2022-2023学年秋季学期

信息论与编码
第四讲

■ 马尔可夫信源

在很多信源的输出序列中，符号之间的依赖关系是有限的，任何时刻信源符号发生的概率只与前边已经发出的若干个符号有关，而与更前面的符号无关。

为了描述这类信源除了信源符号集外还要引入状态集。这时，信源输出消息符号还与信源所处的状态有关。

若一个信源满足下面两个条件，则称为马尔可夫信源：

（1）某一时刻信源输出的符号的概率只与当前所处的状态有关，而与以前的状态无关；

（2）信源的当前状态由当前输出符号和前一时刻信源状态唯一确定。

所谓“状态”，指与当前输出符号有关的前m个随机变量序列$(X_1X_2\ldots X_m)$的某一具体消息，用s_i表示，把这个具体消息看作是某个状态。

m时刻的状态

$$s_i = \{a_{k_1}, a_{k_2}, \ldots, a_{k_m}\}$$

$$a_{k_1}, a_{k_2}, \ldots, a_{k_m} \in \{a_1, a_2, \ldots, a_n\}$$

$$k_1, k_1, \ldots, k_m = 1, 2, \ldots, n$$

$$i = 1, 2, \ldots, n^m$$

当信源在(m+1)时刻发出符号$a_{k_{m+1}}$时，我们可把s_j看成另一种状态：

m+1时刻的状态

$$s_j = s_{i+1} = \{a_{k_2}, \ldots, a_{k_m}, a_{k_{m+1}}\}$$

$$a_{k_2}, \ldots, a_{k_m}, a_{k_{m+1}} \in \{a_1, a_2, \ldots, a_n\}$$

$$k_2, \ldots, k_m, k_{m+1} = 1, 2, \ldots, n$$

$$i = 1, 2, \ldots, n^m$$

➢ 某一时刻信源输出的符号的概率只与当前所处的状态有关，而与以前的状态无关。或者说只与此前已输出的若干个符号有关。即 $p(a_{k_{m+1}} \mid s_i)$

$$\ldots, X_1, X_2, \ldots, X_{m-1}, X_m, X_{m+1}, \ldots$$

s_i s_j

把此前已输出的符号视为状态

➢ 信源的当前状态由当前输出符号和前一时刻信源状态唯一确定。

该条件表明，若信源处于某一状态 s_i ，当它发出一个符号后，所处的状态就变了，一定转移到另一状态。状态的转移依赖于发出的信源符号，因此任何时刻信源处在什么状态完全由前一时刻的状态和当前发出的符号决定。将信源输出符号的不确定性问题变换为讨论信源状态转换问题。状态之间的一步转移概率为:

$$p(s_j|s_i)$$

由上可知，m阶马尔可夫信源符号集共有n个符号，则信源共有 n^m 个不同状态。信源在某一时刻时，必然处于某一种状态，等到下一个符号输出时，转移到另外一个状态。

从而得到马尔可夫信源的状态空间。

$$\begin{bmatrix} s_1 & s_2 & \cdots & s_{n^m} \\ p(s_1) & p(s_2) & \cdots & p(s_{n^m}) \end{bmatrix}$$

其状态转移图如下页。在状态转换图中，把信源的每一种状态用圆圈表示，用有向箭头表示信源发出某一符号后由一种状态到另一状态的转移。

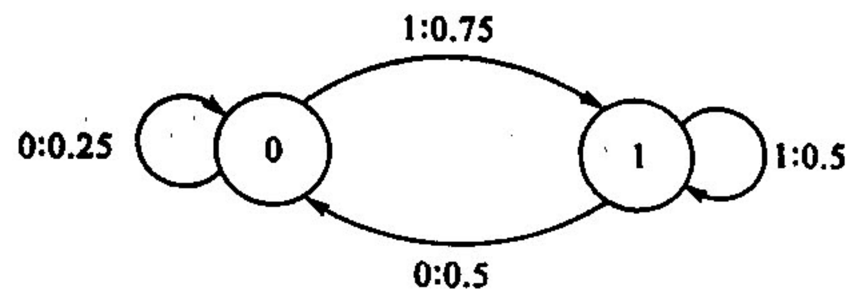
例：设有一二进制一阶马尔可夫信源，信源符号集为{1,0}，

条件概率定为：$P(0|0)=0.25$

$P(0|1)=0.50$

$P(1|0)=0.75$

$P(1|1)=0.50$



由于信源符号数n=2，因此二进制一阶信源仅有两个状态：

$S_0=0, S_1=1.$

由条件概率求得信源的状态转移概率为：

$P(S_1|S_1)=0.25$；$P(S_1|S_2)=0.50$；

$P(S_2|S_1)=0.75$；$P(S_2|S_2)=0.50.$

例、设信源符号 $X\in\{x_1, x_2, x_3\}$，信源所处的状态 $S\in\{e_1, e_2, e_3, e_4, e_5\}$。

各状态之间的转移情况由下图给出。

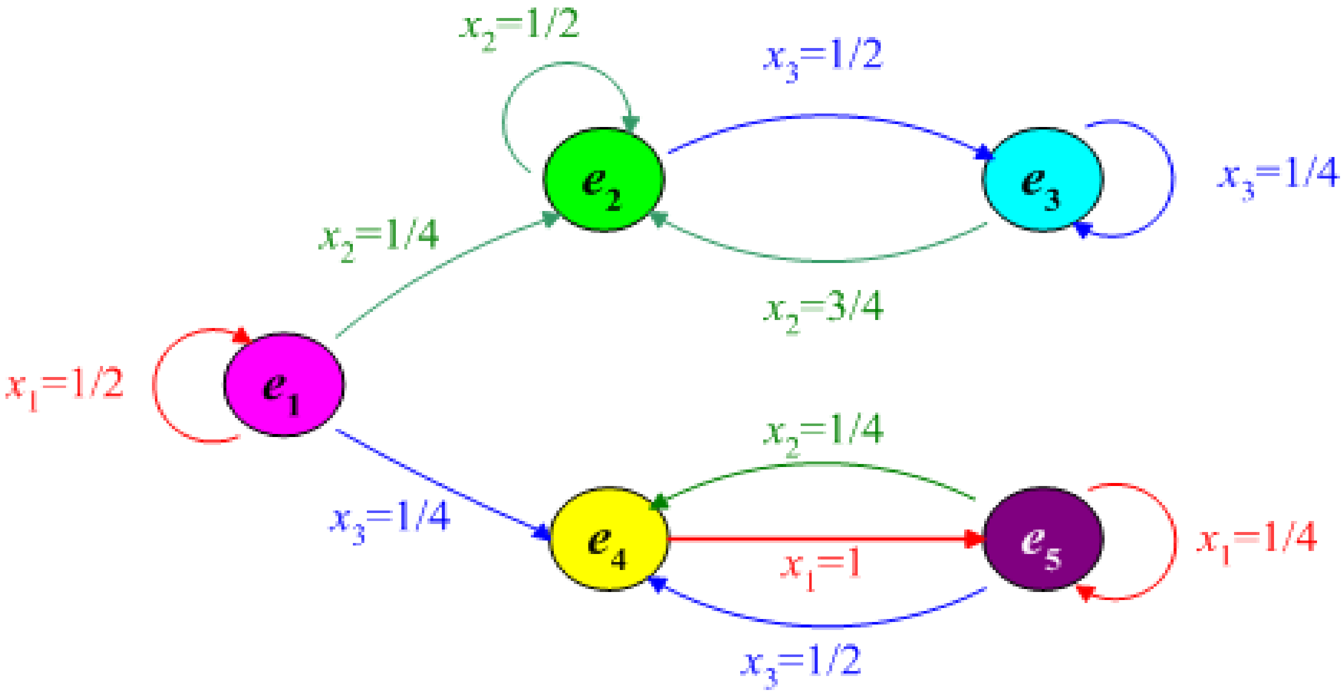


图2.2.1 状态转移图

将图中信源在e_i状态下发符号x_k 的条件概率$p(x_k /e_i)$用矩阵表示

可以看出 $\sum_{k=1}^{3} p(x_k \mid e_i) = 1, \quad i = 1,2,3,4,5$

$$
\begin{array}{c|ccc}
 & x_1 & x_2 & x_3 \\
\hline
e_1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\
e_2 & 0 & \frac{1}{2} & \frac{1}{2} \\
e_3 & 0 & \frac{3}{4} & \frac{1}{4} \\
e_4 & 1 & 0 & 0 \\
e_5 & \frac{1}{4} & \frac{1}{4} & \frac{1}{2}
\end{array}
$$

$$
\begin{cases}
P(S_l = e_2 \mid X_l = x_1, S_{l-1} = e_1) = 0 \\
P(S_l = e_1 \mid X_l = x_1, S_{l-1} = e_1) = 1 \\
P(S_l = e_2 \mid X_l = x_2, S_{l-1} = e_1) = 1 \\
P(S_l = e_3 \mid X_l = x_2, S_{l-1} = e_1) = 0 \\
\cdots
\end{cases}
$$

