《信息论与编码理论基础》
第四章
信道容量

马猛
北京大学，信息科学技术学院

提纲

- 一、通信系统模型
- 二、信道编码定理
- 三、信道容量示例分析

一、通信系统模型

- 通信的概念
  - A与B通信是指A的物理行为使B产生一种需要的物理状态。
- 无误传输*
  - 20世纪40年代早期，人们普遍认为，以正速率发送信息而忽略误差概率是不可能做到的。
  - 一种简单的提高可靠性的传输方法是多次重复传输相同的信息。若要通过这种方法是错误概率趋近于零，则传输速率也将趋近于零。
  - Shannon证明了只要通信速率低于信道容量，总可以使误差概率接近于零。

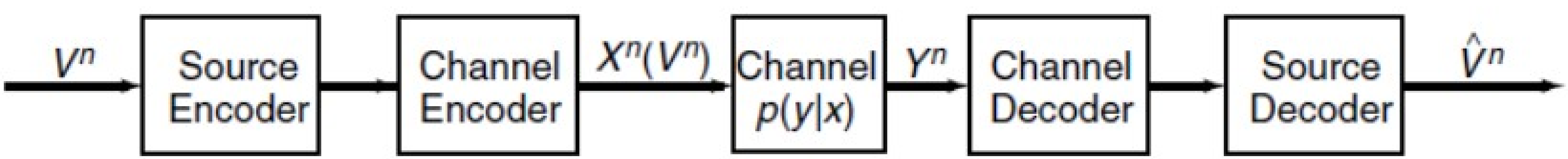


FIGURE 7.15. Separate source and channel coding.

* 请思考无误传输在通信系统中是必要的吗？

一、通信系统模型

- **离散信道定义**：离散信道是由输入字母表χ，输出字母表Y和概率转移矩阵$p(y|x)$构成的系统，其中$p(y|x)$表示发送字符x的条件下收到输出字符y的概率。
- **无记忆信道定义**：如果输出的概率分布仅依赖于它所对应的输入，而与先前信道的输入或者输出条件独立，就称这个信道是无记忆的。
- 离散无记忆信道的信道容量定义为

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

这里的取最大值取自所有可能的输入分布$p(x)$

二、信道编码定理

- 信道编码定理预览
  - 对于每个输入的（典型的）n长序列，会有大约$2^{nH(Y|X)}$个可能的Y序列与之对应，并且所有这些序列是等可能的
  - 无误传输需要确保没有两个序列能够产生相同的Y输出序列
  - 所有可能的（典型的）Y输出序列的总数约等于$2^{nH(Y)}$
  - 需要将接收序列集合分割成大小为$2^{nH(Y|X)}$的许多小集合，不相交集的总数等于$2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$
  - 因此，最多可以传输$2^{nI(X;Y)}$个可区分的n长序列

二、信道编码定理

- **信道编码定理**：对于离散无记忆信道，小于信道容量C的所有码率都是可达的。具体来说，对任意码率$R < C$，存在一个$(2^{nR}, n)$码*序列，它的最大误差概率为$\lambda^{(n)} \to 0$。反之，任何满足$\lambda^{(n)} \to 0$的$(2^{nR}, n)$码序列必定有$R \le C$。

- 证明：小于C的码率是R可达的

固定$p(x)$，根据分布$p(x)$随机生成一个$(2^{nR}, n)$码，即根据分布$p(x^n) = \prod_{i=1}^{n} p(x_i)$独立生成$2^{nR}$个码字。

* 有2^{nR}个码字，每个码字长度为n

二、信道编码定理

- 考虑下面的系列事件：
  - 1. 服从分布$p(x)$的随机码C生成
  - 2. 将码C告知给发送者和接收者，并且假定二者都知道该信道的信道转移矩阵$p(y|x)$
  - 3. 依如下的均匀分布选取一条消息W

    $$P_r(W=w)=2^{-nR}, w=1,2,\cdots,2^{nR}$$
  - 4. 第w个码字$X^n(w)$是C的第w个码字，通过信道被发送
  - 5. 接收者收到的序列服从分布$P(y^n|x^n(w))=\prod_{i=1}^{n}p(y_i|x_i(w))$
  - 6. 接收者猜测所发送的消息是什么。（使误差概率达到最小的最优方法是最大后验概率译码，也就是说，接收者应该根据选择后验概率最大的消息）

