# 第 2 章 离散信源及其信息度量

1、同时掷两个正常的骰子，也就是各面呈现的概率都是 l/6，求：

(1) “3 和 5 同时出现”事件的自信息量；

(2)“两个 1 同时出现”事件的自信息量；

(3)两个点数的各种组合（无序对）的熵或平均信息量；

(4) 两个点数之和(即 2，3，…，12 构成的子集）的熵；

(5)两个点数中至少有一个是 1 的自信息。

解：（1）P（3、5 或 5、3）＝P（3、5）+P（5、3）＝1/18

I＝log2（18）＝ 4.1699bit。

（2）P（1、1）＝l/36。I＝log2（36）＝5.1699bit。

（3）相同点出现时（11、22、33、44、55、66）有 6 种，概率 1/36。

不同点出现时有 15 种，概率 1/18。

H（i，j）＝6*1/36*$\log_2$（36）+15*1/18*$\log_2$（18）＝4.3366bit/事件。

（4）

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| i+j | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| P(i+j) | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |

H(i+j)=H(1/36 2/36 3/36 4/36 5/36 6/36 5/36 4/36 3/36 2/36 1/36)

=3.2744bit/事件。

（5）P（1、1or1、j or i、1）＝1/36+5/36+5/36＝11/36。

I＝log2（36/11）＝1.7105bit/

2、 居住某地区的女孩中有 25％是大学生，在女大学生中有 75％身高为 1.6m 以上，而女孩中身高 1.6m 以上的占总数一半。假如得知“身高 1.6m 以上的某女孩是大学生”的消息，问获得多少信息量？、

解：P（女大学生）＝1/4；P（身高>1.6m / 女大学生）=3/4；P（身高>1.6m）＝1/2；

P（女大学生 / 身高>1.6m）＝P（身高>1.6m、女大学生）/P（身高>1.6m ）

＝3/4*1/4*2＝3/8

I＝log2（8/3）＝1.4150bit。

3、两个实验 $X=\{x_1,x_2,x_3\}$ 和 $Y=\{y_1,y_2,y_3\}$，联合概率 $p(x_iy_j)=p_{ij}$ 为

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 7/24 & 1/24 & 0 \\ 1/24 & 1/4 & 1/24 \\ 0 & 1/24 & 7/24 \end{bmatrix}$$

（1）如果有人告诉你 $X$ 和 $Y$ 的实验结果，你得到的平均信息量是多少？

（2）如果有人告诉你 $Y$ 的实验结果，你得到的平均信息量是多少？

（3）在已知 $Y$ 的实验结果的情况下，告诉你 $X$ 的实验结果，你得到的平均信息量是多少？

解：

| P(x,y) | | Y | | | .x |
|---|---|---|---|---|---|
| | | y1 | y2 | y3 | |
| X | x1 | 7/24 | 1/24 | 0 | 1/3 |
| | x2 | 1/24 | 1/4 | 1/24 | 1/3 |
| | x3 | 0 | 1/24 | 7/24 | 1/3 |
| .y | | 1/3 | 1/3 | 1/3 | |

（1）
$$\begin{aligned} H(X,Y) &= -\sum_{i=1}^{3}\sum_{j=1}^{3} p(x_i, y_j)\log P(x_i, y_j) \\ &= 2.301 bit/symbol \end{aligned}$$

（2）
$$\begin{aligned} H(Y) &= -\sum_{j=1}^{3} p(y_j)\log p(y_j) \\ &= 1.5894 bit/symbol \end{aligned}$$

（3）
$$\begin{aligned} H(X \mid Y) &= H(X,Y) - H(Y) \\ &= 2.301 - 1.5894 \\ &= 0.7151 bit/symbol \end{aligned}$$

4、某一无记忆信源的符号集为$\{0,1\}$，已知 $p_0 = 1/4$， $p_1 = 3/4$。

（1）求信源符号的平均信息量；

（2）由 100 个符号构成的序列，求某一特定序列（例如有 $m$ 个 0 和$100-m$ 个 1）的信息量的表达

（3）计算（2）中的序列熵。

解：（1）因为信源是无记忆信源，所以符号的平均熵

$$H(X) = H\left(\frac{1}{4},\frac{3}{4}\right) = \frac{1}{4}\times 2 + \frac{3}{4}\times 0.415 = 0.81 \text{bit}/\text{符号}$$

（2）某一特定序列（例如：m 个 0 和 100-m 个 1）出现的概率为

$$P(X^L) = P(X_1, X_2, \cdots, X_{100}) = [P(0)]^m [P(1)]^{100-m} = \left(\frac{1}{4}\right)^m \left(\frac{3}{4}\right)^{100-m}$$

所以，自信息量为

$$I\left(X_1,X_2,\cdots,X_{100}\right)=-\log P\left(X^L\right)=-\log\left\{\left(\frac{1}{4}\right)^m\left(\frac{3}{4}\right)^{100\text{-}m}\right\}$$

$$=200-(100-m)\log_2 3 \quad (bit)$$

（3）序列的熵

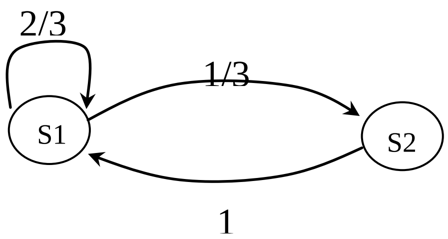
$$\mathrm{H}\left(\mathrm{X}^{\mathrm{L}}\right)=100\mathrm{H}(\mathrm{X})=81\mathrm{bit}/\text{序列}$$

5、　有一个马尔可夫信源，已知转移概率为

$$P(S_1\,|\,S_1)=\frac{2}{3},\,P(S_2\,|\,S_1)=\frac{1}{3},\,P(S_1\,|\,S_2)=1,\,P(S_2\,|\,S_2)=0\text{。}$$

试画出状态转移图，并求出信源熵。

解：（1）由题意可得状态转移图




由状态转移图可知：该马尔可夫链具有遍历性，平稳后状态的极限分布存在。

一步转移矩阵 $\mathrm{P}=\begin{bmatrix}\frac{2}{3} & \frac{1}{3}\\ 1 & 0\end{bmatrix}$

由 $\sum_i \mathrm{W_i p_{ij}}=\mathrm{W_j}$ 和 $\sum_j \mathrm{p_{ij}}=1$ 可得方程组

$$\begin{cases}\mathrm{W_1}=\frac{2}{3}\mathrm{W_1}+\mathrm{W_2}\\ \mathrm{W_2}=\frac{1}{3}\mathrm{W_1}\\ \mathrm{W_1}+\mathrm{W_2}=1\end{cases}$$

解方程组得到各状态的稳态分布概率 $\begin{cases}\mathrm{W_1}=3/4\\ \mathrm{W_2}=1/4\end{cases}$，

因为 $\mathrm{H}(\mathrm{X}/\mathrm{S_1})=\mathrm{H}\left(\frac{2}{3},\frac{1}{3}\right),\quad \mathrm{H}(\mathrm{X}/\mathrm{S_2})=\mathrm{H}(1,0)=0$，

所以信源的熵

$$\mathrm{H}_\infty(\mathrm{X})=\sum_i \mathrm{p(s_i)H(X/s_i)}=\frac{3}{4}\mathrm{H}\left[\frac{2}{3},\frac{1}{3}\right]=\frac{3}{4}\times 0.92=0.69\mathrm{bit}/\text{符号}$$

6、有一个一阶马尔可夫链 $X_1, X_2, \cdots, X_r, \cdots$, 各 $X_r$ 取值于集 $A=\{a_1, a_2, \cdots, a_q\}$，已知起始概率为 $p_1 = P(X_1 = x) = \dfrac{1}{2}, p_2 = p_3 = \dfrac{1}{4}$，其转移概率如下：

| $i \backslash j$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 1/2 | 1/4 | 1/4 |
| 2 | 2/3 | 0 | 1/3 |
| 3 | 2/3 | 1/3 | 0 |

（1）求 $X_1 X_2 X_3$ 的联合熵和平均符号熵；

（2）求这个链的极限平均符号熵；

（3）求 $H_0$、$H_1$、$H_2$ 和它们对应的冗余度。

解：（1）

<u>方法一、</u>

因为 $P(x_1 x_2 x_3) = P(x_1) P(x_2/x_1) P(x_3/x_1 x_2) = P(x_1) P(x_2/x_1) P(x_3/x_2)$

