有噪信道编码

信息论与编码原理不挂科 第七讲

有噪信道编码

信息论与编码原理不挂科 第七讲

4大模块

4道题目

信息论与编码原理不挂科 第七讲

有噪信道编码

- **模块1** 信道编码概述
- **模块2** 错误概率与译码规则
- **模块3** 错误概率与编码方法
- **模块4** 有噪信道编码定理

信道编码概述

小节1 信道编码概述

信道编码概述

- 信道编码的目标：**提高通信系统的可靠性**，尽可能将消息通过传输前后所发生的错误率降到最低。
- 信道编码：按照一定的规则给信源编码的码符号序列增加一定的**冗余信息**，使其变为具有一定数学规律的码符号序列。
- 信道译码：接收到码符号序列后，按照与信道编码器相同的数学规律，去掉符号序列中的冗余符号。

本章探究如何让使有噪信道传输错误最低，对应的最大信息传输率是多少。

错误概率与译码规则

与错误概率有关的因素

在有噪信道中，传输消息会发生错误，为减少错误，提高可靠性，首先分析与错误概率有关的因素。

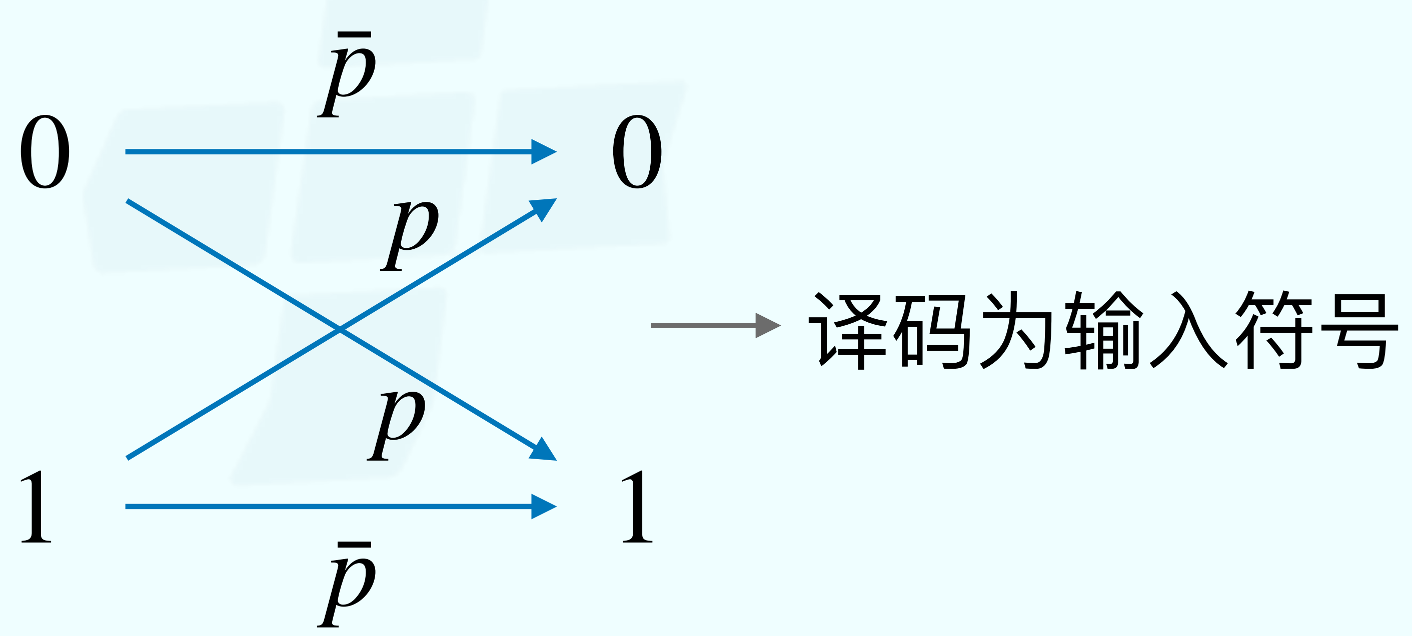
- 错误概率与**信道统计特性**有关，用信道的输入概率与信道传递矩阵（正确/错误传递概率）描述。
- 错误概率与**译码过程**有关：译码过程和译码规则对系统的错误概率影响很大。