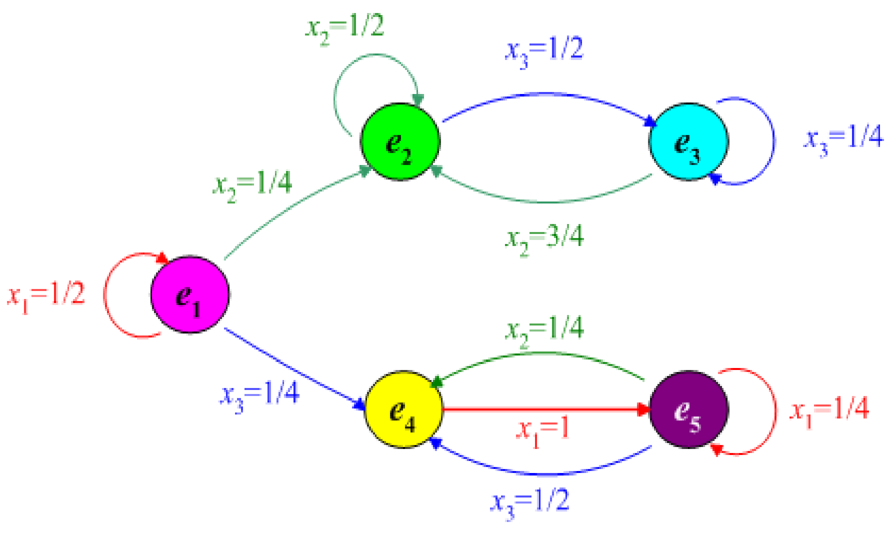


图2.2.1 状态转移图

$$
\begin{array}{c|ccccc}
 & e_1 & e_2 & e_3 & e_4 & e_5 \\
\hline
e_1 & \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\
e_2 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
e_3 & 0 & \frac{3}{4} & \frac{1}{4} & 0 & 0 \\
e_4 & 0 & 0 & 0 & 0 & 0 \\
e_5 & 0 & 0 & 0 & \frac{3}{4} & \frac{1}{4}
\end{array}
$$

- **M阶马尔可夫信源的平均信息量，即，信源的极限熵**

$$H_{\infty} = \lim_{N \to \infty} H(X_N \mid X_1, X_2, \ldots, X_{N-1}) = H(X_{m+1} \mid X_1 X_2 \cdots X_m)$$

$$= H_{m+1} = \sum_{i=1}^{n^m} p(S_i) H(X \mid S_i) --- m\text{阶条件熵}$$

$$= \sum_{i=1}^{n^m} \sum_{j=1}^{n} p(S_i) p(a_j \mid S_i) \log_2 p(a_j \mid S_i)$$

$$= \sum_{i=1}^{n^m} \sum_{j=1}^{n^m} p(S_i) p(S_j \mid S_i) \log_2 p(S_j \mid S_i)$$

称马尔可夫链为齐次的，如果马尔可夫链的统计特性不随时间而改变。

➢马尔可夫链的平稳分布

若齐次马尔可夫链对一切i, j存在不依赖于i的极限

$$\lim_{n\to\infty} p_n(s_j \mid s_i) = p(s_j)$$

且满足

$$p(s_j) > 0,\quad \sum_{j=1}^{n^m} p(s_j) = 1$$

则称其具有遍历性，$\boldsymbol{p}(\boldsymbol{s}_\mathrm{i})$称为平稳分布。

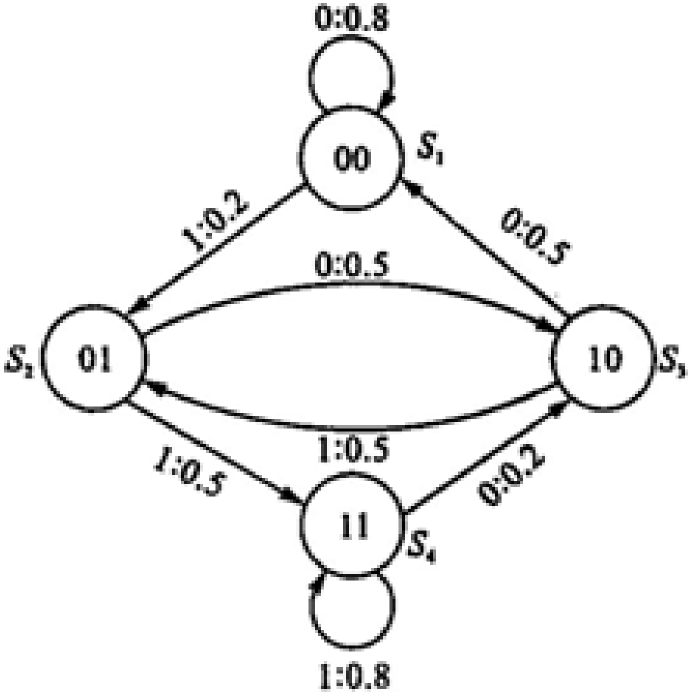
$$p(s_j) = \sum_{i=1}^{n^m} p(s_i) p(s_j / s_i),\quad j = 1, 2, \cdots, n^m$$

[例] 二元2阶马尔可夫信源，原始信号X的符号集为$\{X1=0, X2=1\}$，其状态空间共有$n^m=2^2=4$个不同的状态$s1,s2,s3,s4$，即

E: $\{s1=00,s2=01,s3=10,s4=11\}$, $p(0|00)=p(1|11)=0.8$;

$p(1|00)=p(0|11)=0.2$; $p(0|01)=p(0|10)=p(1|01)=p(1|10)=0.5$

状态转移图见右图所示。



解: $p(s1/s1)= p(x1/s1)=p(0/00)=0.8$

$p(s2/s1)= p(x2/s1)=p(1/00)=0.2$

$p(s3/s2)= p(x1/s2)=p(0/01)=0.5$

$p(s4/s2)= p(x2/s2)=p(1/01)=0.5$

$p(s1/s3)= p(x1/s3)=p(0/10)=0.5$

$p(s2/s3)= p(x2/s3)=p(1/10)=0.5$

$p(s3/s4)= p(x1/s4)=p(0/11)=0.2$

$p(s4/s4)= p(x2/s4)=p(1/11)=0.8$

对于由二元信源$X\in\{0,1\}$得到的状态空间(s_1,s_2,s_3,s_4)，容易验证

$$\sum_{j=1}^{4} p(s_j \mid s_i) = 1, \qquad p(s_i) > 0$$

称为状态极限概率

由 $$p(s_j) = \sum_{i=1}^{4} p(s_i) p(s_j \mid s_i) \qquad j = 1,2,3,4$$

求出稳定状态下的$p(s_j)$ 。将一步转移概率代入上式得

$$p(s_1) = 0.8\, p(s_1)+0.5\, p(s_3)$$

$$p(s_2) = 0.2\, p(s_1)+0.5\, p(s_3)$$

$$p(s_3) = 0.5\, p(s_2)+0.2\, p(s_4)$$

$$p(s_4) = 0.5\, p(s_2)+0.8\, p(s_4)$$

解方程组得 $p(s_1)= p(s_4)=5/14$， $p(s_2)= p(s_3)=1/7$

极限熵

$$H_\infty = H_{2+1} = -\sum_{i=1}^{4}\sum_{j=1}^{4} p(s_i) p(s_j \mid s_i) \log_2 p(s_j \mid s_i) = 0.801 \text{ (bit/sign)}$$

■ 信源的相关性和冗余度

平均符号熵随N增加是非递增的, 也就是说信源输出符号间的相关程度越长，信源的实际熵越小，趋近于极限熵；若相关程度减小，信源实际熵增大。

$$0 \le H_N(X) \le H_{N-1}(X) \le \cdots \le H_1(X) \le H_0(X) \le \infty$$

当信源输出符号间彼此不存在依存关系，且为等概率分布时，信源实际熵趋于最大熵H_0

$$H_0(X) = \log_2 n$$

为了衡量信源的相关程度，引入信源冗余度的概念。

称$\eta = H_\infty / H_0$为相对熵率。

冗余度（多余度、剩余度）表示给定信源在实际发出消息时所包含的多余信息。

冗余度 $R = 1 - \eta$

冗余度来自两个方面：

① 信源符号间的相关性。由于信源输出符号间的依赖关系使得信源熵减小，这就是信源的相关性。相关程度越大，信源的实际熵越小，趋于极限熵$H_\infty(X)$；反之相关程度减小，信源实际熵就增大。

② 信源符号分布的不均匀性。当等概率分布时信源熵最大。而实际应用中大多是不均匀分布，使得实际熵减小。当信源输出符号间彼此不存在依赖关系且为等概率分布时，信源实际熵趋于最大$H_0(X)$。