二、信道编码定理

- 基于最大后验概率的译码过程很难分析。因此，我们选用联合典型译码来进行分析。
- **联合典型定义：**服从分布$p(x,y)$的联合典型序列$\{(x^n,y^n)\}$所构成的集合$A_\varepsilon^{(n)}$是指其经验熵与真实熵接近的长序列构成的集合，即：

$$A_\varepsilon^{(n)}=\{(x^n,y^n)\in X^n\times Y^n:$$
$$\left|-\frac{1}{n}\log p(x^n)-H(X)\right|<\varepsilon$$
$$\left|-\frac{1}{n}\log p(y^n)-H(Y)\right|<\varepsilon$$
$$\left|-\frac{1}{n}\log p(x^n,y^n)-H(X,Y)\right|<\varepsilon\}$$

其中$p(x^n,y^n)=\prod_{i=1}^{n}p(x_i,y_i)$

- 如果满足下面两个条件，则接收者认为$\hat{w}$就是所发送的下标
  - $(X(\hat{w}),Y^n)$是联合典型的，该联合典型与信道转移概率有关
  - 不存在其它的下标满足联合典型的条件

二、信道编码定理

- 如果实际发送的码字为X^n，接收序列为Y^n，则x^n与y^n应具有相关性；反之则两者应该独立
- 考虑译码错误事件，即两个独立的序列x^n与y^n却满足联合典型条件的概率是多少？
  - 大约有$2^{nH(X)}$个典型的X序列和大约有$2^{nH(Y)}$个典型的Y序列（共形成序列对组合$2^{n(H(X)+H(Y))}$个），而联合典型序列只有$2^{nH(X,Y)}$个。
  - 因此，随机选取的输入信号X^n与Y^n为联合典型的概率大约等于$2^{n(H(X,Y)-H(X)-H(Y))} = 2^{-nI(X,Y)}$
- 发射码集共有2^{nR}个随机序列
  - 其中一个为真实发送的序列，该序列与Y^n不是联合典型的概率随着n的增加可以任意小（由典型集的概念），即小于ε。
  - 其它随机序列与Y^n联合典型的概率约为$2^{n(R-I(X,Y))}$，如果$R < C$，则总可以找到适当的n满足$2^{n(R-I(X,Y))} < \varepsilon$
- 综上，错误概率$< 2\varepsilon$是可实现的

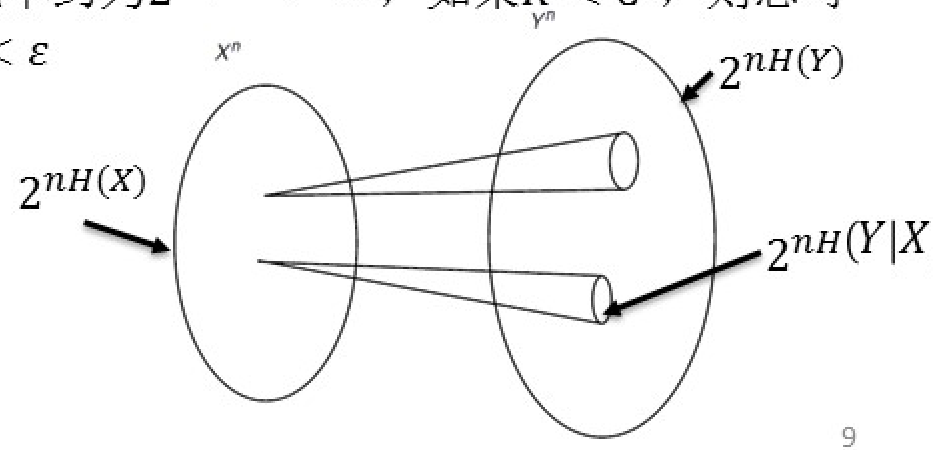


FIGURE 7.7. Channels after n uses.