错误概率与译码规则

译码规则对错误概率的影响

以二进制对称信道（BSC）为例，设输入分布为 $P_X = \begin{bmatrix} 0 & 1 \\ 0.3 & 0.7 \end{bmatrix}$，

信道传递矩阵 $P = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$，



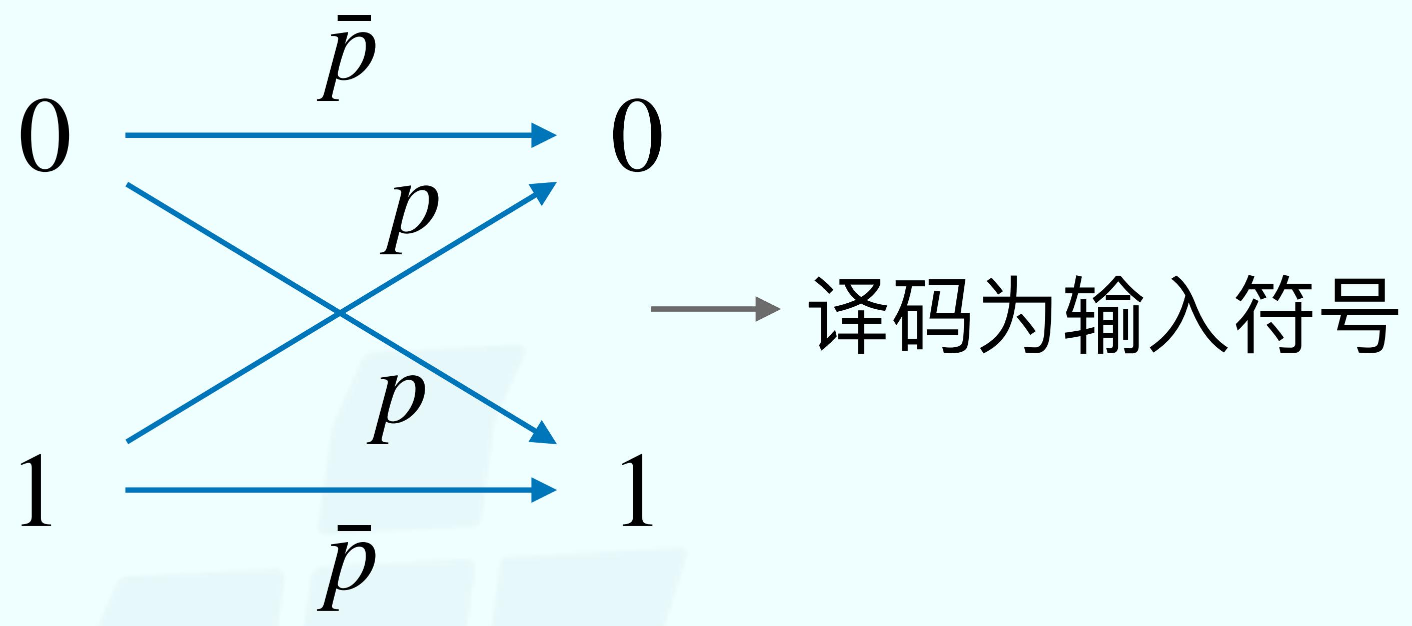
则传输错误概率为 $p_e = p(01) + p(10) = p(0)p(1|0) + p(1)p(0|1) = 0.3 \times 0.1 + 0.7 \times 0.1 = 0.1$，

传输正确概率为 $\bar{p}_e = p(00) + p(11) = p(0)p(0|0) + p(1)p(1|1) = 0.3 \times 0.9 + 0.7 \times 0.9 = 0.9$。

译码规则对错误概率的影响

以二进制对称信道（BSC）为例，设输出分布为 $P_x = \begin{bmatrix} 0 & 1 \\ 0.3 & 0.7 \end{bmatrix}$，

信道传递矩阵 $P = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix}$。



若确定如下两种译码规则：

- 规则1：信道译码器收到信号"0" → 译为"0"，则正确译码概率为 0.9；
  信道译码器收到信号"1" → 译为"1"，则正确译码概率为 0.9。
- 规则2：信道译码器收到信号"0" → 译为"1"，则正确译码概率为 0.1；
  信道译码器收到信号"1" → 译为"0"，则正确译码概率为 0.1。

规则1 的正确概率 $\bar{p}_e = p(00) + p(11) = 0.3 \times 0.9 + 0.7 \times 0.9 = 0.9$，

规则2 的正确概率 $\bar{p}_e = p(01) + p(10) = 0.3 \times 0.1 + 0.7 \times 0.1 = 0.1$。

可见，译码规则不同，正确或错误概率也就不同。

错误概率与译码规则

译码规则

设信道的输入符号集为 $X=\{x_1,x_2,\ldots,x_r\}$，输出符号集为 $Y=\{y_1,y_2,\ldots,y_s\}$，若对每一个输出符号 y_j，都有一个确定的函数 $F(y_j)$ 对应于**唯一的一个输入符号** x_i，则称此函数为译码规则。

$$F(y_j)=x_i\ (i=1,2,\ldots,r\ ;j=1,2,\ldots,s)$$

每个输出有 r 种译码可能，故共有 r^s 种译码规则。

错误概率与译码规则

错误概率

设译码规则为 $F(y_j)=x_i \longrightarrow \begin{cases} \text{当输入符号为 } x_i \text{ 时，译码正确。} \\ \text{当输入符号除 } x_i \text{ 以外的 } (r-1) \text{ 种符号时，译码错误。} \end{cases}$

此时正确译码的概率为 $p[F(y_j)\,|\,y_j]=p(x_i\,|\,y_j)$，

错误译码的概率为 $p(e\,|\,y_j)=1-p(x_i\,|\,y_j)=1-p[F(y_j)\,|\,y_j]=\sum_{k\neq i}p(x_k\,|\,y_j)$。

上式取平均，可求得平均正确译码概率与平均错误译码概率：

- 平均错误译码概率 $P_E=\sum_{j=1}^{s}p(y_j)p(e\,|\,y_j)=\sum_{j=1}^{s}p(y_j)\{1-p[F(y_j)\,|\,y_j]\}$

$$=1-\sum_{j=1}^{s}p(y_j)p[F(y_j)\,|\,y_j]=1-\sum_{j=1}^{s}p(x_iy_j)$$

- 平均正确译码概率 $\bar{P}_E=\sum_{j=1}^{s}p(y_j)p[F(y_j)\,|\,y_j]=\sum_{j=1}^{s}p(y_j)p(x_i\,|\,y_j)=\sum_{j=1}^{s}p(x_iy_j)$