我们可以压缩剩余度来压缩信源，提高通信的效率。

自然语言的熵

（1）对于英文字母

$$R=1-\eta=1-\frac{H_\infty}{H_0}=1-\frac{1.4}{\log 27}=0.71$$

（2）对于中文

$$R=1-\eta=1-\frac{H_\infty}{H_0}=1-\frac{9.733}{\log 10^4}=0.264$$

增加冗余

原始信息的冗余 → 恢复传输中被损坏的信息

信息压缩可能

提高符号的信息携带量

减少冗余，提高传输率

这是信源编码的作用

提高信息的抗干扰能力

这是信道编码的作用

（冗余小，有效）　　（冗余大，可靠）

中国 ⇔ 中华人民共和国

2.4 离散无失真信源编码定理

❍信源问题

❍信源输出的信息计算问题

前面主要讨论的

❍信源输出的有效表示问题

离散信源符号
- 概率分布存在不均匀性
- 符号间相关性

冗余度

传输效率低

信源效率低

解决办法：对信源进行压缩编码，减少冗余。

信源编码：

- 编码以提高通信有效性为目的。
- 通常通过压缩信源的冗余度来实现。

■ 离散信源的无失真编码

将信源输出的离散消息符号 → **转变** → 适合信道传输的信道符号

该变换满足：在不失真的前提下，用较短的码字来代表每一条消息。

信源编码器

将信源输出符号列 → 数学规则 → 信道传输的符号列(称为码序列)

- **编码**：对信源输出的原始符号按照一定的数学规则进行的变换。
- **编码器**：完成编码功能的器件。

接收端对应有译码器

- **注**：各种码字长度可以相同也可以不同。
  - 等长编码：码字长度相同。
  - 变长编码：码字长度不同。

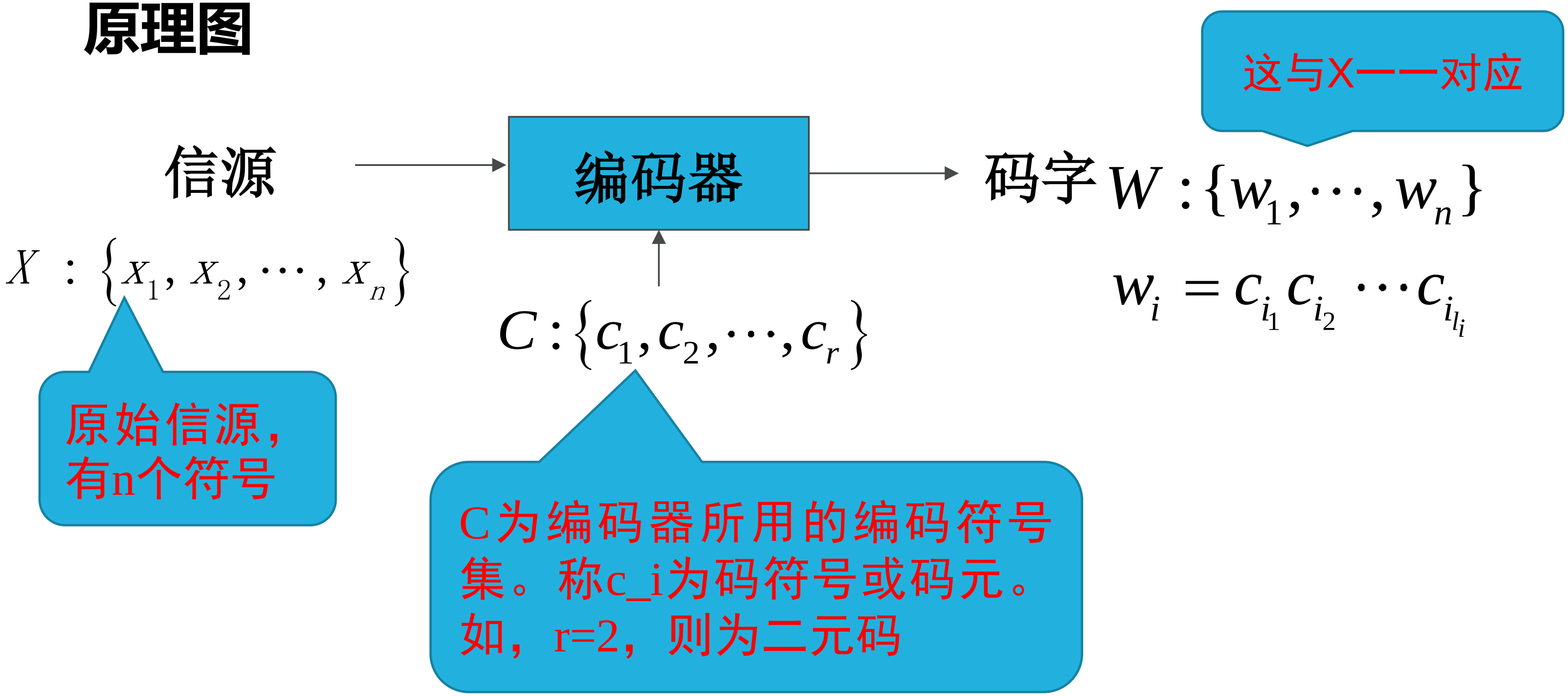
编码可以减少信源的冗余度，但在编码效率相同的前提下，等比编码比变长更复杂。

■ 唯一可译码

若某种码的任意一串有限长的符号序列只能被唯一地译成所对应的信源符号，则该码称为唯一可译码。

单符号离散信源

原理图



- ✓W为编码器输出的码字集合，共有n个码字 $w_i = c_{i_1}c_{i_2}\cdots c_{i_{l_i}}$
- ✓$|w_i|=l_i$,称l_i为w_i的码长。
- ✓全体码字集合称为码，平均码长为 $\bar{L} = \sum_{i=1}^{n} p(x_i) l_i$

编码过程：

按照一定的规则，将信源中每个原始符号xi用码字wi表示，满足：

① 无失真传送；

② 唯一可译：任意有限长的码元序列只能被唯一地分割成一个个码字，并被译成所对应的信源符号。

例

对于信源$\begin{pmatrix} X \\ p(X) \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ 0.4 & 0.1 & 0.3 & 0.2 \end{pmatrix}$，

若对应编码$W = \{1, 0001, 01, 001\}$，则$\overline{L} = \sum_{i=1}^{4} p(x_i) l_i = 2$；

若对应编码$W = \{0001, 1, 001, 01\}$，则$\overline{L} = \sum_{i=1}^{4} p(x_i) l_i = 3$。

信源符号与编码对应关系决定效率。即，大概率符号赋予短码，小概率符号赋予长码(反概率搭配)。

给定一个信源空间，一定有一种平均码长最小的信源编码，称此最小的平均码长为信源的最小平均码长。

N次扩展信源无失真编码基本原理

信源消息共有n^N个，相应的码字有n^N个，码元集合为$\{c_1,\cdots,c_r\}$。

对于X^N中消息$\vec{\alpha_i}$的编码w_i的长为l_i，则

对X^N中消息$\{\vec{\alpha_i}\}$的编码$\{w_i\}$的平均码长为$\overline{L}_N=\sum\limits_{i=1}^{n^N}p(\vec{\alpha_i})l_i$。

于是，原始信源 X 中各符号编码的平均码长为$\frac{\overline{L}_N}{N}$。

信源编码

N长序列 → **信源编码器** → K长码字

$$X = (X_1 X_2 \cdots X_l \cdots X_N)$$

$$X_l \in \{a_1, a_2, \cdots, a_i, \cdots, a_n\}$$

$$\mathcal{C} = \{c_1, \cdots, c_m\}$$

$$W = (W_1 W_2 \cdots W_k \cdots W_K)$$

$$W_k \in \{c_1, c_2, \cdots, c_m\}$$

- ✓ **原始信源X的信息熵为H(X)，即，每个信源符号携带信息量。**
- ✓ **每个码符的最大信息量log m， 即，$\mathcal{C}$ 的最大熵。**
- ✓ 平均码长为$K = \overline{L}_N$， $\frac{\overline{L}_N}{N}$个码符平均携带原始信源的一个符号的平均信息量H(X)。

用C的不确定性（信息量）表示X的不确定性。

(1)信息率(编码速率)

$$R = (K/N)\log_2 m \quad \text{bit/信源符号}$$

$\log_2 m$ —— 每个码符号的最大熵(bit/码符号)

$K\log_2 m$ —— 每个码符号序列最大熵(bit/码序列)

$(K/N)\log_2 m$— 编码后，平均每个信源符号所能载荷的最大信息量

(2)编码效率

$$\eta = \frac{H(X)}{R}$$

$H(X)$—编码前，平均每个信源符号包含的信息量

R——编码后，平均每个信源符号所能传送的最大信息量

若$R = H(X) + \varepsilon$, 于是

$$\eta = \frac{H(X)}{H(X)+\varepsilon}$$

例 设离散无记忆信源

$$\begin{pmatrix} X \\ p(X) \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix},$$ 计算

1) H(X)

2) 用二元码符号$\{0,1\}$对 X 进行编码：$x_1 \to 0$与$x_2 \to 1$。
计算平均码长、编码效率和信息率。

3) 给出对二次扩展信源 X^2 的编码，计算 X^2 平均码长、编码效率和信息率。

类似可计算X^3,X^4等的编码及效率。

由此可见，随着信源的扩展次数N的增加，编码效率越来越接近1，编码后得到的信息率也越来越接近原始信源的最大熵。

于是，增加信源的扩展次数，可以非常有效地提高信源的编码效率。

❍Shannon第一定理

给定离散无记忆信源为

$$\begin{pmatrix} X \\ P(X) \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ p(x_1) & p(x_2) & \cdots & p(x_n) \end{pmatrix},$$