二、信道编码定理

- 思考：如果$R > C$，错误概率随n的增长如何变化？
- 二元对称信道容量分析
  - 发送和接收信号均为0，1等概率出现的二进制序列
  - 信道错误概率为p，这里假设$p = 0.1$
  - 当序列长度n足够大时，一个真实的发送序列与接收序列大约有 $0.1n$个错误
  - 接收机将所有发射码集中的序列与接收序列一一比较，两者误差在$0.1n + \varepsilon < P_e < 0.1n + \varepsilon$范围内的即判为发送序列
  - 发射码集中的非真实发送序列与接收序列独立，因而大概率事件是发送序列与接收序列不同的符号数大约为$0.5n$，两者误差为$0.1n$是小概率事件
  - 但是，当发送序列数量很大时，有一个发射序列与Y^n误差为$0.1n$的概率可能由小概率时间转化为大概率事件，分界点为$R = C$

二、信道编码定理

$$I(X;Y)=H(Y)-H(Y|X)=H(Y)-\left(-p\log p-(1-p)\log(1-p)\right)$$
$$=H(Y)-H(p)$$

上式第一项为接收序列的熵，第二项完全由信道的错误概率决定，即对于给定的信道该项为定值。因此，最大化$I(X;Y)$等价于最大化$H(Y)$。

如下图所示，每个输入序列会有大约$2^{nH(Y|X)}$=$2^{nH(p)}$个可能的接收序列，要使可分辨的X^n序列最多（即传输速率最大），则应使右边的Y^n集合最大，以容纳更多的大小为$2^{nH(p)}$的小集合，即信道容量为互信息的最大值

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

由第一章例题1.2.1，均匀分布时熵最大为1比特，所以

$$I(X;Y)\le 1-H(p)=C$$

- (M,n)码的码率R为$R=\frac{\log_2 M}{n}$bit/transmission
- 容量C的单位和码率R的单位一样，可以用bit/transmission

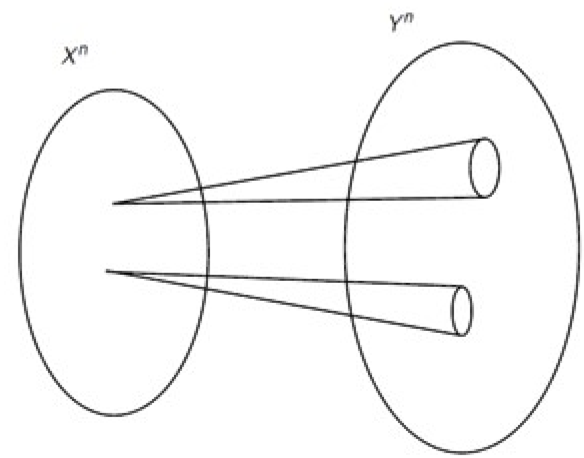


FIGURE 7.7. Channels after n uses.

三、信道容量示例分析

- 无噪声二元信道
  - 二元输入序列能够无误地传输到接收端
  - 相当于二元对称信道的错误概率$p=0$，则$H(p)=0$（基于约定$0\log 0=0$）
- 信道容量计算：$C=\max\limits_{p(x)} I(X;Y)=\max\limits_{p(x)}[H(Y)-H(Y|X)]=\max\limits_{p(x)} H(Y)=1\text{ bit}$
  当$p(x)=\left(\frac{1}{2},\frac{1}{2}\right)$时达到
- 思考题：请画出此时X^n到Y^n映射的图像，确定各集合的大小。

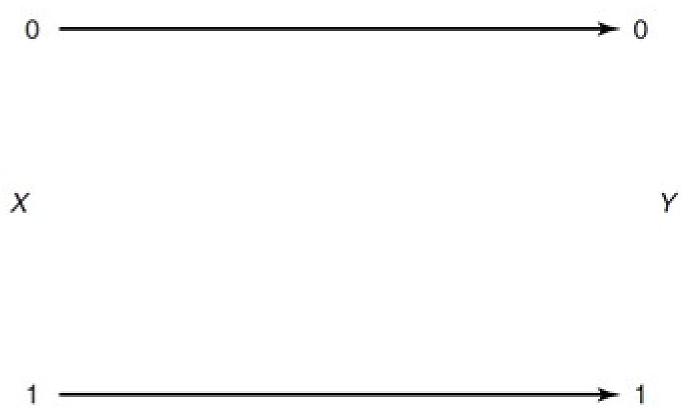


FIGURE 7.2. Noiseless binary channel. $C=1$ bit.