错误概率与译码规则

译码规则的选取原则

译码规则的选取规则：使平均错误译码概率 P_E 最小。

$P_E = \sum_{j=1}^{s} p(y_j)p(e|y_j) = 1 - \sum_{j=1}^{s} p(y_j)p[F(y_j)|y_j]$　　P_E 最小时，需让$p[F(y_j)|y_j]$最大化。

我们给出两个常用的重要译码规则：**最大后验概率准则**与**极大似然译码准则**。

最大后验概率（MAP）准则

令 $F(y_j)=x^*, x^*\in X$， 而 x^* 应满足 $p(x^*|y_j)\geq p(x_i|y_j), x_i\in X, x_i\neq x^*$，

则称上述条件的译码函数对应的译码规则为最大后验概率准则（或最小错误概率译码准则）。

极大似然译码（ML）准则

最大后验概率 $p(x^*|y_j)$ 通常未知，我们从而介绍极大似然译码准则。

令 $F(y_j)=x^*, x^*\in X$，而 x^* 应满足 $p(x^*)p(y_j|x^*)\geq p(x_i)p(y_j|x_i), x_i\in X, x_i\neq x^*$，**当输入符号等概分布时，** $p(y_j|x^*)\geq p(y_j|x_i)$，此时可直接从信道矩阵的传递概率中选定译码函数，即收到 y_j 后译码成矩阵 P 中第 j 列最大元素对应的信源符号。

错误概率分析

$$P_E = \sum_{j=1}^{s} p(y_j)p(e|y_j) = \sum_{j=1}^{s} p(y_j)[1 - p(x^*|y_j)]$$

$$= 1 - \sum_{Y,X^*} p(y_j)p(x_i|y_j) = \boxed{1 - \sum_{Y,X^*} p(x_i y_j)}$$

MAP准则下，需利用联合概率矩阵使 P_E 最小化，此时 P_E 为联合概率矩阵除去每列最大元素之外的其他元素之和。

$$= 1 - \sum_{Y,X^*} p(y_j|x_i)p(x_i) = \sum_{Y,\underline{X-X^*}} p(y_j|x_i)p(x_i)$$

对 $F(y_j) = x^*$ 之外

当输入等概分布时，$P_E = 1 - \frac{1}{r}\sum_{Y,X} p(y_j|x_i) = \frac{1}{r}\sum_{Y,X-X^*} p(y_j|x_i)$ 。

ML准则下，需利用信道矩阵使 P_E 最小化，选取信道矩阵中每一列的最大元素作为译码规则，P_E 为其余元素之和乘以 $\frac{1}{r}$ 。

两种译码规则的使用方法

- MAP准则：$p(x^* | y_j) \geq p(x_i | y_j)$
  - 转移概率矩阵每行乘以 $p(x_i)$ 得**联合概率分布矩阵**；
  - 以**每列**矩阵中的**最大概率**对应的输入概率 x_i 作为译码准则；
  - 所有译码结果对应的联合概率之和为正确概率，矩阵中其余元素的和为错误概率。

- ML准则：$p(y_j | x^*) \geq p(y_j | x_i)$
  - **输入符号为等概分布**或**先验概率未知**时采用；
  - 以**转移概率矩阵**每列中的**最大的一个元素**对应的 x_i 作为译码准则；
  - 所有译码准则所对应的转移概率之和**乘以** $\frac{1}{r}$ 为正确概率，其余矩阵元素之和**乘以** $\frac{1}{r}$ 为错误概率。

例题7-1 设信道矩阵为 $P=\begin{bmatrix}0.5 & 0.3 & 0.2\\ 0.3 & 0.3 & 0.4\\ 0.2 & 0.3 & 0.5\end{bmatrix}$，在输入符号为下列情况时，求译码规则与平均错误译码概率。

（1）输入符号概率分布为 $\begin{bmatrix}\frac{1}{3} & \frac{1}{2} & \frac{1}{6}\end{bmatrix}$。

（2）输入符号等概分布。

解析7-1 （1）输入符号不等概，采用最大后验概率译码准则。

$$P_{XY}=\begin{bmatrix}0.5\times\frac{1}{3} & 0.3\times\frac{1}{3} & 0.2\times\frac{1}{3}\\ 0.3\times\frac{1}{2} & 0.3\times\frac{1}{2} & 0.4\times\frac{1}{2}\\ 0.2\times\frac{1}{6} & 0.3\times\frac{1}{6} & 0.5\times\frac{1}{6}\end{bmatrix}=\begin{array}{c c}& \begin{array}{ccc} y_1 & y_2 & y_3\end{array}\\ \begin{array}{c}x_1\\ x_2\\ x_3\end{array} & \begin{bmatrix}\frac{1}{6} & \frac{1}{10} & \frac{1}{15}\\ \frac{3}{20} & \frac{3}{20} & \frac{1}{5}\\ \frac{1}{30} & \frac{1}{20} & \frac{1}{12}\end{bmatrix}\end{array}$$

译码规则为 $\begin{cases}F(y_1)=x_1\\ F(y_2)=x_2\\ F(y_3)=x_2\end{cases}$。

$$P_E=\sum_{Y,X-X^*}p(x_iy_j)=(\frac{3}{20}+\frac{1}{30})+(\frac{1}{10}+\frac{1}{20})+(\frac{1}{15}+\frac{1}{12})=0.483$$

