《信息论基础》

第4章 离散信道容量

吉小鹏

E-mail: 003163@nuist.edu.cn

南京信息工程大学 电子与信息工程学院 尚贤楼209

在实际通信系统中，为了提高传输效率，往往需要把信源的大量冗余进行压缩，即所谓**信源编码**。

但是考虑通信中的抗干扰问题，则需要信源具有一定的冗余度。因此在传输之前通常加入某些特殊的冗余度，即所谓**信道编码**，以达到通信系统中理想的传输有效性和可靠性。

信道是构成信息流通系统的重要部分，其任务是以信号形式传输和存储信息。在物理信道一定的情况下，人们总是希望传输的信息越多越好。这不仅与信源输出符号的统计特性和载荷信息的信号形式有关，还与物理信道本身的特性有关。

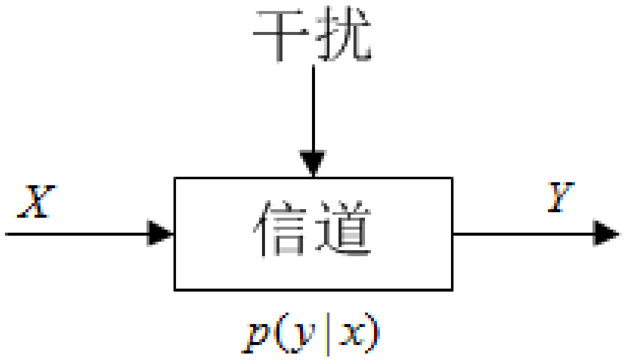
提纲

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

信道模型



信道可以看成一个变换器，将输入事件 x 变换成输出事件 y 。由于干扰存在，一个输入事件总是以一定的概率变换成各种可能的输出事件，所以观测者只能从统计的观点来判断输出事件。

输入事件的概率空间以 $\begin{bmatrix} X \\ P \end{bmatrix}$ 表示；输出事件的概率空间以 $\begin{bmatrix} Y \\ P \end{bmatrix}$ 表示。

输入事件 x 通过信道变换成输出事件 y ，可以采用条件概率分布函数 $\mathrm{p}(y\,|\,x)$ 描述。

4.1 互信息量和平均互信息量

■ 信道分类

根据时间特性和事件集分类：

- 离散信道：X和Y均为离散事件集（数字信道）
- 连续信道：X和Y均为连续事件集（模拟信道）
- 半连续信道：X和Y，一个离散，一个连续
- 时间离散的连续信道：有限或无限可数个取自连续集
- 波形信道：信道的输入和输出都是时间的实函数

根据输入、输出个数分类：

- 两端信道：输入和输出都只有一个事件集
- 多端信道：输入输出至少有一个有两个及以上事件集

4.1 互信息量和平均互信息量

信道分类

根据信道的统计特性分类：

恒参信道：信道的参数（统计特性）不随时间变化
随参信道：信道的参数（统计特性）随时间变化

根据信道的记忆特性分类：

无记忆信道：信道输出集Y仅与当前输入集X有关
有记忆信道：Y不仅与当前X有关，还与过去的X有关

4.1 互信息量和平均互信息量

■ 离散信道的数学模型

设离散信道输入空间A={a_1, a_2, …, a_r}，相应概率分布为{p_i}；

输出空间B ={b_1, b_2, …, b_s}，相应概率分布为{q_j}；

信道的输入序列X={X_1, X_2, …, X_N}，取值为x={x_1,x_2,…,x_N}，

相应的输出序列Y={Y_1, Y_2, …, Y_N}，取值为y={y_1,y_2,…,y_N}，

信道特性可用转移概率来描述

$$p(\mathbf{y}\,|\,\mathbf{x}) = p(y_1 y_2 \cdots y_N \,|\, x_1 x_2 \cdots x_N)$$

信道的数学模型可表示为：

$$\{\mathbf{X},\, p(\mathbf{y}\,|\,\mathbf{x}),\, \mathbf{Y}\}$$

4.1 互信息量和平均互信息量

■ 离散信道的数学模型

根据信道的统计特性即条件概率 $p(\mathbf{y}|\mathbf{x})$ 的不同，离散信道又可分成三种情况：

(1)无干扰（无噪）信道

信道中不存在随机干扰，或者干扰很小可忽略不计，输出符号 y_{n} 与输入符号 x_{n} 有**确定的一一对应的关系**，即 $y_{\mathrm{n}} = f(x_{\mathrm{n}})$，且

$$p(y_n \mid x_n) = \begin{cases} 1 & y_n = f(x_n) \\ 0 & y_n \neq f(x_n) \end{cases} \quad (1 \leq n \leq N)$$

离散信道的数学模型

(2) 有干扰无记忆信道

信道中存在随机干扰，输出符号与输入符号之间无确定的对应关系；但信道中任一时刻的输出符号仅统计以来与对应时刻的输入符号，而与非对应时刻的输入符号及其他任意时刻的输出符号无关。可用下面的条件概率表示

$$p(\mathbf{y}\mid\mathbf{x})=p(y_1y_2\cdots y_N\mid x_1x_2\cdots x_N)=\prod_{n=1}^{N}p(y_n\mid x_n)$$

(3) 有干扰有记忆信道

实际的信道既有干扰，又有记忆。

离散信道的数学模型

定义：若离散信道对任意 N 长的输入、输出序列有：

$$p(\mathbf{y}\,|\,\mathbf{x})=\prod_{n=1}^{N}p(y_n\,|\,x_n)$$

则称它为**离散无记忆信道**。

注：在任何时刻，信道的输出只与此时的信道输入有关，而与以前的输入无关。

定义：对任意 n 和 m，若离散无记忆信道还满足：

$$p(y_n=j\,|\,x_n=i)=p(y_m=j\,|\,x_m=i)$$

则称此信道为**平稳**的或**恒参**的。

注：**平稳信道下的转移概率不随时间变化。**

4.1 互信息量和平均互信息量

■ 单符号离散信道

单符号离散信道的输入随机变量为X，其取值为x， $x \in A = \{a_1, a_2, \dots, a_r\}$，输入概率空间为

$$\begin{bmatrix} X \\ P(x) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \cdots & a_r \\ P(a_1) & P(a_2) & \cdots & P(a_r) \end{bmatrix},$$

输出随机变量为Y，其取值为y， $y \in B = \{b_1, b_2, \dots, b_s\}$。信道的传递概率为$p(y|x) = p(Y = b_j | X = a_i) = p(b_j | a_i)$

信道的传递概率（又称转移概率）是一个条件概率，且

$$p(b_j|a_i) \ge 0,\ \sum_{j=1}^{s} p(b_j|a_i) = 1 \quad i = 1, 2, \cdots, r;\ \ j = 1, 2, \cdots, s.$$

由于信道中存在干扰，输入符号a_i在传输中可能会产生错误，这种干扰对传输的影响可用$p(b_j|a_i)$, $i = 1, 2, \dots, r$; $j = 1, 2, \dots, s$ 来描述。

4.1 互信息量和平均互信息量

单符号离散信道

信道传递概率实际上是一个传递概率矩阵，称为信道矩阵 $\boldsymbol{P}$。

$$P = \{p(b_j \mid a_i), i = 1, 2, \cdots, r; j = 1, 2, \cdots, s\}$$

或

$$P = \begin{bmatrix} P(b_1 \mid a_1) & P(b_2 \mid a_1) & \cdots & P(b_s \mid a_1) \\ P(b_1 \mid a_2) & P(b_2 \mid a_2) & \cdots & P(b_s \mid a_2) \\ \vdots & \vdots & \ddots & \vdots \\ P(b_1 \mid a_r) & P(b_2 \mid a_r) & \cdots & P(b_s \mid a_r) \end{bmatrix}$$

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1s} \\ p_{21} & p_{22} & \cdots & p_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1} & p_{r2} & \cdots & p_{rs} \end{bmatrix}$$

$$0 \le p_{ij} \le 1,$$

$$\sum_{j=1}^{s} p_{ij} = 1 \quad (i = 1, 2, \cdots, r)$$

4.1 互信息量和平均互信息量

- 单符号离散信道

(1) 输入/输出符号的联合概率为

$$P(x=a_i, y=b_j)=P(a_ib_j)$$

则有

$$P(a_ib_j)=P(a_i)\cdot P(b_j\,|\,a_i)=P(b_j)\cdot P(a_i\,|\,b_j)$$

其中

$P(b_j\,|\,a_i)$ 是信道传递概率，即发送为 a_i，通过信道传递，接收到为 b_j 的概率（**前向概率**）。它是由信道噪声引起的，所以描述了信道噪声的特性。