可以计算得到

$P(a_1 a_1 a_1) = P(a_1) P(a_1/a_1) P(a_1/a_1) = \dfrac{1}{8}$, $\quad P(a_1 a_2 a_1) = \dfrac{1}{12}$,

$P(a_1 a_1 a_2) = P(a_1) P(a_1/a_1) P(a_2/a_1) = \dfrac{1}{16}$, $\quad P(a_1 a_2 a_2) = 0$,

$P(a_1 a_1 a_3) = P(a_1) P(a_1/a_1) P(a_3/a_1) = \dfrac{1}{16}$, $\quad P(a_1 a_2 a_3) = \dfrac{1}{24}$,

$P(a_1 a_3 a_1) = \dfrac{1}{12}$, $\quad P(a_2 a_1 a_1) = \dfrac{1}{12}$, $\quad P(a_2 a_2 a_1) = 0$, $\quad P(a_2 a_3 a_1) = \dfrac{1}{18}$,

$P(a_1 a_3 a_2) = \dfrac{1}{24}$, $\quad P(a_2 a_1 a_2) = \dfrac{1}{24}$, $\quad P(a_2 a_2 a_2) = 0$, $\quad P(a_2 a_3 a_2) = \dfrac{1}{36}$,

$P(a_1 a_3 a_3) = 0$, $\quad P(a_2 a_1 a_3) = \dfrac{1}{24}$, $\quad P(a_2 a_2 a_3) = 0$, $\quad P(a_2 a_3 a_3) = 0$,

$P(a_3a_1a_1)=\frac{1}{12}$, $P(a_3a_2a_1)=\frac{1}{18}$, $P(a_3a_3a_1)=0$,

$P(a_3a_1a_2)=\frac{1}{24}$, $P(a_3a_2a_2)=0$, $P(a_3a_3a_2)=0$,

$P(a_3a_1a_3)=\frac{1}{24}$, $P(a_3a_2a_3)=\frac{1}{36}$, $P(a_3a_3a_3)=0$,

所以，

$$H(X_1X_2X_3)=-\sum_{X_1}\sum_{X_2}\sum_{X_3}P(x_1x_2x_3)\log P(x_1x_2x_3)$$

$$=\frac{1}{8}\times\log 8+2\times\frac{1}{16}\times\log 16+4\times\frac{1}{12}\times\log 12+6\times\frac{1}{24}\times\log 24+2\times\frac{1}{18}\log 18+2\times\frac{1}{36}\times\log 36$$

$=3.967$bit/三个符号

所以，平均符号熵 $H_3(X^3)=\frac{1}{3}H(X_1X_2X_3)=1.322\text{bit}/$符号

方法二、

$$H(X_1X_2X_3)=H(X_1)+H(X_2/X_1)+H(X_3/X_2)$$

$=1.5+1.209+1.26$

$=3.967$bit/三个符号

所以，平均符号熵 $H_3(X^3)=\frac{1}{3}H(X_1X_2X_3)=1.322\text{bit}/$符号

（2）因为这个信源是一阶马尔可夫链，其状态极限概率分布就是信源达到平稳后的符号概率分布.

由题意得到一步转移矩阵 $P=\begin{bmatrix}\frac{1}{2}&\frac{1}{4}&\frac{1}{4}\\ \frac{2}{3}&0&\frac{1}{3}\\ \frac{2}{3}&\frac{1}{3}&0\end{bmatrix}$

由 $\sum_i W_ip_{ij}=W_j$ 和 $\sum_j p_{ij}=1$ 可得方程组

$$\begin{cases}W_1=\frac{1}{2}W_1+\frac{2}{3}W_2+\frac{2}{3}W_3\\ W_2=\frac{1}{4}W_1+\frac{1}{3}W_3\\ W_3=\frac{1}{4}W_1+\frac{1}{3}W_2\\ W_1+W_2+W_3=1\end{cases}$$

解方程组得到各状态的稳态分布概率$\begin{cases} W_1 = 4/7 \\ W_2 = 3/14 \\ W_3 = 3/14 \end{cases}$，

所以信源平稳后的概率分布为$\begin{cases} P(a_1) = 4/7 \\ P(a_2) = 3/14 \\ P(a_3) = 3/14 \end{cases}$

因为信源为一阶马尔可夫信源，所以信源的熵

$$H_\infty(X) = H_2 = H(X_2/X_1)$$
$$= \frac{4}{7}H\left[\frac{1}{2},\frac{1}{4},\frac{1}{4}\right] + \frac{3}{14}H\left[\frac{2}{3},0,\frac{1}{3}\right] + \frac{3}{14}H\left[\frac{2}{3},\frac{1}{3},0\right] = 1.251\text{bit/符号}$$

（3）$H_0 = \log 3 = 1.585\text{bit/符号}$

$$H_1 = H\left[\frac{4}{7},\frac{3}{14},\frac{3}{14}\right] = 1.414\text{bit/符号}$$

$$H_2 = H(X_2/X_1) = 1.251\text{bit/符号}$$

$$H_\infty = H(X_2/X_1) = 1.251\text{bit/符号}$$

对应的冗余度分别为

$$\gamma_0 = 1 - \frac{H_0}{H_0} = 0$$

$$\gamma_1 = 1 - \frac{H_1}{H_0} = 0.054$$

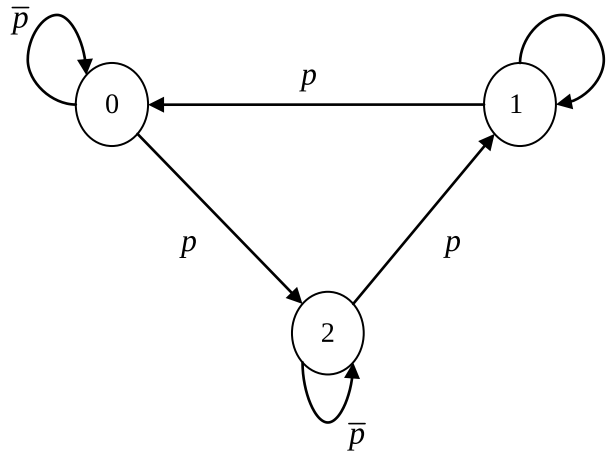
$$\gamma_2 = 1 - \frac{H_2}{H_0} = 0.145$$

7、 一阶马尔可夫信源的状态如图所示，信源 $X$ 的符号集为$\{0,1,2\}$。

（1）求平稳后的信源的概率分布；

（2）求信源熵$H_\infty$；

（3）求当 $p=0$ 和 $p=1$ 时信源的熵，并说明其理由。




解：（1）由状态转移图可得状态一步转移矩阵

$$P=\begin{bmatrix}\overline{p} & 0 & p\\ p & \overline{p} & 0\\ 0 & p & \overline{p}\end{bmatrix}$$

由状态转移图可知：该马尔可夫链具有遍历性，平稳后状态的极限分布存在。

由$\sum_i W_i p_{ij}=W_j$和$\sum_j p_{ij}=1$可得方程组

$$\begin{cases}W_1=\overline{p}W_1+pW_2\\ W_2=\overline{p}W_2+pW_3\\ W_3=pW_1+\overline{p}W_3\\ W_1+W_2+W_3=1\end{cases}$$

解方程组得到各状态的稳态分布概率$\begin{cases}W_1=1/3\\ W_2=1/3\\ W_3=1/3\end{cases}$，

所以信源平稳后的概率分布为$\begin{cases}p(0)=1/3\\ p(1)=1/3\\ p(2)=1/3\end{cases}$

（2）因为信源为一阶马尔可夫信源，所以信源的熵

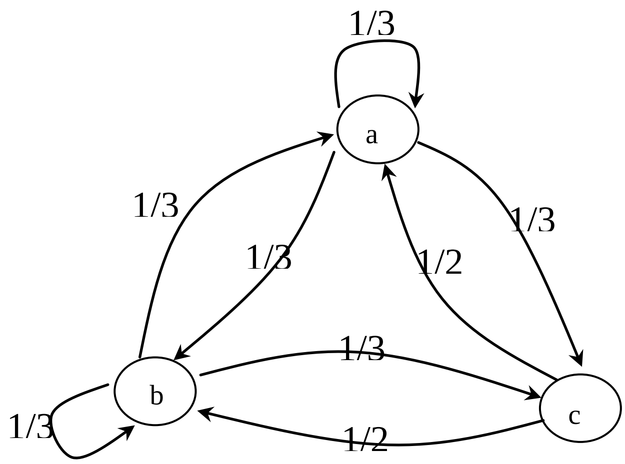
$$H_\infty(X)=\sum_i p(s_i)H(X/s_i)=p(0)H(X/0)+p(1)H(X/1)+p(2)H(X/2)$$

$$=\frac{1}{3}H[\overline{p},0,p]+\frac{1}{3}H[p,\overline{p},0]+\frac{1}{3}H[0,p,\overline{p}]=H[\overline{p},p]=H[p]$$