（或 $P_E=1-(\frac{1}{6}+\frac{3}{20}+\frac{1}{5})=0.483$）

例题7-1 设信道矩阵为 $P=\begin{bmatrix}0.5 & 0.3 & 0.2\\ 0.3 & 0.3 & 0.4\\ 0.2 & 0.3 & 0.5\end{bmatrix}$，在输入符号为下列情况时，求译码规则与平均错误译码概率。

（1）输入符号概率分布为 $\begin{bmatrix}\frac{1}{3} & \frac{1}{2} & \frac{1}{6}\end{bmatrix}$。

（2）输入符号等概分布。

解析7-1 （2）输入符号等概分布时，采用极大似然译码准则。

$$P=\begin{bmatrix}0.5 & 0.3 & 0.2\\ 0.3 & 0.3 & 0.4\\ 0.2 & 0.3 & 0.5\end{bmatrix}$$

译码规则为 $\begin{cases}F(y_1)=x_1\\ F(y_2)=x_2\\ F(y_3)=x_3\end{cases}$。

$$P_E=\frac{1}{r}\sum_{Y,X-X^*}p(y_j|x_i)=\frac{1}{3}[(0.3+0.2)+(0.3+0.3)+(0.2+0.4)]\approx 0.57$$

（或 $P_E=1-\frac{1}{3}(0.5+0.3+0.5)=0.57$）

例题7-2 设有一离散信道，其信道传递矩阵为$\begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$，并设 $p(x_1)=\frac{1}{2}$ ，$p(x_2)=p(x_3)=\frac{1}{4}$ 。试分别按照最小错误概率译码准则与最大似然译码准则确定译码规则并计算平均错误概率。

解析7-2 MAP准则：求出联合概率分布

$$P_{XY}=\begin{bmatrix} \frac{1}{2}\times\frac{1}{2} & \frac{1}{3}\times\frac{1}{2} & \frac{1}{6}\times\frac{1}{2} \\ \frac{1}{6}\times\frac{1}{4} & \frac{1}{2}\times\frac{1}{4} & \frac{1}{3}\times\frac{1}{4} \\ \frac{1}{3}\times\frac{1}{4} & \frac{1}{6}\times\frac{1}{4} & \frac{1}{2}\times\frac{1}{4} \end{bmatrix}=\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} \frac{1}{4} & \frac{1}{6} & \frac{1}{12} \\ \frac{1}{24} & \frac{1}{8} & \frac{1}{12} \\ \frac{1}{12} & \frac{1}{24} & \frac{1}{8} \end{bmatrix} \end{array}$$

译码规则为 $\begin{cases} F(y_1)=x_1 \\ F(y_2)=x_1 \\ F(y_3)=x_3 \end{cases}$ 。

$$P_E=\sum_{Y,X-X^*}p(x_iy_j)=(\frac{1}{24}+\frac{1}{12})+(\frac{1}{8}+\frac{1}{24})+(\frac{1}{12}+\frac{1}{12})=\frac{11}{24}$$

（或 $P_E=1-(\frac{1}{4}+\frac{1}{6}+\frac{1}{8})=\frac{11}{24}$ ）

例题7-2 设有一离散信道，其信道传递矩阵为$\begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$，并设 $p(x_1)=\frac{1}{2}$，$p(x_2)=p(x_3)=\frac{1}{4}$。试分别按照最小错误概率译码准则与最大似然译码准则确定译码规则并计算平均错误概率。

解析7-2

ML准则：$P=\begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$

译码规则为$\begin{cases} F(y_1)=x_1 \\ F(y_2)=x_2 \\ F(y_3)=x_3 \end{cases}$。

$$P_E=\frac{1}{r}\sum_{Y,X-X^*}p(x_i|y_j)=\frac{1}{3}[(\frac{1}{6}+\frac{1}{3})+(\frac{1}{3}+\frac{1}{6})+(\frac{1}{6}+\frac{1}{3})]=\frac{1}{2}$$