其熵为 H(X)。设其 N 次扩展信源为

$$\begin{pmatrix} X^N \\ P(X^N) \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & \cdots & a_{n^N} \\ p(a_1) & p(a_2) & \cdots & p(a_{n^N}) \end{pmatrix},$$

其熵为$H(X^N)$，码符集为$\mathcal{C} = \{c_1, \cdots, c_r\}$。现在对信源$X^N$进行编码，总可以找到一种编码方法构成唯一可译码，使信源 X 中每个信源符号所需的码字平均长度满足

$\frac{1}{N} + \frac{H(X)}{\log r} > \frac{\overline{L}_N}{N} \geq \frac{H(X)}{\log r}$，或$\frac{1}{N} + H_r(X) > \frac{\overline{L}_N}{N} \geq H_r(X)$。

当$N \to \infty$时，$\lim\limits_{N\to} \frac{\overline{L}_N}{N} = H_r(X)$。

这里$\overline{L}_N$是无记忆扩展信源X^N中每个信源符号a_i所对应的平均码长，

$\frac{\overline{L}_N}{N}$表示散无记忆信源 X 中每个信源符号x_i所对应的平均码长。

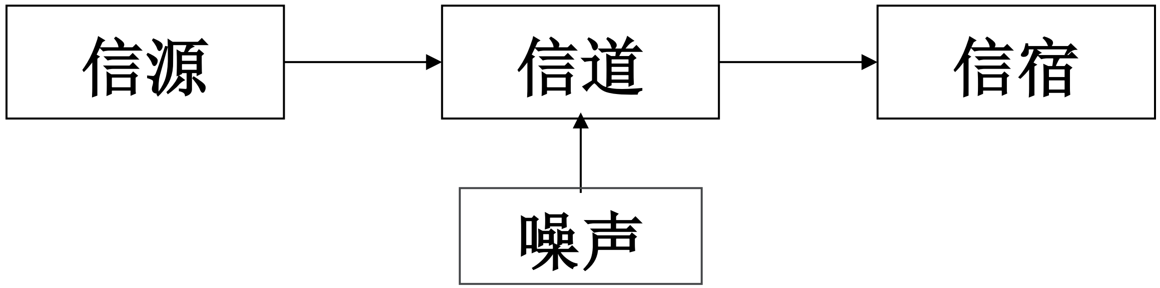
说明：$\overline{L}_1$与$\overline{L}_N/N$都是原始信源符号所需的码符的平均数，但不同的是：

1) $\overline{L}_N/N$是对 N 次扩展信源符号序列的编码长度，然后对 N 求平均。
2) $\overline{L}_1$是单符号编码的长度。

说明：

1) 要做到无失真的信源编码，每个信源符号平均所需最少的 r-元码元数为$H_r(X)$，即，$H_r(X)$为无失真信源压缩的极限值。
2) 若编码的平均码长小于$H_r(X)$，则唯一可译码不存在，在译码时必然会带来失真或差错。
3) 通过增加信源的扩展次数 N，即，让输入编码器的信源消息分组长度 N 增大，可以使编码的平均码长$\overline{L}_N/N$趋于下限。显然，减少平均码长所付出的代价是增加编码的复杂性。
4) 编码效率η 不超过 1。当$\overline{L}_N/N$达到$H_r(X)$时，编码效率η =1，此时编码达到最高效率。

第三章 信道+容量



通信系统的简化模型

通信的本质：
信息通过信道得以传输，实现异地间的交流。

- **信道的任务**：以信号方式传输信息和存储信息。

❍研究内容：

- ➢ **描述、度量和分析不同类型的信道；**
- ➢ **计算其能够传输的最大信息量，即，信道容量。**
- ➢ **分析信道的特性。**

通信中信道按其物理组成有不同信道，信息在其中传输时遵循不同的物理规律。通信技术必须研究这些规律以获得信号在这些信道中的传输特性。

❍我们介绍：

- ❍不介绍怎样获得这些传输特性。
- ❍假定传输特性已知，研究信息的传输问题，抽象地将信道用数学模型来描述。

❍信道模型

- 信道的数学模型：

$$X \longrightarrow \boxed{P(Y|X)} \longrightarrow Y$$

- X : 输入事件的集合, 概率空间为[X,P(X)]
- Y : 输出事件的集合, 概率空间为[Y,P(Y)]

- 信道：在输入已知的情况下，用输出的条件概率分布$P(Y|X)$表示。

❍信道表示为 [X,P(Y|X),Y]

➢信道分类

1、根据信道用户的多少，分为：

- ✓单用户信道---输入、输出均只有一个
- ✓多用户信道---输入、输出有多个

2、根据输入输出信号的特点，分为：

- ✓离散信道---输入、输出随机变量均离散取值
- ✓连续信道---输入、输出随机变量均连续取值
- ✓半离散(连续)信道--- 一为离散，另一为连续

3、根据输入、输出随机变量的个数

- ✓单符号信道---输入、输出均用随机变量表示
- ✓多符号信道---输入、输出用随机矢量表示

4、根据信道上有无噪声(干扰)：

- ✓有噪(扰)信道
- ✓无噪(扰)信道

5、根据信道有无记忆特性

- ✓无记忆信道---输出仅与当前输入有关，与先前输入无关
- ✓有记忆信道---输出不仅与当前输入有关，还与先前输入有关

离散无记忆信道(DMC)

信道在任意时刻的输出只与此时刻信道的输入有关，而与其它时刻的输入和输出无关，则称之为离散无记忆信道。

- 单符号离散信道的数学模型

设离散信道的输入空间为 $X \in \{a_1, a_2, \ldots, a_n\}$

概率分布为 $\{p_i\}, i = 1, 2, \ldots, n$

输出空间为 $Y \in \{b_1, b_2, \ldots, b_m\}$

概率分布为 $\{q_j\}, j = 1, 2, \ldots, m$

并有条件概率 $P(y \mid x) = P(b_j \mid a_i), (i = 1, 2, \cdots, n; j = 1, 2, \cdots, m)$

称条件概率为信道的传递概率或转移概率。

$$\mathbf{X}\left\{\begin{matrix} a_1 \\ \vdots \\ a_n \end{matrix}\right. \longrightarrow \boxed{P(b_j \mid a_i)} \longrightarrow \left.\begin{matrix} b_1 \\ \vdots \\ b_m \end{matrix}\right\}\mathbf{Y}$$

将所有转移概率以矩阵方式列出，得：

$$[P(Y|X)]=\begin{bmatrix} p(b_1|a_1) & p(b_2|a_1) & \dots & p(b_m|a_1) \\ p(b_1|a_2) & p(b_2|a_2) & \dots & p(b_m|a_2) \\ \dots & \dots & \dots & \dots \\ p(b_1|a_n) & p(b_2|a_n) & \dots & p(b_m|a_n) \end{bmatrix}$$

其中 $p(b_j|a_i)\ge 0$ $\sum_{j=1}^{m} p(b_j|a_i)=1$

- **该矩阵完全描述了信道在干扰作用下的统计特性，称为信道矩阵(n行m列)。**
  - **行表示输入x，列表示输出y；**
  - **每个元素均非负，每一行元素之和为1。**
  - **每列若有多个非零，则疑义度H(X|Y)非零。**

离散信道中的概率关系

（1）联合概率 $P(a_i b_j) = P(a_i)P(b_j / a_i) = P(b_j)P(a_i / b_j)$

其中 $P(b_j / a_i)$ 称为前向概率，描述信道的噪声特性

$P(a_i / b_j)$ 称为后向概率，有时也把 $P(a_i)$ 称为先验概率

$P(a_i / b_j)$ 称为后验概率

（2）输出符号的概率 $P(b_j) = \sum_{i=1}^{n} p(a_i) p(b_j / a_i)$

（3）后验概率 $P(a_i / b_j) = \dfrac{P(a_i b_j)}{P(b_j)}$ $\sum_{i=1}^{n} P(a_i / b_j) = 1$

表明输出端收到任一符号，必定是输入端某一符号输入所致。

- 互信息I(X;Y)表示输出端只能收到的信息量。即，在平均意义下，每传送一个符号，流经信道的平均信息量。
- 信息传输率R：信道中平均每个符号所能传送的信息量。