（3）当$p=0$或$p=1$时，信源的熵为 0。因为此时它表明信源从某一状态出发转移到另一状态的情况是一定发生或一定不发生，即是确定的事件。

8、设有一信源，它在开始时以 $P(a)=0.6, P(b)=0.3, P(c)=0.1$ 的概率发出 $X_1$，如果 $X_1$ 为 $a$ 时，则 $X_2$ 为 $a,b,c$ 的概率为 $\frac{1}{3}$；如果 $X_1$ 为 $b$ 时，则 $X_2$ 为 $a,b,c$ 的概率为 $\frac{1}{3}$；如果 $X_1$ 为 $c$ 时，则 $X_2$ 为 $a,b$ 概率为 $\frac{1}{2}$，为 $c$ 的概率为 0。而且后面发出 $X_i$ 的概率只与 $X_{i-1}$ 有关。有 $P(X_i \mid X_{i-1}) = P(X_2 \mid X_1),\ i \ge 3$。试利用马尔可夫信源的图示法画出状态转移图，并且计算信源熵 $H_\infty$。

解：（1）由题目可知，这个信源为一阶马尔可夫信源，状态空间就等于信源符号集合{a,b,c},其状态转移图为




（2）由状态转移图可知：该马尔可夫链具有遍历性，平稳后状态的极限分布存在。

一步转移矩阵 $P = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$

由 $\sum_i W_i p_{ij} = W_j$ 和 $\sum_j p_{ij} = 1$ 可得方程组

$$\begin{cases} W_1 = \frac{1}{3}W_1 + \frac{1}{3}W_2 + \frac{1}{2}W_3 \\ W_2 = \frac{1}{3}W_1 + \frac{1}{3}W_2 + \frac{1}{2}W_3 \\ W_3 = \frac{1}{3}W_1 + \frac{1}{3}W_2 \\ W_1 + W_2 + W_3 = 1 \end{cases}$$

解方程组得到各状态的稳态分布概率$\begin{cases} W_1 = 3/8 \\ W_2 = 3/8 \\ W_3 = 1/4 \end{cases}$，

因为信源为一阶马尔可夫信源，所以信源的熵

$$H_\infty(X) = \sum_i p(s_i)H(X/s_i) = p(a)H(X/a) + p(b)H(X/b) + p(c)H(X/c)$$

$$= \frac{3}{8}H\left[\frac{1}{3},\frac{1}{3},\frac{1}{3}\right] + \frac{3}{8}H\left[\frac{1}{3},\frac{1}{3},\frac{1}{3}\right] + \frac{1}{4}H\left[\frac{1}{2},\frac{1}{2}\right] = 1.439\text{bit}/\text{符号}$$

# 第 3 章 离散信道及其信道容量

1、设二进制对称信道的概率转移矩阵为$\begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix}$,

(1)若 $p(x_0) = 3/4, p(x_1) = 1/4$，求 H(X), H(X|Y), H(Y|X)和 I(X;Y).

(2)求该信道的信道容量及其达到信道容量时的输入概率分布.

解:(1) $H(X) = -\sum_{i=0}^{1} p(x_i)\log p(x_i) = 0.811$ bit/符号

$$p(y_0) = p(x_0)p(y_0 \mid x_0) + p(x_1)p(y_0 \mid x_1) = 7/12$$

$$p(y_1) = 1 - p(y_0) = 5/12$$

$$H(Y) = -\sum_{i=0}^{1} p(y_i)\log p(y_i) = 0.980 \text{ bit/符号}$$

$$H(Y \mid X) = -\sum_{i=0}^{1}\sum_{j=0}^{1} p(x_i)p(y_j \mid x_i)\log p(y_j \mid x_i) = 0.918 \text{ bit/符号}$$

$$H(XY) = H(X) + H(Y \mid X) = 1.729 \text{ bit/符号}$$

$$H(X \mid Y) = H(XY) - H(Y) = 0.749 \text{ bit/符号}$$

$$I(X;Y) = H(X) - H(X \mid Y) = 0.062 \text{ bit/符号}$$

(2)该信道为对称 DMC 信道,因此在 $p(x_0) = p(x_1) = 1/2$时,取得信道容量为:

$$C = \log 2 + \sum_{j=1}^{2} p_{ij}\log p_{ij} = 0.082 \text{ bit/符号}$$

2、 设某对称离散信道的信道矩阵为

$$\mathbf{P}=\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

求其信道容量。

解：

$$C=\log s-H(\mathbf{P}\text{的行矢量})$$

$$=\log 4-H(\frac{1}{3},\frac{1}{3},\frac{1}{6},\frac{1}{6})=2+\frac{1}{3}\log\frac{1}{3}+\frac{1}{3}\log\frac{1}{3}+\frac{1}{6}\log\frac{1}{6}+\frac{1}{6}\log\frac{1}{6}$$

$$=0.0817 \qquad \text{比特/符号}$$

在这个信道中，每个符号平均能够传输的最大信息为 0.0817 比特，而且只有当信道输入是等概分布时才能达到这个最大值。

3、设有扰信道的传输情况分别如图 3.17 所示。试求这种信道的信道容量。

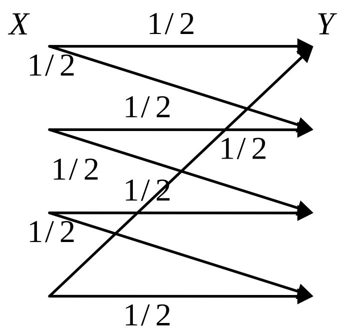



图 3.17　题 3.13 的信道转移图

解：该信道对应的转移矩阵为：

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

从该转移矩阵中可看出，该信道是对称信道，故其信道容量为：

$$C=H(X)-H(\frac{1}{2},\frac{1}{2},0,0)=\log 4-\left[\frac{1}{2}\log 2+\frac{1}{2}\log 2\right]=1bit/\text{符号}$$

4、 求下列两个信道的信道容量，并加以比较。

（1）$\begin{bmatrix} \bar{p}-\varepsilon & p-\varepsilon & 2\varepsilon \\ p-\varepsilon & \bar{p}-\varepsilon & 2\varepsilon \end{bmatrix}$ （2）$\begin{bmatrix} \bar{p}-\varepsilon & p-\varepsilon & 2\varepsilon & 0 \\ p-\varepsilon & \bar{p}-\varepsilon & 0 & 2\varepsilon \end{bmatrix}$

其中 $p+\bar{p}=1$。

解：1)该信道为准对称信道，将其划分成两个互不相交的子集：

$$\begin{bmatrix} 1-p-\varepsilon & p-\varepsilon \\ p-\varepsilon & 1-p-\varepsilon \end{bmatrix}, \quad \begin{bmatrix} 2\varepsilon \\ 2\varepsilon \end{bmatrix}$$

直接应用准对称信道的信道容量公式进行计算：

$$C_1 = \log r - H(1-p-\varepsilon, p-\varepsilon, 2\varepsilon) - \sum_{k=1}^{2} N_k \log M_k$$

其中 r=2, $N_1 = M_1 = 1-2\varepsilon, N_2 = 2\varepsilon, M_2 = 4\varepsilon$ ,所以

$$C_1 = \log 2 + (1-p-\varepsilon)\log(1-p-\varepsilon) + (p-\varepsilon)\log(p-\varepsilon) + 2\varepsilon\log 2\varepsilon$$
$$-(1-2\varepsilon)\log(1-2\varepsilon) - 2\varepsilon\log 4\varepsilon$$
$$=(1-2\varepsilon)\log\frac{2}{1-2\varepsilon} + (1-p-\varepsilon)\log(1-p-\varepsilon) + (p-\varepsilon)\log(p-\varepsilon)$$