错误概率与译码规则

费诺不等式

平均错误译码概率 P_E 与译码准则有关，译码准则与信道特性有关，故 P_E 与信道统计特性存在联系，且平均错误概率与信道疑义度 $H(X|Y)$ 满足不等式：

$$H(X|Y) \le H(P_E) + P_E \log(r-1) \quad \text{费诺不等式}$$

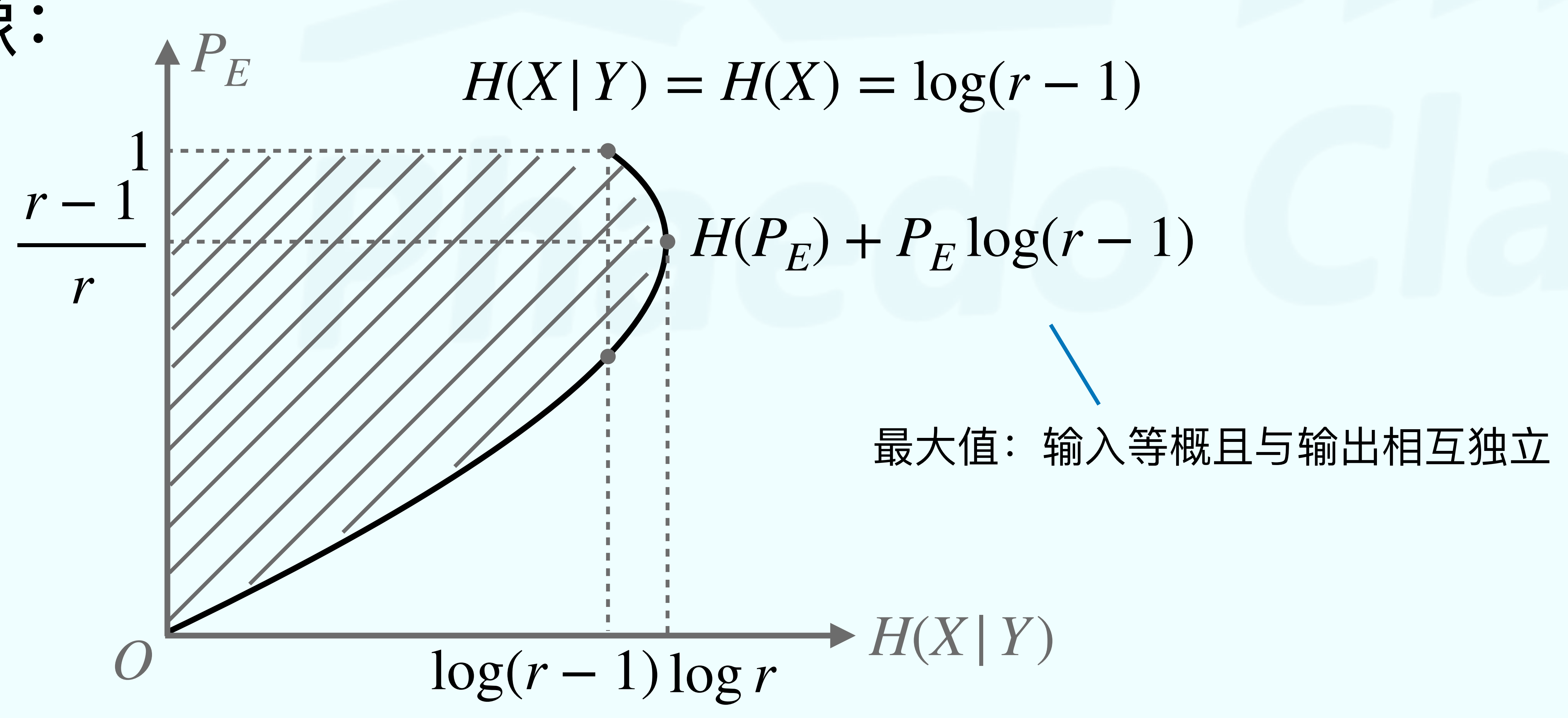
费诺不等式的解读 $H(X|Y) \le H(P_E) + P_E \log(r-1)$

- 不论采用什么译码准则，费诺不等式均成立。
- 信道疑义度由两部分组成：

$H(P_E)$：是否发生错误的不确定性，即收到 y 后产生值为 P_E 的平均错误概率的平均不确定度。

当错误发生后，确定由 $(r-1)$ 个输入符号中哪一个引起错误的不确定，其最大值为 $P_E \log(r-1)$。

- 图像：



- 物理意义：当信源、信道给定时，信道疑义度就给定了译码错误的下限。

错误概率与译码方法

小节1 简单重复编码

小节2 汉明距离

错误概率与译码方法

- 对于给定信道，输入符号概率一定时，选择译码规则可使 P_E 最小。
- 一般数字通信系统要求平均错误概率 P_E 在 10^{-6} 和 10^{-9} 数量级甚至更低。
- 费诺不等式表明，要进一步降低 P_E，仅仅制定译码规则已经不够了，需通过对信道的输入符号进行编码。

错误概率与译码方法

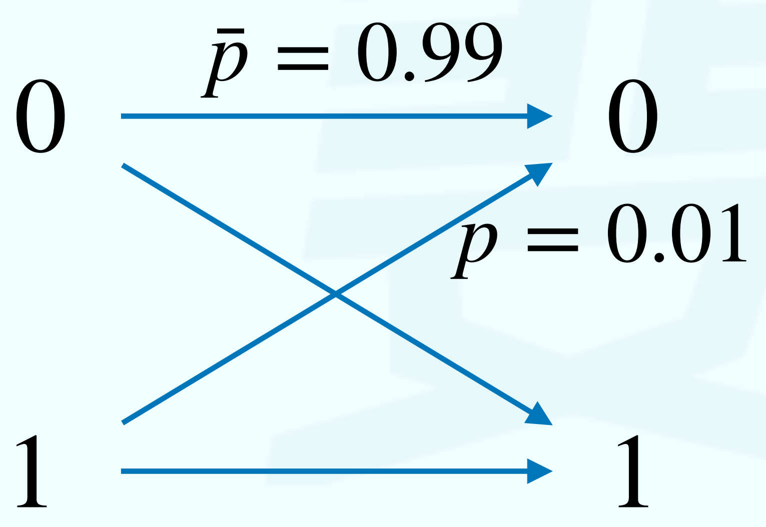
小节1 简单重复编码

小节2 汉明距离

简单重复编码

在发送端把消息多重复几遍，可以使接收端接收消息时错误概率减小。

以二元对称信道为例，假设输入等概分布，



信道传递矩阵为 $P = \begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} \end{array}$ (其中 $\bar{p} \gg p$) 。

译码规则 $\begin{cases} F(0) = 0 \\ F(1) = 1 \end{cases}$ 。

$$P_E = 1 - \frac{1}{r} \sum_{Y,X^*} p(y_j | x_i) = 1 - \frac{1}{2}(\bar{p} + \bar{p}) = 10^{-2}$$

