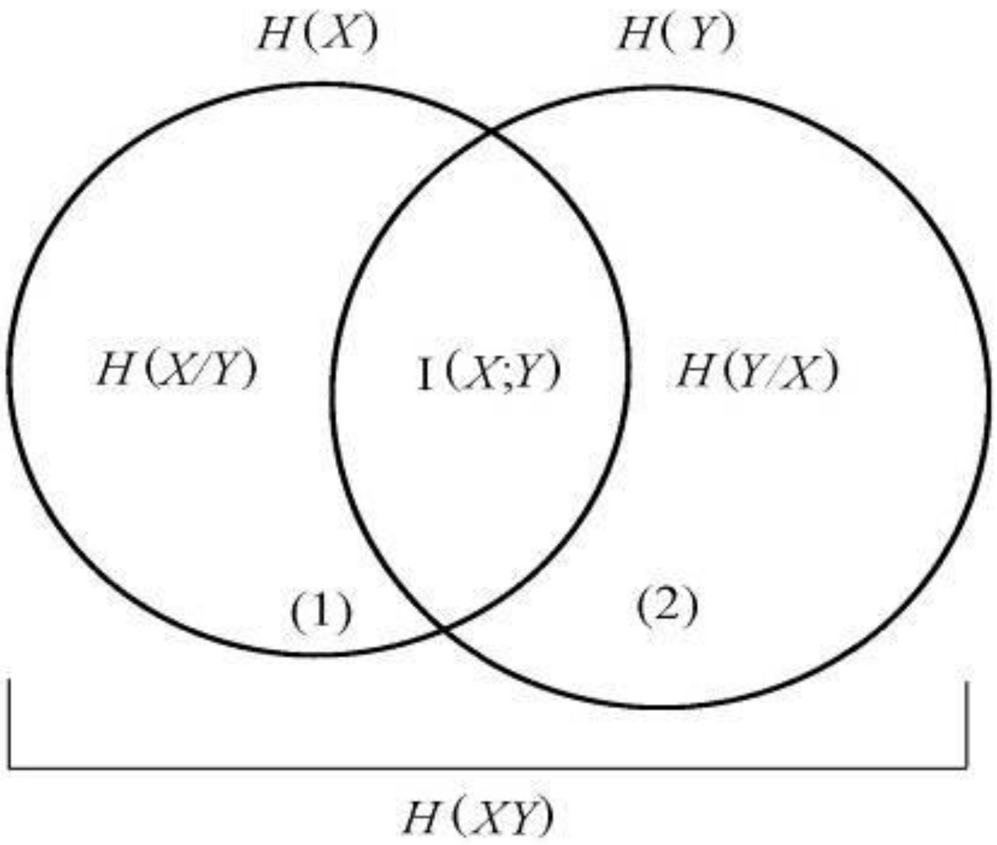
$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\left[\log p(b_j/a_i)-\log p(b_j)\right]=H(Y)-H(Y/X)$$

噪声熵

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\left[\log p(a_ib_j)-\log p(a_i)-\log p(b_j)\right]$$

$$=H(X)+H(Y)-H(XY)$$





$$I(X;Y)=H(X)-H(X/Y)$$

$$I(X;Y)=H(Y)-H(Y/X)$$

$$I(X;Y)=H(X)+H(Y)-H(XY)$$

2 平均互信息量的物理意义

1. $$I(X;Y)=H(X)-H(X/Y)$$

平均互信息量是收到Y前后，关于X的平均不确定度减少的量，即由Y获得的关于X的平均信息量。

2. $$I(X;Y)=H(Y)-H(Y/X)$$

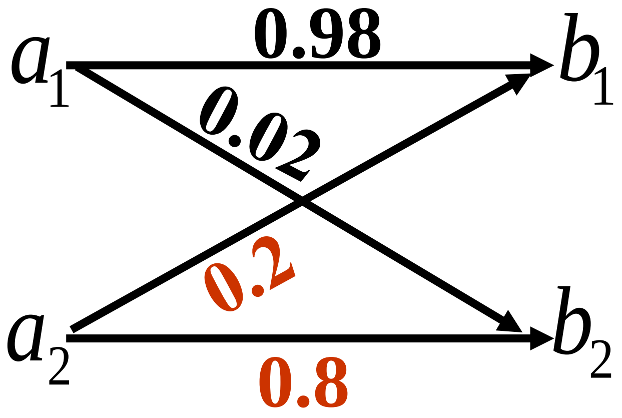
平均互信息量是发送X前后，关于Y的平均不确定度减少的量，即由X获得的关于Y的平均信息量。

3. $$I(X;Y)=H(X)+H(Y)-H(XY)$$

平均互信息量等于通信前后，整个系统不确定度减少的量，即信道中流通信息量的整体测度。

例4.1.2 信源X接入图示信道

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} a_1 & a_2 \\ 0.5 & 0.5 \end{Bmatrix}$$

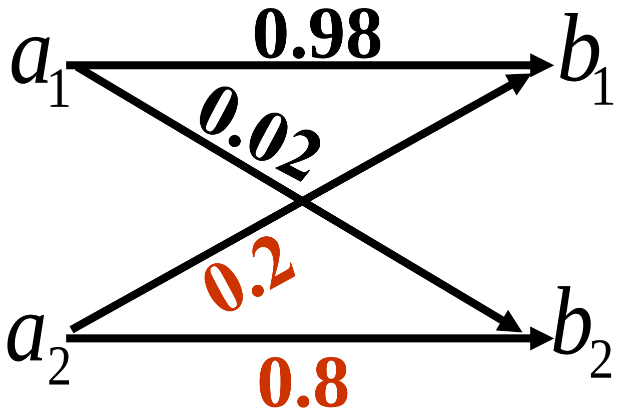


$$p(b_1/a_1) = 0.98, \quad p(b_2/a_1) = 0.02, \quad p(b_1/a_2) = 0.2, \quad p(b_2/a_2) = 0.8$$

求：$H(X),\ H(Y),\ H(Y/X),\ H(X/Y),\ H(XY),\ I(X;Y)$

例4.1.2 信源X接入图示信道

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} a_1 & a_2 \\ 0.5 & 0.5 \end{Bmatrix}$$



$$p(b_1/a_1) = 0.98, \quad p(b_2/a_1) = 0.02, \quad p(b_1/a_2) = 0.2, \quad p(b_2/a_2) = 0.8$$

$$p(a_i b_j) = p(a_i) p(b_j/a_i)$$

$$p(a_1 b_1) = p(a_1) p(b_1/a_1) = 0.5 \times 0.98 = 0.49$$

$$p(a_1 b_2) = 0.5 \times 0.02 = 0.01, \ p(a_2 b_1) = 0.5 \times 0.2 = 0.1, \ p(a_2 b_2) = 0.5 \times 0.8 = 0.4$$

$$p(b_j) = \sum_{i=1}^{2} p(a_i b_j)$$

$$p(b_1) = p(a_1 b_1) + p(a_2 b_1) = 0.1 + 0.49 = 0.59$$

$$p(b_2) = p(a_1 b_2) + p(a_2 b_2) = 0.01 + 0.4 = 0.41$$

$$p(a_i / b_j) = \frac{p(a_i b_j)}{p(b_j)}$$

$$p(a_1 / b_1) = \frac{p(a_1 b_1)}{p(b_1)} = \frac{0.49}{0.59} = 0.831, \quad p(a_2 / b_1) = 1 - p(a_1 / b_1) = 0.169$$

$$p(a_1 / b_2) = \frac{p(a_1 b_2)}{p(b_2)} = \frac{0.01}{0.41} = 0.024, \quad p(a_2 / b_2) = 1 - p(a_1 / b_2) = 0.976$$

$$H(X) = -0.5\log 0.5 - 0.5\log 0.5 = 1(bit/sign)$$

$$H(Y) = -0.59\log 0.59 - 0.41\log 0.41 = 0.98(bit/sign)$$

$$H(XY) = -0.49\log 0.49 - 0.01\log 0.01 - 0.1\log 0.1 - 0.4\log 0.4 = 1.43(bit/sign)$$

$$\begin{aligned} H(X/Y) &= -\sum_{i=1}^{2}\sum_{j=1}^{2} p(a_i b_j)\log p(a_i / b_j) \\ &= -0.49\log 0.831 - 0.01\log 0.024 - 0.1\log 0.169 - 0.4\log 0.976 \\ &= H(XY) - H(Y) = 0.45(bit/sign) \end{aligned}$$

$$H(Y/X) = H(XY) - H(X) = 1.43 - 1 = 0.43(bit/sign)$$

$$I(X;Y) = H(X) - H(X/Y) = 1 - 0.45 = 0.55(bit/sign)$$

3 平均互信息量的性质

① 对称性 $I(X;Y)=I(Y;X)$

$$\because I(a_i;b_j)=I(b_j;a_i)$$

$$\therefore I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)I(a_i;b_j)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)I(b_j;a_i)=I(Y;X)$$

2 非负性 $I(X;Y)\geq 0$

$$-I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\log\left[\frac{p(a_i)p(b_j)}{p(a_ib_j)}\right]$$

$$\leq\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\left[\frac{p(a_i)p(b_j)}{p(a_ib_j)}-1\right]\log e$$

$$\leq\left[\sum_{i=1}^{n}p(a_i)\sum_{j=1}^{m}p(b_j)-\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\right]\log e=0$$

非负性表明：从整体和平均的意义上来说，信道每传输一条消息，总能提供一定的信息量；或者说接收端每收到一条消息，总能提取到关于信源的信息量。也可以说从一个事件提取关于另一个事件的信息，最坏的情况是0，不会由于知道了一个事件，反而对另一个事件的不确定性增加。

3 极值性 $I(X;Y) \le H(X)$; $I(Y;X) \le H(Y)$

1 X、Y 一一对应时：

$$p(a_i/b_j)=\begin{cases}1 & i=j\\ 0 & i\ne j\end{cases},\quad H(X/Y)=0 \text{ 故 } I(X;Y)=H(X)$$