$P(a_i\,|\,b_j)$ 是已知信道输出端接收到的符号为 b_j，但发送端的输入符号为 a_i 的概率（**后向概率**）。

4.1 互信息量和平均互信息量

■ 单符号离散信道

(2) 根据全概率公式可得输出符号的概率

$$P(b_j)=\sum_{i=1}^{r}P(a_i)\cdot P(b_j\,|\,a_i),\ j=1,2,\cdots,s$$

或矩阵形式
$$\begin{bmatrix}P(b_1)\\P(b_2)\\\vdots\\P(b_s)\end{bmatrix}=P^{\mathrm{T}}\begin{bmatrix}P(a_1)\\P(a_2)\\\vdots\\P(a_r)\end{bmatrix}$$

(3) 根据贝叶斯定理，可得后验概率 $P(a_i\,|\,b_j)=\dfrac{P(a_ib_j)}{P(b_j)}\quad (P(b_j)\neq 0)$

$$=\frac{P(a_i)\cdot P(b_j\,|\,a_i)}{\sum_{i=1}^{r}P(a_i)\cdot P(b_j\,|\,a_i)}$$

且 $\sum_{i=1}^{r}P(a_i\,|\,b_j)=1\ \ (i=1,2,\cdots,s)$

$\sum_{i=1}^{r}P(a_i\,|\,b_j)=1$ 说明：信道输出端接收的任一符号 b_j，必是输入符号 $a_1,a_2,\ldots,a_r$ 中的某一个送入信道的。

4.1 互信息量和平均互信息量

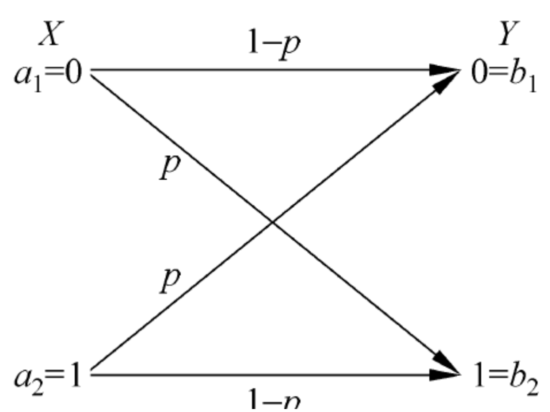
■ 二元离散信道

二元离散信道模型由一个允许输入值的集合 $X=\{0,1\}$ 和可能输出值的集合 $Y=\{0,1\}$，以及一组表示输入、输出关系的条件概率（转移概率）组成。最简单的二元离散信道是**二元对称信道**（binary symmetric channel，**BSC**）。它是一种无记忆信道。转移概率为：

$$p(b_1|a_1)=p(0|0)=1-p=\bar{p}，\quad p(b_2|a_2)=p(1|1)=1-p=\bar{p}$$

$$p(b_1|a_2)=p(0|1)=p，\quad p(b_2|a_1)=p(1|0)=p$$

$$[P]=\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$



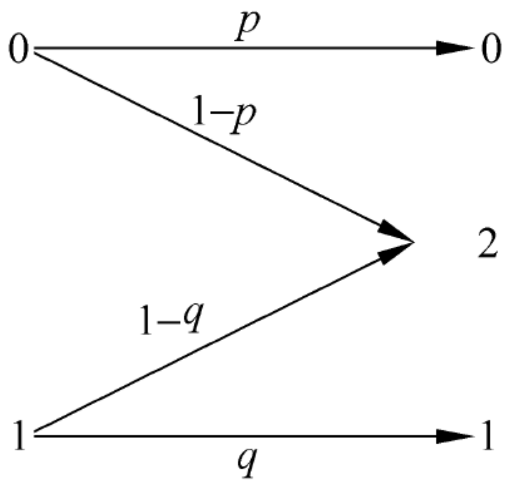
■ 二元离散信道

二元删除信道（binary erase channel，**BEC**）模型由一个允许输入值的集合 $X=\{0,1\}$ 和可能输出值的集合Y={0,1,2}。它也是一种无记忆信道。转移概率矩阵为：

$$[P]=\begin{bmatrix} p & 1-p & 0 \\ 0 & 1-q & q \end{bmatrix}$$

$\sum_{j=1}^{3} p(b_j \mid a_i)=1, i=1,2$，$p$，$q$ 表示单符号无错误传输的概率；

1-p，1-q 表示单符号传输错误的概率。



互信息量

定义：对两个离散随机事件集 X 和 Y，事件 y_i 的发生给出关于事件 x_i 的信息量，定义为互信息，其定义式为：

$$\text{互信息 } I(x_i; y_j) = \log \frac{p(x_i \mid y_j)}{p(x_i)} = \log \frac{\text{后验概率}}{\text{先验概率}}$$

注：

① 互信息量单位与自信息量单位一样，取决于对数的底。

② 由定义可知：

$$I(x_i; y_i) = \log_2 \frac{p(x_i \mid y_i)}{p(x_i)} = \log_2 \frac{1}{p(x_i)} - \log_2 \frac{1}{p(x_i \mid y_i)}$$
$$= I(x_i) - I(x_i \mid y_i)$$

如何理解 互信息量等于自信息量减去条件自信息量 ？

■ 互信息量

举例：

A预先知道他的三位朋友B、C、D中必有一人于某晚来他家，且三人来的可能性相同，其先验概率为

$$P(B)=P(C)=P(D)=1/3,$$

但这天上午A接D电话说因故不来了。若把上午接电话事件定为E，则有后验概率

$$P(D|E)=0,\ P(B|E)=P(C|E)=1/2,$$

下午A又接到C的电话，晚上有事来不了了，若把此次电话当做事件F，则后验概率

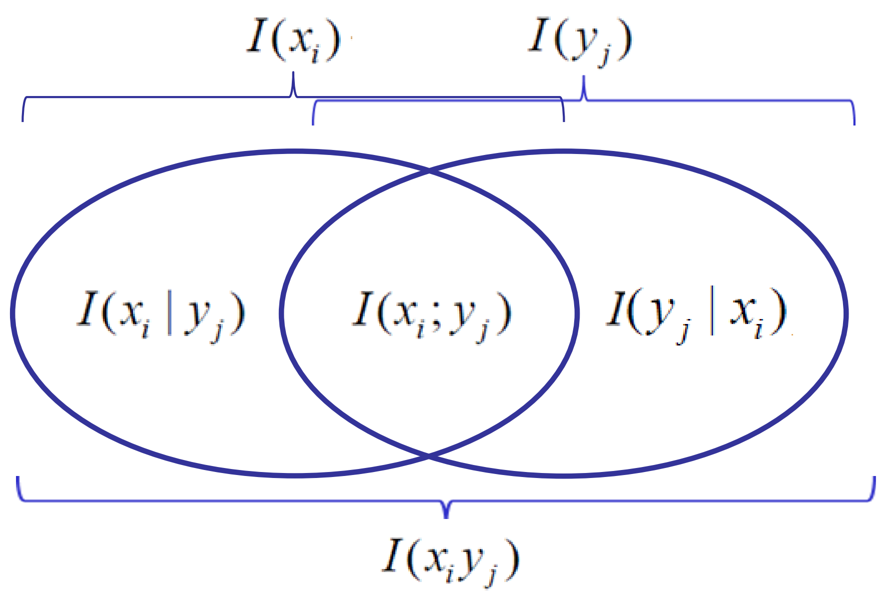
$$P(C|EF)=P(D|EF)=0,\ \text{而}P(B|EF)=1$$

互信息量

$$I(x_i;y_j) = -\log p(x_i) + \log p(x_i \mid y_j) = I(x_i) - I(x_i \mid y_j)$$

$$= \log \frac{p(x_i \mid y_j)p(y_j)}{p(x_i)p(x_j)}$$

$$= I(x_i) + I(y_j) - I(x_i y_j)$$



互信息量

例：某地二月份天气构成的信源为 $\begin{bmatrix} X \\ P(X) \end{bmatrix} = \begin{Bmatrix} a_1(\text{晴}), & a_2(\text{阴}), & a_3(\text{雨}), & a_4(\text{雪}) \\ \frac{1}{2}, & \frac{1}{4}, & \frac{1}{8}, & \frac{1}{8} \end{Bmatrix}$