由于平均互信息I(X;Y)的含义是接收到符号Y后，平均每个符号获得的关于X的信息量，因此信道信息传输率就是平均互信息。

$$R = I(X;Y) = H(X) - H(X|Y)$$

信息传输速率Rt：单位时间内平均传输的信息量

$$R_t = \frac{1}{t} I(X;Y)$$ 单位：比特/秒

最大信息传输速率Ct：单位时间内平均传输的最大信息量

$$C_t = \frac{C}{t} = \frac{1}{t} \max_{p(x_i)} I(X;Y)$$

3.2单符号离散信道的信道容量

信道容量的定义

平均互信息I(X;Y)是信源分布p(x)的上凸函数，是信道传递概率p(y|x)的下凸函数。

对于一个固定的信道，总存在一种信源，使传输每个符号平均获得的信息量I(X;Y)最大。

对于固定信道，总有一个最大信息传输率

信道容量定义为平均互信息的最大值：

$$C = \max_{p(x_i)} I(X;Y) \qquad \text{bit/信道符号}$$

信道容量C只与信道的统计特性p(y|x)有关，而与信源的分布p(x)无关。它是信道的特征参数，反应的是信道的最大的信息传输能力。

含义：给定信道时，

① 对应各种信源分布，求取的最大平均互信息；

② 理论上能传输的最大信息量，表征信道传送信息的最大能力。

- **说明：**
  1. C是信道传送消息的最大能力的度量；仅与信道的统计特性有关，而与输入信源的概率分布无关。即，信道固定，则C固定。
  2. 信道实际传送的信息量不超过C；否则，发生错误。

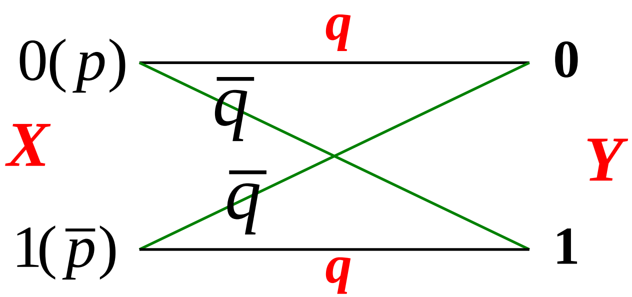
- **核心问题**

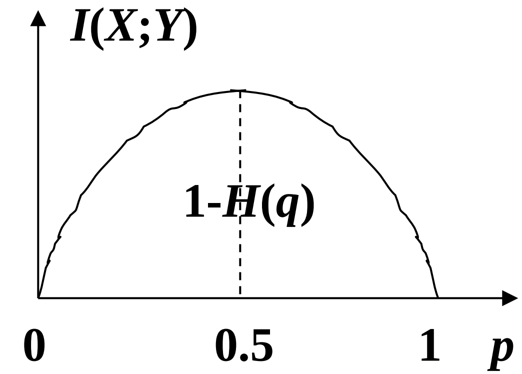
  计算信道容量，确定达到信道容量的信源概率分布。

  - **计算一般信道容量十分复杂：对平均互信息求极大值。**
  - **计算特殊信道容量，简化计算：**
    - 利用信道的特点；
    - 运用信息论的基本概念。

- **下面考虑几种特殊离散信道计算**

例：对于二元对称信道



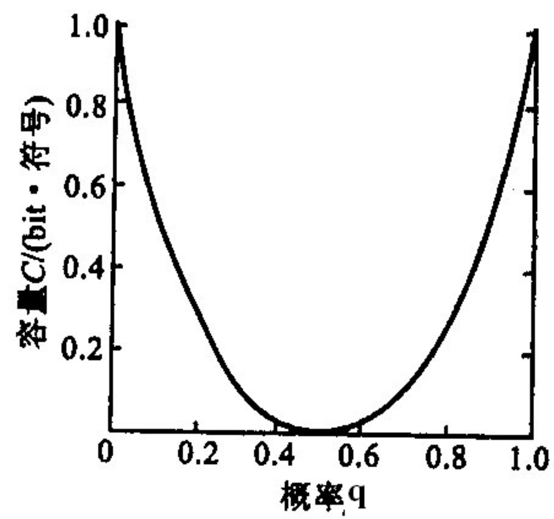


如果信源分布X={p, 1−p}，则

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

信道容量为：

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



二元对称信道的信道容量

3.2单符号离散信道的信道容量

■ 离散无噪信道：输出Y和输入X有确定关系

➢ 无损无噪确定信道：一对一（n=m）

$$\begin{array}{ccc} \boldsymbol{a_1} & \text{———} & \boldsymbol{b_1} \\ \boldsymbol{a_2} & \text{———} & \boldsymbol{b_2} \\ \boldsymbol{a_3} & \text{———} & \boldsymbol{b_3} \\ \cdots & \cdots & \cdots \\ \boldsymbol{a_n} & \text{———} & \boldsymbol{b_n} \end{array}$$

信道矩阵：

$$\boldsymbol{P} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$p(b_j \mid a_i) = p(a_i \mid b_j) = \begin{cases} 0,(i \neq j) \\ 1,(i = j) \end{cases}$$

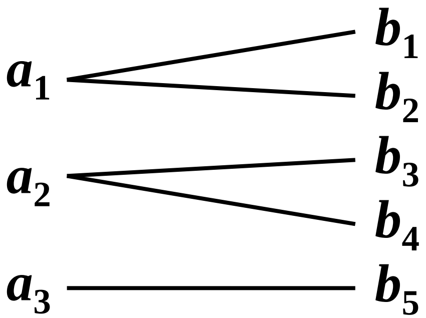
损失熵 H(X/Y)=0； 噪声熵 H(Y/X)=0；H(X) = H(Y)

I(X;Y)=H(X)-H(X|Y)=H(Y)-H(Y|X)

$$C = \max_{p(x_i)} I(X;Y) = \max_{p(x_i)} H(X) = \max_{p(x_i)} H(Y) = \log_2 n = \log_2 m$$

无损信道：一对多（n<m）

具有扩展性的无噪信道



信道矩阵：

每列只有一个非0元素，不全是0、1

$$\boldsymbol{P}=\begin{bmatrix} p_{11} & p_{12} & 0 & 0 & 0 \\ 0 & 0 & p_{23} & p_{24} & 0 \\ 0 & 0 & 0 & 0 & p_{35} \end{bmatrix},\ p(a_i\,|\,b_j)=\begin{cases} 0,(b_j\notin B_i) \\ 1,(b_j\in B_i) \end{cases}$$

信道疑义度（损失熵）H(X|Y)=0；噪声熵 H(Y|X)>0

I(X;Y)=H(X)-H(X|Y)=H(Y)-H(Y|X)=H(X)

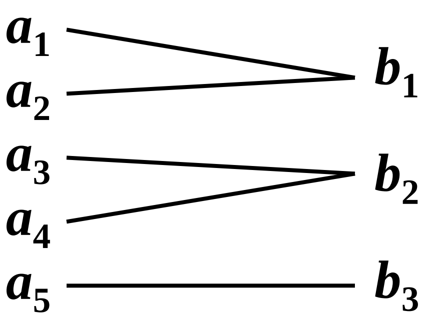
H(X)<H(Y)

发出X，不能确定收到符号。

$$C=\max_{p(x_i)} I(X;Y)=\max_{p(x_i)} H(X)=\log_2 n$$

➢ 确定信道：多对一（n>m）

具有归并性的无噪信道



信道矩阵：

$$P=\begin{bmatrix}1&0&0\\1&0&0\\0&1&0\\0&1&0\\0&0&1\end{bmatrix},\ p(b_j\,|\,a_i)=\begin{cases}0,(a_i\notin A_j)\\1,(a_i\in A_j)\end{cases}$$

每行只有一个元素“1”，其它全是0

接收到符号却不能确切判断发出的符号。

损失熵 H(X/Y)>0；噪声熵 H(Y/X)=0

I(X;Y)=H(X)-H(X|Y)=H(Y)-H(Y|X)=H(Y)

H(X) > H(Y)

$$C=\max_{p(x_i)} I(X;Y)=\max_{p(x_i)} H(Y)=\log_2 m$$

思考：p(x)应该怎样取值？

- **说明：**
  - 在计算信道容量时，要调整的始终是输入端的概率分布p(x)。在C=logm中，虽然信道容量等于输出端符号熵H(Y)，但是在求极大值时调整的仍是输入端的概率p(x)，而不能用输出端p(y)替代调整。
- **小结**
  - 以上三种信道，求信道容量C的问题实现如下归约：

计算信道容量 → 计算I(X;Y)的极值 → 计算H(X)和H(Y)的极大值

此时，信道容量C只决定于信道的输入符号数n或输出符号数m，与具体信源无关，表示了信道的统计特性。

下面考虑离散对称信道

3.2单符号离散信道的信道容量

- 强对称离散信道(均匀信道)
  - 信道特点

    信道输入、输出均为n元

    每符号正确传输概率均为 $\bar{p}=1-p$

    其他符号错误传输概率为p/(n-1)
  - 矩阵特点

    (1)n×n阶对称阵

    (2)每行和为1，每列和为1

$$\boldsymbol{P}=\begin{bmatrix} \bar{p} & \frac{p}{n-1} & \cdots & \frac{p}{n-1} \\ \frac{p}{n-1} & \bar{p} & \cdots & \frac{p}{n-1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{p}{n-1} & \frac{p}{n-1} & \cdots & \bar{p} \end{bmatrix}$$