2 X、Y 相互独立，即 $H(X|Y)=H(X)$

$$I(X;Y)=H(X)-H(X/Y)=0$$

4 凸函数性

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\log\frac{p(b_j/a_i)}{p(b_j)},$$

$$p(a_ib_j)=p(a_i)p(b_j/a_i)\qquad p(b_j)=\sum_{i=1}^{n}p(a_i)p(b_j/a_i)$$

$$I(X;Y)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_i)p(b_j/a_i)\log\frac{p(b_j/a_i)}{\sum_{i=1}^{n}p(a_i)p(b_j/a_i)}$$

$$\Downarrow$$

$$I(X;Y)=f\left[p(a_i),p(b_j/a_i)\right]$$

信源　信道

$$I(X;Y)=f\left[p(a_i),p(b_j/a_i)\right]=\begin{cases} f\left[p(a_i)\right] & \text{固定信道} \\ f\left[p(b_j/a_i)\right] & \text{固定信源} \end{cases}$$

1 $I(X;Y)$是信源$\{p(a_i)\}$的 上凸函数

令 $0<\alpha<1$

$$I\left[\alpha p_1(a_i)+(1-\alpha)p_2(a_i)\right]\geq \alpha I\left[p_1(a_i)\right]+(1-\alpha)I\left[p_2(a_i)\right]$$

是研究信道容量的理论基础。

2 $I(X;Y)$是$\{p(b_j/a_i)\}$的 下凸函数

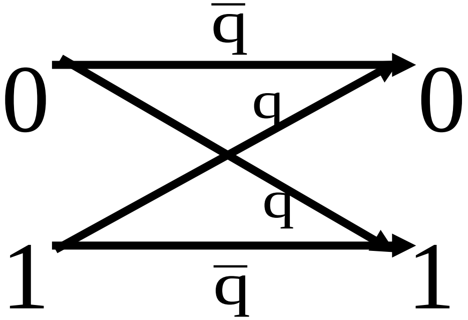
令 $0<\alpha<1$

$$I\left[\alpha p_1(b_j/a_i)+(1-\alpha)p_2(b_j/a_i)\right]\le \alpha I\left[p_1(b_j/a_i)\right]+(1-\alpha)\left[p_2(b_j/a_i)\right]$$

是研究信源的信息率失真函数的理论基础。

例4.1.3 二元信源X接入对称信道，求$I(X;Y)$。

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} 0 & 1 \\ p & \bar{p} \end{Bmatrix}$$



$$\Rightarrow P_{Y|X} = \begin{Bmatrix} \bar{q} & q \\ q & \bar{q} \end{Bmatrix}$$

解：$I(X;Y) = H(X) - H(X/Y) = \underline{H(Y) - H(Y/X)} = H(X) + H(Y) - H(XY)$

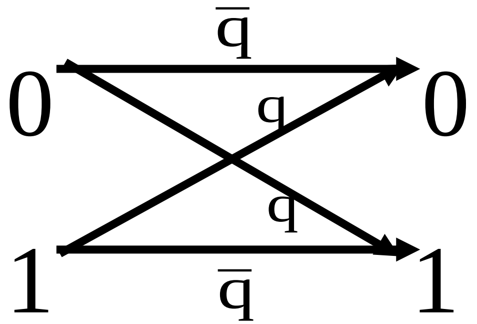
$$H(Y/X) = -\sum_{i=1}^{2}\sum_{j=1}^{2} p(a_i)p(b_j/a_i)\log p(b_j/a_i)$$

$$= -\left[p\left(\bar{q}\log\bar{q} + q\log q\right) + \bar{p}\left(q\log q + \bar{q}\log\bar{q}\right)\right]$$

$$= -q\log q - \bar{q}\log\bar{q} = H(q)$$

例4.1.3 二元信源X接入对称信道，求$I(X;Y)$。

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} 0 & 1 \\ p & \bar{p} \end{Bmatrix}$$



$$\Rightarrow P_{Y|X} = \begin{Bmatrix} \bar{q} & q \\ q & \bar{q} \end{Bmatrix}$$

$$I(X;Y) = H(X) - H(X/Y) = \underline{H(Y) - H(Y/X)} = H(X) + H(Y) - H(XY)$$

$$p(b_j) = \sum_{i=1}^{2} p(a_i) p(b_j/a_i) \quad p(b_1) = p\bar{q} + \bar{p}q \quad p(b_2) = pq + \bar{p}\bar{q}$$

$$H(Y) = -\sum_{j=1}^{2} p(b_j) \log p(b_j)$$

$$= -(p\bar{q} + \bar{p}q)\log(p\bar{q} + \bar{p}q) - (pq + \bar{p}\bar{q})\log(pq + \bar{p}\bar{q}) = H(p\bar{q} + \bar{p}q)$$

$$H(Y/X) = -q\log q - \bar{q}\log\bar{q} = H(q)$$

$$H(Y) = H(p\bar{q} + \bar{p}q)$$

$$\Rightarrow I(X;Y) = H(Y) - H(Y/X) = H(p\bar{q} + \bar{p}q) - H(q)$$

$$\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} 0 & 1 \\ p & \bar{p} \end{Bmatrix} \qquad P_{Y|X} = \begin{Bmatrix} \bar{q} & q \\ q & \bar{q} \end{Bmatrix}$$

接下来验证平均互信息的凸性。

（1）固定信道（即q）：

p=0时，I(X;Y)=?

p=1时，I(X;Y)=?

p=？时，I(X;Y)取最大值，最大值是多少？

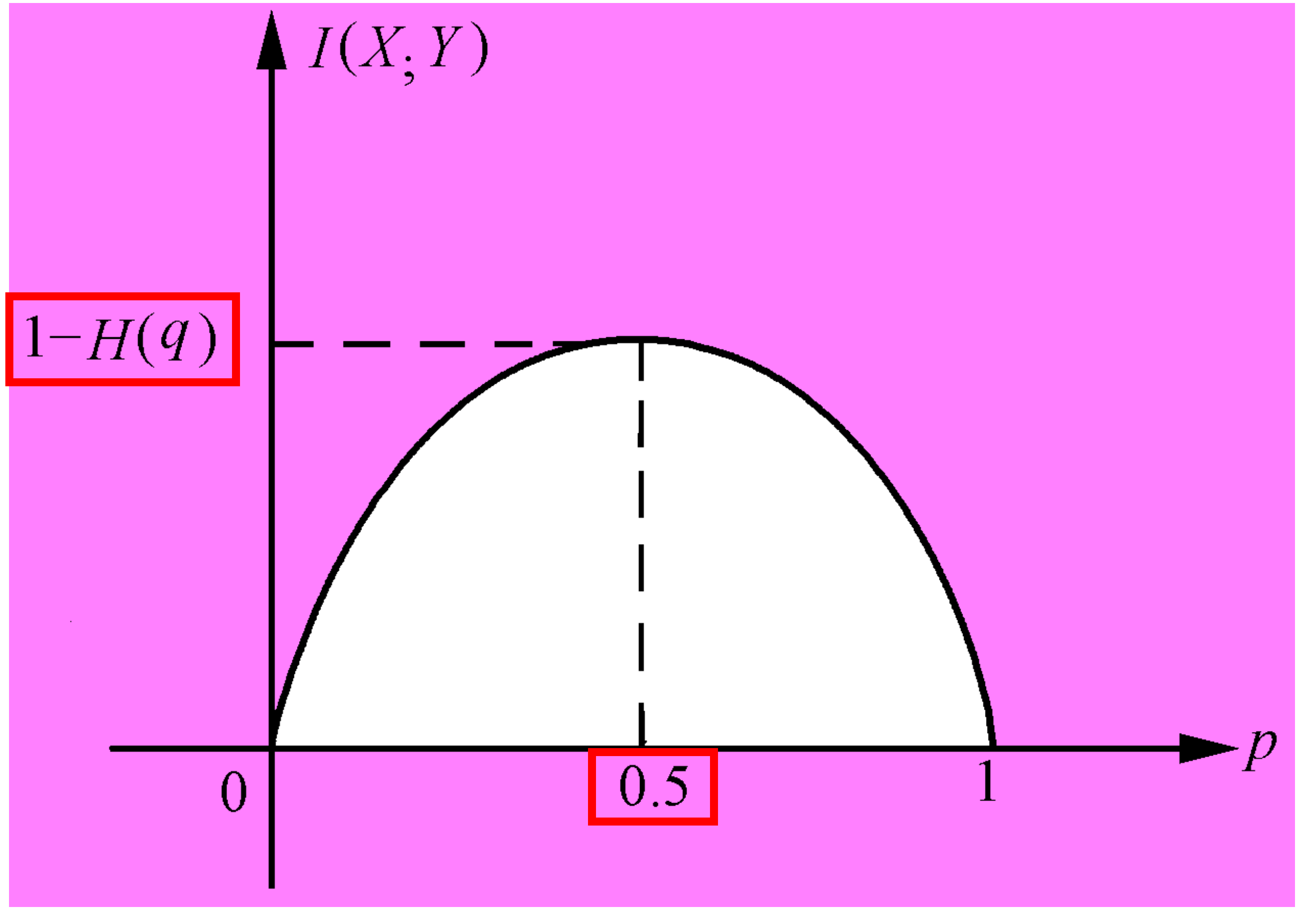
（2）固定信源（即p）：

q=0时，I(X;Y)=?

q=1时，I(X;Y)=?