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

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

依据公式，可以计算出 b_1 与各种天气之间的互信息量。

对天气 a_1，因 $p(a_1|b_1)=0$，不必再考虑 a_1 与 b_1 之间的互信息量。

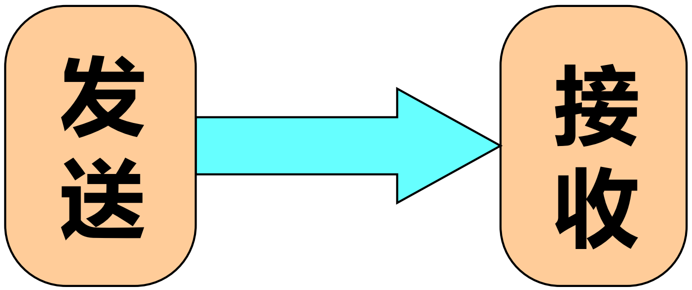
对天气 a_2，可计算出

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

4.1 互信息量和平均互信息量

- 互信息量

$$I(x_i;y_i)=\log_2\frac{p(x_i\mid y_i)}{p(x_i)}=\log_2\frac{1}{p(x_i)}-\log_2\frac{1}{p(x_i\mid y_i)}$$
$$=I(x_i)-I(x_i\mid y_i)$$



$$b_j=a_i,i=j?$$

理想情况：$I(a_i;b_j)=I(a_i)$

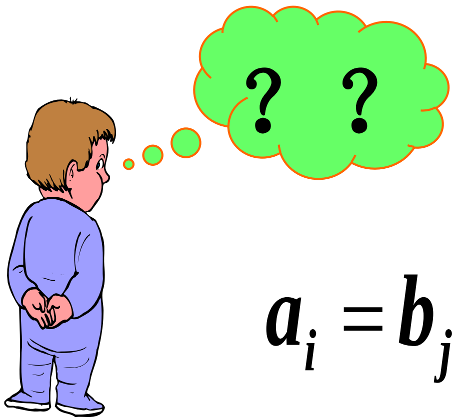
收到 b_j 后 a_i 仍有不确定性，但比原来的不确定性发生了一些变化。不确定性变化的部分，即是观察者从接收端获得的关于发送端的信息量。

4.1 互信息量和平均互信息量

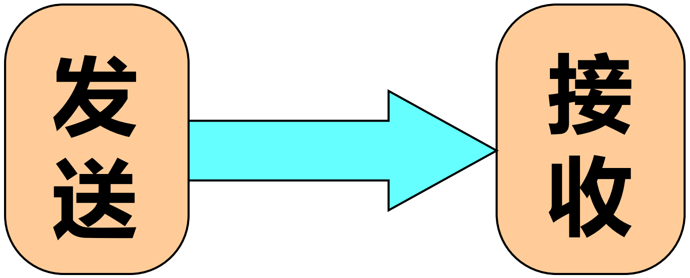
- 互信息量

$$I(x_i;y_i)=\log_2\frac{p(x_i\mid y_i)}{p(x_i)}=\log_2\frac{1}{p(x_i)}-\log_2\frac{1}{p(x_i\mid y_i)}$$

$$=I(x_i)-I(x_i\mid y_i)$$



$$a_i=b_j,i=j?$$



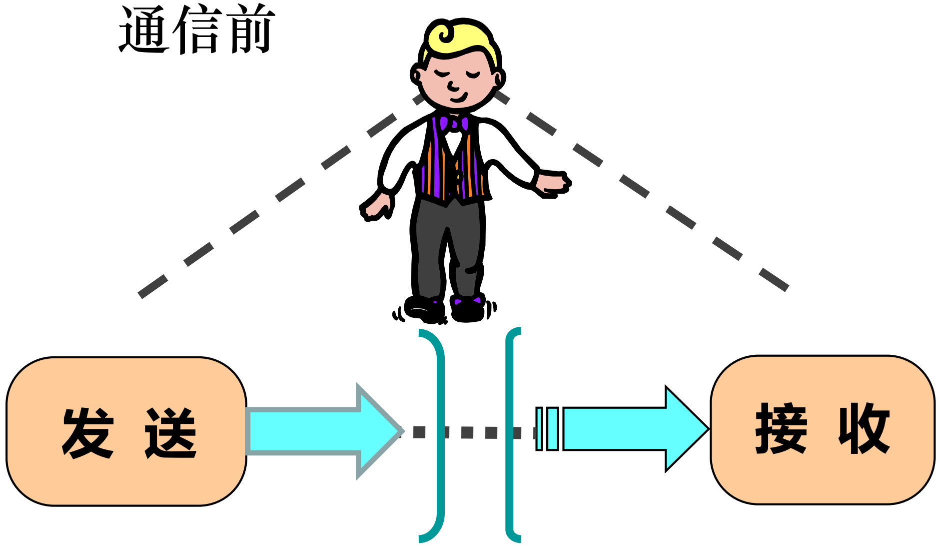
理想情况：$I(b_j;a_i)=I(b_j)=I(a_i)$

发送 a_i 后 b_j 仍有不确定性，但比原来的不确定性发生了一些变化。不确定性变化的部分，即是观察者从发送端获得的关于接收端的信息量。

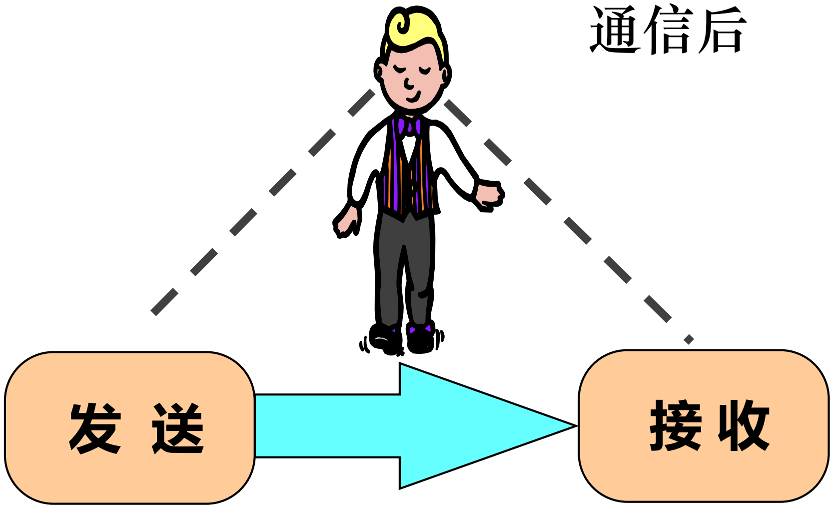
4.1 互信息量和平均互信息量

■ 互信息量

通信前



通信后



通信前的联合概率

$$\boxed{p(a_i b_j) = p(a_i)\,p(b_j)}$$

通信后的联合概率

$$\boxed{p(a_i b_j) = p(a_i)\,p(b_j/a_i)}$$

先验不定度(联合自信息量)

$$I'(a_i b_j) = \log \frac{1}{p(a_i)\,p(b_j)}$$

后验不定度

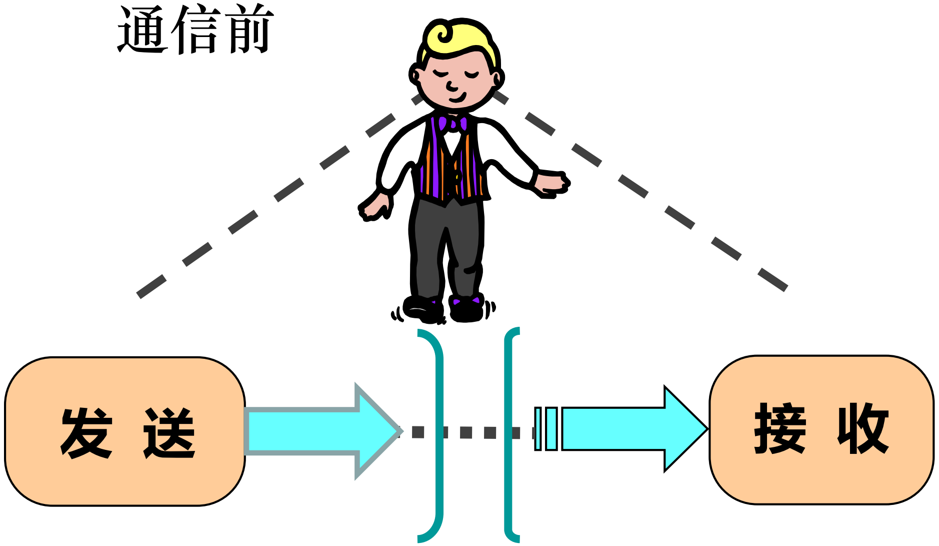
$$I''(a_i b_j) = \log \frac{1}{p(a_i b_j)}$$