输入概率等概率时达到此信道容量。

2)此信道也是准对称信道，将其划分成两个互不相交的子集：

$$\begin{bmatrix} 1-p-\varepsilon & p-\varepsilon \\ p-\varepsilon & 1-p-\varepsilon \end{bmatrix}, \quad \begin{bmatrix} 2\varepsilon & 0 \\ 0 & 2\varepsilon \end{bmatrix}$$

这两矩阵为对称矩阵，其中 $r=2, N_1 = M_1 = 1-2\varepsilon, N_2 = M_2 = 2\varepsilon$

所以，

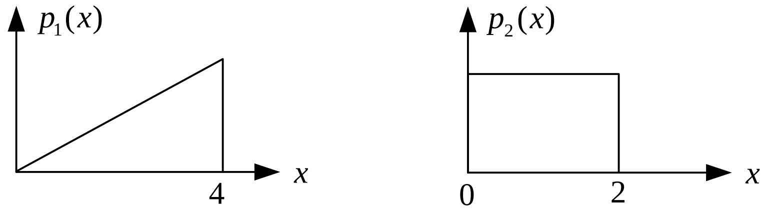
$$C_2 = \log 2 + (1-p-\varepsilon)\log(1-p-\varepsilon) + (p-\varepsilon)\log(p-\varepsilon) + 2\varepsilon\log 2\varepsilon$$
$$-(1-2\varepsilon)\log(1-2\varepsilon) - 2\varepsilon\log 2\varepsilon$$
$$=(1-2\varepsilon)\log\frac{2}{1-2\varepsilon} + (1-p-\varepsilon)\log(1-p-\varepsilon) + (p-\varepsilon)\log(p-\varepsilon) + 2\varepsilon\log 2$$
$$= C_1 + 2\varepsilon$$

输入等概时，达到此信道容量，比较两个信道容量，可知它们之间的关系为

$$C_2 = C_1 + 2\varepsilon$$

# 第 4 章 连续信源与连续信道

1 两个一维随机变量的概率分布密度函数分别如图 6.7(a)和(b)所示。问哪一个熵大？




(a)　　　　　　　　(b)

图 6.7　习题 6.1 图

解：（a）$h(X)=-\int_{\mathbf{R}} p(x)\log p(x)dx=\int_0^4 \frac{x}{8}\log\frac{x}{8}dx=1.72$ bit

（b）$h(X)=-\int_a^b \frac{1}{b-a}\log\frac{1}{b-a}dx=\log(b-a)$ =1bit

所以图（a）所示一维随机变量的概率分布密度函数所对应的连续信源的熵较大。

2 计算机终端通过带宽为 3400Hz 的信道传输数据。

（1）设要求信道的 $S/N=30dB$，试求该信道的信道容量是多少？

（2）设线路上的最大信息传输速率为 4800b/s，试求所需最小信噪比为多少？

解：(1) 因为 S/N =30dB,即 $10\log_{10}\frac{S}{N}=30dB$，

得：S/N=1000

由香农公式得信道容量

$$C=B\log_2(1+\frac{S}{N})$$

$$=3400\times\log_2(1+1\quad 1$$

$$\approx 33.89\times 10^3\ bit$$

（2）因为最大信息传输速率为 4800b/s，即信道容量为 4800b/s。由香农公式

$$C=B\log_2(1+\frac{S}{N})$$

得：$\frac{S}{N}=2^{\frac{C}{B}}-1=2^{\frac{4800}{3400}}-1\approx 2.66-1=1.66$。

则所需最小信噪比为 1.66。

3 已知某一信道的信息传输速率为 $6kbit/s$，噪声功率谱 $\frac{N_0}{2}=10^{-4}$ W/Hz，在带宽为 6kHz 的高斯信道中进行传输。试计算无差错传输需要的最小输入信号功率。当带宽为 6MHz 时，重新计算最小输入信号功率。

解：当 $C\geq R_t$ 时，才能无差错传输。即

$$C = B\log(1+SNR) = B\log(1+\frac{P_s}{N_0 B}) \ge R_t$$

所以，当带宽为 6kHz 时，最小输入信号功率为

$$P_s \ge N_0 B\left\{\exp\left(\frac{R_t}{B}\right)-1\right\} = 2.052\text{W}$$

当带宽为 6MHz 时，最小输入信号功率为

$$P_s \ge N_0 B\left\{\exp\left(\frac{R_t}{B}\right)-1\right\} = 1.2\times 10^3\left(e^{0.001}-1\right)$$

4 已知一个平均功率受限的连续信号，通过带宽 $B = 1MHz$ 的高斯白噪声信道，试求

（1）若信噪比为 10，信道容量为多少？

（2）若信道容量不变，信噪比降为 5，信道带宽应为多少？

（3）若信道通频带减为 0.5MHz 时，要保持相同的信道容量，信道上的信号与噪声的平均功率比值应等于多大？

解：1)信道容量为：$C = B\log(1+\frac{S}{N}) = \log(1+10) = 3.46 Mbits/s$

2)在信道容量不变的条件下，信噪比变为 5，其带宽为：

$$B = \frac{C}{\log(1+\frac{S}{N})} = \frac{3.46}{\log(1+5)} = \frac{3.46}{2.58} = 1.34MHz$$

3)在信道容量不变的条件下，带宽变为 0.5MHz，其信噪比变为：

$$C = 0.5\times\log(1+\frac{S}{N}) = 3.46 \Rightarrow 1+\frac{S}{N} = 121 \Rightarrow \frac{S}{N} = 120$$

# 第 5 章 无失真信源编码

1、对信源$\begin{bmatrix} \mathrm{S} \\ \mathrm{P} \end{bmatrix} = \begin{bmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 \\ 0.2 & 0.19 & 0.18 & 0.17 & 0.15 & 0.10 & 0.01 \end{bmatrix}$进行二元编码，编码方案为

| 信息符号 | 对应码字 |
|---|---|
| $s_1$ | 000 |
| $s_2$ | 001 |
| $s_3$ | 011 |
| $s_4$ | 100 |
| $s_5$ | 101 |
| $s_6$ | 1110 |
| $s_7$ | 1111110 |

（1）计算平均码长$\overline{\mathrm{L}}$；

（2）编码后信息传输率$\mathrm{R}$；

（3）编码信息率$\mathrm{R}'$；

（4）编码效率$\eta$。

解：（1）$\overline{\mathrm{L}}=\sum_{i=1}^{q}p(s_i)\cdot L_i=3.14$（码元/信源符号）

（2）$\mathrm{H}(\mathrm{S})=2.61$（比特/信源符号）

$R=\dfrac{H(\mathrm{S})}{\overline{\mathrm{L}}}=\dfrac{2.61}{3.14}=0.831$（bit/码元）

（3）$\mathrm{R}'=\overline{\mathrm{L}}\log r$=3.14（ bit/信源符号）

（4）$\eta=\dfrac{\mathrm{R}}{\mathrm{R}_{\max}}=0.831$

或者$\eta=\dfrac{\mathrm{H}(\mathrm{S})}{\mathrm{R}'}=0.831$

2、 设离散无记忆信源的概率空间为$\begin{bmatrix}S\\P\end{bmatrix}=\begin{bmatrix}s_1 & s_2\\ \frac{3}{4} & \frac{1}{4}\end{bmatrix}$，若对信源采取等长二元编码，要求编码效率$\eta=0.96$，允许译码错误概率$\delta\leq 10^{-5}$，试计算需要的信源序列长度N为多少？

解：信源熵为

$H(\mathrm{S})=\frac{1}{4}\log 4+\frac{3}{4}\log\frac{4}{3}=0.811$（bit/符号）

自信息量的方差

$$\sigma^2(\mathrm{S})=\sum_{i=1}^{q}p_i(\log p_i)^2-[\mathrm{H}(\mathrm{S})]^2=\frac{3}{4}\left(\log\frac{3}{4}\right)^2+\frac{1}{4}\left(\log\frac{1}{4}\right)^2-0.811^2=0.4715$$