简单重复编码

输入端三次重复，发 0 时，发送序列 $\alpha_1 = 000$，发1 时，发送序列 $\alpha_2 = 111$。

由于信道存在干扰，则输出存在 $2^3 = 8$ 种情况，即可视为三次无记忆扩展信道。

$\alpha_1 = 000$

$\alpha_2 = 111$

二元对称信道的三次扩展信道

$\beta_1 = 000 \quad \beta_2 = 001$

$\beta_3 = 010 \quad \beta_4 = 011$

$\beta_5 = 100 \quad \beta_6 = 101$

$\beta_7 = 110 \quad \beta_8 = 111$

$$P = \begin{array}{c} \\ 0 \\ 1 \end{array}\!\!\begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} \end{array}$$

此时信道矩阵为
$$P = \begin{array}{c} \\ 000 \\ 111 \end{array}\!\!\begin{array}{c} \begin{array}{cccccccc} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \end{array} \\ \begin{bmatrix} \bar{p}^3 & \bar{p}^2 p & \bar{p}^2 p & \bar{p} p^2 & \bar{p}^2 p & \bar{p} p^2 & \bar{p} p^2 & p^3 \\ p^3 & \bar{p} p^2 & \bar{p} p^2 & \bar{p}^2 p & \bar{p} p^2 & \bar{p}^2 p & \bar{p}^2 p & \bar{p}^3 \end{bmatrix} \end{array}$$
，且 $\bar{p} \gg p$。

假定输入等概分布，根据ML准则，译码准则为

$$\begin{cases} F(000) = F(001) = F(010) = F(100) = 000 = \alpha_1 \\ F(011) = F(101) = F(110) = F(111) = 111 = \alpha_2 \end{cases}$$

$$P_E = 1 - \frac{1}{r} \sum_{Y,X^*} P(y_j | x_i) \approx 3 \times 10^{-4}$$ 降低两个数量级

简单重复编码的分析

- 在输入符号集（r 个信源符号）等概的条件下，每个符号平均携带的最大信息量是 $\log r$ 。
- 当用 n 长码元符号来传输信源符号时，每个码符号携带的平均信息量，即信道编码信息率/信息传输率 $R=\dfrac{H(S)}{\bar{L}}=\dfrac{\log r}{n}=\dfrac{\log 2^k}{n}=\dfrac{k}{n}$ $bit/$ 码元符号。

  ①不重复编码时，$n=k$，$R=1$ $bit/$ 码符号。

  ②重复编码 $k=1$ 时，若三次重复，则 $R=\dfrac{1}{3}$ $bit/$ 码符号。

k : 信息位个数　　n : 编码后总位长

$S_1=0$, $S_2=1$ → 信道编码 → $\alpha_1=000$, $\alpha_2=111$

- n 增大时，P_E 减小，但 R 也在减小，因此效率 η 降低，$n\to\infty$ 时，$P_E\to 0$，但码率 $R\to 0$ 。

如果把简单重复编码的概念推广为：在扩展信源的 r^N 个码符号序列中任意选 M 个作为信道输入，以代表 M 个信源消息。如，上述例子中若不选用 000 和 111 作为输入，而选用 000 、001 作为输入，P_E 会高很多。

我们需要一个较为简便的方法，选择平均错误概率最小的 M 个序列，从而引入汉明距离。

错误概率与译码方法

小节1 简单重复编码

小节2 汉明距离

汉明距离

设 $\alpha_i = x_{i_1}x_{i_2}\dots x_{i_n}$ 和 $\beta_j = y_{j_1}y_{j_2}\dots y_{j_n}$ 表示两个长度为 n 的码字序列。定义

$$D(\alpha_i, \beta_j) = \sum_{k=1}^{n} x_{i_k} \oplus y_{j_k}$$

称 $D(\alpha_i, \beta_j)$ 为码字 α_i 和 β_j 之间的汉明距离。

例如：$\alpha_i = \underline{0}\ \underline{0}\ \underline{0}\ \underline{1}\ 1\ 0\ \underline{0}\ \underline{1}\ 0\ \underline{1}$