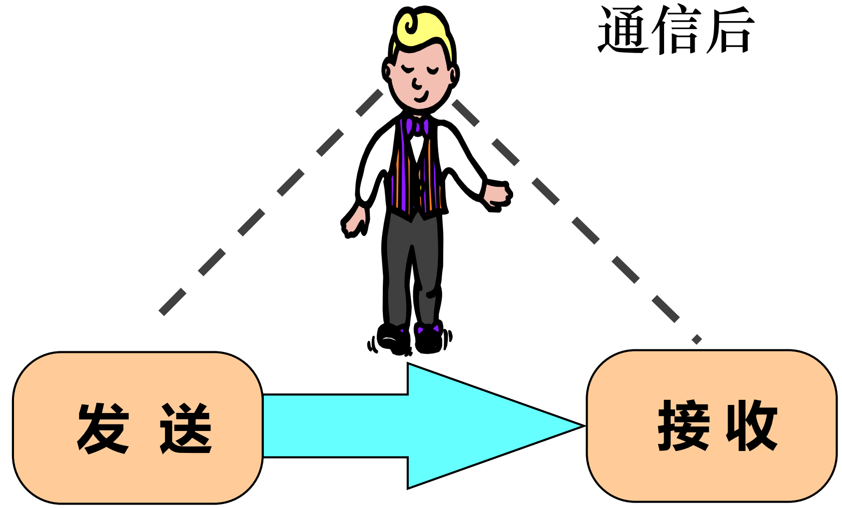
4.1 互信息量和平均互信息量

■ 互信息量

通信前

通信后





先验不定度(联合自信息量)

$$I'(a_i b_j) = \log \frac{1}{p(a_i)p(b_j)}$$

后验不定度

$$I''(a_i b_j) = \log \frac{1}{p(a_i b_j)}$$

这样，通信后流经信道的信息量，等于通信前后不定度的差。

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

■ 互信息量

互信息量的性质：

（1）对称性（互易性）

$$I(x_i;y_j)=I(y_j;x_i);\quad （即：\log\frac{p(x_i\mid y_j)p(y_j)}{p(x_i)p(y_j)}=\log\frac{p(y_j\mid x_i)p(x_i)}{p(y_j)p(x_i)}）$$

（2）互信息量可为0（当二者独立）

$$I(x_i;y_j)=I(x_i)+I(y_j)-I(x_iy_j)=I(x_i)+I(y_j)-[I(x_i)+I(y_j)]=0$$

（3）互信息量可正，可负

（4）任何两个事件之间的互信息量不可能大于其中任一事件的自信息量。

4.1 互信息量和平均互信息量

■ 互信息量

多变量间的互信息：

符号 x_i 与符号对 $y_j z_k$ 之间的互信息量定义为：$I(x_i;y_j,z_k)=\log\dfrac{p(x_i\mid y_j,z_k)}{p(x_i)}$

$I(x_i;y_j\mid z_k)$ 是在给定 z_k 条件下，x_i 与 y_j 之间的条件互信息量

$$I(x_i;y_j\mid z_k)=\log\frac{p(x_i\mid y_j,z_k)}{p(x_i\mid z_k)}$$

则有：$I(x_i;y_j,z_k)=I(x_i;z_k)+I(x_i;y_j\mid z_k)$

证明：

$$\begin{aligned}
I(x_i;y_jz_k)&=\log\frac{p(x_i\mid y_jz_k)}{p(x_i)}\\
&=\log\frac{p(x_i\mid y_jz_k)}{p(x_i)}\cdot\frac{p(x_i\mid z_k)}{p(x_i\mid z_k)}\\
&=\log\frac{p(x_i\mid y_jz_k)}{p(x_i\mid z_k)}+\log\frac{p(x_i\mid z_k)}{p(x_i)}\\
&=I(x_i;z_k)+I(x_i;y_j\mid z_k)
\end{aligned}$$

4.1 互信息量和平均互信息量

■ 互信息量

多变量间的互信息：

$$I(x_i;y_j,z_k)=I(x_i;z_k)+I(x_i;y_j\mid z_k)$$

说明：一个联合事件 y_jz_k 出现后所提供的有关 x_i 的信息量 $I(x_i;y_j,z_k)$等于 z_k 事件出现后提 供的有关 x_i 的信息量 $I(x_i;z_k)$,加上在给定 z_k 条件下再出现 y_j 事件后所提供的有关 x_i 的信息量 $I(x_i;y_j|z_k)$。

同理，还有以下公式：

$$I(x_i;y_jz_k)=I(x_i;y_j)+I(x_i;z_k\mid y_j)$$

$$I(x_i;y_j,z_k)=I(x_i;z_k,y_j)$$

■ 平均互信息量

1) 互信息量 $I(x_i;y_j)$ 在 X 集合上的统计平均值为

$$I(X;y_j)=\sum_i p(x_i|y_j)I(x_i;y_j)=\sum_i p(x_i|y_j)\log\frac{p(x_i|y_j)}{p(x_i)}$$

$I(X;y_j)$是在联合集 XY 上，由 y_j 提供的关于集 X 的平均条件互信息量，它是当接收到符号 y_j 后所能获得的关于集 X 的平均信息量，由于 y_j 只是随机变量的一个取值，因此 $I(X;y_j)$仍是随机变量。

2) 上述 $I(X;y_j)$在 Y 集合上的概率加权统计平均值:

$$I(X;Y)=\sum_j p(y_j)I(X;y_j)=\sum_{i,j} p(y_j)p(x_i|y_j)\log\frac{p(x_i|y_j)}{p(x_i)}$$

$$=\sum_{i,j} p(x_iy_j)\log\frac{p(x_i|y_j)}{p(x_i)}$$

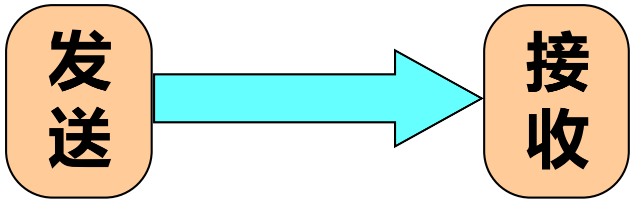
4.1 互信息量和平均互信息量

■ 平均互信息量

Y 对X 的***平均互信息量***，也称平均交互信息量、交互熵。

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

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



$$\boldsymbol{b_j = a_i, i = j}\ ?;$$

$$\boldsymbol{i} = 1,\cdots,\boldsymbol{n};\ \boldsymbol{j} = 1,\cdots,\boldsymbol{m}$$

a. 在通信系统中，若发端的符号是 X，而收端的符号是 Y， I(X;Y)就是**在接收端收到 Y 后所能获得的关于 X 的信息**。

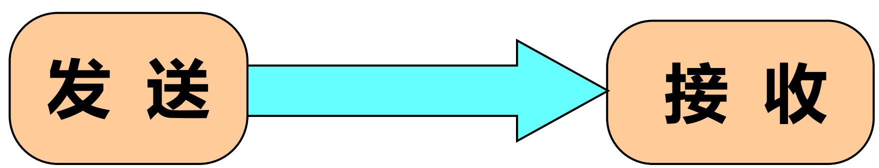
b. 若干扰很大， Y 基本上与 X 无关，或说 X 与 Y 相互独立，那时就收不到任何关于X的信息。

c. 若没有干扰， Y 是 X 的确知一一对应函数，那就能完全收到 X 的信息 $H(X)$。

- **平均互信息量**

  同理，X 对 Y 的平均互信息量

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

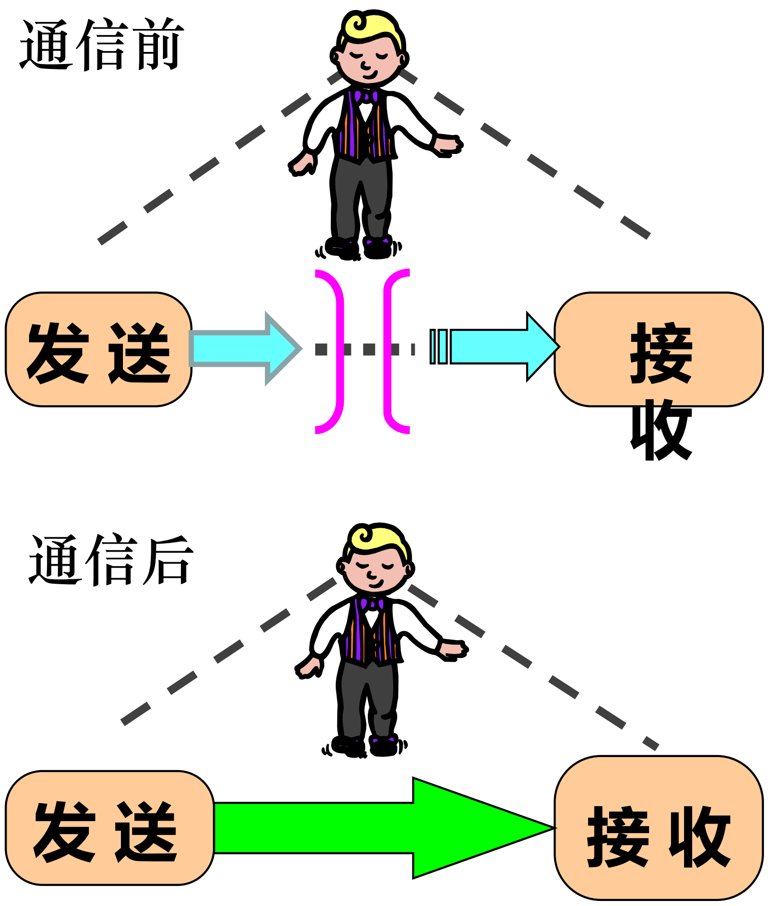


$$a_i=b_j, i=j?\ i=1,\cdots,n; j=1,\cdots,m$$

平均互信息量

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

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



信道中流通信息量的整体测度。

■ 平均互信息量

平均互信息量的物理意义

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

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

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

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。

4.1 互信息量和平均互信息量

- 平均互信息量

③ $$I(Y;X)=\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\left[\log p(a_ib_j)-\log p(a_i)-\log p(b_j)\right]$$
$$=H(X)+H(Y)-H(XY)$$