因为编码效率$\eta=0.96$，由

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

可得

$$\varepsilon=\frac{1-\eta}{\eta}\mathrm{H}(\mathrm{S})=\frac{0.04}{0.96}\times 0.811=0.3379$$

可得

$$\mathrm{N}\geq\frac{\sigma^2(\mathrm{S})}{\varepsilon^2\delta}=\frac{0.4715}{0.3379^2\times 10^{-5}}=4.13\times 10^7$$

所以，信源序列长度达到$4.13\times10^{7}$以上，才能实现给定的要求，因此等长编码没有实际的意义，一般统计编码都是采用不等长编码。

3、设离散无记忆信源的概率空间为$\begin{bmatrix} S \\ P \end{bmatrix}=\begin{bmatrix} s_1 & s_2 \\ 0.9 & 0.1 \end{bmatrix}$，对信源进行 N 次扩展，采用霍夫曼编码。当 N=1，2，∞时的平均码长和编码效率为多少？

解：

（1）N=1 时，将$s_1$编成 0，$s_2$编成 1，则

$$L_1=1$$

又因为信源熵

$$H(S)=-\sum_{i=1}^{q}p(s_i)\log p(s_i)=0.469\ \text{bit/符号}$$

所以

$$\eta_1=\frac{H(S)}{L_1}=0.469$$

（2）N=2 时，编码过程如下

| $S^2$ | 概率 | 霍夫曼编码 |
|---|---|---|
| $s_1s_1$ | 0.81 | 1 |
| $s_1s_2$ | 0.09 | 01 |
| $s_2s_1$ | 0.09 | 000 |
| $s_2s_2$ | 0.01 | 001 |

所以

$$\overline{L}_2=1\times0.81+2\times0.09+3\times(0.01+0.09)=$$

则

$$\frac{\overline{L}_2}{2}=0.645$$

所以

$$\eta_2=\frac{H(X)}{0.645}=$$

（3）N=∞时，由香农第一定理可知，必然存在唯一可译码，使

$$\lim_{N\to\infty}\frac{\overline{L}_N}{N}=H_r(S)$$

而霍夫曼编码为最佳码，即平均码长最短的码，故

$$\lim_{N\to\infty}\frac{\overline{L}_N}{N}=H_r(S)=H(S)=0.469$$

即

$$\lim_{N\to\infty}\eta_N=1$$

4、已知信源共 7 个符号消息，其概率空间为

$$\begin{bmatrix} S \\ P(x) \end{bmatrix}=\begin{bmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 \\ 0.2 & 0.19 & 0.18 & 0.17 & 0.15 & 0.10 & 0.01 \end{bmatrix}$$

试进行香农编码。并计算编码后的信息传输率和编码效率。

解：下面以消息 $s_5$ 为例来介绍香农编码。

计算 $-log(p_5)=-\log 0.15=2.74$，取整数 $L_5=3$ 作为 $s_5$ 的码长。计算 $s_1,s_2,s_3,s_4$ 的累加概率，有

$$P_5=\sum_{k=1}^{4}p_k=0.2+0.19+0.18+0.17=0.74$$

将 0.74 变换成二进制小数 $(0.74)_{10}=(0.1011110)_2$，取小数点后面三位 101 作为 $s_5$ 的代码。

其余消息的代码可以用相同的方法计算得到，如表所示。

| 信息符号 | 符号概率 | 累加概率 | $-\log p_i$ | 码字长度 | 码字 |
|---|---|---|---|---|---|
| $s_1$ | 0.20 | 0 | 2.34 | 3 | 000 |
| $s_2$ | 0.19 | 0.20 | 2.41 | 3 | 001 |
| $s_3$ | 0.18 | 0.39 | 2.48 | 3 | 011 |
| $s_4$ | 0.17 | 0.57 | 2.56 | 3 | 100 |
| $s_5$ | 0.15 | 0.74 | 2.74 | 3 | 101 |
| $s_6$ | 0.10 | 0.89 | 3.34 | 4 | 1110 |
| $s_7$ | 0.01 | 0.99 | 6.6 | 7 | 11 |

| | | | | | |
|---|---|---|---|---|---|
| | | | 6 | | 11110 |

信源熵： $H(S)=-\sum_{i=1}^{q}p(s_i)\log p(s_i)=2.61\text{bit/符号}$

平均码长： $\overline{L}=\sum_{i=1}^{q}p(s_i)\cdot L_i=3.14\text{码元/符号}$

信息传输率 $R=\dfrac{H(S)}{\overline{L}}=\dfrac{2.61}{3.14}=0.831$

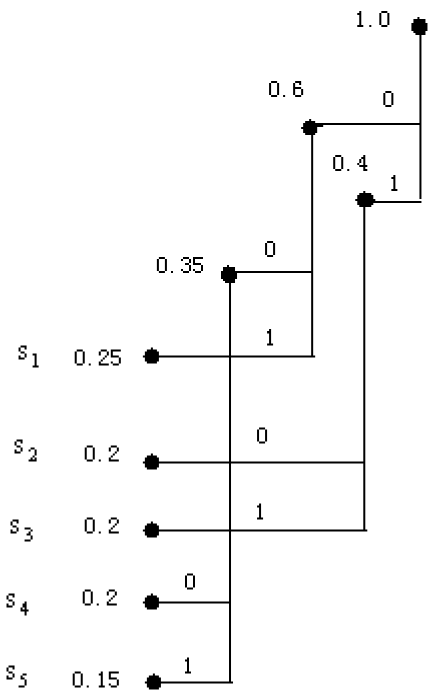
编码效率： $\eta=\dfrac{H(S)}{\overline{L}}=\dfrac{2.61}{3.14}=0.831$

5、已知信源

$$\begin{bmatrix}S\\P\end{bmatrix}=\begin{bmatrix}s_1 & s_2 & s_3 & s_4 & s_5\\0.25 & 0.2 & 0.2 & 0.2 & 0.15\end{bmatrix}$$

用霍夫曼编码法编成二进制变长码，计算平均码长和编码效率。

解：




| 信源符号 | 码字 | 码长 |
|---|---|---|
| $s_1$ | 01 | 2 |
| $s_2$ | 10 | 2 |
| $s_3$ | 11 | 2 |
| $s_4$ | 000 | 3 |
| $s_5$ | 001 | 3 |

平均码长： $\overline{L}=\sum_{i=1}^{q}p(s_i)\cdot L_i=0.65*2+0.35*3=2.36\text{码元/符号}$

信息传输率 $R=\dfrac{H(S)}{\overline{L}}$

编码效率： $\eta=\dfrac{H(S)}{\overline{L}}$

6、已知信源

$$\begin{bmatrix} S \\ P \end{bmatrix} = \begin{bmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \\ 0.32 & 0.22 & 0.18 & 0.16 & 0.08 & 0.04 \end{bmatrix}$$

（1）求符号熵 H(S)。

（2）用香农编码法编成二进制变长码，计算其编码效率。

（3）用费诺编码法编成二进制变长码，计算其编码效率。

（4）用霍夫曼编码法编成二进制变长码，计算其编码效率。

（5）用霍夫曼编码法编成三进制变长码，计算其编码效率。

（6）若用逐个信源符号来编定长二进制码，要求不出差错译码，求所需要的每符号的平均信息率和编码效率。

（7）当译码差错小于$10^{-3}$的定长二进制码要达到（4）中的霍夫曼编码效率时，估计要多少个信源符号一起编才能办到？

解：（1）$H(S) = -\sum_{i=1}^{6} p(s_i)\log p(s_i) = 2.352$ bit/符号

（2）香农编码过程如下：

| $s_i$ | $p(s_i)$ | 累加概率 $P_i$ | $-\log p(s_i)$ | $L_i$ | 码字 |
|---|---|---|---|---|---|
| $s_1$ | 0.32 | 0 | 1.644 | 2 | 00 |
| $s_2$ | 0.22 | 0.32 | 2.184 | 3 | 010 |
| $s_3$ | 0.18 | 0.54 | 2.474 | 3 | 100 |
| $s_4$ | 0.16 | 0.72 | 2.644 | 3 | 101 |
| $s_5$ | 0.08 | 0.88 | 3.644 | 4 | 1110 |
| $s_6$ | 0.04 | 0.96 | 4.644 | 5 | 11110 |