q=？时，I(X;Y)取最小值，最小值是多少？

$$I(X;Y)=H(p\bar{q}+\bar{p}q)-H(q)$$



$I(X;Y)$随信源变化的曲线

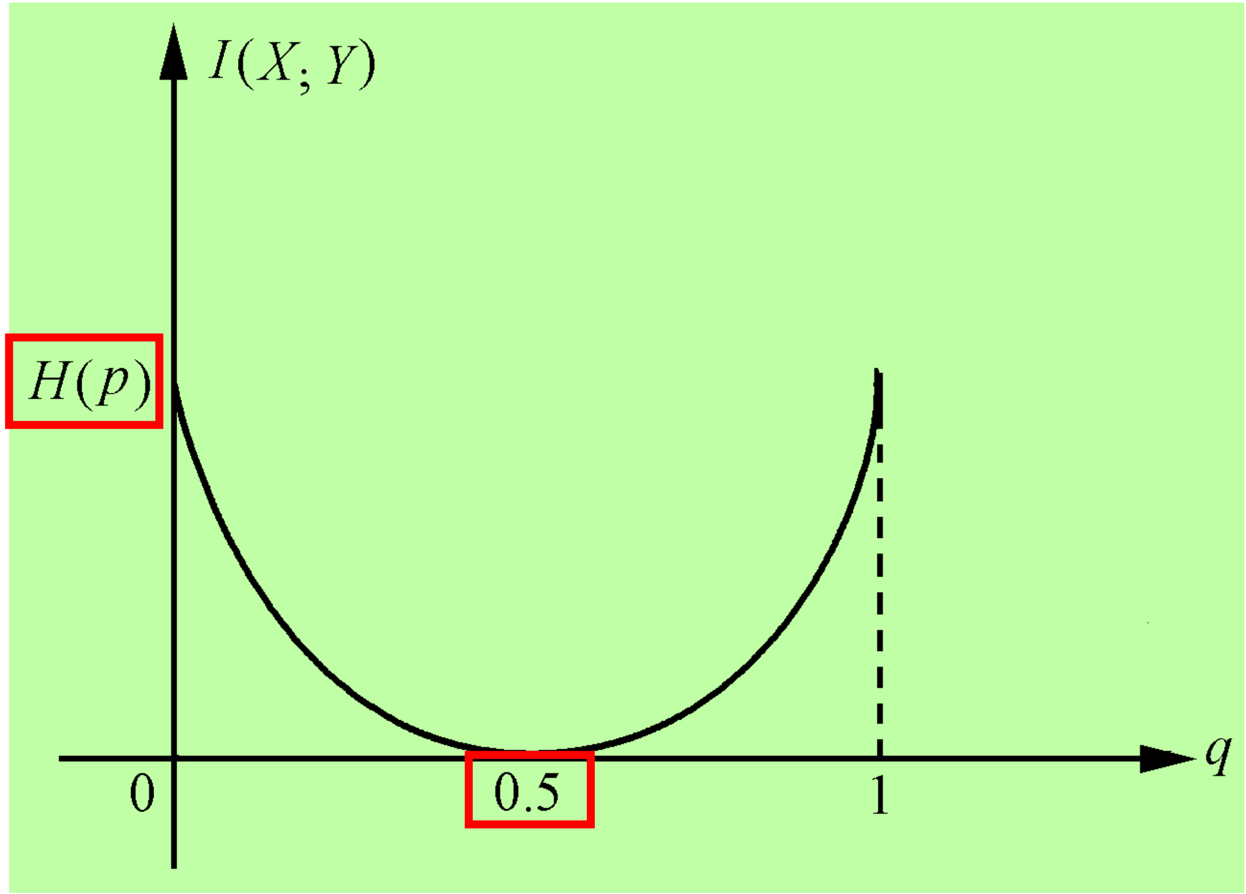
$$\max I(X;Y)=1-H(q)$$

信道容量

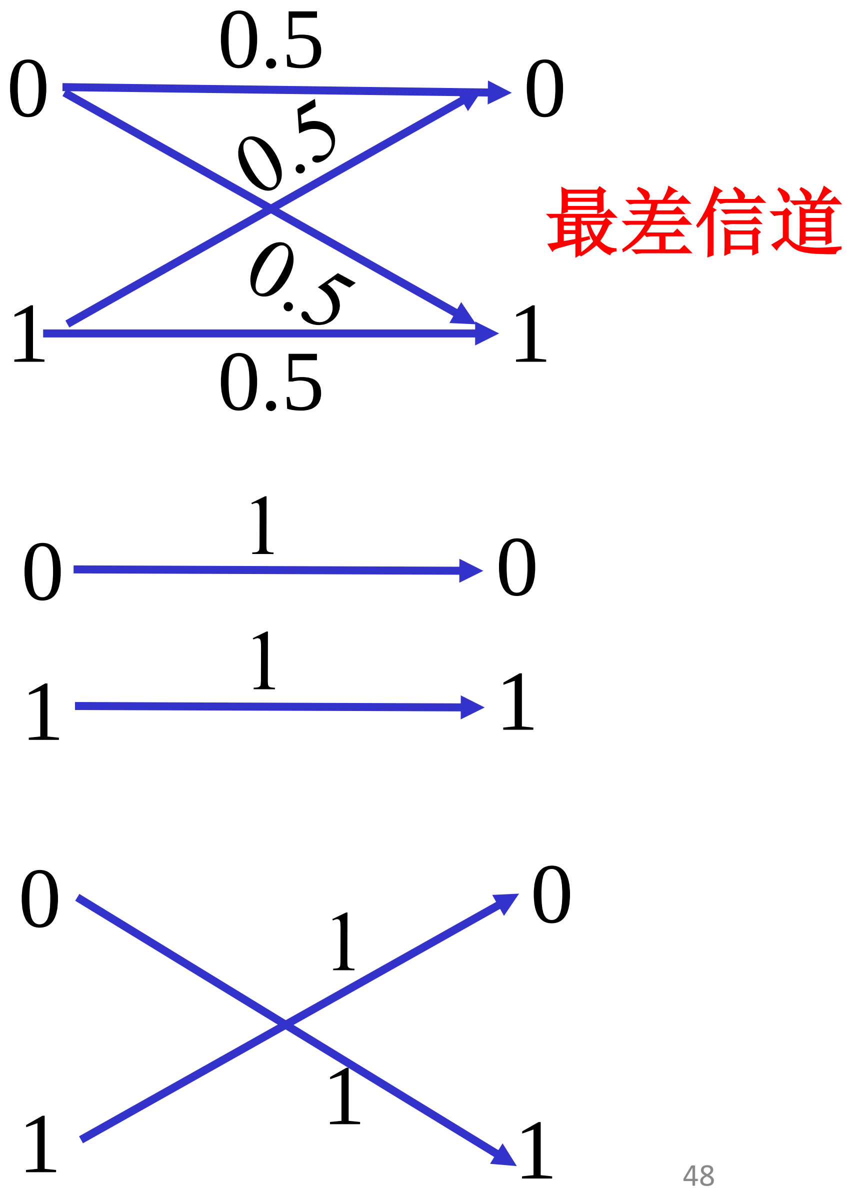
$$\begin{bmatrix} X \\ P(X) \end{bmatrix}=\begin{Bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{Bmatrix}$$

最佳分布

$$I(X;Y) = H(p\bar{q} + \bar{p}q) - H(q)$$

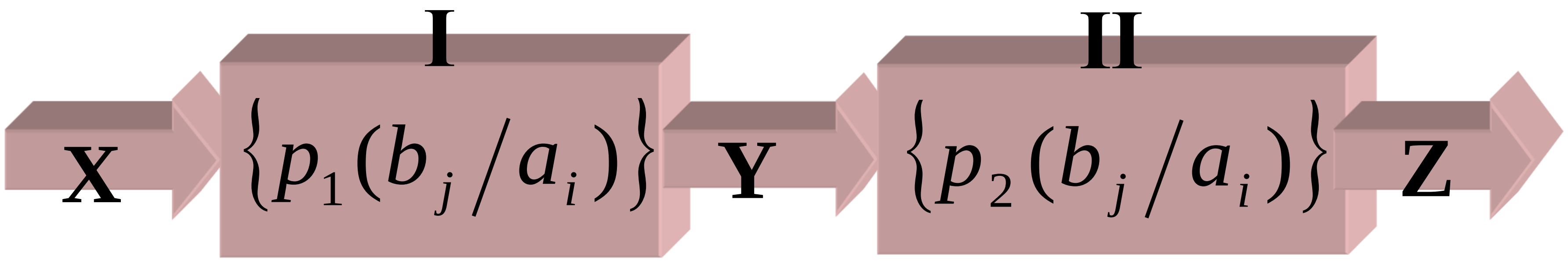


I(X;Y)随信道变化的曲线



5 数据处理定理

当消息经过多级处理后，随着处理器数目的增加，输入与输出之间的平均互信息量趋于变小。



数据处理模型

假定Y条件下X、Z相互独立

$$I(X;Z)\le I(Y;Z)$$

$$I(X;Z)\le I(X;Y)$$

假定Y条件下X、Z相互独立，则 $I(X;Z) \le I(Y;Z)$

$$I(X;Z) \le I(X;Y)$$

证明：

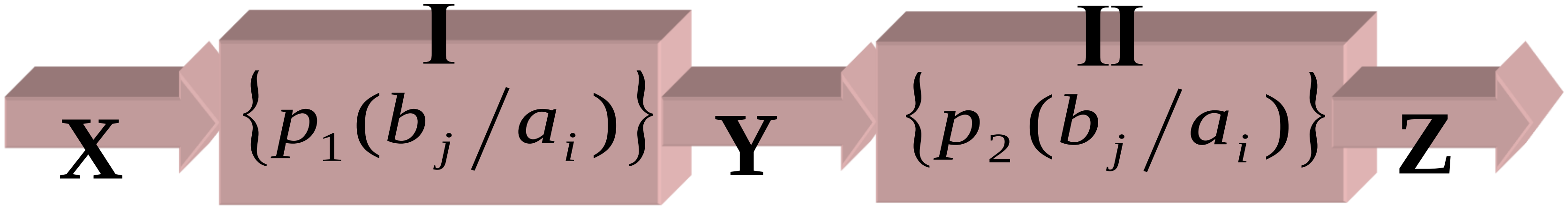
$$\left.\begin{array}{l} I(X;YZ) = H(X) - H(X/YZ) \\ I(X;Y/Z) = H(X/Z) - H(X/YZ) \end{array}\right\} I(X;Z) = I(X;YZ) - I(X;Y/Z)$$