3.2单符号离散信道的信道容量

- 信道容量 $$C=\max_{p(x_i)}[H(Y)-H(Y/X)]$$

$$令H_{ni}=-\sum_j p(y_j/x_i)\log_2 p(y_j/x_i)$$

$$得H_{ni}=-\bar{p}\log_2\bar{p}-p\log_2\frac{p}{n-1},与i无关$$

$$于是H(Y/X)=\sum p(x_i)H_{ni}=H_{ni}$$

$$C=\max_{p(x_i)}[H(Y)-H(Y/X)]=\max_{p(x_i)}[H(Y)-H_{ni}]$$

$$=\log n+\bar{p}\log\bar{p}+p\log\frac{p}{n-1}$$

$$=\log n-p\log(n-1)-H(p)\ \text{bit/信道符号}$$

可以看出，当输入为等概率分布时，输出等概分布，即H(Y)=log$\boldsymbol{n}$，达到信道容量。

3.2单符号离散信道的信道容量

那么，在什么样的信源输入情况下，信道输出能等概分布呢？可以证明，输入等概分布时，离散对称信道的输出也为等概分布

$$p(b_j)=\sum_{i=1}^{n}p(a_i)p(b_j\mid a_i),\ j=1,2,\ldots,n$$

$$\begin{bmatrix} p(b_1)\\ p(b_2)\\ \vdots\\ p(b_n)\end{bmatrix}=\boldsymbol{P}_{n\times n}^{T}\begin{bmatrix} p(a_1)\\ p(a_2)\\ \vdots\\ p(a_n)\end{bmatrix}=\begin{bmatrix} \bar{p} & \frac{p}{n-1} & \cdots & \frac{p}{n-1}\\ \frac{p}{n-1} & \bar{p} & \cdots & \frac{p}{n-1}\\ \vdots & \vdots & \vdots & \vdots\\ \frac{p}{n-1} & \frac{p}{n-1} & \cdots & \bar{p}\end{bmatrix}\begin{bmatrix} p(a_1)\\ p(a_2)\\ \vdots\\ p(a_n)\end{bmatrix}$$

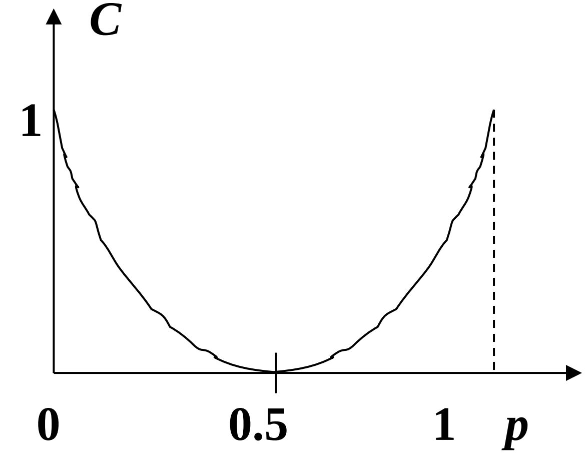
[结论]对强对称信道，输入等概→输出等概，可达到C

➢ 当n=2时，称为二元对称信道，记为B.S.C

$$C = \log_2 n + \bar{p}\log_2 \bar{p} + p\log_2 \frac{p}{n-1}$$

$$= 1 + \bar{p}\log_2 \bar{p} + p\log_2 p = 1 - H(p)$$

$$H(p) = -p\log_2 p - \bar{p}\log_2 \bar{p}$$



离散有对称性的信道

离散输入对称信道

如果一个离散无记忆信道的信道矩阵中，每一行都是其他行的同一组元素的不同排列，则称此类信道为离散输入对称信道。

$$\boldsymbol{P}=\begin{bmatrix} 1/3 & 1/3 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/3 & 1/3 \end{bmatrix}$$

离散输出对称信道

如果一个离散无记忆信道的信道矩阵中，每一列都是其他列的同一组元素的不同排列，则称此类信道为离散输出对称信道。

$$\boldsymbol{P}=\begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix}$$

离散对称信道

如果一个离散无记忆信道的信道矩阵中，每一行都是其他行的同一组元素的不同排列，并且每一列都是其他列的同一组元素的不同排列，则称此类信道为离散对称信道。

$$\boldsymbol{P}=\begin{bmatrix} 1/3 & 1/3 & 1/6 & 1/6 \\ 1/6 & 1/6 & 1/3 & 1/3 \end{bmatrix} \qquad \boldsymbol{P}=\begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 1/6 & 1/2 & 1/3 \\ 1/3 & 1/6 & 1/2 \end{bmatrix}$$

离散对称信道

行可排列——矩阵每行各元素都来自同一集合Q

$Q \in \{q_1, q_2, \ldots, q_m\}$（每行是其它行的排列）

列可排列——矩阵每列各元素都来自同一集合P

$P \in \{p_1, p_2, \ldots, p_n\}$（每列是其它列的排列）

称这样的离散信道为离散对称信道。

（1）$m=n$ 时，Q、P为同一集合

$m \neq n$时，Q、P中，一个必为另一个的子集

（2）输入等概→输出等概

➢ 离散对称信道的信道容量

定理：如果一个离散对称信道具有n个输入符号，m个输出符号，则当输入为等概率分布时，达到信道容量C。

$$令H_{mi}=-\sum_{j=1}^{m}p(y_j|x_i)\log_2 p(y_j|x_i)$$

$$\text{So } H(Y|X)=-\sum_{i=1}^{n}\sum_{j=1}^{m}p(x_i)p(y_j|x_i)\log_2 p(y_j|x_i)=\sum_{i=1}^{n}p(x_i)H_{mi}$$

对于行可排列情况，$\boldsymbol{H_{mi}}$与$\boldsymbol{i}$ 无关

$$\therefore H(Y|X)=H_{mi}=H(q_1,q_2,\cdots,q_m)--常数$$

$$C=\max_{p(x_i)}[H(Y)-H(Y/X)]=\log_2 m-H(q_1,q_2,\cdots,q_m)$$

可以看出，当输出等概分布时，即$H(Y)=\log_2 m$时达到信道容量。

由于对称信道的特点，输入等概率分布→输出等概率分布。

这是为什么？

准对称离散信道

信道矩阵的行可排列，列不可排列

但把该矩阵在水平方向上分割，则各子矩阵皆具可排列性

[例]

$$\boldsymbol{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8}\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8}\end{bmatrix}\qquad \boldsymbol{P}_1=\begin{bmatrix}\frac{1}{2} & \frac{1}{4}\\ \frac{1}{4} & \frac{1}{2}\end{bmatrix}\qquad \boldsymbol{P}_2=\begin{bmatrix}\frac{1}{8} & \frac{1}{8}\\ \frac{1}{8} & \frac{1}{8}\end{bmatrix}$$

[关键]——$\boldsymbol{H(Y)}$最大（改变$\boldsymbol{p(x_i)}$时）

$$C=\max_{p(x_i)}[H(Y)]-H(Y\mid X)=\max_{p(x_i)}[H(Y)]-H(q_1,q_2,\cdots,q_m)$$

思考：$\boldsymbol{p(x_i)}$应该怎样取值？

将H(Y)中的m项分解成s个子集$M_1,M_2,\ldots,M_s$,各个子集分别有$m_1,m_2,\ldots,m_s$个元素($m_1+m_2+\ldots+m_s=m$),则

$$H(Y)=-\sum_{j=1}^{m}p(b_j)\log_2 p(b_j)=-\sum_{k=1}^{s}\sum_{p(b_j)\in M_k}^{m}p(b_j)\log_2 p(b_j)$$

$$=-\sum_{p(b_j)\in M_1}p(b_j)\log_2 p(b_j)\cdots-\sum_{p(b_j)\in M_s}p(b_j)\log_2 p(b_j)$$

则

$$\overline{p}(b_k)=\frac{\sum\limits_{p(b_j)\in M_k}p(b_j)}{m_k},k=1,2,\ldots,s \qquad \frac{\overline{p}(b_k)}{p(b_j)}>0$$

由 $\ln x \le x-1 (x>0)$ $\log_2 x = \ln x \bullet \log_2 e$

$$\sum_{p(b_j)\in M_k} p(b_j)\log_2 \frac{\overline{p}(b_k)}{p(b_j)} = \sum_{p(b_j)\in M_k} p(b_j)\ln \frac{\overline{p}(b_k)}{p(b_j)}\log_2 e \le$$

$$\sum_{p(b_j)\in M_k} p(b_j)\left[\frac{\overline{p}(b_k)}{p(b_j)}-1\right]\log_2 e = \left[\overline{p}(b_k)m_k - \sum_{p(b_j)\in M_k} p(b_j)\right]\log_2 e = 0$$

故有 $$\sum_{p(b_j)\in M_k} p(b_j)\log_2 \overline{p}(b_k) \le \sum_{p(b_j)\in M_k} p(b_j)\log_2 p(b_k)$$