$$\overline{L} = \sum_{i=1}^{6} p(s_i)L_i = 2.84 \text{ 码元/符号}$$

$$\eta = \frac{H(S)}{\overline{L}} = 0.828$$

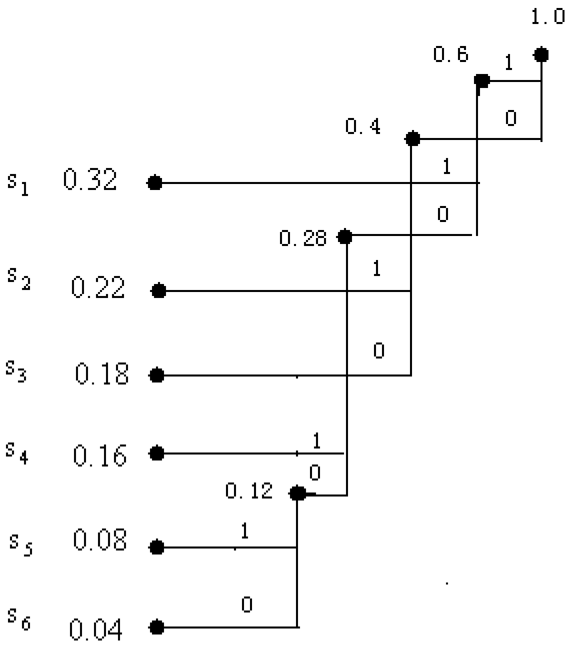
（3）费诺编码过程如下：

<table>
<tr><th>消息符号</th><th>消息概率</th><th>第一次分组</th><th>第二次分组</th><th>第三次分组</th><th>第四次分组</th><th>二元码字</th><th>码长</th></tr>
<tr><td>$s_1$</td><td>0.32</td><td rowspan="2">0</td><td>0</td><td rowspan="3"></td><td rowspan="4"></td><td>00</td><td>2</td></tr>
<tr><td>$s_2$</td><td>0.22</td><td>1</td><td>01</td><td>2</td></tr>
<tr><td>$s_3$</td><td>0.18</td><td rowspan="4">1</td><td>0</td><td>10</td><td>2</td></tr>
<tr><td>$s_4$</td><td>0.16</td><td rowspan="3">1</td><td>0</td><td>110</td><td>3</td></tr>
<tr><td>$s_5$</td><td>0.08</td><td rowspan="2">1</td><td>0</td><td>1110</td><td>4</td></tr>
<tr><td>$s_6$</td><td>0.04</td><td>1</td><td>1111</td><td>4</td></tr>
</table>

$$\overline{\mathrm{L}} = \sum_{\mathrm{i}=1}^{6} p(\mathrm{s}_i)\mathrm{L}_i = 2.4 \text{ 码元/符号}$$

$$\eta = \frac{H(\mathrm{S})}{\overline{\mathrm{L}}} = 0.98$$

（4）霍夫曼二进制编码过程如下：

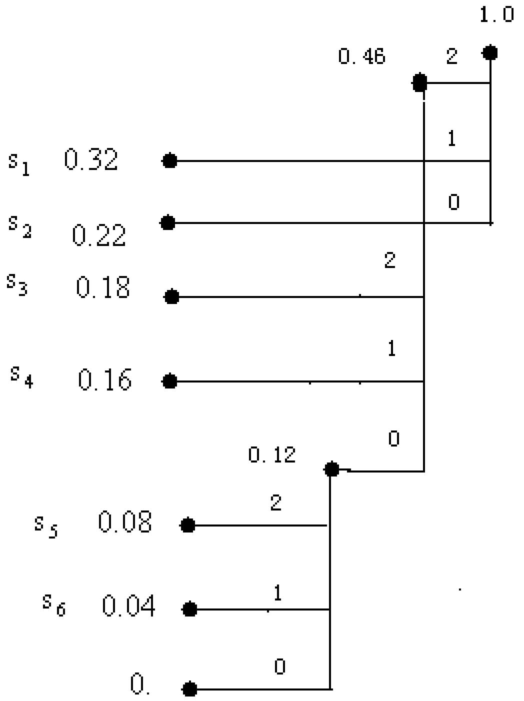



| 信源符号 | 码字 | 码长 |
|---|---|---|
| $s_1$ | 11 | 2 |
| $s_2$ | 01 | 2 |
| $s_3$ | 00 | 2 |
| $s_4$ | 101 | 3 |
| $s_5$ | 1001 | 4 |
| $s_6$ | 1000 | 4 |

$$\overline{\mathrm{L}} = \sum_{\mathrm{i}=1}^{6} p(\mathrm{s}_i)\mathrm{L}_i = 2.4 \text{ 码元/符号}$$

$$\eta = \frac{H(\mathrm{S})}{\overline{\mathrm{L}}} = 0.98$$

（5）霍夫曼三进制编码过程如下：




| 信源符号 | 码字 | 码长 |
|---|---|---|
| $s_1$ | 1 | 1 |
| $s_2$ | 0 | 1 |
| $s_3$ | 22 | 2 |
| $s_4$ | 21 | 2 |
| $s_5$ | 202 | 3 |
| $s_6$ | 201 | 3 |

按 3 为底的信源熵为：

$$H'(\mathrm{S}) = -\sum_{i=1}^{6} p(x_i)\log_3 p(x_i) = 1.484 \text{ 单位/符号}$$

$$\overline{\mathrm{L}} = \sum_{\mathrm{i}=1}^{6} p(\mathrm{s}_i)\mathrm{L}_i = (0.32+0.22)+(0.18+0.16)\times 2+(0.08+0.04)\times 3 = 1.58 \text{ 码元/符号}$$

$$\eta = \frac{H'(\mathrm{S})}{\overline{\mathrm{L}}} = 0.939$$

（6）因为有 6 个符号，因此需要 3bit 来进行编码，故需要的符号平均信息率为 3bit/符号。

$$\eta = \frac{H(\mathrm{S})}{\overline{\mathrm{L}}} = 0.784$$

（7）$\eta = \dfrac{H(\mathrm{S})}{H(\mathrm{S})+\varepsilon} = 0.98$ 解得 $\varepsilon = 0.048$

$$\sigma^2(\mathrm{S}) = \sum_{i=1}^{6} p_i(\log p_i)^2 - [H(\mathrm{S})]^2 = 0.527 bit^2$$

$$L \ge \frac{\sigma^2(\mathrm{S})}{\varepsilon^2\delta} = \frac{0.527}{0.048^2\times 10^{-3}} = 2.29\times 10^5$$

因此，使用定长码则需要 $2.29\times 10^5$ 个符号连在一起编。

7、 设信源 S 的 N 次扩展信源为$S^N$，采用最佳编码对它进行编码，而码符号为

$X=\{x_1,x_2,\cdots x_r\}$，编码后所得的码符号可以看作一个新信源

$$\begin{bmatrix} X \\ P \end{bmatrix}=\begin{bmatrix} x_1 & x_2 & \cdots & x_r \\ p_1 & p_2 & \cdots & p_r \end{bmatrix}$$

求证：当$N\to\infty$时，新信源 X 符号集的概率分布趋于等概分布。

证明：由香农第一定理知

$$\frac{H(S)}{\log r}\le\frac{\overline{L}_N}{N}\le\frac{H(S)}{\log r}+\frac{1}{N}$$

当平均码长达到极限值时，编码效率

$$\frac{\overline{L}_N}{N}=\frac{H(S)}{\log r}$$

这时编码后的信道信息传输率（新信源的信息传输率）

$$R=\log r$$

即新信源的 r 个符号（即码符号）独立等概分布，达到最大熵。

8、设某二元无记忆信源

$$\begin{bmatrix} S \\ P \end{bmatrix}=\begin{bmatrix} 0 & 1 \\ 0.8 & 0.2 \end{bmatrix}$$

每秒钟发出 2.5 个信源符号。将此信源的输出符号送入无噪无损信道中进行传输，而信道每秒钟只传送两个二元符号。

（1） 如果不通过编码，信源能否在此信道中进行无失真传输？。

（2） 如果通过适当编码，信源能否在此信道中进行无失真传输？如何进行信源编码？

解：

（1）如果不对信源进行编码，这时信源输出$2.5$（二元信源符号/秒），而信道传输速率为 2（二元信道符号/秒），因为$2.5>2$，所以不通过编码，该信源不能在信道中进行无失真编码。