$$I(X;Y) = I(X;YZ) - \underbrace{I(X;Z/Y)}_{=0}$$

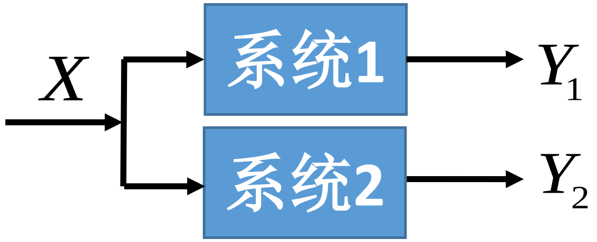
所以 $\boxed{I(X;Z) \le I(X;Y)}$

同理可证 $\boxed{I(X;Z) \le I(Y;Z)}$

数据处理定理告诉我们多次处理信息量将减少——信息不增原理



多次测量



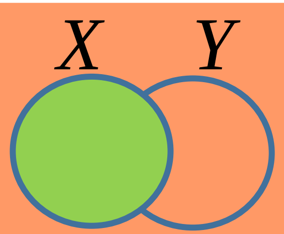
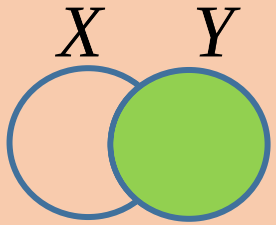
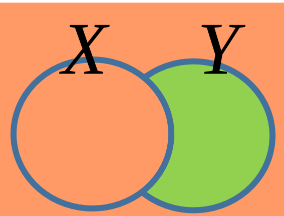
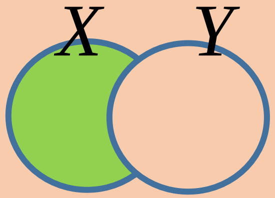
$$I(X;Y_1)=H(X)-\boxed{H(X/Y_1)}$$

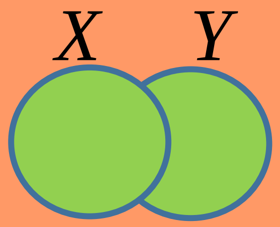
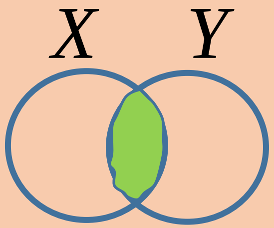
$$I(X;Y_1Y_2)=H(X)-\boxed{H(X/Y_1Y_2)}$$

$$I(X;Y_1Y_2)\geq I(X;Y_1)$$

多次测量要比单次测量能获得更多X的信息量

4.1.4 各种熵之间的关系

名称	符 号	关 系 式	图 示
无条件熵	$H(X)$		
	$H(Y)$	$H(Y)=H(Y/X)+I(X;Y)$ $=H(XY)-H(X/Y)\ge H(Y/X)$	
条件熵	$H(Y/X)$	$H(Y/X)=H(XY)-H(X)=H(Y)-I(X;Y)$	
	$H(X/Y)$	$H(X/Y)=H(XY)-H(Y)=H(X)-I(X;Y)$	

名称	符号	关系式	图示
联合熵	$H(XY)$	$H(XY)=H(X)+H(Y/X)=H(Y)+H(X/Y)$ $=H(X)+H(Y)-I(X;Y)$ $=H(X/Y)+H(Y/X)+I(X;Y)$	
交互熵	$I(X;Y)$ $=I(Y;X)$	$I(X;Y)=H(X)-H(X/Y)=H(Y)-H(Y/X)$ $=H(XY)-H(Y/X)-H(X/Y)$ $=H(X)+H(Y)-H(XY)$	

第4章

4.1 互 信 息 量 和 平 均 互 信 息 量

4.2 单 符 号 离 散 信 道 的 信 道 容 量

4.3 多 符 号 离 散 信 道 的 信 道 容 量

4.4 网 络 信 息 论

4.2.1 单符号离散信道容量定义

$X \in \{a_1, a_2, \cdots, a_n\}$ $\rightarrow$ $p(b_j/a_i)$ $\rightarrow$ $Y \in \{b_1, b_2, \cdots, b_m\}$

$$i = 1, 2, \cdots, n$$

$$j = 1, 2, \cdots, m$$

单符号离散信道的数学模型

单符号离散信道的信息测度：

$$I(X;Y)$$

用信道转移概率矩阵表述信道输入和输出之间的依存关系：

$$\begin{array}{c} \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{array} \begin{array}{c} \begin{array}{cccc} b_1 & b_2 & \cdots & b_m \end{array} \\ \begin{bmatrix} p(b_1/a_1), p(b_2/a_1), \cdots, p(b_m/a_1) \\ p(b_1/a_2), p(b_2/a_2), \cdots, p(b_m/a_2) \\ \cdots\cdots \\ p(b_1/a_n), p(b_2/a_n), \cdots, p(b_m/a_n) \end{bmatrix} \end{array}$$

$$I(X;Y)=f\left[p(a_i),p(b_j/a_i)\right]=\begin{cases} f\left[p(a_i)\right] & \text{固定信道} \\ f\left[p(b_j/a_i)\right] & \text{固定信源} \end{cases}$$

固定信道，$I(X;Y)$ 是$\{p(a_i)\}$的上凸函数，因此存在一种概率分布$\{p(a_i)\}$（即某一种信源），使信道所能传送的信息率$I(X;Y)$**最大**。

信道容量定义

$$C=\max_{p(a_i)} I(X;Y)$$

$$C = \max_{p(a_i)} I(X;Y)$$

损失熵

噪声熵

$$= \max_{p(a_i)} \left[H(X) - H(X/Y) \right] = \max_{p(a_i)} \left[H(Y) - H(Y/X) \right]$$

- 信道容量是描述信道特性的参量，是信道的最大传信能力。
- 与信源无关。

4.2.2 几种特殊离散信道的信道容量

1 离散无噪信道的信道容量

1 具有一一对应关系的无噪无损信道

$$X \in \{a_1, a_2, \cdots, a_n\} \qquad Y \in \{b_1, b_2, \cdots, b_n\}$$

$$\begin{bmatrix} 100\ldots\ldots 0 \\ 010\ldots\ldots 0 \\ \ldots\ldots \\ 000\ldots\ldots 1 \end{bmatrix}$$

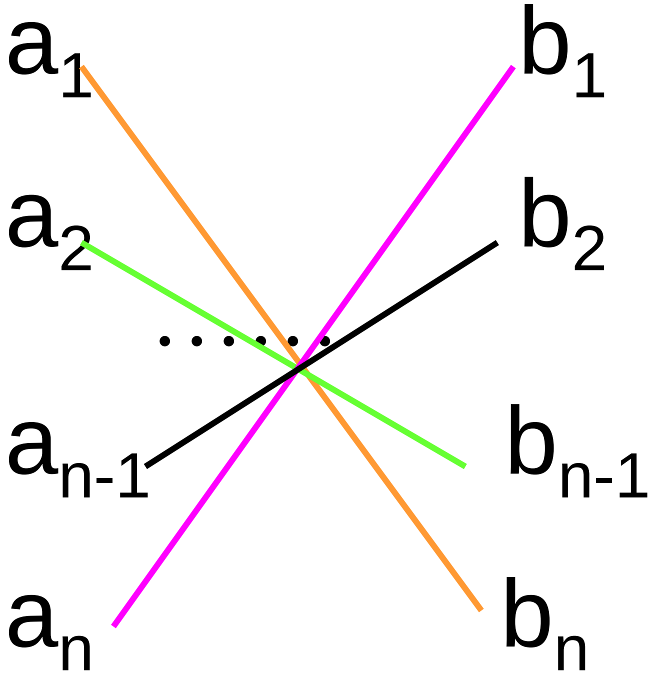
a_1 —— b_1

a_2 —— b_2

……

a_n —— b_n

$$\begin{bmatrix} 000\ldots\ldots 01 \\ 000\ldots\ldots 10 \\ \ldots\ldots \\ 010\ldots\ldots 00 \\ 100\ldots\ldots 00 \end{bmatrix}$$



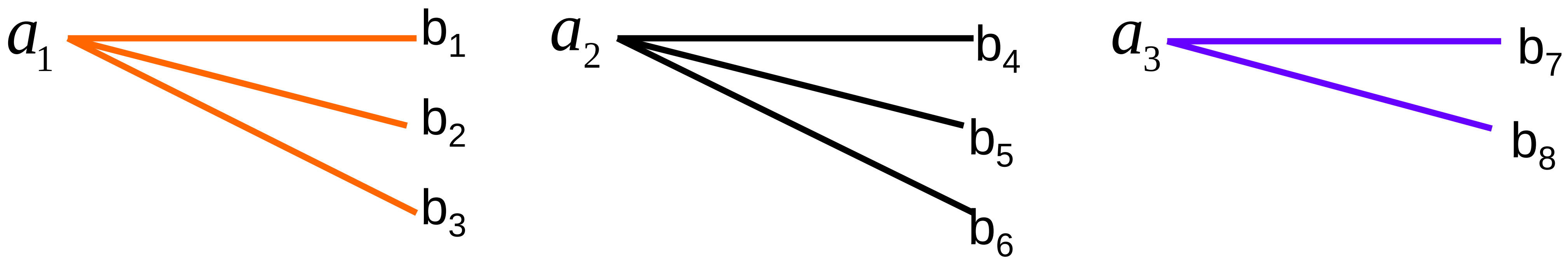
$$H(X/Y)=0,\ H(Y/X)=0 \Rightarrow I(X;Y)=H(X)=H(Y)$$

$$C=\max_{\boxed{p(a_i)}} I(X;Y)=\log n$$

信道容量只取决于信道的输入符号数n

信源概率分布如何时，可使得信道取到信道容量值？

② 具有扩展性能的无损有噪信道　　一个输入对应多个输出



$$\begin{bmatrix} p\left(b_1/a_1\right) & p\left(b_2/a_1\right) & p\left(b_3/a_1\right) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p\left(b_4/a_2\right) & p\left(b_5/a_2\right) & p\left(b_6/a_2\right) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & p\left(b_7/a_3\right) & p\left(b_8/a_3\right) \end{bmatrix}$$