通信前: $H(XY)=H(X)+H(Y)$　　通信后: $H(XY)=H(X)+H(Y/X)$

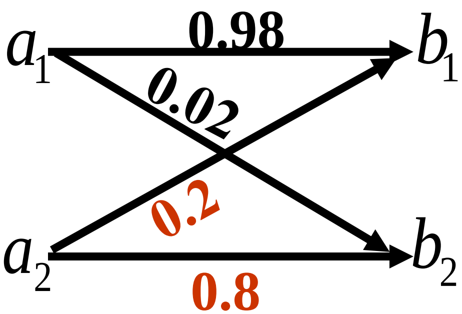
平均互信息量等于通信前、后，整个系统不确定度减少的量。

信息就是负熵——从一个事件获得另一个事件的平均互信息需要消除不确定度，一旦消除了不确定度，就获得了信息。

4.1 互信息量和平均互信息量

■ 平均互信息量

例：信源 X 接入图示信道



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

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

1

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

$$p(a_1b_1)=p(a_1)p(b_1/a_1)=0.5\times0.98=0.49$$

同理

$$p(a_1b_2)=0.5\times0.02=0.01,\ p(a_2b_1)=0.5\times0.2=0.1,\ p(a_2b_2)=0.5\times0.8=0.4$$

4.1 互信息量和平均互信息量

2

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

$$p(b_1)=p(a_1b_1)+p(a_2b_1)=0.1+0.49=0.59$$

$$p(b_2)=p(a_1b_2)+p(a_2b_2)=0.01+0.4=0.41$$

3

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

$$p(a_1/b_1)=\frac{p(a_1b_1)}{p(b_1)}=\frac{0.49}{0.59}=0.831,\ \ p(a_2/b_1)=1-p(a_1/b_1)=0.169$$

$$p(a_1/b_2)=\frac{p(a_1b_2)}{p(b_2)}=\frac{0.01}{0.41}=0.024,\ \ p(a_2/b_2)=1-p(a_1/b_2)=0.976$$

4.1 互信息量和平均互信息量

4

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

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

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

等概率信源的熵最大。

5

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

$$= -0.49\log 0.831 - 0.01\log 0.024 - 0.1\log 0.169 - 0.4\log 0.976$$

$$= H(XY) - H(Y) = 0.45(bit/sign)$$

6 $I(X;Y)=H(X)-H(X/Y)=1-0.45=0.55(bit/sign)$

7 $H(Y/X)=H(XY)-H(X)=1.43-1=0.43(bit/sign)$

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。 $I(X;Y)\ge 0$

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

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

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

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

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

4.1 互信息量和平均互信息量

1 X、Y一一对应。

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

2 X、Y 相互独立， $p(a_i/b_j)=p(a_i)$

即

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

$$=-\sum_{i=1}^{n}p(a_i)\log p(a_i)+\sum_{i=1}^{n}\sum_{j=1}^{m}p(a_ib_j)\log p(a_i)$$

$$==H(X)-H(X)=0$$

4 凸函数性

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

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

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

$$\boxed{I(X;Y)=f\left[p(a),p(b/a)\right]}$$

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

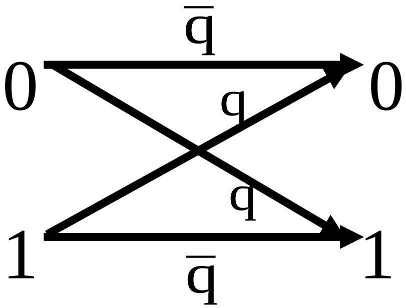
1 $I(X;Y)$是信源分布$\{p(a)\}$ 的 **上凸函数**

$$I\left[\alpha p_1(a)+(1-\alpha)p_2(a)\right]\ge \alpha I\left[p_1(a)\right]+(1-\alpha)I\left[p_2(a)\right]$$

$$0<\alpha<1$$

例：二元信源X接入对称信道，求平均互信息量 $I(X;Y)$。

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



4.1 互信息量和平均互信息量

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

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

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

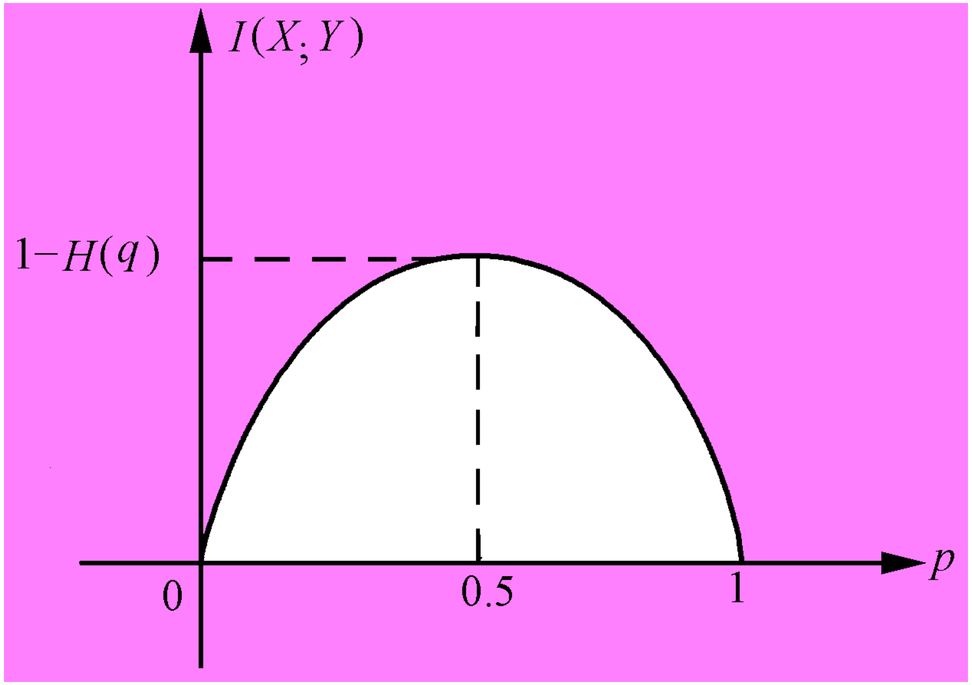
$$p(b_j) = \sum_{i=1}^{2} p(a_i)p(b_j/a_i) \quad p(b_1) = p\bar{q} + \bar{p}q \quad p(b_2) = pq + \bar{p}\bar{q}$$