（2）信源熵

$$H(S)=0.722bit/\text{信源符号}$$

二元无噪无损信道的最大信息传输率

$$C=1\ \text{比特}/\text{信道符号}$$

而信道每秒钟传送 2 个符号，所以该信道的最大信息传输速率

$$C_t=2\ \text{比特}/\text{秒}$$

如果信源每秒钟发送 2.5 个信源符号，则信源输出的信息速率

$$R_t = 2.5 \times H(S) = 1.805 \text{比特/秒}$$

则

$$R_t < C_t$$

所以，通过适当编码，信源能够在此信道中进行无失真传输。如何进行编码呢？我们将对 N 次扩展信源进行信源编码。当 N=2 时，对二次扩展信源进行霍夫曼编码。编码过程如下

| $S^2$ | 概率 | 霍夫曼编码 |
|---|---|---|
| $s_1s_1$ | 0.64 | 1 |
| $s_1s_2$ | 0.16 | 01 |
| $s_2s_1$ | 0.16 | 000 |
| $s_2s_2$ | 0.04 | 001 |

则单个符号的平均码长

$$\overline{L} = \frac{\overline{L}_2}{2} = 0.78 \text{二元码符号/信源符号}$$

所以，二次扩展编码后，送入信道的传输速率为

$0.78 \times 2.5 = 1.95$（二元码符号/秒）

信源编码得到的二元码符号进入信道，即信道符号就是二元码符号，由题意可知，信道每秒钟可以传送两个符号。因为$1.95 < 2$，此时就可以在信道中进行无失真编码。

# 第 6 章　有噪信道编码

1、假设二进制对称信道的差错率$P_e = 10^{-1}$，当（5，1）重复码通过该信道时，采用大数准则进行译码，译码错误概率为多少？

解：$P_E = 1 - [(1-P_e)^5 + \binom{5}{1} p_e (1-P_e)^4 + \binom{5}{2} P_e^2 (1-P_e)^3] = 7 \times 10^{-3}$

2、已知一线性码的全部码字为$(000000)$、$(001110)$、$(010101)$、$(011011)$、$(100011)$、$(101101)$、$(110110)$、$(111000)$。

（1）计算码率（假设码字等概率分布）；

（2）计算最小汉明距离；

（3）若用于检错，能检出几位错码？

（4）若用于纠错，能纠正几位错码？

解：（1）码率

$$R = \frac{\log M}{n} = \frac{\log 8}{5} = \frac{3}{5}$$

（2）因为线性码的最小距离等于非零码字的最小码重。所以最小距离为 3。

（3）因为最小距离为 3，所以可以检测 2 位错码。

（4）因为最小距离为 3，所以可以纠正 1 位错码。

3、设线性码的生成矩阵为

$$\boldsymbol{G} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

（1）确定 $(n,k)$ 码中的 $n$ 和 $k$。

（2）写出监督矩阵。

（3）写出该 $(n,k)$ 码的全部码字。

（4）说明纠错能力。

解：（1）因为生成矩阵为 3 行 6 列，所以 n=6，k=3

（2）监督矩阵为

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

（3）因为 $\boldsymbol{C} = \boldsymbol{M}\,\boldsymbol{G}$，可以得到该 $(n,k)$ 码的全部码字为：

000000，001011，010110，011101，100101，101110，110011，111000。

（4）因为线性码的最小距离等于非零码字的最小码重。所以最小距离为 3，所以可以纠正 1 位错码。

4、已知（7，4）线性分组码的生成矩阵为

$$\boldsymbol{G} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

（1） 写出标准的生成矩阵和监督矩阵。

（2） 写出监督子与错码位置的对应关系。

（3） 如果接收码字为（1111111），（1010111），试计算监督子，并进行译码。

解：（1）$G=\begin{bmatrix}1&0&0&0&1&0&1\\0&1&0&0&1&1&1\\0&0&1&0&1&1&0\\0&0&0&1&0&1&1\end{bmatrix}$，$H=\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$

（2）监督子与错码位置的对应关系

| $S_1S_2S_3$ | 错码位置 | | $S_1S_2S_3$ | 错码位置 |
|---|---|---|---|---|
| 001 | $r_0$ | | 110 | $r_4$ |
| 010 | $r_1$ | | 111 | $r_5$ |
| 100 | $r_2$ | | 101 | $r_6$ |
| 011 | $r_3$ | | 000 | 无错 |

（4） 如果接收码字为（1111111），则监督子

$$S_1S_2S_3=RH^T=(1111111)\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}=000$$

所以无错，译码结果为 1111111。

如果接收码字为（1010111），则监督子

$$S_1S_2S_3=RH^T=(1010111)\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}=100$$

所以 $r_2$ 错，译码结果为 1010011。

5、设某二元无记忆信源

$$\begin{bmatrix}S\\P\end{bmatrix}=\begin{bmatrix}s_1&s_2\\0.8&0.2\end{bmatrix}$$

将此信源的输出符号送入无噪无损信道中进行传输，而信道每秒钟只传送两个二元符号。试问

（1） 信源每秒钟发出 2.5 个信源符号。不通过编码，信源能够和信道直接连接；

（2） 信源每秒钟发出 2.5 个信源符号。通过适当编码，信源是否能够在此信道中进行无失真传输？试说明如何进行适当编码。如何进行信源编码？

（3） 信源每秒钟发出 3 个信源符号，信源是否能够在此信道中进行无失真传输？如果不能，允许失真度为多少时，信源就可以在此信道中传输（假设失真函数采用汉明失真）？

解：

（1）如果不对信源进行编码，这时信源输出 2.5（二元信源符号/秒），而信道传

输速率为 2（二元信道符号/秒），因为$2.5 > 2$，所以不通过编码，该信源不能在信道中进行无失真编码。

（2）信源熵

$$H(S) = 0.722 bit / \text{信源符号}$$

二元无噪无损信道的最大信息传输率

$$C = 1 \ \text{比特} / \text{信道符号}$$

而信道每秒钟传送 2 个符号，所以该信道的最大信息传输速率

$$C_t = 2 \ \text{比特} / \text{秒}$$

如果信源每秒钟发送 2.5 个信源符号，则信源输出的信息速率

$$R_t = 2.5 \times H(S) = 1.805 \text{比特/秒}$$

则

$$R_t < C_t$$

所以，通过适当编码，信源能够在此信道中进行无失真传输。如何进行编码呢？我们将对 N 次扩展信源进行信源编码。当 N=2 时，对二次扩展信源进行霍夫曼编码。编码过程如下

| $S^2$ | 概率 | 霍夫曼编码 |
|---|---|---|
| $s_1 s_1$ | 0.64 | 1 |
| $s_1 s_2$ | 0.16 | 01 |
| $s_2 s_1$ | 0.16 | 000 |
| $s_2 s_2$ | 0.04 | 001 |

则单个符号的平均码长

$$\overline{L} = \frac{\overline{L}_2}{2} = 0.78 \text{二元码符号} / \text{信源符号}$$

所以，二次扩展编码后，送入信道的传输速率为

$$0.78 \times 2.5 = 1.95 \text{（二元码符号/秒）}$$

信源编码得到的二元码符号进入信道，即信道符号就是二元码符号，由题意可知，信道每秒钟可以传送两个符号。因为$1.95 < 2$，此时就可以在信道中进行无失真编码。

（2）如果信源每秒钟发送 3 个信源符号，则信源输出的信息速率

$$R_t = 3 \times H(S) = 2.166 \ \text{比特/秒}$$

则

$$R_t > C_t$$

所以，信源不能够在此信道中进行无失真传输。

因为该信源为二元信源，失真函数采用汉明失真，所以

$$R(D)=H(0.8,0.2)-H(D)=0.722-H(D)\text{（比特/信源符号）}$$

$$R_t(D)=3\times R(D)\text{（比特/秒）}$$

当

$$R_t(D)\le C_t$$

则此信源不会因为信道而增加新的失真，总的失真就是信源压缩而造成的允许失真 $D$ 。此时

$$3\times[0.722-H(D)]\le 2$$

即

$$H(D)\approx 0.055$$

$$D\approx 0.06$$

所以，允许失真度 $D\approx 0.06$ 时，信源就可以在此信道中传输。

# 第 7 章 限失真信源编码

1、设信源 $U$＝{0,1}，接收变量 $V$＝{0,1,2}，定义失真函数为 $d(0,0)$＝$d(1,1)$＝0，$d(0,1)$＝$d(1,0)$＝1，$d(0,2)$＝$d(1,2)$＝0.5，计算失真矩阵。