三、信道容量示例分析

- 无重叠输出的有噪声信道
  - 输入序列为二进制，输出序列为四进制，如下图所示
  - 根据输出信号可以无误地恢复发送信号，因而其容量为1bit，且当$p(x) = \left(\frac{1}{2}, \frac{1}{2}\right)$时达到
  - 思考题：请画出此时X^n到Y^n映射的图像，确定各集合的大小。

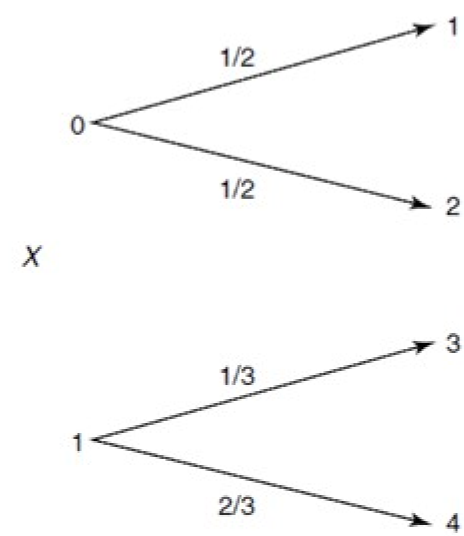


FIGURE 7.3. Noisy channel with nonoverlapping outputs. $C = 1$ bit.

三、信道容量示例分析

- 二元对称信道
  - 输入输出皆为二元变量，错误概率为p
- 互信息

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

$$= H(Y) - \sum p(x)H(Y|X=x) = H(Y) - \sum p(x)H(p)$$

$$= H(Y) - H(p) \le 1 - H(p)$$

当输入均匀分布时取等号

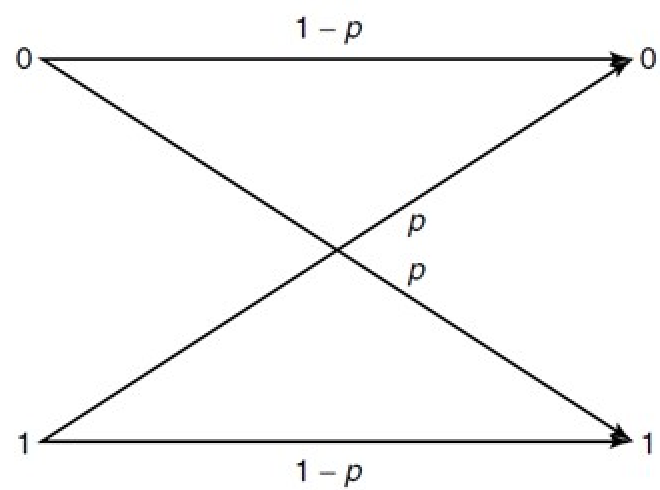


FIGURE 7.5. Binary symmetric channel. $C = 1 - H(p)$ bits.

三、信道容量示例分析

- 习题7.15 联合典型序列。设二元对称信道的交叉概率为0.1。达到信道容量的输入分布为均匀分布，即$p(x) = (0.5,0.5)$，此时产生的联合分布为

$X\backslash Y$	0	1
0	0.45	0.05
1	0.05	0.45

（a）计算$H(X)$，$H(Y)$，$H(X,Y)$和$I(X;Y)$

（b）设$X_1, X_2, \cdots, X_n$为i.i.d服从Bernoulli$\left(\frac{1}{2}\right)$分布。在长度为n的2^n个可能的输入序列中，哪些是典型的（即$\varepsilon = 0.2$时，$A_\varepsilon^{(n)}(X)$中的元素）？$A_\varepsilon^{(n)}(Y)$中的典型序列又是什么？

（c）该信道可以建模为$Y = X \oplus Z$，其中Z为二元序列，等于1的概率为p，且独立于X。对$n = 25, \varepsilon = 0.2$,计算$A_\varepsilon^{(n)}(Z)$的大小？

（d）对于固定的码字$x^n(i)$，使$(x^n(i), y^n)$为联合典型的接收序列y^n的概率为多少？

（e）考虑特定的接收序列$y^n = 000000\cdots0$。假定在长度为n的所有2^n个可能的二元序列上，随机均匀地选取一个序列X^n.选取的序列与这个y^n为联合典型的概率是多少？

（f）考虑由512个码字组成的编码，每个码字都是随机均匀取自长度为$n = 25$的2^n个序列。设其中一个码字为发送序列，对应一个接收序列y^n，其余码字与y^n联合典型的概率是多少？

（g）译码时，当发送序列与接收序列不联合典型，或者另一个码字与接收序列联合典型都导致错误，求错误概率。

三、信道容量示例分析

- 解：（a）$H(X) = H(Y) = 1\text{bit}$

$H(X,Y) = H(X) + H(Y|X) = 1 + H(p) = 1 - 0.9\log 0.9 - 0.1\log 0.1 = 1.469\text{bit}$