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

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

4.1 互信息量和平均互信息量

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



$I(X;Y)$随信源变化的曲线

下凸函数

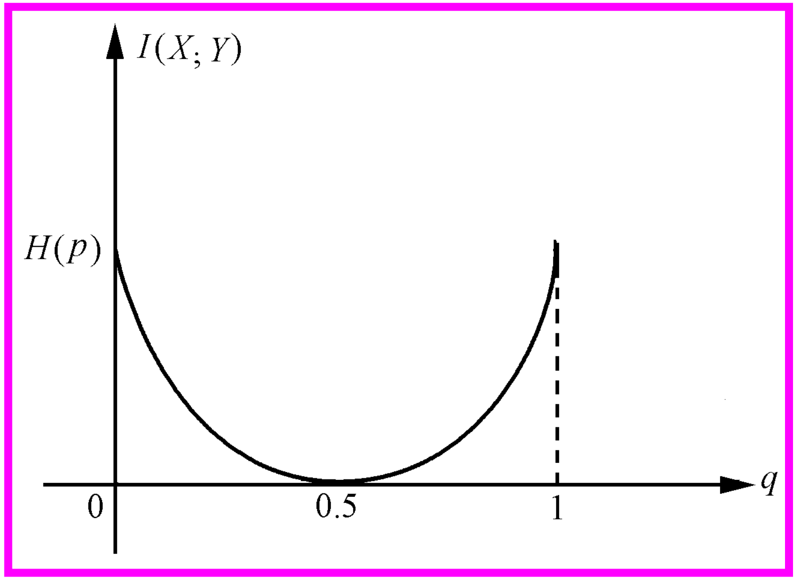
2 $I(X;Y)$是信道传递概率$\{p(b_j/a_i)\}$的

令 $0<\alpha<1$

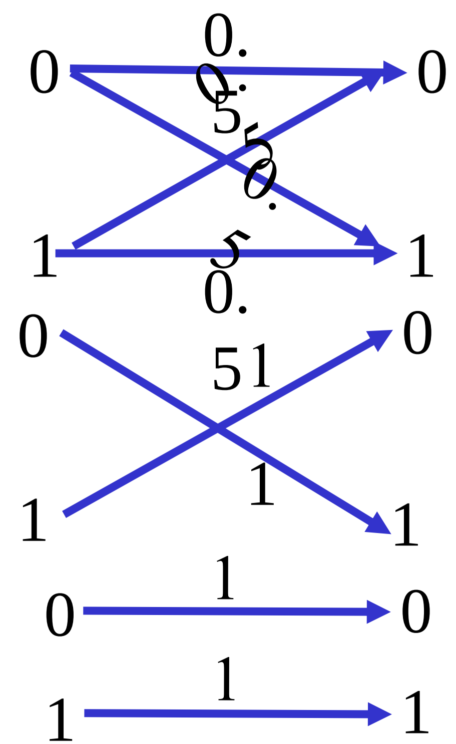
$$p_3(b_j/a_i)=\alpha p_1(b_j/a_i)+(1-\alpha)p_2(b_j/a_i)$$

$$I\left[p_3(b_j/a_i)\right]=I\left[\alpha p_1(b_j/a_i)+(1-\alpha)p_2(b_j/a_i)\right]$$
$$\leq \alpha I\left[p_1(b_j/a_i)\right]+(1-\alpha)\left[p_2(b_j/a_i)\right]$$

4.1 互信息量和平均互信息量



$I(X;Y)$随信道变化的曲线



5 数据处理定理

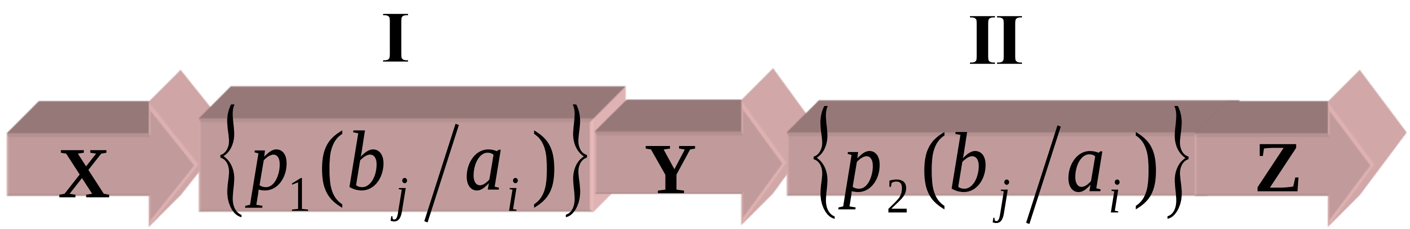


图4.1.7 数据处理模型

假定Y条件下X、Z相互独立

$$I(X;Z) \le I(Y;Z) \quad (4.1.36)$$

$$I(X;Z) \le I(X;Y) \quad (4.1.37)$$

多次处理信息量将减少

4.1 互信息量和平均互信息量

$$I(X;YZ)=\sum_i\sum_j\sum_k p(a_ib_jc_k)\log\frac{p(a_i/b_jc_k)}{p(a_i)}$$

定义

$$=\sum_i\sum_j\sum_k p(a_ib_jc_k)\log p(a_i/b_jc_k)-\sum_i\sum_j\sum_k p(a_ib_jc_k)\log p(a_i)$$

$$=H(X)-H(X/YZ)$$

$$I(X;Y/Z)=\sum_i\sum_j\sum_k p(a_ib_jc_k)\log\frac{p(a_i/b_jc_k)}{p(a_i/c_k)}$$

$$I(X;Z)=I(X;YZ)-I(X;Y/Z) \quad I(X;Y)=I(X;YZ)-I(X;Z/Y)$$

$\boxed{I(X;Z)\le I(X;Y)}$ 同理: $\boxed{I(X;Z)\le I(Y;Z)}$

4.1 互信息量和平均互信息量

多次测量

$$I(X;Y_1)=H(X)-H(X/Y_1)$$

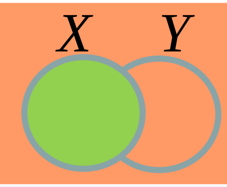
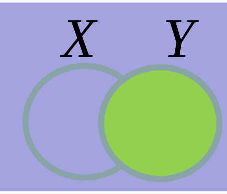
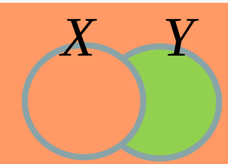
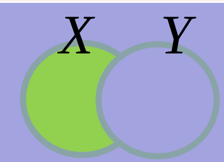
$$I(X;Y_1Y_2)=H(X)-H(X/Y_1Y_2)$$

多次测量的互信息量要比单次测量的互信息量大。

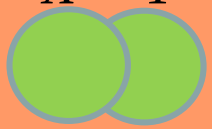
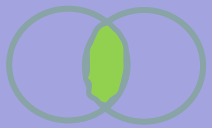
$$I(X;Y_1Y_2)\geq I(X;Y_1)$$

4.1 互信息量和平均互信息量

- 各种熵之间的关系

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

■ 平均互信息量

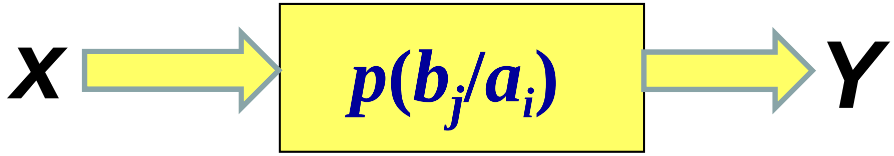
名称	符 号	关 系 式	图 示
联合熵	$H(XY)$	$H(XY)=H(X)+H(Y/X)=H(Y)+H(X/Y)$ $=H(X)+H(Y)-I(X;Y)$ $=H(X/Y)+H(Y/X)+I(X;Y)$	X Y 
交互熵	$I(X;Y)$ $=I(Y;X)$	$I(X;Y)=H(X)-H(X/Y)=H(Y)-H(Y/X)$ $=H(XY)-H(Y/X)-H(X/Y)$ $=H(X)+H(Y)-H(XY)$	X Y 

4.2 单符号离散信道的信道容量

■ 单符号离散信道容量的定义

设单符号离散信道输入空间 $\boldsymbol{X}=\{a_1, a_2, \ldots, a_r\}$，相应的输出空间$\boldsymbol{Y}=\{b_1, b_2, \ldots, b_s\}$，信道特性可用转移概率来描述

$$P=\begin{bmatrix} P(b_1|a_1) & P(b_2|a_1) & \cdots & P(b_s|a_1) \\ P(b_1|a_2) & P(b_2|a_2) & \cdots & P(b_s|a_2) \\ \vdots & \vdots & \ddots & \vdots \\ P(b_1|a_r) & P(b_2|a_r) & \cdots & P(b_s|a_r) \end{bmatrix}$$



$$i=1,2,\cdots,r$$
$$j=1,2,\cdots,s$$

信道的数学模型可表示为: $\{X,P,Y\}$

信道传输**率**（信息率）$R=I(X;Y)=H(X)-H(X|Y)$

信息传输**速率** $R_t=\frac{1}{t}I(X;Y)$

■ 单符号离散信道容量的定义

信道传输率（信息率）：$R=I(X;Y)=H(X)-H(X|Y)=f[p(a),p(b|a)]$

当信道特性 $p(b|a)$ 固定后，总能找到一种信源概率分布 $p(a)$，使得信道所能传输的信息率最大。

信道容量，定义为这个最大的信息传输率，即

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

信道的最大信息传输速率 $C_t=\dfrac{1}{t}\max\limits_{p(a)} I(X;Y)$

4.2 单符号离散信道的信道容量

- 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信道

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

$$\begin{bmatrix} 100\ldots\ldots0 \\ 010\ldots\ldots0 \\ \ldots\ldots \\ 000\ldots\ldots1 \end{bmatrix}$$