此时 $H(X/Y)=0,\ H(Y/X)\neq 0,$ 且 $H(X)<H(Y)\ \Rightarrow I(X;Y)=H(X)$

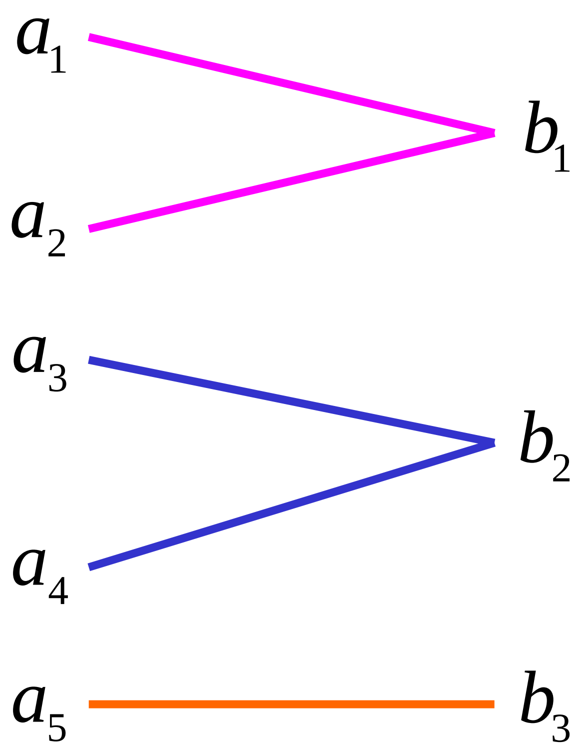
$$\boldsymbol{C}=\max_{p(a_i)} \boldsymbol{H}(\boldsymbol{X})=\log \boldsymbol{n}$$

信道容量只取决于信道的输入符号数n

信源概率分布如何时，可使得信道取到信道容量值？

③ 具有归并性能的无噪有损信道　多个输入对应一个输出

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



$$H(X/Y)\neq 0,\ \ \mathrm{H}(Y/X)=0$$

$$\Rightarrow I(X;Y)=H(Y)$$

$$\boldsymbol{C}=\max_{p(a_i)} \boldsymbol{H}(\boldsymbol{Y})=\log \boldsymbol{m}$$

信道容量只取决于信道的输出符号数m

信源概率分布如何时，可使得信道取到信道容量值？

$$\boldsymbol{P}_{5\times 2}=\begin{bmatrix}1 & 0 & 0\\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$p(b_j)=\sum_{i=1}^{5}p(a_ib_j)=\sum_{i=1}^{5}p(a_i)p(b_j/a_i)$$

$$P_Y=P_XP_{Y/X}$$

$$\begin{bmatrix}p(b_1) & p(b_2) & p(b_3)\end{bmatrix}_{1\times 3}$$

$$=\begin{bmatrix}p(a_1) & p(a_2) & \ldots & p(a_5)\end{bmatrix}_{1\times 5}\begin{bmatrix}p(b_1/a_1) & p(b_2/a_1) & p(b_3/a_1)\\ p(b_2/a_2) & p(b_2/a_2) & p(b_3/a_2)\\ \cdots & \cdots & \cdots\\ p(b_1/a_5) & p(b_2/a_5) & p(b_3/a_5)\end{bmatrix}_{5\times 2}$$

$$\begin{cases}\frac{1}{3}=p(a_1)+p(a_2)\\ \frac{1}{3}=p(a_3)+p(a_4)\\ \frac{1}{3}=p(a_5)\end{cases}$$

有无穷组解

2 强对称离散信道的信道容量

$$X \in \{a_1, a_2, \ldots\ldots a_n\} \qquad Y \in \{b_1, b_2, \ldots\ldots b_n\}$$

$$P_{n\times n} = \begin{bmatrix} 1-p & \frac{p}{n-1} & \cdots\cdots & \frac{p}{n-1} \\ \frac{p}{n-1} & 1-p & \cdots\cdots & \frac{p}{n-1} \\ & \cdots\cdots & \cdots\cdots & \\ \frac{p}{n-1} & \frac{p}{n-1} & \frac{p}{n-1} & 1-p \end{bmatrix}$$

行：所含元素相同，排列不同

列：所含元素相同，排列不同

p： 总体错误概率

$$C=\max_{p(a_i)} I(X;Y)=\max_{p(a_i)}[H(Y)-H(Y/X)]$$

$$P_{n\times n}=\begin{bmatrix} 1-p & \frac{p}{n-1} & \cdots\cdots & \frac{p}{n-1} \\ \frac{p}{n-1} & 1-p & \cdots\cdots & \frac{p}{n-1} \\ & \cdots\cdots & \cdots\cdots & \\ \frac{p}{n-1} & \frac{p}{n-1} & \frac{p}{n-1} & 1-p \end{bmatrix}$$

$$H(Y/X)=-\sum_{i=1}^{n}\sum_{j=1}^{n}p(a_i)p(b_j/a_i)\log p(b_j/a_i)=-\sum_{i=1}^{n}p(a_i)\left[\sum_{j=1}^{n}p(b_j/a_i)\log p(b_j/a_i)\right]$$

$$=H(1-p,\frac{p}{n-1},\ldots,\frac{p}{n-1})\triangleq H_{ni}$$

$$=H_{ni}\sum_{i=1}^{n}p(a_i)$$

$$=H_{ni}$$

只要转移概率矩阵行可排列，就可得到此式

由于转移矩阵各行元素相同，因此对于任意i, H_{ni} 相等

$$H_{ni}=H(1-p,\frac{p}{n-1},\ldots,\frac{p}{n-1})$$

$$=-[(1-p)\log(1-p)+(\frac{p}{n-1}\log\frac{p}{n-1})(n-1)]$$

$$=-[(1-p)\log(1-p)+p\log\frac{p}{n-1}]$$

$$H(Y/X)=H_{ni}=H(1-p,\frac{p}{n-1},\dots,\frac{p}{n-1})$$

$$P_{n\times n}=\begin{bmatrix} 1-p & \frac{p}{n-1} & \dots\dots & \frac{p}{n-1} \\ \frac{p}{n-1} & 1-p & \dots\dots & \frac{p}{n-1} \\ & \dots\dots & \dots\dots & \\ \frac{p}{n-1} & \frac{p}{n-1} & \frac{p}{n-1} & 1-p \end{bmatrix}$$

$$\Downarrow$$

$$C=\max_{p(a_i)} I(X;Y)=\max_{p(a_i)}[H(Y)-H(Y/X)]$$

$$=\max_{p(a_i)}[H(Y)-H_{ni}]=\log n-H_{ni}$$

相应的 $p(a_i)=\frac{1}{n}$

其中 $H_{ni}=-\sum_{j=1}^{n} p(b_j/a_i)\log p(b_j/a_i)$

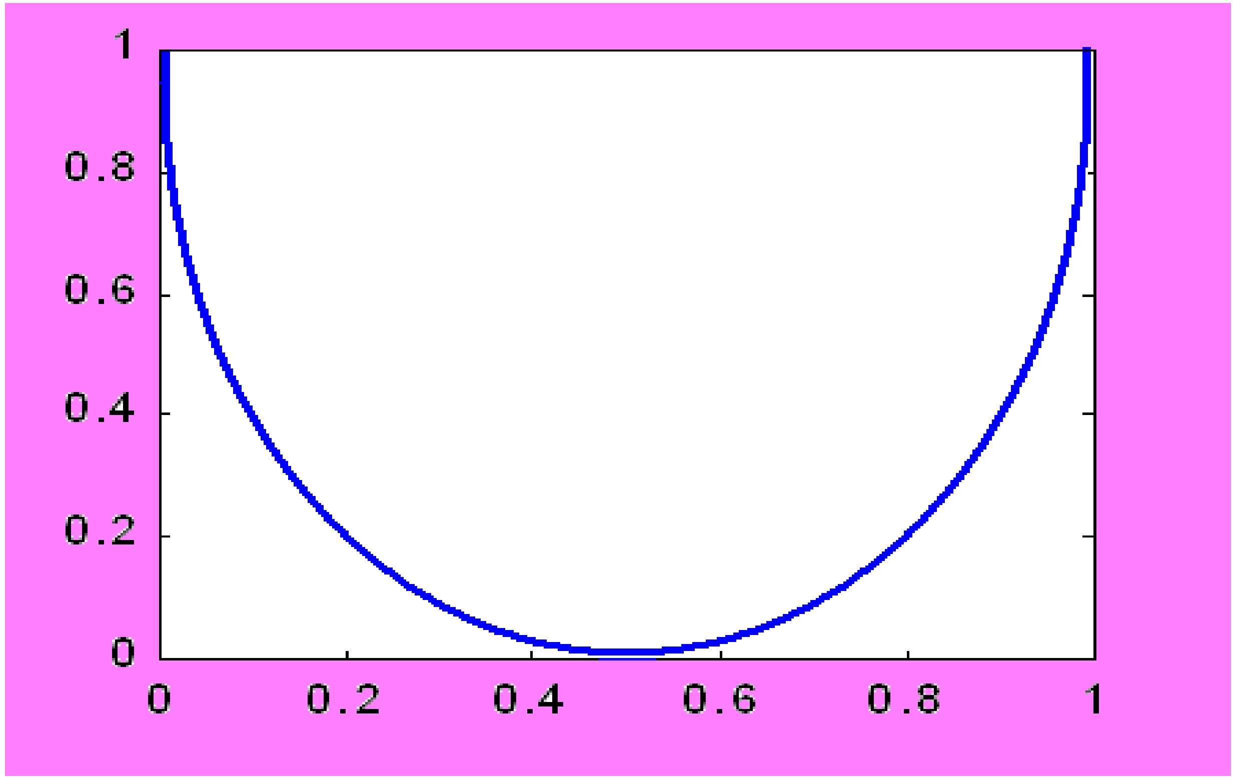
$$=H(1-p,\frac{p}{n-1},\dots,\frac{p}{n-1})$$