即 $$-\sum_{p(b_j)\in M_k} p(b_j)\log_2 p(b_k) \le -\sum_{p(b_j)\in M_k} p(b_j)\log_2 \overline{p}(b_k)$$

$$= -\log_2 \overline{p}(b_k)[m_k \overline{p}(b_k)] = -m_k \overline{p}(b_k)\log_2 \overline{p}(b_k)$$

$$-\sum_{p(b_j)\in M_k} p(b_j)\log_2 p(b_k) \le -m_k \bar{p}(b_k)\log_2 \bar{p}(b_k)$$

$$p(b_j) \to \bar{p}(b_k)$$

各个子矩阵具有可排列性, 只要信源呈等概率分布，即可使第k个子集中的输出概率相等，即达到其均值。

$$H(Y) \le -\sum_{k=1}^{s} m_k \bar{p}(b_k)\log_2 \bar{p}(b_k)$$

相应的准对称信道的容量为

$$C = -\sum_{k=1}^{s} m_k \bar{p}(b_k)\log_2 \bar{p}(b_k) - H(q_1, q_2, \cdots, q_m)$$

［例］ $$\boldsymbol{P}=\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{8}\\ \frac{1}{4} & \frac{1}{2} & \frac{1}{8} & \frac{1}{8}\end{bmatrix} \qquad \boldsymbol{P}_1=\begin{bmatrix}\frac{1}{2} & \frac{1}{4}\\ \frac{1}{4} & \frac{1}{2}\end{bmatrix} \qquad \boldsymbol{P}_2=\begin{bmatrix}\frac{1}{8} & \frac{1}{8}\\ \frac{1}{8} & \frac{1}{8}\end{bmatrix}$$

$$\overline{p}(b_1)=\frac{\sum_{p(b_j)\in M_1} p(b_j)}{m_1}=\frac{\sum_{p(b_j)\in M_1}\sum_{i=1}^{2} p(a_i)p(b_j\mid a_i)}{m_1} \qquad \overline{p}(b_2)=\frac{1}{8}$$

$$=\frac{\sum_{j=1}^{2}\sum_{i=1}^{2} p(a_i)p(b_j\mid a_i)}{m_1}=\frac{1}{2}[\frac{1}{2}+\frac{1}{4}]=\frac{3}{8}$$

$$C=-\sum_{k=1}^{s} m_k\overline{p}(b_k)\log_2\overline{p}(b_k)-H(q_1,q_2,\cdots,q_m)=0.06125(bit/sign)$$

$$p(a_1)=p(a_2)=\tfrac{1}{2}$$

■ 离散信道容量的一般计算方法

对于固定的信道，平均互信息$I(X;Y)$是信源概率分布$p(x_i)$的上凸函数, 所以极大值一定存在。

$I(X;Y)$是n个变量$p(a_i)(i=1,2,\ldots,n)$的多元函数，并满足

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

所以可以用拉格朗日乘子法计算条件极值:

$$\varPhi = I(X;Y) - \lambda\left[\sum_{i=1}^{n} p(a_i) - 1\right]$$

其中，λ 为拉格朗日乘子

解方程组 $\dfrac{\partial \boldsymbol{\Phi}}{\partial p(a_i)} = \dfrac{\partial\left\{I(X;Y) - \lambda[\sum_{i=1}^{n} p(a_i) - 1]\right\}}{\partial p(a_i)} = 0$

可得一般信道容量C。

由 $p(b_j) = \sum_{i=1}^{n} p(a_i) p(b_j \mid a_i) \longrightarrow \dfrac{dp(b_j)}{dp(a_i)} = p(b_j \mid a_i)$

$$\frac{\partial}{\partial p(a_i)}\left\{-\sum_{j=1}^{m} p(b_j)\log_2 p(b_j) + \sum_{i=1}^{n}\sum_{j=1}^{m} p(a_i)p(b_j \mid a_i)\log_2 p(b_j \mid a_i)\right.$$

$\boldsymbol{H(Y)}$ $\boldsymbol{H(Y|X)}$

$$\left.-\lambda[\sum_{i=1}^{n} p(a_i) - 1]\right\} = 0 \longrightarrow$$

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

$$-\sum_{j=1}^{m}[p(b_j \mid a_i)\log_2 p(b_j) + p(b_j \mid a_i)\log_2 e] + \sum_{j=1}^{m} p(b_j \mid a_i)\log_2 p(b_j \mid a_i) - \lambda = 0$$

$$\longrightarrow \sum_{j=1}^{m} p(b_j \mid a_i)\log_2 \frac{p(b_j \mid a_i)}{p(b_j)} = \log_2 e + \lambda$$

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

***I(X;Y)*最大值** $C=\log_2 e+\lambda$

$$\sum_{j=1}^{m} p(b_j\,|\,a_i)\log_2 p(b_j\,|\,a_i)=\sum_{j=1}^{m} p(b_j\,|\,a_i)\log_2 p(b_j)+C$$

$$=\sum_{j=1}^{m} p(b_j\,|\,a_i)[\log_2 p(b_j)+C]$$

令 $\beta_j=\log_2 p(b_j)+C$

则 $\sum_{j=1}^{m} p(b_j\,|\,a_i)\beta_j=\sum_{j=1}^{m} p(b_j\,|\,a_i)\log_2 p(b_j\,|\,a_i)$ **若*m*=*n*,则可求解**

$$\beta_j\longrightarrow p(b_j)=2^{\beta_j-C}\longrightarrow 2^C=\sum_{j=1}^{m}2^{\beta_j}\longrightarrow C=\log_2(\sum_{j=1}^{m}2^{\beta_j})$$

一般离散信道容量的求解步骤：

(1) $$\sum_{j=1}^{m} p(b_j \mid a_i)\beta_j = \sum_{j=1}^{m} p(b_j \mid a_i)\log_2 p(b_j \mid a_i) \longrightarrow \beta_j$$

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

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

(4) $$p(b_j) = \sum_{i=1}^{n} p(a_i)p(b_j \mid a_i) \longrightarrow p(a_i)$$

需要确认所有的 $p(a_i) \geq 0$ ，所求的***C***才存在。

例：

$$P=\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4}\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{2}\end{bmatrix}$$

可列方程组：

$$\begin{cases}\frac{1}{2}\beta_1+\frac{1}{4}\beta_2+\frac{1}{4}\beta_4=\frac{1}{2}\log\frac{1}{2}+\frac{1}{4}\log\frac{1}{4}+\frac{1}{4}\log\frac{1}{4}\\ \beta_2=0\\ \beta_3=0\\ \frac{1}{4}\beta_1+\frac{1}{4}\beta_3+\frac{1}{2}\beta_4=\frac{1}{4}\log\frac{1}{4}+\frac{1}{4}\log\frac{1}{4}+\frac{1}{2}\log\frac{1}{2}\end{cases}$$

3.2单符号离散信道的信道容量

解之得：　$\beta_2 = \beta_3 = 0,\ \beta_1 = \beta_4 = -2$

$$C = \log(2^{-2} + 2^0 + 2^0 + 2^{-2}) = \log\frac{5}{2} = \log 5 - 1$$

$$P(b_1) = P(b_4) = 2^{-2-\log 5+1} = \frac{1}{10} \qquad P(b_2) = P(b_3) = 2^{0-\log 5+1} = \frac{4}{10}$$

$$P(a_1) = P(a_4) = \frac{4}{30},\ \ P(a_2) = P(a_3) = \frac{11}{30}$$

3.3多符号离散信道的信道容量

■ 多符号离散信道的数学模型

➢ 基本概念

多符号离散信源$X=X_1X_2\ldots X_k \ldots X_N$在$N$个不同时刻分别通过单符号离散信道$\{X,P(Y/X),Y\}$，在输出端出现$Y=Y_1Y_2\ldots Y_k \ldots Y_N$，形成一个新的信道，此即多符号离散信道。由于新信道相当于单符号离散信道在N个不同的时刻连续运用了N次，也称为单符号离散信道的N次扩展信道

3.3多符号离散信道的信道容量

- 多符号离散信道的数学模型

$$X=X_1X_2\dots X_k\dots X_N \longrightarrow \boxed{\begin{matrix} P(Y/X) \\ \text{或}\{p(b_j/a_i)\} \end{matrix}} \longrightarrow Y=Y_1Y_2\dots Y_k\dots Y_N$$

X_k取值：$a_1,a_2,\dots,a_n$，则X共有n^N种α_i，$i=1,\dots,n^N$

Y_k取值：$b_1,b_2,\dots,b_m$，则Y共有m^N种β_j，$j=1,\dots,m^N$

$$\alpha_i=(a_{i_1}a_{i_2}\dots a_{i_N}),\qquad a_{i_1},a_{i_2},\dots,a_{i_N}\in\{a_1,a_2,\dots,a_n\}$$

$$\beta_j=(b_{i_1}b_{i_2}\dots b_{i_N}),\qquad b_{i_1},b_{i_2},\dots,b_{i_N}\in\{b_1,b_2,\dots,b_n\}$$