a_1 —— b_1

a_2 —— b_2

……

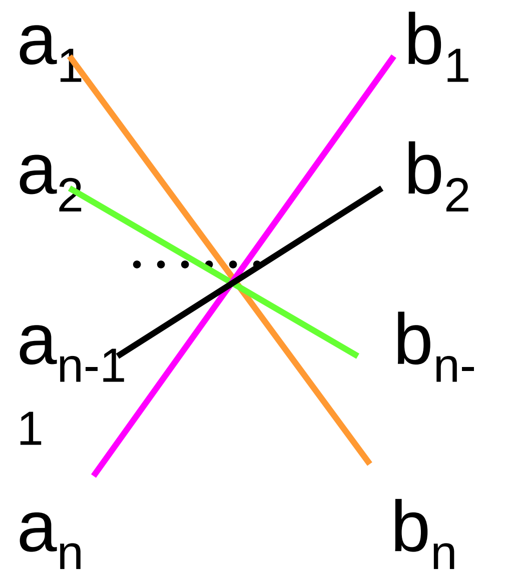
a_n —— b_n

图4.2.2 一一对应的无噪信道

4.2 单符号离散信道的信道容量

■ 几种特殊离散信道的信道容量

$$\begin{bmatrix} 000\ldots\ldots01 \\ 000\ldots\ldots10 \\ \ldots\ldots \\ 010\ldots\ldots00 \\ 100\ldots\ldots00 \end{bmatrix}$$



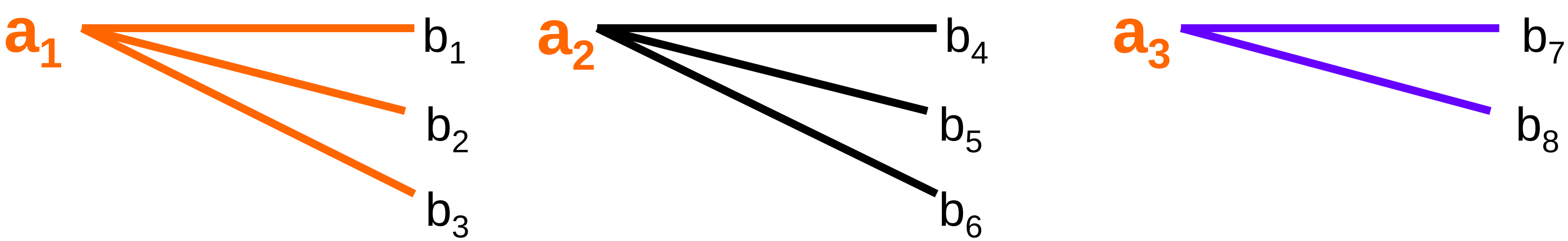
一一对应的无噪信道

$$H(X|Y)=0$$

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

■ 几种特殊离散信道的信道容量

② 具有扩展性能的无噪信道



一个输入对应多个输出

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

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

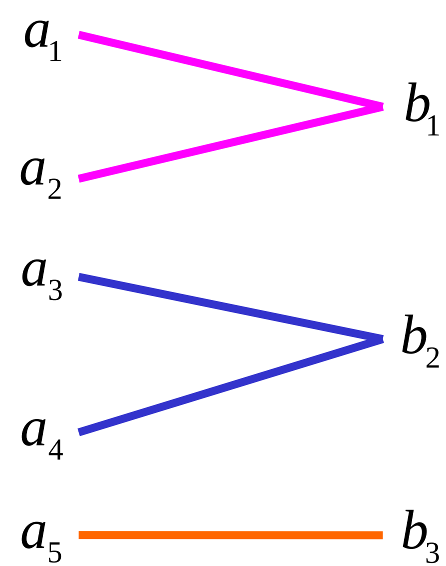
$$\boxed{C=\max_{p(a_i)} H(X)=\log n}$$

4.2 单符号离散信道的信道容量

■ 几种特殊离散信道的信道容量

③ 具有归并性能的无噪信道

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



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

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

多个输入变成一个输出

几种特殊离散信道的信道容量

2 强对称离散信道的信道容量

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

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

p：总体错误概率

■ 几种特殊离散信道的信道容量

$$
\begin{aligned}
H(Y/X) &= -\sum_{i=1}^{n}\sum_{j=1}^{n} p(a_i)p(b_j/a_i)\log p(b_j/a_i) = -\sum_{i=1}^{n} p(a_i)\sum_{j=1}^{n} p(b_j/a_i)\log p(b_j/a_i) \\
&= -\sum_{i=1}^{n} p(a_i)[(1-p)\log(1-p) + (\frac{p}{n-1}\log\frac{p}{n-1})(n-1)] \\
&= -\sum_{i=1}^{n} p(a_i)[(1-p)\log(1-p) + p\log\frac{p}{n-1}] \\
&= -[(1-p)\log(1-p) + p\log\frac{p}{n-1}] = H_{ni}
\end{aligned}
$$

$$
\begin{aligned}
C &= \max_{p(a_i)} I(X;Y) = \max_{p(a_i)}[H(Y) - H(Y/X)] \\
&= \max_{p(a_i)}[H(Y) - H_{ni}] = \log n - H_{ni}
\end{aligned}
$$

相应的 $p(a_i) = \dfrac{1}{n}$

4.2 单符号离散信道的信道容量

■ 几种特殊离散信道的信道容量

二进制均匀信道容量 $\boxed{C=1-H(p)}$,

其中 $H(p)=-((1-p)\log(1-p)+p\log p)$

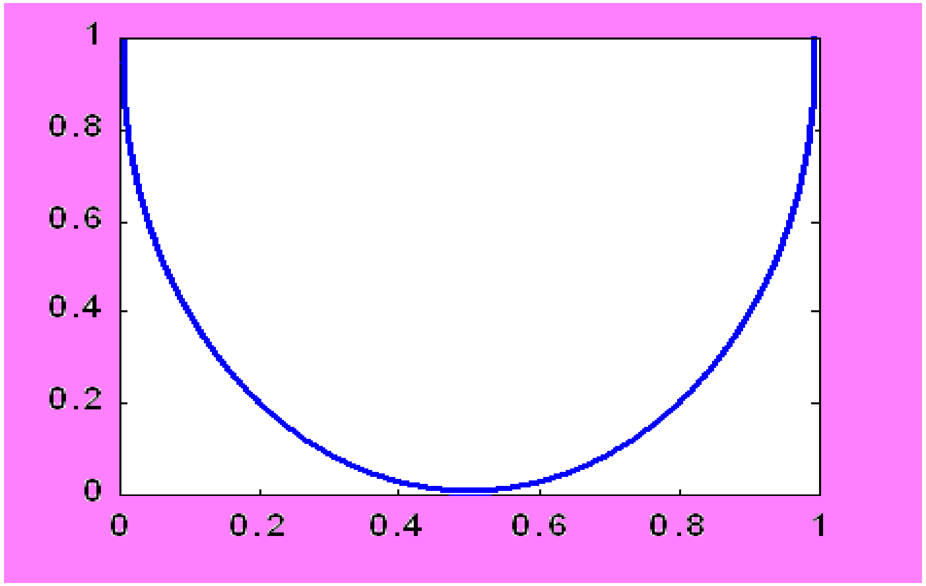


图4.2.5 二进制均匀信道容量曲线

4.2 单符号离散信道的信道容量

几种特殊离散信道的信道容量

3 对称离散信道的信道容量

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

矩阵中的每列都是集合$Q=\{q_1, q_2, \cdots\cdots, q_m\}$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，$P=Q$。

4.2 单符号离散信道的信道容量

- 几种特殊离散信道的信道容量

练习：判断下列矩阵表示的信道是否是对称信道？

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

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

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

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

4.2 单符号离散信道的信道容量

- 几种特殊离散信道的信道容量

对称离散信道的信道容量

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

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

$$=H_{mi}$$

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

相应的 $p(a_i)=\dfrac{1}{n}$

4.2 单符号离散信道的信道容量

- 几种特殊离散信道的信道容量

强对称信道与对称信道比较

强对称	对称
n=m	**n**与**m**未必相等
矩阵对称	矩阵未必对称
P=Q	**P**与**Q**未必相等
行之和，列之和均为**1**	行之和为**1**

■ 几种特殊离散信道的信道容量

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。例如下面的矩阵：

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8} \end{bmatrix}$$

$$H(Y/X) = H_{mi}$$

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

■ 几种特殊离散信道的信道容量

假设此时将矩阵的列分为S个子集，每个子集的元素个数分别是m_1，m_2，……，m_s。

$$H(Y)=-\sum_{j}^{m}p(b_j)\log p(b_j)=-\sum_{j_1}^{m_1}p(b_{j_1})\log p(b_{j_1})-\cdots-\sum_{j_s}^{m_s}p(b_{j_s})\log p(b_{j_s})$$