$I(X;Y) = H(Y) - H(Y|X)$=0.531bit

（b）对于均匀分布，每个序列的概率都满足$-\frac{1}{n}\log p(x^n) = 1 = H(X)$,因此所有的序列都是典型的。同样，所有序列$y^n$也是典型的。

（c）$H(Z) = H(0.1) = 0.469$.对于$\varepsilon = 0.2$需要寻找范围(0.269,0.669)，对应$k = 1,2,3,4$，序列落入典型集的概率为0.830216，个数为15275

（d）落入典型集的概率为0.830216

k	$\binom{n}{k}$	$\binom{n}{k}p^k(1-p)^{n-k}$	$-\frac{1}{n}\log p(x^n)$
0	1	0.071790	0.152003
1	25	0.199416	0.278800
2	300	0.265888	0.405597
3	2300	0.226497	0.532394
4	12650	0.138415	0.659191
5	53130	0.064594	0.785988
6	177100	0.023924	0.912785
7	480700	0.007215	1.039582
8	1081575	0.001804	1.166379
9	2042975	0.000379	1.293176
10	3268760	0.000067	1.419973
11	4457400	0.000010	1.546770
12	5200300	0.000001	1.673567

三、信道容量示例分析

（e）随机选取的X^n与给定的y^n为联合典型的概率为

$$\frac{2^{nH(X|Y)}}{2^{nH(X)}} = 2^{-nI(X;Y)} = \frac{2^{nH(Y|X)}}{2^{nH(Y)}} = \left|A_\varepsilon^{(n)}(Z)\right| \times 2^{-n}$$

$$= 15275 \times 2^{-25} = 4.552 \times 10^{-4}$$

（f）其余的511个序列中没有一个与y^n联合典型的概率为$(1 - 4.552 \times 10^{-4})^{511} = 0.79241$,至少有一个联合典型的概率为$1 - 0.79241 = 0.20749$

（g）考虑（d）和（f）的结果，错误概率上界为

0.1698 + 0.2075 = 0.3773

错误概率较高的原因是该题的码字长度n还不够大

三、信道容量示例分析

- 二元擦除信道，信道模型如下图所示
- 容量计算

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

$$= \max_{p(x)} [H(Y) - H(\alpha)]$$

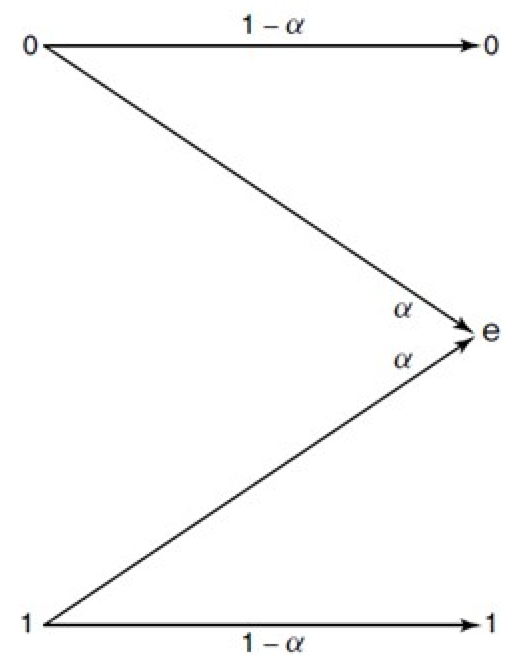


FIGURE 7.6. Binary erasure channel.

三、信道容量示例分析

设E代表事件$\{Y=E\}$，且由链式法则：

$$H(Y)=H(Y,E)=H(E)+H(Y|E)$$

设$\Pr(X=1)=\pi$，考虑到$H(Y|E=e)=0$，$H(Y|E\neq e)=H(\pi)$，由上式得

$$H(Y)=H(\alpha)+(1-\alpha)H(\pi)$$

则：
$$C=\max_{p(x)}[H(Y)-H(\alpha)]$$

$$=\max_{p(x)}[(1-\alpha)H(\pi)+H(\alpha)-H(\alpha)]$$

$$=\max_{p(x)}[(1-\alpha)H(\pi)]=1-\alpha$$

三、信道容量示例分析

- 习题7.1 输出的预处理。如果一个统计学家面对具有转移概率为$p(y|x)$且信道容量$\max_{p(x)} I(X,Y)$的通信信道，他会对输出做出很有帮助的预处理：$\tilde{Y} = g(Y)$，并且断定这样做能够严格地改进容量。