3.3多符号离散信道的信道容量

- 信道矩阵

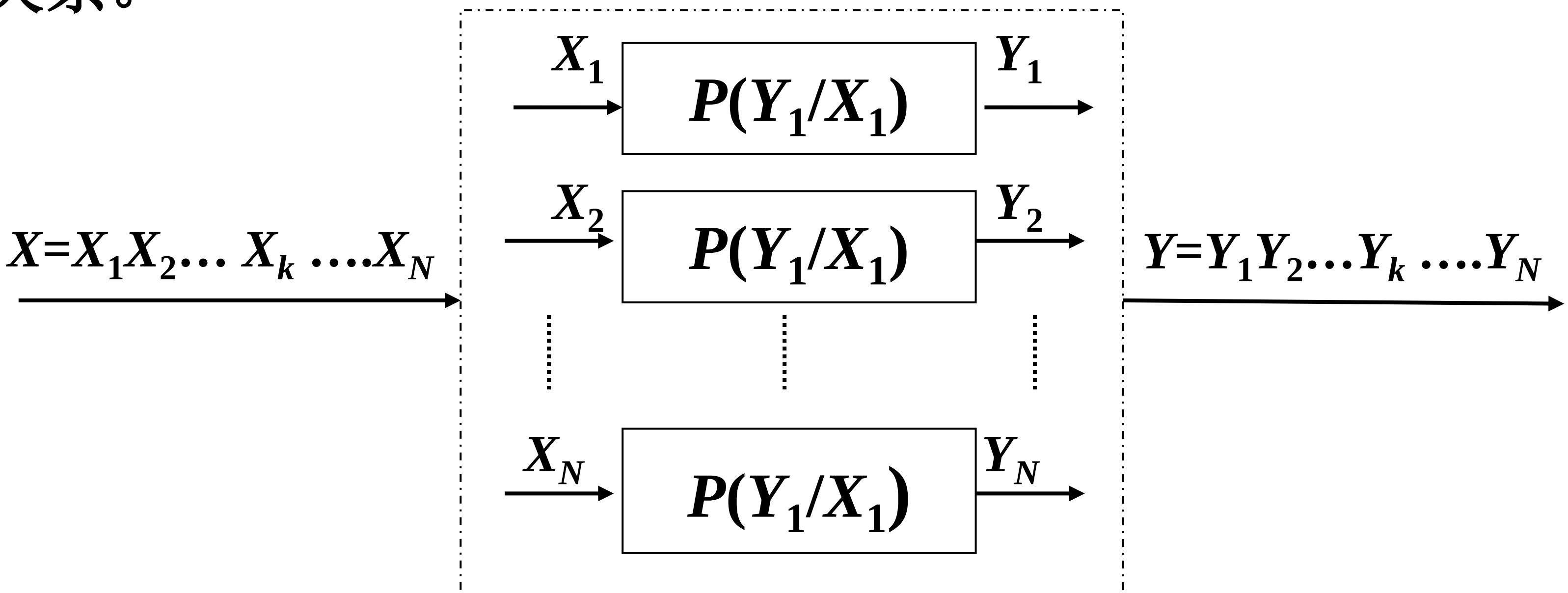
$$\boldsymbol{P}(\boldsymbol{Y}/\boldsymbol{X})=\begin{bmatrix} p(\beta_1/\alpha_1) & p(\beta_2/\alpha_1) & \cdots & p(\beta_{m^N}/\alpha_1) \\ p(\beta_1/\alpha_2) & p(\beta_2/\alpha_2) & \cdots & p(\beta_{m^N}/\alpha_2) \\ \vdots & \vdots & \vdots & \vdots \\ p(\beta_1/\alpha_{n^N}) & p(\beta_2/\alpha_{n^N})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

每行元素之和等于1

3.3多符号离散信道的信道容量

■ 离散无记忆扩展信道的信道容量

把多符号离散信道理解成单符号离散信道在每一个单位时间传递一个随机变量的时候，需要考虑k时刻的输出变量Y_k与时刻k之前的输入变量$X_1X_2\ldots X_{k-1}$和输出变量$Y_1Y_2\ldots Y_{k-1}$之间有无依赖关系。



单符号离散信道的N次扩展信道的数学模型

若多符号离散信道的转移概率满足

$$\boldsymbol{P}(\boldsymbol{Y}\mid\boldsymbol{X}) = P(Y_1Y_2\ldots Y_N \mid X_1X_2\ldots X_N) =$$

$$P(Y_1\mid X_1)P(Y_2\mid X_2)\ldots P(Y_N\mid X_N) = \prod_{k=1}^{N} p(Y_k\mid X_k)$$

则称之为离散无记忆信道的N次扩展信道

[解释]扩展信道的转移概率

=各时刻单符号信道转移概率的连乘

无记忆性——$\boldsymbol{k}$时刻输出$\boldsymbol{Y_k}$只与$\boldsymbol{k}$ 时刻输入$\boldsymbol{X_k}$有关，与$\boldsymbol{k}$时刻之前输入$\boldsymbol{X_1X_2\ldots X_{k-1}}$无关

无预感性——$\boldsymbol{k}$时刻之前输出$\boldsymbol{Y_1Y_2\ldots Y_{k-1}}$只与$\boldsymbol{k}$时刻之前输入$\boldsymbol{X_1X_2\ldots X_{k-1}}$有关，与$\boldsymbol{X_k}$无关

[结论] 离散无记忆N次扩展信道——无记忆，无预感

➢ 互信息和信道容量

离散无记忆N次扩展信道两端的平均互信息

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

由于信道无记忆 $H(\boldsymbol{Y}\mid\boldsymbol{X})=\sum_{k=1}^{N}H(Y_k\mid X_k)$

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

当第$\boldsymbol{k}$个随机变量$\boldsymbol{X_k}$单独通过单符号离散信道时，在输出端得到的关于$\boldsymbol{X_k}$的平均互信息量

$$I(X_k;Y_k)=H(Y_k)-H(Y_k\mid X_k),\qquad k=1,2,\ldots,N$$

$$\sum_{k=1}^{N}I(X_k;Y_k)=\sum_{k=1}^{N}[H(Y_k)-H(Y_k\mid X_k)]=\sum_{k=1}^{N}H(Y_k)-\sum_{k=1}^{N}H(Y_k\mid X_k)$$

相减可得 $I(\boldsymbol{X};\boldsymbol{Y})-\sum_{k=1}^{N}I(X_k;Y_k)=H(Y_1Y_2\ldots Y_N)-\sum_{k=1}^{N}H(Y_k)\le 0$

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

离散无记忆信道的N次扩展信道的平均互信息，不大于N个随机变量$X_1X_2\ldots X_N$单独通过信道{X,P(Y|X),Y}的平均互信息之和。

当且仅当信源$X=X_1X_2\ldots X_N$无记忆,或信源X是离散无记忆信源X的N次扩展信源时，等号成立。

事实上， $$H(\boldsymbol{Y})=H(Y_1Y_2\ldots Y_N)=\sum_{k=1}^{N}H(Y_k)$$

离散无记忆信道的N次扩展信道，当输入端的N个输入随机变量统计独立时，信道的总平均互信息等于这N个变量单独通过信道的平均互信息之和。

对于离散无记忆信源的N次扩展信源

$$X_k \in \{a_1, a_2, ..., a_n\}, \qquad \{p(a_1), p(a_2), ..., p(a_n)\}$$

通过同一个离散无记忆信道信道$\{\boldsymbol{X}, \boldsymbol{P(Y|X)}, \boldsymbol{Y}\}$
在信道输出端，随机变量序列$\boldsymbol{Y=Y_1Y_2\ldots Y_N}$中的随机变量$\boldsymbol{Y_k}$

$$Y_k \in \{b_1, b_2, ..., b_m\}, \qquad \{p(b_1), p(b_2), ..., p(b_m)\}$$

$$I(X_k;Y_k) = I(X;Y) \longrightarrow I(\boldsymbol{X};\boldsymbol{Y}) = \sum_{k=1}^{N} I(X_k;Y_k) = NI(X;Y)$$

$$C^N = NC$$

离散无记忆信道的N次扩展信道，如果信源也是离散无记忆信源的N次扩展信源，则信道总的平均互信息量是单符号离散无记忆信道的平均互信息量的N倍。

■独立并联信道

➢含义

信道输入序列的各随机变量取值于不同符号集

信道输出序列的各随机变量亦取值于不同符号集

也称为独立并列信道、独立平行信道或积信道

➢信道容量

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

当N个输入随机变量之间统计独立，并且每个输入随机变量X_k的概率分布为达到各自信道容量C_k的最佳分布时，C^N达到最大值

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

N个独立并联信道的信道容量等于各个信道容量之和

信道编码定理

- 信道编码定理(Shannon第二定理)

对离散平稳无记忆信道，其信道容量为***C***，输入序列长度为***L***。只要实际信息率***R***<***C***，就必可找到一种编码，当***L***足够长时，译码差错概率***Pe*** < ***ε***，***ε***为任意大于零的正数。反之，若实际信息率***R***>***C***，则对任何编码，***Pe***必大于零；且当 $L \to \infty$ 时，$P_e \to 1$ 。

[说明]

①给出了信息传输率的极限

只要***R***<***C***，必可无失真传输

若***R***>***C***，必为有失真传输

②存在性定理

失真不可避免，甚至必要；允许一定程度的失真，可以压缩信息。