例4.2.1

信道矩阵 $[P]=\begin{bmatrix}0.5 & 0.25 & 0.125 & 0.125\\ 0.25 & 0.5 & 0.125 & 0.125\end{bmatrix}$

分成子矩阵 $[P_1]=\begin{bmatrix}0.5 & 0.25\\ 0.25 & 0.5\end{bmatrix}$ $[P_2]=\begin{bmatrix}0.125 & 0.125\\ 0.125 & 0.125\end{bmatrix}$

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

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

$$\boxed{C=0.0612(bit/sign)}$$

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。采用拉各朗日乘子法来计算。

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

$$\frac{\partial}{\partial p(a_i)}\left\{-\sum_{j=1}^{m} p(b_j)\log p(b_j) + \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j/a_i)\log p(b_j/a_i) - \lambda\left[\sum_{i=1}^{n} p(a_i) - 1\right]\right\} = 0$$

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

■ 离散信道容量的一般计算方法

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

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

$$\sum_{j=1}^m p(b_j/a_i)\log p(b_j/a_i) - \sum_{j=1}^m p(b_j/a_i)\log p(b_j) = \log e + \lambda \qquad (4.2.23)$$

两边乘$p(a_i)$，并求和有：

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

(4.2.24)

4.2 单符号离散信道的信道容量

- 离散信道容量的一般计算方法

$$I(X;Y)=\log_2 e+\lambda \quad \Longleftrightarrow \quad C=\log_2 e+\lambda \qquad (4.2.25)$$

将(4.2.25)代入(4.2.23)，则有：

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

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

令
$$\beta_j=\log p(b_j)+C \qquad (4.2.26)$$

则
$$\sum_{j}^{m} p(b_j/a_i)\log p(b_j/a_i)=\sum_{j}^{m} p(b_j/a_i)\beta_j \qquad (4.2.27)$$

4.2 单符号离散信道的信道容量

■ 离散信道容量的一般计算方法

由(4.2.27)求出 β_j，再由(4.2.26)求出 $p(b_j)$。

$$p(b_j) = 2^{\beta_j - C} \qquad (4.2.28)$$

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

$$\boxed{\boldsymbol{c = \log_2 \sum_{j=1}^{m} 2^{\beta_j}}} \qquad (4.2.29)$$

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

。

- 离散信道容量的一般计算方法

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 $c=\log_2\sum_{j=1}^{m} 2^{\beta_j}$ 求C；

3 由$p(b_j)=2^{\beta_j-C}$ 求 $p(b_j)$；

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

离散信道容量的一般计算方法

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

4.2 单符号离散信道的信道容量

- 离散信道容量的一般计算方法

3

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

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

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

4

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

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

4.2 单符号离散信道的信道容量

■ 离散信道容量的一般计算方法

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

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

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

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

- 多符号离散信道容量的数学模型

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

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

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

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

$$i_1, i_2, \cdots, i_N = 1, 2, \cdots, n$$

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

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

$$\beta_j = \{b_{j_1} b_{j_2} \cdots b_{j_N}\}$$

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

$$j_1, j_2, \cdots, j_N = 1, 2, \cdots, m$$

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

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

4.3 多符号离散信道的信道容量

■ 多符号离散信道容量的数学模型

$$\begin{bmatrix} p(\beta_1/\alpha_1) & p(\beta_2/\alpha_1) & \cdots & p(\beta_{m^N}/\alpha_1) \\ p(\beta_1/\alpha_2) & p(\beta_2/\alpha_2) & \cdots & p(\beta_{m^N}/\alpha_2) \\ & \cdots & \cdots & \\ p(\beta_1/\alpha_{n^N}) & p(\beta_2/\alpha_{n^N}) & \cdots & p(\beta_{m^N}/\alpha_{n^N}) \end{bmatrix} \tag{4.3.6}$$

4.3 多符号离散信道的信道容量

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

$$\boldsymbol{I}(\boldsymbol{X};\boldsymbol{Y}) = \sum_{i=1}^{n^N}\sum_{j=1}^{m^M} \boldsymbol{p}(\alpha_i\beta_j)\boldsymbol{I}(\alpha_i;\beta_j) = \sum_{i=1}^{n^N}\sum_{j=1}^{m^M} \boldsymbol{p}(\alpha_i\beta_j)\frac{\boldsymbol{p}(\alpha_i/\beta_j)}{\boldsymbol{p}(\alpha_i)} \quad (4.3.7)$$

$$\boldsymbol{I}(\boldsymbol{X};\boldsymbol{Y}) = \sum_{i_1=1}^{n}\cdots\sum_{i_N=1}^{n}\sum_{j_1=1}^{m}\cdots\sum_{j_M}^{m} p(a_{i_1}\cdots a_{i_N}b_{j_1}\cdots b_{j_M})\log_2\frac{p(a_{i_1}\cdots a_{i_N}/b_{j_1}\cdots b_{j_M})}{p(a_{i_1}\cdots a_{i_N})} \quad (4.3.8)$$

定义多符号离散信道容量为

$$\boxed{\boldsymbol{C} = \max_{\{\boldsymbol{p}(\alpha_i)\}}\boldsymbol{I}(\boldsymbol{X};\boldsymbol{Y}),\ \ \boldsymbol{i}=1,2,\cdots,\boldsymbol{n}^N} \quad (4.3.11)$$

■ 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

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

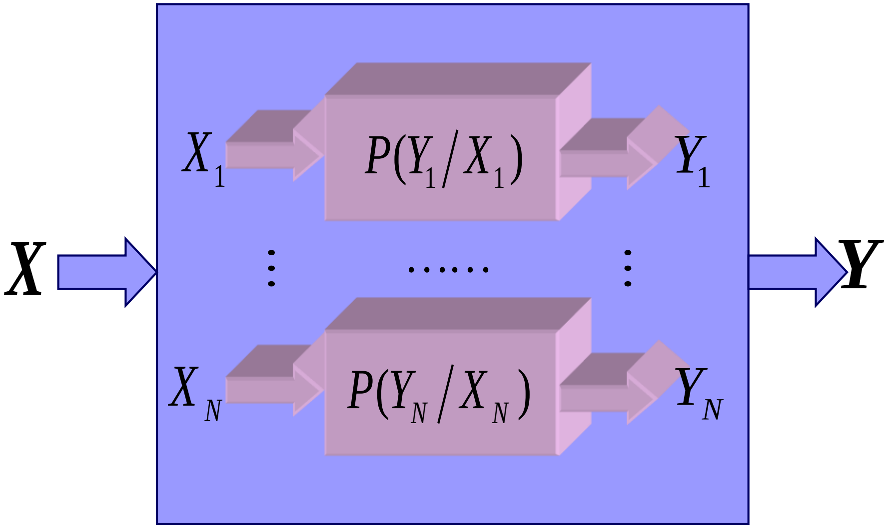


图4.3.2　单符号离散信道的N次扩展信道的实现模型

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

4.3 多符号离散信道的信道容量

■ 多符号离散信道容量定义

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

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

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

$$= H(Y_1/X_1)+H(Y_2/X_2)+\ldots+H(Y_K/X_K)$$

$$= \sum_{K=1}^{N} H(Y_K/X_K) \tag{4.3.13}$$

$$\boldsymbol{I}(\boldsymbol{X};\boldsymbol{Y}) = \boldsymbol{H}(\boldsymbol{Y})-\boldsymbol{H}(\boldsymbol{Y}/\boldsymbol{X}) = H(Y_1Y_2\cdots Y_N)-\sum_{k=1}^{N} H(Y_k/X_k) \tag{4.3.14}$$

4.3 多符号离散信道的信道容量

- 多符号离散信道容量定义

$$H(Y_1Y_2\cdots Y_N)\le\sum_{k=1}^{N}H(Y_k) \qquad (4.3.17)$$

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

如果输出端$Y_1Y_2\cdots Y_N$相互独立，式(4.3.17)等号成立。若输入端$X_1X_2\cdots X_N$也是无记忆的，则有

$$\boldsymbol{I}(\boldsymbol{X};\boldsymbol{Y})=NI(X;Y) \qquad (4.3.24)$$

离散无记忆信道N次扩展信道的信道容量为

$$C^N=NC \qquad (4.3.25)$$

4.3 多符号离散信道的信道容量

■ 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每个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

$$C^N \le C_1 + C_2 + \cdots + C_N = \sum_{k=1}^{N} C_k \qquad (4.3.26)$$

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，对独立并联信道容量最大

$$C_{\max}^N = \sum_{k=1}^{N} C_k \qquad (4.3.27)$$