证明见P72 $P_Y=P_X P_{Y/X}$

信道转移概率矩阵若列可排列，则 信源等概→信宿等概

$$P_{2\times 2}=\begin{bmatrix}1-p & p\\ p & 1-p\end{bmatrix}$$ 二元对称信道容量 $\boxed{C=1-H(p)}$

$$H(p)=-\left[(1-p)\log(1-p)+p\log p\right]$$



二元对称信道容量曲线

3 对称离散信道的信道容量

矩阵中的每行都是集合$Q = \{q_1, q_2, \ldots\ldots, q_m\}$ 中的诸元素的不同排列，称矩阵的行是可排列的。

矩阵中的每列都是集合$P = \{p_1,\ p_2, \cdots\cdots, p_n\}$中的诸元素的不同排列，称矩阵的列是可排列的。

如果矩阵的行和列都是可排列的，称矩阵是可排列的。如果一个信道矩阵具有可排列性，则它所表示的信道称为 **对称信道**

对称信道中，当$n<m$时，P是Q的子集；当$n>m$，Q是P的子集；当$n=m$时，$P=Q$。

练习：判断下列矩阵表示的信道是否是对称信道？

$$[p_1]=\begin{bmatrix}\frac{1}{3}&\frac{1}{3}&\frac{1}{6}&\frac{1}{6}\\\frac{1}{6}&\frac{1}{6}&\frac{1}{3}&\frac{1}{3}\end{bmatrix}$$

√

$$[p_2]=\begin{bmatrix}\frac{1}{2}&\frac{1}{3}&\frac{1}{6}\\\frac{1}{3}&\frac{1}{6}&\frac{1}{2}\\\frac{1}{6}&\frac{1}{2}&\frac{1}{3}\end{bmatrix}$$

√

$$[p_3]=\begin{bmatrix}\frac{1}{3}&\frac{1}{3}&\frac{1}{6}&\frac{1}{6}\\\frac{1}{3}&\frac{1}{6}&\frac{1}{3}&\frac{1}{6}\end{bmatrix}$$

×

$$[p_4]=\begin{bmatrix}0.7&0.2&0.1\\0.1&0.2&0.7\end{bmatrix}$$

×

对称信道特点：

不一定是方阵

行可排列

列可排列

行和为1

列和不一定为1

对称离散信道的信道容量

$$\begin{array}{c} \text{行可排列} \\ \text{列可排列} \end{array} \begin{bmatrix} p(b_1/a_1), p(b_2/a_1), \cdots, p(b_m/a_1) \\ p(b_1/a_2), p(b_2/a_2), \cdots, p(b_m/a_2) \\ \cdots\cdots \\ p(b_1/a_n), p(b_2/a_n), \cdots, p(b_m/a_n) \end{bmatrix}$$

$$H(Y/X) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j/a_i)\log p(b_j/a_i)$$

$$= -\sum_{i=1}^{n} p(a_i)[\sum_{j=1}^{m} p(b_j/a_i)\log p(b_j/a_i)]$$

$$= H_{mi}$$

其中 $H_{mi} = -\sum_{j=1}^{m} p(b_j/a_i)\log p(b_j/a_i) = H(q_1, q_2, \ldots, q_m)$

（只要行可排列，就可得到此式）

$$C = \max_{p(a_i)}[H(Y) - H_{mi}] = \log m - H_{mi}$$

相应的 $p(a_i) = \dfrac{1}{n}$

信道转移概率矩阵若列可排列，
则 信源等概→信宿等概

强对称信道与对称信道比较

强对称信道	对称信道
n=m	n与m未必相等
矩阵对称	矩阵未必对称
行元素与列元素相同	行元素与列元素未必相同
行之和、列之和均为1	行之和为1

$$C = \log n - H_{ni}$$

相应的 $p(a_i) = \dfrac{1}{n}$

$$C = \log m - H_{mi}$$

相应的 $p(a_i) = \dfrac{1}{n}$

$$H(Y/X) = H(q_1, q_2, \cdots, q_m)$$

只要行可排列，就可得到此式

信道转移概率矩阵若列可排列，则 信源等概→信宿等概

4 准对称离散信道的信道容量

若信道矩阵的行是可排列的，但列不可排列，如果把列分成若干个不相交的子集，且由n行和各子集的诸列构成的各个子矩阵都是可排列的，则称相应的信道为准对称信道。例如：

$$\left[\begin{array}{cc|cc} \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{array}\right]$$

特点：

不一定是方阵

行可排列

列不可排列

行和为1

列和不一定为1

按列分割后的子阵 →

行可排列

列可排列

假设此时将矩阵的列分为s个子集，每个子集的元素个数（列数）分别是$m_1,m_2,\ldots,m_s$。对应的把 $p(b_j)$ 也分为s个子集: $M_1,M_2,\ldots,M_s$。

$$
\begin{array}{c|cccc}
 & b_1 & b_2 & b_3 & b_4 \\
\hline
a_1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\
a_2 & \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8}
\end{array}
$$

$m_1 = 2 \qquad m_2 = 2$

$M_1 = \{p(b_1), p(b_2)\}$

$M_2 = \{p(b_3), p(b_4)\}$

$$I(X;Y) = H(Y) - H(Y/X)$$

$H(Y/X) = H_{mi}$ （只要行可排列，就可得到此式）

$$C = \max_{p(a_i)}[H(Y) - H_{mi}]$$

$$= \max_{p(a_i)} H(Y) - H_{mi} \overset{?}{=} \log m - H_{mi}$$

问题：达到信道容量的信源的概率分布是怎样的呢?

$$H(Y) = -\sum_{j}^{m} p(b_j)\log p(b_j)$$

$$= -\sum_{p(b_j)\in M_1} p(b_j)\log p(b_j) - \cdots - \sum_{p(b_j)\in Ms} p(b_j)\log p(b_j)$$

可以证明（P74）

$$-\sum_{p(b_j)\in M_k} p(b_j)\log p(b_j) \le -m_k \bar{p}(b_k)\log \bar{p}(b_k)$$

其中 $\bar{p}(b_k) = \frac{1}{m_k}\sum_{p(a_j)\in M_k} p(b_j) \quad k = 1,2,\ldots,s$

是子集 M_k 中概率的平均值

且当子集 M_k 中 $p(b_j) = \bar{p}(b_k)$ 时上式取等号。

信道转移概率矩阵若局部列可排列，
则 信源等概→信宿局部等概

$$H(Y) = -\sum_{p(b_j)\in M_1} p(b_j)\log p(b_j) - \cdots - \sum_{p(b_j)\in Ms} p(b_j)\log p(b_j)$$

$$\leq -\sum_{k=1}^{s} m_k \bar{p}(b_k)\log \bar{p}(b_k)$$

$$C = \max_{p(a_i)} H(Y) - H_{mi} = -\sum_{k=1}^{s} m_k \bar{p}(b_k)\log \bar{p}(b_k) - H_{mi}$$

只要信源等概分布，就可使得达到信道容量。

例4.2.1 求如下信道的信道容量。

$$[P]=\begin{bmatrix} 0.5 & 0.25 & 0.125 & 0.125 \\ 0.25 & 0.5 & 0.125 & 0.125 \end{bmatrix} \qquad n=2, m=4$$

解：分成子矩阵 $[P_1]=\begin{bmatrix} 0.5 & 0.25 \\ 0.25 & 0.5 \end{bmatrix}$ $\quad [P_2]=\begin{bmatrix} 0.125 & 0.125 \\ 0.125 & 0.125 \end{bmatrix}$ $\quad m_1=m_2=2$

$$\overline{p}(b_1)=\frac{1}{m_1}\sum_{j=1}^{2} p(b_j)=\frac{1}{m_1}\sum_{j=1}^{2}\sum_{i=1}^{2} p(a_i)p(b_j/a_i)=0.5\times(0.5+0.25)=0.375$$

$$\overline{p}(b_2)=0.5\times(0.125+0.125)=0.125$$

$$H_{mi}=-\sum_{j=1}^{4} p(b_j/a_i)\log p(b_j/a_i)=H(0.5,0.25,0.125.0.125)=1.75$$

$$C=-\sum_{k=1}^{2} m_k\overline{p}(b_k)\log\overline{p}(b_k)-H_{mi}=0.0612(bit/sign)$$

强对称信道

$$C = \log n - H_{ni} = \log n - H(1-p, \tfrac{p}{n-1}, \dots, \tfrac{p}{n-1}) \qquad p(a_i) = \tfrac{1}{n}$$

对称信道

$$C = \log m - H_{mi} = \log m - H(q_1, q_2, \dots, q_m) \qquad p(a_i) = \tfrac{1}{n}$$

准对称信道

$$C = -\sum_{k=1}^{s} m_k \bar{p}(b_k) \log \bar{p}(b_k) - H(q_1, q_2, \dots, q_m) \qquad p(a_i) = \tfrac{1}{n}$$

将准对称信道按列划分成 s 个行列可排列的子集合

第 k 个子集合中概率的平均值为 $\bar{p}(b_k)$

思考： 准对称信道的信道容量还有别的计算方法吗?