（a）请证明他错了

（b）在什么条件下他不会严格地减小容量？

解：（a）回顾第一章介绍的数据处理不等式。因为$X \to Y \to \tilde{Y}$构成马尔可夫链，因此

$$I(X;Y) \geq I\left(X;\tilde{Y}\right)$$

令$\tilde{p}(x)$是最大化$I\left(X;\tilde{Y}\right)$的$x$的分布函数，则

$$C = \max_{p(x)} I(X;Y) \geq I(X;Y)_{p(x)=\tilde{p}(x)} \geq I\left(X;\tilde{Y}\right)_{p(x)=\tilde{p}(x)} = \max_{p(x)} I\left(X;\tilde{Y}\right) = \tilde{C}$$

因此，该统计学家的方法无法提高容量

（b）当$I\left(X;Y\middle|\tilde{Y}\right) = 0$时等号成立。

当$X \to \tilde{Y} \to Y$也构成一个马尔可夫链（即在给定$\tilde{Y}$时，Y独立于X）时，等式成立，即容量不减小（此时称$\tilde{Y}$为充分统计量，见教材2.9节）。例如当函数$g(Y)$可逆时。

三、信道容量示例分析

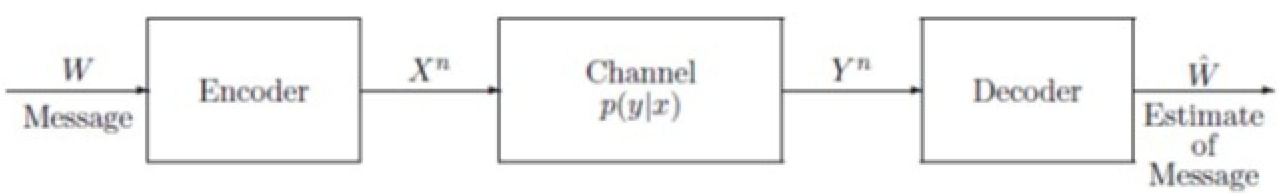
- 习题7.16 编码器与解码器作为信道的一部分。考虑交叉概率为0.1的二元对称信道。对于这个信道，考虑两个长度为3的码字。可能的方案是将消息a_1编码为000，将消息a_2编码为111。对此编码方案，进一步将编码器、信道和译码器组合起来考虑，从而形成一个新的BSC，其两个输入为a_1和a_2，两个输出也为a_1和a_2。

（a）计算该信道的交叉概率

（b）该信道的信道容量为多少？（量纲为比特/原信道传输）

（c）交叉概率为0.1的原始BSC的信道容量为多少？

（d）证明下面关于信道的一般结论：将编码器、信道和译码器组合起来考虑，形成一个消息到被估计消息的新信道，这种方式不会增加信道容量（量纲为比特/原信道传输）



三、信道容量示例分析

解：（a）接收判决为 a_1 的码字包括：000、100、010、001，正确概率为：$p_r = 0.9^3 + 3 \times 0.9^2 \times 0.1 = 0.972$，错误概率为0.028

接收判决为a_2的码字包括：111、110、101、011，错误概率也为0.028。该信道仍为二元对称信道，交叉概率为0.028

（b）$C = 1 - H(p = 0.028) = 1 - 018426 =$ 0.81574bit/codeword=0.2719bit/transmission

（c）原信道容量为$C = 1 - H(p = 0.1) =$ 0.531bit/transmission

（d）由前面介绍的数据处理不等式可得$I\left(W;\widehat{W}\right) \leq I(X^n;Y^n)$，因此

$$C_W = \frac{1}{n}\max_{p(w)} I\left(W;\widehat{W}\right) \leq \frac{1}{n}\max_{p(x^n)} I(X^n;Y^n) = C$$

因此，任何加入编码器和解码器的信道不会增加信道容量。

- 思考题：考虑一个基带有线数据通信系统，可以利用电平高低传输信息。

（1）请你根据香侬信道容量的思想设计该系统使得其可以达到无误传输的最大速率。

（2）请你分析该方案在系统实现的过程中会遇到哪些困难？

习题

- 阅读教材7.1，7.3，7.4，7.5，7.6，7.7
- 习题7.4，7.6（a，b），7.8，7.9，7.23