解：由题意得到失真矩阵为

$$\boldsymbol{D}=\begin{bmatrix}0 & 1 & 0.5\\ 1 & 0 & 0.5\end{bmatrix}$$

2、设有删余信源符号集 $U=\{u_1,u_2,\cdots,u_r\}$，接收符号集 $V=\{v_1,v_2,\cdots,v_s\}$， $s=r+1$ 。定义它的单个符号失真度为

$$d(u_i,v_j)=\begin{cases}0 & i=j\\ 1 & i\neq j, j\neq s\\ 1/2 & i\neq j, j=s\end{cases}$$

其中，接收信号 $v_s$ 作为一个删除符号，计算 r=3 时的失真矩阵。

解：由题意得到失真矩阵为

$$\boldsymbol{D}=\begin{bmatrix}0 & 1 & 1 & \frac{1}{2}\\ 1 & 0 & 1 & \frac{1}{2}\\ 1 & 1 & 0 & \frac{1}{2}\end{bmatrix}$$

3、 设有对称信源（$r=s=4$），其失真函数$d(u_i,v_j)=(u_i-v_j)^2$，求失真矩阵。

解：由题意得到失真矩阵为

$$\boldsymbol{D}=\begin{bmatrix}0 & 1 & 4 & 9\\ 1 & 0 & 1 & 4\\ 4 & 1 & 0 & 1\\ 9 & 4 & 1 & 0\end{bmatrix}$$

4、 已知无记忆信源

$$\begin{bmatrix}U\\ P(u)\end{bmatrix}=\begin{bmatrix}0 & 1 & 2\\ 0.2 & 0.3 & 0.5\end{bmatrix}$$

失真矩阵为

$$\boldsymbol{D}=\begin{bmatrix}4 & 2 & 1\\ 0 & 3 & 2\\ 2 & 0 & 1\end{bmatrix}$$

计算$D_{min}$和$D_{max}$。

解：

$$D_{min}=\sum_{i=1}^{r}P(u_i)\min_j d(u_i,v_j)=0.2\times 1=0.2$$

$$D_{max}=\min_V\sum_U P(u)d(u,v)=\min(1.8,\ \ 1.3,\ \ 1.3)=1.3$$

5、 假设离散无记忆信源输出 N 维随机序列$\boldsymbol{U}=(U_1U_2U_3)$，其中$U_i(i=1,2,3)$取自符号集$\{0,1\}$，通过信道传输到信宿，接收 N 维随机序列$\boldsymbol{V}=(V_1V_2V_3)$，其中$V_i(i=1,2,3)$取自符号集$\{0,1\}$，定义失真函数

$$\begin{aligned}d(0,0)=d(1,1)=0\\ d(0,1)=d(1,0)=\alpha\end{aligned}$$

求符号序列的失真矩阵。

解：由 N 维信源序列的失真函数的定义得

$$\mathrm{d_N}(\boldsymbol{u},\boldsymbol{v}) = \mathrm{d_N}(\alpha_{\mathrm{i}},\beta_{\mathrm{j}}) = \frac{1}{\mathrm{N}}\sum_{l=1}^{\mathrm{N}} \mathrm{d}(\mathrm{u}_{i_l},\mathrm{v}_{j_l}) \qquad \boldsymbol{u}\in\boldsymbol{U},\boldsymbol{v}\in\boldsymbol{V}$$

所以

$$\mathrm{d_N}(000,000) = \frac{1}{3}[\mathrm{d}(0,0)+\mathrm{d}(0,0)+\mathrm{d}(0,0)] = 0$$

$$\mathrm{d_N}(000,001) = \frac{1}{3}[\mathrm{d}(0,0)+\mathrm{d}(0,0)+\mathrm{d}(0,1)] = \frac{\alpha}{3}$$

类似计算其他元素值，得到信源序列的失真矩阵为

$$\boldsymbol{D}_{\mathrm{N}} = \begin{bmatrix} 0 & \frac{\alpha}{3} & \frac{\alpha}{3} & \frac{2\alpha}{3} & \frac{\alpha}{3} & \frac{2\alpha}{3} & \frac{2\alpha}{3} & \alpha \\ \frac{\alpha}{3} & 0 & \frac{2\alpha}{3} & \frac{\alpha}{3} & \frac{2\alpha}{3} & \frac{\alpha}{3} & \alpha & \frac{2\alpha}{3} \\ \frac{\alpha}{3} & \frac{2\alpha}{3} & 0 & \frac{\alpha}{3} & \frac{2\alpha}{3} & \alpha & \frac{\alpha}{3} & \frac{2\alpha}{3} \\ \frac{2\alpha}{3} & \frac{\alpha}{3} & \frac{\alpha}{3} & 0 & \alpha & \frac{2\alpha}{3} & \frac{2\alpha}{3} & \frac{\alpha}{3} \\ \frac{\alpha}{3} & \frac{2\alpha}{3} & \frac{2\alpha}{3} & \alpha & 0 & \frac{\alpha}{3} & \frac{\alpha}{3} & \frac{2\alpha}{3} \\ \frac{2\alpha}{3} & \frac{\alpha}{3} & \alpha & \frac{2\alpha}{3} & \frac{\alpha}{3} & 0 & \frac{2\alpha}{3} & \frac{\alpha}{3} \\ \frac{2\alpha}{3} & \alpha & \frac{\alpha}{3} & \frac{2\alpha}{3} & \frac{\alpha}{3} & \frac{\alpha}{3} & 0 & \frac{\alpha}{3} \\ \alpha & \frac{2\alpha}{3} & \frac{2\alpha}{3} & \frac{\alpha}{3} & \frac{2\alpha}{3} & \frac{\alpha}{3} & \frac{\alpha}{3} & 0 \end{bmatrix}$$

6、 设某二元无记忆信源

$$\begin{bmatrix} \mathrm{S} \\ P \end{bmatrix} = \begin{bmatrix} \mathrm{s}_1 & \mathrm{s}_2 \\ \frac{3}{4} & \frac{1}{4} \end{bmatrix}$$

每秒钟发出 3 个信源符号。将此信源的输出符号送入无噪无损信道中进行传输，而信道每秒钟只传送两个二元符号。

（1） 试问信源能否直接在此信道中进行无失真传输？

（2） 如此信源失真度测度定义为汉明失真，问允许信源平均失真多大时，此信源就是否可以在此信道中传输？

解：（1）如果不对信源进行编码，这时信源输出 3（二元信源符号/秒），而信道传输速率为 2（二元信道符号/秒），因为$3>2$，所以不通过编码，该信源不能在信道中进行无失真编码。

（2）信源熵

$$H(\mathrm{S}) = \frac{1}{4}\log 4 + \frac{3}{4}\log\frac{4}{3} = 0.811 bit/\text{信源符号}$$

如果信源每秒钟发送 3 个信源符号，则信源输出的信息速率

$$\mathrm{R_t} = 3\times \mathrm{H}(\mathrm{S}) = 2.433 \ \text{比特/秒}$$

二元无噪无损信道的最大信息传输率

$$C = 1 \text{ 比特/信道符号}$$

而信道每秒钟传送 2 个符号，所以该信道的最大信息传输速率

$$C_t = 2 \text{ 比特/秒}$$

则

$$R_t > C_t$$

所以，信源不能够在此信道中进行无失真传输。

因为该信源为二元信源，失真函数采用汉明失真，所以

$$R(D) = H\left(\frac{1}{4}, \frac{3}{4}\right) - H(D) = 0.811 - H(D) \text{（比特/信源符号）}$$

$$R_t(D) = 3 \times R(D) \text{（比特/秒）}$$

当

$$R_t(D) \le C_t$$

则此信源不会因为信道而增加新的失真，总的失真就是信源压缩而造成的允许失真 $D$ 。此时

$$3 \times [0.811 - H(D)] \le 2$$

即

$$H(D) \approx 0.144$$

$$D \approx 0.02$$

所以，允许失真度 $D \approx 0.02$ 时，信源就可以在此信道中传输。