4.2.3 离散信道容量的一般计算方法

对一般离散信道而言，求信道容量，就是在固定信道的条件下，对所有可能的输入概率分布$\{p(a_i)\}$，求平均互信息的极大值。采用拉格朗日乘子法来计算。

$$\phi = I(X;Y) - \lambda\left[\sum_{i}^{n} P(a_i) - 1\right],\ \text{令}\ \frac{\partial \phi}{\partial P(a_i)} = 0,\ \text{则有}$$

$$\frac{\partial}{\partial p(a_i)}\left\{-\sum_{j=1}^{m} p(b_j)\log p(b_j) + \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j/a_i)\log p(b_j/a_i) - \lambda\left[\sum_{i=1}^{n} p(a_i) - 1\right]\right\}$$

$$\because\ p(b_j) = \sum_{i=1}^{n} p(a_i)p(b_j/a_i),\ \frac{dp(b_j)}{dp(a_i)} = p(b_j/a_i),\ \log x = \ln x \log_2 e$$

$$\therefore \quad \frac{\partial\phi}{\partial p(a_i)} = -\left\{\sum_j^m \left[p(b_j/a_i)\log p(b_j) + p(b_j/a_i)\log_2 e \right]\right\}$$

$$+\sum_j^m p(b_j/a_i)\log p(b_j/a_i) - \lambda = 0$$

$$\Downarrow$$

$$-\sum_{j=1}^m p(b_j/a_i)\log p(b_j) + \sum_{j=1}^m p(b_j/a_i)\log p(b_j/a_i) = \log_2 e + \lambda \quad (4.2.16)$$

两边乘$p(a_i)$，并求和有：

$$\underbrace{\boxed{-\sum_{i=1}^n\sum_{j=1}^m p(a_i)p(b_j/a_i)\log_2 p(b_j)}}_{H(Y)} + \underbrace{\boxed{\sum_{i=1}^n\sum_{j=1}^m p(a_i)p(b_j/a_i)\log_2 p(b_j/a_i)}}_{-H(Y/X)} = \log_2 e + \lambda$$

$$I(X;Y) = \log_2 e + \lambda \quad \Longleftrightarrow \quad C = \log_2 e + \lambda \tag{4.2.17}$$

将(4.2.25)代入(4.2.23)，则有:

$$\sum_{j=1}^{m} p(b_j/a_i)\log p(b_j/a_i) = \sum_{j=1}^{m} p(b_j/a_i)\log p(b_j) + C$$

$$= \sum_{j=1}^{m} p(b_j/a_i)[\log p(b_j) + C]$$

令 $$\beta_j = \log p(b_j) + C \tag{4.2.18}$$

则 $$\sum_{j=1}^{m} p(b_j/a_i)\log p(b_j/a_i) = \sum_{j=1}^{m} p(b_j/a_i)\beta_j \tag{4.2.19}$$

方程组: 有m个未知数，n个方程

由(4.2.19)求出β_j， 再由(4.2.18)求出$p(b_j)$。

$$p(b_j) = 2^{\beta_j - C}$$

$$\sum_j^m p(b_j) = \sum_j^m 2^{\beta_j - C} = 1 \rightarrow 2^C = \sum_j^m 2^{\beta_j}$$

$$\boxed{C = \log_2 \sum_{j=1}^m 2^{\beta_j}}$$

由$p(b_j) = \sum_{i=1}^n p(a_i) p(b_j / a_i)$求出$p(a_i)$。

如果$p(a_i)$满足概率约束条件，则C正确，求解结束。

总结C的求法，过程如下：

1. 由$\sum_{j}^{m} p(b_j/a_i)\log p(b_j/a_i) = \sum_{j}^{m} p(b_j/a_i)\beta_j$ 求 β_j ；

2. 由$C = \log_2\left(\sum_{j=1}^{m} 2^{\beta_j}\right)$求$C$;

3. 由 $p(b_j) = 2^{\beta_j - C}$ 求 $p(b_j)$；

4. 由$\sum_{j}^{m} p(b_j/a_i)\log p(b_j/a_i) = \sum_{j}^{m} p(b_j/a_i)\beta_j$ 求出 $p(a_i)$，并验证。

例4.2.2 有一信道矩阵 $\begin{bmatrix} 1 & 0 \\ \varepsilon & 1-\varepsilon \end{bmatrix}$ 求C。

1 $1\times\beta_1+0\times\beta_2=1\times\log 1+0\times\log 0=0 \rightarrow \beta_1=0$

$$\varepsilon\beta_1+(1-\varepsilon)\beta_2=\varepsilon\log\varepsilon+(1-\varepsilon)\log(1-\varepsilon)$$

$$\therefore\ \beta_2=\frac{\varepsilon}{1-\varepsilon}\log\varepsilon+\log(1-\varepsilon)=\log\left[(1-\varepsilon)\cdot\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}\right]$$

2 $C=\log_2\left(\sum_{j=1}^{m}2^{\beta_j}\right)=\log_2\left[1+(1-\varepsilon)\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}\right]$

3

$$p(b_j) = 2^{\beta_j - C}$$

$$p(b_1) = 2^{\beta_1 - C} = 2^{-C} = \frac{1}{1 + (1-\varepsilon)\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}$$

$$p(b_2) = 1 - p(b_1) = \frac{(1-\varepsilon)\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}{1 + (1-\varepsilon)\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}$$

4

$$p(b_1) = p(a_1)p(b_1/a_1) + p(a_2)p(b_1/a_2)$$

$$p(b_2) = p(a_1)p(b_2/a_1) + p(a_2)p(b_2/a_2)$$

$$\begin{bmatrix} 1 & 0 \\ \varepsilon & 1-\varepsilon \end{bmatrix}$$

$$p(b_1) = p(a_1) + \varepsilon p(a_2) \quad p(b_2) = (1-\varepsilon) p(a_2)$$

$$\therefore\ p(a_1) = \frac{1-\varepsilon^{\frac{1}{1-\varepsilon}}}{1+(1-\varepsilon)\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}$$

$$p(a_2) = \frac{\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}{1+(1-\varepsilon)\varepsilon^{\frac{\varepsilon}{1-\varepsilon}}}$$

$$\because 0 \le \varepsilon \le 1 \quad \Longrightarrow \quad p(a_1), p(a_2) \ge 0$$

第4章

4.1 互信息量和平均互信息量

4.2 单符号离散信道的信道容量

4.3 多符号离散信道的信道容量

4.4 网络信息论

4.3.1 多符号离散信道容量的数学模型

$$\vec{X} = X_1 X_2 \cdots X_K \cdots X_N$$

$$X_K \in \{a_1, a_2, \cdots, a_n\}$$

$$\alpha_i = (a_{i_1} a_{i_2} \cdots a_{i_N})$$

$$i = 1, 2, \cdots, n^N$$

$$i_1, i_2, \cdots, i_N \in \{1, 2, \cdots, n\}$$

$$\vec{Y} = Y_1 Y_2 \cdots Y_K \cdots Y_M$$

$$Y_K \in \{b_1, b_2, \cdots, b_m\}$$

$$\beta_j = (b_{j_1} b_{j_2} \cdots b_{j_M})$$

$$j = 1, 2, \cdots, m^M$$

$$j_1, j_2, \cdots, j_M \in \{1, 2, \cdots, m\}$$

输入 $\vec{X} \rightarrow \boxed{P(\vec{Y}/\vec{X})} \rightarrow \vec{Y}$ 输出

$$\left\{ \vec{X} \quad P(\vec{Y}/\vec{X}) \quad \vec{Y} \right\}$$

转移概率矩阵

$$
\boldsymbol{P}(\vec{\boldsymbol{Y}}/\vec{\boldsymbol{X}}) = \begin{bmatrix} \boldsymbol{p}(\beta_1/\alpha_1) & \boldsymbol{p}(\beta_2/\alpha_1) & \cdots & \boldsymbol{p}(\beta_{\boldsymbol{m}^M}/\alpha_1) \\ \boldsymbol{p}(\beta_1/\alpha_2) & \boldsymbol{p}(\beta_2/\alpha_2) & \cdots & \boldsymbol{p}(\beta_{\boldsymbol{m}^M}/\alpha_2) \\ & \cdots & \cdots & \\ \boldsymbol{p}(\beta_1/\alpha_{\boldsymbol{n}^N}) & \boldsymbol{p}(\beta_2/\alpha_{\boldsymbol{n}^N}) & \cdots & \boldsymbol{p}(\beta_{\boldsymbol{m}^M}/\alpha_{\boldsymbol{n}^N}) \end{bmatrix}
$$

$$
0 \le p(\beta_j/\alpha_i) \le 1, \quad \sum_{j=1}^{m^M} p(\beta_j/\alpha_i) = 1, i = 1, 2, \cdots, n^N
$$

4.3.2 多符号离散信道容量定义

多符号离散信道的平均互信息量

$$I(\vec{X};\vec{Y})=\sum_{i=1}^{n^N}\sum_{j=1}^{m^M}p(\alpha_i\beta_j)I(\alpha_i;\beta_j)=\sum_{i=1}^{n^N}\sum_{j=1}^{m^M}p(\alpha_i\beta_j)\log\frac{p(\alpha_i/\beta_j)}{p(\alpha_i)}$$

$$I(\vec{X};\vec{Y})=\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_M}^{m}p(a_{i_1}\cdots a_{i_N}b_{j_1}\cdots b_{j_M})\log\frac{p(a_{i_1}\cdots a_{i_N}/b_{j_1}\cdots b_{j_M})}{p(a_{i_1}\cdots a_{i_N})}$$