$\beta_j = 1\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 0\ 0$

$D(\alpha_i, \beta_j) = 7$ （有几个相异，汉明距离就是几）

最小汉明距离

在二元码 C 中，任意两个码字之间的汉明距离的最小值，被称为码 C 的最小汉明距离。

$$D_{\min} = \min[D(w_i, w_j)] \ (w_i \neq w_j,\ w_i, w_j \in C)$$

例如：设有 $n=3$ 的两组码，分别求它们的最小汉明距离。

	C_1	C_2
α_1	0 0 0	0 0 0
α_2	0 1 1	0 0 1
α_3	1 0 1	0 1 0
α_4	1 1 0	1 0 0

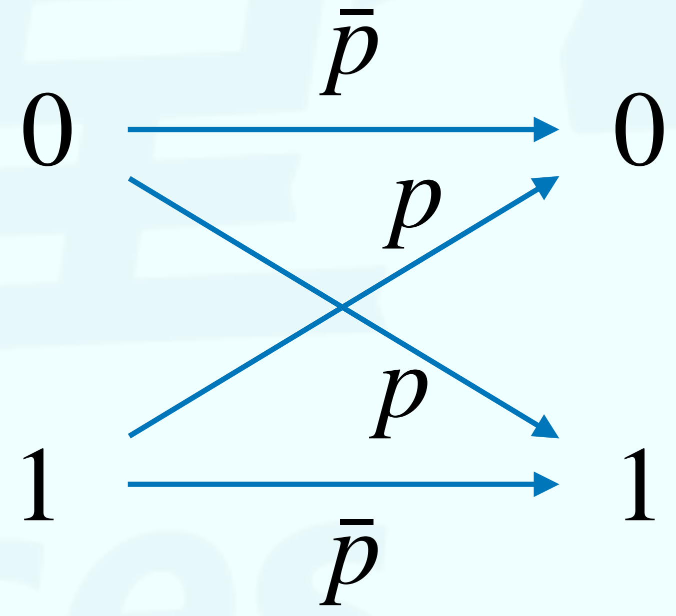
C_1 的最小汉明距离为 2 。

C_2 的最小汉明距离为 1 。

最小汉明距离译码准则MD

设 $\alpha_i = x_{i_1}x_{i_2}\ldots x_{i_n}$ 和 $\beta_j = y_{j_1}y_{j_2}\ldots y_{j_n}$ 表示两个长度为 n 的码字序列。α_i 为信道输入，β_j 表示信道输出，α_i 和 β_j 的汉明距离为 D。

对于离散无记忆二元对称信道，有 $\bar{p} > p$。



$$p(\beta_j|\alpha_i) = p(y_{j_1}y_{j_2}\ldots y_{j_n}|x_{i_1}x_{i_2}\ldots x_{i_n}) = \prod_{k=1}^{n} p(y_{j_k}|x_{i_k})$$
$$= (p)^D(\bar{p})^{n-D}$$

根据极大似然译码准则，$p(\beta_j|\alpha^*) \geq p(\beta_j|\alpha_i), \forall i, \bar{p} > p$，则 D 越小，$p(\beta_j|\alpha_i)$ 越大，

即 $D(\alpha^*, \beta_j) = D_{\min}(\alpha_i, \beta_j)$。

最小汉明距离译码准则MD

以简单重复编码的信道矩阵为例：

$$P=\begin{array}{c} \\ 000 \\ 111 \end{array}\begin{array}{c} \begin{array}{cccccccc} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \end{array} \\ \begin{bmatrix} \bar{p}^3 & \bar{p}^2 p & \bar{p}^2 p & \bar{p} p^2 & \bar{p}^2 p & \bar{p} p^2 & \bar{p} p^2 & p^3 \\ p^3 & \bar{p} p^2 & \bar{p} p^2 & \bar{p}^2 p & \bar{p} p^2 & \bar{p}^2 p & \bar{p}^2 p & \bar{p}^3 \end{bmatrix} \end{array}$$

化为码字之间的汉明距离

$$P=\begin{array}{c} \\ 000 \\ 111 \end{array}\begin{array}{c} \begin{array}{cccccccc} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 2 & 1 & 2 & 2 & 3 \\ 3 & 2 & 2 & 1 & 2 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

$$P_E = 1 - \frac{1}{2}\sum_{Y\cdot X^*} P(y_j|x_i) \approx 3\times 10^{-4}$$

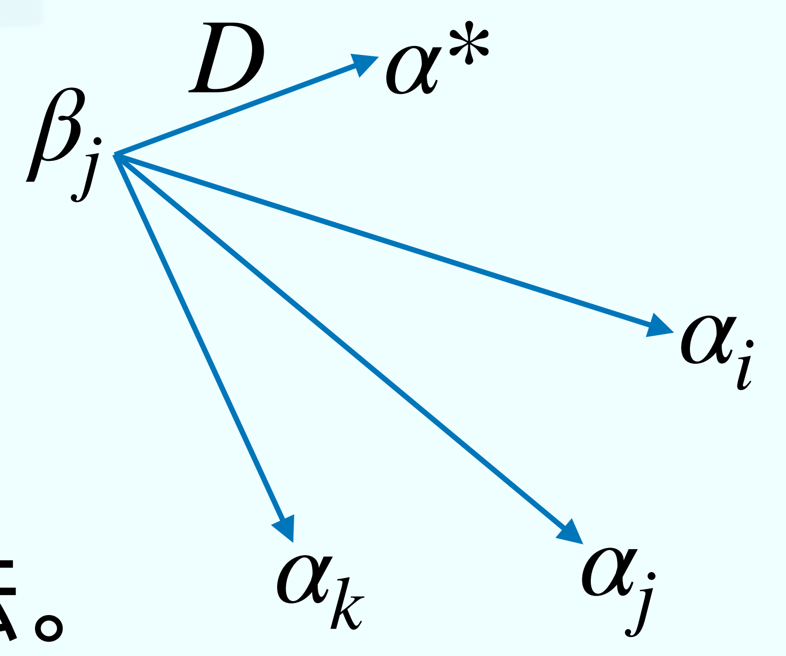
二者等价

$$P_E = 1 - \frac{1}{2}(2\bar{p}^3 + 6\bar{p}^2 p) \approx 3\times 10^{-4}$$

最小汉明距离译码准则的解读

- 收到 β_j 后，译成与之距离最近的输入码字 α^*，即把 β_j 译成与之最邻近的那个码字 α^*，使 P_E 最小。

- 应使其他码字 $\alpha_i \neq \alpha^*$ 与 β_j 的距离尽可能远，即使 M 个许用码字中的任何两个不同码字之间的距离尽量大，这样可保持一定的信息传输率 R 。