定义多符号离散信道容量为

$$\boxed{C=\max_{\{p(\alpha_i)\}}I(X;Y)}$$

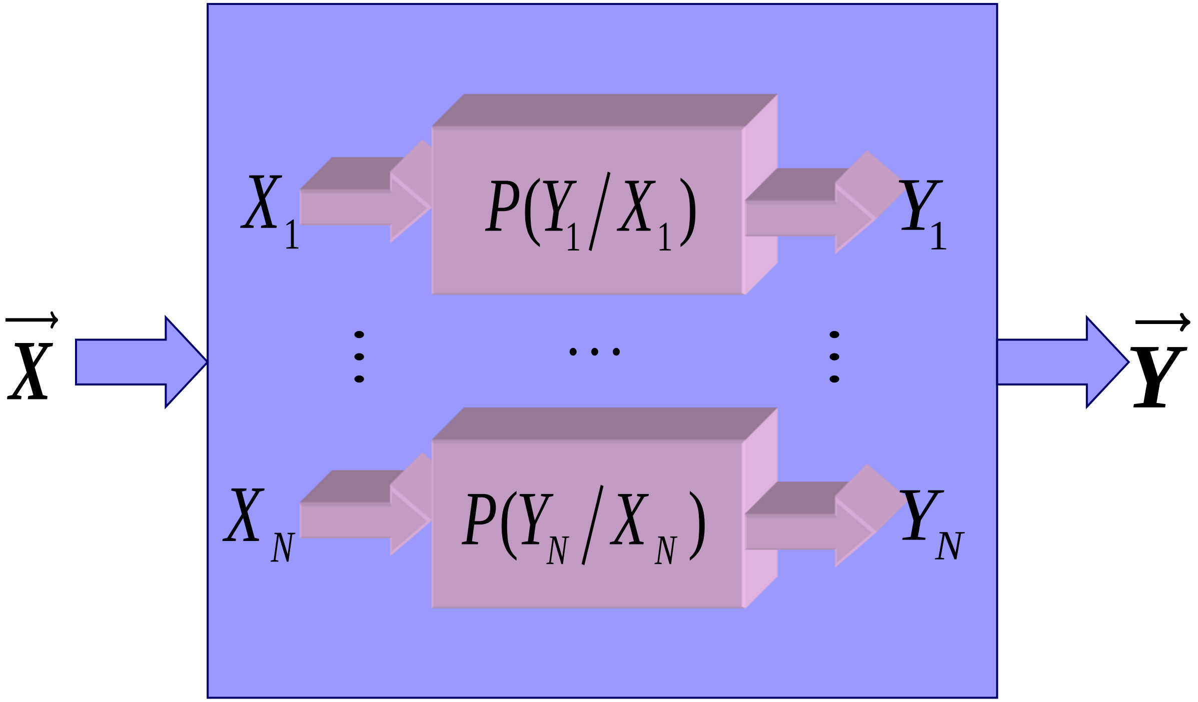
4.3.3 离散无记忆扩展信道的信道容量

无记忆：Y_k仅与X_k有关

$$\boldsymbol{P}(\vec{\boldsymbol{Y}}/\vec{\boldsymbol{X}}) = P(Y_1Y_2\cdots Y_N/X_1X_2\cdots X_N)$$
$$= P(Y_1/X_1)P(Y_2/X_2)\cdots P(Y_N/X_N)$$
$$= \prod_{k=1}^{N} P(Y_k/X_k)$$

$$\boldsymbol{I}(\vec{\boldsymbol{X}};\vec{\boldsymbol{Y}}) = \boldsymbol{H}(\vec{\boldsymbol{Y}}) - \boldsymbol{H}(\vec{\boldsymbol{Y}}/\vec{\boldsymbol{X}})$$

$$\boldsymbol{I}(\boldsymbol{X}_i;\boldsymbol{Y}_i) = \boldsymbol{H}(\boldsymbol{Y}_i) - \boldsymbol{H}(\boldsymbol{Y}_i/\boldsymbol{X}_i)$$



单符号离散信道的N次扩展信道的实现模型

$$H(\vec{Y}/\vec{X}) = -\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\ldots a_{i_N})\boxed{p(b_{j_1}b_{j_2}\ldots b_{j_N}/a_{i_1}a_{i_2}\ldots a_{i_N})}\cdot\log_2\boxed{p(b_{j_1}b_{j_2}\ldots b_{j_N}/a_{i_1}a_{i_2}\ldots a_{i_N})}$$

$$= -\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\cdots a_{i_N})p(b_{j_1}/a_{i_1})\cdots p(b_{j_N}/a_{i_N})\cdot\boxed{\log_2\left[p(b_{j_1}/a_{i_1})\cdots p(b_{j_N}/a_{i_N})\right]}$$

$$= \boxed{-\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\cdots a_{i_N})p(b_{j_1}/a_{i_1})\cdots p(b_{j_N}/a_{i_N})\cdot\log_2 p(b_{j_1}/a_{i_1})} - \ldots$$

$$-\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\cdots a_{i_N})p(b_{j_1}/a_{i_1})\cdots p(b_{j_N}/a_{i_N})\cdot\log_2 p(b_{j_N}/a_{i_N})$$

$$-\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\cdots a_{i_N})\underline{p(b_{j_1}/a_{i_1})}\cdots p(b_{j_N}/a_{i_N})\cdot\underline{\log_2 p(b_{j_1}/a_{i_1})}$$

$$=-\sum_{i_1=1}^{n}\sum_{j_1=1}^{m} p(b_{j_1}/a_{i_1})\log_2 p(b_{j_1}/a_{i_1})\left[\sum_{i_2=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_2=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\cdots a_{i_N})p(b_{j_2}/a_{i_2})\cdots p(b_{j_N}/a_{i_N})\right]$$

$$=-\sum_{i_1=1}^{n}\sum_{j_1=1}^{m} p(b_{j_1}/a_{i_1})\log_2 p(b_{j_1}/a_{i_1})\left[\underbrace{\sum_{i_2=1}^{n}\cdots\sum_{i_N=1}^{n} p(a_{i_1}a_{i_2}\cdots a_{i_N})}_{=p(a_{i_1})}\underbrace{\sum_{j_2=1}^{m} p(b_{j_2}/a_{i_2})}_{=1}\cdots\underbrace{\sum_{j_N=1}^{m} p(b_{j_N}/a_{i_N})}_{=1}\right]$$

$$=-\sum_{i_1=1}^{n}\sum_{j_1=1}^{m} p(a_{i_1})p(b_{j_1}/a_{i_1})\log_2 p(b_{j_1}/a_{i_1})$$

$$=H(Y_1/X_1)$$

$$H(\vec{Y}/\vec{X}) = -\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\ldots a_{i_N})\boxed{p(b_{j_1}b_{j_2}\ldots b_{j_N}/a_{i_1}a_{i_2}\ldots a_{i_N})}\cdot\log_2\boxed{p(b_{j_1}b_{j_2}\ldots b_{j_N}/a_{i_1}a_{i_2}\ldots a_{i_N})}$$

$$= -\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\cdots a_{i_N})p(b_{j_1}/a_{i_1})\cdots p(b_{j_N}/a_{i_N})\cdot\boxed{\log_2\left[p(b_{j_1}/a_{i_1})\cdots p(b_{j_N}/a_{i_N})\right]}$$

$$= \boxed{-\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\cdots a_{i_N})p(b_{j_1}/a_{i_1})\cdots p(b_{j_N}/a_{i_N})\cdot\log_2 p(b_{j_1}/a_{i_1})} - \ldots$$

$$-\sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_N=1}^{m} p(a_{i_1}a_{i_2}\cdots a_{i_N})p(b_{j_1}/a_{i_1})\cdots p(b_{j_N}/a_{i_N})\cdot\log_2 p(b_{j_N}/a_{i_N})$$

$$= H(Y_1/X_1) + H(Y_2/X_2) + \ldots + H(Y_N/X_N) = \sum_{k=1}^{N} H(Y_k/X_k)$$

$$\boldsymbol{I(\vec{X};\vec{Y}) = H(\vec{Y}) - H(\vec{Y}/\vec{X})} = H(Y_1Y_2\cdots Y_N) - \sum_{k=1}^{N} H(Y_k/X_k)$$

$$H(Y_1Y_2\cdots Y_N) = H(Y_1) + H(Y_2/Y_1) + \ldots + H(Y_N/Y_1\ldots Y_{N-1}) \le \sum_{k=1}^{N} H(Y_k)$$

如果输出端 $Y_1Y_2\cdots Y_N$ 相互独立，上式等号成立。

$$\boldsymbol{I}(\vec{\boldsymbol{X}};\vec{\boldsymbol{Y}})\le\sum_{k=1}^{N}H(Y_k)-\sum_{k=1}^{N}H(Y_k/X_k)=\sum_{k=1}^{N}[H(Y_k)-H(Y_k/X_k)]=\sum_{k=1}^{N}I(X_k;Y_k)$$

如果输出端 $Y_1Y_2\cdots Y_N$ 相互独立，则上式等号成立。

若输入端 $X_1X_2\cdots X_N$ 也是无记忆的，且是同分布的，则有

$$I(X_k;Y_k)=I(X;Y) \Longrightarrow \boldsymbol{I}(\vec{\boldsymbol{X}};\vec{\boldsymbol{Y}})=NI(X;Y)$$

离散无记忆信道*N*次扩展信道的**信道容量**为

$$\boxed{C^N=NC}$$

单符号离散信源的*N*次扩展信源的**信源熵**

$$H(X^N)=NH(X)$$

4.3.4 独立并联信道的信道容量

将离散无记忆信道的N次扩展信道加以推广，即令信道输入和输出序列中的各随机变量取值于不同的符号集合，就构成了独立并联信道。独立并联信道中输出端随机变量Y_k仅与对应的输入随机变量X_k有关，两者构成了一条独立的单符号离散信道，用C_k表示，则

$$C^N \le C_1 + C_2 + \cdots + C_N = \sum_{k=1}^{N} C_k$$

当N个输入随机变量相互独立，且每个输入随机变量的概率分布达到各自信道容量的最佳分布时，独立并联信道容量最大

$$C_{\max}^N = \sum_{k=1}^{N} C_k$$

第4章 小 结

1. 掌握单符号离散信道的数学模型

   转移概率矩阵

2. 掌握互信息的定义、计算公式、物理意义、性质

3. 掌握平均互信息的定义、计算公式、物理意义、性质

4. 掌握熵与平均互信息的关系

5. 掌握单符号离散信道的**信道容量**

- 定义
- 几种特殊信道的信道容量
  - 三种无噪信道
  - 强对称信道
  - 对称信道
  - 准对称信道
- 信道容量的一般计算方法（考试不作要求）

6. 多符号离散信道的信道容量

- 了解定义
- 掌握单符号无记忆扩展信道的数学模型与信道容量
- 了解独立并联信道的信道容量