- 综上所述：P_E 与各种编码、译码方法有关。

①编码可采用选择发端的 M 个消息所对应的码字间的最小距离 $d_{\min}$ 尽可能大的方法。

②译码时采用将接收序列 β_j 译成与之距离最近的那个码字 α^* 的译码规则。

最小汉明距离译码准则的解读

- 纠错能力：可纠正 1 位及以内的错误。

$$P = \begin{array}{c|cccccccc} & 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ \hline 000 & 0 & 1 & 1 & 2 & 1 & 2 & 2 & 3 \\ 111 & 3 & 2 & 2 & 1 & 2 & 1 & 1 & 0 \end{array}$$

纠正回 000 : $C_3^0\bar{p}^3p^0 + C_3^1\bar{p}^2p^1$

纠正回 111 : $C_3^0\bar{p}^3p^0 + C_3^1\bar{p}^2p^1$

此时 $P_E = 1 - \bar{P}_E = 1 - \frac{1}{2}(C_3^0p^0\bar{p}^3 + C_3^1p^1\bar{p}^2 + C_3^0p^0\bar{p}^3 + C_3^1p^1\bar{p}^2) \quad = 1 - (C_3^0p^0\bar{p}^3 + C_3^1p^1\bar{p}^2)$

提前引入检、纠错能力的概念（第八讲还会介绍）

编码可以纠正u个以内错误的充要条件是$d_{\min} = 2u + 1$，本例最小汉明距离为3，因此可以纠1位错。

编码可以检测t个以内错误的充要条件是 $d_{\min} = t + 1$ ，本例最小汉明距离为3，因此可以检2位错。

- $\bar{p} \gg p$时，输入是否等概的影响可忽略不计，我们完全可以直接按照最大似然准则（或最小汉明距离准则）进行译码，不过译码错误概率就不能按照等概分布的公式计算了。

例题7-3 计算码长为 $n=5$ 的二元重复编码的错误概率，假设输入码字等概分布，无记忆二元对称信道中正确传递概率为 $\overline{p}$，错误传递概率 $p=1-\overline{p}$，问此码可以检测出多少错误？又能纠正多少错误？若错误传递概率 $p=0.01$，求错误概率。

解析7-3 长度为5的二元重复编码也就是00000，11111，故可以求得汉明距离为5

因此，该二元重复编码可以检测4位及以内的错误，输入等概分布时，纠正2位及以内的错误；

错3位的概率为 $C_5^3 p^3 \overline{p}^2$

错4位的概率为 $C_5^4 p^4 \overline{p}$

错5位的概率为 $C_5^5 p^5$

此时译码错误概率为 $p_e = C_5^3 p^3 \overline{p}^2 + C_5^4 p^4 \overline{p} + C_5^5 p^5 \approx 1.03 \times 10^{-5}$

例题7–4 设某二元码为 $C=\{11100,01001,10010,00111\}$。

(1) 计算此码的最小距离 $d_{\min}$；

(2) 计算此码的码率 R，假设码字等概分布；

(3) 采用最小距离译码准则，试问接收序列 10000、01100 和 00100 应译成什么码字？

(4) 此码能纠正几位码元的错误？

解析7–4 (1)

$$
\left.\begin{aligned}
11100 \oplus 01001 &= 3\\
11100 \oplus 10010 &= 3\\
11100 \oplus 00111 &= 4\\
01001 \oplus 10010 &= 4\\
01001 \oplus 00111 &= 3\\
10010 \oplus 00111 &= 3
\end{aligned}\right\} \longrightarrow d_{\min} = 3
$$

(2) $R=\dfrac{H(S)}{\bar{L}}=\dfrac{\log M}{n}=\dfrac{\log 2^k}{n}=\dfrac{k}{n}=\dfrac{2}{5}$ $bit/$码符号。

(3) 根据最小汉明距离译码准则，

10000 应译为 10010，01100 应译为 11100，00100 应译为 11100 或 00111 均可。

(4) 此码可纠正 1 位码元的错误，$d_{\min}=3=2\times1+1$。

有噪信道编码定理

小节1 香农第二定理

香农第二定理

内容：设有一离散无记忆平稳信源，其信道容量为C，若编码信息率 $R < C$。当码长n足够大时，则至少存在一种编码，使译码错误概率任意小；相反，若信息传输率 $R > C$，则码长n无论多长，总也找不到使译码错误概率任意小的编码。

「说明」香农第二定理指出来“高效率、高可靠性”的信道编码的存在性，但未指出具体方法。该定理指出了信道编码的极限性能，为信道编码的研究指明方向。

① 高效率：信息传输率接近信道容量。

② 高可靠性：译码差错任意小。

③ 存在这种信道编码的必要条件是 $R < C$。

香农三大定理总结

在信源编码部分，我们介绍了香农第一定理与香农第三定理。

香农第一定理给出了本课程的第一个理论极限：实现无失真信源编码时平均码长的压缩下限为信源熵。

香农第三定理给出了本课程的第二个理论极限：在允许失真D的条件下，信源最小的、可达的信息传输率是信源的信息率失真函数$R(D)$，信息率失真函数提供了压缩的下界。

在信道编码部分，我们介绍了香农第二定理。

香农第二定理给出了本课程的第三个理论极限：实现高可靠性信道编码的信息率极限是信道容量。

香农信息论的三个概念：信源熵、信道容量和信息率失真函数，都是临界值，是从理论上衡量通信能够满足要求的重要极限。